Political Agency in Democracies and Dictatorships

A dissertation presented by

Georgy Vladimirovich Egorov

to

The Department of Economics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Economics

Harvard University
Cambridge, Massachusetts

May 2009
Political Agency in Democracies and Dictatorships

ABSTRACT

This dissertation consists of three chapters. In the first chapter, I study how heterogeneity of voter preferences affects political accountability in democratic regimes. I introduce a common agency model, with voters as principals and the politician as the agent, and multiple policy dimensions. I identify several new effects resulting from the heterogeneity in voters’ preferences. In particular, there is a non-monotonic effect of transparency on political accountability. The model also implies that small groups may be more successful in inducing politicians to choose policies in line with their preferences, and provides a novel mechanism for the underprovision of public goods, whereby voters who equally care about a public good may nonetheless fail to induce the politician to provide it.

The second chapter studies the dynamic selection of governments under different political institutions, with a special focus on institutional flexibility. The competence of the government in office determines collective utilities, and each individual derives additional utility from being part of the government. Then perfect democracy, where current members of the government do not have an incumbency advantage, always leads to the emergence of the most competent government in equilibrium. However, any deviation from perfect democracy destroys this result. Moreover, a greater degree of democracy may lead to worse governments. In contrast, in the presence of stochastic shocks or changes in the environment, greater democracy increases the probability that competent governments will come to power. This suggests that a particular advantage of democratic regimes may be their greater adaptability to changes rather than their performance under given conditions.

The third chapter studies the principal-agent interactions in nondemocratic regimes. Consider a dictator who may be betrayed by a close associate. More competent advi-
sors are better able to discriminate among potential plotters, and this makes them more risky subordinates. To avoid this, rulers, especially those which are weak and vulnerable, sacrifice the competence of their agents, hiring mediocre but loyal subordinates. The static model allows us to characterize, under what conditions incompetent advisors will be chosen. The dynamic model allows us to focus on the succession problem that insecure dictators face.
# TABLE OF CONTENTS

Abstract ................................................................. iii

Acknowledgements ..................................................... vi

1. Political Accountability under Special Interest Politics .......... 1
   1.1 Introduction ...................................................... 1
   1.2 Relation to Literature ......................................... 4
   1.3 Example .......................................................... 8
      1.3.1 Transparency ............................................... 8
      1.3.2 Public Good Provision ..................................... 10
   1.4 Formal Model ................................................... 11
   1.5 Equilibria ....................................................... 14
   1.6 Effects of Transparency ......................................... 19
      1.6.1 Transparency and Accountability ......................... 19
      1.6.2 Optimal Degree of Supermajority ......................... 22
   1.7 Polarization ..................................................... 25
   1.8 Deliberation and Delegation .................................... 28
   1.9 Public Goods and Lobbying ..................................... 31
      1.9.1 Public Goods Provision ................................... 31
      1.9.2 Lobbying .................................................... 34
   1.10 Conclusion ..................................................... 35
   1.11 Appendix ........................................................ 38

2. Political Selection and Persistence of Bad Governments .......... 48
   2.1 Introduction ...................................................... 48
   2.2 Model .............................................................. 55
   2.3 Political Equilibria in Nonstochastic Environments ............ 59
      2.3.1 Political Equilibrium ....................................... 60
      2.3.2 Characterization ............................................. 62
   2.4 Characterization of Nonstochastic Transitions ................. 65
   2.5 Equilibria in Stochastic Environments .......................... 71
      2.5.1 Stochastic Political Equilibria ............................ 72
      2.5.2 The Structure of Stochastic Transitions ................. 74
   2.6 Conclusion ...................................................... 79
   2.7 Appendix A ........................................................ 82
   2.8 Appendix B ....................................................... 92
      2.8.1 Dynamic Game ............................................... 92
      2.8.2 Strategies and Definition of Equilibrium ................. 95
      2.8.3 Nonstochastic Characterization ............................ 97
ACKNOWLEDGMENTS

I am deeply indebted to my thesis committee Philippe Aghion, Daron Acemoglu, and Attila Ambrus for their support and encouragement. I am especially indebted to Daron for his contagious enthusiasm which encouraged me to work in theoretical political economy, and for his help, guidance, advice, and collaboration on joint projects. I am grateful to Philippe for his insightful comments and attention to my research and my progress. I thank Attila for his advice and for encouragement to keep up with research in economic theory.

This dissertation would be different without the help of other faculty members. In particular, I would like to thank Susan Athey, Kenneth Shepsle, and Aleh Tsyvinski. Aleh deserves a special mention, as both taking his class, working as his research assistant at the NBER, and getting his advice in general was much more important than he expected and even intended to be. I would also like to thank Alberto Alesina, Oliver Hart, Roger Myerson, James Robinson, and Andrei Shleifer for the conversations and comments over the years I spent at Harvard.

I am especially grateful to Konstantin Sonin, who played a vital role in my decision to take the academic path in life. Since I started working with him in 2003, he has been a great co-author and friend. He played a key role in teaching me how to do research in economics. I would also like to thank him, together with Daron Acemoglu, Sergei Guriev, and James Robinson for their encouragement to do research in political economy.

Most importantly, I thank my family for their support and encouragement during my years at Harvard and before that. My studies, my research, my formation as an economist and a person would not be possible without their help and their love.
1. POLITICAL ACCOUNTABILITY UNDER SPECIAL INTEREST POLITICS

1.1 Introduction

Two common arguments in favor of democratic electoral systems are the following. First, democratic elections provide an effective method of aggregating the dispersed and heterogeneous information and interests of citizens. Second, elections make politicians accountable to voters. There are large literatures which study both these roles separately, but not much research on their interaction. On the one hand, many papers study how voters’ preferences are aggregated, both normatively and positively. On the other hand, there is an equally fruitful literature about political accountability, i.e., the mechanisms and strategies the voters should use to distinguish between a well-performing and a poorly performing politician, and to make sure that a poorly performing one does not stay in office. This paper suggests why the aggregation of information about heterogeneous interests may interfere with the political accountability role of democracies.

The fact that conflicts of interests between voters exist, and are likely to play a role in the political process, was understood as far back as the 18th century by the Founding Fathers of the U.S. Known for their concern about voters’ control over politicians, the balance of power and the non-degeneration towards dictatorship, they did recognize the existence and possible impact of the misalignment of voters’ preferences. In Federalist papers, James Madison writes: “The latent causes of faction are thus sown in the nature of man; and we see them everywhere brought into different degrees of activity, according to the different circumstances of civil society.” He then goes on to discuss the implications of factions on which constitution would work best for the new country.

In this paper, I study the interaction between the two tasks of democracy. More precisely, I ask the following question: How does the conflict, or the misalignment, of
voters’ interests impact the accountability of a politician to voters at large? I take a theoretical approach, and introduce a framework to study the interaction between conflict of interests and political accountability. I build a common agency model in which the voters are principals and the politician is an agent. I make the assumption that the only decision a voter makes is to vote for or against the incumbent politician in future elections; in other words, monetary transfers are ruled out. I characterize the equilibrium, which under certain assumptions is simple and natural, and then study the comparative statics of such equilibria with respect to several parameters. I find several new effects which would be missing if voters had aligned preferences. These effects demonstrate that the conflict of interests among voters does have an important and non-trivial impact on the problem of making the politician accountable to the voters.

In particular, I show the following:

1. In contrast to the standard models of political accountability, which imply that more transparency makes the politician more accountable, I show that there is a non-monotonic effect, and an optimal degree of transparency. This effect only exists when voters care about different policy dimensions.

2. Small groups may be more successful in influencing the policies than large groups. This result is consistent with Olson’s (1971) observation, although contradicts much of the previous literature on voting. The effect also comes from that voters may care about more than one policy issue, and large groups have a disadvantage here. It is worth noting that the effect does not come from the ability of small groups to overcome the free-rider problem and lobby more efficiently than large groups; here, it arises even though voting is the only instrument available to voters.

3. The underprovision of public goods is typical when citizens make voluntary contributions. In voting models, however, the standard result is that the choice of the median voter will be implemented. I show that once voters are assumed to care about other policy issues as well as the public good, the public good will again be underprovided. The conflict of interests between voters on other policy dimen-
sions makes them condition their voting on the politician’s performance on those dimensions, while putting less emphasis on the provision of the public good.

4. Higher polarization over some policy issues may make the politician exert more, rather than less, effort, and thus make the non-partisan voters better off. This goes against Besley’s (2005) conclusion that higher polarization decreases the ability of voters to induce the politician to exert effort on the dimension all voters care about equally, and thus decreases the utility of non-partisan voters. The reason is that, in my model, higher polarization over some dimensions decreases the politician’s ability to get a lot of votes by taking some position along that dimension; consequently, to get enough votes to get reelected, he must exert effort in other dimensions.

5. Supermajority (and, in the extreme, unanimity) rules are better when (a) satisfying most of the voters is not too costly for the politician, and (b) the environment is transparent, i.e. the voters have a good understanding of whether the politician tried to do what they wanted. In that case, a politician who works hard will get the voters’ support, and will not find it too expensive to satisfy the demands of many voters. If these conditions are not satisfied, a supermajority requirement will induce the politician to shirk because he will view his chance of reelection as very small, regardless of his effort.

6. I also discuss the possibility of deliberation (meaning the exchange of information about the politician’s performance) between voters with aligned interests prior to voting. While each group of voters would always prefer to exchange information, this may hurt the society as a whole; this may happen because the votes will then effectively become more informative, and too much information possessed by the voters may hurt the society at large. This also enables me to compare direct elections, where all citizens vote together, with two-stage elections, such as presidential elections in the U.S. It turns out that direct elections are better when voters are well-informed, while two-stage elections have an advantage when the voters are poorly informed.
The rest of the paper is organized as follows. Section 1.2 contains a literature review. In Section 1.3, I provide an example which illustrates some of the main effects and trade-offs in this paper: the non-monotonic effect of transparency and the underprovision of public goods. Section 1.4 introduces the formal model, and Section 1.5 studies its equilibria. Then, I apply these results to study the effects of transparency in Section 1.6, and also discuss when a high degree of supermajority is optimal. The effects of polarization are considered in Section 1.7. I study whether it is optimal for voters with similar interests to deliberate prior to voting in Section 1.8. Finally, I extend the model to the case where voters may care about a public policy issue in Section 1.9. Section 1.10 concludes.

1.2 Relation to Literature

The voters’ problem of providing incentives to politicians has been at the forefront of research in political economy since the seminal works of Barro (1973) and Ferejohn (1986). Since then, there has been an extensive research on political accountability using prospective voting (both Barro, 1973, and Ferejohn, 1986, employed retrospective voting), where citizens observe the past actions of the politician and update their beliefs on his preferences or abilities (see Austen-Smith and Banks, 1999, Persson and Tabellini, 2000, and Besley, 2005, for recent surveys). In this paper, I take the retrospective voting approach by assuming that voters are able to commit to rewarding or punishing the politician. While admittedly a simplification of the complex reality, retrospective voting plays an important role in the U.S. congressional elections (Fiorina, 1981).

Historically, most of voting models assumed a single-dimensional policy space, and focused on convergence of the policy outcome to the preferences of the median voter. Downs (1957) provides an argument for the convergence of political platforms when politicians are free to choose their electoral positions. Other models, where the politician is both policy- and office-oriented (e.g., Wittman, 1973), or can invest in valence (e.g., Ashworth and Bueno de Mesquita, 2007), yield non-convergence. A more general case with multiple policy issues is typically intractable for deriving comparative statics, and the focus is instead on the general existence and uniqueness of equilibria. Another part of the literature
on voting models challenges the assumption that voters vote automatically and sincerely. Voters may have strategic considerations: for instance, they may take into account that their vote matters only if they are pivotal. This effect is introduced in Feddersen and Pesendorfer (1996); in a later paper (Feddersen and Pesendorfer, 1998a), the authors study the effect of voting rule on information revelation in juries’ verdicts. This literature, however, does not consider the problem of providing incentives to elected politicians, rather focusing on political campaigns or aggregation of information available to voters.

The agency relation between voters and the politician links this paper to a vast literature on principal-agent problems and, most importantly, common agency. The common agency approach was initially introduced in the seminal paper of Bernheim and Whinston (1986), and later applied to problems in political economy, such as lobbying (Dixit, Grossman, and Helpman 1997) and legislative control of bureaucrats (Martimort, 1996 and 2004). In games of common agency the agent chooses an action after observing the “menus” provided by the multiple principals. The focus in this literature is on “truthful” equilibria, in which the multiple principals (e.g., the lobbyists) present contracts which leave the agent (e.g., the politician) indifferent between choosing different actions, and the socially efficient action is implemented. Bernheim and Whinston (1986) also show that such equilibria are coalition-proof, and any coalition-proof equilibrium is payoff-equivalent to some truthful one.

This paper is different from this literature in that it assumes away monetary transfers, both from the voters to the politician (and vice versa) and between voters. Instead, I assume that voting is the only instrument available to voters. This rules truthful equilibria out, except for trivial cases. However, the equilibria are more straightforward and intuitive: for example, in many cases, the only equilibrium involves voters rewarding the politician with their votes if they got a positive signal, and punishing him if they got a negative one. Such simple equilibrium structure may be compared with “natural” equilibria in common agency games, introduced in Kirchsteiger and Prat (2001). They show that such equilibria always exist, and while they do not have some nice properties that truthful equilibria do, like coalition-proofness, they are easier to compute and are even more likely to be played.
In short, a natural equilibrium is one where each principal makes a non-zero transfer to the agent only for one realization of the signal. In common agency games, both truthful and natural equilibria exist, and the question of proper equilibrium refinement remains. In this paper, the simple voting strategy is typically a unique trembling-hand perfect equilibrium.

The comparative statics connects my model to vast recent literatures on transparency, polarization, lobbying, and public good provision. An immediate, and very intuitive, consequence of standard models of political accountability (Persson, Roland, and Tabellini, 1997, Persson and Tabellini, 2000) and similar models is that transparency is required for accountability: The more able are voters to observe politician’s performance, the more powerful incentives they can provide. Feddersen and Pesendorfer (1998b) argue that poorly informed voters introduce noise in the election, and make the argument for better informed voters. In line with this research, Glaeser, Ponzetto, and Shleifer (2007) argue that for democracy to function, citizens must be well-educated.

Yet a number of recent papers highlighted the potential adverse impact of transparency. Prat (2005) considers a political career-concerns model, where the incumbent politician wants to signal that he is of a higher type in order to get reelected. The ultimate effect of transparency depends on the extent to which the voters get information both about politician’s actions and his type. More precise information about the politician’s actions is always welfare-improving, while more precise information about his type has two countervailing effects. On one hand, it allows the voters to choose better politicians, once they observe their types with less noise. On the other hand, transparency destroys politician’s incentives to signal his type. As a result, transparency may have a negative overall effect on the social welfare. (See also Levy, 2007, for applications of this idea to optimal voting systems in committees.)

In my model, all politicians are of the same type, and, unlike the Prat’s (2005) model, more information about the policy outcomes may hurt the voters. The absence of career concerns leads to different policy implications than those in the Prat’s paper. While in his model there is no downside of monitoring old politicians, whose types are already known to the voters, the model in this paper suggests a reason to restrict monitoring even of old
A different model where full transparency is suboptimal is in Mattozzi and Merlo (2007). The authors build upon the citizen-candidate model pioneered by Osborne and Slivinski (1996) and Besley and Coate (1997) where citizens endogenously choose to run for a public office. Whenever the society monitors the politicians too closely, they are unable to get sufficient rents, and as a result most able citizens do not opt to become politicians, which is costly for the society. Mattozzi and Merlo model suggests that, other things equal, young politicians should be monitored less than old and known ones. When a citizen chooses his career, he discounts the future rents of the office; therefore, allowing young politicians to earn higher rents makes political career more attractive to most able young people. At the same time, if old politicians are monitored closely, they earn fewer rents and the society is better off, but this hardly affects the decision of young citizens. Again, the model in this paper is free from this counterintuitive prediction: here, the optimal degree of transparency is given by the preferences of the politician and of different voters.

Also, this paper obtains new results about aggregating voters’ preferences and lobbying which is different from the existing literature. Both in Downsian voting models and in common-agency model with monetary transfers, larger groups of voters with aligned interests might be more powerful politically, than smaller groups. E.g., this effect is present in the ‘protection for sale’ model of Grossman and Helpman (1994). This is also true in collective bargaining models, where voters become agenda-setters at random and make proposals (e.g., Banks and Duggan, 2000, and Acemoglu, Egorov, and Sonin, 2008). However, as Mancur Olson has famously argued in “The Logic of Collective Action” and Esteban and Ray (2001) demonstrated in a formal model, small lobbies may be more politically powerful than large ones due to their superior ability to overcome the collective action problem. In the former paper, the result rests on the assumption that members of the lobbying groups may make campaign contributions or bribing the sitting politician, i.e., there are monetary transfers of some kind. Therefore, there is a plausible argument that a purely democratic procedure (voting) always result in the prevalence of large groups
over the small ones. I show, however, that small groups may be politically powerful even if transfers to the politician are completely ruled out.

1.3 Example

1.3.1 Transparency

This example illustrates the main trade-off considered in the paper. There are three voters 1, 2, 3 and one politician. Each voter cares about one policy issue (voter 1 cares about issue 1, etc.), and he wants the politician to exert effort to solve the problem along that dimension. Assume that the effort that the politician can exert in solving the problem is binary in each dimension, thus he can choose any of the eight combinations from \( \{0,1\}^3 \), and so \( e = (e_1, e_2, e_3) \in \{0,1\}^3 \). Effort level 0 is costless, while effort 1 costs \( c = \frac{1}{8} \) (and the total effort cost is therefore \( (e_1 + e_2 + e_3) c \)).

After the politician exerts effort, the policy outcome \( \pi = (\pi_1, \pi_2, \pi_3) \in \{0,1\}^3 \) is realized and observed by the voters. In this example, I assume that the politician has perfect control over the policy outcomes, i.e., \( \pi_j = e_j \) for sure. However, the voters may not observe \( \pi \) precisely at the time of voting. Namely, voter \( i \) gets signal \( s_i \) such that

\[
\Pr (s_i = 1 \mid \pi_i = 1) = \Pr (s_i = 0 \mid \pi_i = 0) = p \in \left[ \frac{1}{2}, 1 \right],
\]

so their signals about the politician’s effort are noisy. Then each voter casts a vote (yes or no), and if there are two or more votes in favor of the politician, he is reelected and gets utility 1; otherwise, he is not reelected. Let us assume that voter \( i \) uses the following simple voting strategy: he votes yes if and only if \( s_i = 1 \). This strategy is natural in the sense that the voter rewards the politician for working on the policy issue that he cares about. (In Section 1.5 below, I prove that this is indeed an equilibrium strategy profile.)

We now consider, how the politician’s effort depends on \( p \), which captures the level of transparency in the environment. Start with the case \( p = \frac{1}{2} \). In this case, the probability of reelection equals \( \frac{1}{2} \) due to symmetry of signals and simple majority rule, and does not depend on the politician’s effort. Since effort is costly, the principal will choose
If \( p = \frac{1}{2} \), the outcome is different. The voters get perfectly precise signals, and the politician will get reelected if and only if he exerts effort in two or three dimensions. Since effort is costly, he will never work on all three dimensions, but at the same time, since the cost \( c \) is low enough, he will work rather than shirk. In this case, he will choose to work on any two dimensions, which is strictly better for the society than in the case \( p = \frac{1}{2} \).

But let us consider the intermediate case, say, with \( p = \frac{3}{4} \). The probabilities of reelection, as a function of the number of dimensions where the politician exerts effort, are given in the table below.

<table>
<thead>
<tr>
<th>( e_1 + e_2 + e_3 )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Pr(\text{reelection}) )</td>
<td>( \frac{5}{32} )</td>
<td>( \frac{11}{32} )</td>
<td>( \frac{21}{32} )</td>
<td>( \frac{27}{32} )</td>
</tr>
</tbody>
</table>

For example, if the politician chooses effort \((1, 1, 1)\), the probability of reelection is \( \left(\frac{3}{4}\right)^3 + 3 \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right) = \frac{27}{32} \). Given the effort cost \( c = \frac{1}{8} \), one can easily see that \( e = (1, 1, 1) \) will indeed be the optimal choice of the politician. We therefore see a non-monotonic dependency of politician’s effort, as well as of social welfare (for social welfare, more effort is better in this example, as the benefit of effort is \( 1 \) and the cost is \( \frac{1}{8} \)), on the degree of transparency \( p \). More precisely, the politician will choose \( e = (1, 1, 1) \) for \( 0.635 < p < 0.919 \); he will choose \( e = (0, 0, 0) \) for \( 0.5 < p < 0.628 \), and he will choose \( e = (1, 1, 0) \) (or working for some other two dimensions) in all other cases.

To understand the reasons for why transparency may hurt in this example, consider the case where voters do not have conflict of interests, and each acts to maximize their total welfare given by \( \pi_1 + \pi_2 + \pi_3 \). In that case, even with \( p = 1 \) they could mimic the behavior of players in the case of \( p = \frac{3}{4} \); it would suffice for each of them to support the politician with probability \( \frac{3}{4} \) if he received a positive signal, and with probability \( \frac{1}{4} \) otherwise. As we saw above, this would implement the first best. When players are self-interested, such cooperation would be impossible, as each voter would prefer to switch to the simple voting strategy (support the politician if and only if he received a positive signal), and the reason is that such voter would be more likely to receive the effort of politician,
as he is more responsive and does not randomize. This illustrates the main trade-off between accountability and preventing targeted policies. With too precise signals, the politician is able to target some of the voters while totally ignoring the rest. Some degree of uncertainty, obtained by lack of transparency, may force the politician to conduct a more balanced policy. If policy space were one-dimensional, such effect would obviously not arise.

1.3.2 Public Good Provision

Let us now modify the example to demonstrate, how conflict of interests among voters may affect the provision of public goods by the politician. In addition to the three policy issues introduced above, there is a public policy issue, about which all voters care equally. This policy issue, indexed by 0, costs the politician \( c = \frac{1}{5} \) (as the other policy issues), but each voter has utility function given by \( \pi_i + \frac{6}{5} \pi_0 \), i.e., each of them wants the public good, and cares about it more than about the private good \( i \). In particular, all voters would prefer to have \( \pi_0 = 1 \) and \( \pi_1 = \pi_2 = \pi_3 = 0 \) to \( \pi_0 = 0 \) and \( \pi_1 = \pi_2 = \pi_3 = 1 \). Voters have the following strategies: in the beginning, each of them announces the conditions under which he will support the politician and under which he will vote against him; assume the voters have the power to commit. The politician learns these announcements and then chooses his effort.

Assume that voter \( i \) gets a perfectly precise signal about \( \pi_i \) (so \( s_i = \pi_i = e_i \)), but learns the correct value of \( \pi_0 \) (signal \( s_i^0 \)) with probability \( \frac{3}{4} \) only. With these precisions, the voters, if they could cooperate, would easily induce the politician to choose \( \pi_0 = 1 \): voter \( i \) would simply have to promise to support the politician if \( s_i^0 = 1 \). The politician then would have no incentive to choose \( e_1, e_2, e_3 \neq 0 \), but he would choose \( e_0 = 1 \) instead of \( e_0 = 0 \), since the probabilities of winning are then \( \frac{27}{32} \) and \( \frac{5}{32} \), respectively.

However, playing these strategies is not an equilibrium. Take voter 1, and suppose that he deviates to the following strategy: support the politician if and only if the politician chose \( \pi_i = 1 \) (recall that he has perfect information on that). The politician will then choose \( e_0 = e_1 = 1, e_2 = e_3 = 0 \): indeed, the probability of reelection if he provides the
public good only is $\frac{9}{16}$, while by choosing $e_1 = 1$, too, he increases it to $\frac{15}{16}$. This deviation is clearly profitable to voter 1, so the above strategy profile cannot be sustained as an equilibrium.

Now let us consider the other extreme: the strategy profile where voter $i$ promises to support the politician if and only if $\pi_i = 1$; in this profile, voters completely ignore the public good. It is easy to see that it is an equilibrium. In this strategy profile, the politician chooses any two of the three voters, and works on the issues that these two workers care about. Now take any voter, say, voter 1. If he deviates to any other strategy and puts non-zero weight on the public good, the only thing he will achieve is that the politician will no longer be indifferent, which two voters to work for. Namely, he will work for voters 2 and 3 as he will know that this will buy their votes for sure, while working for voter 1, which costs the same (regardless of whether this means choosing $e_0 = 1$ or $e_1 = 1$), will still leave some probability that voter 1 will vote against him. Consequently, this is not a profitable deviation for voter 1. We have thus found an equilibrium, in which the voters fail to induce the politician to provide the public good, and instead try to influence the politician to work on the policy issue that only they like. One can verify that this is the only symmetric equilibrium, provided that costs of effort are stochastic (this is a technical condition which is formalized and discussed below).

This example illustrates that even if all voters have the same preferences about the optimal level of public good, they may fail to provide the proper incentives to the politician. The reason is that voters who pay attention to the amount of public good when voting are in a disadvantage, as the politician has fewer incentives to satisfy their private interests. The competition between the voters then makes all of them choose to ignore the public good in their voting decision; as a result, the politician has no incentive to provide the public good.

### 1.4 Formal Model

There are $n$ voters $1, \ldots, n \in V$ (indexed by $i$) and a politician $P$. Throughout the paper, voters may be interpreted as individuals, or as interest groups of equal size that have
solved any conflicts within themselves. Policy space $\mathcal{P}$ is $k$-dimensional (indexed by $j$), and in each dimension there are two possible outcomes, 0 and 1. The natural interpretation for 0 and 1 is status-quo and reform, respectively, but these outcomes may simply stay for two possible outcomes of the reform. The policy outcome in dimension $j \in \mathcal{P}$ is therefore $\pi_j \in \{0, 1\}$, and these values form the policy outcome vector $\pi = (\pi_1, \pi_2, \ldots, \pi_k) \in \{0, 1\}^k$.

To allow for the possibility that the politician has imperfect control of the policy outcome, I assume that he chooses an effort vector $e$ rather than policy outcomes $\pi$ directly. Vector $e$ has $k$ dimensions, corresponding to policy dimensions. Each component $e_j, j \in \{1, 2, \ldots, k\}$ is binary, and can have two values, 0 and 1. Assume that choosing effort $e_j = 1$ increases the chance that the realized policy outcome $\pi_j$ will equal 1. More precisely,

$$\Pr (\pi_j = 1 \mid e_j = 1) = \Pr (\pi_j = 0 \mid e_j = 0) = q_j,$$

where $\frac{1}{2} < q_j \leq 1$ for each $q_j$. Naturally, a policy issue $j$ such that $q_j = 1$ may be interpreted as an issue that the politician has complete control of, such as taking a position in some debate; e.g., a politician has full control over his position about abortions or guns. Similarly, a smaller $q_j$ may correspond to a reform which may fail even if the politician tries hard, such as a reform of healthcare or education. Effort $e_j = 0$ is normalized to be costless to the politician, while effort $e_j = 1$ costs $c_j$. I allow for the possibility that $c_j$ is positive or negative; the latter may correspond to the case where the politician likes outcome $\pi_j = 0$ better than $\pi_j = 1$. Denote $c = (c_1, c_2, \ldots, c_k)$. The cost vector $c$ is assumed to be stochastic (with continuous density and full support) from the voters’ viewpoint, but the politician knows the exact realization of $c$. This assumption is made to make sure that all possible effort vectors will be chosen and all possible policy outcomes will be realized with a positive probability along the equilibrium path. This will facilitate the characterization of optimal voting rules, as this will ensure that a small deviation by any of the voters will have some impact on the politician’s behavior for some cost vector for any strategies played by other voters.

In addition to the costs of efforts, the politician also values being reelected on the elections that happen in the end of the game. I normalize his utility from losing the
elections to 0 and utility from winning the elections to 1. The politician’s utility is, therefore,

\[ U_P = \mathbb{E} \left( I_P \text{ is reelected} - \sum_{j=1}^{k} c_j e_j \right). \]  

(1.2)

In order to get reelected, the politician needs to get the support of a certain share of voters \( \mu; \mu = \frac{1}{2} \) corresponds to a simple majority rule, \( \mu = 1 \) corresponds to a unanimity rule, etc.

Voter \( i \) cares about policy issue \( j = d(i) \) and does not care about other policy issues (this last assumption will be relaxed in Section 1.9). More precisely, his utility is given by

\[ \Pr (\pi_{d(i)} = b_i), \]  

(1.3)

where \( b_i \) is the outcome of policy \( d(i) \) that the voter likes best. I allow for the possibility that many voters care about the same policy issue (so \( d(i) = d(i') \) for \( i \neq i' \)), and they may have similar or opposite preferences. After the policy outcome is realized, but before the elections take place, each voter \( i \) gets a (noisy) binary signal about the policy issue that he cares about. Denote this signal by \( s_i \) and assume that

\[ \Pr (s_i = 1 \mid \pi_{d(i)} = 1) = \Pr (s_i = 0 \mid \pi_{d(i)} = 0) = p_i, \]  

(1.4)

where \( \frac{1}{2} < p_i \leq 1 \) for all voters \( i \). The natural interpretation of \( p_i \) is informativeness of signal that voter \( i \) gets, or information transparency. Furthermore, all signals are independent, conditional on the realization of \( \pi \).

After observing the signal, voter \( i \) casts one of two possible votes, \( \text{no} \) or \( \text{yes} \). At the time of voting, each voter is indifferent between these two actions. Following the standard retrospective voting models (as in Barro, 1973), I assume that in the beginning of the game, each voter \( i \) announces a voting strategy, which is a mapping from the information set available to him by the time of voting (i.e., the set of signals \( S_i = \{0,1\} \)) to \( \Delta \), which is a probability distribution over the set \( \{\text{no, yes}\} \). Since voters are indifferent between the two voting actions, it is incentive compatible for the voters to fulfil their announcements.
Let us denote the probability that voter $i$ votes yes after receiving signal $s_i$ by $M_i(s_i)$. 

The timing of the game is as follows.

1. Each voter $i$ announces, to the politician, voting strategy $M_i$. The announcements are made simultaneously.

2. Politician $P$ observes the vector of costs $c$ and chooses the multidimensional effort $e$.

3. Policy outcome vector $\pi$ is realized, each voter $i$ gets a noisy signal $s_i$.

4. The elections take place. Each voter automatically casts vote $v_i = yes$ with probability $M_i(s_i)$ and vote $v_i = no$ with probability $1 - M_i(s_i)$.

5. If $|\{i \in V : v_i = yes\}| \geq \mu n$, then politician $P$ stays in office, otherwise he is fired. Politician gets his payoff.

6. Voters finally learn the true realization of $\pi$ and get their payoffs.

The timing implies that the model involves retrospective voting, as in Barro (1973) and Ferejohn (1986). This is one way of modeling political accountability. Implicitly, the assumption is that the voters use the politician’s past performance to judge his future actions and, as a result, they support him if he performed well, and punish him otherwise. Assuming that voters have commitment power, in the sense that they can commit to a certain voting strategy, makes the model simple and tractable. As I show, in many interesting cases, the equilibrium voting strategies are simple and natural, which further justifies the retrospective voting approach.

### 1.5 Equilibria

The equilibrium notion I will use here is Trembling-Hand Perfect equilibrium. There are two stages in the game (voters announce voting strategies, and then the politician chooses effort; the voting itself is assumed to be automatic). However, since each agent acts only once, the standard definition is applicable. In a THPE, as opposed to a SPE, each
voter considers himself pivotal with a positive probability. This allows to refine “herding” equilibria where, for example, all voters promise to vote against the politician for any signals they get.

For any policy issue $j$, let $A_j$ denote the set of voters who care about this issue:

$$A_j = \{ i \in \mathcal{V} : d(i) = j \}.$$  \hspace{1cm} (1.5)

Without loss of generality, $A_j \neq \emptyset$ for any $j$ (we could otherwise ignore the existence of this policy issue). We get the following equilibrium characterization result. All proofs are in the Appendix.

**Theorem 1** Suppose that for each policy $j$, $A_j$ is a singleton (there is one voter, or interest group, per policy issue). Then there exists a unique THPE, in which each voter chooses $M_i(s_i = b_i) = 1$, $M_i(s_i \neq b_i) = 0$.

Theorem 1 suggests that when only one voter, or interest group, cares about each policy issue, then each voter uses a simple and natural voting strategy. The voter decides to support the politician if and only if he got a signal which is more likely if the politician did what the voter wanted him to do, and to vote against the politician otherwise. Indeed, from (1.1) and (1.4) it follows that

$$\Pr(s_i = b_i \mid e_{d(i)} = b_i) > \Pr(s_i = b_i \mid e_{d(i)} \neq b_i).$$  \hspace{1cm} (1.6)

The intuition for this result is that voter $i$ is the only person who can provide incentives for the politician to exert a certain level of effort on policy issue $d(i)$; if voter $i$ does not do this, nobody will. This enables voter $i$ to reward the politician as strongly as possible (support him with probability 1) if he gets a positive signal about the politician’s performance, and punish him as much as he can otherwise. This is a powerful result, which is especially well applicable to the case where voters represent interest groups.

It would seem natural for Theorem 1 to hold even if there are multiple voters who care about the same policy issue. It turns out, however, that this is not the case. Below,
I provide Example 1.5, where some voters care about the same policy dimension and, moreover, prefer the same outcome, but, nevertheless, simple voting strategies do not form an equilibrium. The intuition of this result is that by voting against the politician in case of a negative signal, the voter hurts a working politician more than a shirking one, because of the different probabilities of him being pivotal in the cases the politician works and shirks. Example 1.5, in particular, shows that the “swing voter’s curse” effect (Feddersen and Pesendorfer, 1996) may appear in retrospective voting models, too.

However, if one introduces more assumptions about the precision of signals or distribution of costs, it is possible to obtain a tight characterization of equilibrium even if many voters care about the same policy issue, which makes Theorem 1 inapplicable.

I now relax the assumption that only one voter may care about each policy issue, and also allow for the possibility that voters observe signals of other voters. Consider the vector of all signals $s = (s_1, s_2, \ldots, s_n)$. For any subset $F \subset V$, denote its projection on the coordinates corresponding to voters in $F$ by $s|_F$. Now assume that, in addition to his own signal $s_i$ about policy issue $d(i)$, voter $i$ can observe the signals of some other voters; denote the set of these voters by $\mathcal{F}_i$. There are multiple reasons for why this may be the case: a voter may observe the well-being of his friends, or read the same newspapers as other voters, or if some voters have the same preferences, they will have a strategic reason to exchange information. In this paper, introducing the possibility of voters observing each other’s signals allows characterization of equilibria in cases where multiple voters care about the same policy issue; this will be important when studying polarization (Section 1.7) and delegation in voting (Section 1.8).

Naturally, $i \in \mathcal{F}_i$ for any voter $i$. Under this assumption, the voting strategy of each voter, which he announces in the beginning of the game, is now a function $M_i$ from $S|_{\mathcal{F}_i} = \{s|_{\mathcal{F}_i}\} = \{0, 1\}^{\mid\mathcal{F}_i\mid}$ to $\Delta$. The following characterization result holds when the signals that voters get carry almost full information about the actions of the politician.

**Theorem 2** Suppose that each voter gets all available information about the policy issue he cares about: $A_{d(i)} \subset \mathcal{F}_i$ for each $i \in V$. Then there exists $p^* < 1$ such that if for each voter $i$ we have $p_i > p^*$, and for all policy issues $q_j > p^*$, then there exists a unique THPE
where voter $i$ votes yes if the signal $s|_{\mathcal{F}_i}$ that he gets satisfies

\[ \Pr \left( s|_{A_{d(i)}} \mid \pi_{d(i)} = b_i \right) > \Pr \left( s|_{A_{d(i)}} \mid \pi_{d(i)} \neq b_i \right), \]

(1.7)

and votes no otherwise.

First of all, note that the voting rule is well-defined: since for each $i$, $A_{d(i)} \subset \mathcal{F}_i$, then signals $s|_{A_{d(i)}}$ are known to voter $i$. The voting rule given by (1.7) voting rule is a generalization of the simple voting strategy from Theorem 1. The condition (1.7) prescribes voter $i$ to support the politician if the signals he observed are more likely if the politician did what voter $i$ wanted him to do rather than if he did not. Under the conditions of Theorem 2, voters pay no attention to signals about policy outcomes which they do not care about even if they observe those signals.

The intuition of this result is the following. Suppose first that voter $i$ observes the full profile $s$ of signals. If all $p_i$ and all $q_j$ are sufficiently high, he can be almost certain about the action of the politician that generated signal $s$. Similarly, from the politician’s viewpoint, the likelihood that the voters get signal $s$ such that condition (1.7) holds is much higher if he chooses $e_{d(i)} = b(i)$ rather than $e_{d(i)} \neq b(i)$. This implies that if voter $i$ did not follow the voting strategy specified in Theorem 2, a small deviation towards this strategy would provide the politician with stronger incentives to do what voter $i$ wants him to do, as this deviation will reward the working politician stronger than the shirking one (and, likewise, punish the shirking politician stronger). The equilibrium characterization immediately follows. If, however, voter $i$ does not observe the signals of all other voters, the intuition is as follows. For any combination of signals that $i$ does not observe, $s|_{\mathcal{V}\setminus\mathcal{F}_i}$, following the above-defined rule is optimal. Hence, it is optimal even if voter $i$ does not know the realization of $s|_{\mathcal{V}\setminus\mathcal{F}_i}$. This reasoning establishes that there is a unique THPE (the only non-uniqueness may be the case because the politician may choose different effort vectors when costs are such that he is indifferent, but this happens with probability 0).

In Section 1.8, I consider a “symmetric” case, where there is an equal number of voters...
caring about each policy issue, and the politician’s costs of working on each project are not very different. This would be another case where the equilibrium of the game is simple and natural (and given by \((1.7)\)). The next example shows, however, that these natural strategies do not always form an equilibrium.

Suppose that there are three voters and only one policy issue. The preferences of all voters are identical: each of them prefers \(\pi = 1\) to \(\pi = 0\). Each voter observes his signal only, and these signals are sufficiently precise; for instance, \(p_i = \frac{3}{4}\) for each voter \(i\). Furthermore, \(q = 1\), i.e., the politician has a perfect control over the policy outcome. The politician prefers \(\pi = 0\), and implementing \(\pi = 1\) costs him \(c = \frac{7}{16}\). As always, politician’s utility from reelection is normalized to 1.

Consider unanimity voting rule. Let us show that the simple voting rule, where each voter supports the politician if and only if he gets a positive signal \(s_i = 1\), does not form an equilibrium. Indeed, in that case, the politician is reelected with probability \((\frac{3}{4})^3 = \frac{27}{64}\) if he works and with probability \((\frac{1}{4})^3 = \frac{1}{64}\) if he shirks. Since \(\frac{27}{64} - \frac{1}{64} < \frac{7}{16} = c\), the politician will find it optimal to shirk.

Now take voter 1, and consider his deviation to strategy “always vote for the politician”. In that case, the politician will need to get the support of two other voters; he is reelected with probability \((\frac{3}{4})^2 = \frac{9}{16}\) if he works and with probability \((\frac{1}{4})^2 = \frac{1}{16}\) if he shirks. The difference is now \(\frac{1}{2} > \frac{7}{16} = c\), so the politician will work. This is clearly a profitable deviation for the voter. Consequently, the simple voting strategy is not an equilibrium in this case.

The intuition for this result is the following. By choosing the simple voting strategy, i.e., to reward the politician for a good signal and to punish him for a bad signal, a voter tries to provide proper incentives. In this particular case, however, voter 1 runs into the following problem: given the strategies of voters 2 and 3, he is much more likely to be pivotal if the politician works (he is then pivotal with probability \(\frac{9}{16}\)) than if he shirks (the probability is then \(\frac{1}{16}\)). As a result, by choosing the simple voting strategy, he is more likely to punish a working politician (the probability is \(\frac{9}{16} = \frac{9}{64}\))
than a shirking politician (he punishes him with probability $\frac{3}{4} \frac{1}{16} = \frac{3}{64}$). So, instead of inducing the politician to work harder, this voter ends up decreasing his incentives to work. Instead, deviating to the strategy “never punish the politician” improves the politician’s incentives, and for $c = \frac{7}{16}$, it makes the politician switch to working.

1.6 Effects of Transparency

1.6.1 Transparency and Accountability

In this section, I use the equilibrium characterization results obtained above to establish the non-monotonic effect of transparency (modeled as the precision of signals that voters get) on social welfare. Assume, for simplicity, that each policy issue concerns only one voter, and that voters observe their signals only. In this case, Theorem 1 is applicable, and it implies that voters will use the simple voting strategy (choose $M_i(s_i) = 1$ if and only if $s_i = b_i$) for any distribution of cost vector $c$. To simplify the formulation of results, I assume that vector $c$ satisfies $c_1 = c_2 = \cdots = c_n = \tilde{c}$ with probability 1; this value $\tilde{c}$, however, may be stochastic. Clearly, this distribution is a limit of distributions with continuous density and full support, so I will assume that simple voting strategies are played in this case as well. Finally, let $q_j = 1$ for all policy issues $j$.

Let us start with a simple observation: each voter would be better off if he had access to a more precise signal. Namely, consider a game identical to the one described in Section 1.4, with the exception that in the beginning of the game, i.e., before announcing voting strategies, all voters choose, simultaneously and independently, their precision parameters $p_i$ from a range of alternatives $p_i \in [p_L, p_H] \subset (1/2, 1]$. We get the following result.

**Proposition 1.1** If the voters are to choose their precision parameters $p_i$, each of them would choose $p_i = p_H$. This does not depend on whether their choices become known to other voters before voting strategies are announced.

Indeed, a voter cannot be worse off if he picks a higher precision $p_i$, holding the strategies of all other voters fixed. Indeed, the voter can always mimic his behavior if he picked a lower precision, which means that a higher precision is at least weakly better
than a lower precision. As it turns out, a higher precision is strictly better, because the politician, when indifferent or close to indifferent between working on the policy issues favored by two voters, will be more likely to choose the more responsive of them, i.e., the one who has a higher precision of signal. It is worth noting that, due to strictly monotonic preferences of each voter over the precision of his signal, the voters would be willing to pay a non-zero amount of money to increase their signals’ precisions.

In the light of the above discussion, it is surprising that more precise signals may make the society as a whole worse off. By social welfare, I mean the sum of the utilities of all voters, and not of the politician: in a dynamic version of the game, if the politician is not reelected, some other politician would enjoy the same benefits of being in power. Formally, we have the following result.

**Proposition 1.2** Suppose that \( p_i = p \) for all voters \( i \), and cost \( \tilde{c} \) has positive density on \([0, \varepsilon]\) for some \( \varepsilon > 0 \). Then the social welfare is monotonically increasing in \( p \) if \( p \) is close to \( \frac{1}{2} \) and monotonically decreasing in \( p \) if \( p \) is close to 1, provided that \( m \neq 1 \), where \( m = \lfloor \mu n \rfloor \), the greatest integer not larger than \( \mu n \). In particular, social welfare is maximized at an interior value of precision \( p \).

Moreover, if cost \( \tilde{c} \) is smaller than \( \varepsilon \) with probability 1 for \( \varepsilon \) sufficiently small, then social welfare is a single-peaked function of \( p \).

Proposition 1.2 is central in the paper. It establishes a trade-off between providing incentives for the politician to exert effort, and preventing him from targeting a smaller group of voters; this trade-off implies non-monotonic effect of transparency. The intuition of this proposition is best understood with the help of the following consideration. Even though voters care about different policy issues, they impose negative externalities on one another in the following way. By providing stronger incentives for the politician to work on his project, each voter reduces the number of votes of other voters that the politician needs to get reelected. Hence, the prisoners’ dilemma intuition holds: even though it is individually rational for each voter to get a more precise signal, this may hurt the society overall. When \( p \) is small, this negative externality effect is weak. However, when \( p \) is high, this effect dominates, and makes \( p = 1 \) suboptimal for the society.
Whenever $p$ is exactly $\frac{1}{2}$, the politician will never exert any effort as long as $\tilde{c} > 0$. For higher $p$, he will be willing to exert effort for positive values of $\tilde{c}$, as this will yield a non-trivial reward in terms of a higher probability of getting reelected. For this reason, $p = \frac{1}{2}$ can never be optimal. But at the same time, $p = 1$ would imply that the politician would choose to work for the minimal needed number of voters only (which equals $m$), except for the cases where $\tilde{c}$ is high enough and therefore the politician will shirk on all dimensions. For a slightly lower $p$, there would be a range of values of $\tilde{c}$ for which the politician will exert effort on more than $m$ dimensions, because he would want to safeguard himself against some of the voters he works for getting wrong signals.

The result of Proposition 1.2 has a number of important implications. First, it suggests a trade-off between holding the politician accountable to voters in general (which implies the necessity of a higher precision of signals, so that shirking is detected), and allowing the politician to target specific voters at the expense of other voters. Consequently, the idea that transparency is good has its limitations: some transparency is better than no transparency, but a fully transparent environment is optimal only when voters have aligned preferences. Second, this result implies the optimality of some rules or institutions which restrict the voters’ ability to obtain information about the performance of the politician: for example, this is another argument for politician’s right to privacy. But the result also highlights that symmetry is important: if only some group of voters is denied access to full information, the other voters would get an unfair advantage, and the politician will exert a suboptimal level of effort, targeting the voters who obtain more precise information. Third, the intuition behind Proposition 1.2 implies a trade-off between ex-post efficient elections (rewarding the politician who performs well and punishing the one who does not) and providing the proper incentives to the politician while he is in office. It is generally thought (e.g., Feddersen and Pesendorfer, 1998b) that having more informed voters is better for the society, as this reduces noise in the elections, which is considered good for the incentives. But Proposition 1.2 suggests that it may be actually good to have imperfectly informed voters, and for precisely the same reason: that they introduce noise, which makes the politician work harder. This noise prevents the politician from targeting...
specific groups, and this benefits the society overall.

1.6.2 Optimal Degree of Supermajority

To show that the non-monotonic effect of transparency does not disappear as the number of players and policy issues goes to infinity, I now consider the limit case, with a continuum of voters such that each cares about his own policy issue. This will also allow the characterization of the optimal degree of supermajority (or submajority): when it is better to have a simple majority rule, and when it is better to have a unanimity rule.

Suppose that there is a continuum of voters and a continuum of policy issues. Since for any finite number of voters the equilibrium involves voters playing the simple voting strategy, assume that for the continuum of voters, again, each voter \( i \) supports the politician (votes \( v_i = yes \)) if and only if \( s_i = b_i \). Furthermore, suppose that the value of cost of politician’s effort, \( \tilde{c} \), is fixed. In the next result, I will take advantage both of the simple voting strategies, following from Theorem 1, and of the simpler algebra due to a continuum of voters.

The choice of the politician is the share of voters whose preferences he is going to satisfy (from his viewpoint, all voters are identical); denote this share by \( |e| \). By the law of large numbers, the share of voters who get their preferred signal \( s_i = b_i \) is then \( p|e| + (1-p)(1-|e|) \). Consequently, in order to get \( \mu \geq \frac{1}{2} \) votes (the case where a submajority is needed is considered below), the politician needs to choose

\[
|e| \geq e^* = \frac{p + \mu - 1}{2p - 1} = 1 - \frac{p - \mu}{2p - 1},
\]

and this immediately shows that in order for the politician to have a chance of reelection, \( p \geq \mu \) must hold. Assuming that this is the case, the actual choice of the politician is as follows. If \( \tilde{c} \leq 0 \), then the politician will simply choose to fulfil policy preferences of all voters, as these coincide with his own preferences and guarantee reelection. If \( \tilde{c} > 0 \), then the politician will have a non-trivial choice between ensuring reelection, at a cost, and shirking and losing the elections. This is summarized in the following Proposition.
**Proposition 1.3** Suppose $\mu \geq \frac{1}{2}$. Then:

(i) The share of voters for whom the politician works, $|e|$, is given by

$$
|e| = \begin{cases} 
1 & \text{if } \tilde{c} \leq 0 \\
\frac{p+\mu-1}{2p-1} & \text{if } 0 < \tilde{c} \leq \frac{2p-1}{p+\mu-1} \\
0 & \text{if } \tilde{c} > \frac{2p-1}{p+\mu-1}
\end{cases}
$$

and is a (weakly) decreasing function of cost $\tilde{c}$.

(ii) If $0 < \tilde{c} < 1$, then $|e|$ is a nonmonotonic function of precision of signal $p$: it equals 0 for $\frac{1}{2} < p < \mu$, and $\frac{p+\mu-1}{2p-1}$ for $p > \mu$, which is decreasing in $p$. The maximum is obtained for $p = \mu$, and is equal to 1. If $1 < \tilde{c} < \frac{1}{\mu}$, then the function equals 0 for $\frac{1}{2} < p < \frac{p-\mu}{2-\mu}$, which is a non-empty interval, and monotonically decreasing for $p > \frac{p-\mu}{2-\mu}$, with maximum equal to $\frac{1}{\tilde{c}}$, which is obtained at $p = \frac{1-\tilde{c}(1-\mu)}{2-\tilde{c}}$. If $\tilde{c} > \frac{1}{\mu}$, the function is a constant 0.

(iii) The optimal precision $p$ is given by (for $0 < \tilde{c} < \frac{1}{\mu}$)

$$
p^{\text{opt}}(\tilde{c}) = \max \left( \mu, \frac{1-\tilde{c}(1-\mu)}{2-\tilde{c}} \right),
$$

and is weakly increasing in $\tilde{c}$. Any $p$ is optimal for other values of $\tilde{c}$.

From Proposition 1.3, it follows that the politician is less willing to work as the cost of effort increases. However, when effort is not too costly, it is suboptimal to monitor the politician with too much precision, for in that case he would only choose $|e| = \mu$, i.e., precisely the effort needed to get reelected. Interestingly, the optimal level of precision is nondecreasing in cost $\tilde{c}$; the reason for that is that if cost increases, but $p$ stays the same, the politician finds it too costly to work for the same number of voters and prefers to shirk. To combat this effect, more responsiveness by the voters, and thus more transparency, is needed.

In the case of a submajority rule ($\mu < \frac{1}{2}$), the following result holds.

**Proposition 1.4** Suppose $\mu < \frac{1}{2}$. Then:

(i) The share of voters for whom the politician works, $|e|$, is given by (1.8) if $\mu > 1 - p$
and by

\[ |e| = \begin{cases} 
1 & \text{if } \tilde{c} \leq 0 \\
0 & \text{if } \tilde{c} > 0 
\end{cases} \]

if \( \mu \leq 1 - p \).

(ii) For \( 0 < \tilde{c} < \frac{1}{\mu} \), the politician’s effort \(|e|\) equals 0 for \( \frac{1}{2} < p < 1 - \mu \) and \( \frac{\mu - 1}{2p - 1} \) otherwise.

(iii) The optimal precision \( p \) is 1; in that case, the politician’s effort is \( \mu \), provided that \( 0 < \tilde{c} < \frac{1}{\mu} \), and 0 if \( \tilde{c} > \frac{1}{\mu} \).

In the case of submajority, more noise makes the politician willing to shirk, because there will be a sufficient number of supporters anyway. Consequently, a more transparent environment is strictly better in this case.

We can now use the results of Propositions 1.3 and 1.4 to find out the optimal degree of (super)majority \( \mu \). A low \( \mu \) is good to induce the politician to choose a non-zero effort, but if it is less than \( \frac{1}{2} \), it also sets an upper bound on the politician’s effort. A low \( \mu \) would therefore be preferred in environments where cost \( \tilde{c} \) is relatively high, provided that it is accompanied by precise monitoring. At the same time, if cost of effort \( \tilde{c} \) is low (for instance, if the politician is close to indifferent between two positions, or if the cost is paid by some third party), it is better to have a high \( \mu \), but less than perfect monitoring. This is summarized in the next proposition.

**Proposition 1.5** Suppose that \( \tilde{c} \) is a singleton, and the society wants to choose both the degree of supermajority \( \mu \) and the transparency \( p \) optimally. Then:

(i) If \( \tilde{c} < 0 \), the politician is benevolent, and the choice of \( \mu \) and \( p \) has no effect.

(ii) If \( 0 < \tilde{c} < 2 \), then politician’s effort is maximized (and equals \( \min \left( 1, \frac{1}{\tilde{c}} \right) \)) if the society chooses \( \mu = p \) to be any number between \( \frac{1}{2} \) and \( \frac{1}{\tilde{c}} \). In particular, for \( 1 < \tilde{c} < 2 \), it is never optimal to have full transparency.

(iii) If \( \tilde{c} > 2 \), a submajority rule works best. The society must choose \( \mu = \frac{1}{\tilde{c}} \) and \( p = 1 \), i.e., full transparency.

We see therefore that even in the case with a continuum of voters (and thus no aggre-
gate uncertainty), choosing less than full transparency may help even if the society can also vary the degree of (super)majority. High transparency may hurt the society when the cost of effort is intermediate: for low cost, the society may want to adjust the majority parameter $\mu$ instead of decreasing $p$, while for extremely high cost, lowering $p$ will shy the politician away from such costly and non-responsive voters. The comparative statics on the optimal degree of majority is more straightforward and quite intuitive: the more costly it is for the politician to exert effort, the less votes he must get in order to get reelected.

If the society is unable to choose $p$, then the decision on optimal $\mu$ is also easy to derive. If the cost of effort is not too high, then it is optimal to choose $\mu = p$: a lower $\mu$ will provide suboptimal incentives to the politician, while a higher $\mu$ will be impossible to achieve, and therefore the politician will shirk. If, however, the cost of effort is very high, a submajority rule will be optimal; but in either case, the optimal degree of majority is increasing in $p$.

To summarize, a higher degree of supermajority is more likely to be optimal if the cost of politician’s effort is small, and the environment is transparent. Otherwise, a simple majority or a submajority rule may be optimal. This intuition generalizes for the case where $\tilde{c}$ is stochastic, though there will be a trade-off between providing high incentives when $\tilde{c}$ is small and risking that the politician will give up and choose zero effort if $\tilde{c}$ is high.

1.7 Polarization

In this section, I investigate the question, whether and when the politician will be more accountable to the citizens if their preferences become more polarized on some dimensions. This analysis complements the one by Besley (2005) for the case where there are multiple policy issues. There, the conclusion is that higher polarization decreases the ability of voters to induce the politician to exert effort in the dimension all voters care about equally; as a result, higher polarization decreases the utility of non-partisan voters. Here, I model polarization as disagreement of voters who care about the same policy issue, rather than
the intensity of their preferences. I show that in this case, higher polarization of voters who care about some policy dimensions will actually benefit the voters who care about different issues. The intuition for that is that the politician will typically prefer to gain support of voters who care about uncontroversial policy issues rather than issues over which the society is highly polarized, because taking a different position on the latter issues will make the politician lose many supporters as well.

More precisely, assume again that there is a continuum of policy issues, and each policy issue concerns a continuum of voters (so there is a continuum of voters totally). Assume that the politician has perfect control over the policy implementation ($q_j = 1$ for each policy $j$), and that voters’ signals are perfectly precise ($p_i = 1$). Furthermore, there is no asymmetry of information between voters, which makes Theorem 2 applicable. In this particular case, each voter will use the simple voting strategy ($M_i = 1$ if and only if $s_i = b_i$) because his signal is perfectly precise.

All policy dimensions are divided into two groups. One group (share $\gamma$) is about spending, staying uncorrupt, or any other issue where the interests of the politician and the citizens are likely to be opposite; call such issues “financial”. Orient such policy dimensions so that voters prefer outcome $\pi_j = 1$, while the politician prefers outcome $\pi_j = 0$. The other group of dimensions (share $1 - \gamma$) are such that there is no built-in conflict between the politician and the voters, and thus the politician is as likely to prefer a particular policy outcome as a random voter. Such issues may include abortions, gun control, etc; call these issues “position-taking” issues. Again, orient such policy dimensions so that a majority of voters who care prefer $\pi_j = 1$, but this majority is not absolute as in the previous case, but has share of voters $\alpha \geq \frac{1}{2}$. Naturally, parameter $\alpha$ captures polarization: a high $\alpha$ (close to 1) implies that there is almost no disagreement about voters who care about a particular issue (although the rest of the population might not care), whereas $\alpha$ close to $\frac{1}{2}$ implies that there is little agreement between the voters. Indeed, a society where voters who care about an issue are split 50 : 50 is a polarized society, while a society where the distribution is 90 : 10 is not. The politician prefers outcome $\pi_j = 1$ with probability $\alpha$ and outcome $\pi_j = 0$ with probability $1 - \alpha$; this
assumption means that the preferences of the politician resemble those of a random voter. Assume that working on each of the financial dimensions costs the politician $c_f > 0$, while exerting effort $e_j = 1$ on position-taking dimensions costs him $c_p > 0$ if he prefers $\pi_j = 0$ and $-c_p$ if he prefers $\pi_j = 1$. Both $c_f$ and $c_p$ are assumed to be sufficiently low, and $\mu$ not too high, so that the politician can get reelected and will prefer to do so.

Let us see, how the politician’s and voters’ welfare changes as polarization changes, and how this depends on the incumbency advantage $\mu$ and the share of financial policy issues $\gamma$. Consider two cases. If $c_p > (2\alpha - 1) c_f$, then a vote of a person who cares about a financial policy issue is cheaper for the politician. Assuming that the share of such persons is sufficiently high, we find that the politician will first choose his favorite outcomes on the position-taking dimensions, which will bring him $(1 - \gamma) \left( \alpha^2 + (1 - \alpha)^2 \right)$ votes, and then choose effort 1 on share of $y$ financial policy dimensions, so that

\[(1 - \gamma) \left( \alpha^2 + (1 - \alpha)^2 \right) + \gamma y \geq \mu. \tag{1.9}\]

If, however, $(2\alpha - 1) c_p < c_f$, he will decide that it is cheaper to satisfy the majority of voters who care about position-taking policy issues, and only after exert effort on financial dimensions. In that case, he will choose $y$ to satisfy

\[(1 - \gamma) \alpha + \gamma y \geq \mu. \tag{1.10}\]

We get the following proposition.

**Proposition 1.6** Suppose $\gamma > \mu$. The total effort of the politician on the financial dimensions is decreasing in $\alpha$, so it is increasing in polarization along position-taking policy issues. The effect is stronger if there are many position-taking issues, i.e., $\gamma$ is small, and if cost of effort on these dimensions is relatively large.

This proposition suggests that higher polarization makes it harder for the politician to satisfy a large number of voters simply by taking his favorite positions. When there are many position-taking dimensions, this effect is stronger, even though the total share
of votes obtained in this way is larger. The same is true if \( c_p \) is sufficiently large: in that case, the politician will try to satisfy the majority in all of the policy-taking dimensions, which will make him more sensitive to increasing or decreasing polarization.

It is interesting to contrast this result with Besley (2005). There, higher polarization implies that citizens pay more attention to the partisan policy issue (which corresponds to position-taking issue in this paper), which allows the politician to ignore the non-partisan issue. I show that with strategic voters, this need not be the case. Here, higher polarization actually induces the politician to ignore the partisan issues, merely taking his preferred positions. He will instead make up for the voters by exerting more effort on the financial issues, which the voters agree upon.

### 1.8 Deliberation and Delegation

In this section, I discuss, whether or not it is optimal for voters with similar preferences to exchange information prior to voting. The intuition, drawing on the results from Section 1.6, is that exchanging information is always optimal for each particular group of voters. However, if voters are sufficiently well-informed, this exchange of information may make the society as a whole worse off, because each vote becomes effectively more informative, and this may have a negative effect on the politician’s effort. This intuition applies to the following related problem. Suppose that, instead of deliberating and exchanging information, the voters in each group have a majority voting, where they decide whether or not to approve the politician, and after that, a delegate of this group casts this vote on behalf of the entire group. Again, in the second case, votes are more informative, which is good if the information available to voters is imprecise, and bad if it is sufficiently precise. Note that the mechanism described above may be interpreted as a simplification of the presidential elections in the U.S. The same intuition implies that if voters are sufficiently well-informed, direct elections are better.

To proceed, we need an additional equilibrium characterization result. I argue that the natural voting strategies that form the equilibrium in Theorem 2 are also an equilibrium in a wider set of situations.
Consider the following “symmetric” case. Assume that \( n \) is a multiple of \( k \), and \( n/k \) voters care about each policy issue. For each policy issue \( j \), \( q_j = q \), and \( p_i = p \) for all voters. Assume that interests of voters who care about policy issue \( j \) are aligned, and all of them prefer \( b_j = 1 \). Furthermore, make the following assumption on the cost vector \( c \). Assume that for each \( i \), \( c_i = c' + \sigma \varepsilon_i \), where \( c' \) has positive density on \( (-\infty, +\infty) \), \( \varepsilon_i \) are i.i.d., independent from \( c' \), and have positive density on \( (-\infty, +\infty) \) and a single-peaked distribution with peak at 0, mean 0 and variance 1. If \( \sigma > 0 \) is sufficiently small, we get the following characterization result.

**Theorem 3** Suppose the assumptions above hold, and suppose \( \mathcal{F}_i = A_{d(i)} \) (so that each voter \( i \) gets signals from all voters who care about the same policy dimension, and only them). Then there exists \( \sigma^* > 0 \) such that if \( \sigma < \sigma^* \) then in the unique symmetric THPE, voter \( i \) votes yes if and only if condition (1.7) holds.

If, instead of equality \( \mathcal{F}_i = A_{d(i)} \), the inclusion \( \mathcal{F}_i \subset A_{d(i)} \) holds for all voters \( i \), then this strategy profile is an equilibrium, provided that all \( p_i \) are bounded away from 1, and the number of policy issues \( n \) and \( k \) is sufficiently large (for a given \( n/k \)).

This theorem establishes that if the projects are roughly equally costly to implement, then the competition for politician’s effort will make voters choose the most responsive voting strategies. Intuitively, if some voter attempts to deviate and play a different voting strategy, this will undermine the politician’s incentives to choose the policy favored by that voter. The reason is that the politician, given that the costs of implementing different policy issues are not too different, will always prefer to satisfy more responsive voters rather than less responsive ones – and responsiveness here is captured by whether or not a voter uses only relevant information, or sometimes conditions his vote on less relevant signals.

Let us now compare the following two cases. In one, voter \( i \) can observe his signal \( s_i \) only; this corresponds to no deliberation or other way of exchanging information. In the other, he can observe signals of all voters with similar preferences, \( s|_{A_{d(i)}} \); this is the case where voters can deliberate, or aggregate their information in some other way. In light of Theorem 3, in the first case, voters will implement the usual simple voting strategy, while
in the second one they will use the rule given by (1.7). The next proposition answers the question whether this information exchange is beneficial for the society.

**Proposition 1.7** Suppose the above assumptions hold, and that the politician needs the approval of more than half of the population ($\mu \geq \frac{1}{2}$). Then:

(i) Within each group, voters are better off if they exchange signals.

(ii) The society is better off from deliberation within each group if $p$ is sufficiently low.

(iii) Deliberation makes the society worse off if $p$ is sufficiently close to 1.

As follows from Proposition 1.7, deliberation increases the responsiveness of a given group of voters: it makes it more rewarding for the politician to work for this group, but at the same time more costly for him (in terms of the expected number of votes) to ignore this group. Any voting group is, other things equal, better off from exchanging information within itself, and this is what would happen in an equilibrium. However, this effectively increases the precision of the signal that each group possesses, and according to which it casts its votes. As we saw above, such an increase in the precision, while rational for each of the groups, may have a negative effect on social welfare if the voters already get sufficiently precise signals, as such aggregation only further induces the politician to work for the minimal number of special interests.

Proposition 1.7 may be interpreted to make the following comparison. Suppose that the politician’s task is to create a legislative proposal which may satisfy or not satisfy different interest groups. One can compare the following two mechanisms. One is having the proposal voted in narrowly specialized committees first, and then accepting it if sufficiently many committees approve the proposal. The other is direct voting over the proposal by all legislators. The last result suggests that when legislators are sufficiently poorly informed, it is better that the proposal be considered in committees first. However, when the legislators are sufficiently aware about the subject, direct voting will be preferred, for otherwise the politician will be too tempted to emphasize some dimensions of the proposal while sacrificing other ones.

This discussion has an interesting application to the issues of federalism, and to the design of the U.S. presidential elections. The very idea of decentralized government rests
on the heterogeneity of voters by geographical location, and hence one can interpret it as that voters living in different provinces, or states, have different preferences. Let us push this interpretation to the extreme, in that the voters in the same state have similar preferences, but those in different states care about different policy issues. We now have the following comparison. When the voters are sufficiently poorly informed, the aggregation of votes, and therefore of information, on the state level increases the politician’s accountability. This makes the electoral college system relatively more preferrable to direct voting. In contrast, if the voters are sufficiently well informed, their voting behavior becomes more predictable to the politician, and the electoral college system induces the politician to target a narrower share of the states while ignoring the states where he is going to lose. While this is by far not the only consideration in the discussion on whether or not electoral college is an outdated or optimal system, it allows a clear comparative statics: The advantage of the electoral college is higher when voters are poorly informed. In particular, if we assume that over time, with the spread of mass media and other means of communication the voters became more informed, the importance of the electoral college must have diminished. This confirms the idea that the Founding Fathers had: the voters may be imperfectly informed, but they can choose electors who are better informed. In fact, this paper shows that the electoral college may be a good solution even if the electors are not better informed than ordinary voters.

1.9 Public Goods and Lobbying

1.9.1 Public Goods Provision

So far, I have considered situations where each voter cares about only one policy issue. Here, I explore a more general case, where, in addition, there is a policy issue which all voters care about, such as provision of a public good. In other words, each voter cares about this public policy issue, as well as some “private” policy issue (I will use this term to distinguish policy dimensions that not all voters care about from the public policy issue).

The first result here is that the public good may be underprovided, for the reason that voters will not reward or punish the politician for his performance on the public policy
issue, and instead reward or punish him only according to their signals about private issues. The underprovision of public goods is a standard result in the literature (Olson, 1971); however, the reason here is different. In the standard models, individuals internalize the costs, but not the benefits of a public good, which induces them to contribute suboptimal amounts of resources for public good provision. However, with a simple voting mechanism one would expect an efficient level of public good (or at least the amount that the median voter prefers, if voters are heterogenous) to be provided, for a very simple reason: if public good is all the voters care about, they will have incentives to reelect the politician who provided at least as much as they want. As I show below, this reasoning fails if voters care about other policy issues as well: in that case, voters’ competition for politician’s effort on their private issues will prevent them from putting any weight on the information about the amount of public good provided by the politician.

Formally, suppose that each voter cares about two policy issues. First, each voter cares about a public project, where he wants the politician to exert effort; this project is the same for all voters. Second, each voter cares about his own private policy issue. For simplicity, assume that voters correspond to private policy issues on a one-to-one basis, so voter $i$ cares about policy issue $i$. Voter $i$ gets two signals: $s_i$ about private project $i$, and $s_i^0$ about the public project. He maximizes

$$u_i = \Pr(\pi_i = 1) + \phi \Pr(\pi_0 = 1),$$

where index 0 denotes the public project and $\phi > 0$ is the weight that voters attribute to the public project.

As before, I assume that all private projects have roughly equal cost for the politician (i.e., the assumptions required for Theorem 3 hold), and that the public project costs $c_0$, which also has a continuous positive density on $(-\infty, +\infty)$ and is distributed independently of $\{c_i\}$. Each voter gets equally precise signals $p$ about the public project, as well as about the private project he cares about. As before, assume that $[\mu n] < 1$, i.e., the politician does not need a unanimous support. Then the following result holds.
Proposition 1.8 Suppose the assumptions above hold. Then for any $\phi$ there exists $\sigma^* > 0$ such that if $\sigma < \sigma^*$ then the only symmetric THPE involves voter $i$ supporting the politician if and only if he got a positive signal ($s_i = 1$) about the outcome of the private project $i$, and voting against the politician if and only if he got a negative signal ($s_i = 0$) about the private project $i$. The signal about the public project is ignored in this equilibrium. The politician chooses $e_0 = 1$ if and only if $c_0 \leq 0$.

The result of Proposition 1.8 is striking at first glance. Naively, one could expect the voters to mix, by giving some weight to the public project and some weight to the private one they are interested in. However, under the assumptions above this is not the case. The intuition for this result is that from a voter’s perspective, it is not very likely that his decision to put more weight on (the signal about) the outcome of the public project will induce the politician to work on that project. But what is more likely to happen is that the politician will switch from working on that voter’s project to working on some other issue, which concerns a more responsive voter. In other words, the voters in equilibrium compete with one another over the private project, and this competition drives away their attention (and, in turn, politician’s effort) from the public project. For obvious reasons, this equilibrium will be inefficient, and if $\phi$ is sufficiently high, it is even worse than what the voters could achieve if each could commit to vote basing on the signal $s_i^0$ about the public project only.

The inefficiency of the voluntary contributions equilibrium is a well-known result due to free-riding. However, in that model, when voters have similar preferences about the optimal level of public good, this problem would not arise, and this is a major argument for voting over public projects rather than raising money through voluntary contributions. Proposition 1.8 says that this need not be the case whenever individuals, while sharing their preferences about the public project, also have conflicting preferences over other issues. In that case, competition for politician’s effort over the private issues may make the voters fully ignore politician’s performance on the public project. It is important to emphasize, however, that under different assumptions, ignoring the private interests, and voting solely on the basis of the outcome of the public project (or just giving nontrivial
weights to the two issues), may also be an equilibrium. For instance, if the costs of individuals' private projects are unlikely to be close to one another, and therefore the voters do not compete with one another fiercely, it may well be the case that they make the politician work on the public project in equilibrium. Another conclusion is that a public good is likely to be provided not when it benefits many people, but when there are many people for which this is the primary policy question.

1.9.2 Lobbying

The assumptions on preferences above may be slightly modified to encompass a lobbying story. The question here is, under what circumstances can a minority be successful in making the politician exercise the policy they prefer, and which the majority dislikes. The conventional wisdom would suggest that large groups are likely to have more influence on the policies, for at least two reasons. First, larger groups tend to have more resources, and therefore can allow their candidate to spend more on political campaigns, or offer a higher bribe to the politician. Second, they also have more voters, which makes the politician inherently more willing to gain support of these voters. However, Olson (1971) suggested an opposite view, i.e., that smaller groups are more likely to succeed in pushing their interests. Esteban and Ray (2001) establish this result formally for the case of lobbying: they show that small groups may be able to gather more resources because the free-rider problem is stronger in larger groups. Below, I confirm Olson's intuition for the case without monetary transfers as well. The argument is as follows. Large groups are more likely to include individuals with heterogenous preferences. Even though their preferences on a single dimension may be perfectly aligned, they may well care about other policy issues. Similarly to the result on public good provisions, each voter from the large group will then have an incentive to ignore their common interest, and condition his vote on policy outcomes that only he cares about. As the small group is less likely to have this problem, it is possible that the politician will respond to the interests of a small group rather than a large one. And indeed, minority groups are sometimes more successful in lobbying in real life, even if only voting is involved. For instance, labor unions reward politicians with
support on elections rather than bribes, but they succeed in making politicians choose the policies they advocate even if union members constitute a minority of the population, and everyone else is naturally opposed to the policy.

Consider the above model, except that the preferences of voter 1 are now different. He cares about a public project, but, unlike other voters, he prefers outcome \( \pi_0 = 0 \) to outcome \( \pi_0 = 1 \), so his preferences are given by

\[
u_1 = \Pr(\pi_0 = 0).
\]

We now have the following result.

**Proposition 1.9** Suppose \( k \geq 3 \). For any \( \phi \) there exists \( \sigma^* > 0 \) such that if \( \sigma < \sigma^* \), then in the unique symmetric THPE, voter \( i > 1 \) support the politician if and only if he got signal \( s_i = 1 \). Voter \( 1 \) supports the politician if and only if he gets the signal that the public project failed (\( s_0^0 = 0 \)). In this equilibrium, the politician will choose \( e_0 = 0 \) for a non-empty set of cost vectors with \( c_0 < 0 \).

This proposition, in line with Proposition 1.8 above, states that the large group of voters (\( i > 1 \)) will not condition their voting on the public project. Because of that, whenever the effort on that project is costly for the politician, the project will not be implemented. Moreover, even if the politician wants the project to be implemented, the minority lobbyist (voter 1) may influence the politician and prevent its implementation. Of course, the opposite story is also possible: a minority may induce the politician to undertake a project which the rest of the society opposes. Such effect is possible because the voting power of the majority is distracted on their private projects, and as a result, the minority becomes the sole lobbyist on the public project dimension.

### 1.10 Conclusion

In this paper, I investigate, how multidimensionality of and heterogeneity in voters’ preferences affects the problem of providing incentives to the politician, in the case where voting is the only instrument they have. Having taken the heterogeneity of voters’ interests
into consideration, one uncovers a number of novel effects. These effects include non-monotonic effect of transparency, possible underprovision of public goods even if all voters have the same preferences about their optimal level, and the possibility that small interest groups succeed in implementing their preferences better than large interest groups. This has potential policy implications. For example, in some cases, the society may want to limit the transparency of political decisions. Another corollary is that a public good is more likely to be provided not when all voters want that public good, but when there is an interest group, albeit small, which cares about this public good more than anything else, and is therefore able to judge the politician on the basis on the amount of this public good provided.

The theoretical framework introduced in this paper is also potentially applicable to a number of other situations, both political and economical. For example, take the problem of legislators who want to provide incentives to a bureaucrat or a government agency. The multiprincipal nature of problem is recognized and emphasized by McCubbins, Noll, and Weingast (1987), who argue that different interests of the principals are actually a virtue, which allows the legislative body to capture interests of different population groups in different localities. They make the conclusion that having multiple principals will facilitate collection of information about the performance of bureaucrats, thereby providing better control. Here in this paper, I show that this need not be the case, as multiplicity of principals with opposing interests also allows the bureaucrats to bias the policies to favor only a narrow subset of principals, be it voters or legislators.

More generally, the results of the paper are applicable to many economic situations with one or many principals and a binary reward for the agent. Such cases include decisions to hire a new worker (assuming, for example, that the wage is set by the market) or fire an existing one, decisions to promote or pay a bonus. A person seeking employment, for instance, must often show different qualities to future colleagues: some care more about professionalism, other care about ability to substitute other workers if needed, while some may merely care about having an easy-going person as a colleague. On the job interviews, therefore, the prospective employee must strike a balance between signaling
different qualities of his. An employee seeking promotion to a higher position needs not only to have the approval of the head of his department, but also be known and valued by heads of other departments with whom he will have to interact in a new position. In this case, a conflict of interest in the employee's current job is created, as he may seek to divert time from his immediate responsibilities to working on projects which are more visible to heads of other departments.

The ideas of the paper may be potentially applied to a corporate governance problem where shareholders have imperfectly aligned interests. Indeed, some shareholders may be interested in payoff of dividends, other in future appreciation of the firm's market value, and some may have environmental friendliness in mind as well. Even more important may be the conflict of interest between minority and large shareholders, where the latter may want to tunnel the capital stock or perhaps have strategic reasons for mergers and acquisitions that go beyond the goal of maximizing the value of this particular firm. In the spirit of the theory in this paper, one may argue that while minority shareholders possess too little information about the CEO's actions, large shareholders possess too much information, and this naturally inclines the CEO to undertake actions that benefit large shareholders. In this case, the board of directors may not only solve the collective action and information acquisition problem of minority shareholders (this is a trivial effect), but also introduce an agency problem between large shareholders and the directors, which may be exactly the friction that would introduce noise in large shareholder's behavior, thus relaxing the problem of too precise signals about CEO's action. In other words, the board of directors may solve the problem of having too little information by some shareholders and too much information by other ones at the same time.

There are also other situations, where the conflict of interests between principals make the framework in this paper relevant. Among them is the problem of a mediator between several interested parties. The mediator gets some reward from the success of negotiations, but is not interested in the precise solution per se. The parties involved have opposing interests, but not exactly so (otherwise there is no room for negotiation), as the job of the mediator is to provide a solution that would satisfy all or most of the parties involved.
The theory predicts that in multilateral negotiations when a consensus is not needed, some noise may be beneficial not only for the probability of success, but also for the quality of the agreement reached with the help of the mediator. I leave these extensions for future research.

1.11 Appendix

Proof of Theorem 1. Take voter $i$ who prefers $e_d(i) = 1$ to $e_d(i) = 0$ (the opposite case is treated similarly). The idea of the proof is to show that it is impossible that the politician switches from $e_d(i) = 1$ to $e_d(i) = 0$ if voter $i$ decides to provide higher-powered incentives. Formally, fix the strategies of all voters other than $i$, and suppose that he supports the politician with probability $x_1$ if $s_i = 1$ and with probability $x_0$ if $s_i = 0$. Suppose, to obtain a contradiction, that either $x_1 < 1$ or $x_0 < 1$. Then let

$$z = (q_d(i)p_i + (1 - q_d(i))(1 - p_i))x_1 + (q_d(i)(1 - p_i) + (1 - q_d(i))p_i)x_0 - (q_d(i)p_i + (1 - q_d(i))(1 - p_i))x_0 - (q_d(i)(1 - p_i) + (1 - q_d(i))p_i)x_1$$

$$= (q_d(i)p_i + (1 - q_d(i))(1 - p_i) - q_d(i)(1 - p_i) - (1 - q_d(i))p_i)(x_1 - x_0);$$

then $z$ is the probability of player $i$ supporting the politician if $e_d(i) = 1$ less the same probability if $e_d(i) = 0$. Since $q_d(i) > \frac{1}{2}$ and $p_i > \frac{1}{2}$, then $z$ is increasing in $x_1 - x_0$. The assumption above states that $x_1 - x_0 < 1$.

Let us first prove that $x_1 - x_0 \leq 0$ cannot be the case in a THPE, i.e., it is impossible that the voter promises to support the politician with a higher probability if he gets a low signal. Indeed, this implies, given (1.6), that the politician is at least as likely to get the support of voter $i$ if he chooses $e_d(i) = 0$ than $e_d(i) = 1$. Consequently, he will only choose $e_d(i) = 1$ if $c_d(i) \leq 0$. Suppose voter $i$ switches to $x_1 - x_0 > 0$. Consider any tremble in the strategies of other players; voter $i$ is then pivotal with a positive probability for any cost vector $c$. This means that the politician will choose $e_d(i) = 1$ not only if $c_d(i) \leq 0$, but also for some cost vectors where $c_d(i) > 0$, as this would by him the support of voter $i$ with a positive probability. Consequently, this deviation is profitable, which means that
To obtain a contradiction, consider the following argument. Suppose that for cost vector \( c \), the politician is indifferent between two effort vectors which he likes best, \( e^1 = \left( e^1_{-d(i)}, 1 \right) \) and \( \left( e^2_{-d(i)}, 0 \right) \), where \( e_{-d(i)} \) denotes the politician’s effort choice on all dimensions except \( d(i) \) (if he is not indifferent between such two vectors, then a small deviation by voter \( i \) will not make him switch from \( e_{d(i)} = 1 \) to \( e_{d(i)} = 0 \) or in the other direction). We know that he weakly prefers these vectors to \( \left( e^1_{-d(i)}, 0 \right) \) and \( \left( e^2_{-d(i)}, 1 \right) \). Consider some tremble of other voters’ strategies, and let \( \rho_1 \) and \( \rho_2 \) be the probabilities that voter \( i \) is pivotal if the politician chooses \( e^1_{-d(i)} \) and \( e^2_{-d(i)} \), respectively (it obviously does not depend on \( e_{d(i)} \)). Since \( e^1 = \left( e^1_{-d(i)}, 1 \right) \) is weakly preferred to \( \left( e^1_{-d(i)}, 0 \right) \), we must have

\[
\rho_1 z \geq c;
\]

indeed, otherwise the politician would be better off switching to \( \left( e^1_{-d(i)}, 0 \right) \). Similarly, we have

\[
\rho_2 z \leq c.
\]

This implies that \( \rho_1 \geq \frac{c}{z} \geq \rho_2 \) (we proved above that \( z > 0 \), since \( x_1 - x_0 > 0 \)).

Suppose now that voter \( i \) deviates by increasing \( x_1 - x_0 \) by some small \( \varepsilon \). For interpretational purposes, suppose that he does so by increasing \( x_1 \) while keeping \( x_0 \) constant. This increases the likelihood that voter \( i \) will support the politician for any action that the politician takes. However, if he chooses vector \( e^1 \), the increment would be higher for two reasons: first, the likelihood of signal \( s_i = 1 \) is higher, because \( e^1_{d(i)} = 1 \) and \( e^2_{d(i)} = 0 \), and because voter \( i \) is (weakly) more likely to be pivotal if the politician chooses \( e_1 \) rather than \( e_2 \). Hence, the politician is no longer indifferent, and strictly prefers \( e^1 \) to \( e^2 \).

The argument above implies that under such deviation, there does not exist a cost vector for which the politician will switch from \( e_{d(i)} = 1 \) to \( e_{d(i)} = 0 \). Given the full support of vector \( c \), the opposite will happen with a positive probability, as there are always cost vectors for which the politician is indifferent between \( e_{d(i)} = 1 \) and \( e_{d(i)} = 0 \). This is a profitable deviation, which means that simple voting strategies are the only
possible THPE. It is now trivial to demonstrate that it is indeed a THPE; indeed, the argument above suggests that the simple voting strategy strictly dominates any other strategy under the conditions of this Theorem. This completes the proof. ■

**Proof of Theorem 2.** Let us first consider the case $F_i = V$. Suppose that strategies $M_i : \{0, 1\}^{|V|} \rightarrow \Delta \{n, y\}$, where $S_i = \{0, 1\}^{|V|}$, form a THPE. This means that if we take a small parameter $\xi$, so that all voters except $i$ put at least probability $\xi$ on each voting strategy, then the best response of voter $i$ is to play strategy $M_i$.

We start by taking a combination of signals $s$ (observed by all voters, since $F_i = V$), such that $\Pr(s | \pi_j = 1) > \Pr(s | \pi_j = 0)$, and proving that a voter $i$ who cares about dimension $j$ and prefers $\pi_j = 1$ to $\pi_j = 0$ should vote $y$ with probability 1 in this equilibrium, provided that the precisions of signals $\{p_l\}_{l \in V}$ of all voters is sufficiently high. If we do that, we would be able to establish, by a similar argument, that in this case voters who prefer $\pi_j = 0$ to $\pi_j = 1$ should vote against the politician, and, similarly, if the combination of signals is such that $\Pr(s | \pi_j = 1) < \Pr(s | \pi_j = 0)$, then the voting strategies of these two kinds of voters are reversed. Since there are only a finite number of voters and combinations of signals, we would then conclude that if all precisions $p_i$ are sufficiently high, then the only THPE may involves voting according to (1.7). It would remain to show that these voting strategies indeed form a THPE if signals are sufficiently precise, but this will be straightforward, in light of the reasoning below.

Suppose $s'$ is such that $\Pr(s' | \pi_j = 1) > \Pr(s' | \pi_j = 0)$, and voter $i$ prefers $\pi_j = 1$ to $\pi_j = 0$. Suppose, to obtain a contradiction, that the probability that he votes in this case for the politician is $\gamma < 1$. Take a small $\varepsilon > 0$ and consider the following deviation: for the combination of signals $s'$, voter $i$ votes for the politician with probability $\gamma + \varepsilon \in (\gamma, 1)$. Let us show that this is a profitable deviation when $\varepsilon$ is sufficiently small and signals are sufficiently precise. Consider all possible vectors $c$ of costs. For almost all vectors, the politician is not indifferent between any two effort vectors neither before nor after the deviation. Because of that, such cost vectors may be classified into one of three categories: where $e_j^1(c) = e_j^2(c)$, where $e_j^1(c) = 1 > 0 = e_j^2(c)$, and where $e_j^1(c) = 0 < 1 = e_j^2(c)$ (here, $e^1(c)$ and $e^2(c)$ denote the politician’s effort choice if the realization of the cost
vector is \( c \) before and after the deviation, respectively). For the deviation to be profitable, we need to show that the probability of the cost vector falling into the second category is higher than the probability of it falling into the third category.

Denote the probability that voter \( i \) is pivotal if vector \( s' \) is realized by \( p_i \). For the combination of signals \( s' \) consider the “most likely” effort vector \( \tilde{e}(s') \), constructed as follows:

\[
\tilde{e}_j(s') = \begin{cases} 
1, & \text{if } \Pr \left( \{s_l\}_{l \in A_j} = \{s'_l\}_{l \in A_j} \mid e_j = 1 \right) > \Pr \left( \{s_l\}_{l \in A_j} = \{s'_l\}_{l \in A_j} \mid e_j = 0 \right) \\
0, & \text{otherwise}
\end{cases}
\]

(1.12)

Now, evidently, for any \( Z > 0 \) there exists a combination of precisions \( \{p_i^0\} \) such that if \( p_i > p_i^0 \) for all \( i \), the following property holds:

\[
\frac{\Pr (s = s' \mid e = \tilde{e}(s'))}{\Pr (s = s' \mid e \neq \tilde{e}(s'))} > Z,
\]

(1.13)
i.e., the probability of obtaining the combination of signals \( s' \) is much higher for \( e = \tilde{e}(s') \) than for any other \( e \). Note that for voter \( i \), the value \( \tilde{e}_{d(i)}(s) \) equals 1 if the equation (1.7) prescribes him to support the politician, and equals 0 otherwise. In other words, equation (1.7) prescribes each voter to use his best guess about the politician’s true effort.

Let us go back to the deviation and verify that the deviation indeed induces the politician to make voter \( i \) better off. This deviation increases the payoff of the politician (the expected probability of being reelected) for realization of signal \( s' \) and does not change it in any other case. Let us compare the measure of cost vectors such that the politician switches from some \( e \) such that \( e_{d(i)} = 0 \) to vector \( \tilde{e}(s') \), and the measure of cost vectors where any other switch occurs. Since \( Z \) may be taken arbitrarily high, and that the density of the cost vector is bounded away from zero on any compact, the second measure may be made negligible, but the first one will remain bounded away from zero (recall that the probability that the voter is pivotal is the same given \( s \)). This implies that the deviation is more likely to make the politician switch from \( e^1 \) such that \( e_{d(i)}^1 = 0 \) to one \( e^2 \) where \( e_{d(i)}^2 = 1 \), and less likely to make him switch in the opposite direction. Hence, the deviation
is profitable, which proves that for this signal \( s \), voter \( i \) is better off announcing that he will vote for the politician with probability 1.

Now consider the case where \( \mathcal{F}_i \supset A_{d(i)} \) holds instead of \( \mathcal{F}_i = \mathcal{V} \). As we showed above, the best strategy of voter \( i \) does not depend on the values of signals of voters who care about other policy issues; the only thing that matters is \( s|_{A_{d(i)}} \). Hence, if voter \( i \) does not support the politician with probability 1 in case (1.7) holds, a deviation where voter \( i \) increases this probability by a small amount will again be profitable. This proves that the only possible THPE is one where each voter uses the rule given in the statement of the Theorem.

Finally, we must verify that this is indeed a THPE. But this is trivial: indeed, if some voter had a profitable deviation, there would exist a profile of signals \( s' \) such that a small deviation further from the candidate equilibrium strategy is profitable. But the reasoning above shows why this is impossible. This simple consideration completes the proof. ■

**Proof of Proposition 1.1.** The idea of the proof is similar to the one of Theorem 1. By choosing a higher \( p_i \), it is impossible to make the politician choose \( e_{d(i)} = 0 \) instead of \( e_{d(i)} = 1 \) for any cost vector \( c \). The reason is that \( z \), given by (1.11), is increasing in \( p_i \), because \( x_1 = 1 > 0 = x_0 \) (this was established in the proof of Theorem 1), and \( q_{d(i)} > \frac{1}{2} \) by assumption. Consequently, if \( p_i \neq p_H \), then voter \( i \) has a profitable deviation to a higher \( p_i \), as this will increase his utility with a positive probability. Therefore, in the only THPE, voters must choose \( p_i = p_H \). This does not depend on whether these choices are known to other voters, because for any precisions of other voters’ signals, playing the simple voting strategy is a dominant strategy for each voter in the continuation game. This completes the proof. ■

**Proof of Proposition 1.2.** Suppose that the politician faces cost \( \tilde{c} \) and compares his utility from working on \( |e| = r + 1 \) and \( |e| = r \) policy dimensions. The difference in his expected utility is then

\[
\begin{align*}
    f(r) &= (2p - 1) \sum_{i=0}^{r} \binom{r}{i} p^i (1 - p)^{r-i} \left( \frac{n-r}{m-1-i} \right) p^{n-r-(m-1)+i} (1 - p)^{m-1-i} - \tilde{c} \quad (1.14)
\end{align*}
\]
(if \( r < m - 1 \), the binomial coefficients are assumed to be zero). Indeed, the politician
gets benefit from working on the policy dimension favored by an extra voter only if this
voter is pivotal, and in this case he increases the probability of reelection from \( 1 - p \) to \( p \).
This function has a unique maximum. The politician’s utility is maximized at the only \( r \)
such that \( f (r - 1) > 0 \) and \( f (r) < 0 \) (where we assume \( f (n) < 0 \)), or at \( r = 0 \).

The function \( f \) is locally increasing in \( p \) if \( p \) is close to \( \frac{1}{2} \), and locally decreasing in \( p \) if
it is close to 1. Consequently, the optimal choice of \(|e| \) may only increase in the first case,
and only decrease in the second. Given the support of distribution of \( \tilde{c} \), for a small \( \varepsilon \), if \( p \) is
close to \( \frac{1}{2} \) or to 1, there is a positive probability that the politician will actually switch to
a different \( r \). This establishes that if \( p \) is close to \( \frac{1}{2} \), the voters’ welfare, is nondecreasing,
and sometimes increasing, in \( p \); for \( p \) close to 1, the effect is opposite.

To prove the last part of the proposition, take \( \tilde{c} \) sufficiently small. Then optimal \(|e| \)
is a single-peaked function of \( p \), and its maximum is \( n \) (i.e., for sufficiently small values of \( \tilde{c} \)
and some values of \( p \), it is optimal for the politician to exert full effort). If the cost \( \tilde{c} \) takes
such values with probability 1, then the \(|e| \) is also a single-peaked function of \( p \). There
may be several values of \( p \) where the maximum is achieved. All of them are those where
\(|e| = n \) for the largest value in the support of \( \tilde{c} \), because the set of \( p \)’s for which \(|e| = n \)
is shrinking in \( \tilde{c} \). This completes the proof. ■

**Proof of Proposition 1.3.** (i) If \( \tilde{c} \leq 0 \), the politician is willing to exert effort \(|e| = 1 \),
since it has negative cost and increases his chance of being reelected. If \( \tilde{c} > 0 \), then he
compares two possibilities. In order to get reelected, he needs to choose \(|e| \geq \frac{p^{p+\mu-1}}{2^{p-1}} \).
Indeed, if he exerts effort \(|e| \), then he gets \(|e| p + (1 - |e|) (1 - p) \) votes, and he needs \( \mu \)
votes. If he decides to save his effort, he will not get reelected (because \( \mu \geq \frac{1}{2} \)), but will
not exert any effort. The precise condition when it is optimal for him to exert effort is
\( \tilde{c} \leq \frac{2^{p-1}}{p^{p+\mu-1}} \).

(ii) The voters’ utility cannot be higher than 1. If \( \tilde{c} < 1 \), this is attainable when \( p = \mu \).
If \( \tilde{c} > 1 \), then the optimal \( p \) should be as low as possible, as long as \( \tilde{c} \leq \frac{2^{p-1}}{p^{p+\mu-1}} \) (this follows
from (1.8). If \( \tilde{c} > \frac{1}{\mu} \), such \( p \) would have to be greater than 1, which is impossible; for such
\( \tilde{c} \), the politician will choose to exert \(|e| = 0 \).
(iii) The formula generalizes the statement of part (ii). Indeed, \( \mu > \frac{1-c(1-\nu)}{2-c} \) is equivalent to \( c < 1 \), provided that \( c < 2 \). This completes the proof. ■

**Proof of Proposition 1.4.** (i) If \( \mu > 1 - p \), the politician who exerts effort \( |e| = 1 \) will be reelected. Hence, the same argument as in the proof of Proposition 1.3 (i) applies, and the optimal effort \( |e| \) is given by (1.8). If, however, \( \mu < 1 - p \), the politician has no chance of getting reelected, given the voters’ strategies. He will then choose the policy outcomes that he prefers best (i.e., choose \( |e| = 1 \) if \( c < 0 \) and \( |e| = 0 \) if \( c > 0 \), and lose elections.

(ii) If \( c > 0 \) and \( \frac{1}{2} < p < 1 - \mu \), the politician does not have a chance of reelection, and will choose \( |e| = 0 \). Otherwise, the same reasoning as in the proof of 1.3 (ii) applies: to get reelected, he needs to choose \( |e| \geq \frac{p+\mu-1}{2p-1} \). It is optimal for him to do as long as \( c < \frac{1}{\mu} \); for higher \( c \), he would prefer \( |e| = 0 \).

(iii) In this case, it is always better to choose higher \( p \), since \( \frac{p+\mu-1}{2p-1} \) is increasing in \( p \) for \( \mu < \frac{1}{2} \). Whenever this formula does not apply, the politician chooses the same effort for any \( p \), hence \( p = 1 \) is always optimal. This completes the proof. ■

**Proof of Proposition 1.5.** (i) Here, the politician will always choose \( |e| = 1 \). Hence, the choices of \( \mu \) and \( p \) do not matter.

(ii) As follows from Proposition 1.3 (ii), it is possible to achieve \( |e| = 1 \) if one chooses \( \mu = p \in \left( \frac{1}{2}, \frac{1}{c} \right) \). The optimal value of \( p \) may be 1 (and in that case, unanimity rule \( \mu = 1 \) should be chosen) if and only if \( c < 1 \). For \( 1 < c < 2 \), \( \frac{1}{c} < 1 \), and choosing \( p = 1 \) is not optimal.

(iii) If \( c > 2 \), then, as follows from Proposition 1.3 (ii), any supermajority rule will make the politician choose \( |e| = 0 \). The society may do better by choosing a submajority rule. In that case, the optimal combination of \( \mu \) and \( p \) is given by \( \mu = \frac{1}{c} \) and \( p = 1 \), as follows from Proposition 1.4. This completes the proof. ■

**Proof of Proposition 1.6.** If \( c_p > (2\alpha - 1) c_f \), then, taking into account that a financial project buys the politician 1 vote and a position-taking project buys him \( \alpha - (1 - \alpha) = 2\alpha - 1 \) votes, then a financial project brings him more votes for a unit cost.
of effort. He then only takes his favorite positions on the position-taking dimensions, and this brings him $\alpha$ votes for every issue where his preferences are aligned with the majority (the share of such issues is $\alpha$) and $1 - \alpha$ votes for every issue where his preferences are misaligned (share $1 - \alpha$). Given that $\gamma > \mu$, there are enough financial issues to get share $\mu$ of votes by working on spending issues only. If, however, $c_p < (2\alpha - 1) c_f$, then the politician will first take the position of the majority on all position-taking issues, and then get the rest of the votes from spending projects. The share of votes he will get from voters who care about position-taking issues is therefore $(1 - \gamma) \alpha$.

Overall, the share of spending projects that the politician works on, $y$, is given by

$$y = \frac{\mu - (1 - \gamma) \left( \alpha^2 + (1 - \alpha)^2 \right)}{\gamma}$$

(1.15)

if $c_p > (2\alpha - 1) c_f$, and by

$$y = \frac{\mu - (1 - \gamma) \alpha}{\gamma}$$

(1.16)

otherwise. In either case, $y$ is decreasing in $\alpha$, and the effect is stronger if $\gamma$ is smaller. In addition, the derivative $\frac{dy}{d\alpha}$ equals $-\frac{1 - \gamma}{\gamma} (4\alpha - 2)$ if $c_p > (2\alpha - 1) c_f$, and $-\frac{1 - \gamma}{\gamma}$ otherwise. The negative effect is clearly stronger if the cost of taking an unfavorable position is stronger than the cost of spending on a financial issue. This completes the proof.

**Proof of Theorem 3.** Take any symmetric THPE. Suppose that the voters are not using the strategies as defined by (1.7). Consider then a small deviation by voter $i$ towards this strategy. By the assumption about costs, this deviation is much more likely to make the politician change the identities of policies that he works while keeping his total effort fixed, rather than his total effort. In that case, the politician will become more likely to choose $e_{d(i)} = 1$, because he is almost indifferent about the costs, and such strategy shifts the distribution of votes to the right, as voter $i$ became more responsive to his effort. If $\sigma$ is sufficiently small, then this benefit will dominate the potential cost that voter's deviation will make the politician choose a smaller total effort. This makes the deviation profitable for voter $i$, and this means that there is a unique symmetric THPE.

If, however, voters do not observe signals of all voters from $A_{d(i)}$, then for the argument
to go through, one needs to require that \( n \) and \( k \) are large. In that case, if \( p < 1 \), then the signals that other voters from the same group get affect the probability of each voter to be pivotal very little. In that case, deviations of each voter towards a more responsive strategy are profitable. This completes the proof.

Proof of Proposition 1.8. (i) Consider two groups, \( A_j \) and \( A_{j'} \); the first one exchanges their signals, the second does not. Trivially, if the politician chooses between exerting \( e_j = 1, e_{j'} = 0 \) and \( e_j = 0, e_{j'} = 1 \), he would choose the first, as this gives him a distribution of votes which stochastically dominates the one in the second case. Because of that, exchanging signals within each group will make the politician more likely to exert effort that they prefer.

(ii), (iii) If \( p \) is low, than deliberation within each group will make it more likely that the politician will exert effort \(|e| > 0\), which makes the society better off. However, if \( p \) is close to 1, then deliberation will make the politician much more certain that he does not need to exert effort higher than \(|e| = m_i\), as the probability that some voters fail to reward him becomes lower, which completes the proof.

Proof of Proposition 1.8. Take any \( \phi \) and suppose that in a symmetric THPE, voters condition their voting decision on their signals about \( \pi_0 \). Consider voter \( i \) and suppose that he deviates to strategy \( M_i = 1 \) if and only if \( s_i = 1 \). This deviation may have two effects: he will ensure \( e_i = 1 \) with a higher probability, but for some cost vectors he will also make the politician switch from \( e_0 = 1 \) to \( e_0 = 0 \). However, if \( \sigma \) is small, the ratio of these two probabilities may be made arbitrarily large. Consequently, this deviation is profitable, and the only THPE involves \( M_i = 1 \) if and only if \( s_i = 1 \). Now, \( \phi < 1 \) ensures that no voter has a profitable deviation from this equilibrium: it is not profitable for any one of them to try to induce the politician to choose \( e_0 = 1 \).

Now that the voters ignore their signals \( s_i^0 \) about \( \pi_0 \), the politician will choose \( e_0 = 1 \) only if \( c_0 \leq 0 \). This completes the proof.

Proof of Proposition 1.9. The proof that in any symmetric THPE, voter \( i > 1 \) will announce \( M_i = 1 \) if and only if \( s_i = 1 \) is similar to the proof of Proposition 1.8 and is
omitted. Voter 1, since there are no other policy issues that he cares about, will announce $M_1 = 1$ if and only if $s_1^0 = 0$. The politician will then never choose $e_0 = 1$ if $c_0 > 0$ (so if the politician dislikes the public project, he will not do it). Obviously, if $c_0 < 0$ is sufficiently small in absolute value, the politician will decide to get the vote of voter 1, even though he would like to see the project to be implemented. In this case, voter 1 lobbies successfully. This completes the proof. ■
2. POLITICAL SELECTION AND PERSISTENCE OF BAD GOVERNMENTS

(joint work with Daron Acemoglu and Konstantin Sonin)

2.1 Introduction

A central role of (successful) political institutions is to ensure the selection of the right (honest, competent, motivated) politicians. Besley (2005, p. 43), for example, quotes James Madison, to emphasize the importance of the selection of politicians for the success of a society:

“The aim of every political Constitution, is or ought to be, first to obtain for rulers men who possess most wisdom to discern, and most virtue to pursue, the common good of society; and in the next place, to take the most effectual precautions for keeping them virtuous whilst they continue to hold their public trust.”

Equally important, but less often emphasized, is the “flexibility” of institutions, meaning their ability to deal with shocks and changing situations, for example by adapting the nature of the government and changing the characteristics of those in power in response to changes in the environment. In this paper, we construct a simple dynamic model of government formation to highlight the potential sources of inefficiency and the selection of governments and to identify features of political processes that create “institutional flexibility”.¹

¹ Even though we model changes in the underlying environment and the competences of different governments as resulting from stochastic shocks, in practice these may also result from deterministic changes in the nature of the economy. For example, authoritarian regimes such as the rule of General Park in South Korea or Lee Kuan Yew in Singapore may be beneficial or less damaging during the early stages of development, while a different style of government, with greater participation, may be necessary as the economy
The “government” is made up of a subset of the citizens (e.g., each three-player group may be a government, etc.). Each (potential) government has a different level of competence, affecting the collective utility it provides to citizens (e.g., the level of public goods). Each individual also receives rents from being part of the government (additional income, utility of office, or rents from corruption). New governments are put in place by a combination of “votes” from the citizens and “consent” from current government members. The extent of necessary consent of current government members is a measure of the “degree of democracy”. For example, a perfect democracy can be thought of as a situation in which current incumbents have no special power and no such consent is necessary. Many political institutions provide additional decision-making or blocking power to current government members, however; in many democracies, there is incumbency advantage, so that a government in power is harder to oust than it would have been to institute had it been out of power (e.g., see Cox and Katz, 1996, for a discussion of such incumbency advantage in mature democracies). In nondemocratic societies, this advantage of current government members is more pronounced, even palpable. For instance, only new governments that include some members of previous governments might be feasible or much more than a simple majority of voters may be necessary to oust the government (as, unfortunately, illustrated by the recent events in Zimbabwe).

The first contribution of our paper is to provide a general and tractable framework for the study of such dynamic political selection issues and to provide the detailed characterization of the structure (and efficiency) of the selection of politicians under different political institutions. Perfect democracies always ensure the emergence of the best (most competent) government. In contrast, under any other arrangement, incompetent and bad governments can emerge and persist despite the absence of information-related challenges to selecting good politicians. For example, even a small departure from perfect democracy, whereby only one member of the current government needs to consent to a new government, may make the worst possible government persist forever. The intuitive explanation develops and becomes more complex. Acemoglu, Aghion and Zilibotti (2006) suggest that “appropriate” institutions may be a function of the distance of an economy to the world technology frontier.
for why even a small degree of incumbency advantage might lead to such outcomes is as follows: improvements away from a bad (or even the worst) government might lead to another potential government that is itself unstable and will open the way for a further round of changes. If this process ultimately leads to a government that does not have any common members with the initial government, then it may fail to get the support of any of the initial government members. In this case, the initial government survives even though it has a low, or even possibly the lowest, level of competence.

This discussion highlights the important dynamic interactions in the process of selecting politicians and governments. Another important implication of this analysis is that, beyond perfect democracy, there is no obvious ranking among different shades of imperfect democracy (and dictatorships). Any of these different regimes may lead to the emergence of better governments in the long run. This result is consistent with the empirical findings in the literature that show no clear-cut relationship between democracy and economic performance (e.g., Przeworski and Limongi, 1997, Barro, 1997, Minier, 1999). In fact, both under imperfect democracies and extreme dictatorships, the competence of the equilibrium government and the success of the society depend strongly on the identity of the initial members of the government. This is consistent with the emphasis in the recent political science and economics literatures on the role that leaders may play under weak institutions (see, for example, Brooker, 2000, or Jones and Olken, 2004, who show that the death of an autocrat leads to a significant change in growth, and this does not happen with democratic leaders).

Our second contribution relates to the study of institutional flexibility. For this purpose, we enrich the above-mentioned framework with shocks that change the competences of different types of governments (thus capturing potential changes in the needs of the society for different types of skills and expertise). Although the systematic and tractable analysis of this class of dynamic games is challenging, we provide a systematic characterization of the structure of equilibria when stochastic shocks are sufficiently infrequent. Using this characterization, we show how the quality (competence level) of governments evolves in the presence of stochastic shocks and how this evolution is impacted by the set
of political institutions. While without shocks, a greater degree of democracy does not necessarily guarantee a better government (beyond perfect democracy), the pattern that emerges when we turn to institutional flexibility is different. In particular, our analysis shows that a greater degree of democracy generally leads to better outcomes in the long run; in particular, it increases the probability that the best government will be in power. This is because a greater degree of democracy enables greater adaptability to changes in conditions (which alter the relative rankings of quality of different governments). This therefore establishes the claim above, that more democratic institutions ensure greater flexibility.\footnote{The stochastic analysis also shows that random shocks to the identity of the members of the government may also lead to better governments in the long run because they destroy stable incompetent governments. Besley (2005) writes: “History suggests that four main methods of selection to political office are available: drawing lots, heredity, the use of force and voting.” Our model suggests why, somewhat paradoxically, drawing lots, which was used in Ancient Greece, might sometimes lead to better long-run outcomes than the alternatives.}

Finally, we also show that in the presence of shocks “royalty-like” nondemocratic regimes, where some individuals must always be in the government, may lead to better long-run outcomes than “junta-like” regimes, where a subset of the current members of the junta can block change even though no member is essential. The royalty-like regimes might sometimes allow greater adaptation to change because one of the members of the initial government is secure in her position. In contrast, as discussed above, without such security the fear of further changes might block all competence-increasing reforms in government.

We now illustrate some of the basic ideas with a simple example.

\textit{Suppose that the society consists of a large number, }n, \textit{of individuals. Assume that any }k = 3\textit{ individuals could form a government. A change in government requires both the support of the majority of the population and the consent of }l = 1\textit{ member of the government, so that there is an imperfect democracy, with a “minimal” degree of incumbency advantage. Suppose that each individual has a level of competence, denoted by }\gamma_j\textit{ for individual }j, \textit{and order them, without loss of any generality, in descending order according to their competence, so }\gamma_1 > \gamma_2 > \ldots > \gamma_n. \textit{The competence of}
a government is the sum of the competences of its three members. Each individual obtains instantaneous utility from the competence level of the government and also a large rent from being in office, so that each prefers to be in office regardless of the competence level of the government. Suppose also that individuals have a sufficiently high discount factor, so that the future matters a lot relative to the present.

It is straightforward to determine the stable governments that will persist and remain in power once formed. Evidently, \( \{1, 2, 3\} \) is a stable government, since it has the highest level of competence, so neither the majority of outsiders nor members of the government would like to initiate a change (some outsiders may want to initiate a change: for example, 4, 5, and 6 would prefer government \( \{4, 5, 6\} \), but they do not have the power to enforce such a change). In contrast, governments of the form \( \{1, i, j\} \), \( \{i, 2, j\} \), and \( \{i, j, 3\} \) are unstable (for \( i, j > 3 \)), meaning that starting with these governments, there will necessarily be a change. In particular, in each of these cases, \( \{1, 2, 3\} \) will receive support from both one current member of government and from the rest of the population, who would be willing to see a more competent government.

Consider next the case where \( n = 6 \) and suppose that society starts with the government \( \{4, 5, 6\} \). This is also a stable government, even though it is the lowest competence government and thus the worst possible option for the society as a whole. This is because any change in government will take the individuals to a new government of the form of either \( \{1, i, j\} \), \( \{i, 2, j\} \), or \( \{i, j, 3\} \), but we know that all of these are unstable. Therefore, any of the more competent governments will ultimately take us to \( \{1, 2, 3\} \), which does not include any of the members of the initial government. Since individuals are assumed to be relatively patient, none of the initial members of the government would support (consent to) a change that will ultimately exclude them. As a consequence, the initial worst government persists forever. This example illustrates how the worst possible government can be stable.

One can also verify easily that \( \{4, 5, 6\} \) is also stable government when \( l = 3 \), since in this case any change requires the support of all three members of government and none of them would consent to a change. Therefore, a greater degree of democracy (lower
l) does not guarantee better outcomes in the long run. In contrast, one can also show that under \( l = 2 \), \( \{4,5,6\} \) is not a stable government, so the lower degree of democracy can improve the quality of the government.

Now consider the same environment as above, but with potential changes in the competences of the agents. For example, individual 4 may see an increase in his competence from \( \gamma_4 \) to \( \gamma_3 \), while the competence of individual 3 declines to \( \gamma_4 \) (or we could allow for any other such switch). Suppose that shocks are sufficiently infrequent that stability of governments in periods without shocks is given by the same reasoning as for the nonstochastic case. Consider the situation starting with the government \( \{4,5,6\} \) and \( l = 1 \). Then, this government will remain in power until there is a shock. Nevertheless, the equilibrium government will eventually converge to \( \{1,2,3\} \); at some point a shock will change the competence of the ruling government to \( \{\gamma_3,\gamma_5,\gamma_6\} \), which is itself unstable, because now individual 4 (with competence \( \gamma_3 \)) would support the emergence of the government \( \{1,2,4\} \) (with the highest competence \( \{\gamma_1,\gamma_2,\gamma_3\} \)). In contrast, when \( l = 3 \), the ruling government will remain in power even after the shock. This simple example thus illustrates starkly how, even though a greater degree of democracy does not ensure better outcomes in the nonstochastic environment, it provides greater flexibility and hence better long-run outcomes in the presence of shocks.

Our paper is related to several different literatures. While much of the literature on political economy focuses on the role of political institutions in providing (or failing to provide) the right incentives to politicians (see, among others, Barro, 1973, Ferejohn, 1986, Besley and Case, 1995, Persson, Roland and Tabellini, 1997, Niskanen, 1971, Shleifer and Vishny, 1993, Acemoglu, Robinson and Verdier, 2004, Padro-i-Miquel, 2007), there is also a small (but growing) literature investigating the selection of politicians, most notably, Banks and Sundaram (1998), Diermeier, Keane, and Merlo (2005), and Besley (2005). The main challenge facing the society and the design of political institutions in these papers is that the ability and motivations of politicians are not observed by voters or outside parties. While such information-related selection issues are undoubtedly important, our paper focuses on the difficulties in ensuring that the “right” government is selected even
when information is perfect and common. Also differently from these literatures, we emphasize the importance of institutional flexibility in the face of shocks.

Besley and Coate (1997, 1998), Caselli and Morelli (2004), Messner and Polborn (2004), and Mattozzi and Merlo (2006) provide alternative and complementary “theories of bad governments/politicians”. For example, Caselli and Morelli (2004) suggest that voters might be unwilling to replace the corrupt incumbent by a challenger whom they expect to be equally corrupt. Mattozzi and Merlo (2006) argue that more competent politicians have higher opportunity costs of entering politics. However, these papers do not develop the potential persistence in bad governments resulting from dynamics of government formation and do not focus on the importance of institutional flexibility. We are also not aware of other papers providing a comparison of different political regimes in terms of the selection of politicians under nonstochastic and stochastic conditions. McKelvey and Reizman (1992) suggest that seniority rules in the Senate and the House may be playing the role of creating an endogenous incumbency advantage, and when this is the case, current members of these bodies will indeed have an incentive to introduce such seniority rules.

Our results are also related to recent work on the persistence of bad governments and inefficient institutions, including Acemoglu and Robinson (2008), Acemoglu, Ticchi, and Vindigni (2006), and Egorov and Sonin (2004).\footnote{Acemoglu (2008) also emphasizes the potential benefits of democracy in the long run, but through a different channel—because the alternative, oligarchy, creates entry barriers and sclerosis.}

More closely related to our work are prior analyses of dynamic political equilibria in the context of club formation as in Roberts (1997) and Barbera, Maschler, and Shalev (2001), as well as dynamic analyses of choice of constitutions and equilibrium political institutions as in Acemoglu and Robinson (2006), Barbera and Jackson (2004), Matthias and Polborn (2004), and Lagunoff (2006). Our recent work, Acemoglu, Egorov, and Sonin (2008), provided a general framework for the analysis of the dynamics of constitutions, coalitions and clubs. The current paper is a continuation of this line of research. It differs from our previous work in a number of important dimensions, however. First, the focus here, which is on the effects of different political institutions on the selection of politicians and
governments, is new and substantively different. Second, this paper extends our previous work by allowing for stochastic shocks and enables us to investigate issues of institutional flexibility. Third, it allows for a structure of preferences for which our previous results cannot be directly applied.

Finally, our paper is in the tradition of citizen-candidate models as in Besley and Coate (1997, 1998) and Osborne and Slivinski (1996), since individuals run for office and act both as potential government officials and citizens voting in favor of or against other governments.

The rest of the paper is organized as follows. Section 2.2 introduces the model. Section 2.3 introduces the concept of (Markov) political equilibrium, which allows a general and tractable characterization of equilibria in this class of games. Section 2.4 provides our main results on the comparison of different regimes in terms of selection of governments and politicians. Section 2.5 extends the analysis to allow for stochastic changes in the competences of the members of the society and in the rules governing elections. It also contains the comparison of different regimes in the presence of stochastic shocks. Section 2.6 concludes. Appendix A contains all of the proofs of the results stated in the text, while Appendix B considers an environment with explicitly specified proposal and voting procedures, and shows the equivalence between the Markov perfect equilibria of this environment and the (simpler) notion of political equilibrium used in Section 2.3.

2.2 Model

We consider a dynamic game in discrete time indexed by \( t = 0, 1, 2, \ldots \). The population is represented by the set \( \mathcal{I} \) and consists of \( n < \infty \) individuals. We refer to non-empty subsets of \( \mathcal{I} \) as coalitions and denote the set of coalitions by \( \mathcal{C} \). Most importantly, we also designate a subset of coalitions \( \mathcal{G} \subset \mathcal{C} \) as the set of feasible governments. For example, the set of feasible governments could consist of all groups of individuals of size \( k_0 \) (for some integer \( k_0 \)) or all groups of individuals of size greater than \( k_1 \) and less than some other integer \( k_2 \). To simplify the discussion, we define \( \bar{k} = \max_{G \in \mathcal{G}} |G| \), so \( \bar{k} \) is the upper bound for the size of any feasible government: for any \( G \in \mathcal{G} \), \( |G| \leq \bar{k} \). It may be natural to
assume that $k < n/2$.

In each period, the society is ruled by one of the feasible governments $G^t \in \mathcal{G}$. The initial government $G^0$ is given as part of the description of the game and $G^t$ for $t > 0$ is determined in equilibrium as a result of the political process described below. The government in power at any date affects three aspects of the society:

1. It influences collective utilities (for example, by providing public goods or influencing how competently the government functions).

2. It determines individual utilities (members of the government may receive additional utility because of rents of being in office or corruption).

3. It indirectly influences the future evolution of governments by shaping the distribution of political power in the society (for example, by creating incumbency advantage in democracies or providing greater decision-making power or veto rights to members of the government under other political institutions).

We now describe each of these in turn. The influence of the government on collective utilities is modeled via its competence. In particular, at each date $t$, there exists a function

$$\Gamma^t : \mathcal{G} \to \mathbb{R}$$

designating the competence of each feasible government $G \in \mathcal{G}$ (at that date). We refer to $\Gamma^t_G \in \mathbb{R}$ as government $G$’s competence, with the convention that higher values correspond to greater competence. In Section 2.4, we will assume that each individual has a level of competence and the competence of a government is a function of the competences of its members. For now, this additional assumption is not necessary. Note also that the function $\Gamma^t$ depends on time. This generality is introduced to allow for changes in the environment (in particular, changes in the relative competences of different individuals and governments).

Individual utilities are determined by the competence of the government that is in power at that date and by whether the individual in question is himself in the government.
More specifically, each individual $i \in I$ at time $t$ has discounted (expected) utility given by

$$U_i^t = E \sum_{t=\tau}^{\infty} \beta^{(t-\tau)} u_i^t,$$

(2.1)

where $\beta \in (0, 1)$ is the discount factor and $u_i^t$ is individual’s instantaneous utility, given by

$$u_i^t = w_i \left( G^t, \Gamma_{G^t}^t \right) = w_i \left( G^t \right),$$

(2.2)

where in the second equality we suppress dependence on $\Gamma_{G^t}^t$ to simplify notation; we will do this throughout unless special emphasis is necessary. We assume that $w_i$ satisfies the following properties.

**Assumption 2.1** The function $w_i$ satisfies the following properties:

1. for each $i \in I$ and any $G, H \in \mathcal{G}$ such that $\Gamma_G^t > \Gamma_H^t$: if $i \in G$ or $i \notin H$, then $w_i (G) > w_i (H)$.

2. for any $G, H \in \mathcal{G}$ and any $i \in G \setminus H$, $w_i (G) > w_i (H)$.

Part 1 of this assumption is a relatively mild restriction on payoffs. It implies that all else equal, more competent governments give higher instantaneous utility. In particular, if an individual belongs to both governments $G$ and $H$, and $G$ is more competent than $H$, then the individual prefers $G$. The same conclusion also holds when the individual is not a member of either of these two government or when he is only a member of $G$ (and not of $H$). Therefore, this part of the assumption implies that the only situation in which an individual may prefer a less competent government to a more competent one is when she is a member of the former, but not of the latter. This simply captures the presence of rents from holding office or additional incomes from being in government due to higher salaries or corruption. The interesting interactions in our setup result from the conflict of interest whereby individuals prefer to be in government even when this does not benefit the rest of the society. Part 2 of the assumption strengthens the first part and imposes that this conflict of interest is always present; that is, individuals receive higher payoffs from governments that include them than from those that exclude them (regardless of
the competence level of the two governments). We impose both parts of this assumption throughout.

We next provide an example that makes some of these notions slightly more concrete.

Suppose that the competence of government $G$, $\Gamma_G$, is the amount of public good produced in the economy under feasible government $G$. Then, we can write $w_i(G)$ as

$$w_i(G) = v_i(\Gamma_G) + b_i I_{i\in G},$$  \hspace{1cm} (2.3)

where $v_i : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function (for each $i \in \mathcal{I}$) corresponding to the utility from public good for individual $i$, $b_i$ is a measure of the rents that individual $i$ obtains from being in office, and $I_X$ is the indicator of event $X$. If $b_i \geq 0$ for each $i \in \mathcal{I}$, then (2.3) satisfies part 1 of Assumption 2.1. In addition, if $b_i$ is sufficiently large for each $i$, then each individual prefers to be a member of the government, even if this government has a very low level of competence, thus part 2 of Assumption 2.1 is also satisfied.

Finally, the government in power influences the determination of future governments whenever consent of some current government members is necessary for a change. We represent the set of individuals (regular citizens and government members) who can implement a change in government by specifying the set of winning coalitions, $\mathcal{W}_G$, as a function of current government $G$ (for each $G \in \mathcal{G}$). This is an economical way of summarizing the relevant information, since the set of winning coalitions is precisely the subsets of the society that are able to force (or to block) a change in government. We only impose a minimal amount of structure on the set of winning coalitions.

**Assumption 2.2** For any feasible government $G \in \mathcal{G}$, $\mathcal{W}_G$ is given by

$$\mathcal{W}_G = \{ X \in \mathcal{C} : |X| \geq m_G \text{ and } |X \cap G| \geq l_G \},$$

where $l_G$ and $m_G$ are integers satisfying $0 \leq l_G \leq |G| \leq \bar{k} < m_G \leq n - \bar{k}$ (where recall that $\bar{k}$ is the maximal size of the government and $n$ is the size of the society).
The minimal restrictions imposed in Assumption 2.2 are intuitive. In particular, they amount to requiring that a new government can be instituted if it receives a sufficient number of votes from the entire society \((m_G \text{ total votes})\) and if it receives support from some subset of the members of the current government \((l_G \text{ of the current government members need to support such a change})\). This definition allows \(l_G\) to be any number between 0 and \(|G|\).

Given this notation, the case where \(l_G = 0\) should be thought of as perfect democracy, where current members of the government have no special power, and the case where \(l_G = |G|\) as extreme dictatorship, where unanimity among government members is necessary for any change. In-between these extremes are imperfect democracies (or less strict forms of dictatorships), which may arise either because there is some form of (strong or weak) incumbency advantage in democracy or because current government (junta) members are able to block the introduction of a new government. Note also that we imposed some mild assumptions on \(m_G\). In particular, less than \(\bar{k}\) individuals is insufficient for a change to take place. This ensures that that a rival government cannot take the power without any support from other individuals (recall that \(\bar{k}\) denotes the maximum size of the government, so the rival government must have less than \(\bar{k}\) members), and \(n - \bar{k}\) individuals are sufficient to implement the change provided that \(l_G\) members of the current government are among them (though less may be sufficient as well). For example, these requirements are naturally met when \(\bar{k} < n/2\) and \(m_G = \lfloor (n + 1)/2 \rfloor\) (i.e., majority rule).

In addition to Assumptions 2.1 and 2.2, we also impose the following genericity assumption, which ensures that different governments have different competences. This assumption simplifies the notation and is without much loss of generality, since if it were not satisfied for a society, any small perturbation of competence levels would restore it.

**Assumption 2.3** For any \(t \geq 0\) and any \(G, H \in \mathcal{G}\) such that \(G \neq H\), \(\Gamma_{G}^{t} \neq \Gamma_{H}^{t}\).

### 2.3 Political Equilibria in Nonstochastic Environments

In this section, we focus on nonstochastic environments, where \(\Gamma^{t} = \Gamma\) (or \(\Gamma_{G}^{t} = \Gamma_{G}\) for all \(G \in \mathcal{G}\)). For these environments, we introduce our equilibrium concept, \((Markov)\)
political equilibrium, and show that equilibria have a simple recursive characterization.  

We return to the more general stochastic environments in Section 2.5.

2.3.1 Political Equilibrium

Our equilibrium concept, (Markov) political equilibrium, imposes that only transitions from the current government to a new government that maximize the discounted utility of a winning coalition will take place; and if no such transition exists, the current government will be stable (i.e., it will persist in equilibrium). The qualifier “Markov” is added since this definition implicitly imposes that transitions from current to a new government depend only on the current government—not on the entire history.

To introduce this equilibrium concept more formally, let us first define the mapping \( \phi : \mathcal{G} \rightarrow \mathcal{G} \), which maps each feasible government \( G \) to the government that would emerge in period \( t + 1 \). In other words, \( \phi \) defines a transition rule for governments. Given \( \phi \), we can write the discounted utility implied by (2.1) for each individual \( i \in \mathcal{I} \) starting from current government \( G \in \mathcal{G} \) recursively as \( V_i(G | \phi) \). Evidently, for each \( i \in \mathcal{I} \), we have

\[
V_i(G | \phi) = w_i(G) + \beta V_i(\phi(G) | \phi) \quad \text{for all } G \in \mathcal{G}. \tag{2.4}
\]

Intuitively, starting from \( G \in \mathcal{G} \), individual \( i \in \mathcal{I} \) receives a current payoff of \( w_i(G) \). Then \( \phi \) (uniquely) determines next period’s government \( \phi(G) \), and thus the continuation value of this individual, discounted to the current period, is \( \beta V_i(\phi(G) | \phi) \).

A government \( G \) is stable given mapping \( \phi \) if \( \phi(G) = G \). In addition, we say that \( \phi \) acyclic if for any (possibly infinite) chain \( H_0, H_1, \ldots \subset \mathcal{G} \) such that \( H_{k+1} \in \phi(H_k) \), and any \( a < b < c \), if \( H_a = H_c \) then \( H_a = H_b = H_c \).

The next definition shows that given (2.4), a political equilibrium can be summarized by the mapping \( \phi \) provided that two simple conditions are met.

**Definition 1** A mapping \( \phi : \mathcal{G} \rightarrow \mathcal{G} \) constitutes a (Markov) political equilibrium if for any \( G \in \mathcal{G} \), the following two conditions are satisfied:

\[4\] Throughout, we refer to this equilibrium concept as “political equilibrium” or simply as “equilibrium”. We do not use the acronym MPE, which will be used for Markov perfect equilibrium in Appendix B.
(i) either the set of players who prefer $\phi(G)$ to $G$ (in terms of discounted utility) forms a winning coalition, i.e., $S = \{i \in I : V_i(\phi(G) | \phi) > V_i(G | \phi)\} \in \mathcal{W}_G$, (or equivalently $|S| \geq m_G$ and $|S \cap G| \geq l_G$); or else, $\phi(G) = G$;

(ii) there is no alternative government $H \in \mathcal{G}$ that is preferred both to a transition to $\phi(G)$ and to staying in $G$ permanently, i.e., there is no $H$ such that $S' = \{i \in I : V_i(H | \phi) > V_i(\phi(G) | \phi)\} \in \mathcal{W}_G$ and $S'' = \{i \in I : V_i(H | \phi) > w_i(G) / (1 - \beta)\} \in \mathcal{W}_G$ (alternatively, $|S'| < m_G$, or $|S' \cap G| < l_G$, or $|S''| < m_G$, or $|S'' \cap G| < l_G$).

This definition states that a mapping $\phi$ constitutes a political equilibrium ("is a political equilibrium") if it maps the current government $G$ to alternative $\phi(G)$ that (unless it coincides with $G$) must be preferred to $G$ by a sufficient majority of the population and a sufficient number of current government members (so as not to be blocked). Part (ii) of the definition requires that there does not exist another alternative $H$ that would have been a "more preferable" transition; that is, there should be no $H$ that is preferred both to a transition to $\phi(G)$ and to staying in $G$ forever by a sufficient majority of the population and a sufficient number of current government members. The latter condition is imposed, since if there exists a subset $H$ that is preferred to a transition to $\phi(G)$ but not to staying in $G$ forever, then at each stage a move to $H$ can be blocked.\(^5\)

We take the definition of political equilibrium given in Definition 1 as a primitive and use it in this and the next section. The advantage of this definition is that it is simple and economical. A possible disadvantage is that it does not explicitly specify how offers for different types of transitions are made and the exact sequences of events at each stage.\(^6\) In Appendix B, we specify the sequences in which offers are made, voting takes place, and when transitions can take place explicitly, and characterize the Markov perfect equilibria of this extended environment. We show that for high discount factors, Markov perfect equilibria exist.

---

\(^5\) The explicit game form in Appendix B clarifies this further.

\(^6\) In this regard, our equilibrium concept is similar to the concept of Markov voting equilibrium in Roberts (1999).
equilibria coincide with our definition of political equilibria.\footnote{More specifically, acyclic Markov perfect equilibria are equivalent to acyclic political equilibria. In addition, we introduce the refinement of “order-independent” equilibria and show that these are also equivalent to acyclic political equilibria.}

### 2.3.2 Characterization

We now prove the existence and provide a characterization of political equilibria. We start with a recursive characterization of the mapping $\phi$ described in Definition 1. Let us enumerate the elements of the set $\mathcal{G}$ as $\{G_1, G_2, ..., G_{|\mathcal{G}|}\}$ such that $\Gamma_{G_x} > \Gamma_{G_y}$ whenever $x < y$. With this enumeration, $G_1$ is the most competent (“best”) government, while $G_{|\mathcal{G}|}$ is the least competent government. In view of Assumption 2.3, this enumeration is well defined and unique.

Now, suppose that for some $q > 1$, we have defined $\phi$ for all $G_j$ with $j < q$ where $q > 1$. Define the set

$$M_q \equiv \{j : 1 \leq j < q, \{i \in \mathcal{I} : w_i (G_j) > w_i (G_q)\} \in \mathcal{W}_{G_q}, \text{ and } \phi (G_j) = G_j\}.$$  \hspace{1cm} (2.5)

Note that this set depends simply on instantaneous utilities in (2.2), not on the discounted utilities defined in (2.4), which are “endogenous” objects. This set can thus be computed easily from the primitives of the model (for each $q$). Given this set, let the mapping $\phi$ be

$$\phi (G_q) = \begin{cases} 
G_q & \text{if } M_q = \emptyset; \\
G_{\min\{j \in M_q\}} & \text{if } M_q \neq \emptyset.
\end{cases} \hspace{1cm} (2.6)$$

Since the set $M_q$ is well-defined, this mapping is also well-defined, and by construction it is single-valued. Theorems 4 and 5 next show that, for sufficiently high discount factors, this mapping constitutes an acyclic political equilibrium and that, under additional mild conditions, it is the unique political equilibrium.

**Theorem 4** Suppose that Assumptions 2.1-2.3 hold and let $\phi : \mathcal{G} \rightarrow \mathcal{G}$ be as defined in (2.6). Then there exists $\beta_0 < 1$ such that for any discount factor $\beta > \beta_0$, $\phi$ is the unique acyclic political equilibrium.
Proof. See Appendix A. ■

Let us now illustrate the intuition for why the mapping $\phi$ constitutes a political equilibrium. Recall that $G_1$ is the most competent (“best”) government. It is clear that we must have $\phi(G_1) = G_1$, since all members of the population not in $G_1$ will prefer it to any other $G' \in \mathcal{G}$ (directly from Assumption 2.1). Assumption 2.2 then ensures that no winning coalition will be in favor of a permanent move to $G'$. However, $G'$ may not persist itself, and it may eventually lead to some alternative government $G'' \in \mathcal{G}$. But in this case, we can apply this reasoning to $G''$ instead of $G'$, and thus the conclusion $\phi(G_1) = G_1$ applies. Next considered a situation starting with government $G_2$ in power. The same argument applies if $G'$ is any one of $G_3, G_4, \ldots, G_{|\mathcal{G}|}$. One of these may eventually lead to $G_1$, thus for sufficiently high discount factor, a sufficient majority of the population may support the transition to such a $G'$ in order to eventually reach $G_1$. However, discounting also implies that a sufficient majority would also prefer a direct transition to $G_1$ to this dynamic path (recall part (ii) of Definition 1). So the relevant choice for the society is between $G_1$ and $G_2$. In this comparison, $G_1$ will be preferred if it has sufficiently many supporters, that is, if the set of individuals preferring $G_1$ to $G_2$ forms a winning coalition within $G_2$, or more formally if

$$\{i \in \mathcal{I} : w_i(G_1) > w_i(G_2)\} \in \mathcal{W}_{G_2}.$$  

If this is the case, $\phi(G_2) = G_1$; otherwise, $\phi(G_2) = G_2$. This is exactly what $\phi$ defined in (2.6) stipulates. Now let us start from government $G_3$. We then only need to consider the choice between $G_1$, $G_2$, and $G_3$. To move to $G_1$, it suffices that a winning coalition within $G_3$ prefers $G_1$ to $G_3$. However, whether the society will transition to $G_2$ depends on the stability of $G_2$. In particular, we may have the situation in which $G_2$ is not a stable government, which, by necessity, implies that $\phi(G_2) = G_1$. Then a transition to $G_2$ will lead to a transition to $G_1$ in the next period. However, this sequence may not be desirable.

---

8 If some winning coalition also prefers $G_2$ to $G_3$, then $G_1$ should still be chosen over $G_2$, because only members of $G_2$ who do not belong to $G_1$ prefer $G_2$ to $G_1$, and Assumption 2.2 ensures that those preferring $G_1$ over $G_2$ (starting in $G_3$) also form a winning coalition. Then a transition to $G_2$ is ruled out by part (ii) of Definition 1.
for those who prefer to move to $G_2$. In particular, we may have that

$$\{i \in I : w_i(G_2) > w_i(G_3)\} \in W_{G_3},$$

but

$$\{i \in I : w_i(G_1) > w_i(G_3)\} \notin W_{G_3}.$$  

If so, the transition from $G_3$ to $G_2$ may be blocked with the anticipation that it will lead to $G_1$ which does not receive the support of a winning coalition starting from $G_3$. This reasoning illustrates that for a transition to take place, not only the target government should be preferred to the current one by a winning coalition (starting from the current government), but also that the target government should be “stable,” i.e., $\phi(G') = G'$. This is exactly the requirement in (2.6). In this light, the intuition for the mapping $\phi$ and thus Theorem 4 is that a government $G$ will persist in equilibrium (will be stable) if there does not exist another stable government receiving support from a winning coalition (a sufficient majority of the population and the required number of current members of government).

Theorem 4 states that $\phi$ in (2.6) is the unique acyclic political equilibrium. However, it does not rule out other cyclic equilibria. We provide an example of a cyclic equilibrium in Appendix B. Cyclic equilibria are unintuitive and also “fragile”. We next strengthen Theorem 4 by providing (relatively mild) conditions under which they can be ruled out and $\phi$ in (2.6) is the unique political equilibrium (among both cyclic and acyclic ones).

**Theorem 5** The mapping $\phi$ defined in (2.6) is uniquely defined and is the unique political equilibrium (and equivalently, any political equilibrium is acyclic) if either of the following conditions holds:

1. For any $G \in \mathcal{G}$, $l_G \geq 1$.

2. $\theta > \varepsilon \cdot |\mathcal{G}|$, where $\theta \equiv \min\{i \in I \text{ and } G, H \in \mathcal{G}: i \in G \setminus H\} \{w_i(G) - w_i(H)\}$ and $\varepsilon \equiv \max\{i \in I \text{ and } G, H \in \mathcal{G}: i \in G \cap H\} \{w_i(G) - w_i(H)\}$.

**Proof.** See Appendix A. \hfill \qed
This theorem states two conditions under which there are no cyclic equilibria (thus making \( \phi \) in (2.6) the unique equilibrium). The first one is that if we always need the consent of at least one member of the current government for a transition to a new government, equilibria must be acyclic. This implies that acyclic equilibria are only possible if starting from some (but not all) governments, there is no incumbency advantage. The second part of the theorem provides a condition on preferences that also rules out cyclic equilibria. In particular, this condition states that if each individual receives sufficiently high utility from being in government (greater than \( \theta \)) and does not care much about the composition of the rest of the government (less than \( \varepsilon \), then all equilibria must be acyclic.

### 2.4 Characterization of Nonstochastic Transitions

In this section, we provide the comparison of different political regimes in terms of their ability to select governments with high levels of competence. To simplify the exposition and focus on the more important interactions, we assume that all feasible governments have the same size, \( k \in \mathbb{N} \), where \( k < n/2 \). More formally, let us define

\[ C^k = \{ Y \in \mathcal{C} : |Y| = k \}. \]

Then, \( \mathcal{G} = C^k \). In addition, we assume that for any \( G \in \mathcal{G} \), \( l_G = l \in \mathbb{N} \) and \( m_G = m \in \mathbb{N} \), so that the set of winning coalitions can be simply expressed as

\[ W_G = \{ X \in \mathcal{C} : |X| \geq m \text{ and } |X \cap G| \geq l \}, \quad (2.7) \]

where \( 0 \leq l \leq k < m \leq n - k \). This specification implies that given \( n, k, \) and \( m \), the number \( l \) corresponds to an inverse measure of democracy. If \( l = 0 \), then all individuals have equal weight and there is no incumbency advantage, thus we have a perfect democracy. In contrast, if \( l > 0 \), the consent of some of the members of the government is necessary for a change, thus we have an imperfect democracy. We thus have strengthened Assumption 2.2 to the following.
Assumption 2' We have that $\mathcal{G} = C^k$, and that there exist integers $l$ and $m$ such that the set a winning coalitions is given by (2.7).

Given this additional structure, equations (2.5) and (2.6) that determine the mapping $\phi$ can be written in a simpler form. Recall that governments are still ranked according to their level of competence, so that $G_1$ denotes the most competent government. Then we have:

$$\mathcal{M}_q = \{j : 1 \leq j < q, |G_k \cap G_q| \geq l, \text{ and } \phi(G_j) = G_j\},$$

(2.8)

and, as before,

$$\phi(G_q) = \begin{cases} G_q & \text{if } \mathcal{M}_q = \emptyset; \\ G_{\min\{j \in \mathcal{M}_q\}} & \text{if } \mathcal{M}_q \neq \emptyset. \end{cases}$$

(2.9)

Naturally, the mapping $\phi$ is again well-defined and unique. Finally, let us also define

$$\mathcal{D} = \{G \in \mathcal{G} : \phi(G) = G\}$$

as the set of stable governments (the fixed points of mapping $\phi$). If $G \in \mathcal{D}$, then $\phi(G) = G$, and this government will persist forever if it is the initial government of the society.

Given this more specific environment, we now investigate the structure of stable governments and how it changes as a function of the underlying political institutions, in particular, the degree of democracy. Throughout this section, we assume that Assumptions 1, 2' and 3 hold and we focus on order-independent MPE as characterized in Theorem 10. We do not add these qualifiers to any of the propositions to economize on space.

Our first proposition first provides an important technical result (part 1). It then uses this results to show that perfect democracy ensures the emergence of the best (most competent) government, but any departure from perfect democracy destroys this result and enables the emergence of highly incompetent/inefficient governments. It also shows that extreme dictatorship makes all initial governments stable, regardless of how low their competences may be.

**Proposition 2.1** The set of stable feasible governments $\mathcal{D}$ satisfies the following proper-
ties.

1. If $G, H \in \mathcal{D}$ and $|G \cap H| \geq l$, then $G = H$. In other words, any two distinct stable governments may have at most $l - 1$ common members.

2. Suppose that $l = 0$, so that the society is a perfect democracy. Then $\mathcal{D} = \{G_1\}$. In other words, starting from any initial government, the society will transition to the most competent government.

3. Suppose $l \geq 1$, so that the society is an imperfect democracy or a dictatorship. Then there are at least two stable governments, i.e., $|\mathcal{D}| \geq 2$. Moreover, the least competent governments may be stable.

4. Suppose $l = k$, so that the society is an extreme dictatorship. Then $\mathcal{D} = \mathcal{G}$, so any feasible government is stable.

Proof. See Appendix A.

Proposition 2.1 shows the fundamental contrast between perfect democracy, where there is no incumbency advantage, and other political institutions, which provide some additional power to “insiders” (current members of the government). With perfect democracy, the best government will necessarily emerge. With any deviation from perfect democracy, there will necessarily exist at least one other stable government (by definition less competent than the best) and even the worst government might be stable. The next example supplements Example 2.1 from the Introduction by showing a richer environment in which the least competent government is stable.

Suppose $n = 9$, $k = 3$, $l = 1$, and $m = 5$, so that a change in government requires support from a simple majority of the society, including at least one member of the current government. Suppose $I = \{1, 2, \ldots, 9\}$, and that instantaneous utilities are given by (2.3) in Example 2.2. Assume also that $\Gamma_{i_1, i_2, i_3} = 1000 - 100i_1 - 10i_2 - i_3$, provided that $i_1 < i_2 < i_3$.

Then $\{1, 2, 3\}$ is the most competent government, and is therefore stable. Any other government that includes 1 or 2 or 3 is unstable. For example, the government $\{2, 5, 9\}$
will transit to \( \{1, 2, 3\} \), as all individuals except 5 and 9 prefer the latter. However, government \( \{4, 5, 6\} \) is stable: any more competent government must include 1 or 2 or 3, and therefore is either \( \{1, 2, 3\} \) or will immediately transit to \( \{1, 2, 3\} \), which means that any such transition will not get support by any of the members of \( \{4, 5, 6\} \). Now, proceeding inductively, we find that any government other than \( \{1, 2, 3\} \) and \( \{4, 5, 6\} \) that contains at least one individual 1, 2, \( \ldots \), 6 is unstable. Consequently, government \( \{7, 8, 9\} \), which is the least competent government, is stable.

The next example shows that, starting with the same government, the long-run equilibrium government may be worse when political institutions are more democratic (as long as we are not in a perfect democracy).

*Take the setup from Example 2.4 (\( n = 9, k = 3, l = 1, \) and \( m = 5 \)), and suppose that the initial government is \( \{4, 5, 6\} \). As we showed there, government \( \{4, 5, 6\} \) is stable, and will therefore persist. Suppose, however, that \( l = 2 \) instead. In that case, \( \{4, 5, 6\} \) is unstable, and \( \phi(\{4, 5, 6\}) = \{1, 4, 5\} \), therefore, there will be a transition to \( \{1, 4, 5\} \). Since \( \{1, 4, 5\} \) is more competent than \( \{4, 5, 6\} \), this is an example where the long-run equilibrium government is worse under \( l = 1 \), when the institutions are more democratic, than under \( l = 2 \), when they are less democratic. Note that if \( l = 3 \), \( \{4, 5, 6\} \) would be stable again.*

When the number of individuals, \( n \), is sufficiently large, we can provide a tighter characterization of the structure of stable governments. In particular, the next proposition shows that for societies larger than a certain threshold (as a function of \( k \) and \( l \)), governments that contain no member of the ideal government and no member of any group of a prespecified size can always be found.

**Proposition 2.2** Suppose \( l \geq 1 \).

1. If

\[
n \geq 2k + k (k - l) \frac{(k - 1)!}{(l - 1)! (k - l)!},
\]

(2.10)
then there exists a stable government $G \in \mathcal{D}$ that contains no members of the ideal government $\mu_1$.

2. Take any $x \in \mathbb{N}$. If

$$n \geq k + x + x(k - 1)(k - l)\frac{(k - 1)!}{(l - 1)! (k - l)!}$$

then for any set of individuals $X$ with $|X| \leq x$ there exists a stable government $G \in \mathcal{D}$ such that $X \cap G = \emptyset$ (so no member of set $X$ belongs to $G$).

**Proof.** See Appendix A.

Let us provide the intuition for Proposition 2.2 when $l = 1$. Recall that $G_1$ is the most competent government. Let $G$ be the most competent government among those that do not include members of $G_1$ (such $G$ exists, since $n > 2k$ by assumption). In this case, Proposition 2.2 implies that $G$ is stable, that is, $G \in \mathcal{D}$. The reason is that if $\phi(G) = H \neq G$, then $\Gamma_H > \Gamma_G$, and therefore $H \cap G_1$ contains at least one element by construction of $G$. But then $\phi(H) = G_1$, as implied by (2.9). Intuitively, if $l = 1$, then once the current government contains a member of the most competent government $G_1$, this member will consent to (support) a transition to $G_1$, which will also receive the support of the population at large. He can do so, because $G_1$ is stable, thus there is no threats that a further round of transitions will harm him. But then, as in Example 2.1 in the Introduction, $G$ itself becomes stable, because any reform away from $G$ will take us to an unstable government.\(^9\)

We will next strengthen our characterization results by putting some more structure on the competences of governments. For this reason, suppose that each individual $i \in \mathcal{I}$ has a level of ability (or competence) given by $\gamma_i \in \mathbb{R}_+$ and suppose that the competence of the government is a strictly increasing function of the abilities of its members. More formally, we impose the following assumption throughout the rest of the analysis (again without explicitly stating it in each proposition).

\(^9\) It is also interesting to note that the upper bound on $X$ in Part 2 of Proposition 2.2 is $O(x)$, meaning that, increasing $x$ does not require an exponential increase in the size of population $n$ for Proposition 2.2 to hold.
Assumption 2.4 Suppose $G \in \mathcal{G}$, and individuals $i, j \in \mathcal{I}$ are such that $i \in G$, $j \notin G$, and $\gamma_i \geq \gamma_j$. Then $\Gamma_G \geq \Gamma_{(G \setminus \{i\}) \cup \{j\}}$.

Assumption 2.3 now implies that $\gamma_i \neq \gamma_j$ whenever $i \neq j$. The canonical form of the competence function consistent with Assumption 2.4 is

$$\Gamma_G = \sum_{i \in G} \gamma_i,$$

though for most of our analysis, we do not need to impose the specific functional form.

Given Assumption 2.4, it is also useful to enumerate individuals according to their abilities, so that $\gamma_i > \gamma_j$ whenever $i < j$. Recall also that $[x]$ denotes the integer part of a real number $x$.

When either $k = 1$ or $k = 2$, we have the following general result.

**Proposition 2.3**

1. If $k = 1$, then either $l = 0$ (perfect democracy), in which case $\phi(G) = \{G_1\} = \{1\}$ for any $G \in \mathcal{G}$, or $l = 1 = k$ (extreme dictatorship), in which case $\phi(G) = G$ for any $G \in \mathcal{G}$.

2. If $k = 2$, and $l = 0$ (perfect democracy), then $\phi(G) = G_1 = \{1, 2\}$ for any $G \in \mathcal{G}$.

If $k = 2$ and $l = 1$ (imperfect democracy), then if $G = \{p, q\}$ with $p < q$, then $\phi(G) = \{p - 1, p\}$ if $p$ is even and $\phi(G) = \{p, p + 1\}$ if $p$ is odd; in particular, $\phi(G) = G$ if and only if $p$ is odd and $q = p + 1$. If $k = 2$ and $l = 2$ (perfect dictatorship), then $\phi(G) = G$ for any $G \in \mathcal{G}$.

**Proof.** See Appendix A. ■

Proposition 2.3, though simple, provides an important insight about the structure of stable governments that will be further exploited in the next section. In the case of an extreme dictatorship, all governments are stable, so the initial government persists forever. In contrast, a perfect democracy ensures the emergence of the most competent feasible government. The case of imperfect democracy, highlighted in the case $k = 2$, lies somewhere in between. In that case, the competence of the limiting government is determined by the more able of the two members of the initial government. This means
that, with rare exceptions, the initial quality of government will be improved to some degree, but large improvements are not possible, because one of the two members must still be part of the ultimate government. Therefore, summarizing these three cases, we can say that with a perfect democracy, the best government will arise; with an extreme dictatorship, there will be no improvement in the initial government; and with an imperfect democracy, there will be some limited improvements in the quality of the government.

When \( k \geq 3 \) and \( l \geq 2 \), the structure of stable governments is more complex, and this is illustrated in the next example.

Suppose that \( k = 3 \) and \( l = 2 \). Then the following are stable governments that include the most able individual, 1: \( \{1, 2, 3\} \), \( \{1, 4, 5\} \), \( \{1, 6, 7\} \), and so on. Similarly, the following are stable governments that include individual 2 but not 1: \( \{2, 4, 6\} \), \( \{2, 5, 7\} \), and so on. The most competent government that does not include 1 and 2 might then be either \( \{3, 4, 7\} \) or \( \{3, 5, 6\} \), depending on which one is more competent.

2.5 Equilibria in Stochastic Environments

In this section, we introduce stochastic shocks to competences of different coalitions (or different individuals) in order to study the flexibility of different political institutions in their ability to adapt the nature and the composition of the government to changes in the underlying environment. Changes in the structure of appropriate governments may result from changes in economic, political, or social environment, which may in turn require different types of government to deal with the newly-emerging problems. Our main results will show the strong links between the degree of democracy and the flexibility to adapt to changing environments.

Changes in the environment are modeled succinctly by allowing changes in the function \( \Gamma_G : \mathcal{G} \rightarrow \mathbb{R} \), which determines the competence associated with each feasible government. Formally, we assume that at each \( t \), with probability \( 1 - \delta \), there is no change in \( \Gamma_G^t \) from \( \Gamma_G^{t-1} \), and with probability \( \delta \), there is a shock and \( \Gamma_G^t \) may change. In particular, following such a shock we assume that there exists a set of distribution functions \( F_t \left( \Gamma_G^t \mid \Gamma_G^{t-1} \right) \) that gives the conditional distribution of \( \Gamma_G^t \) at time \( t \) as functions of \( \Gamma_G^{t-1} \). We will focus
on the case where $\delta$ is small.

We first generalize our definition of (Markov) political equilibrium to this stochastic environment and generalize Theorem 4. We then provide a systematic characterization of political transitions in this stochastic environment and illustrate the links between the degree of democracy and institutional flexibility.

2.5.1 Stochastic Political Equilibria

The structure of stochastic political equilibria is complicated in general because individuals need to consider the implications of current transitions on future transitions under a variety of scenarios. Nevertheless, when the likelihood of stochastic shocks is sufficiently small, as we have assumed here, then political equilibria must have the logic similar to those given in Definition 1 in Section 2.3. Motivated by this reasoning, we introduce a simple definition of equilibrium here for stochastic political equilibria (would low probabilities of changes in the environment). Appendix B shows that when the discount factor is high and stochastic shocks are sufficiently infrequent, Markov perfect equilibria and our notions of (stochastic) political equilibrium are again equivalent.

To introduce the notion of \textit{(stochastic Markov) political equilibrium}, let us first consider a set of mappings $\phi_{\{\Gamma_G\}} : \mathcal{G} \rightarrow \mathcal{G}$ defined as in (2.6), but now separately for each $\{\Gamma_G\}_{G \in \mathcal{G}}$. These mappings are indexed by $\{\Gamma_G\}$ to emphasize this dependence. Essentially, if the configuration of competences of different governments were given by $\{\Gamma_G\}_{G \in \mathcal{G}}$ applied forever, we would be in a nonstochastic environment and $\phi_{\{\Gamma_G\}}$ would be the political equilibrium as shown by Theorems 4 and 5. The idea underlying our definition for this stochastic environment with infrequent changes is that while the current configuration is $\{\Gamma_G\}_{G \in \mathcal{G}}$, $\phi_{\{\Gamma_G\}}$ will still determine equilibrium behavior, because the probability of a change in competences is sufficiently small (see Appendix B).

**Definition 2** Let the set of mappings $\phi_{\{\Gamma_G\}} : \mathcal{G} \rightarrow \mathcal{G}$ (a separate mapping for each configuration $\{\Gamma_G\}_{G \in \mathcal{G}}$) be defined by the following two conditions. When the configuration of competences is given by $\{\Gamma_G\}_{G \in \mathcal{G}}$, we have that for any $G \in \mathcal{G}$:

(i) the set of players who prefer $\phi_{\{\Gamma_G\}}(G)$ to $G$ (in terms of discounted utility) forms
a winning coalition, i.e., $S = \{i \in I : V_i(\phi_{\{G\}}(G) | \phi_{\{G\}}) > V_i(G | \phi_{\{G\}})\} \in W_G$;

(ii) there is no alternative government $H \in \mathcal{G}$ that is preferred both to a transition to $\phi_{\{G\}}(G)$ and to staying in $G$ permanently, i.e., there is no $H$ such that $S' = \{i \in I : V_i(H | \phi_{\{G\}}) > V_i(\phi_{\{G\}}(G) | \phi_{\{G\}})\} \in W_G$ and $S'' = \{i \in I : V_i(H | \phi_{\{G\}}) > w_i(G)/(1-\beta)\} \in W_G$ (alternatively, $|S'| < m_G$, or $|S' \cap G| < l_G$, or $|S''| < m_G$, or $|S'' \cap G| < l_G$).

Then a set of mappings $\phi_{\{G\}} : \mathcal{G} \to \mathcal{G}$ constitutes a (stochastic Markov) political equilibrium for an environment with sufficiently infrequent changes if when $\{\Gamma_G\}_{G \in \mathcal{G}} = \{\Gamma_G\}_{G \in \mathcal{G}}$, there is a transition to government $G_{t+1}$ at time $t$ (starting with government $G_t$) if and only if $G_{t+1} = \phi_{\{G\}}(G_t)$.

Therefore, a political equilibrium with sufficiently infrequent changes involves the same political transitions (or the stability of governments) as that implied by the mappings $\phi_{\{G\}}$ defined in (2.6), applied separately for each configuration $\{\Gamma_G\}$.

The next theorem provides the general characterization of stochastic political equilibria in environments with sufficiently infrequent changes.

**Theorem 6** Suppose that Assumptions 2.1-2.3 hold and let $\phi_{\{G\}} : \mathcal{G} \to \mathcal{G}$ be the mapping defined by (2.6) applied separately for each configuration $\{\Gamma_G\}$. Then there exists $\beta_0 < 1$ such that for any discount factor $\beta > \beta_0$, $\phi_{\{G\}}$ is the unique acyclic political equilibrium.

**Proof.** See Appendix A. ■

The intuition for this theorem is straightforward. When shocks are sufficiently infrequent, the same calculus that applied in the nonstochastic environment still determines preferences because all agents put most weight on the events that will happen before such a change. Consequently, a stable government will arise and will remain in place until a stochastic shock occurs and changes the configuration of competences. Following such a shock, the stable government for this new configuration of competences emerges. Therefore, Theorem 6 provides us a tractable way of characterizing stochastic transitions. In the next subsection, we use this result to study the links between different political regimes and institutional flexibility.
2.5.2 The Structure of Stochastic Transitions

In the rest of this section, we provide further characterization results and the comparisons of this and regimes in this environment. In the results that follow, we always assume that Assumptions 1, 2, 3 and 2.4 hold, and we focus on acyclic political equilibria. We do not add these qualifiers to economize on notation in the statement of the propositions.

We also impose some additional structure on the distribution $F_T (\Gamma^t_G \mid \Gamma^{t-1}_G)$ by assuming that any shock corresponds to a rearrangement of the abilities of different individuals (and thus $F_T (\Gamma^t_G \mid \Gamma^{t-1}_G)$ gives the induced distribution of government competences according to Assumption 2.4). Put differently, we assume throughout this subsection that there is a fixed vector of abilities, say $a = \{a_1, ..., a_n\}$, and the actual distribution of abilities across individuals at time $t$, $\{\gamma^t_j\}_{j=1}^n$, is given by some permutation $\varphi^t$ of this vector $a$. We adopt the convention that $a_1 > a_2 > ... > a_n$. Intuitively, this captures the notion that a shock will change which individual is the best placed to solve certain tasks and thus most effective in government functions.

The next proposition shows the difference in flexibility implied by different political regimes. Throughout the rest of this section, our measure of “flexibility” is the probability with which the best government will be in power (either at given $t$ or as $t \to \infty$). More formally, let $\pi_t (l, k, n \mid G, \{\Gamma_G\})$ be the probability that in a society with $n$ individuals under a political regime characterized by $l$ and $k$, configuration of competence is given by $\{\Gamma_G\}$, and current government $G \in \mathcal{G}$, the most competent government will be in power at the time $t$. We will think of a regime characterized by $l'$ and $k'$ as more flexible than one characterized by $l$ and $k$ if $\pi_t (l', k', n \mid G, \{\Gamma_G\}) > \pi_t (l, k, n \mid G, \{\Gamma_G\})$ for all $G$ and $\{\Gamma_G\}$ for any $t$ following a stochastic shock. Similarly, we can think of the regime as asymptotically more flexible than another, if $\lim_{t \to \infty} \pi_t (l', k', n \mid G, \{\Gamma_G\}) > \lim_{t \to \infty} \pi_t (l, k, n \mid G, \{\Gamma_G\})$ for all $G$ and $\{\Gamma_G\}$ (provided that these limits are well defined). Clearly, “being more flexible” is a partial order.

**Proposition 2.4**

1. If $l = 0$ (i.e., perfect democracy), then a shock immediately leads to the replacement of the current government by the new most competent government.
2. If $l = 1$ (i.e., imperfect democracy), the competence of the government following a shock never decreases further; instead, it increases with probability no less than

$$1 - \frac{(k-1)! (n-k)!}{(n-1)!} = 1 - \binom{n-1}{k-1}^{-1}.$$ 

Starting with any $G$ and $\{\Gamma_G\}$, the probability that the most competent government will ultimately come to power as a result of a shock is

$$\lim_{t \to \infty} \pi_t (l, k, n | G, \{\Gamma_G\}) = \pi (l, k, n) \equiv 1 - \binom{n-k}{k} \binom{n}{k}^{-1} < 1.$$ 

For fixed $k$ as $n \to \infty$, $\pi (l, k, n) \to 0$.

3. If $l = k$ (i.e., extreme dictatorship), then a shock never leads to a change in government. The probability that the most competent government is in power at any given period (any $t$) after the shock is

$$\pi_t (l = k, k, n | \cdot, \cdot) = \binom{n}{k}^{-1}.$$ 

This probability is strictly less than $\pi_t (l = 0, k, n | G, \{\Gamma_G\})$ and $\pi_t (l = 1, k, n | G, \{\Gamma_G\})$ for any $G$ and $\{\Gamma_G\}$.

**Proof.** See Appendix A. ■

Proposition 2.4 contains a number of important results. A perfect democracy does not create any barriers against the installation of the best government at any point in time. Hence, under perfect democracy every shock is, flexibly, met by a change in government according to the wishes of the population at large (which here means that the most competent government will come to power). As we know from the analysis of Section 2.4, this is no longer true as soon as members of the governments have incumbency advantage. In particular, we know that without stochastic shocks, arbitrarily incompetent governments may come to power and remain in power. However, in the presence of shocks the evolution of equilibrium governments becomes more complex.

Even though the immediate effect of a shock may be a deterioration in government
competence, there are forces that increase government competence in the long run. This is most clearly illustrated in the case where \( l = 1 \). With this set of political institutions, there is zero probability that there will be a further decrease in government competence following a shock. Moreover, there is a positive probability that competence will improve and in fact a positive probability that, following a shock, the most competent government will be instituted. This is intuitive: a shock may make the current government unstable, and in this case, there will be a transition to a new stable government. A transition to a less competent government would never receive support from the population. The change in competences may be such that the only stable government after the shock, starting with the current government, may be the best government. Nevertheless, the probability of the most competent government coming to power, though positive, may be arbitrarily low. Proposition 2.4 also shows that when political institutions take the form of an extreme dictatorship, there will never be any transition, thus the current government can deteriorate following shocks (in fact, it can do so significantly).

Most importantly, Proposition 2.4 also shows that imperfect democracy has a higher degree of flexibility than dictatorship, ensuring better long-run outcomes (and naturally perfect democracy has the highest degree of flexibility). This unambiguous ranking between imperfect democracy and dictatorship in the presence of stochastic shocks contrasts with the results in Section 2.4, which showed that general comparisons between imperfect democracy and dictatorship are not possible in the nonstochastic case. This highlights that a distinct advantage of more democratic regimes might be their flexibility in the face of changing environments.

The next proposition strengthens the conclusions of Proposition 2.4. In particular, it establishes that the probability of having the most competent government in power is increasing in the degree of democracy more generally (i.e., it is decreasing in \( l \)).

---

10 This conclusion is not true for “expected competence” of the government, since we have not made “cardinal” assumptions on abilities. In particular, it is possible that some player is not a member of any stable government for some \( l \) and becomes part of a stable government for some \( l' < l \). If this player has very low ability, then expected competence under \( l' \) may be lower. In Appendix B, we provide an example illustrating this point, and we also show that expected competence of government is monotone in \( l \) when \( l \) is close to 0 or to \( k \).
Proposition 2.5 The probability of having the most competent government in power after a shock (for any \( t \)), \( \pi_t(l, k, n | G, \{ \Gamma_G \}) \), is decreasing in \( l \) for any \( k, n, G \) and \( \{ \Gamma_G \} \). That is, more democratic regimes are more likely to produce the most competent government.

Proof. See Appendix A. ■

The results of Propositions 2.4 and 2.5 can be strengthened further, when shocks are limited so that only the abilities of two (and in the second part of the proposition \( x \)) individuals and society are swapped. The next proposition contains these results.

Proposition 2.6 Suppose that any shock swaps the abilities of \( x \) individuals in the society.

1. If \( x = 2 \) (so that the abilities of two individuals are swapped at a time), then the competence of the government in power is nondecreasing in time \( (\pi_t(l, k, n | G, \{ \Gamma_G \}) \) is increasing in \( t \) for any \( l, k, n, G \) and \( \{ \Gamma_G \} \). Moreover, if the probability of swapping of abilities between any two individuals is positive, then the most competent government will be in power as \( t \rightarrow \infty \) with probability 1, that is, \( \lim_{t \rightarrow \infty} \pi_t(l, k, n | G, \{ \Gamma_G \}) = 1 \) (for any \( l, k, n, G \) and \( \{ \Gamma_G \} \)).

2. If \( x > 2 \), then the results in part 1 hold provided that \( l \leq k - \lfloor x/2 \rfloor \).

Proof. See Appendix A. ■

An interesting application of Proposition 2.6 is that when shocks are (relatively) rare and limited in their scope, relatively democratic regimes will gradually improve over time and institute the most competent government in the long run. This is not true for the most autocratic governments, however. This proposition, therefore, strengthens the conclusions of Propositions 2.4 and 2.5 in highlighting the flexibility benefits of more democratic regimes.

The political institutions considered so far are “junta-like” in the sense that no single member is essential. Incumbency advantage takes the form of the requirement that some members of the current government must consent to change. The alternative is a “royalty-like” environment where one or several members of the government are irreplaceable. All else equal, this can be conjecture to be a negative force, since it would mean that a
potentially low ability person must always be part of the government. However, the situation is more complex, because such an irreplaceable member (the member of the “royalty”) is also unafraid of changes, whereas, as we have seen, junta members would resist certain changes because of the further transitions that these will unleash.

More formally, we change Assumption 2.2 and the structure of the set of winning coalitions $W_G$ to accommodate “royalty-like” situations. We assume that there are $l$ royalty individuals whose votes are always necessary for a transition to be implemented (regardless of whether they are current government members). We denote the set of these individuals by $Y$. So, the new set of winning coalitions becomes

$$W_G = \{X \in C : |X| \geq m \text{ and } Y \subset X\}.$$ 

We also assume that all royal individuals are members of the initial government, that is, $Y \subset G^0$. The next proposition characterizes the structure of equilibrium in this case, with the focus on the expected competence of the government.

**Proposition 2.7** Suppose that we have a royalty-system with $1 \leq l < k$ and competences of governments are given by (2.12), so that the $l$ royals are never removed from the government. Then if $\{a_1, ..., a_n\}$ is sufficiently “convex” meaning that $\frac{a_1-a_2}{a_2-a_n}$ is sufficiently large, then the expected competence of the government under the royalty system is greater than under the original, junta-like system (with the same $l$). The opposite conclusion holds if $\frac{a_1-a_n-1}{a_{n-1}-a_n}$ is sufficiently low and $l = 1$.

**Proof.** See Appendix A. ■

Proposition 2.7 shows that royalty-like systems perform better in the face of shocks than junta-like systems when $\{a_1, ..., a_n\}$ is highly “convex,” which increases the benefit to society from having the highest ability individual in government. As discussed above, juntas are unlikely to lead to such high quality governments because of the fear of a change leading to a further round of changes, excluding all initial members of the junta. Royalty-like systems avoid this fear. When the ability of the royals is fixed (part 2), then they necessarily perform better. When the ability of the royals also changes, then they have a
disadvantage also, because they may keep a low ability royalty as part of the government. For this reason in part 1, we look at expected competence and show that when \( \{a_1, ..., a_n\} \) is sufficiently convex (so as to outweigh the loss of expected competence because of the presence of a potentially low ability royal), expected competence is higher under the royalty-like system. This result is interesting because it suggests that different types of dictatorships may have quite distinct implications for long-run quality of government and performance, and regimes that provide security to certain members of the incumbent government may be better at dealing with changes and in ensuring relatively high-quality governments in the long run.

2.6 Conclusion

In this paper, we provided a tractable dynamic model of political selection. The main barrier to the selection of good politicians and to the formation of good governments in our model is not the difficulty of identifying competent or honest politicians, but the incumbency advantage of current governments. Our framework shows how a small degree of incumbency advantage can lead to the persistence of highly inefficient and incompetent governments. This is because incumbency advantage implies that one of (potentially many) members of the government needs to consent to a change in the composition of government. However, all current members of the government may recognize that any change may unleash a further round of changes, ultimately unseating themselves. In this case, they will all oppose any change in government, even if such changes can improve welfare significantly for the rest of the society, and highly incompetent governments can remain in power.

Using this framework, we study the implications of different political institutions for the selection of governments both in nonstochastic and stochastic environments. A perfect democracy corresponds to a situation in which there is no incumbency advantage, thus citizens can nominate alternative governments and vote them to power without the need for the consent of any member of the incumbent government. In this case, we show that the most competent government will always come to power. However, interestingly,
any deviation from perfect democracy breaks this result and the long-run equilibrium government can be arbitrarily incompetent (relative to the best possible government). In extreme dictatorship, where any single member of the current government has a veto power on any change, the initial government always remains in power and this can be arbitrarily costly for the society. Perhaps more surprisingly, the same is true for any political institution other than perfect democracy. Moreover, there is no obvious ranking between different sets of political institutions (other than perfect democracy and extreme dictatorship) in terms of what they will imply for the quality of long run government. Even though no such ranking across political institutions is possible, we provide a fairly tight characterization of the structure of stable governments in our benchmark nonstochastic society.

In contrast, in stochastic environments, more democratic political regimes have a distinct advantage because of their greater “flexibility”. In particular, in stochastic environments, either the abilities and competences of individuals or the needs of government functions change, shuffling the ranking of different possible governments in terms of their competences and effectiveness. A greater degree of democracy then ensures greater “adaptability” or flexibility. Perfect democracy is most flexible and immediately adjusts to any shock by instituting the new government that has the greatest competence after the shock. Extreme dictatorship is at the other extreme and again leads to no change in the initial government. Therefore, shocks that reduce the competence of the individuals currently in power can lead to significant deterioration in the quality of government. Most interestingly, more democratic political institutions allow a greater degree of institutional flexibility in response to shocks. In particular, we show that shocks cannot lead to the emergence of a worse government (relative to the competence of the government the power following the shock). They may, however, destabilize the current government and induce the emergence of a more competent government. We show that political institutions with a greater degree of democracy have higher probability of improving the competence of the government following a shock and ultimately instituting the most competent government.

Finally, we also compare “junta-like” and “royalty-like” regimes. The former is our
benchmark society, where change in government requires the consent or support of one or multiple members of the current government. The second corresponds to situations in which one or multiple individuals are special and must always be part of the government (hence the title “royalty”). If royal individuals have low ability, royalty-like regimes can lead to the persistence of highly incompetent governments. However, we also show that in stochastic environments royalty-like regimes may lead to the emergence of higher quality governments in the long run than junta-like regimes. This is because royal individuals are not afraid of changes in governments, because their powers are absolute. In contrast, members of the junta may resist changes in government even if this increases its quality because such changes may lead to another round of changes, ultimately excluding all members of the initial government.

There are various interesting research areas highlighted by our analysis. Most important is an extension of this framework that combines the dynamic considerations emphasized here with the asymmetric information issues emphasize in the previous literature. For example, we can generalize the environment in this paper such that the ability of an individual is not observed until he becomes part of the government. In this case, to institute high quality governments, it is necessary to first “experiment” with different types of governments. The dynamic interactions highlighted by our analysis will then become a barrier to such experimentation. In this case, the set of political institutions that will ensure high quality governments must exhibit a different type of flexibility, whereby some degree of “churning” of governments can be guaranteed even without shocks. Another interesting area is to introduce additional instruments, so that some political regimes can provide incentives to politicians to take actions in line with the interests of the society at large. In that case, successful political institutions must ensure both the selection of high ability individuals and the provision of incentives to to these individuals once they are in government. We view these directions as interesting and important challenges for future research.
2.7 Appendix A

In the appendix, we use the following notation. First, we introduce the following binary relation on the set of feasible governments $\mathcal{G}$. For any $G, H \in \mathcal{G}$ we write

$$H \succ G \text{ if and only if } \{i \in \mathcal{I} : w_i(H) > w_i(G)\} \in W_G, \quad (2.13)$$

in other words, $H \succ G$ if and only if a winning coalition in $G$ prefers $H$ to $G$. Also, define set $D$ as

$$D = \{G \in \mathcal{G} : \phi(G) = G\}.$$

The next two lemmas summarize the properties of payoff functions and mapping $\phi$.

**Lemma 7** Suppose that $G, H \in \mathcal{G}$ and $\Gamma_G > \Gamma_H$. Then:

1. If for $i \in \mathcal{I}$, $w_i(G) < w_i(H)$, then $i \in H \setminus G$.
2. $H \not\succ G$.
3. $|\{i \in \mathcal{I} : w_i(G) > w_i(H)\}| > n/2 \geq \bar{k}$.

**Proof of Lemma 7. Part 1.** If $\Gamma_G > \Gamma_H$ then, by Assumption 2.1, $w_i(G) > w_i(H)$ whenever $i \in G$ or $i \notin H$. Hence, $w_i(G) < w_i(H)$ is possible only if $i \in H \setminus G$ (note that $w_i(G) = w_i(H)$ is ruled out by Assumption 2.3). At the same time, $i \in H \setminus G$ implies $w_i(G) < w_i(H)$ by Assumption 2.1, hence the equivalence.

**Part 2.** We have $|H \setminus G| \leq |H| \leq \bar{k} \leq m_G$; since by Assumption 2.2 $\bar{k} \leq m_G$, then $H \setminus G \notin W_G$, and $H \not\succ G$ by definition (2.13).

**Part 3.** We have $\{i \in \mathcal{I} : w_i(G) > w_i(H)\} = \mathcal{I} \setminus \{i \in \mathcal{I} : w_i(G) < w_i(H)\} \supset \mathcal{I} \setminus (H \setminus G)$, hence, $|\{i \in \mathcal{I} : w_i(G) > w_i(H)\}| \geq n - \bar{k} \geq n - n/2 = n/2 \geq \bar{k}$. ■

**Lemma 8** Consider the mapping $\phi$ constructed in Section 2.3 and let $G, H \in \mathcal{G}$. Then:

1. Either $\phi(G) = G$ (and then $G \in D$) or $\phi(G) \succ G$.
2. $\Gamma_{\phi(G)} \geq \Gamma_G$. 

82
3. If $\phi(G) \succ G$ and $H \succ G$, then $\Gamma_{\phi(G)} \geq \Gamma_H$.

4. $\phi(\phi(G)) = \phi(G)$.

**Proof of Lemma 8.** Straightforward. ■

**Proof of Theorem 4.** As before, we enumerate elements of $G$ as $\{G_1, G_2, \ldots, G_{|G|}\}$ such that $\Gamma_{G_x} > \Gamma_{G_y}$ whenever $x < y$. First, let us prove that $\phi$ constitutes a political equilibrium. Take any $G_q$, $1 \leq q \leq |G|$. By (2.6), either $\phi(G_q) = G_q$ or $\phi(G_q) \in M_q$. In the second case, the set of players who prefer $\phi(G_q)$ to $G_q$ is a winning coalition in $G_q$ by (2.5), so $V_i(\phi(G_q)) > V_i(G_q)$ for a winning coalition of players. Hence, in either case condition (i) of Definition 1 is satisfied.

Now, suppose, to obtain a contradiction, that condition (ii) of Definition 1 is violated, and $X, Y \in \mathcal{W}_{G_q}$ are winning coalitions such that $V_i(H) > \frac{w_i(G_q)}{1 - \beta}$ for all $i \in X$ and $V_i(H) > V_i(\phi(G_q))$ for all $i \in Y$. Suppose first that $\Gamma_{\phi(H)} \neq \Gamma_{\phi(G_q)}$, then $V_i(H) > V_i(\phi(G_q))$ would imply $w_i(\phi(H)) > w_i(\phi(G_q))$ as $\beta$ is close to 1, and hence the set of players who prefer $\phi(H)$ to $\phi(G_q)$ would be a winning coalition in $G_q$. This is impossible if $\Gamma_{\phi(H)} < \Gamma_{\phi(G_q)}$ (only players in $\phi(H)$ would possibly prefer $\phi(H)$, and they are less than $m_{G_q}$). If $\Gamma_{\phi(H)} > \Gamma_{\phi(G_q)}$, we would get a government $\phi(H)$ which is $\phi$-stable by construction of $\phi$ and which is preferred to $G_q$ by at least $m_{G_q}$ players (all except perhaps members of $G_q$) and at least $l_{G}$ members of $G_q$ (indeed, at least $l_{G}$ members of $G_q$ – those in coalition $X$ – had $w_i(\phi(H)) > \frac{w_i(G_q)}{1 - \beta}$, which then means they belong to $\phi(H)$, and hence must have $w_i(\phi(H)) > w_i(G_q)$) as $\phi(H)$ is stable. This would imply that $\phi(H) \in M_q$ by (2.5), but in that case $\Gamma_{\phi(H)} \geq \Gamma_{\phi(G_q)}$ would contradict (2.6).

Finally, consider the case $\Gamma_{\phi(H)} = \Gamma_{\phi(G_q)}$, which by Assumption 2.3 implies $\phi(H) = \phi(G_q)$. Now $V_i(H) > V_i(\phi(G_q))$ implies $w_i(H) > w_i(\phi(G_q))$ for all $i \in Y$. Since this includes at least $m$ players, we must have that $\Gamma_H \geq \Gamma_{\phi(G_q)}$. But $\Gamma_{\phi(H)} \geq \Gamma_H$ by (2.6), so $\Gamma_{\phi(H)} \geq \Gamma_H \geq \Gamma_{\phi(G_q)}$, which contradicts $\Gamma_{\phi(H)} = \Gamma_{\phi(G_q)}$. This contradiction proves that mapping $\phi$ satisfies both conditions Definition 1, and thus forms a political equilibrium.

To prove uniqueness, suppose that there is another acyclic political equilibrium $\psi$. For each $G \in G$, define $\chi(G) = \psi(G)$; due to acyclicity, $\chi(G)$ is $\psi$-stable for all $G$. We
prove the following sequence of claims. First, we must have that $\Gamma_{\chi(G)} \geq \Gamma_G$, for then condition (i) of Definition 1 would not be satisfied for large $\beta$.

Second, let us prove that all transitions must take place in one step, i.e., $\chi(G) = \psi(G)$ for all $G$. If this were not the case, then, due to finiteness of any chain of transitions there would exist $G \in \mathcal{G}$ such that $\chi(G) \neq \psi(G)$ but $\chi(G) = \psi^2(G)$. Take $H = \chi(G)$. Then $\Gamma_H > \Gamma_{\psi(G)}$, $\Gamma_H > \Gamma_G$ and $\psi(H) = H$. Since $\beta$ is large enough, the condition $V_i(H) > w_i(G)/(1-\beta)$ is automatically satisfied for the winning coalition of players who had $V_i(\phi(G)) > w_i(G)/(1-\beta)$. Let us prove that $V_i(H) > V_i(\phi(G))$ for a winning coalition of players in $G$. Note that this condition is equivalent to $w_i(H) > w_i(\phi(G))$. The fact that at least $m_G$ players prefer $H$ to $\phi(G)$ follows from $\Gamma_H > \Gamma_{\psi(G)}$. Moreover, since $\chi(G) = H$, at least $l_G$ members of $G$ must also be members of $H$; naturally, they prefer $H$ to $\phi(G)$. Consequently, condition (ii) of Definition 1 is violated. This contradiction proves that $\chi(G) = \psi(G)$ for all $G$.

Finally, let us prove that $\psi(G)$ coincides with $\phi(G)$ defined in (2.6). Suppose not, i.e., that $\phi(G) \neq \psi(G)$. Without loss of generality, we may assume that $G$ is the most competent government that satisfies this condition, i.e., $\phi(H) = \psi(H)$ whenever $\Gamma_H > \Gamma_G$. By Assumption 2.3, we have that $\Gamma_{\phi(G)} \neq \Gamma_{\psi(G)}$. Suppose that $\Gamma_{\phi(G)} > \Gamma_{\psi(G)}$ (the case $\Gamma_{\phi(G)} < \Gamma_{\psi(G)}$ is treated similarly). As $\psi(G)$ forms a political equilibrium, it must satisfy condition (ii) of Definition 1. Take $H = \phi(G)$. Since $\phi(G)$ is a political equilibrium, it must be that $w_i(\phi(G)) > w_i(G)$, and thus $V_i(H | \phi) > w_i(G) / (1-\beta)$, by a winning coalition of players. At least $m_G$ players prefer $H$ to $\psi(G)$ and, moreover, at least $l_G$ members of government $G$ have $V_i(H | \psi) > V_i(\psi(G) | \psi)$; the latter follows from that $V_i(H | \psi) = V_i(H | \phi) = w_i(H) / (1-\beta)$ as $H$ is both $\phi$-stable and $\psi$-stable by assumption, and from that the intersection of $H$ and $G$ contains at least $l_G$ members. The existence of such $H$ leads to a contradiction which completes the proof. ■

**Proof of Theorem 5. Part 1.** Suppose there is a cycle, so that there are $k$ governments $H_1, \ldots, H_k$ such that $\phi(H_j) = H_{j+1}$ for all $1 \leq j < k$, and $\phi(H_k) = H_1$. Without loss of generality, let $H_1$ be the most competent of these governments. Suppose that $i \in H_1$. In that case, $V_i(H_1) > V_i(H_2)$, as player $i$ gets the highest utility under $H_1$. 84
Since this holds for all members of \( H_1 \) and \( l_{H_1} \geq 1 \), it is impossible that for a winning coalition of players in \( H_1 \) the condition \( V_i(H_2) > V_i(H_1) \) is satisfied. This contradiction completes the proof.

**Part 2.** Again, suppose there is a cycle \( H_1, \ldots, H_k \) such that \( \phi(H_j) = H_{j+1} \) for all \( 1 \leq j < k \), and \( \phi(H_k) = H_1 \). Without loss of generality, suppose \( H_2 \) is the least competent of these governments. Consider government \( H_1 \); we have \( \phi(H_1) = H_2 \). Notice that at least \( l_G \) members of \( H_1 \) must be members of all governments \( H_1, \ldots, H_k \), for otherwise \( V_i(H_2) > V_i(H_1) \) would not hold for a winning coalition of players. However, these \( l_{H_1} \) players must have \( V_i(H_3) > V_i(H_2) \), as \( H_2 \) is the least competent government and they get the least utility there. This condition also holds for at least \( m_{H_1} \) players (all except, perhaps, members of \( H_2 \)). It remains to prove that \( V_i(H_3) > \phi_i(H_1) / (1 - \beta) \) for a winning coalition of players. However, we know that \( V_i(H_2) > \phi_i(H_1) / (1 - \beta) \) for a winning coalition, since \( \phi \) is a political equilibrium. We know that at most \( |H_2| \) could have preferred to move to \( H_2 \) because of \( H_2 \) rather than of continuation game, hence, at least \( m_{H_1} \) players have \( V_i(H_3) > \phi_i(H_1) / (1 - \beta) \). Moreover, the players who are members of all governments in the cycle have \( V_i(H_3) > \phi_i(H_1) / (1 - \beta) \) satisfied for the same reason (they preferred to move to \( H_2 \) in order to move further along the cycle). This proves that there exists \( H = H_3 \) which violates condition (ii) of Definition 1. This contradiction completes the proof. ■

**Proof of Proposition 2.1.** Part 1. Suppose, to obtain a contradiction, that \( |G \cap H| \geq l \), but \( G \neq H \); by Assumption 2.3 we need to have \( \gamma_G < \gamma_H \) or \( \gamma_G > \gamma_H \); without loss of generality assume the former. Then \( H > G \) by Lemma 7, since \( |G \cap H| \geq l \).

Note that \( G = G_q \) for some \( q \) and \( H = G_j \) for some \( j \) such that \( j < q \). Since \( H \) is stable, \( \phi(G_j) = G_j \), but then \( M_q \neq \emptyset \) by (2.5), and so \( \phi(G_q) \neq G_q \), as follows from (2.6).

However, this contradicts the hypothesis that \( G_q = G \in \mathcal{D} \), and thus completes the proof.

Part 2. By definition of mapping \( \phi \), \( \phi(G_1) = G_1 \), so \( G_1 \in \mathcal{D} \). Take any government \( G \in \mathcal{D} \); since \( |G \cap G_1| \geq 0 = L \), we have \( G = G_1 \) by part 1. Consequently, \( \mathcal{D} = \{G_1\} \), so \( \mathcal{D} \) is a singleton.

Part 3. As before, the most competent government, \( G_1 \), is stable, i.e. \( G_1 \in \mathcal{D} \). Now
consider the set of governments which intersect with $G_1$ by fewer than $l$ members:

$$B = \{ G \in \mathcal{G} : |G \cap G_1| < l \}.$$

This set is non-empty, because $n > 2k$ implies that there exist a government which does not intersect with $G_1$; obviously, it is in $B$. Now take the most competent government from $B$, $\mu_j$ where

$$j = \min \{ q : 1 \leq q \leq |G| \text{ and } G_q \in B \}.$$

Obviously, $G_j \neq G_1$, because $G_1 \notin B$. Let us show that $G_j$ is stable. Note that any government $G_q$ such that $\gamma_{G_q} > \gamma_{G_j}$ does not belong to $B$ and therefore has at least $l$ common members with stable government $G_1$. Hence, $\phi(G_q) = G_1$ (see (2.6)), and therefore $G_q$ is unstable, except for the case $q = 1$. Now we observe that set $\mathcal{M}_j$ is empty: for each government $\mu_q$ with $1 < q \leq j$ either the first condition in (2.8) is violated (if $q = 1$) or the second one (otherwise). But this implies that $\phi(\mu_j) = \mu_j$, so $\mu_j$ is stable. This proves that if $l \geq 1$, $\mathcal{D}$ contains at least two elements. Finally, note that this boundary is achieved: for example, if $l = 1$ and $n < 3k$.

Part 4. Take $l = 1$. If $n = ak$ where $a$ is an integer, then every individual is part of a stable government. More precisely, there are exactly $a$ stable governments in this case. Indeed, part 1 implies that no two stable governments can intersect, so the number of individuals who belong to some stable government is a multiple of $k$. If it is less than $n$, then there are at least $k$ individuals who are not part of any stable government, so we can form at least one government consisting of such individuals. Pick among such governments the most competent one; it is straightforward to show (like we did in part 3) that this government is stable, which is a contradiction to the assumption that these individuals belong to no stable government. Now assume that $n$ is not a multiple of $k$. Then again, any two stable governments must intersect, and therefore there must be a individual who is not part of a stable government. This completes the proof of Proposition 2.1 ■

Proof of Proposition 2.2. Part 1. We prove the more general part 2, then the statement of part 1 will be a corollary: to obtain (2.10), one only needs to substitute $b = k$
Let us prove the existence of such stable government. Define a set-valued function \( \chi : \mathcal{C}^L \rightarrow \mathcal{C}^{K-L} \cup \{ \emptyset \} \) by

\[
\chi (S) = \begin{cases} 
G \setminus S & \text{if } G \in \mathcal{D} \text{ and } S \subset G; \\
\emptyset & \text{if there exists no } G \in \mathcal{D} \text{ such that } S \subset G.
\end{cases}
\] (2.14)

In words, for any coalition of \( l \) individuals, function \( \chi \) assigns a coalition of \( k-l \) individuals such that their union is a stable government whenever such other coalition exists or an empty set when it does not exist. Note that \( \chi (S) \) is a well-defined single-valued function: indeed, there cannot be two different stable governments \( G \) and \( H \) which contain \( S \), for this would violate Proposition 2.1 (part 1).

Let \( Y_{l-1} \) be the coalition of \( l-1 \) individuals such that \( X \cap Y_{l-1} = \emptyset \); denote these individuals by \( i_1, \ldots, i_{l-1} \). We now add \( k-l+1 \) individuals to this coalition one by one. Let \( X_{l-1} = X \), and let

\[
X_l = X \cup Y_{l-1} \cup \left( \bigcup_{i \in X} \chi (Y_{l-1} \cup \{ i \}) \right).
\] (2.15)

Intuitively, we take the set of individuals which are either forbidden to join the government under construction (\( X \)) or are already there (\( Y_{l-1} \)), and add all individuals which can be in the same government with all individuals from \( Y_{l-1} \) and at least one individual from \( X_{l-1} = X \). Now take some individual \( i_l \in \mathcal{I} \setminus X_l \) (below we show that such individual exists) and let \( Y_l = Y_{l-1} \cup \{ i_l \} \). At each subsequent step \( z, l + 1 \leq z \leq k \), we choose \( z \)th individual for the government under construction as follows. We first define

\[
X_z = X \cup Y_{z-1} \cup \left( \bigcup_{S \subset Y_{z-1} : |S| = l-1; i \in X} \chi (S \cup \{ i \}) \right)
\] (2.16)

and then take

\[
i_z \in \mathcal{I} \setminus X_z
\] (2.17)

(we prove that we can do that later) and denote \( Y_z = Y_{z-1} \cup \{ i_z \} \). Let the last government
obtained in this way be denoted by $Y = Y_k$.

We now show that $\phi(Y) \cap X = \emptyset$. Suppose not, then there is individual $i \in \phi(Y) \cap X$. By (2.6) we must have that $|\phi(Y) \cap X| \geq l$; take the individual $i_j$ with the highest $j$ of such individuals. Clearly, $j \geq l$, so individual $i_j$ could not be a member of the initial $Y_{l-1}$ and was added at a later stage. Now let $S$ be a subset of $(\phi(Y) \cap X) \setminus \{i_j\}$ such that $|S| = l - 1$. Since government $\phi(Y)$ is stable and contains the entire $S$ as well as $i \in X$ (and $i \notin S$ because $S \subset Y$ and $X \cap Y = \emptyset$), we must have $\chi(S \cup \{i\}) = \phi(Y)$. Consequently, if we consider the right-hand side of (2.16) for $z = j$, we will immediately get that $\phi(Y) \subset X_j$, and therefore $i_j \in X_j$. But we picked $i_j$ such that $i_j \in I \setminus X_j$, according to (2.17). We get to a contradiction, which proves that $\phi(Y) \cap X = \emptyset$, so $\phi(Y)$ is a stable government which contains no member of $X$.

It remains to show that we can always pick such individual, we need to show that the number of individuals in $X_z$ is less than $n$ for any $z : l \leq z \leq k$. Note that the union in the inner parentheses of (2.16) consists of at most

$$(k - l) \left( \frac{z - 1}{l - 1} \right) b \leq (k - l) \left( \frac{k - 1}{l - 1} \right) b$$

individuals, while $z - 1 \leq k - 1$. Therefore, it is sufficient to require that

$$n > b + k - 1 + (k - l) \left( \frac{k - 1}{l - 1} \right) b$$

$$= b + k - 1 + b(k - l) \frac{(k - 1)!}{(l - 1)!(k - l)!}.$$

Because we are dealing with integers, this implies (2.11), which completes the proof.

Proof of Proposition 2.3. Part 1. By Assumption 2.2, $0 \leq l \leq k$, so either $l = 0$ or $l = 1$. If $l = 0$, then Proposition 2.1 (part 2) implies that the only stable government is $G_1$, so $\phi(G) = G_1$ for all $G \in \mathcal{G}$, where $G_1 = \{i_1\}$. If $l = 1$, then Proposition 2.1 (part 4) implies that any $G$ is stable.

Part 2. In this case, either $l = 0$, $l = 1$, or $l = 2$. If $l = 0$ or $l = 2$, the proof is similar to that of part 1 and follows from Proposition 2.1 (parts 2 and 4). If $l = 1$,
then \( \{i_1, i_2\} \) is the most competent, and hence stable, government. By 2.1 (part 1), any other government containing \( \{i_1, i_2\} \) is unstable. Hence, \( \{i_3, i_4\} \), the most competent government not containing \( i_1 \) or \( i_2 \), is stable. Proceeding likewise, we find that the only stable governments are \( \{i_{2j-1}, i_{2j}\} \) for \( 1 \leq j \leq n/2 \).

By the construction of mapping \( \phi \), either \( \phi(G) = G \) or \( |\phi(G) \cap G| = 1 \). If \( G = \{i_p, i_q\} \) with \( p < q \), then \( \phi(G) \) will include either \( i_p \) or \( i_q \). Now it is evident that \( \phi(G) \) will be the stable state which includes \( i_p \), because it is more competent than the one which includes \( i_q \) if the latter exists and is different. Full characterization follows.

**Proof of Theorem 6.** The proof follows immediately from Theorem 4 and the assumption that changes are sufficiently infrequent. Indeed, in the latter case, all the strict inequalities in Definitions 1 and 2 are preserved.

**Proof of Proposition 2.4. Part 1.** If \( l = 0 \), then by Proposition 2.1 for any \( G \),

\[
\phi^l(G) = G_1^l, \text{ where } G_1^l \text{ is the most competent government } \{i_1^l, \ldots, i_k^l\}.
\]

**Part 2.** Suppose \( l = 1 \), then Proposition 2.1 provides a full characterization. There are \( \lfloor n/k \rfloor \leq n/k \) stable governments. Each consists of \( k \) individuals, so the probability that a random new government coincides with any given stable government is \( 1 / \binom{n}{k} = \frac{k!(n-k)!}{n!} \).

The probability that it coincides with any stable government is \( \lfloor n/k \rfloor / \binom{n}{k} \leq \frac{n}{k} \frac{k!(n-k)!}{(n-1)!} = \frac{(k-1)!(n-k)!}{(n-1)!} = 1/(\binom{n-1}{k-1}) \). The government will change to a more competent one if and only if it is unstable, which happens with probability greater than or equal to \( 1 - 1/(\binom{n-1}{k-1}) \).

The most competent government will be installed if and only if after the shock, the government contains at least 1 of the \( k \) most competent members. The probability that it does not contain any of these equals \( \binom{n-k}{k} / \binom{n}{k} \) (this is the number of combinations that do not include \( k \) most competent members divided by the total number of combinations).

We have

\[
\binom{n-k}{k} / \binom{n}{k} = \frac{(n-k)! (n-k)!}{k! (n-2k)!(n-k)!} = \frac{(n-k)! (n-k)!}{(n-2k)! n!} = \prod_{j=0}^{k} \frac{n-k-j}{n-j}.
\]
Since each of the \( k \) factors tends to 1 as \( n \to \infty \), so does the product. Hence, the probability that the most competent government will arise, \( 1 - \left( \frac{n-k}{k} \right) \), tends to 0 as \( n \to \infty \).

**Part 3.** If \( l = k \), then \( \phi^t(G) = G \) for any \( t \) and \( G \). Hence, the government will not change. It will be the most competent if it contains \( k \) most competent individuals, which happens with probability \( \frac{1}{k} \). This is less than 1, which is the corresponding probability for \( l = 0 \). If \( k \geq 2 \), it is also less than the corresponding probability for \( l = 1 \): in the latter case, there are at least two governments which will lead to the most competent one: \( \{i_1^t, \ldots, i_k^t\} \) and \( \{i_1^t, \ldots, i_{k-1}^t, i_{k+1}^t\} \), which completes the proof. ■

**Proof of Proposition 2.5.** The probability of having the most competent government \( G_1^t \) is the probability that at least \( l \) members of \( G^t \) are members of \( G_1^t \). This probability equals (from hypergeometric distribution)

\[
\sum_{q=1}^{k} \binom{k}{q} \binom{n-k}{k-q} \frac{n}{\binom{n}{k}},
\]

and therefore is strictly decreasing in \( l \). ■

**Proof of Proposition 2.6.** Part 1. Any such swapping (or, more generally, any transposition \( \sigma \), where \( \sigma (i) \) is the individual whose former competence individual \( i \) now has) induces a one-to-one mapping that maps government \( G \) to government \( \rho(G) \): \( i \in \rho(G) \) if and only if \( \sigma (i) \in G \). By construction, \( \Gamma_{G^t}^{t-1} = \Gamma_{\rho(G)^t}^t \), and, by construction of mapping \( \phi \), \( \phi^{t-1}(G) = \phi^t(\rho(G)) \) for all \( G \). If all transitions occur in one stage, and a shock triggers a period of instability, then with probability 1 all shocks arrive at times \( t \) where government \( G^{t-1} \) is \( \phi^{t-1} \)-stable.

If abilities of only two individuals are swapped, then \(|G \cap \rho(G)| \geq k - 1 \geq l \). But \( G \) is \( \phi^{t-1} \)-stable with probability 1, hence, \( \rho(G) \) is \( \phi^t \)-stable. Consider two cases. If \( \Gamma_G^t \geq \Gamma_{\rho(G)}^t \), then \( \Gamma_{\phi(G)}^t \geq \Gamma_G^t \geq \Gamma_{\rho(G)}^t \). If \( \Gamma_G^t < \Gamma_{\rho(G)}^t \), then again \( \Gamma_{\phi(G)}^t \geq \Gamma_{\rho(G)}^t \), since there is a \( \phi^t \)-stable government \( \rho(G) \) which has with \( G \) at least \( l \) common members and the competence of which is \( \Gamma_{\rho(G)}^t \). Hence, \( \phi^t(G) \) is either \( \rho(G) \) or a more competent government. Hence, the competence of government cannot decrease. However, it may
increase, unless $G$ contains $k$ most competent members. Indeed, in that case there exist $i, j \in I$ with $i < j$ such that $i \notin G$ and $j \in G$. Obviously, swapping the abilities of these individuals increases the competence of $G$: $\Gamma^t_G > \Gamma^{t-1}_G$, and thus the stable government that will evolve will satisfy $\Gamma^t_{\phi(G)} \geq \Gamma^t_G > \Gamma^{t-1}_G$. Since the probability of this swapping is non-zero, eventually the competence of government will improve. Since there is a finite number of possible values of current government’s competence, then with probability 1 the most competent government will emerge.

**Part 2.** This follows from an argument of part 1, taking into account that if abilities of $x$ individuals changed, then $|G \cap \rho(G)| \geq k - \lfloor x/2 \rfloor \geq l$. Indeed, if $|G \cap \rho(G)| < k - \lfloor x/2 \rfloor$, we would have $|G \cap \rho(G)| \leq k - \lfloor (x + 1)/2 \rfloor$ since the numbers of both sides are integers, and thus $|(G \setminus \rho(G)) \cup (\rho(G) \setminus G)| > 2 \lfloor (x + 1)/2 \rfloor \geq x$. However, all individuals in $(G \setminus \rho(G)) \cup (\rho(G) \setminus G)$ changed their abilities, so the last inequality contradicts the assumption that no more than $x$ individuals did. This contradiction completes the proof.

**Proof of Proposition 2.7.** The probability of having the most able player in the government under the royalty system is 1. Indeed, government $\phi^t(G^t)$ will for any $t$ and any $G^t$ consist of $l$ irreplaceable members and $k - l$ most competent members. Since $l < k$, this always includes the most competent player. In the case of junta-like system, there is a positive probability that a government that does not include player $i^t_1$ is stable. If $\frac{a_1 - a_2}{a_2 - a_n}$ is sufficiently large, any government that includes player $i^t_1$ is more competent that a government that does not. The first part follows. Now consider the probability that the least competent player, $i^t_n$, is a part of the government. In a royalty system, this can happen if and only if the irreplaceable member is the least competent, i.e., with probability $1/n$. In a junta-like system, the probability that the most competent government is installed is the probability that one of the players $i^t_1, \ldots, i^t_k$ is in the government immediately after the shock. This probability is higher than the probability that any given player is among $i^t_1, \ldots, i^t_k$, which is $k/n \geq 2/n$. This completes the proof.
2.8 Appendix B

This Appendix contains the setup of the dynamic game, the proofs of results related to the dynamic game, Example 2.8.7 (which is also related to the dynamic game) and Example 2.8.7.

2.8.1 Dynamic Game

We have so far specified the set of winning coalitions and preferences over feasible governments and from being in office. To complete the description of the environment, we need to specify the order in which moves are made and how a government can be replaced by an alternative.

We first introduce an additional state variable, denoted by $v^t$, which determines whether the current government can be changed. In particular, $v^t$ takes two values: $v^t = s$ corresponds to a “sheltered” political situation (or “stable” political situation, though we will use the term stable for another purpose below) and $v^t = u$ designates an unstable situation. The government can only be changed during unstable times. A sheltered political situation destabilizes (becomes unstable) with probability $r$ in each period, that is, $P(v^t = u \mid v^{t-1} = s) = r$. These events are independent across periods and we also assume that $v^0 = u$. An unstable situation becomes sheltered when an incumbent government survives a challenge or is not challenged (as explained below).

We next describe the procedure for challenging an incumbent government. We start with some government $G^t$ at time $t$. If at time $t$ the situation is unstable, then all individuals $i \in \mathcal{I}$ are ordered according to some sequence $\eta_{G^t}$. Then each individual, in this order, nominates a subset of alternative governments $A^t_i \subset \mathcal{G} \setminus \{G^t\}$ that will be part of the primaries. An individual may choose not to nominate any alternative government, in which case he may choose $A^t_i = \emptyset$. All nominated governments (except the incumbent) make up the set $\mathcal{A}^t$, so

$$\mathcal{A}^t = \{G \in \mathcal{G} \setminus \{G^t\} : G \in A_i \text{ for some } i \in \mathcal{I}\}. \quad (2.18)$$
If $A^t \neq \emptyset$, then all alternatives in $A^t$ take part in the primaries at time $t$. The primaries take place as follows. All of the alternatives in $A^t$ are ordered $\pi_{G^t}^A (1), \pi_{G^t}^A (2), \ldots, \pi_{G^t}^A (|A^t|)$ according to some pre-specified order (depending on $A^t$ and the current government $G^t$). We refer to this order as the protocol, $\pi_{G^t}^A$. The primaries are then used to determine the challenging government $G' \in A^t$. In particular, we start with $G'_1$ given by the first element of the protocol $\pi_{G^t}^A (1)$. At the second step, $G'_1$ is voted against the second element, $\pi_{G^t}^A (2)$. We assume that all votings are sequential (and show in the Appendix that the sequence in which votes take place does not have any affect on the outcome). If more than $n/2$ of individuals support the latter, then $G'_2 = \pi_{G^t}^A (2)$; otherwise $G'_2 = G'_1$. Proceeding in order, $G'_3, G'_4, \ldots,$ and $G'_{|A^t|}$ are determined, and $G'$ is equal to the last element of the sequence, $G'_{|A^t|}$. This ends the primary.

After the primary, the challenger $G'$ is voted against the incumbent government $G^t$. $G'$ wins if and only if a winning coalition of individuals (i.e., a coalition that belongs to $W^t_{G^t}$) supports $G'$. Otherwise, we say that the incumbent government $G^t$ wins. If $A^t = \emptyset$ to start with, then there is no challenger and the incumbent government is again the winner.

If the incumbent government wins, it stays in power, and moreover the political situation becomes sheltered, that is, $G^{t+1} = G^t$ and $v^{t+1} = s$. Otherwise, the challenger becomes the new government, but the situation remains unstable, that is, $G^{t+1} = G'$ and $v^{t+1} = v^t = u$. All individuals receive instantaneous payoff $w_i (G^t)$ (we assume that the new government starts acting from the next period on).

More formally, the exact procedure is as follows.

- Period $t = 0, 1, 2, \ldots$ begins with government $G^t$ in power. If the political situation is sheltered, $v^t = s$, then each individual $i \in I$ receives instantaneous utility $u^t_i (G^t)$; in the next period, $G^{t+1} = G^t$, $v^{t+1} = v^t = s$ with probability $1 - r$ and $v^{t+1} = u$ with probability $r$.
- If the political situation is unstable, $v_t = u$, then the following events take place:
  1. Individuals are ordered according to $\eta_{G^t}$, and in this sequence, each individual
i nomi nates a subset of feasible governments $A_i^t \subset G \setminus \{G^t\}$ for the primaries. These determine the set of alternatives $A^t$ as in (2.18).

2. If $A^t = \emptyset$, then we say that the incumbent government wins, $G^{t+1} = G^t$, $v^{t+1} = s$, and each individual receives instantaneous utility $u^t_i (G^t)$. If $A^t \neq \emptyset$, then the alternatives in $A^t$ are ordered according to protocol $\pi^{A^t}_{G^t}$.

3. If $A^t \neq \emptyset$, then the alternatives in $A^t$ are voted against each other. In particular, at the first step, $G'_1 = \pi^{A^t}_{G^t} (1)$. If $\lvert A^t \rvert > 1$, then for $2 \leq j \leq \lvert A^t \rvert$, at step $j$, alternative $G'_{j-1}$ is voted against $\pi^{A^t}_{G^t} (j)$. Voting in the primary takes place as follows: all individuals vote yes or no sequentially according to some pre-specified order, and $G'_j = \pi^{A^t}_{G^t} (j)$ if and only if the set of the individuals who voted yes, $Y'_j$, is a simple majority (i.e., if $\lvert Y'_j \rvert > n/2$); otherwise, $G'_j = G'_{j-1}$.

The challenger is determined as $G' = G'_{\lvert A^t \rvert}$.

4. Government $G'$ challenges the incumbent government $G^t$, and voting in the election takes place. In particular, all individuals vote yes or no sequentially according to some pre-specified order, and $G'$ wins if and only if the set of the individuals who voted yes, $Y^t$, is a winning coalition in $G^t$ (i.e., if $Y^t \in W^t_{G^t}$); otherwise, $G^t$ wins.

5. If $G^t$ wins, then $G^{t+1} = G^t$, $v^{t+1} = s$; if $G'$ wins, then $G^{t+1} = G'$, $v^{t+1} = u$. In either case, and each individual gets instantaneous utility $u^t_i (G^t)$.

There are several important features about this dynamic game that are worth emphasizing. First, the set the winning coalitions, $W^t_{G^t}$ when the government is $G^t$, determines which proposals for governmental change are accepted. Second, to specify a well-defined game we had to introduce the pre-specified order $\eta_G$ in which individuals nominate alternatives for the primaries, the protocol $\pi^{A^t}_{G^t}$ for the order in which alternatives are considered, and also the order in which votes are cast. Ideally we would like these orders not to have a major influence on the structure of equilibria, since they are not an essential part of the economic environment and we do not have a good way of mapping the specific orders to reality. We will see that this is indeed the case in the equilibria of interest. Finally, the
rate at which political situations become unstable, \( r \), has an important influence on payoffs by determining the rate at which opportunities to change the government arise. In what follows, we will typically suppose that \( r \) is relatively small, so that political situations are not unstable most of the time. Here, it is also important that political instability ceases after the incumbent government withstands a challenge (or if there is no challenge). This can be interpreted as the government having survived a “no-confidence” motion. We will also focus on situations in which the discount factor \( \beta \) is large.

2.8.2 Strategies and Definition of Equilibrium

We define strategies and equilibria in the usual fashion. In particular, let \( h^{t,Q^t} \) denote the history of the game up to period \( t \) and stage \( Q^t \) in period \( t \) (there are several stages in period \( t \) if \( v^t = u \)). This history includes all governments, all proposals, votes and stochastic events up to this time. The set of histories is denoted by \( \mathcal{H}^{t,Q^t} \). A history \( h^{t,Q^t} \) can also be decomposed into two parts. We can write \( h^{t,Q^t} = (h^t, Q^t) \) and correspondingly, \( \mathcal{H}^{t,Q^t} = \mathcal{H}^t \times Q^t \), where \( h^t \) summarizes all events that have taken place up to period \( t - 1 \) and \( Q^t \) is the list of events that have taken place within the time instant \( t \) when there is an opportunity to change the government.

A strategy for individual \( i \in \mathcal{I} \), denoted by \( \sigma_i \), maps \( \mathcal{H}^{t,Q^t} \) (for all \( t \) and \( Q^t \)) into a proposal when \( i \) nominates an alternative government (i.e., at the first stage of the period where \( v^t = u \)) and a vote for each possible proposal at each possible decision node (recall that the ordering of alternatives is automatic and is done according to a protocol). A Subgame Perfect Equilibrium (SPE) is a strategy profile \( \{\sigma_i\}_{i \in \mathcal{I}} \) such that the strategy of each \( i \) is the best response to the strategies of all other individuals for all histories.

Throughout the rest of the paper, we focus on the Markovian subset of SPEs, and equilibrium refers to Markov Perfect Equilibrium (MPE) in pure strategies. More formally:

**Definition 3** A Markov Perfect Equilibrium is a SPE profile of strategies \( \{\sigma^*_i\}_{i \in \mathcal{I}} \) such that \( \sigma^*_i \) for each \( i \) in each period \( t \) depends only on \( G^t \), \( \Gamma^t \), \( W^t \), and \( Q^t \) (previous actions taken in period \( t \)).

MPEs are natural in such dynamic games, since they enable individuals to condition
on all of the payoff-relevant information, but rule out complicated trigger-like strategies, which are not our focus in this paper. It turns out that even MPEs potentially lead to a very rich set of behavior. For this reason, it is also useful to consider subsets of MPEs, in particular, acyclic MPEs and order-independent MPEs. Loosely speaking, an equilibrium is acyclic if cycles (changing the initial government but then reinstalling it at some future date) do not take place along the equilibrium path. Cyclical MPEs are both less realistic and also more difficult to characterize, motivating our main focus on acyclic MPEs. Formally, we have:

**Definition 4** An MPE $\sigma^*$ is cyclic if the probability that there exist $t_1 < t_2 < t_3$ such that $G^{t_3} = G^{t_1} \neq G^{t_2}$ along the equilibrium path is positive. An MPE $\sigma^*$ is acyclic if it is not cyclic.

We also refer to acyclic MPEs as acyclic equilibria. Another relevant subset of MPEs, order-independent MPEs or simply order-independent equilibria, is introduced by Moldovanu and Winter (1995). These equilibria impose that strategies should not depend on the order in which certain events, in particular here the order of proposal-making, unfold. Here we generalize (and slightly modify) their definition for our present context. For this purpose, let us denote the above-described game when the set of protocols is given by $\pi = \{\pi^G\}_{G \in \mathcal{G}, A^t \in \mathcal{P}(G), G \in \mathcal{A}^t}$ as $\text{GAME}[\pi]$ and denote the set of feasible protocols by $\Pi$.

**Definition 5** Consider $\text{GAME}[\pi]$. Then $\sigma^*$ is an order-independent equilibrium for $\text{GAME}[\pi]$ if for any $\pi' \in \Pi$, there exists an equilibrium $\sigma'^*$ of $\text{GAME}[\pi']$ such that $\sigma^*$ and $\sigma'^*$ lead to the same distributions of equilibrium governments $G^\tau \mid G^t$ for $\tau > t$.

We will establish the relationship between acyclic and order-independent equilibria in Theorem 10.\(^\text{11}\)

\(^1\) One could also require order independence with respect to $\eta$ as well as with respect to $\pi$. It can be easily verified that the equilibria we focus on already satisfy this property and hence, this is not added as a requirement of “order independence” in Definition 5.
Recall the mapping defined by $\phi : \mathcal{G} \to \mathcal{G}$ be the mapping defined by (2.6). We use the next theorem to establish the equivalence between political equilibria and MPE in the dynamic game.

**Theorem 9** Consider the game described above. Suppose that Assumptions 2.1-2.3 hold and let $\phi : \mathcal{G} \to \mathcal{G}$ be the political equilibrium given by (2.6). Then there exists $\varepsilon > 0$ such that if $\beta < 1 - \varepsilon$ and $r/(1 - \beta) > \varepsilon$ then for any protocol $\pi \in \Pi$:

1. There exists an acyclic MPE in pure strategies $\sigma^*$.

2. Take an acyclic MPE in pure or mixed strategies $\sigma^*$. Then under $\sigma^*$, we have that:
   - if $\phi (G^0) = G^0$, then there are no transitions; and
   - if $\phi (G^0) \neq G^0$, then with probability 1 there exists a period $t$ where the government $\phi (G^0)$ is proposed, wins the primaries, and wins the power struggle against $G^t$. After that, there are no transitions, so $G^\tau = \phi (G^0)$ for all $\tau \geq t$.

The intuition for this theorem is provided by the construction of the mapping $\phi$. The hypothesis that $r$ is sufficiently small ensures that stable political situations are sufficiently stable, so that if the government passes a “no-confidence” voting, it stays for some time. As it turns out, this is crucial to ensure that an equilibrium in pure strategies exists (which in turn allows us to obtain a characterization of equilibria). Example 2.8.7 in Appendix B illustrates the potential problem. We also assumed that the discount factor $\beta$ is sufficiently large; the reason for this assumption is that we want the individuals to be forward-looking and take into account the transitions that will happen further along the equilibrium path.

We should note at this point that the acyclicity requirement is not redundant. Example 2.8.4 in the next section shows that there may be cyclic equilibria. for a certain set of protocols $\pi \in \Pi$. Cyclic equilibria are less natural than the equilibria we focus on and this motivates our focus on acyclic equilibria, while the discussion in the previous paragraph motivates us to consider the case with high $\beta$ and small $r$. 
It is also noteworthy that part 2 of Theorem 9 ensures that in any acyclic equilibrium there will eventually be a transition to government $\phi(G^0)$, but such a transition can be slow and in multiple steps. In the next section, we establish that order-independent equilibria not only ensure acyclicity, but also that all transitions take place "rapidly," i.e., at the first unstable period (which is period $t = 0$ by assumption).

2.8.4 Cycles, Acyclicity, and Order-Independent Equilibria

We first show that cyclic MPEs are possible. The next example illustrates this.

Consider a society consisting of five individuals ($n = 5$). The only feasible governments are $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}$. Suppose that there is “full democracy,” so that in terms of Assumption 2.2 $l_G = l = 0$ for $G \in \mathcal{G}^k$, and that voting takes the form of a simple majority rule, so that again with the same notation $m_G = m = 3$ for all $G$. Suppose also that the competences of different feasible governments are given by $f_1 g, f_2 g, f_3 g, f_4 g, f_5 g$. Suppose also that the competences of different feasible governments are given by

$$\Gamma_{\{i\}} = 5 - i,$$

so $\{1\}$ is the best government.

Assume also that instantaneous utilities are given as in Example 2.2. In particular,

$$w_i(G) = \Gamma_G + 100I_{\{i \in G\}}.$$

These utilities imply that each individual receives a high value from being part of the government relative to the utility she receives from government competence.

Finally, we define the protocols $\pi^A_G$ as follows. If $G = \{1\}$, then $\pi^G_\{\{1\}\} = \pi^\{\{2\},\{3\},\{4\},\{5\}\} = (\{3\}, \{4\}, \{5\}, \{2\})$ and $\pi^A_{\{1\}}$ for $A \neq (\{3\}, \{4\}, \{5\}, \{2\})$ is obtained from $\pi^\{2,3,4,5\}_{\{1\}}$ by dropping governments which are not in $A$: for example, $\pi^\{2,3,5\}_{\{1\}} = (\{3\}, \{5\}, \{2\})$. For other governments, we define $\pi^\{1\},\{3\},\{4\},\{5\}\}_{\{2\}} = (\{4\}, \{5\}, \{1\}, \{3\}), \pi^\{1\},\{2\},\{4\},\{5\}\}_{\{3\}} = (\{5\}, \{1\}, \{2\}, \{4\}), \pi^\{1\},\{2\},\{3\},\{5\}\}_{\{4\}} = (\{1\}, \{2\}, \{3\}, \{5\})$ and $\pi^\{1\},\{2\},\{3\},\{4\}\}_{\{5\}} = (\{2\}, \{3\}, \{4\}, \{1\})$, and for other $A$.
again define $\pi^A$ by dropping the governments absent in $A$. Then there exists an equilibrium where the governments follow a cycle of the form $\{5\} \rightarrow \{4\} \rightarrow \{3\} \rightarrow \{2\} \rightarrow \{1\} \rightarrow \{5\} \rightarrow \cdots$.

To verify this claim, consider the following nomination strategies by the individuals. If the government is $\{1\}$, two individuals nominate $\{2\}$ and other three nominate $\{5\}$; if it is $\{2\}$, two individuals nominate $\{3\}$ and three nominate $\{1\}$; if it is $\{3\}$, two nominate $\{4\}$ and three nominate $\{2\}$; if it is $\{4\}$, two nominate $\{5\}$ and three nominate $\{3\}$; if it is $\{5\}$, two nominate $\{1\}$ and three nominate $\{4\}$.

Let us next turn to voting strategies. Here we appeal to Lemma 1 from Acemoglu, Egorov and Sonin (2008), which shows that in this class of games, it is sufficient to focus on strategies in which individuals always vote for the alternative yielding the highest payoff for them at each stage. Then, note that in equilibrium, any alternative government which wins the primaries, on or off equilibrium path, subsequently wins against the incumbent government. In particular, in such an equilibrium supporting the incumbent government will break a cycle, but only one person (the member of the incumbent government) is in favor of it. We next show if only one individual deviates at the nomination stage, then next government in the cycle still wins in the primaries. Suppose that the current government is $\{3\}$ (other cases are treated similarly). Then by construction, governments $\{2\}$ and $\{4\}$ are necessarily nominated, and perhaps $\{1\}$ or $\{5\}$ also are. Notice that if the last voting in the primaries is between $\{2\}$ and $\{4\}$, then $\{2\}$ wins: indeed, all individuals know that both alternatives can take over the incumbent government, but $\{2\}$ is preferred by individuals 1, 2, and 5 (because they would want to be government members earlier rather than later). If, however, the last stage involves voting between $\{4\}$ on the one hand and either $\{1\}$ or $\{5\}$ on the other, then $\{4\}$ wins for similar reason. Now, if either $\{1\}$ or $\{5\}$ is nominated, then in the first voting it is voted against $\{2\}$. All individuals know that accepting $\{2\}$ will ultimately lead to a transition to $\{2\}$, whereas supporting $\{1\}$ or $\{5\}$ will lead to $\{4\}$. Because of that, at least three individuals (1, 2, 5) will support $\{2\}$. This proves that $\{2\}$ will win against the incumbent government $\{3\}$, provided that $\{2\}$ and
participate in the primaries, which is necessarily the case if no more than one individual deviates. This, in turn, implies that nomination strategies are also optimal in the sense that there is no profitable one-shot deviation for any individual. We can easily verify that this holds for other incumbent governments as well.

We have thus proved that the strategies we constructed form a SPE; since they are also Markovian, it is a MPE as well. Along the equilibrium path, the governments follow a cycle \(5 \to 4 \to 3 \to 2 \to 1 \to 5 \to \cdots\). We can similarly construct a cycle that moves in the other direction: \(1 \to 2 \to 3 \to 4 \to 5 \to 1 \to \cdots\) (though this would require different protocols). Hence, for some protocols, cyclic equilibria are possible.

Intuitively, a cycle enables different individuals that will not be part of the limiting (stable) government to enjoy the benefits of being in power. This example, and the intuition we suggest, also highlight that even when there is a cyclic equilibrium, an acyclic equilibrium still exists (see, in particular, Theorem 10). Moreover, the acyclic and order-independent equilibria have additional desirable properties.

Example 2.8.4 shows that the order in which proposals are made may be crucial for supporting the cyclic equilibrium. In particular, it is straightforward to verify that the equilibrium does not survive if we take \(\pi_G^{1t}\) to be the same for all \(G\)’s. We can also construct examples in which there are no cycles, but the emergence of a stable government takes several transitions. This is illustrated in the next example. Theorem 10 below will show that when we focus on order-independent MPE, both cyclic equilibria and equilibria with multi-step transitions will be ruled out.

Take the setup of Example 2.8.4, with the exception that \(l_{\{1\}} = 1\) (so that consent of individual 1 is needed to change the government when the government is \(\{1\}\)). It is then easy to check that the strategy profile constructed in Example 2.8.4 is a MPE in this case as well. However, since individual 1 will vote against any alternative which wins the primaries, the difference is that alternative \(\{5\}\) will not be accepted in equilibrium and government \(\{1\}\) will persist. Hence, in equilibrium, the transitions are as follows: \(5 \to 4 \to 3 \to 2 \to 1\).
We now establish that order-independent equilibria always exist, are always acyclic, and lead to rapid (one-step) equilibrium transitions. As such, this theorem will be the starting point of our more detailed characterization of the influence of political institutions on the selection of politicians and governments in the next section. The proof of this theorem also requires a slightly stronger version of Assumption 2.3, which we now introduce.

**Assumption 3’** For any \( i \in I \) and any sequence of feasible governments, \( H_1, H_2, \ldots, H_q \in \mathcal{G} \) (for \( q \geq 2 \)), we have
\[
  w_i (H_1) \neq \frac{\sum_{j=2}^{q} w_i (H_j)}{q-1}.
\]

Recall that Assumption 2.3 imposed that no two feasible governments have exactly the same competence. Assumption 3’ strengthens this and requires that the competence of any government should not be the average of the competences of other feasible governments. Like Assumption 2.3, Assumption 3’ is satisfied “generically,” in the sense that if it were not satisfied for a society, any small perturbation of competence levels would restore it.

**Theorem 10** Consider the game described above. Suppose that Assumptions 2.1, 2.2 and 3’ hold and let \( \phi : \mathcal{G} \rightarrow \mathcal{G} \) be the political equilibrium defined by (2.6). Then there exists \( \varepsilon > 0 \) such that if \( \beta < 1 - \varepsilon \) and \( r/(1 - \beta) > \varepsilon \) for any protocol \( \pi \in \Pi \):

1. There exists an order-independent MPE in pure strategies \( \sigma^* \).
2. Any order-independent MPE in pure strategies \( \sigma^* \) is acyclic.
3. In any order-independent MPE \( \sigma^* \), we have:
   - if \( \phi (G^0) = G^0 \), then there are no transitions and government \( G^t = G^0 \) for each \( t \);
   - if \( \phi (G^0) \neq G^0 \), then there is a transition from \( G^0 \) to \( \phi (G^0) \) in period \( t = 0 \), and there are no more transitions: \( G^t = \phi (G^0) \) for all \( t \geq 1 \).
4. In any order-independent MPE $\sigma^*$, the payoff of each individual $i \in \mathcal{I}$ is given by

$$u_i^0 = w_i(G^0) + \frac{\beta}{1 - \beta} w_i(\phi(G^0)).$$

2.8.5 Stochastic Characterization

In addition, throughout this section, we simplify the discussion by assuming that Assumption 2.4 holds and by focusing on order-independent MPEs. The next theorem generalizes Theorem 10 to this stochastic environment.

**Theorem 11** Consider the above-described stochastic environment. Suppose that Assumptions 2.1, 2.2, 3, and 2.4 hold. Let $\phi^s : \mathcal{G} \to \mathcal{G}$ be the political equilibrium defined by (2.6) for $\Gamma^s_G$. Then there exists $\varepsilon > 0$ such that if $\beta < 1 - \varepsilon$, $r/(1 - \beta) > \varepsilon$ and $\delta > \varepsilon$ then for any protocol $\pi \in \Pi$, we have the following results.

1. There exists an order-independent MPE in pure strategies.

2. Suppose that between periods $t_1$ and $t_2$ there are no shocks. Then in any order-independent MPE in pure strategies, the following results hold:

   - if $\phi(G^{t_1}) = G^{t_1}$, then there are no transitions between $t_1$ and $t_2$;
   - if $\phi(G^{t_1}) \neq G^{t_1}$, then alternative $\phi(G^{t_1})$ is accepted during the first period of instability (after $t_1$).

2.8.6 Proofs

**Proof of Theorem 9. Part 1.** In this proof, we use of Lemma 7. We take $\beta_0$ such that for any $\beta > \beta_0$ the following inequalities are satisfied:

$$\forall G, G', H, H' \in \mathcal{G} \text{ and } i \in \mathcal{I} : w_i(G) < w_i(H) \text{ implies } \left(1 - \beta^{\mathcal{G}_i}\right) w_i(G') + \beta^{\mathcal{G}_i} w_i(G) < \left(1 - \beta^{\mathcal{G}_i}\right) w_i(H') + \beta^{\mathcal{G}_i} w_i(H).$$  \hspace{1cm} (2.19)
For each $G \in \mathcal{G}$, define the following mapping $\chi_G : \mathcal{G} \to \mathcal{G}$:

$$
\chi_G (H) = \begin{cases} 
\phi (H) & \text{if } H \neq G \\
G & \text{if } H = G
\end{cases}.
$$

Take any protocol $\pi \in \Pi$. Now take some node of the game in the beginning of some period $t$ when $\nu^t = u$. Consider the stages of the dynamic game that take place in this period as a finite game by assigning the following payoffs to the terminal nodes:

$$
v_i (G, H) = \begin{cases} 
\frac{1+r\beta}{1-\beta(1-r)} w_i (H) + \frac{\beta}{1-\beta} w_i (\phi (H)) & \text{if } H \neq G \\
\frac{1+r\beta}{1-\beta(1-r)} w_i (G) + \frac{r\beta^2}{(1-\beta)(1-\beta(1-r))} w_i (\phi (G)) & \text{if } H = G
\end{cases},
$$

(2.20)

where $H = G^{t+1}$ is the government that is scheduled to be in power in period $t + 1$, i.e., the government that defeated the incumbent $G^t$ if it was defeated and $G^t$ itself if it was not. For any such period $t$, take a SPE in pure strategies $\sigma_G^* = \sigma_{G^t}^*$ of the truncated game, such that this SPE is the same for any two nodes with the same incumbent government; the latter requirement ensures that once we map these SPEs to a strategy profile $\sigma^*$ of the entire game $GAME[\pi]$, this profile will be Markovian. In what follows, we prove that for any $G \in \mathcal{G}$, (a) if $\sigma_G^*$ is played, then there is no transition if $\phi (G) = G$ and there is a transition to $\phi (G)$ otherwise and (b) actions in profile $\sigma^*$ are best responses if continuation payoffs are taken from profile $\sigma^*$ rather than assumed to be given by (2.20). These two results will complete the proof of part 1.

We start with part (a); take any government $G$ and consider the SPE of the truncated game $\sigma_G^*$. First, consider the subgame where some alternative $H$ has won the primaries and challenges the incumbent government $G$. Clearly, proposal $H$ will be accepted if and only if $\phi (H) \succ G$. This implies, in particular, from the construction of mapping $\phi$, that if $\phi (G) = G$, then no alternative $H$ may be accepted. Second, consider the subgame where nominations have been made and the players are voting according to protocol $\pi_{G^t}^A$. We prove that if $\phi (G) \in \mathcal{A}$, then $\phi (G)$ wins the primaries regardless of $\pi$ (and subsequently wins against $G$, as $\phi (\phi (G)) = \phi (G) \succ G$. This is proved by backward induction: suppose
that \( \phi(G) \) has number \( q \) in the protocol, let us show that if it makes its way to \( j \)th round, where \( q \leq j \leq |A| \), and then it will win this round. The base is evident: if \( \phi(G) \) wins in the last round, players will get \( v(G, \phi(G)) = \chi_G(\phi(G)) = \frac{1}{1-\beta}w(\phi(G)) \) (we drop the subscript for player to refer to \( w \) and \( v \) as vectors of payoffs), while if it loses, they either get \( v(G, H) \) for \( H \neq \phi(G) \). Clearly, voting for \( \phi(G) \) is better for a majority of population, and thus \( \phi(G) \) wins the primaries and defeats \( G \). The step is proven similarly, hence, in the subgame which starts from \( q \)th round, \( \phi(G) \) will defeat the incumbent government.

Since this holds irrespective of what happens in previous rounds, this concludes the second step. Third, consider the stage where nominations are made, and suppose, to obtain a contradiction, that \( \phi(G) \) is not proposed. Then, in the equilibrium, players get a payoff vector \( v(G, H) \), where \( H \neq \phi(G) \). But then, clearly, any member of \( \phi(G) \) has a profitable deviation, which is to nominate \( \phi(G) \) instead of or in addition to what he is nominating in profile \( \sigma^*_G \). Since in a SPE there should be no profitable deviations, this competes the proof of part (a).

Part (b) is fairly obvious. Suppose that the incumbent government is \( G \). If some alternative \( H \) defeats government \( G \), then from part (a), the payoffs that players get starting from next period are given by \( \frac{1}{1-\beta}w_i(H) \) if \( \phi(H) = H \) and \( w_i(H) + \frac{\beta}{1-\beta}w_i(\phi(H)) \) otherwise; in either case, the payoff is exactly equal to \( v_i(G, H) \). If no alternative defeats government \( G \), then the \( \nu^{t+1} = s \) (the situation becomes stable), and after that, government \( G \) stays until the situation becomes unstable, and government \( \phi(G) \) is in power in all periods ever since; this again gives the payoff \( \frac{1+r\beta}{1-\beta(1-\gamma)}w_i(G) + \frac{r\beta^2}{(1-\beta)(1-\beta(1-\gamma))}w_i(\phi(G)) \).

This implies that the continuation payoffs are indeed given by \( v_i(G, H) \), which means that if in the entire game profile \( \sigma^* \) is played, no player has a profitable deviation. This proves part 1.

**Part 2.** Suppose \( \sigma^* \) is an acyclic MPE. Take any government \( G = G^t \) at some period \( t \) in some node on or off the equilibrium path. Define binary relation \( \rightarrow \) on set \( G \) as follows: \( G \rightarrow H \) if and only if either \( G = H \) and \( G \) has a positive probability of staying in power when \( G^t = G \) and \( \nu^t = u \), or \( G \neq H \) and \( G^{t+1} = H \) with positive probability if \( G^t = G \) and \( \nu^t = u \). Define another binary relation \( \leftrightarrow \) on \( G \) as follows:
\( G \mapsto H \) if any only if there exists a sequence (perhaps empty) of different governments \( H_1, \ldots, H_q \) such that \( G \to H_1 \to H_2 \to \cdots \to H_q = H \) and \( H \to H \). In other words, \( G \mapsto H \) if there is an on-equilibrium path that involves a sequence of transitions from \( G \) to \( H \) and stabilization of political situation at \( H \). Now, since \( \sigma^* \) is an acyclic equilibrium, there is no sequence that contains at least two different governments \( H_1, \ldots, H_q \) such that \( H_1 \to H_2 \to \cdots \to H_q \to H_1 \). Suppose that for at least one \( G \in \mathcal{G} \), the set \( \{ H \in \mathcal{G} : G \mapsto H \} \) contains at least two elements. From acyclicity it is easy to derive the existence of government \( G \) with the following properties: \( \{ H \in \mathcal{G} : G \mapsto H \} \) contains at least two elements, but for any element \( H \) of this set, \( \{ H' \in \mathcal{G} : H \mapsto H' \} \) is a singleton.

Consider the restriction of profile \( \sigma^* \) on the part of the game where government \( G \) is in power, and call it \( \sigma_G^* \). The way we picked \( G \) implies that some government may defeat \( G \) with a positive probability, and for any such government \( H \) the subsequent evolution prescribed by profile \( \sigma^* \) does not exhibit any uncertainty, and the political situation will stabilize at the unique government \( H' \neq G \) (but perhaps \( H' = H \)) such that \( H \mapsto H' \) in no more than \( |\mathcal{G}| - 2 \) steps. Given our assumption (2.19) and the assumption that \( r \) is small, this implies that no player is indifferent between two terminal nodes of this period which ultimately lead to two different governments \( H'_1 \) and \( H'_2 \), or between one where \( G \) stays and one where it is overthrown. But players act sequentially, one at a time, which means that the last player to act on the equilibrium path when it is still possible to get different outcomes must mix, and therefore be indifferent. This contradiction proves that for any \( G \), government \( H \) such that \( G \mapsto H \) is well-defined. Denote this government by \( \psi(G) \).

To finish the proof, we must show that \( \psi(G) = \phi(G) \) for all \( G \). Suppose not; then, since \( \psi(G) \succ G \) (otherwise \( G \) would not be defeated as players would prefer to stay in \( G \)), we must have that \( \Gamma_{\phi(G)} > \Gamma_{\psi(G)} \). This implies that if some alternative \( H \) such that \( H \mapsto \phi(G) \) is nominated, it must win the primaries; this is easily shown by backward induction. If no such alternative is nominated, then, since there is a player who prefers \( \phi(G) \) to \( \psi(G) \) (any member of \( \phi(G) \) does), such player would be better off deviating and nominating \( \psi(G) \). A deviation is not possible in equilibrium, so \( \psi(G) = \phi(G) \) for all \( G \).
By construction of mapping $\psi$, this implies that there are no transitions if $G = \phi(G)$ and one or more transitions ultimately leading to government $\phi(G)$ otherwise. This completes the proof. ■

**Proof of Theorem 10. Part 1.** In part 1 of Theorem 9, we proved that for any $\pi \in \Pi$ there exists a MPE in pure strategies, and from part 2 of Theorem 9 it follows that these MPE constructed for different $\pi \in \Pi$ have the same equilibrium path of governments. The existence of an order-independent equilibrium follows.

**Part 2.** Suppose, to obtain a contradiction, that order-independent MPE in pure strategies $\sigma^*$ is cyclic. Define mapping $\chi : \mathcal{G} \rightarrow \mathcal{G}$ as follows: $\chi(G) = H$ if for any node on equilibrium path which starts with government $G^i = G$ and $\nu^i = u$, the next government $G^{i+1} = H$. Since the equilibrium is in pure strategies, this mapping is well-defined and unique. The assumption that equilibrium $\sigma^*$ is acyclic implies that there is a sequence of pairwise different governments $H_1, H_2, \ldots, H_q$ (where $q \geq 2$) such that $\chi(H_j) = H_{j+1}$ for $1 \leq j < q$ and $\chi(H_q) = H_1$. Without loss of generality, assume that $H_2$ has the least competence of all governments $H_1, H_2, \ldots, H_q$. If $q = 2$, then the cycle has two elements, of which $H_2$ is the worse government. However, this implies that $H_2$ cannot defeat $H_1$ even if it wins the primaries, since all players, except, perhaps, those in $H_2 \setminus H_1$, prefer $H_1$ to an eternal cycle of $H_1$ and $H_2$. This immediate contradiction implies that we only need to consider the case $q \geq 3$.

If $q \geq 3$, then, by the choice of $H_2$, $\Gamma_{H_1} > \Gamma_{H_2}$ and $\Gamma_{H_3} > \Gamma_{H_2}$. Without loss of generality, we may assume that the protocol is such that if the incumbent government is $H_1$, $H_3$ is put at the end (if $H_3$ is nominated); this is possible since $\sigma^*$ is an order-independent equilibrium. By definition, we must have that proposal $H_2$ is nominated and accepted in this equilibrium along the equilibrium path.

Let us first prove that alternative $H_3$ will defeat the incumbent government $H_1$ if it wins the primaries. Consider a player $i$ who would have weakly preferred $H_2$, the next equilibrium government, to win over $H_1$ if $H_2$ won the primaries; since $H_2$ defeats $H_1$ on the equilibrium path, such players must form a winning coalition in $H_1$. If $i \notin H_2$, then $H_2$ brings $i$ the lowest utility of all governments in the cycle; hence, $i$ would be willing to
skip $H_2$; hence, such $i$ would be strictly better off if $H_3$ defeated $H_1$. Now suppose $i \in H_2$. If, in addition, $i \in H_1$, then he prefers $H_1$ to $H_2$. Assume, to obtain a contradiction, that $i$ weakly prefers that $H_3$ does not defeat $H_1$; it is then easy to see that since he prefers $H_1$ to $H_2$, he would strictly prefer $H_2$ not to defeat $H_1$ if $H_2$ won the primaries. The last case to consider is $i \in H_2$ and $i \notin H_1$. If $\beta$ is sufficiently close to 1, then, as implied by Assumption 3', player $i$ will either prefer that both $H_2$ and $H_3$ defeat $H_1$ or that none of them does. Consequently, all players who would support $H_2$ also support $H_3$, which proves that $H_3$ would be accepted if nominated.

Let us prove that in equilibrium $H_3$ is not nominated. Suppose the opposite, i.e., that $H_3$ is nominated. Then $H_2$ cannot win the primaries: indeed, in the last voting, $H_2$ must face $H_3$, and since, as we showed, only members of $H_2$ may prefer that $H_2$ rather than $H_3$ is the next government, $H_3$ must defeat $H_2$ in this voting. This means that in equilibrium $H_3$ is not nominated.

Consider, however, what would happen if all alternatives were nominated. Suppose that some government $G$ then wins the primaries. It must necessarily be the case that $G$ defeats $H_1$: indeed, if instead $H_1$ would stay in power, then $G \neq H_3$ (we know that $H_3$ would defeat $H_1$), and this implies that in the last voting of the primaries, $H_3$ would defeat $G$. Let us denote the continuation utility that player $i$ gets if some government $H$ comes to power as $v_i(H)$. If there is at least one player with $v_i(G) > v_i(H_2)$, then this player has a profitable deviation during nominations: he can nominate all alternatives and ensure that $G$ wins the primaries and defeats $H_1$. Otherwise, if $v_i(G) \leq v_i(H_2)$ for all players, we must have that $v_i(G) < v_i(H_3)$ for a winning coalition of players, which again means that $G$ cannot win the primaries. This contradiction proves that for the protocol we chose, $H_2$ cannot be the next government, and this implies that there are no cyclic order-independent equilibria in pure strategies.

**Part 3.** The proof is similar to the proof of part 2. We define mapping $\chi$ in the same way and choose government $H$ such that $\chi(\chi(H)) \neq \chi(H)$, but $\chi(\chi(\chi(H))) = \chi(\chi(H))$. We then take a protocol which puts government $\chi(\chi(H))$ at the end whenever it is nominated and come to a similar contradiction.
Part 4. This follows straightforwardly from part 3, since the only transition may happen at period $t = 0$.

Proof of Theorem 11. Part 1. If $\delta$ is sufficiently small, then the possibility of shocks does not change the ordering of continuation utilities at the end of any period any for any player. Hence, the equilibrium constructed in the proof of part 1 of Theorem 9 proves this statement as well.

Part 2. If $\delta$ is sufficiently small, the proof of Theorem 10 (parts 2 and 3) may be applied here with minimal changes, which are omitted.

2.8.7 Examples

Suppose that the society consists of five individuals $1, 2, 3, 4, 5$ ($n = 5$). Suppose each government consists of two members, so $k = 2$. There is “almost perfect” democracy ($l = 1$), and suppose $m = 3$. Assume

$$\Gamma_{\{i, j\}} = 30 - \min \{i, j\} - 5 \max \{i, j\}.$$ 

Moreover, assume that all individuals care a lot about being in the government, and about the competence if they are not in the government; however, if a individual is a member of two different governments, he is almost indifferent. In this environment, there are two fixed points of mapping $\phi$: $\{1, 2\}$ and $\{3, 4\}$.

Let us show that there is no MPE in pure strategies if $v^t = u$ for all $t$ (so that the incumbent government is contested in each period). Suppose that there is such equilibrium for some protocol $\pi$. One can easily see that no alternative may win if the incumbent government is $\{1, 2\}$: indeed, if in equilibrium there is a transition to some $G \neq \{1, 2\}$, then in the last voting, when $\{1, 2\}$ is challenged by $G$, both 1 and 2 would be better off rejecting the alternative and postpone transition to the government (or a chain of governments) that they like less. It is also not hard to show that any of the governments that include 1 or 2 (i.e., $\{1, 3\}, \{1, 4\}, \{1, 5\}, \{2, 3\}, \{2, 4\}, \text{and } \{2, 5\}$) lose the contest for power to $\{1, 2\}$ in equilibrium. Indeed, if $\{1, 2\}$ is included in the
primaries, it must be the winner (intuitively, this happens because \(\{1, 2\}\) is the Condorcet winner in simple majority votings). Given that, it must always be included in the primaries, for otherwise individual 1 would have a profitable deviation and nominate \(\{1, 2\}\). We can now conclude that government \(\{3, 4\}\) is stable: any government which includes 1 or 2 will immediately lead to \(\{1, 2\}\) which is undesirable for both 3 and 4, while \(\{3, 5\}\) and \(\{4, 5\}\) are worse than \(\{3, 4\}\) for 3 and 4 as well; therefore, if there is some transition in equilibrium, then 3 and 4 are better off staying at \(\{3, 4\}\) for an extra period, which makes a profitable deviation.

We now consider the governments \(\{3, 5\}\) and \(\{4, 5\}\). First, we rule out the possibility that from \(\{3, 5\}\) the individuals move to \(\{4, 5\}\) and vice versa. Indeed, if this was the case, then in the last voting when the government is \(\{3, 5\}\) and the alternative is \(\{4, 5\}\), individuals 1, 2, 3, 5 would be better off blocking this transition (i.e., postponing it for one period). Hence, either one of governments \(\{3, 5\}\) and \(\{4, 5\}\) is stable or one of them leads to \(\{3, 4\}\) in one step or \(\{1, 2\}\) in two steps. We consider these three possibilities for the government \(\{3, 5\}\) and come to a contradiction; the case of \(\{4, 5\}\) may be considered similarly and also leads to a contradiction.

It is trivial to see that a transition to \(\{1, 2\}\) (in one or two steps) cannot be in an equilibrium. If this was the case, then in the last voting, individuals 3 and 5 would block this transition, since they are better off staying in \(\{3, 5\}\) for one more period (even if the intermediate step to \(\{1, 2\}\) is a government that includes either 3 or 5). This is a profitable deviation which cannot happen in an equilibrium. It is also trivial to check that \(\{3, 5\}\) cannot be stable. Indeed, if this was the case, then if alternative \(\{3, 4\}\) won the primaries, it would be accepted, as individuals 1, 2, 3, 4 would support it. At the same time, any alternative that would lead to \(\{1, 2\}\) would not be accepted, and neither will be alternative \(\{4, 5\}\), unless it leads to \(\{3, 4\}\). Because of that, alternative \(\{3, 4\}\) would make its way through the primaries if nominated, for it is better than \(\{3, 5\}\) for a simple majority of individuals. But then \(\{3, 4\}\) must be nominated, for, say, individual 4 is better off if it were, since he prefers \(\{3, 4\}\) to \(\{3, 5\}\). Consequently, if \(\{3, 5\}\) were stable, we would get a contradiction, since we proved that in this case,
\{3,4\} must be nominated, win the primaries and take over the incumbent government \{3,5\}.

The remaining case to consider is where from \{3,5\} the individuals transit to \{3,4\}. Note that in this case, alternative \{1,2\} would be accepted if it won the primaries: indeed, individuals 1 and 2 prefer \{1,2\} over \{3,4\} for obvious reasons, but individual 5 is also better off if \{1,2\} is accepted, for he is worse off under \{3,4\} than under \{1,2\}, even if the former grants him an extra period of staying in power (recall that the discount factor \(\beta\) is close to 1). Similarly, any alternative which would lead to \{1,2\} in the next period must also be accepted in the last voting. This implies, however, that such alternative (\{1,2\} or some other one which leads to \{1,2\}) must necessarily win the primaries if nominated (by the previous discussion, \{4,5\} may not be a stable government, and hence the only choice the individuals make is whether to move ultimately to \{3,4\} or to \{1,2\}, of which they prefer the latter). This, in turn, means that \{1,2\} must be nominated, for otherwise, say, individual 1 would be better off doing that. Hence, we have come to a contradiction in all possible cases, which proves that for no protocol \(\pi\) there exists an MPE in pure strategies. Note that we have proven that both cyclic and acyclic equilibria do not exist.

We prove this by constructing an example. Suppose \(n = 9, k = 4, l_1 = 3, m = 5\). Let the individuals be denoted 1, 2, 3, 4, 5, 6, 7, 8, 9, with decreasing ability. Namely, suppose that abilities of individuals 1, \ldots, 8 are given by \(\gamma_i = 2^{8-i}\), and \(\gamma_9 = -10^6\). Then the
14 stable governments, in the order of decreasing competence, are given in the table.

<table>
<thead>
<tr>
<th>1234</th>
<th>2358</th>
</tr>
</thead>
<tbody>
<tr>
<td>1256</td>
<td>2367</td>
</tr>
<tr>
<td>1278</td>
<td>2457</td>
</tr>
<tr>
<td>1357</td>
<td>2468</td>
</tr>
<tr>
<td>1368</td>
<td>3456</td>
</tr>
<tr>
<td>1458</td>
<td>3478</td>
</tr>
<tr>
<td>1467</td>
<td>5678</td>
</tr>
</tbody>
</table>

(Note that this would be the list of stable governments for any decreasing sequence \( \{\gamma_i\}_{i=1}^{9} \), except for that, say, \( \Gamma_{\{1368\}} \) may become than \( \Gamma_{\{1458\}} \).) Now consider the same parameters, but take \( l_2 = 2 \). Then there are three stable governments 1234, 1567, and 2589. For a random initial government, the probability that individual 9 will be a part of the stable government that evolves is \( \frac{9}{126} = \frac{1}{16} \): of \( \binom{9}{4} = 126 \) feasible governments there are 9 governments that lead to 2589, which are 2589, 2689, 2789, 3589, 3689, 3789, 4589, 4689, and 4789. Clearly, the expected competence of government for \( l_2 = 2 \) is negative, whereas for \( l_1 = 3 \) it is positive, as no stable government includes the least competent individual 9.
3. DICTATORS AND THEIR VIZIERS: ENDOGENIZING THE LOYALTY–COMPETENCE TRADE-OFF

(joint work with Konstantin Sonin)

3.1 Introduction

Authoritarianism is one of the oldest forms of government (see Tullock, 1987, Olson, 1993), yet it attracted little attention of political economists until very recently. While the number of democratic countries increased significantly during the last decades of the 20th century, dictatorships still account for almost one half of the world’s current population. There is also a non-negligible number of relatively new dictatorships, especially in countries of the former Soviet Union such as Turkmenistan, Uzbekistan, and Belarus. In this paper, we focus on the internal structure of the most personalized dictatorships, especially those that have exhibited strikingly poor governance (for the most recent examples, see, e.g., Stepan and Linz, 1978, Bueno de Mesquita et al., 2003, Haber, 2005).

Why do dictators, who presumably fear that poor economic performance or the inability to carry out political repressions might cost them their position in power, appoint incompetent cronies or relatives to crucial economic and military positions in government? While incompetent ministers are arguably not unusual in democratic countries, historians and political scientists would agree that personalized dictatorships are especially marred by incompetence. (Section 2 discusses broad historic and political evidence.) In such regimes, the main problem for an autocrat might be not the incompetence, but the pos-

1 For an overview of the state of the art in the formal work on dictatorship see Acemoglu and Robinson (2005), Bueno de Mesquita et al. (2003), and Wintrobe (2002). Acemoglu (2003), Acemoglu, Robinson, and Verdier (2004), Acemoglu and Robinson (2005), and Galliego and Pitchik (2004) suggest dynamic frameworks for studying modern dictatorships. Recent empirical studies of dictatorships include Epstein et al. (2004), Przeworski et al. (2000), Przeworski and Limongi (1993), Gandi (2005), and Wantchekon (2000).
sible disloyalty, of a vizier. Haber (2005) observes that “virtually all constituents and colleagues in dictatorships – at least those who value their necks – profess loyalty to the dictator, even as they conspire against him”. Brooker (2000) finds that military coups are attempted against dictators at least twice as often as against democratic rulers. Not surprisingly, Wintrobe (2000) concludes that paranoia is “the most likely personality characteristics possessed by dictators”. Most recently, military experts pointed out that even under increasing outside threat, Saddam Hussein placed incompetent loyalists to the crucial positions and blocked communication between field-level commanders as he feared a plot (Gordon and Tailor, 2006).2

We present a simple contract-theory model to analyze the loyalty vs. competence trade-off in a dictatorial political environment. Facing a potential challenger of an unknown strength, a dictator hires a lieutenant (a vizier) who is more competent in determining the extent of the threat than the dictator himself; to do this, the dictator chooses from a pool of lieutenants of varying competence. Appointing a more competent vizier, the dictator benefits from both successful defense, and the detrimental effect on enemy’s incentives to launch a challenge. However, the very competence of the vizier makes him more prone to treason: better knowledge of odds of winning allows him to participate in a plot even if he is offered a lower reward than a less competent person would require. A cunning vizier acts as a discriminating monopolist for possible enemies of the dictator, while an uninformed first minister acts as a normal monopolist. Assuming that the willingness of the first minister to accept a bribe is increasing both in the size of the bribe and the probability of success of the plot or a foreign invasion, the ruler trades off the loyalty of his first minister (lower willingness to accept a bribe) and his competence (higher willingness to accept bribes for treason). The trade-off a personalistic dictator faces is especially pronounced as any use of incentive schemes by a dictator is constrained by the consideration that all rewards and punishments are necessarily conditional upon the dictator’s own survival.3


3 North and Weingast (1989) (see also Shepsle, 1991 and Acemoglu and Robinson, 2001, 2005) made the commitment issue central in political science, demonstrating in particular that a constitutionally restrained
The trade-off between loyalty and competence is not new in the corporate governance literature, where the principal-agent conflict was first studied. Though the commitment problem is not so extreme in the corporate world – there are contracts and courts that enforce them – a top manager concerned with the possibility of “betrayal” by a hired agent might be willing to hire a mediocre agent rather than one of either high or low ability. Prendergast and Topel (1996) show that a principal who values the power to affect his subordinate’s welfare does not necessarily appoint the most competent agents (see also FreiBel and Raith, 2004, on how these considerations affect optimal information flows inside the firm). Though the principal’s primary concern is somewhat akin to that of our dictator, the analysis does significantly rely on the fact that there is a third party (the firm’s management). In Glazer (2002), when agents with high ability to run a firm also possess superior skills in internal rent-seeking, the owner might be willing to hire a low-ability agent. In contrast, our model highlights that it might not be that the same person has two complementary qualities, but the very quality for which the agent is hired (competence) might be the source of potential disutility to the principal.

Prendergast (1993) demonstrates that if a subordinate’s activity is rewarded based on subjective performance evaluation, then high-powered incentives, while inducing the subordinate to work harder, make her conform to the opinion of the principal (see also recent works by Morris, 2003, and Wagner, 2004). Though some of the features are similar to those of our model (e.g., that relevant information possessed by the agent is lost for the principal in equilibrium), the approaches are very different. First, we do not make use of subjective performance evaluation: if the plot fails, the dictator gets all the relevant information. Second, the vizier has no need to conform to the dictator’s opinion, which he knows for sure. Burkart, Panunzi, and Shleifer (2003) investigate the trade-off between competence of a hired manager and the loyalty of a family member generally lacking that competence. In our model, loyalty and incompetence are two sides of the same token.

The rest of the paper is organized as follows. Section 3.2 discusses evidence from
historical and political sources. Section 3.3 introduces the setup and starts the formal analysis. Section 3.4 considers self-selection of would-be viziers, while Section 3.5 is focused on the dynamic perspective and its implications. Section 3.6 concludes.

### 3.2 Agency Problems in Dictatorships

Any historic and political evidence related to the loyalty-competence trade-off is bound to be anecdotal. While for loyalty there are relatively secure *ex post* estimates, this is much less so for competence.\(^4\) This is not to say that the issue is intractable. In this section, we discuss some systematic studies of government spanning across centuries and regimes such as Gibbon (1781), Finer (1997), Chehabi and Linz (1998), Dominguez (2002) as well as individual-regime studies such as Conquest (1967), Gregory (2004), Montefiore (2003), and Petrov and Skorkin (1999) on Stalin’s Russia, Hartlyn (1998) on Trujillo’s Haiti, Kiernan (2004) on Pol Pot’s Cambodia, Lewis (1978) on Salazar’s Portugal, Speer (1970) on Hitler’s Germany, Young and Turner (1985) on Mobutu’s Zaire, etc.

The word “vizier” comes from the Ottoman Empire, where the Grand Vizier was, essentially, a prime-minister appointed by the sole sovereign (sultan). Viziers typically played a crucial role in palace revolutions throughout the 12\(^{th}\)–19\(^{th}\) centuries (Finer, 1997). We use this word as our model can be most directly applied to “the palace” regimes in Finer’s typology of governments (Finer, 1997, I:38) and neo-sultanistic regimes (Chehabi and Linz, 1998), their modern counterpart (see below). Historical examples include ancient kingdoms and empires such as Byzantine, Roman, Persian, and Chinese, and absolutist European monarchies of 17\(^{th}\) and 18\(^{th}\) centuries. Still, palace courts played enormous role even in modern dictatorships such as that of Hitler (Speer, 1970) and Stalin (Montefiore, 2003).

As early as in 1965 BC, i.e. almost four thousand years ago, King Sesostris I of Egypt

\(^4\) Some econometric evidence comes from empirical work on military campaigns, where definition of success or failure (and thus competence and incompetence of those responsible) is relatively clear. Peceny, Beer, and Sanchez-Terry (2002) and Reiter and Stam (2003) found that personalist dictatorships are especially prone to open military conflicts with democracies (the latter paper also demonstrates that unconstrained dictators are more likely to challenge a democracy than vice-versa), despite the fact that personalist dictatorships fight wars poorly (at least since 1945).
warned future kings in his instruction: “Be on your guard against all subordinates, because you cannot be sure who is plotting against you” (Rindova and Starbuck, 1997). Han Fei Tzu, a Chinese philosopher of the 3rd century BC, advised rulers to distrust subordinates and inspire fear in them. Finer (1997: I, 545) notes that until Trajan most Roman emperors lived in real or imaginary terror of enemies. Many ancient and medieval rulers hired foreign bodyguards, who were less able to take power themselves than the local military. Finer (1997, I:18): “A ruler [in contrast with the political regime, of which he is a focal part] might fancy himself more secure when surrounded by a band of foreign mercenaries.”

In The Twelve Caesars, Suetonius (110CE, 1979) mentions foreign bodyguards protecting Caligula, the foremost example of an unconstrained ruler fearing betrayal.

The institute of eunuchs, infertile human males, is a specific example of how ruler’s fears of betrayal and undesired succession have been institutionalized. The very idea of having eunuch as a close subordinate is related to the fact that under no circumstances a eunuch can be a legitimate head of the state, exactly for the reason he is infertile. Throughout history, they have been a significant part of courts in Persia, Egypt, Ancient Greece and the Roman Empire. In China, by the end of Ming Dynasty, the Imperial Palace employed no less than seventy thousand eunuchs. Finer (1997, I:575) observes that the careers of eunuch Grand Chamberlains Eusebius, Eutropius, Chrysaphius, and Eutherius demonstrate that each was the principal minister of their respective emperor. Hopkins (1978, 179) points out that not only these eunuchs, but many much humbler ones have had “tremendous and sustained influence” on court decision’s and emperors’ succession. Discussing the rise and institutionalization of eunuchs, Finer (1997, I:575) notes a close parallel between the Chinese Han empire and the decline of the Roman Empire; as an emperor became more sacred (which required more and more layers of bureaucrats between him and his subjects) and subsequently less accessible, he became more vulnerable. Gibbon (1781, II:Ch. 19) argues at length that Constantius II “feared his generals and mistrusted his ministers” and that was true to all of the Roman emperors in 4th century.

Two millenia from Romans, Portugal’s Salazar and Germany’s Hitler, Soviet Russia’s
Stalin and Cambodia’s Pol Pot, Zaire’s Mobutu and Iraq’s Hussein, Romania’s Ceausescu and Turkemenistan’s Niyazov might have exhibited considerably different life-paths and individual habits if taken individually. However, apart from being exceptional political manipulators, they were strikingly similar in organization of their courts and broad regimes.

Considering the rule of Salazar, the Portuguese dictator in the period 1932–1968, Lewis (1978) identifies the problem of self-selection of subordinates: “[Salazar] was ... intolerant of those who did not share [his own views] to the last degree. That discouraged many talented young men from entering the government service.” With Salazar’s power increasingly secure, ‘the patterns of recruitment show the regime evolving from its military and semi-fascist beginnings in the direction of a modern technocratic state’ (Lewis, 1978).

Petrov and Skorkin (1999) provide a striking illustration of the loyalty-competence trade-off through an examination of the educational level of top officers of the NKVD, Stalin’s praetorians that had been a major political authority during the era of “great purges” (1934–1941). Literally, NKVD stands for the “Ministry of Internal Affairs”, but in practice it, as the “armed hand” of the Communist Party, exercised control of almost all aspects of life in the USSR (Conquest, 1967, Gregory, 2004). In 1934, 41 percent of top officers (39 officers) had less than 7 years of total schooling, and another 42 percent had less than 10 years. In 1937, the peak year of purges, these numbers were 37 and 43 percent, respectively.

Montefiore (2003) focuses his work on Stalin’s inner circle, and provides overwhelming archival evidence on fear of betrayal, which was typical for Stalin and his closest subordinates (both Montefiore, 2003, and Gregory, 2004, use top-secret archival documents made available for researchers only recently). One result was a total devastation of Russian military in the period immediately preceding World War II (Volkogonov, 1998: 370-71, 405-409; Spahr, 1997:231). At the same time, a sudden change in circumstances might bring a sudden change in the dictator’s strategy. In October 1941, when the German forces were advancing to Moscow with a speed exceeding all forecasts, Stalin brought back hundreds of top military commanders, purged 3-5 years before, from concentration camps.
and prisons (Spahr, 1997, Glazer, 2002). (Table 1 in Section 4 illustrates choices by both safe and vulnerable dictators in different circumstances.)

In his memoirs, Albert Speer, once a second-highest ranked official in the Third Reich and a confidant of Hitler, used the words “negative selection” in his description of Hitler’s court, discussing at length the ignorance and incompetence of Hitler’s closest subordinates (Speer, 1970). Further into the World War II, the more military commanders, even those who were successful in the battlefield, e.g. Field-Marshal Rommel, were replaced by Nazi loyalists; with stakes rising, the competence became less important than personal loyalty to the Fuhrer. During the Nuremberg trials, one of the foremost Nazi loyalist, the foreign minister Ribbentrop produced examples of both his unquestionable loyalty (he claimed that he is still eager to carry any Hitler’s orders) and outright stupidity.

In more recent times, one dictator that apparently valued loyalty higher than even basic education was Pol Pot, the leader of the Khmer Rouge movement in Cambodia, who upon taking power in 1975 attempted to execute all public servants, teachers, and anyone with a higher education or being a public servant or a teacher (Kiernan, 2004). Pol Pot’s regime lasted for three years and cost Cambodia over 1.7 millions of lives, more than a fifth of the entire population. (Cook, 2004, Heder and Tittemore, 2004).

Chehabi and Linz (1998a) identify a certain type of dictatorships, which they dub, following Max Weber, the “sultanistic regimes”. These include, among others, the regimes of Idi Amin in Uganda, Nguema in Equatorial Guinea, Duvalier in Haiti, Batista in Cuba, Trujillo in the Dominican Republic, Rezva Shah in Iran, Mobutu in Zaire, Marcos in the Philippines. Each of the regimes has been characterized by selection of dictator’s subordinates basing on personal loyalty; economic performance, at least at their respective age of maturity, was dismal. Such different personalities as Mobutu Sese-Seko, Ferdinand Marcos, and Jean-Claude Duvalier brought up corps of technocrats in their governments

---

5 Apparently, Speer does consider himself a counterexample that only proves the rule.

6 Prosecution: “What further pressure could you put on the head of a country beyond threatening him that your Army would march in, in overwhelming strength, and your air force would bomb his capital?” Ribbentrop: “War, for instance.” Nuremberg Trial materials, April 1, 1946.

7 If in 1998, when the book edited by Chehabi and Linz came to print, the Turkmenian regime of Saparmurat Niyazov was a “sultanistic regime candidate”; by 2006, it is certainly so.
so that they could, in words of Marcos himself, “sit back and let technocrats run things”, but “the political leadership then allowed the unconstrained introduction of exceptions that made complete mockery of the spirit and letter of the plans” (Hutchcroft, 1991). At the same time, Mobutu encouraged military officers to report rumors to him personally; in 1978, nearly 10 percent of officer corps were dismissed overnight since the dictator doubted their loyalty (Young and Turner, 1985). Katouzian (1998) observes that the Shah’s fear of a military plot made the Iranian army incapable to carry most simple tasks.

Thompson (1998) bases his study of the Marcos’s regime on an account by Primitivo Mijares, who was the Philippines chief propagandist and a close advisor of the dictator before publicly defecting and exposing the sultanistic nature of the regime. The “traitor” was later either lured or tricked into a return journey to the Philippines, where he was executed alongside with his son. For the Marcos regime, appointment of loyal and incompetent cronies to key government positions was a rule, rather than an exception (e.g., Marcos’s chauffeur was appointed the chief of all security forces, Thompson, 1998).

The ultimate manifestation of emphasis on loyalty at the expense of competence is appointment of close relatives to the government position which require either political or professional competence – that is, making the regime less competitive and less stable. Chehabi and Linz (1998b) consider this a characteristic trait of sultanistic regimes. In socialist Romania, in a reminiscent of medieval regimes, key government position were held by four brothers, wife (as a second-in-line in the ruling party leadership and the President of the Romanian Academy of Science), and son (as a designated successor famous for his gambling and drinking habits) of President Ceausescu; in a secular dictatorship of Saddam Hussein, most of government posts were occupied by his relatives. (Stalin’s son Vassily was made a commander of an elite Air Force division, Montefiore, 2003.)

Hartlyn (1998) observes that the role of Ramphis, Trujillo’s son who was made a full colonel at the age of four and brigadier general at the age of nine, “highlighted the fact that the morale, autonomy and abilities of the armed forces were affected by nepotism”. Analyzing both general patterns and specific examples of replacement of professionals in armed forces by personal loyalists in Batista’s Cuba, Dominguez (1998) concludes: “It was
these unprofessional, politicized officers who faced Fidel Castro’s insurgency and lost.”

In a search for a “perfect dictatorship”, Domínguez (2002) studies the most successful dictatorships of the 20th century: Mexico, Brazil, Chile and South Korea. He argues that all of them employed talented people, at least in the early years of the regime. However, in maturing regimes, personalist and institutionalized dictatorships diverged with respect to political competence. In a personalist dictatorship such as South Korea under Park Chung Hee’s rule, competent people were driven out of the government with the establishment of the loyalty-based Yushin system in 1971. Brazil and Mexico, with their institutionalized succession within an authoritarian regime, have had significantly fewer of these problems.

### 3.3 Model

There are three strategic players in the model: the incumbent Dictator, Vizier, who is Dictator’s agent, and Enemy, who wants to oust Dictator. The threat to Dictator’s power depends on Enemy’s type, which might might be either strong, \( t = \text{Strong} \), or weak, \( t = \text{Weak} \). A strong enemy type will be the victor in the fight with Dictator, unless Vizier remains loyal; a weak enemy type will lose the fight. Obviously, \( t \) does not need to be interpreted exclusively as Enemy’s attribute; it might easily reflect the vulnerability of Dictator in a given period. To allow for both interpretations, we assume that none of the players know Enemy’s type for sure until the end of the game, but get noisy signals about it.

For an Enemy, the *ex-ante* probability of being of type \( t = \text{Strong} \) is \( p \). Before deciding whether or not to launch a coup, Enemy acquires information about his chances to oust the incumbent Dictator. Formally, Enemy receives a signal \( z \in [0, 1] \), which is random with the following distribution. If \( s = \text{Strong} \), then p.d.f of \( z \) is given by \( f_s(z) = 2z \); if \( s = \text{Weak} \), p.d.f is given by \( f_w(z) = 2(1-z) \). Hence, if signal \( z \) is received, Enemy’s posterior probability that he is strong equals

\[
\mu(z, p) = \frac{2pz}{2zp + 2(1-z)(1-p)}.
\]
For the sake of simplicity, we will assume that the \textit{ex-ante} probability $p$ is equal to $\frac{1}{2}$; then the posterior probability of being strong equals $\mu \left( z, \frac{1}{2} \right) = z$.

To protect himself, Dictator needs to make a decision on whether or not he needs to take costly precautionary measures. Without these measures, the coup against Dictator succeeds if Enemy is strong, and fails if Enemy is weak. If Dictator survives the potential coup, he receives utility of $Y$; if not, his utility is normalized to 0. Precautionary measures cost Dictator $C$, $0 < C < Y$, and ensure Dictator’s safety. Thus, Dictator trades off the benefits of taking precautionary measures and their costs.

From Dictator’s point of view, the probability that Enemy is strong is $P(t = \text{Strong}) = q$, which in turn is determined by Enemy’s equilibrium strategy. Dictator is unable to get any information on the whether or not Enemy is strong himself, so he hires an agent, Vizier. An agent himself might be imperfectly informed; however, Dictator has access to a pool of viziers with varying level of competence, $\theta \in [0, 1]$. Conditional on Enemy’s true type being weak, $t = \text{Weak}$, Vizier of type $\theta$ gets the signal $s = \text{Weak}$ with probability $\theta$:

$$P(s = \text{Weak} \mid t = \text{Weak}) = \theta.$$  

while the strong Enemy is recognized with certainty: $P(s = \text{Strong} \mid t = \text{Strong}) = 1$.\footnote{The assumption is made for expositional simplicity: the results go through for any $P(s = \text{Strong} \mid t = \text{Strong})$.}

Vizier reports the signal $a \in \{\text{Strong}, \text{Weak}\}$ to Dictator. However, Vizier may deliberately misinform Dictator by reporting that Enemy is weak when the signal is $s = \text{Strong}$, or that Enemy is strong when $s = \text{Weak}$. We say that Vizier \textit{betrays} Dictator whenever $a \neq s$.

Vizier has incentives to betray as he expects to be rewarded by Enemy conditional on a successful coup. Enemy’s ability to pay for Vizier’s betrayal is denoted by $\mathcal{B}$, and is stochastic from Dictator’s viewpoint. The p.d.f. is $g(x) = 1/\mathcal{B}$ if $x \in [0, \mathcal{B}]$, and $g(x) = 0$ otherwise.

Dictator provides Vizier with an incentive contract, in which payments are conditioned on all the information that Dictator possesses if he stays in power. If the enemy decided
not to mount a coup, Dictator pays $w_n$ to Vizier. If the coup is unsuccessful, Dictator learns that Enemy’s type was weak (for otherwise Dictator would not have survived); if Vizier betrayed, Dictator knows this. Thus, an incentive contract includes four payments: $w_f, w_b, w_e, w_n$, standing for “fight”, “betray”, “economize” and “no enemy”, respectively. We normalize the most severe punishment to zero, so $w_f, w_b, w_e, w_n \geq 0$ must be satisfied.

After Enemy observes whether or not Dictator takes precautionary measures, he decides whether or not to mount a coup. The payoffs conditional on the outcome of the game are as follows. If measures were taken and Enemy decides not to attack, Dictator gets $Y - C - w_f$, Vizier gets his wage $w_f$, and Enemy gets 0. If there were no measures and Enemy does not attack, dictator gets $Y - w_n$, Vizier gets $w_n$ unless he has betrayed, in which case Vizier gets 0. If Dictator takes measures and Enemy attacks, Dictator gets again $Y - C - w_f$ and Vizier gets $w_f$, but Enemy gets $-D$ (which is nonpositive). If no measures were taken and Enemy attacks, then the outcome depends on Enemy’s strength. If Enemy is strong, then Dictator is ousted, and Vizier gets the agreed-upon reward $B$ from Enemy. If Enemy is weak, Dictator remains in power, and Vizier gets either $w_e$ (for economizing on the cost of defense), or $w_b$ (for betrayal).

**Assumption 3.1** Throughout the paper, we assume $C < Y < A$.

This assumption says, first, that a defense is not too costly, so that Dictator, if he had full information, would prefer to defend against a strong Enemy. Second, Enemy has a chance to accumulate more resources ($A$ is high enough) than the dictator is willing to pay Vizier for his loyalty, so if Vizier is perfectly competent, sometimes there will be betrayals, as Dictator is never willing to pay Vizier more than $Y$.

**Timing**

The timing of the game is as follows. (The setup of the dynamic succession game is relegated to Section 3.5).

1. Enemy gets an imperfect signal $z$ about his type $t$, which parametrizes Enemy’s odds of ousting the incumbent Dictator. Knowing $z$, Enemy decides whether or not to
mount a coup. Dictator chooses a Vizier of type $\theta$ and offers him an incentive contract $w = (w_f, w_e, w_b)$ which specifies payments conditional on information available to Dictator in all three possible outcomes, and $w_n$ if there is no coup. Enemy and Dictator make their choices simultaneously.

2. If Enemy’s choice was to prepare a coup, he acquires information about $\theta$ and $w$, and also accumulates certain amount of resources $\overline{B}$, which is uniformly distributed on $[0, A]$. If Enemy decides not to prepare a coup, the game ends.

3. If Enemy prepares a coup, Vizier recognizes Enemy and gets a signal $s$ about Enemy’s type $t$. Enemy approaches Vizier with a take-it-or-leave-it offer $B \leq \overline{B}$ which Vizier would get, conditional on Dictator’s removal, if he gives Dictator certain report $a$.

4. Vizier reports the signal to Dictator, who then follows Vizier’s advice.$^{9}$

5. The residual uncertainty is realized; Dictator, Vizier, and Enemy receive their payoffs.

We will analyze perfect Bayesian equilibria of this game; we rule out equilibria in which Dictator ignores the Vizier’s advice.

3.4 Analysis

3.4.1 Optimal Contract for Vizier

We solve for a perfect Bayesian equilibrium by backward induction. Therefore, we start by considering Vizier’s actions as a function of his competence $\theta$ and an incentive contract $w = (w_f, w_e, w_b)$ offered by Dictator. In doing this, we consider Enemy characteristics given; below, we endogenize Enemy’s behavior, and feed the results into the solution of the Dictator vs. Vizier game.

$^{9}$ In equilibrium, Vizier never accepts money if $s = Weak$ and accepts with a non-zero probability if $s = Strong$; Enemy offers Vizier the amount that makes Vizier indifferent when $s = Strong$.

$^{10}$ This assumption is without loss of generality; in equilibrium, Dictator would never appoint a Vizier whose advise he does not intend to follow. We make this assumption to economize on pay-off specification for some of out-of-equilibrium paths.
If Vizier receives a weak signal, he has two options. If Vizier chooses to fight Enemy, he would get $w_f$. Alternatively, Vizier may allow the coup to unfold without taking precautionary measures. In this case, which is essentially a bet on the chance that Enemy is indeed weak, Vizier gets $w_e$. Thus, if Dictator wants to save $C$ when Enemy is weak, the incentive contract should satisfy the constraint

$$w_e \geq w_f.$$

If Vizier receives a strong signal, he again has two options. Taking precautionary measures brings him $w_f$. Betraying, on the other hand, gives Vizier $B$ if Enemy was indeed strong, which happens, by Bayes formula, with probability

$$P(t = Strong \mid s = Weak) = \frac{q}{q + (1 - \theta)(1 - q)},$$

and $w_b$ with complementary probability

$$P(t = Weak \mid s = Weak) = \frac{q + (1 - \theta)q}{q + (1 - \theta)(1 - q)}.$$

Vizier remains loyal to Dictator (that is, advises to take precautionary measures when warranted) if and only if the reward for loyalty in face of strong Enemy exceeds the expected reward of betrayal:

$$w_f \geq \frac{qB + (1 - \theta)(1 - q)w_b}{q + (1 - \theta)(1 - q)}.$$

This is equivalent to

$$B \leq w_f + (1 - \theta)\frac{1 - q}{q}(w_f - w_b). \quad (3.1)$$

We denote the expected benefit of staying loyal for Vizier, the right-hand side of inequality (3.1) by

$$L = w_f + (1 - \theta)\frac{1 - q}{q}(w_f - w_b).$$

The expected benefit of loyalty, $L$, is decreasing in Vizier’s competence: this is a con-
sequence of the fact that Dictator choose an optimal contract to keep Vizier loyal, and incentivizing a more competent Vizier requires a higher payment to him. Obviously, the dictator needs to keep the Vizier’s payoff as low as possible in the case of betrayal.

From Dictator’s viewpoint, a Vizier that gets a strong signal obeys with probability

$$G(L) = \frac{1}{A} \left( w_f + (1 - \theta) \frac{1-q}{q} (w_f - w_b) \right).$$

Dictator’s payoff as a function of Vizier’s competence and the terms of incentive contract is determined as follows. If Vizier gets signal $s = Weak$ (which happens with probability $\theta (1-q)$), then Dictator gets

$$Y - w_e$$

(provided that $w_e \geq w_f$). If Vizier gets signal $s = Strong$ and Enemy is indeed strong, $t = Strong$ (which happens with probability $q$), then the Dictator’s payoff is

$$(Y - C - w_f) G(L) + 0 \cdot (1 - G(L)) = \frac{1}{A} (Y - C - w_f) L.$$ 

If Vizier gets signal $s = Strong$, but actual type $t = Weak$, which happens with probability $(1 - \theta) (1-q)$, then Dictator gets

$$(Y - C - w_f) G(L) + (Y - w_b) (1 - G(L)) = \frac{1}{A} (w_b - w_f - C) L + Y - w_b.$$ 

**Lemma 12** The incentive contract that maximizes Dictator’s utility given that Vizier betrays if and only if inequality (3.1) holds, has $w_n = w_b = 0$ and $w_f = w_e > 0$.

Lemma 12 allows us to denote $w = w_f = w_e$. Now the payoff of Dictator equals

$$U^D = \theta (1-q) (Y - w) + q (Y - C - w) G(L) + (1 - \theta) (1-q) (Y - (w + C) G(L) ),$$

(3.2)

where

$$G(L) = \frac{L}{A} = w \left( 1 + (1 - \theta) \frac{1-q}{q} \right).$$
Solving Dictator’s optimization problem,

$$\max_{\theta, w} U_D (\theta, w; Y, C, A, q)$$

we arrive to the following result (for the sake of tractability, we consider the case where all solutions are interior only; all qualitative results go through for corner solutions as well).

**Proposition 3.1** Given the ex-ante probability that Enemy is strong, \(q\), the level of optimal competence for Vizier is

$$\theta^* (q) = 1 - \frac{1}{1 - q} \left( \sqrt{\frac{Aq}{C}} - q \right). \quad (3.3)$$

The dictator pays such Vizier optimal wage

$$w^* = \frac{A + Y}{2} \sqrt{\frac{qC}{A}} - C, \quad (3.4)$$

and the bribe that Enemy needs to offer is

$$B^* = \frac{A + Y}{2} - \sqrt{\frac{AC}{q}}. \quad (3.5)$$

The formal results of Proposition 3.1 allow to analyze comparative statics. The equilibrium competence \(\theta^*\) decreases in \(A\), which parametrizes Enemy’s expected ability to compensate Vizier for betrayal, and increases in \(C\), which is a cost of precautionary measures. This is intuitive: when prospective challengers have access to significant resources, Dictator has to put more emphasis on loyalty, rather than on competence. A more loyal Vizier has lower benefits of betrayal as he is more uncertain about Enemy’s prospects, and thus requires a higher bribe for cooperation with Enemy. With respect to the cost of taking precautionary measures, \(C\), the economic intuition is straightforward. A more competent Vizier allocates Dictator’s resources better than a less competent one. Thus, increasing costs of protection requires a more competent Vizier to deal with them. Machiavelli wrote in *The Prince*: “Let him [the vizier] see that he cannot stand alone, so that
many honors not make him desire more, many riches make him wish for more, and that
many cares may make him dread changes. ... otherwise, the end will always be disastrous
for either one [the prince] or the other [the vizier].”

When are the challengers able to commit to reward the traitor? One such situation
might be that the Vizier has his own political base, be it a certain ethnic or military faction,
with a certain degree of affinity with the dictator’s potential enemies. Thus, a dictator
who thinks of bringing a local warlord to the central government might be interested in
increasing the vizier’s loyalty. For instance, the practice of treating family members as
de-facto hostages was common not only in medieval Khorezm and 12th century England
(Bartlet, 2002), but was widely practiced in most totalitarian dictatorships of the 20th
century.

Proposition 3.2 summarizes the above discussion and adds additional insights.

**Proposition 3.2** In a unique Bayesian perfect equilibrium, the optimal competence of
Vizier, \( \theta^* \), chosen by Dictator (i) decreases with the maximal amount of resources, \( A \),
available to Enemy; (ii) increases with the cost of precautionary measures, \( C \), and (iii)
decreases with the probability that Enemy is strong relative to incumbent Dictator, \( q \). The
wage of the vizier that he gets in the case of no betrayal, \( w^* \), and the equilibrium bribe
\( B^* \) both increase in the dictator’s utility from being at power \( Y \), and in the likelihood of a
strong Enemy type \( q \). Furthermore, bribe level \( B^* \) is decreasing in \( C \).

Proposition 3.2 states that a less competent lieutenant is more likely to be chosen when
either the dictator is weak relative to potential challengers or the dictator values his power
relative to the cost he bears when precautionary measures are taken. Indeed, a higher \( C \),
the cost of fighting that subtracts from Dictator’s value of staying in office, implies a higher
competence of the vizier. Moreover, high \( C \) increases Dictator’s tolerance to betrayals,
as sometimes Vizier gets a wrong signal, and betrayal saves Dictator the cost of defense;
as a result, Dictator is willing to allow Vizier to be bribed more easily. An increase in
the probability that Enemy is strong enough to unseat the incumbent, \( q \), makes a Vizier’s
betrayal more likely; thus Dictator has to put more emphasis on loyalty at the expense of
competence. Furthermore, Dictator has to compensate Vizier more for his loyalty.
One instance where the above analysis has direct implications is the conduct of international negotiations by authoritarian regimes. Unlike their predecessors, modern dictators rarely negotiate on their own. Thus, the choice of a negotiator involves the trade-off we explore: the dictator has to choose a negotiator who is competent enough to bring agreement on favorable terms, yet too much competence might make the negotiator more sensitive to personal alternatives provided by the other side. Kydd (2002) argues that a biased mediator might be more effective in conveying a message to a party in negotiations as a mediator biased towards the recipient of the signal can deliver more credible threats. Dictators, however, often treat suggestions to negotiate as treason, and this may be fatal for subordinates who offer to negotiate. (Since Kydd, 2002, considers mediator’s loyalty as an exogenously parameter, this problem does not arise in his model.)

3.4.2 Enemy’s Problem

At the previous step of the backward induction, we treated Enemy’s behavior as given. Now we consider Enemy’s decision to launch a coup against the incumbent as a function of information that he has about the stability of the regime. Thus, the optimal strategies of Dictator and Vizier, chosen by Dictator, which we derived above, determine Enemy’s payoffs.

Both Dictator and Vizier know that Enemy’s \textit{ex-ante} probability of being strong is \( p = \frac{1}{2}, \) and that Enemy makes his decision to launch a coup conditional on a private signal \( z \) he receives. Suppose that both Dictator and Vizier believe that upon receiving a signal of at least \( z, \) Enemy attacks. The \textit{ex-ante} probability of getting such signal is

\[
\frac{1}{2} \int_{\tau}^{1} 2z \, dz + \frac{1}{2} \int_{\tau}^{1} 2(1 - z) \, dz = \frac{1}{2} \int_{\tau}^{1} 2 \, dz = 1 - \tau.
\]

Conditional on the fact that Enemy does launch a coup, the probability that Enemy is strong, \( q, \) equals \( q (\tau) = \frac{1}{2} (1 + \tau) \) (because the enemy’s signal has a uniform distribution...
as we assumed $p = \frac{1}{2}$). As follows from Proposition 3.1,

$$
\theta^*(q(z)) = 1 - \frac{1}{1 - q(z)} \left( \sqrt{\frac{Aq(z)}{C}} - q(z) \right),
$$

$$
w^*(q(z)) = \frac{A + Y}{2} \sqrt{\frac{q(z)}{C}} - C,
$$

$$
B^*(q(z)) = B^* = \frac{A + Y}{2} - \sqrt{\frac{AC}{q(z)}}.
$$

Let us compute Enemy’s utility if his type is $z$, and enemies with types at least $z$ attack. Enemy wins if the following conditions are met: he is strong (type $t = \text{Strong}$) and Vizier chooses to help Enemy; otherwise he fails. The probability of winning, from Enemy’s perspective is, therefore, $z \cdot \Pr(B \geq B^*) = z \left( 1 - \frac{B^*}{A} \right)$: indeed, Vizier will help a strong Enemy whenever the latter has sufficient money to bribe Vizier, and on the other hand, Enemy will bribe Vizier whenever he has enough resources (see Appendix for a formal proof). Enemy, whenever he has the money to bribe Vizier, will offer him $B^*$, while keeping $B - B^*$ for himself. Due to uniform distribution, the average payoff Enemy gets, conditional on winning, is $Y_e + (A + B^*)/2$. Conditional on losing, Enemy gets his punishment $-D$. Summing up, the utility of Enemy equals

$$
U(z, z) = z \left( 1 - \frac{B^* (q(z))}{A} \right) \left( Y_e + \frac{A + B^* (q(z))}{2} \right) - \left( 1 - z \left( 1 - \frac{B^* (q(z))}{A} \right) \right) \cdot D \quad \text{(3.6)}
$$

$$
= z \left( 1 - \frac{B^* (q(z))}{A} \right) \left( Y_e + D + \frac{A + B^* (q(z))}{2} \right) - D.
$$

This formula is clearly continuous in $z$; hence, type $z$ of enemy must be indifferent between attacking and not attacking. This gives us the following equilibrium condition (for an interior equilibrium, i.e., one where $z \in (0, 1)$; in the Appendix we prove that under Assumption 3.1 all equilibria are interior).

$$
z \left( 1 - \frac{B^* (z)}{A} \right) \left( Y_e + D + \frac{A + B^* (z)}{2} \right) = D. \quad \text{(3.7)}
$$

Before proceeding, let us discuss the following trade-off for the enemy. If fewer Enemy types attack, the average attacker is perceived, by both Dictator and Vizier, to be stronger.
This inspires Vizier, all things equal, to betray with a higher probability. Dictator has two means to counter this: raise Vizier’s wage, or choose a dumber Vizier. As we know from Proposition 3.2, in this case Dictator uses both tools: he chooses a less competent Vizier and pays him more. These two effects unambiguously increase the equilibrium bribe that Enemy has to pay, and this makes each Enemy’s type worse off. Potentially, this could lead to multiple equilibria: if many Enemy types attack, an average plot is weak, Vizier is smart and gets a relatively small wage – so attacking is profitable even for relatively weak Enemies. If few Enemy types attack, each plot is a real threat, so the dictator chooses a smart Vizier and pays him a lot. This is an interesting possibility, and for some parameter values this may be the case. However, under very natural assumptions the equilibrium is unique, and we focus on this case.

**Proposition 3.3** There is a unique equilibrium characterized by a threshold $\tau \in [0,1]$, where Enemies with types higher than $\tau$ attack and Enemies with types lower than $\tau$ do not attack. This threshold $\tau$ is increasing in $D$ (higher punishment implies that only very confident Enemies attack), increases in $Y$ and decreases in $Y_e$ (a plot is less likely if Dictator values power a lot and Enemy values power a little, but on average, plots are relatively dangerous). It also decreases as $C$ increases (more costly defense implies Dictator is more reluctant to use it, and this gives more Enemy types a chance).

The result of Proposition 3.3 is very intuitive. The equilibrium is unique and is characterized by an interior threshold $\tau \in (0,1)$. A higher threshold $\tau$ means less frequent, but more fierce plots. Not surprisingly, this would be the case if Enemy is strongly punished if he fails, or does not value power very much, while Dictator values power (this would be the case, for example, if Dictator is an autocrat and Enemy is a collective body, or a potential democratic ruler who would not be able to get all the benefits of power). The comparative statics with respect to the value of power, $Y$, $Y_e$, and $C$, follow from the fact that the left-hand side of (3.7) is increasing in $q$.

We now combine the results of Proposition 3.2 and Proposition 3.3 to get our main comparative statics result.
Proposition 3.4  Vizier’s competence level, $\theta^*$, increases in $C$ and $Y_e$, but decreases in $D$ and $Y$.

3.5 Extensions

3.5.1 Succession

Once an absence of ordered continuance was considered a major drawback of non-hereditary dictatorship as a form of government (Herz, 1952, Spearman, 1939, Olson, 1993). However, in the first half of the 20th century a number of once-dictatorial regimes survived the death of their founding fathers (e.g., Lenin in Russia and Kemal Atatürk in Turkey). Nowadays, the technology of succession appears to be advanced enough to produce successful transitions in such diverse countries as Syria in 2000, North Korea in 1994, China in 1989, and Congo in 2003. Our model predicts that a ruler with a longer time horizon, e.g., resulting from the assurance of a desired succession, has more incentives to hire the most able agents. Additional support for this result comes from the the last years of kings from the largest European monarchies in England, France, and Russia before they were executed by revolutionaries. At the time they lost the crown, all these monarchs had very young heirs incapable of grasping power if their fathers were dead. And, the last years of each of these rulers were marred by the colossal incompetence of their prime ministers.

Consider the following extension to the previous formal analysis. There is an infinite number of periods, and in each period there is a ruler. The rulers live for one period; however, they may care about the next ruler (their dynasty). Specifically, the successor may be either desirable for the dictator or not. In the first case, the successor’s utility is added to that of the dictator with a discount factor $\beta < 1$. In the latter case, the dictator does not care about his successor’s utility at all. We may interpret $\beta$ as a measure of affinity between the dictator and his successor. It is natural to think that $\beta$ is high in the case of monarchy, but low in the case of army colonels succeeding one another.

---

11 Twenty years after Herz’s article was completed, the last of the European dictators that came into power in 1930s, General Franco of Spain, successfully transferred power to a designated successor.
Each period is characterized by the incumbent dictator’s ability to ensure succession to a desirable heir. Specifically, the state either provides for a secure succession ($S$) or not ($I$). If Dictator wins his encounter with Enemy, he is able to ensure that the successor is (i) desired and (ii) the next state is $S$. However, if Dictator loses in state $S$, then his successor is the Dictator’s heir with probability $P_S < 1$, and the state becomes $I$. If Dictator loses to Enemy in state $I$, his successor is the Dictator’s heir with probability $P_I$, $P_I < P_S$, and the next state is $I$. With complementary probabilities ($1 - P_S$ if the current state is $S$ and $1 - P_I$ if the current state is $I$), Enemy takes over and becomes the next dictator; the next state is $I$.

**Proposition 3.5** In the dynamic game, there exists a unique Markov Perfect equilibrium. In this equilibrium, $U_S > U_I$, and thus the competence of Vizier is lower in state $I$, than in the state $S$. If Dictator cares less about his predecessors ($\beta$ is smaller), the competence of the Vizier is higher both in secure and insecure states.

Proposition 3.5 asserts that a less sure succession leads to less competent agents. The first result may be used to explain the poor governance of monarchs whose immediate heirs are small children, or have other contenders for the throne (i.e., relatives they do not like). Then, Proposition 3.5 demonstrates that less desired succession leads to better agents. It helps to explain the difference between ‘party-machine’ dictatorships such as Mexico in 1940–90, where members of a non-representative selectorate succeed each other as leaders of the country, but have neither desire nor possibility to pass this post to their children, and monarchies, whose rulers have such desire and possibility. The model predicts that a personalist dictatorship is less likely to witness competent advisors than an institutionalized dictatorship. Domínguez (2002) reaffirms that “the most successful authoritarian regimes, namely, historical bureaucratic empires, had means of succession from one monarch to the next and featured bureaucratic organizations for the sharing and exercise of power”. Not coincidentally, Mexico, probably the most successful of 20th century dictatorships, had a well-institutionalized procedure for succession for almost six decades.
3.5.2 Negative Selection

Political scientists (e.g., Lewis, 1978, Linz and Chehabi, 1998) and retired politicians (Speer, 1970) have been long aware that dictatorial rule keeps able people from joining high-level politics. For an economist, this is a familiar case of the Akerlof adverse selection problem: the more severely the dictator punishes those who betrayed him (if he survives the betrayal), the less the ability of agents applying for the job. Hence, the dictator faces a trade-off between high incentives for agents already on the job, which are provided by harsh punishment for betrayal, and low incentives to encourage potential applicants to apply for the job. Indeed, the harsher the punishment for betrayal is, the lower is the expected utility of a competent advisor. Since it is the agent’s competence that allows him to discriminate among potential plotters, he would never need to use his competence when the price of betrayal is infinite disutility.

Formula (3.1) allows us to extend the analysis to the environment where potential viziers have a choice whether or not to enter politics at all. The expected benefit of staying loyal for Vizier is equal to \( L = w_f + (1 - \theta) \frac{1}{q} (w_f - w_b) \). Above, we focused on the case where \( w_b \geq 0 \). However, Dictator may induce more loyalty on Vizier’s behalf if he can actually punish Vizier for treason should he survive, i.e. have \( w_b < 0 \). Though providing better incentives for those who are on the job, this reduces expected utility of competent viziers. Thus, if there is some reservation utility of an agent of type \( \theta \), \( H(\theta) \), which is a continuously differentiable function of \( \theta \) and \( H'(\theta) > 0 \), then certainly there is a trade-off between “loyalty” and “attractiveness of the job.” Then, for a sufficiently high punishment, only relatively incompetent advisors are self-selected.\(^\text{12}\)

Thus, a dictator has incentives to commit to an optimal punishment that is less than execution. However, the very nature of dictatorships precludes such commitment; in a democracy, it is easier to commit to a mild punishment. In a democracy, though a punishment for political betrayal might be politically severe, it rarely brings significant personal harm. A U.S. President is bound by laws not to kill a cabinet member who pursues his own presidential ambitions, as was the case with Treasury Secretary Salmon.

\(^{12}\) The working paper version contains formal assertions and proofs.
Chase in Lincoln’s first cabinet (e.g., Dudley, 1932), or Attorney General Robert Kennedy in the first cabinet of President Johnson.\textsuperscript{13} Betraying a dictator such as Saddam, Castro, or Marcos might have been more costly for their ministers.\textsuperscript{14} The following proposition summarizes the above discussion. Thus, in a relatively mild autocracy, leaders are more exposed to political treason, but the pool of applicants to the agent’s position is likely to be better. Conversely, the bloodiest dictator may feel relatively safe from betrayal, but the agents he will have to choose from will be extremely incompetent.

One potential counter argument is that the dictator could enter the private labor market and selectively depress rewards for competence, say, by threatening the family members of potential agents if the agent refuses to enter his service.\textsuperscript{15} While this argument certainly does have merit when applied to a single agent, this approach seems impossible on a large scale.\textsuperscript{16} Mass emigration is a most clear indication of unfavorable circumstances for talented people. In the first five years of the Mussolini regime, one and half million people left Italy (Cannistraro and Rosoli, 1979). In Haiti, Trujillo “used all means at his disposal to reinforce the natural isolation” (Hartlyn, 1998). Furthermore, the exile of the political and intellectual elite, which is a tiny fraction of any country’s population, might not be easily detected by crude statistical data. For example, the departure of Albert Einstein,

\textsuperscript{13} In a non-technical note, Edwards (2001) points to the same loyalty-vs.-competence trade-off in recent low-level presidential appointments in the U.S.

\textsuperscript{14} In January 1984, the honorary title of ‘Hero of the Cuban Republic’ was conferred upon Gen. Arnaldo Ochoa in recognition of his extraordinary contributions to the insurrection against Batista, to the consolidation of the nation’s defense, and for his service in international missions. In June 1989, MINFAR Minister Raul Castro explained that Gen. Ochoa ‘was no longer the rebel soldier, the invader of Camilo’s column, the internationalist in Venezuela, the commander of our troops in Ethiopia.’ In July 1989, the prosecutor’s closing remarks stated that ‘it became evident that we were confronted with a crime of treason committed against the fatherland, against the people, against his superiors, and against the very idea of what a revolutionary, a military chief, and a Cuban internationalist fighter really is.’ In accordance with the ‘sentence dictated by the Special Military Court, Case No. 1 of 1989,’ Gen. Arnaldo Ochoa and three others faced a firing squad in July 1989. (Alfonso, 1989 and sources cited therein.)

\textsuperscript{15} Gershenson and Grossman (2001) analyze how both cooptation and repression were employed to encourage loyalty to the Soviet regime.

\textsuperscript{16} Political scientists working on modern dictatorships have long been aware of the problem to find a rationale for either ‘random’ terror against population or ‘purges’ against loyal members of the regime (see Friedrich and Brzezinski, 1956, p. 150-151 for discussion of the difference; also, Arendt, 1951). The idea of suppressing the reservation utility of those who might have choose to deliberately abstain from politics/government might provide such a rationale. Another possible policy is restrictions on emigration, a common feature of many authoritarian regimes.
Joseph Schumpeter, Thomas Mann, and John von Neumann preceded mass emigration of the intellectual elite from Europe, but might have had a more profound impact on the intellectual, and by implication, political life of their home countries. Thus, even for individual geniuses, providing the incentives to work in a certain political environment might be a complicated task for the dictator.  

3.6 Conclusion

In a recent inquiry into the dynamic nature of dictatorships, Acemoglu, Robinson, and Verdier (2004) suggest that “while the academic study of strongly-institutionalized polities is well advanced, there are few studies, and less of a consensus, on the nature of weakly-institutionalized polities.” Poor governance in, and the degeneration of, mature dictatorships allow for a number of plausible explanations. These include the greediness and selfishness of the dictator, as well as his personal incompetence and inability to listen and follow advice. We use the formal apparatus of economic theory to investigate agency problems in dictatorships as compared to democracies. We demonstrate that it is the unwillingness and inability of the dictator, fearful of opportunistic behavior by the agent and potential betrayal, to surround himself with competent associates that causes the poor performance of dictatorships in the long run. Since the definition of competence we use is, in a sense, all-encompassing, the resulting incompetence will sooner or later have an adverse effect on the policies carried out and consequently on economic performance and social welfare. The most profound effect to be felt in neo-sultanistic regimes (Chehabi and Linz, 1998) and “control regimes” (Bates, 2004), which rely on government intervention as the primary mechanism of resource allocation.

In the much less frightening circumstances of the last decades of the Soviet rule, talented young Russians chose mathematics and the natural sciences, generally avoiding politics (and, e.g., political science) as an occupation. One result, besides flourishing science, was that political positions were often occupied by profoundly mediocre appointees.

The economic performance of dictatorships is a recurring topic for empirical research (e.g., see Epstein et al., 2004, and Przeworski et al., 2000, for opposing views; see also Persson, 2002, Gandhi, 2005). While acknowledging that the emerging consensus in recent empirical studies does emphasize the advantages of democracy, Glaeser et al. (2004) argue that a dictatorial rule might be conducive for economic growth. For medieval times with unreliable estimates of the economic growth, De Long and Shleifer (1993) find that “a region ruled by an absolutist prince saw its total urban population shrink by one hundred thousand
There is a strand in the literature on dictatorships that argues that the dictators have an advantage in choosing the most able man for government positions, while in democracies the first-best choice may be impossible (e.g., de Tocqueville, 2000, originally published in 1831). Though there is a certain merit to this point, the circumstances in which a dictator has this advantage are limited. One such situation appears when a new dictator emerges after years of political stagnation or political turmoil, bringing a whole class of politically young and able people with him. However, though emergence of new faces in politics or government may coincide with the accession of a dictator, it might be the same political wave that removed the former elite that both made a new dictator possible and extended the opportunities of other talented individuals. One example might be Napoleon’s famous marshals, a group of brilliant military officers of plebeian origin who pursued their army careers to a point previously reserved to people of noble origin only. Though their military glory came in full under Napoleon’s command in the early 1800s, it was the French revolution of the earlier decade that made a dramatic break in their careers possible.19

Probably the most prominent modern theory highlighting the advantages of dictatorship is Mancur Olson’s “stationary bandit” paradigm (Olson, 1993).20 As Wintrobe (2000) rightly observes (see also Haber 2005), “the problem with Olson’s analysis is that, comparing dictatorships, the worst regimes in human history appear to be precisely those such as Nazi Germany, Soviet Russia, or Cambodia which appear to have been the most encompassing”. The agency theory of dictatorships suggests an explanation why even a benevolent dictator may fail to implement a socially desirable policy. The loyalty-competence trade-off, which is much more severe when commitment mechanisms are weak, is in a sense

---

19 Supporting our basic story, Napoleon’s Marshal Jean-Batist Bernadotte – in Napoleon’s opinion, one of the two marshals with a war talent equal to that of the Emperor himself – left Napoleon not out of fear of ultimate defeat, but at the zenith of Napoleon’s power in 1811. The other military genius, General Moreau, left Napoleon – via participation in an unsuccessful plot – even earlier, in 1803.

20 The idea of a stationary bandit can of course be traced back to Hobbes, who appraised monarchy as a system where public affairs are run perfectly because they are actually private. A formal model can be found in McGuire and Olson (1996). Epstein and Rosendorff (2004) analyze what might prevent an autocrat from pursuing a growth-enhancing policy.
an indispensable feature of any unconstrained dictatorship. Even if a powerful dictator
reads this paper, and understands the logic, it gives no help to him if he is insistent on
keeping his power unrestricted. Until he opts for a sustainable delegation of power to other
political institutions, he will have no opportunity to improve the quality of his ministers.

Appendix

Proof of Proposition 12. The payment to vizier when the enemy decides not to attack,
w_n, should certainly be set to the minimal value 0. Suppose \( w_b > 0 \). Then the dictator
would be better off by decreasing \( w_b \), as this would increase \( L \), thus making the vizier
more loyal and decreasing the dictator’s payment to vizier in case of betrayal. Hence,
w_b = 0 in the optimal contract. Suppose \( w_f > w_e \). In that case, the vizier, when facing
a weak plot (i.e., getting signal \( s = \text{Weak} \), which unambiguously means \( t = \text{Weak} \)) will
strictly prefer to fight (in which case he gets \( w_f \)) rather than to allow the plot to unravel
(in which case he gets \( w_e \), as this is not a betrayal). If, however, the vizier gets a strong
signal, he may betray the dictator. Given that, the dictator would be better off choosing
to fight regardless of vizier’s advice. If, instead, \( w_e > w_f \), then the dictator would be
better off if he decreased \( w_e \) slightly. Finally, consider the possibility \( w_f = w_e = 0 \). But
then the vizier would betray the dictator every time he receives a strong signal, i.e., he
would never defend. However, then the dictator could as well never defend himself, i.e.,
without any help of the vizier. This completes the proof. ■

Proof of Proposition 3.1. Denote \( k = 1 + (1 - \theta) \frac{1 - q}{q} \geq 1 \). Then the dictator’s utility
(if the enemy chooses to attack) then equals

\[
U^D = Y \left( 1 - q + \frac{1}{A} kw \right) - (1 - kq) w - \frac{1}{A} k^2 qw (C + w).
\]

Here, the first term reflects the benefits of survival times the probability of survival, the
second term is the wage paid if the vizier got a weak signal and did not defend, while the
last one is the expenses on defense and wage if the vizier chose to fight. We now have
The first two equations give us the first-order conditions, while the last three ones allow us to compute the hessian, which equals

\[
\frac{2}{A}qw(C + w) \left( \frac{2}{A}k^2q - \left( \frac{1}{A}(Aq + Yq - 2Ckq - 4kqw) \right)^2 \right).
\]

We need to verify that if is negative, provided that the first-order conditions hold. Substituting \( k = \frac{1}{2} \frac{A + Y}{C + w} \) (which follows from the condition \( \frac{\partial U}{\partial k} = 0 \)), the Hessian reduces to

\[
Cw \frac{(A + Y)^2q^2}{(C + w)^2 A^2} > 0.
\]

Consequently, the solution to the first-order conditions \( \frac{\partial U}{\partial k} = 0, \frac{\partial U}{\partial w} = 0 \) gives a global maximum, provided that it is unique. To verify uniqueness, let us find the optimal \( w \) for a given \( k \) from \( \frac{\partial U}{\partial w} = 0 \); we obtain

\[
w = \frac{Akq + Ykq - Ck^2q - A}{2k^2q}.
\]

Plugging this into the equation

\[A + Y - 2Ck - 2kw = 0,\]
which follows from $\frac{\partial U}{\partial k} = 0$, we get

$$\frac{1}{kq} (A - Ck^2q) = 0.$$  

Consequently,

$$k = \sqrt{\frac{A}{Cq}}.$$  

From this we obtain the required formulas, which give the unique solution to the dictator’s maximization problem:

$$\theta^*(q) = 1 - \frac{1}{1 - q} \left( \sqrt{\frac{Aq}{C}} - q \right),$$

$$w^* = \frac{A + Y}{2} \sqrt{\frac{C}{A}} - C,$$

$$B^* = \frac{A + Y}{2} - \frac{AC}{q}.$$  

This completes the proof.  

Proof of Proposition 3.2. The comparative statics of vizier’s competence $\theta^*$ with respect to $A$ and $C$ follows straightforwardly from (3.3). To find that it decreases in $q$, compute the derivative

$$\frac{d\theta^*(q)}{dq} = -\frac{1}{2Aq (1 - q)^2} \left( 1 - 2\sqrt{\frac{C}{A}}q + q \right) \sqrt{\frac{q}{AC}}.$$  

We have

$$1 - 2\sqrt{\frac{C}{A}}q + q > 1 - 2\sqrt{q} + q = (1 - \sqrt{q})^2 \geq 0,$$

since $C < A$ by assumption. Consequently, $\theta^*$ is decreasing in $q$.

The vizier’s wage and the equilibrium bribe are increasing in $Y$ and $q$, as immediately follows from (3.4) and (3.5). Similarly, the equilibrium bribe is decreasing in $C$, as follows from (3.5). This completes the proof.
Proof of Proposition 3.3. Since equilibrium threshold $\pi$ is defined by (3.7) and $\pi = 2q - 1$, to prove uniqueness it suffices to show that

$$(2q - 1) \left(1 - \frac{B^* (q)}{A}\right) \left(Y_e + D + \frac{A + B^* (q)}{2}\right)$$

is increasing in $q$. From Proposition 3.2 we know that $B^* (q)$, and thus the entire last factor, is increasing in $q$. Consequently, it is sufficient to prove that $(2q - 1) \left(1 - \frac{B^* (q)}{A}\right)$ is increasing in $q$. We have

$$(2q - 1) \left(1 - \frac{B^* (q)}{A}\right) = (2q - 1) \left(1 - \frac{\frac{A + Y - \sqrt{AC}}{2}}{A}\right) = (2q - 1) \left(\frac{1}{2} - \frac{Y}{2A} + \sqrt{\frac{C}{Aq}}\right).$$

Differentiating with respect to $q$ yields

$$\frac{\partial}{\partial q} \left((2q - 1) \left(\frac{1}{2} - \frac{Y}{2A} + \sqrt{\frac{C}{Aq}}\right)\right) = \frac{1}{2Aq} \left(A \sqrt{\frac{1}{A} \frac{C}{q}} + \frac{2AQ - 2Yq + 2AQ}{2Aq} \sqrt{\frac{1}{A} \frac{C}{q}}\right) > 0,$$

since by Assumption 3.1 $A > Y$. Hence, equation (3.7) has at most one solution. If it does, this is the only equilibrium, since $\pi = 0$ cannot be an equilibrium threshold (enemies who get signal $z = 0$ will never find it optimal to attack), while if $\pi = 1$ so that no enemy attacks, the strongest enemies would deviate and attack, since then the right-hand side of (3.7) is larger than $D$. If $D$ is large enough, equation (3.7) may fail to have solutions on $[0, 1]$, and then $\pi = 1$ provides an equilibrium threshold, so no enemy attacks because of prohibitively high punishment. This proves the uniqueness of equilibrium threshold.

To show that the solution to (3.7) is increasing in $D$, observe that

$$\frac{\partial}{\partial D} \left(\pi \left(1 - \frac{B^* (q (\pi))}{A}\right) \left(Y_e + D + \frac{A + B^* (q (\pi))}{2}\right) - D\right) = \pi \left(1 - \frac{B^* (q (\pi))}{A}\right) - 1 < 0.$$

So, if $D$ increases, a higher $\pi$ is required to satisfy (3.7). Similarly, the left-hand side of (3.7) is increasing in $Y_e$ (since $B^* (q (\pi))$ does not depend on $Y^e$), so $\pi$ is decreasing in $Y_e$. 

140
Observe now that the left-hand side is decreasing in $B^*$, as it may be rewritten as

$$
\tau \left( \left( 1 - \frac{B^*}{A} \right) (Y_e + D) + \frac{1}{2A} \left( A^2 - (B^*)^2 \right) \right).
$$

Consequently, it is decreasing in $Y$ and increasing in $C$, as follows from Proposition 3.2. This implies that $\tau$ is increasing in $Y$ and decreasing in $C$, which completes the proof. ■

**Proof of Proposition 3.5.** We start by writing down the utility functions of and enemy in each of the two situations. For $X = S, I$ we have

$$
U_X = f_X (U_S, U_I) = (2q_X - 1) (Y + \beta U_S) + (Y + \beta U_S - \beta P_X U_I) (1 - q_X + q_X G_X) (2 - 2q_X)
$$

$$
-w_X (\theta_X (1 - q_X) + (q_X + (1 - \theta_X) (1 - q_X)) G_X) (2 - 2q_X)
$$

$$
-C (q_X + (1 - \theta_X) (1 - q_X)) G_X + \beta (P_X U_I) (2 - 2q_X),
$$

$$
0 = (2q_X - 1) \left( 1 - \frac{B^*(q_X)}{A} \right) \left( (1 - P_X) \beta U_I + D + \frac{A + B^*(q_X)}{2} \right) - D.
$$

To show that the equilibrium is unique, suppose that $U_I$ increases. This means, using the results of Proposition 3.3, that both $q_S$ and $q_I$ should be lower for the latter equation to hold, as more enemies would now prefer to attack. However, both $f_S$ and $f_I$ are increasing in $q_S$ and $q_I$, respectively, since higher $q$ implies that fewer enemies actually attack. Consequently, if $q_S$ and $q_I$ decrease, it must be that $f_I$ decreases. But in equilibrium it equals $U_I$; hence, $U_I$ is uniquely defined in equilibrium. But then $q_S$ and $q_I$ are uniquely defined, and the equation $U_S = f_S (U_S, U_I)$ defines a contracting mapping for $U_S$, which is thus also uniquely determined. But since utilities on the right-hand sides are uniquely defined, we immediately obtain uniqueness as an immediate corollary of Proposition 3.3.

To show that in this equilibrium, $U_S > U_I$, consider the mapping for the equilibrium values of $q_S$ and $q_I$ given by $U_S = f_S (U_S, U_I), U_I = f_I (U_S, U_I)$. The mapping is contracting (in sup-metrics), and iterations converge to the equilibrium values of $U_S$ and $U_I$. It is easy to see, however, that the inequality $U_S > U_I$ is preserved under this mapping, and so in the limit $U_S \geq U_I$. However, as $U_S \neq U_I$, we must have that $U_S > U_I$. Now
Proposition 3.2, together with inequality $P_S > P_I$, implies that $\theta_S > \theta_I$.

Finally, we need to show that lower $\beta$ leads to higher competence of the vizier. Suppose, to obtain a contradiction, that $\beta U_I$ increases. Then $q_S$ and $q_I$ are lower, and, as before, we find that both $U_S$ and $U_I$ must decrease (lower $\beta$ reinforces the effect). Hence, $\beta U_I$ decreases, which implies that $q_S$ and $q_I$ are higher. Now again consider the mapping $(f_S, f_I)$ for the new lower $\beta$ and higher $q_S$ and $q_I$. For any arguments $U_S$ and $U_I$ such that $U_S > U_I$, $f_S$ and $f_I$ are now lower. This implies that in the new equilibrium, $U_S$ and $U_I$ are lower. Moreover, since $\theta_S > \theta_I$ and thus $q_S < q_I$, $U_S$ is decreasing faster than $U_I$, and so $Y + \beta U_S - \beta P_S U_I$ and $Y + \beta U_S - \beta P_I U_I$ are decreasing. But then the competence of the vizier becomes higher as $\beta$ becomes lower both in state $S$ and in state $I$. This completes the proof. $\blacksquare$


145


147


149


Skocpol, Theda (1979) *States and Social Revolutions: A Comparative Analysis of France, Russia and China*, Cambridge: Cambridge Univ. Press.


