

Apolte, Thomas

**Conference Paper**

## A Theory of Autocratic Transition. Prerequisites to Self-Enforcing Democracy

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2018: Digitale Wirtschaft - Session: Institutions, No. A17-V1

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Apolte, Thomas (2018) : A Theory of Autocratic Transition. Prerequisites to Self-Enforcing Democracy, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2018: Digitale Wirtschaft - Session: Institutions, No. A17-V1, ZBW - Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/181500>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# A Theory of Autocratic Transition

Prerequisites to Self-Enforcing Democracy

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

July 13, 2018

## Abstract

This paper aims at contributing to a better understanding of the conditions of self-enforcing democracy by analyzing the recent wave of autocratic transitions. Based on a game-theoretic framework, we work out the conditions under which governments may induce the diverse public authorities to coordinate on extra-constitutional activities, eventually transforming the politico-institutional setting into one of autocratic rule. We find three empirically testable characteristics that promote this coordination process, namely: populism and public support, corruption, and a lack in the separation of powers. By contrast, low degrees of corruption and strongly separated powers can be viewed as prerequisites to self-enforcing democracy.

**JEL classification:** D02; D72; D74; P48

**Keywords:** self-enforcing democracy; political regimes; autocratic transition.

# 1 Introduction

The wave of democratization following the collapse of the Soviet Union and the Eastern-European communist world has shaped a remarkably optimistic view toward the future of democratic rule around the world. This notwithstanding, a non-negligible share of newly established democracies have not survived their first years of existence. The cases of failed democratic transitions have long been outshined by the success stories, particularly those in Middle and Eastern Europe. That only changed when even some of the early success stories turned dubious. While Belarus, Uzbekistan and Kazakhstan went more or less straight into plain dictatorship following the Soviet collapse, the other three Central Asian former Soviet republics—the countries of Caucasasia as well as Ukraine and Moldova—remained partly democratic at best. On a worldwide basis, a number of countries that initially seemed to evolve into successful young democracies eventually backslid into autocratic structures. Perhaps the most prominent examples of this are Russia, Venezuela, Thailand, Turkey and potentially even Hungary and Poland.

On a rebased scale of the Freedom House democracy index<sup>1</sup> that ranges from 1 (pure dictatorship) to 10 (full-fledged democracy), Figure 1 gives an overview of those countries, since 1990, that lost more than three index points after having reached a maximum. The greatest loss was experienced by The Gambia, which was up at 9.25 in 1992, then dramatically dropped to 1.75 in 1994 and finally scored 2.5 in 2016. Most of the countries summarized in Figure 1 reached their maximum in the early and mid 1990s. Some of them—most notably perhaps Thailand—was able to stay there for some years, while others—like The Gambia but also Belarus—fell back to autocracy as fast as they had risen to democracy.<sup>2</sup>

A high rate of failure of democratic constitutions is by no means only a recent phenomenon. To the contrary, the history of modern-age democracy

---

1. As an average of the subindices “political rights” and “civil liberties”.

2. Appendix A gives a more detailed description of countries that underwent at least a partial autocratic transition since 1990.

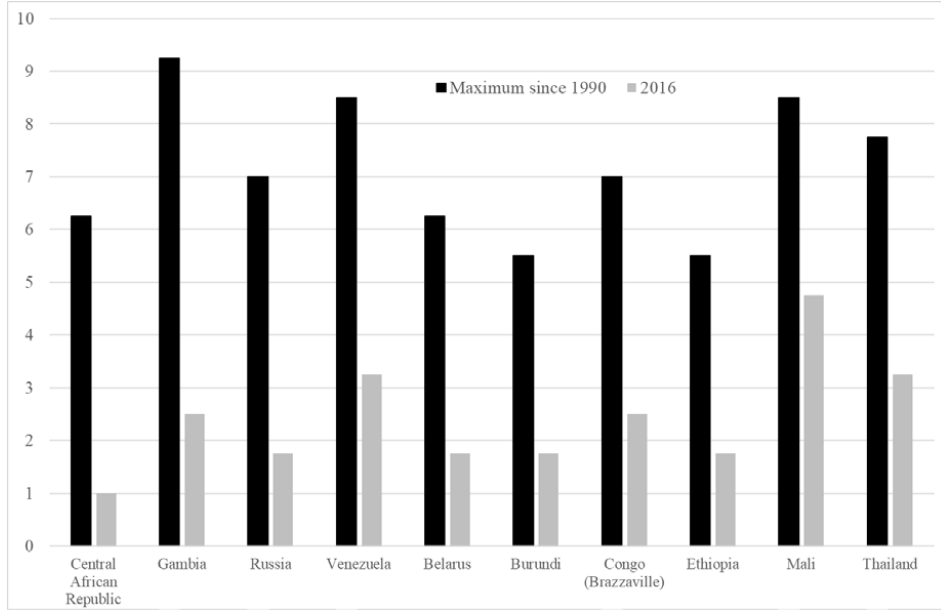


Figure 1: Autocratic Transition

is full of failing endeavors to establish democratic constitutions. According to Ferguson (2001, Appendix E), more than 50 percent of the European interwar democracies failed (see also Weingast 2005, 89). Perhaps more surprising, then, is the development of seemingly established, though still young, democracies in Europe, like those in Hungary and Poland. Recently, however, their governments became notorious for openly challenging widely held convictions about indispensable institutional traits of democracy; remarkably, these governments have enjoyed extensive public support while doing so. This applies in an even more pronounced way to the government leaders of Turkey and Russia, who have already turned their countries into plain autocracies. These leaders started by attacking standard matters of course in established democracies, such as the freedom of the press and an independent judiciary. At the same time, these leaders denounced such components of democracy as instruments secretly established by some internal or external conspirators. Further, they claim that these instruments are a threat to national values as well as to vital national interests and sovereignty. This, then, is the basis for how these governments justify restricting activities by national and international NGOs as well as by insubordinate media represen-

tatives. While some democracies have been able to sustain attacks like these, some others have already fallen victim to them. Russia and Turkey seem to be clear-cut cases of the fallen democracies while the fate of the Hungarian and Polish democracies is still open to date.

These observations represent the point of departure for this paper. Instead of determining what conditions promote autocratic takeovers, this paper derives empirically testable conditions under which democracies sustain or fall victim to autocratic transition. In doing so, the paper aims at contributing to the heretofore relatively narrow literature on self-enforcing democracy and self-enforcing constitutions as it has been developed by authors like Przeworski, Weingast, Fearon and others. Based on a simple game-theoretic analysis, we carve out the conditions under which a democratic constitution can be expected to sustain autocratic attacks by sitting government leaders. To be sure, we do *not* aim to explain how a democratic structure evolves into an institutional equilibrium on which at least the decisive individuals somehow coordinate. Rather, we take such an equilibrium as our point of departure in order to establish the conditions under which a government leader may be able to challenge the existing equilibrium. Also, we do not aim to analyze autocratic transitions pursued by the military. We rather contribute to the question of under what conditions governments will adhere to the constitutional rules by virtue of their individual incentives and, by contrast, under what conditions they may embark on a process toward autocratic transition.

Our central hypothesis rests on a strategic interaction between a government leader on the one hand and a number of other key government officials, like heads of executive offices, judges, chief commanders of the police or the military, and so forth. We do not assume the general public to be capable of credibly threatening the government with public uprisings or the like, mainly for reasons of their collective-action problems. However, in our model, their vote may nevertheless become potentially challenging to a government leader that aims at autocratic transition. This challenge rests on existing rules like the publicly known time schedule of public elections in combination with the

key government officials to mutually monitor each other.

The rest of the paper is organized as follows: We give a brief overview of the relevant literature in section 2. In section 3, we develop a limited version of the model. We expand the transition model to a full model in section 4. In section 5, we discuss our results and derive some normative implications as well as a set of empirical implications. We conclude in section 6.

## 2 Self-Enforcing Democracy

The concept of self-enforcing democracy was introduced by Przeworski in a book on democratization in Eastern Europe and Latin America (see Przeworski 1991). According to this initial concept, a democratic constitution is self-enforcing if an incumbent expects to be better off when stepping down following a lost election rather than sticking with his position by force and, hence, breaching the rules of the constitution. If the incumbent has a sufficiently good chance of being reelected in the due course of a subsequent election and/or will be entitled to a sufficiently generous pension, then this condition is likely to be satisfied. If not, the constitution may lose its self-enforcing character.

Weingast (1997, 2005) models self-enforcing democracy as a game, whereby the players of the game are a group of citizens that face a government leader who may transgress against the citizens' constitutional rights. In a one-shot game, the citizens are trapped in a prisoners' dilemma situation that precludes coordinated action against the government leader. This is different in a sequence of repeated games, since a sequence allows for revolting behavior against the incumbent in equilibrium strategies (see also Mittal and Weingast 2011).

Some 15 years after having introduced his initial concept of self-enforcing democracy, Przeworski came up with an entirely different approach (see Przeworski 2005; Benhabib and Przeworski 2006). The point of departure was the observation of a narrow correlation between per-capita income and demo-

cratic resilience. At the heart of this approach is the assumption that per-capita income is subject to decreasing marginal utility, while utility of living in a democracy is assumed to be independent of income levels. Above some income level, then, marginal utility of the poor from income redistribution will drop below the utility of living under democracy. As a result, the poor will shy away from a revolt against the rich since there is a risk that such a revolt might terminate democracy. The rich, in turn, have an incentive to provide redistribution of income in order to reduce the incentive of the poor to revolt.

This more recent approach has not remained undisputed. Traversa (2015) argues that the results cannot be generalized because the model rests on too narrow a specification of the utility function of the poor. He demonstrates that the central finding disappears altogether when the utility function is slightly different.

Fearon (2011) introduced a model in which the general public implicitly threatens that it will revolt if the government transgresses against the constitutional rules of a democracy. Different from Weingast's approach, he assumes that the general public's strategic interaction with the government is a coordination game rather than a conflict game like the prisoners' dilemma. As a consequence, all the public needs to coordinate on a revolt equilibrium is a distinctive signal in combination with some "warm-glow benefits" of participation. In Fearon's view, a sufficiently distinctive signal might be electoral fraud.

Like Fearon, Przeworski's (2005) more recent approach reduces the issue to a coordination problem. It is indeed akin to approaches that explain institutional or constitutional stability by assuming that the population will coordinate on mutually consistent patterns of behavior (Calvert 1995; Hardin 2006; Przeworski 2006). Note, however, that coordination on a set of mutually benefiting constitutional rules is different from coordination on a rebellion equilibrium. This applies at least if there is a positive expected value of individual participation costs even when such a rebellion is successful. In a long-established democracy, participation costs might be negligible. Elec-

toral fraud is then very likely to spark an instant wave of protests that sweeps away the fraudulent government (Hyde and Marinov 2014). However, in the case of a democracy on the edge of autocratic transition, things are likely to be different. Electoral fraud may be perceived as a signal to rebel by only a part of the population, while another segment of the population might even appreciate when the security forces violently suppress protests. In such an environment, participation in rebellious activity implies a considerable risk of being injured or arrested; this turns a successful rebellion into a public good. Public resistance to an autocratic transition can then better be described as a prisoners' dilemma rather than as a pure coordination problem.

Weingast (1997) acknowledges the public-goods problem and solves it by referring to repeated games. However, almost any equilibrium can be generated in repeated games, and this is why we follow a different approach still. We model autocratic transition versus democratic stability within the structure of a game, the players of which are the government leader on the one hand and a number of key government officials on the other. Autocratic transition evolves within a certain time period. During this period, a government leader always faces the risk of being removed from office by constitutional means, be it on the basis of public elections or on the basis of court procedures. Both may well be accompanied or even initiated by public protests, but that does not need to be the case. Indeed, the efforts of autocratic transition by the government leader may even be accompanied by considerable and visible support by at least part of the public. Hence, the public may either be supportive of or threatening to a government leader who pursues autocratic transition, depending on the government leader's popularity.

We depart from most of the literature around self-enforcing democracy in two ways. The first is that we do not model the strategic interaction between a political elite on the one hand and the citizens on the other. Rather, we analyze the micro-structure within the political and administrative elite of a country. Once a democratic equilibrium is established, each member of this elite finds himself trapped in a structure of mutually enforcing control mechanisms within the elite which does not make it worthwhile for each of the



individual elite members to depart from the established constitutional equilibrium strategy. As long as this equilibrium is unique, a government leader aiming at autocratic transition cannot expect the other key government officials to follow him on a path of extra-constitutional action. There are, however, conditions under which such an equilibrium is unique, which implies the risk of a switch away from an existing democratic equilibrium.

In order to handle complexity, we split the analysis of the entire autocratic transition process into two sets of subgames that we later combine into a full picture. We refer to the first set of subgames as the *transition game* and the second set as the *post-transition game*. We will demonstrate that the players in the *transition game* are plagued by a time-inconsistency problem that they might or might not be able to solve in the *post-transition game*.

### 3 The Transition Game

Consider a country whose political system starts as a democracy and may eventually be subject to autocratic transition. All players are government actors. For simplicity, we assume that the government  $G$  as the top of the political hierarchy consists of only a single individual that is elected and might eventually be voted out of office. We refer to  $G$  as the *government leader*. Below the government leader, we have a number  $N$  of key government officials  $O_i$  with  $i \in \{1, 2, \dots, N\}$ . These individuals are appointed civil servants on lifetime positions that cannot be voted out of office. Upon having acted in an illegal way, however, a government official can still find himself dismissed. The number of all government actors taken together is  $\bar{N} = N + 1$ . Among the government officials, we may consider the leading representatives of the different branches of government as well as leaders of the police, the military, or some secret service. While  $G$  is the formally inaugurated head of the country's government, it is the players  $O_i$  that effectively run the country; each of these players decides to do so strictly on the basis of the existing legal system and, within that restriction, on the basis of the government

leader's orders. However, each of these players may also decide to transgress against some of these rules. Moreover, each of them can, at least to a certain extent, refuse to follow  $G$ , again either within or beyond the limits of the constitutional rules.

We assume the government leader to signal his intentions by either abiding to the constitutional rules or by violating them. Upon having observed extra-constitutional conduct by the government leader, a share  $h \in [0, 1]$  of the government officials effectively follows  $G$  in violating the constitution, while a share  $1 - h$  does not. A necessary condition for a full autocratic transition is that  $h \geq h^c \in (0, 1)$ . Note that  $h^c$  is equivalent to the critical mass in the multi-equilibrium setting of the "Granovetter type" (Granovetter 1978; Marwell and Oliver 1993), in which the expectation of a certain minimum share of a population to participate motivates individual participation, thus turning the expectation into a self-fulfilling prophecy.

At the end of a completed autocratic transition process, the government leader  $G$  will no longer be challenged in his position by democratic means, though he will at least potentially be challenged by the government officials, specifically the military or the police, for example. However, we assume the transition process to take some time until its full completion. During this time of autocratic transition, public elections do still take place. While these elections might not be binding with certainty, there will at least be a chance that some government authorities still enforce their results, possibly even against the will of some of their respective heads. Since there is a non-zero probability that public elections will be held during this time, we can safely assume a positive probability that the government leader will be deposed by legal means during the autocratic transition period. As his chance to survive the transition period in office hinges, *inter alia*, on his reelection probability, it is generally positively correlated with his public popularity.

If a government official follows along with the extra-constitutional activities of a government leader, he may gain career benefits, but he may also suffer significant costs, such as losing his job or facing additional formal penalties. Furthermore, in this situation government officials face two major sources of

uncertainty. First, to determine what decisions their colleagues are making and effectively executing, they must rely on sufficiently informative signals; second, they face a non-zero probability that the government leader will be deposed during the autocratic transition period, which may backfire on their decision to follow the government leader. Hence, a government leader who signals his intentions by purposefully violating the democratic rules of the game creates a delicate environment full of ambiguities for the government officials.<sup>3</sup>

To be precise, holding an office in a democratic constitutional environment provides a payoff of  $U_G = U_i = 1$  to both the government leader and each of the government officials. The payoff can be broadly understood as an indicator of individual income or wealth. But it can also be understood as the level of influence one has on certain political programs or the like. On top of that, autocracy generates an additional pie  $R$  of payoffs for all government officials taken together, which we shall refer to as the autocratic rent. In the case where government actors share the autocratic rents, each of them would receive a share  $\frac{R}{N}$  and hence an autocratic payoff of  $U_{G,i} = s := \frac{R}{N} + 1$ . Payoff  $s$  is hence a multiplier of a government official's reservation utility. However, we also allow for an unequal distribution of the autocratic rents. Such an unequal distribution of the autocratic rent may, for example, end up in payoffs  $U_i = 1 \ \forall i$  and  $U_G = R + 1$ .

Each player may either respect the (initially) democratic constitution, in which case his choice is referred to as  $c_{i,G}$  (comply); or he might start mixing extra-constitutional measures into his activities, in which case his choice is referred to as  $d_{i,G}$  (defect). Players  $O_i$  can observe  $G$ 's choice prior to their own choice, but they cannot mutually observe their respective choices.

We assume the government officials to have less than perfect control over the respective public authorities they preside. In particular, the execution of their decisions may effectively be undermined from within the bureaucracy. This assumption reflects two crucial aspects of reality: First, there is usually

---

3. For an early but still instructive non-formal analysis, see Tullock (1987).

some managerial discretion left to subordinates in bureaucracies. Second, bureaucrats that are in a leading position but still subordinate to the respective government official may be reluctant to execute extra-constitutional decisions, and for two reasons: it violates social norms and it might bear unforeseeable and hence hardly calculable personal consequences. In brief, orders to display extra-constitutional activities cause potential loyalty conflicts which at least some of the bureaucrats might solve by effectively undermining the respective order's intention.

Borrowing from Bénabou and Tirole (2011), we may indeed conceive of these leading bureaucrats as the *alter ego* of the respective government official himself. Given, then, that the strictly rational ego finds transgressing against the constitution to be the optimal choice, the rule-following *alter ego* might simply shy away from actually taking the implied illegal action.

In order to capture these aspects, we define a probability  $\epsilon$  that a decision  $d_i$  of a government official to transgress against the constitution will become effective. We assume this probability to be common knowledge. Hence, if  $\tilde{h}$  is the share of government officials that choose  $d_i$  in principle, we have  $h = \epsilon^{\tilde{h}N} \tilde{h}$  as the share of government officials that *effectively* choose  $d_i$ , where effectively means that the respective government official's decision will indeed be executed and become effective. We assume the personal consequences of an individual decision  $d_i$  by a government official to materialize independently of whether this individual decision becomes effective or not. By contrast, a decision  $d_i$  contributes to an autocratic transition process only with probability  $\epsilon$  that it becomes effective.

Having chosen  $c_G$  leaves the government leader  $G$  with an expected payoff of  $\omega \in (0, 1)$ , which is his reelection probability under democratic rule. By contrast, upon having chosen  $d_G$ , he will be impeached unless at least a share  $h^c$  of the government officials  $O_i$  effectively follows him in choosing  $d_i$ . Should that happen, however, the public is called for an election. As described above, the government leader may be outvoted and effectively forced to step down. That happens with probability  $1 - \sigma$ , so that his reelection probability in its broadest sense is  $\sigma$  if he transgresses against the constitution, but  $\omega$  if he

abides to its rules. As  $O_i$  are not elected politicians but rather appointed civil servants, they will not necessarily be fired but rather tried for extra-constitutional activity and then dismissed with probability  $1 - \pi$ .

In the reelection case, the autocratic-transition process is completed and the constitution loses all of its hitherto existing binding character, if any. No government activity will henceforth be restricted by constitutional rules of the game. In the limited scope of the transition game, we assume  $G$  to be capable of credibly committing to an announcement according to which he equally shares the autocratic rent with those  $\tilde{h}N$  government officials that had participated in the autocratic transition. That would imply payoffs  $U_{G,i} = s^h := \frac{R}{hN+1} + 1 \geq s \ \forall i$ . By contrast, we will endogenize the distribution of the autocratic rent  $R$  in the full model presented in the subsequent section.

Finally, should more than  $(1 - h^c)N$  government officials  $O_i$  fail to effectively follow  $G$  in acting extra-constitutionally, the autocratic-transition attempt fails,  $G$  will be impeached and left with payoff zero. Those government officials that followed  $G$  will again be tried for extra-constitutional activity and fired with probability  $1 - \pi$ . By contrast, those government officials that had refrained from extra-constitutional activity will stay in office with payoff  $U_i = 1$ .

The timing of the transition game is as follows:

1. Government leader  $G$  chooses among actions  $\{c_G, d_G\}$ . Should  $G$  choose  $c_G$ , he will be reelected with probability  $\omega$ , the government officials remain in their respective position, and the game ends with payoffs  $U_G = \omega$  and  $U_i = 1 \ \forall i$ .
2. Upon having observed choice  $d_G$  by  $G$ , players  $O_i$  choose among actions  $\{c_i, d_i\}$ .
3. If  $h < h^c$ , the autocratic-transition attempt fails.  $G$  will be deposed and the government officials  $O_i$  that chose  $d_i$  will lose their position with probability  $1 - \pi$ . The game ends with payoffs  $U_G = 0$ ,  $U_i = \pi$ ,

and  $U_{j \neq i} = 1$ .

4. If  $h \geq h^c$ , a public election is called for.  $G$  will win with probability  $\sigma$  and effectively lose with  $1 - \sigma$ .
5. If  $G$  loses the election, the game ends with  $G$  being fired for sure, all  $O_i$  that chose  $d_i$  will be fired with probability  $1 - \pi$ , and all  $O_{j \neq i}$  that chose  $c_j$  will stay in office, implying payoffs  $U_G = 0$ ,  $U_i = \pi \ \forall i$  and  $U_j = 1 \ \forall j \neq i$ .
6. If  $G$  wins the election,  $G$  equally shares the autocratic rent  $R$  with all  $O_i$  that chose  $d_i$ , implying  $U_{G,i} = s^h$ . By contrast, those  $O_j \neq i$  that effectively chose  $c_j$  will have  $U_j = 1$ .

This gives us the following expected payoff functions for  $G$  as well as for each  $O_i$ :

$$U_G = \begin{cases} \omega & \text{for } c_G; \\ 0 & \text{for } d_G \ \wedge \ h < h^c; \\ \sigma s^h & \text{for } d_G \ \wedge \ h \geq h^c; \end{cases} \quad (1)$$

and:

$$U_i = \begin{cases} 1 & \text{for } c_i; \\ \pi & \text{for } d_{G,i} \ \wedge \ h < h^c; \\ \sigma s^h + (1 - \sigma)\pi & \text{for } d_{G,i} \ \wedge \ h \geq h^c. \end{cases} \quad (2)$$

This leads to:

**Proposition 1.** *There are either one or two Nash-equilibria in pure strategies  $\zeta_{i,G}$  with strategy profile:*

$$\zeta_{i,G} = \begin{cases} d_{i,G} & \text{if } s > s^c := \max\left(\frac{1 - (1 - \epsilon^{h^c N})\pi}{\epsilon^{h^c N} \sigma}; \frac{\omega}{\epsilon^{h^c N} \sigma}\right); \\ c_{i,G}. \end{cases}$$

*Proof: see appendix B ■*

We refer to  $s^c$  as the critical value of the autocratic multiplier  $s$  and hence distinguish two cases  $s \leq s^c$  and  $s > s^c$ .

### **Case $s \leq s^c$ : Self-enforcing Democracy**

There is only one Nash-equilibrium  $\{c_G, c_i\}$  in this case. All government actors abstain from extra-constitutional activity and accept the constitutional rules of the game. We refer to this case as *self-enforcing democracy*. For any given autocratic multiplier  $s$ , the case of self-enforcing democracy is the more likely the higher is its critical value  $s^c$ . From proposition 1, we get the following partial derivatives of  $s^c$ :

$$s^c(\pi, \epsilon, \sigma) < 0; s^c(\omega, h^c, N) > 0. \quad (3)$$

The partial derivatives determine the influence of key institutional parameters on the likelihood of a democratic system to be self-enforcing. We will return to the results further below.

### **Case $s > s^c$ : Tenuous Democracy**

There are two Nash-equilibria  $\{c_i, c_G; d_i, d_G\}$  in this case. There is hence a problem of equilibrium selection (see Harsanyi and Selten 1988). The government leader and the government officials might still coordinate on an equilibrium  $(c_i, c_d)$ , but they might as well coordinate on an equilibrium  $(d_G, d_i)$ . Whether they would want to do the former or the latter cannot unambiguously be determined.<sup>4</sup> Since the government actors might coordinate on extra-constitutional action  $(d_i, d_G)$  and hence initiate a process of autocratic transition, we refer to a democracy under the conditions of case  $s > s^c$  as *tenuous democracy*. The degree to which a tenuous democracy is indeed vulnerable to attempts of autocratic transition depends on the same institutional characteristics that determine whether a democracy is self enforcing or not. This can be demonstrated as follows.

---

4. For a detailed discussion, see van Damme (2002).

It is straightforward that if  $(d_{i,G})$  is a Nash-equilibrium, then it is also the payoff-dominant Nash-equilibrium. The reason is that the conditions are the same for both in our model, namely that the expected payoff of  $(d_{i,G})$  exceeds unity. By contrast, the Nash-equilibrium  $(c_{i,G})$  may be risk dominant. Define, for that matter,  $E(d_i|\delta)(E(c_i|\delta))$  as the expected payoff for any  $O_i$  from playing  $d_i(c_i)$ , conditional on a probability  $\delta$  of some  $O_{j \neq i}$  to play  $d_j$ . Then we have  $E(d_i|\delta) = \delta [\epsilon^{h^c N} \sigma s + (1 - \epsilon^{h^c N}) \pi] + (1 - \delta) \pi$  and  $E(c_i|\delta) = 1$ . In the case of only two players  $O_i$  and  $O_j$ , the probability  $\delta_r$  that equalizes  $E(d_i|\delta)$  and  $E(c_i|\delta)$  and hence makes  $O_i$  indifferent between  $c_i$  and  $d_i$  is the risk factor of equilibrium  $(d_{i,G})$ . In our case of  $N$  players and a critical mass  $h^c N$ , the risk factor is  $\delta_r^{h^c N}$ . In equilibrium  $(d_{i,G})$ , we hence find:

$$\delta_r = \left[ \frac{1 - \pi}{\epsilon^{h^c N} (\sigma s - \pi)} \right]^{\frac{1}{h^c N}} \quad (4)$$

By Harsanyi and Selten's (1988, 216) definition, Nash-equilibrium  $(c_{i,G})$  is risk dominant if  $\delta_r^{h^c N} > 0.5$ . Note, however, that  $(d_{i,G})$  remains the payoff-dominant equilibrium independently of whether  $(c_{i,G})$  is risk dominant, and that rational actors do not necessarily coordinate on the risk-dominant equilibrium. On the other hand, we know from experiments and other empirical observations that real-world actors are the more likely to choose the risk-dominant strategy the higher is the risk factor  $\delta_r^{h^c N}$ .<sup>5</sup> We are safe to assume, hence, that a tenuous democracy is the more vulnerable to attempts of autocratic transition the lower is the risk factor. As a result, the partial derivatives of  $\delta_r$  in equation 4 give us the institutional determinants of the degree of vulnerability:

$$\delta'_r(\pi, \epsilon, s, \sigma) < 0; \delta'_r(h^c, N) > 0. \quad (5)$$

The implications of these partial derivatives are fully in line with those that determine the case of self-enforcing democracy, as given by condition 3. For

---

5. See, inter alia, Cooper et al. (1992), Straub (1995), Clark, Kay, and Sefton (2001) and Février and Linnemer (2006, 170).



both cases, taken together, we hence find: A democracy is less likely to be self-enforcing and, if not, the government actors are more likely to coordinate on an extra-constitutional action, when

- there is a high probability  $\sigma$  that the government leader will survive the autocratic transition attempt in office, which is the more likely the more popular he is;
- there is a high autocratic multiplier  $s = \frac{R}{N} + 1$  of the reservation utility, and hence autocratic rents  $R$  are high;
- there is a low reelection probability  $\omega$  for  $G$  under constitutional rule;
- there is a high probability  $\pi$  that a government official will remain in his position after having acted extra-constitutionally even in the case of a failed transition attempt; and
- there is not much separation of powers, both formally and effectively, as indicated by  $N$  and  $h^c$ , respectively.

The final criterion is of particular importance since the probability of government officials to coordinate on extra-constitutional action quickly drops in  $h^c N$  for any given  $\epsilon$ . Note that  $N$  indicates the number of further government officials, while  $h^c$  gives the critical share of government officials that need to be capable of effectively supporting the government leader. If  $h^c$  is low, then regardless of how many government officials there are only a few of them—say a single army commander plus a single commander of the police—would really count. Democracies that exhibit little separation of powers are less likely to be self enforcing; and in tenuous democracies, little separation of powers makes coordination on the payoff-dominant equilibrium  $d_{i,G}$  and hence autocratic transition more likely.

Note, however, that our results are preliminary since we operate within a model setting of limited scope. In particular, the government leader is assumed to be capable of committing to his promise of equally sharing the autocratic rent with the government officials. We will argue below that this is far from realistic. In the following section, we relax this assumption.

## 4 The Full Game

In the full game, we endogenize the potential struggle for autocratic rents  $R$  among the government officials under the conditions of an abolished constitution. To that end, we add another set of subgames to the transition game as it was developed in the previous section. We aggregate this additional set of games to the *post-transition game*. As demonstrated in the previous section, the government leader's payoff will be  $s^h$  if he shares the autocratic-government rents with the government officials. But, if he fails to share and leaves the government officials with their initial payoff of  $U_i = 1$ , the government leader's payoff will instead be  $S := R + 1 > s^h$ . Hence, an announcement that the government leader intends to share the autocratic rents with the government officials lacks credibility. Indeed, in the environment of a freshly abolished constitution, it is not clear in the first place who would have the power to allocate government rents.

One might naturally think that this power would go to the government leader himself, since, after all, he is now the dictator. But the power of a dictator does not fall like manna from heaven. Rather, it rests with his capacity to play off the government officials against each other, particularly those who administer government decisions (Bueno de Mesquita et al. 2005; Svobik 2012; Tullock 1987). It is therefore crucial to add the post-transition game to the transition game, since rational actors will build expectations about their respective future positions in the post-transition struggle for autocratic rents when making their choices in the transition period.

In order to consider the post-transition interactions, we add the following steps to our game. Upon a successful transition, the government leader may decide to equally share the autocratic rents with the further government officials, thus raising the payoff of each of his followers along with his own payoff to a level of  $s^h > 1$ . Alternatively, he may as well renege on his promise, leaving all further government officials with their initial payoff  $U_i = 1$  and raising his own payoff to  $S = R + 1 > s^h$ .

If the government leader reneges on his promise, those  $\tilde{h}N$  government officials that chose  $d_i$  in the transition game may withdraw their support for  $G$  conditional on their expected capability to coordinate on rebellious action. We define  $\tilde{\gamma} \in [0, 1]$  as the share of  $\tilde{h}N$  government officials that withdraw their support of  $G$  in the post-transition game. As in the transition game, we assume limited control of the government officials over their respective public authorities or, for that matter, over their *alter ego*. We assume the probability of a decision  $d_i$  by an individual  $O_i$  to withdraw his support of the government leader to be executed with probability  $\rho$ . Then,  $\gamma = \rho^{\tilde{h}N} \tilde{\gamma}$  is the share of government officials that effectively rebel against the government leader.

Define now  $g := \gamma h$  as the share of government officials that first effectively follow the government leader by choosing extra-constitutional action in the transition game *and* then effectively withdraw their support for the government leader in the post-transition game. In the post-transition game, the government leader will be deposed and substituted by some individual from outside if he reneges on his sharing obligation *and* if  $\gamma$  reaches at least some critical value  $\gamma^c \in (0, 1)$ . In the latter case, his successor will be forced to share all rents  $R$  equally with the  $\tilde{g}N$  government officials, where  $\tilde{g} := \tilde{h}\tilde{\gamma}$ .

By contrast, should the government officials fail to mobilize a share  $\gamma \geq \gamma^c$  against the cheating government leader, then the leader will stay in office along with those government officials that continued to support him. The disobedient government officials, in turn, will be removed from office and be replaced by newly appointed individuals from outside. The disobedient officials will be left with payoff zero, while  $G$  continues to claim all autocratic rents and hence a payoff  $S = R + 1$ , so that all remaining government officials, along with the newly appointed ones, will be left with payoff  $U_i = 1$ .

The time line of the full game is as follows:

1. Government leader  $G$  chooses among actions  $\{c_G, d_G\}$ . Should  $G$  choose  $c_G$ , he will be reelected with probability  $\omega$ , the government officials

remain in their respective positions, and the game ends with payoffs  $U_G = \omega$  and  $U_i = 1 \ \forall i$ .

2. Upon having observed choice  $d_G$  by  $G$ , players  $O_i$  choose among actions  $\{c_i, d_i\}$ .
3. If  $h < h^c$ , the autocratic-transition attempt fails.  $G$  will be deposed and those further government officials  $O_i$  that chose  $d_i$  will lose their position with probability  $1 - \pi$ . The game ends with payoffs  $U_G = 0$ ,  $U_i = \pi$ , and  $U_{j \neq i} = 1$ .
4. If  $h \geq h^c$ , a public election is called for.  $G$  will win with probability  $\sigma$  and effectively lose with  $1 - \sigma$ .
5. If  $G$  loses the election, the game ends with  $G$  being fired for sure, all  $O_i$  that chose  $d_i$  will be fired with probability  $1 - \pi$ , and all  $O_{j \neq i}$  that chose  $c_j$  will stay in office, implying payoffs  $U_G = 0$ ,  $U_i = \pi \ \forall i$  and  $U_j = 1 \ \forall j \neq i$ .
6. Upon having won the election,  $G$  chooses among actions  $\{c_G, d_G\}$ ;  $c_G$  (comply) implies meeting his promise to equally share the autocratic rent with the  $hN$  officials that followed him in the transition. By contrast,  $d_G$  (defect) implies seizing all autocratic rents  $R$  so as to reap a payoff  $S = R + 1$  and leave the government officials with their initial payoff  $U_i = 1$ .
7. In the case of  $G$  choosing  $c_G$ , the game ends with payoffs  $U_G = U_i = s^h$  and  $U_j = 1$  with  $j \neq i$  indicating those government officials  $O_j$  that had failed to follow  $G$  in the transition game.
8. Upon having observed  $G$  to choose  $d_G$ , by contrast, each government official chooses among actions  $\{c_i, d_i\}$ , where  $c_i$  implies continuing support of  $G$  while  $d_i$  implies withdrawing support of  $G$ .
9. In the case of  $\gamma \geq \gamma^c$ ,  $G$  will be deposed and substituted by an outsider  $G'$ . The game ends with payoff  $U_G = 0$  as well as  $U_{G',i} = s^g := \frac{R}{h\tilde{\gamma}N+1} + 1 \geq s$ .

10. In the case of  $\gamma \in (0, \gamma^c)$ , all rebelling government officials will be replaced by outsiders. The game ends with payoff  $U_G = S$ ,  $U_i = 0$  for all rebelling officials and  $U_j = 1$  for all obedient government officials and all newly appointed government officials.

Note that steps one to five of this time-line are identical to those of the limited transition game. Each player  $i$  and  $G$  needs to decide over a strategy  $\zeta_{i,G}$  that consists of a sequence  $k \in \{1, 2\}$  of actions  $a_{i,G}^k \in \{c_{i,G}, d_{i,G}\}$ . For the government leader, the action profile is  $a_G \in \{c_G, dc_G, dd_G\}$ , while it is  $a_i \in \{c_i, dd_i, dc_i\}$  for each further government official  $O_i$ . The payoffs are as follows:

$$U_G = \begin{cases} \omega & \text{for } c_G; \\ 0 & \text{for } d_G \wedge h < h^c; \\ \sigma s^g & \text{for } dc_G \wedge h \geq h^c; \\ \sigma S & \text{for } dd_G \wedge h \geq h^c \wedge g < g_c; \\ 0 & \text{for } dd_G \wedge h \geq h^c \wedge g \geq g_c. \end{cases} \quad (6)$$

$$U_i = \begin{cases} 1 & \text{for } c_i; \\ \pi & \text{for } d_i \wedge h < h^c; \\ \sigma s^g + (1 - \sigma)\pi & \text{for } d_i \wedge h \geq h^c \wedge dc_G; \\ \sigma s^g + (1 - \sigma)\pi & \text{for } dd_i \wedge g \geq g^c; \\ \sigma + (1 - \sigma)\pi & \text{for } dc_i \wedge h \geq h^c \wedge dd_G; \\ (1 - \sigma)\pi & \text{for } dd_i \wedge h \geq h^c \wedge \gamma < \gamma^c \wedge dd_G. \end{cases} \quad (7)$$

This leads to:

**Proposition 2.** *There are either one or two Nash-equilibria in pure strategies*

$\zeta_{i,G}$  with strategy profile:

$$\zeta_{i,G} = \begin{cases} dd_i, dc_G & \text{if } s > s^c := \max(\frac{1-(1-\epsilon^{h^c N})\pi}{\epsilon^{h^c N}\rho\gamma^{c h^c N}\sigma}; \frac{\omega}{\epsilon^{h^c N}\sigma}); \\ c_{i,G}. \end{cases}$$

*Proof:* see appendix B ■

We have again two cases  $s \leq s^c$  and  $s > s^c$ .

### Case $s \leq s^c$ : Self-enforcing Democracy

There is again only one Nash-equilibrium  $\{c_i, c_G\}$  in this case. All government actors abstain from extra-constitutional activity and accept the constitutional rules of the game. The institutional determinants of a realization of this case follow from the partial derivatives of  $s^c$  in proposition 2:

$$s^{c'}(\pi, \epsilon, \sigma, \rho) < 0; s^{c'}(\omega, h^c, \gamma^c, N) > 0. \quad (8)$$

### Case $s > s^c$ : Tenuous Democracy

There are again two Nash-equilibria in this case. Similar to the transition game, we find that if  $(dc_G; dd_i)$  is a Nash-equilibrium, it is also the payoff-dominant equilibrium. To determine the risk factor, we define  $E(dd_i|\delta^{\gamma^{c h^c N}})$  ( $E(c_i|\delta^{\gamma^{c h^c N}})$ ) to be the expected payoff of government official  $O_i$  from choosing  $dd_i$  ( $c_i$ ), conditional on the probability  $\delta^{\gamma^{c h^c N}}$  of a critical mass  $\gamma^{c h^c N}$  to choose  $d_i$  in both the transition game and the post-transition game. Then, the risk factor is the probability  $\delta_r^{\gamma^{c h^c N}}$  for which  $E(dd_i|\delta^{\gamma^{c h^c N}}) = E(c_i|\delta^{\gamma^{c h^c N}})$ , which is given by:

$$\delta_r^{\gamma^{c h^c N}} [\epsilon^{h^c N} \sigma \rho \gamma^{c h^c N} s + (1 - \epsilon^{h^c N} \pi)] + (1 - \delta_r^{\gamma^{c h^c N}}) \pi = 1. \quad (9)$$

Solving for  $\delta_r$  yields:

$$\delta_r = \left[ \frac{1 - \pi}{\epsilon^{h^c N} (\sigma \rho s - \pi)} \right]^{\frac{1}{\gamma^{c h^c N}}} \quad (10)$$

Remember that the likelihood of the government actors to coordinate on an equilibrium  $\{dd_i; dc_G\}$  drops in a rise of the risk factor  $\delta_r^{\gamma^c h^c N}$ . The partial derivatives of equation 10 hence determine the likelihood of autocratic transition. They are as follows:

$$\delta_r'(\pi, \epsilon, s, \sigma, \rho) < 0, \quad \delta_r'(h^c, \gamma^c, N) > 0. \quad (11)$$

The partial derivatives are again fully in line with those that determine the case  $s \leq s^c$  of a self-enforcing democracy, as given by condition 8. The main result of the transition game thus carries over to the full game. The differences in the result of the full game as compared to the transition game are determined by the following variables: first the critical share  $\gamma^c$  of government officials that need to effectively choose to rebel against the government leader in the post-transition period; and second the probability  $\rho$  of an individual government official to have a decision  $d_i$  in the post-transition period executed by their respective public authority (or by their *alter ego*, for that matter). These differences do not turn the results of the model upside down. But they are meaningful in another another sense: The full game encompasses two significant collective-action problems that the actors face and that hence pinpoint the core difference between a democracy that is self-enforcing and one that is not. The difference arises because the government officials face problems in coordinating *twice* in the full model: first in the transition game where they need to coordinate on extra-constitutional activity, and second in the post-transition game where they need to potentially coordinate in order to credibly threaten to depose the newly established dictator for failing to share the autocratic rents.

What appears as a negligible modification of the conditions in the model may turn out to be decisive in reality. A potential autocratic government leader that signals his intention to switch to extra-constitutional activity is reliant on his cooperation with the government officials. They, in turn, have a coordination problem, since failing to unite to follow the leader may have painful consequences for each of them. Additionally, however, they need

to trust in the government leader's promise to share the autocratic rents, although they do not have any natural reason to do so. Rather, they need to trust in their own capability to keep the government leader in check, which requires that they must solve a second and no less severe collective-action problem. Short of that, it is a better choice for each of the government officials to stick to constitution-abiding behavior.

In order to better grasp the intuition behind these results, we use the definition of the threshold value  $s^c$  of the government officials from proposition 2:

$$s^c := \max\left(\frac{1 - (1 - \epsilon^{h^c N})\pi}{\epsilon^{h^c N} \rho^{\gamma^c h^c N} \sigma}; \frac{\omega}{\epsilon^{h^c N} \sigma}\right). \quad (12)$$

The left-hand term in the brackets represents a necessary condition for the government officials to coordinate on extra-constitutional action. The right-hand term in the brackets gives a necessary condition for the government leader to switch to extra-constitutional action as a first-mover. Note that the latter requires the government leader to expect the government officials to follow suit, so that he must always see both conditions satisfied for even considering extra-constitutional action. The condition that is numerically higher is thus the binding one. Solving the left-hand term and the right-hand term of condition 12 for  $\epsilon$  yields:

$$\epsilon = \left[ \frac{1 - \pi}{s^c \sigma \rho^{\gamma^c h^c N} - \pi} \right]^{\frac{1}{h^c N}}; \quad \epsilon = \left[ \frac{\omega}{s^c \sigma} \right]^{\frac{1}{h^c N}}, \quad (13)$$

respectively. Each of the two equations represents combinations of probabilities  $\rho$  and  $\epsilon$  that form a threshold. If a given combination of  $\rho$  and  $\epsilon$  is beyond both thresholds, then the constitution ceases to be self-enforcing.

The region to the north-east of the downward-sloping solid line in figure 2 represents combinations of  $\epsilon$  and  $\rho$  that satisfy the left-hand condition of condition 13. The region above the horizontal solid line represents levels of  $\epsilon$  that satisfy the right-hand condition of condition 13. If one of these



conditions does not hold, then there is only one Nash-equilibrium  $(d_{i,G})$ , and that applies to the region to the south-west of the bold and solid line. Hence, any point to the south-west of that line represents a self-enforcing democracy.

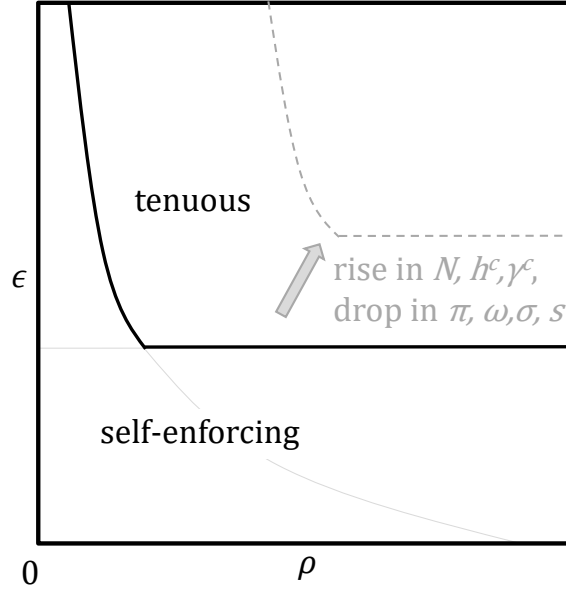


Figure 2: Stability Regions

By contrast, any point that lies both to the north-east of the downward-sloping solid line *and* to the north of the horizontal solid line represents combinations of  $\epsilon$  and  $\rho$  that satisfy both conditions in equation 12. Hence, we have a tenuous democracy in this region of the graph. Note that this does not mean that the government actors will in any case coordinate on extra-constitutional action and thus launch a process of autocratic transition with certainty. What it does mean, however, is that they face two Nash-equilibria rather than one as soon as they cross the bold solid line. Whenever that happens, each government actor needs to evaluate the likelihood of each equilibrium to be established prior to his individual choice between obeying or breaking the rules of the constitution. A thus established environment makes the constitution vulnerable to autocratic transition, since it might happen that a government leader attempts a first step into an autocratic-transition process, and it might then happen that the government officials

coordinate on following the leader in his attempt. As a consequence, the solid line in Figure 2 can be viewed as some sort of a constitutional firewall. Any lack in a second equilibrium, as it is illustrated in the lower left area, makes that firewall effective and the democracy self-enforcing.

Some comparative statics can immediately be derived. According to equation 13, it is the following parameters that shift the solid line in Figure 2 outwards: the probability  $\sigma$  of the government leader to win the terminal election; the probability  $\pi$  of the government officials to remain in office in the case of a failed autocratic-transition attempt; and the autocratic multiplier  $s$  of the government official's reservation utility. By contrast, the following parameters shift the line inwards: the probability  $\omega$  of the government leader to win an election under democratic rule; the number  $N$  of government officials; and the threshold shares  $h^c$  as well as  $\gamma^c$  necessary for the government officials to coordinate in both the transition game and the post-transition game. Note that an inward shift of the solid line in Figure 2 shrinks the “self-enforcing” area and enhances the “tenuous” area.

The probability of autocratic transition to actually happen in the “tenuous” region is lowest at points close to the threshold line. It rises when we move from the threshold line toward the north-east of the diagram and hence when the probabilities  $\epsilon$  and  $\rho$  rise. The latter indicates a rise in the capability of the individual government officials to have both of the following decisions executed by their respective authorities—or accepted by their *alter ego*: a decision to break the constitutional rules in the transition period ( $\epsilon$ ); a decision to (potentially) rebel against the government leader in the post-transition game ( $\rho$ ).

## 5 Discussion

While none of the parameters that determine the likelihood of autocratic transition may appear particularly surprising in a mere technical sense, the parameters  $\sigma$ ,  $\omega$ ,  $\pi$ , as well as  $N$  and  $h^c, \gamma^c$  deserve some closer inspec-

tion:

- The ratio  $\frac{\sigma}{\omega}$  represents the relation between the probabilities of either winning an election under the conditions of autocratic transition or under those of democratic rule. It is hence an indication of how popular an autocratic or populist policy is in relation to a policy that is strictly grounded in the respect for democratic constitutional rules.
- One of the most important determinants of the probability  $\pi$  is corruption. In an environment of due process, government officials that had seriously transgressed against the constitutional rules will have to be sued and typically found personally unsuitable for their government position. Corruption is among the most important diluting forces of this important principle and precisely so because corruption raises the probability that government officials will remain in office despite having transgressed against constitutional or legal rules.
- The number  $N$  of government officials is a proxy for the formal degree of the separation of powers, and the threshold values  $h^c$  and  $\gamma^c$  are proxies for the effective degree of the separation of powers. The higher both the formal and effective degrees of the separation of powers are on the horizontal and—in federations—on the vertical level, the more difficult it becomes for the government officials to coordinate with each other.

The ratio  $\frac{\sigma}{\omega}$  indicates that it is not necessarily the personal popularity of some individual politician that counts. What rather counts is the policy such a politician pursues. Should the public prefer policies strictly based on democratic rule, then  $\frac{\sigma}{\omega}$  should be low. By contrast, should the public call for some “strong” leader, possibly in times perceived as particularly threatening to the nation, then  $\frac{\sigma}{\omega}$  should rise. Modern populists typically gain popularity by referring to loftier goals which they—and only they—are not willing to sacrifice for the sake of some “fussy” statutory stipulations in the constitution; or they gain popularity by pursuing an allegedly desperate battle against external conspirators who exploit constitutional rules for the sake of

hollowing out the cohesive bonds of the nation<sup>6</sup>. However, for such claims to resonate, specific institutional or political backgrounds are required.

Russia in the late 1990s may be an example. Vladimir Putin became President in 1999 following a rather chaotic period under Putin's predecessor. The privatization of former Soviet companies quickly spiraled out of control and led into an oligarchic structure that skyrocketed the wealth of a few but left the masses with virtually nothing. Corruption, crime, and Mafia structures flourished in parallel to rising poverty and unemployment. The face behind this undesirable development was that of then President Boris Yeltsin, whose character as *bon vivant*, and increasingly so as an alcoholic became symbols of both his personal incapability and the chaos of the country he led. And, becoming more obvious was the deeper problem behind the transition from a former superpower to a socially dissolving country shaken by corruption, financial distress, crime, economic stagnation and inequality: the surrendering formerly honorable values to the primary antagonist of the now-perished glorious empire, namely to Western democracy.

Putin was the man to help right the ship, but he quickly made clear that his help would not be grounded in Western democratic rules. What is more, as the oligarchs, the corruption, the financial chaos and the lost empire came ostensibly out of Western democracy, a return to the *status quo ante* appeared almost natural to many. Indeed, Putin was able to curb the symptoms of many issues related to the allegedly decadent Western system: He stopped the disintegration of the Russian federation, he arrested prominent oligarchs, and he brought at least central parts of the economy back under his control. And he did all this with measures decidedly different from what Western consultants and representatives of liberal democracy repeatedly recommended. It was this policy that made him extremely popular with a large part, if not most, of the Russian population. This is what a high ratio for  $\frac{\sigma}{\omega}$  catches in formal terms. A high  $\frac{\sigma}{\omega}$  was hence an important precondition for Putin's path to autocratic transition, and it was high due to the particular unfortunate developments that Russia underwent during the 1990s under Boris

---

6. For an extensive discussion, see Mounk (2018).

Yeltsin.

That alone, however, was not sufficient for Russia; and it is generally not a sufficient condition for any country to embark on such a path. Rather, what is additionally needed is a high degree of mutual trust of the government elites regarding their ability to coordinate on a particular equilibrium; this factor relates to the degree of the separation of powers. Note that the separation of powers is important for two reasons. First, it exacerbates collusion of the government officials against the constitution in the transition game; second, it aggravates the collective-action problem that the government officials face if the government leader fails to share the autocratic rents. Hence, while a populist policy might be capable of safeguarding public support on the path to autocracy, such a path might still remain impassable for the government officials if the degree of the separation of powers is too high.

Choosing to transgress or not to transgress against the constitution implies a far-reaching collective-choice problem for the government officials. They must choose between two competing provisions that could protect their wealth: one provision stems from the existing set of constitutional rules, at least as long as these rules remain self-enforcing; the other that stems from the presumed collective capability of the government officials to keep a future dictator in check by permanently threatening a coup in an environment lacking constitutional rules. If a sufficiently high share of the government officials expects that a sufficiently high share of their respective colleagues will not cooperate in both stages of the game, then individually keeping clear of the autocracy path is each government official's best choice. At the collective level of all government officials, then, this implies that the entire government sector will remain within the boundaries of the existing democratic constitution. By contrast, should at least a share  $g^c$  of the government officials be expected to effectively cooperate on both stages, then it is individually rational for each individual government official to embark on the autocratic path. The constitution will cease to be self-enforcing.

A truly fateful historical example of how a collective choice beyond the limits of self-enforcing democracy can go wrong is that of the highest-ranking

officials of the German army (Wehrmacht) in the early 1930s. Feeling deeply degraded and undervalued under the conditions of the first German democracy following the disastrous World War I, the army officials trusted in their own apparent capability to deliberately exploit the political movement of a devilishly charismatic leader whom they did not like either. Although they disrespected Hitler because of his lack in military rank and nobility, they nevertheless pursued a hands-off approach regarding the Nazi's path to autocracy following Hitler's appointment as chancellor of the Weimar Republic by president Hindenburg on January 30, 1933. The army officials' aim was clear: Let Hitler do the dirty work of destroying the hated Weimar Constitution, eventually removing the Nazis in general, and Hitler in particular, from office, and then installing themselves to allegedly rescue the political culture of the German nation. In that way, they hoped for the resurrection of the pre-WW I German empire; or at least some successor that was adequate according to their elitist views.

But the army officials dramatically underestimated Hitler's capability to undermine their capacity to coordinate on coup activities following the destruction of the Weimar democracy. Para-military organizations like "SS" and "SA" under the firm control of the Nazi party enhanced the complexity of the newly established system of security forces. In combination with a set of additional measures, that split the security forces into numerous competing groups and subgroups. The result was a grossly reduced trust among a hitherto homogeneous elite of army commanders that were once bound by family ties and nobility membership.

After Hindenburg's death, Hitler urged the army commanders to transfer their vow of fidelity from the deceased president to Hitler himself. The binding power of that vow grossly added to the commanders' difficulties in conspiring against Hitler within the complexity of competing security forces, eventually resulting in the failure of each further coup event all the way up to the end of WW II. It is safe to say that the army officials would have likely been more reluctant to exploit Hitler for the sake of destroying the Weimar Republic if they realized they would not collectively be able to keep

the dictator in check.

The most important normative implication of our analysis is this: Constitutional choice should pay particular attention to aspects like the formal and effective disentanglement of government authorities (separation of powers) as well as to federalism (Figueiredo Jr and Weingast 2005); competition among the different branches of government and among different federal levels; independent rules for appointing heads of the executive branch, the members of parliament and, most importantly, the judges of high-ranking courts. Failure to consider these factors may have been the most important driver of autocratic transition in the aftermath of the fourth wave of democratization. As such, many of the newly established democratic constitutions did not survive attacks by leaders like Lukashenko in Belarus, Nazarbayev in Kazakhstan, Karimov in Uzbekistan or, more recently, Putin in Russia.

By contrast, the constitution of the United States, with its deeply rooted separation of powers in various dimensions, has survived numerous attacks from leaders that were certainly not democrats but that did not even dare to openly confess their reservations against the pluralist democratic system; this applies to even very recent developments. In a similar fashion, Hungary and Poland may have a much better chance of surviving as democracies despite the strong attacks by their government leaders and by further decisive politicians because these countries are members of the EU, which enhances the complexity of their separation of powers. This is admittedly somewhat speculative, but our analysis would at least support this supposition.

As the most important empirical implication of our analysis, we have identified three main characteristics that either make a democratic constitution vulnerable to autocratic transition or transition proof. These are both the formal and effective degrees of the separation of powers, the level of corruption, and the popularity of transgressions against constitutional rules by (populist) government leaders. The one characteristic that can best be influenced in constitutional choice is the degree of the separation of powers. While this is indeed not a new aspect, it has possibly attracted less attention as a fundamental characteristic of sustainable democracies than it deserves. The

principle *divide et impera* was meant to protect autocratic leaders against competitors that strive to attack the leader’s regency. Its capacity to protect democracy against attacks on their constitutions, however, appears to rest on precisely the same mechanism.

A normative implication of our model is that constitutions matter. In this regard, it departs from Przeworski (2005, 2006). To him, a “constitution is neither sufficient nor necessary for democracy to survive” (Przeworski 2005, 267). In this view, a constitution is not necessary because actors would agree to the constitutional rules if they were an equilibrium anyway; and it is not sufficient because actors would break the rules if they did not describe an equilibrium. This implication follows from a specific modeling in which democracy is a unique equilibrium, conditional on the relevant parameters. Given these parameters, none of the relevant actors could be better off in any feasible alternative. Depending on the respective parameter setting, then, democracy is either a unique equilibrium or no equilibrium at all, implying in the latter case that democracy is not sustainable.

By contrast, there are potentially two equilibria in our model, of which democracy is but one, and not necessarily the one that is most preferred by the government officials. It is the task of the constitutional rules to structure mutually enforcing control mechanisms of all government actors that leave room for only one equilibrium on which all actors shall eventually coordinate—the democratic equilibrium (Calvert 1995; Hardin 2013). In that sense, then, a constitution can indeed be capable of making democracy “the only game in town” (Przeworski 2006, 324). Hence, our approach defines the conditions under which democracy is self-enforcing as it structures the constitutional rules in a way as to allow for only one, namely the democratic equilibrium; by the same token, our approach defines the conditions under which democracy—if it exists—is not self-enforcing as it structures the constitutional rules in a way as to allow for two equilibria on each of which the government actors might coordinate.

Note that a deep and diverse separation of powers is key among the rules that exclude the non-democratic equilibrium; as such, we may refer to our model



as a post-Montesquieu approach. Since the depth and diversity of the separation of powers is obviously correlated with the level of per-capita income, our model provides an alternative explanation of Przeworski's empirical point of departure, namely the startling correlation between per-capita income and the probability of democracy to survive. However, while per-capita income is doubtlessly strongly correlated with the survival of democracy, it is still not causal in our model. It is rather correlated with those forces that are indeed causal for the survival of democracy.

## 6 Conclusions

We have developed a model of autocratic transition pursued by a sitting government leader. The aim was to identify the conditions under which autocratic transition is likely and, conversely, under which conditions democratic constitutions are effectively protected by mechanisms of self-enforcement. As a key element, each government official needs to expect coordination in two critical stages of the transition process in order to find their own participation worthwhile: Firstly, in an early period of transition, each government official must expect a sufficiently large share of the officials to participate in extra-constitutional activity. Second, the further government officials must, if necessary, expect a sufficient share of the government officials to participate in a coup against the newly established dictator.

Failure of this second condition to be satisfied leads to the non-credibility of the government leader's initial promise to share the autocratic rents with the government officials. Put in general terms, if the government officials fail to establish a credible and permanent collective coup threat, then they will not see themselves as protected against the exploitative power of the government leader in an autocracy.

As a result, we have identified three major testable empirical implications of our analysis. In particular, the likelihood of an autocratic transition is higher when (1) both the formal and the effective degree of the separation of powers

is low, (2) the degree of corruption is high, and (3) the government leader that transgresses against constitutional rules for the sake of some allegedly higher goal—like protecting the nation against external or internal enemies or creating some glorious empire or the like—is highly popular.

As a normative implication, we find that constitutional choice in young democracies should focus on measures that divide the different powers both personally and institutionally, including separating the procedures for appointing government officials in different branches of the government and—if possible—on different federal levels.

## References

- Bénabou, Roland, and Jean Tirole. 2011. “Identity, morals, and taboos: Beliefs as assets.” *The Quarterly Journal of Economics* 126 (2): 805–855.
- Benhabib, Jess, and Adam Przeworski. 2006. “The political economy of redistribution under democracy.” *Economic Theory* 29 (2): 271–290.
- Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2005. *The Logic of Political Survival*. The MIT Press.
- Calvert, Randall. 1995. “The rational choice theory of social institutions: Cooperation, coordination, and communication,” edited by Jeffrey S Banks and Eric A Hanushek, 216–268. Cambridge University Press Cambridge.
- Clark, Kenneth, Stephen Kay, and Martin Sefton. 2001. “When are Nash equilibria self-enforcing? An experimental analysis.” *International Journal of Game Theory* 29 (4): 495–515.
- Cooper, Russell, Douglas V DeJong, Robert Forsythe, and Thomas W Ross. 1992. “Selection criteria in coordination games: Some experimental results.” *The Quarterly Journal of Economics* 107 (2): 739–771.
- Fearon, James D. 2011. “Self-enforcing democracy.” *The Quarterly Journal of Economics* 126 (4): 1661–1708.

- Ferguson, Niall. 2001. *The cash nexus: Money and power in the modern world, 1700-2000*. New York: Basic Books.
- Février, Philippe, and Laurent Linnemer. 2006. "Equilibrium selection: Pay-off or risk dominance? The case of the "Weakest Link." *Journal of Economic Behavior and Organization* 60 (2): 164–181.
- Figueiredo Jr, Rui JP de, and Barry R Weingast. 2005. "Self-enforcing federalism." *Journal of Law, Economics, and Organization* 21 (1): 103–135.
- Granovetter, Mark. 1978. "Threshold models of collective behavior." *American Journal of Sociology*: 1420–1443.
- Hardin, Russell. 2006. "Constitutionalism." In *The Oxford handbook of political economy*, edited by Barry R Weingast and Donald A Wittman, 289–311.
- . 2013. "Why a constitution." In *The social and political foundations of constitutions*, edited by Dennis J Galligan and Mila Versteeg, 51–72. New York: Cambridge University Press.
- Harsanyi, John C, and Reinhard Selten. 1988. *A general theory of equilibrium selection in games*. Cambridge, Mass.: The MIT Press.
- Hyde, Susan D, and Nikolay Marinov. 2014. "Information and self-enforcing democracy: The role of international election observation." *International Organization* 68 (2): 329–359.
- Marwell, Gerald, and Pamela Oliver. 1993. *The critical mass in collective action*. Cambridge University Press.
- Mittal, Sonia, and Barry R Weingast. 2011. "Self-enforcing constitutions: with an application to democratic stability in America's first century." *The Journal of Law, Economics, & Organization* 29 (2): 278–302.
- Mounk, Yascha. 2018. *The People Vs. Democracy: Why Our Freedom is in Danger and how to Save it*. Cambridge, Mass.: Harvard University Press.

- Przeworski, Adam. 1991. *Democracy and the market: Political and economic reforms in Eastern Europe and Latin America*. Cambridge University Press.
- . 2005. “Democracy as an equilibrium.” *Public Choice* 123 (3): 253–273.
- . 2006. “Self-enforcing democracy,” edited by Barry R Weingast and Donald A Wittman, 312–28. Oxford University Press New York.
- Straub, Paul G. 1995. “Risk dominance and coordination failures in static games.” *The Quarterly Review of Economics and Finance* 35 (4): 339–363.
- Svolik, Milan. 2012. *The politics of authoritarian rule*. Cambridge: Cambridge University Press.
- Traversa, Federico. 2015. “Income and the stability of democracy: Pushing beyond the borders of logic to explain a strong correlation?” *Constitutional Political Economy* 26 (2): 121–136.
- Tullock, Gordon. 1987. *Autocracy*. Dordrecht: Kluwer.
- van Damme, Eric. 2002. “Strategic equilibrium,” edited by Robert Aumann and Sergiu Hart, 3:1521–1596. Amsterdam: North-Holland.
- Weingast, Barry R. 1997. “The political foundations of democracy and the rule of the law.” *American Political Science Review* 91 (2): 245–263.
- . 2005. “The constitutional dilemma of economic liberty.” *The Journal of Economic Perspectives* 19 (3): 89–108.

## A Appendix A

Country	Year of Highest Value	Highest Value	Value 2016	Change	Highest Status	Status 2016	Status steps
Gambia	1992	9.25	2.50	-6.75	free	not free	-2
Venezuela	1990	8.50	3.25	-5.25	free	not free	-2
Russia	1991	7.00	1.75	-5.25	partly free	not free	-1
CAR*	1993	6.25	1.00	-5.25	partly free	not free	-1
Thailand	1990	7.75	3.25	-4.50	free	not free	-2
Congo	1992	7.00	2.50	-4.50	partly free	not free	-1
Belarus	1992	6.25	1.75	-4.50	partly free	not free	-1
Mali	1996	8.50	4.75	-3.75	free	partly free	-1
Burundi	2005	5.50	1.75	-3.75	partly free	not free	-1
Ethiopia	1998	5.50	1.75	-3.75	partly free	not free	-1
Nicaragua	1998	7.75	4.75	-3.00	free	partly free	-1
Jordan	1992	7.00	4.00	-3.00	partly free	partly free	0
Kyrgyzstan	1992	7.00	4.00	-3.00	partly free	partly free	0
Gabon	1991	6.25	3.25	-3.00	partly free	not free	-1
Libya	2012	4.75	1.75	-3.00	partly free	not free	-1
Yemen	1993	4.75	1.75	-3.00	partly free	not free	-1
Eritrea	1995	4.00	1.00	-3.00	partly free	not free	-1
Hungary	2004	10.00	7.75	-2.25	free	free	0
Solomon Islands	1990	10.00	7.75	-2.25	free	free	0
Ecuador	1990	8.50	6.25	-2.25	free	partly free	-1
Bangladesh	1991	7.75	5.50	-2.25	free	partly free	-1
Honduras	1990	7.75	5.50	-2.25	free	partly free	-1
Zambia	1991	7.75	5.50	-2.25	free	partly free	-1
Turkey	1990	7.00	4.75	-2.25	partly free	partly free	0
Guinea-Bissau	1994	6.25	4.00	-2.25	partly free	partly free	0
Maldives	2009	6.25	4.00	-2.25	partly free	partly free	0
Algeria	1990	5.50	3.25	-2.25	partly free	not free	-1
Mauritania	2007	5.50	3.25	-2.25	partly free	not free	-1
Uganda	1996	5.50	3.25	-2.25	partly free	not free	-1
Kazakhstan	1991	4.75	2.50	-2.25	partly free	not free	-1
Azerbaijan	1991	4.00	1.75	-2.25	partly free	not free	-1
Bahrain	2002	4.00	1.75	-2.25	partly free	not free	-1
Chad	1998	4.00	1.75	-2.25	partly free	not free	-1
Tajikistan	1991	4.00	1.75	-2.25	partly free	not free	-1
South Sudan	2011	3.25	1.00	-2.25	not free	not free	0
Turkmenistan	1991	3.25	1.00	-2.25	not free	not free	0
Uzbekistan	1991	3.25	1.00	-2.25	not free	not free	0
Average	1995	6.25	3.19	-3.06			

\* CAR: Central African Republic; \*\* Congo (Brazzaville)

Source: Freedom House; own calculations.

Figure 3: Autocratic Transitions

## B Appendix B

### Proof of Proposition 1

Player  $G$  will play  $d_G$  if and only if he expects  $\tilde{h} \geq \tilde{h}^c$ , and if and only if  $\epsilon^{h^c N} \sigma s > \omega$ .<sup>7</sup> Since players  $O_i$  can observe player  $G$ 's choice prior to their respective choice being due, their choice problem reduces to the question as to whether at least  $h^c N$  of them effectively coordinate on  $d_i$  upon having observed  $d_G$ .

Now suppose that  $\tilde{h} \geq \tilde{h}^c$ . Upon having observed  $d_G$ , government official  $O_i$  has  $U_i = 1$  if he chooses  $c_i$  but  $U_i = \epsilon^{h^c N} \sigma s^h + (1 - \epsilon^{h^c N})\pi$  if he chooses  $d_i$ . He hence prefers  $(d_i | d_G, \tilde{h} \geq \tilde{h}^c)$  over  $(c_i | d_G, \tilde{h} \geq \tilde{h}^c)$  if and only if  $\epsilon^{h^c N} \sigma s^h + (1 - \epsilon^{h^c N})\pi > 1$ , or if  $s^h > \frac{1 - (1 - \epsilon^{h^c N})\pi}{\epsilon^{h^c N} \sigma}$ . At the same time,  $G$  has  $U_G = \epsilon^{h^c N} \sigma s^h$  if he chooses  $d_G$  but  $U_G = \omega$  if he chooses  $c_G$ . He hence prefers  $(d_G | \tilde{h} \geq \tilde{h}^c)$  over  $(c_G | \tilde{h} \geq \tilde{h}^c)$  if and only if  $\epsilon^{h^c N} \sigma > \frac{\omega}{s^h}$ . As a result, an action profile  $\{d_i, d_G\}$  is a Nash-equilibrium if and only if  $s^h > \max(\frac{1 - (1 - \epsilon^{h^c N})\pi}{\epsilon^{h^c N} \sigma}; \frac{\omega}{\epsilon^{h^c N} \sigma})$ . Note that  $\tilde{h} = 1$  in a Nash-equilibrium, which implies  $s^h = s$  and hence  $s > s^c := \max(\frac{1 - (1 - \epsilon^{h^c N})\pi}{\epsilon^{h^c N} \sigma}; \frac{\omega}{\epsilon^{h^c N} \sigma})$ .

Alternatively, suppose that  $\tilde{h} < \tilde{h}^c$ . Upon having observed  $d_G$ , government official  $O_i$  has  $U_i = 1$  if he chooses  $c_i$  but  $U_i = \pi$  if he chooses  $d_i$ . He thus always prefers  $(c_i | d_G, \tilde{h} < \tilde{h}^c)$  over  $(d_i | d_G, \tilde{h} < \tilde{h}^c)$ . The government leader  $G$ , in turn, has  $U_G = 0$  if he chooses  $d_G$  but  $U_G = \omega$  if he chooses  $c_G$ . He thus always prefers  $(c_G | \tilde{h} < \tilde{h}^c)$  over  $(d_G | \tilde{h} < \tilde{h}^c)$ . As a result, an action profile  $\{c_i, c_G\}$  is always a Nash-equilibrium ■

### Proof of Proposition 2

Suppose  $G$  to choose a sequence  $dc_G$ , and a share  $\tilde{g} \geq \tilde{g}^c$ . A player  $O_i$  will then have  $U_i(c_i | \tilde{g} \geq \tilde{g}^c; dc_G) = 1$  as well as  $U_i(dc_i | \tilde{g} \geq \tilde{g}^c; dc_G) = \epsilon^{h^c N} \rho^{\gamma^c h^c N} \sigma s^g + (1 - \epsilon^{h^c N})\pi$  and  $U_i(dd_i | \tilde{g} \geq \tilde{g}^c; dc_G) = \epsilon^{h^c N} \rho^{\gamma^c h^c N} \sigma s^g + (1 - \epsilon^{h^c N})\pi$ .  $O_i$  will then prefer both  $(dd_i | g \geq g^c; dc_G)$  and  $(dc_i | g \geq g^c; dc_G)$  over  $(c_i | g \geq g^c; dc_G)$  if and only if  $\epsilon^{h^c N} \rho^{\gamma^c h^c N} \sigma s^g + (1 - \epsilon^{h^c N})\pi > 1$  or if  $s^g > \frac{1 - (1 - \epsilon^{h^c N})\pi}{\epsilon^{h^c N} \rho^{\gamma^c h^c N} \sigma}$ .  $G$ , in

---

7. By convention,  $d_{i,G}$  is chosen *iff*  $U(d) > U(c)$ .

turn, will have  $U_G(dc_G|\tilde{g} \geq \tilde{g}^c) = \epsilon^{h^c N} \sigma s$  as well as  $U_G(dd_G|\tilde{g} \geq \tilde{g}^c) = 0$  and  $U_G(c_G|\tilde{g} \geq \tilde{g}^c) = \omega$ . Player  $G$  will hence never choose  $(dd_G|\tilde{g} \geq \tilde{g}^c)$ , and he will prefer  $(dc_G|\tilde{g} \geq \tilde{g}^c)$  over  $(c_G|\tilde{g} \geq \tilde{g}^c)$  if and only if  $s^g > \frac{\omega}{\epsilon^{h^c N} \sigma}$ .

On the other hand,  $G$  has  $U_G(dd_G|\tilde{h} \geq \tilde{h}^c; \tilde{\gamma} < \tilde{\gamma}^c) = \epsilon^{h^c N} \sigma S$ , but he will only have  $U_G(dc_G|\tilde{h} \geq \tilde{h}^c; \tilde{\gamma} < \tilde{\gamma}^c) = \epsilon^{h^c N} \sigma s^g < \epsilon^{h^c N} \sigma S$ . His announcement to play a sequence  $dc_G$  is hence only credible in the case that  $g \geq g^c$ . As a result, an action profile  $\{dd_i, dc_G\}$  is a subgame-perfect Nash-equilibrium if and only if  $s^g > s^c := \max(\frac{1-(1-\epsilon^{h^c N})\pi}{\epsilon^{h^c N} \sigma}; \frac{\omega}{\epsilon^{h^c N} \sigma})$ . Note that  $\tilde{g} = 1$  in the subgame-perfect Nash-equilibrium and hence  $s = s^g$ , so that the condition for a subgame-perfect Nash-equilibrium can also be written as  $s > s^c := \max(\frac{1-(1-\epsilon^{h^c N})\pi}{\epsilon^{h^c N} \sigma}; \frac{\omega}{\epsilon^{h^c N} \sigma})$ .

Alternatively, suppose that  $\tilde{h} < \tilde{h}^c$ . Then each  $O_i$  has  $U_i(dd_i|\tilde{h} < \tilde{h}^c) = U_i(dc_i|\tilde{h} < \tilde{h}^c) = \pi$  and  $U_i(c_i|\tilde{h} < \tilde{h}^c) = 1$ , so that each  $O_i$  prefers  $(c_i|\tilde{h} < \tilde{h}^c)$  over any alternative.  $G$ , in turn, has  $U_G(dd_G|\tilde{h} < \tilde{h}^c; dd_G) = U_G(dc_G|\tilde{h} < \tilde{h}^c) = 0$  and  $U_G(c_G) = \omega$  and will thus prefer  $(c_G|\tilde{h} < \tilde{h}^c)$  over any alternative. As a result, an action profile  $\{c_i, c_G\} \forall i$  is always a Nash-equilibrium in pure strategies ■