Why do Autocrats Disclose? Economic Transparency and Inter-Elite Politics in the Shadow of Mass Unrest[*]

James R. Hollyer[†1], B. Peter Rosendorff[2], and James Raymond Vreeland[3]

[1]University of Minnesota
[2]New York University
[3]Georgetown University

May 2016

Word Count (excluding appendix): 9,897

Abstract

Autocratic governments hold a preference for opacity. Autocracies are less transparent than democracies and a closed informational environment preserves autocratic regimes from popular opposition. Yet, autocracies vary widely in the extent to which they disclose economic information to their publics. In this paper, we offer an explanation for why some autocrats choose to disclose. We contend that, paradoxically, autocratic leaders may benefit from increasing the mobilizational capacity of the populace. By boosting this capacity, transparency acts as a mechanism through which autocratic leaders may threaten rival members of the elite, reducing the risk of coup and increasing their freedom of maneuver. We formalize these intuitions and demonstrate empirically that leaders in transparent autocracies suffer a reduced hazard of removal via coup relative to their opaque counterparts. Moreover, personalistic dictators and entrenched autocrats – who suffer the smallest risk of sanctioning by their regime – are particularly unlikely to disclose information.

Why do autocratic leaders, free from electoral checks on their behavior, choose to disclose information – particularly information pertaining to economic performance – to their publics?[1] Governments, it is often argued, have a taste for opacity. An absence of information facilitates rent-seeking and increases leaders' freedom of maneuver. Aside from these benefits to opacity, disclosure entails costs in terms of finances and personnel hours. National statistical offices must be staffed, press releases and other documentation crafted, all of which costs time, money and attention. Presumably, leaders who face little pressure from the public to disclose would eschew such costs.

Transparency's role in facilitating mass mobilization under autocratic rule stacks the deck still further against disclosure. Government disclosures – particularly of economic information – help members of the populace form shared beliefs regarding their leaders' performance, easing mass mobilization against the regime. Hollyer, Rosendorff and Vreeland (2015a) demonstrate that autocratic regimes that disclose large volumes of information are more likely to collapse, due to mass protest or processes leading to democratization, than regimes that fail to disclose.

Yet autocratic governments do, sometimes, disclose information to their publics. The HRV Index (Hollyer, Rosendorff and Vreeland, 2014), which measures the extent to which governments disclose credible economic information, ranks such autocratic regimes as pre-transition Korea and Mexico as more transparent than the mean democracy in their sample. Though democracies, on average, disclose substantially more than autocracies; the latter vary considerably in the extent of their opacity (Hollyer, Rosendorff and Vreeland, 2011).

In this paper, we provide an explanation for why autocrats choose to disclose. We contend that, even as disclosure renders autocratic regimes more vulnerable to mass protest, it insulates autocratic leaders from opposition that emerges *within* the regime. Indeed, transparency plays this role precisely because it facilitates mobilization by the public.

To be more precise, we contend that transparency increases the risk rival autocratic elites face when seeking to act against their leadership. Attempts by rival members of the elite to oust their leaders – which typically involve either an explicit or implicit threat of violence – may act as a focal point for mobilization by the mass populace. Such mobilization is threatening to the elite, since it may result the sweeping away of many, or all, members of the incumbent regime. In acting against their leadership, elites jeopardize their own privileged positions.

Because transparency serves to enhance the mobilizational capacity of the populace, it increases the danger the elite faces in staging a *coup d'état*. Individual citizens are better able to coordinate on engaging in protest in the wake of a coup when they recognize that their perceptions of the incumbent regime are widely shared. This might be, for instance, because the populace generally favors the (ousted) leader, and views a the coup as an indication of policy changes in an undesired direction. Conversely, it might be

---

[1]Here and throughout, we concern ourselves with the disclosure of economic information by the government to the public. We use the terms transaprency and disclosure interchageably. We elaborate on our definition and empirical operationalization of transparency in greater detail below.

because the policies of the incumbent regime are despised, and the coup is seen as a moment of regime weakness to be exploited. Regardless, any given citizen is more likely to engage in protest when she recognizes that her perceptions of the regime are widely shared by others – i.e., when the informational environment is relatively rich.

Knowing this, the elite are less likely to act against their leaders when these leaders choose to disclose. Transparency essentially acts as a tool through which leaders use the threat posed by the masses to cow recalcitrant rival members within the regime.

Leaders, therefore, are most likely to disclose when they perceive threats as emerging from within their ruling coalition, rather than from the broader population. This threat is likely to be greatest in institutionalized autocracies – i.e., regimes with designated succession mechanisms and in which the regime possesses popular legitimacy that is distinct from that of its constituent members – and when a given leader is newly installed in power. Leaders are least likely to disclose when the regime has been personalized – such that legitimacy derives from the identity of the leader, rather than the institutions of the regime – and after a given leader has become entrenched in power.[2] Empirically, we interpret these claims as holding that personalistic leaders – as defined in by Geddes, Wright and Frantz (2014) – are less likely to disclose than their analogues in non-personalistic party-based, military or monarchical regimes. We further contend that leaders are most likely to disclose when recently installed in office.

1   Argument

Our argument begins with the contention that authoritarian leaders must navigate two threats to their rule. One threat emerges from within the regime itself. The other is the danger of mass mobilization on the part of the public. We share this framework with much of the recent literature on autocratic institutions (see particularly, Svolik, 2012). Models of autocratic rule often find these threats to be strategic substitutes – policies aimed at alleviating the threat of mass revolt may increase the risk of a coup (Svolik, 2013b). Here, we contend that the reverse holds: policies that increase the mobilizational capacity of the populace may force members of the regime to toe the line set by the leadership.

This contention rests on the assumption that attempts by elites to discipline their leaders tend to destabilize the regime as a whole. Attempts to unseat the leader may serve as a focal point for mobilizing mass unrest. This unrest has the potential to lead to the upending of the regime more generally.[3]

The extent to which removing an autocratic leader destabilizes the regime is a function of the institutional foundations of the autocracy and the time a leader has served in office (Besley and Kudamatsu, 2007; Francois, Rainer and Trebbi, 2014; Gehlbach and Malesky, 2010; Svolik, 2012). Highly institutional-

---

[2]We use the male pronoun for autocratic leaders, given that such leaders historically have been overwhelmingly male.

[3]Analogously, Slater (2009) contends that opposition by communal elites, who may or may not have been previously co-opted by the regime, may facilitate unrest and democratization.

ized regimes, in which the leader is likely to have a designated successor, in which a formal voice is given to regime officials, or in which legitimacy derived from institutional positions rather than the leader's personal authority, face a low risk of mass unrest following a change in leadership. Leadership replacement in such regimes tends to be regularized and is less informative as a signal to the public of intra-regime conflict and weakness. By contrast, in personalized regimes, a changing-of-the-guard typically reflects intense intra-regime conflict and a moment of regime weakness. Under these circumstances, leadership change is particularly likely to provide a focal point for unrest, producing mass protest or setting in motion conflicts likely to upend the regime as a whole (Geddes, 1999).

Similarly, the removal of a leader who recently assumed office is less likely to produce mass unrest than the removal of a long-serving autocrat. New leaders are less likely to have had time to build up a patronage network within the regime, hence their removal is less likely to require sweeping purges of official ranks. They are less likely to have built up a cult of personality with the mass public than leaders who have long been ensconced in power (Bueno de Mesquita et al., 2002; Svolik, 2012).[4] Moreover, the time a leader has served in office may serve as a proxy for unobserved factors correlating with the extent of institutionalized, as opposed to personalized, authority. Term limits and rotation in office are often key components of regime institutionalization and, for instance, were an important aspect of Deng's attempts to institutionalize authority in the People's Republic of China (see Shih, Whan and Liu, 2010; Svolik, 2012).

Thus, members of the regime, when dealing with a leader that acts against their interest, face a choice of whether to remove this leader. Removing the leader, on the one hand, opens up the possibility that his replacement will prove more amenable to the elite. But, this is a costly gamble – removing the leader increases the risk of regime collapse, potentially costing these same elite individuals their privileged positions (Besley and Kudamatsu, 2007; Bueno de Mesquita et al., 2003; Gehlbach and Malesky, 2010; Padró i Miquel, 2007).

We contend that transparency alters elites' decision calculus. It does so by increasing the mobilizational capacity of the populace, boosting the probability that attempts to remove the leader cause the regime to collapse.[5] This pushes members of the elite toward inaction. As the danger of insurrection mounts, members of the elite grow more complacent in the face of an uncooperative leadership.

Transparency, under autocratic rule, tends to facilitate mass mobilization by the populace against the regime. This effect arises due to collective action problems inherent in mass protest. Protests that draw widespread participation are likely to succeed in forcing the regime to change policy or leadership, or even

---

[4]Francois, Rainer and Trebbi (2014) find that the hazard functions of African autocrats are monotonically decreasing over time. Guriev and Treisman (2015) argue that this pattern follows an informational logic – a dictator's survival in office causes the public to update positively with regard to his competence – an argument that suggests the leader's legitimacy, and the consequences for his removal, are both rising over time.

[5]We say 'mobilizational capacity' here to emphasize that large numbers of citizens are not yet already in the streets. The strategic logic of elites in deciding whether to stage a coup may differ if protests have *already* begun (Casper and Tyson, 2014).

in bringing about regime change. But, smaller protests are likely to be put down, at considerable cost to participants. Any individual's decision to protest thus depends on her beliefs about the willingness of others to similarly turn out. In this context, public information is likely to play an important role (Morris and Shin, 2002). Information that is witnessed by all citizens, and *known* by all citizens to be publicly observed, allows individuals not only to update their beliefs about the performance of government, but also their higher-order beliefs about the expectations held by others. In the absence of such information, uncertainty about the willingness of others to mobilize may render mass protest impossible, whereas, public disclosures by the government may render protest feasible. Hollyer, Rosendorff and Vreeland (2015*a*) formalize this argument and demonstrate, both theoretically and empirically, that transparency renders mass protest more likely under autocratic rule.[6]

We contend that the increased mobilizational capacity of the populace in a transparent environment boosts the risk that a coup might act as a focal point for protest (on a related point, see Casper and Tyson, 2014). Coups act as signals to the citizenry of (1) incipient policy changes and (2) intra-regime conflict and weakness. Since mass mobilization entails strategic complementarities, citizens are only likely to take to the streets following a coup when they recognize that others share their perception of the implications of the coup – e.g., of the desirability of policy change or of the level of support/disdain for the regime. Since transparency helps ensure that (1) these perceptions are widely shared, and (2) *known* to be widely shared, citizens are more likely to respond to a coup with protest in a transparent environment than an opaque.

Autocratic leaders determine the level of disclosure with these effects in mind. Disclosure constitutes a risky gamble for such a leader. Elites are rendered more complacent by virtue of the increased mobilizational capacity of the populace. Yet, empowering the populace in this manner is hazardous: citizens may depose the regime even as the elite toes the leadership's line. Leaders are thus placed in the position of trading off the threat they face from the populace against the threat from the elite. When the latter is high, the leader will choose to increase the popular threat to reduce that posed by the elite. By contrast, when the internal threat is low and the popular threat relatively high there is little incentive to disclose. Any gains in internal regime cohesion are more than offset by increases in threat of popular mobilization.

1.1 *Glasnost* and *Perestroika* in the USSR: An Illustrative Example

Our theory rests on an unintuitive claim: autocratic leaders may deliberately take steps that destabilize the regime so as to increase their control over members of their governing coalition. To give this claim greater empirical grounding, we illustrate the mechanisms of our argument with an example: Gorbachev's implementation of the *glasnost* and *perestroika* reforms in the former Soviet Union between 1985 and 1991. We discuss this case for illustrative purposes only. We do not contend that these reforms were *only*

---

[6]See also Bueno de Mesquita (2010), Edmond (2013), Little, LaGatta and Tucker (2015), Shadmehr and Bernhardt (2014) and Shadmehr and Bernhardt (2014).

Figure 1: HRV Transparency Scores Time Series: the USSR
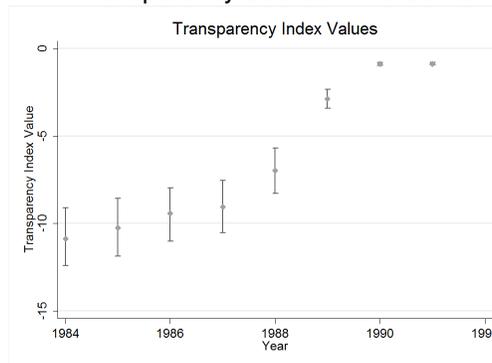
Mean HRV transparency index scores (Hollyer, Rosendorff and Vreeland, 2014), and 95 percent credible intervals, are plotted on the y-axis. Time, measured annually, is plotted on the x-axis. Mean scores are denoted by diamonds, credible intervals are denoted by whiskers. HRV Index values range from -10.9 to 10, in the full sample, with higher scores denoting greater disclosure.

the result of the mechanisms we describe in our model – for instance, Gorbachev was also likely motivated by a desire to reduce bureaucratic slack among subordinate officials (Egorov, Guriev and Sonin, 2009)[7] – nor do we seek to present a full blown case study of the former USSR. We conduct large-N empirical tests of the propositions derived from our theory below.

Mikhail Gorbachev came to power in 1985 faced with a stagnating economy and popular ideological disaffection (Dallin, 1992; Kotz and Weir, 1997). To address these threats, Gorbachev promoted a series of economic and political reforms – labeled *perestroika*[8] – which sought to reduce the authority of the Soviet industrial ministries over state enterprises, introduce limited marketization of the economy, and decentralize authority within the ruling Communist Party (Brown, 2007; Whitefield, 1993). Coupled with these structural reforms, Gorbachev instituted a variety moves aimed at political liberalization. These policies, which were grouped under the rubric of 'socialist democracy,' would come to include multi-candidate elections and increased transparency and freedom to debate (*glasnost*) (Kotz and Weir, 1997). As Figure 1 demonstrates, these 'democratizing' reforms appear as positive jumps in transparency, as measured by the HRV Transparency Index.

Most scholars have interpreted *glasnost* and socialist democracy as, in part, an attempt to overcome resistance among Party members to the economic restructuring entailed in *perestroika*. For instance, Kotz and Weir (1997, 96) state that, "As resistance mounted to [the leadership's] program of economic and social reform, Gorbachev apparently concluded that democratization was the way to break this resistance and

---

[7]We would like to thank Scott Gehlbach for brining this point to our attention.

[8]This term may be translated either as 'restructuring' or 'reconstruction'.

prevent perestroika from being stopped in its tracks," (see also Brown, 2007; Reddaway and Glinski, 2001; Whitefield, 1993). Such 'democratization' helped to ensure the compliance of Party-members through two methods. First, Gorbachev attempted to employ newly available avenues of mobilization to rally the populace to his reformist cause. Second, the creation of such avenues increased the risk that opposition groups might form within the populace and threaten continued Soviet rule. As these threats emerged, conservative Party elites were less able to take action against Gorbachev without endangering their own positions.

Scholars agree that Gorbachev's liberalizing reforms played a critical role in giving rise to a popular opposition. Aleksandr Yakovlev, Politburo member, chief of ideology and close associate of Gorbachev, remarked that in their efforts at reform, "We [the leadership] created an opposition to ourselves," (as quoted in Reddaway and Glinski, 2001, 122). Brown (2007, 18-19) describes *perestoika* as a reform "from above," and notes that prior to these reforms "[n]either a mass movement for reform nor (still less) a revolution was remotely in the cards." Brown (2007, 92-3) further notes that *glasnost* was a "double-edged" sword which aided reformers by publicizing the deficiencies of the Soviet state, but that greater access to information also risked increasing public discontent and mobilization against the leadership.

Nor were these risks merely incidental – the dangers posed by a popular opposition increased Gorbachev's authority within the Party. At times, it seems that Gorbachev acted to deliberately increase these threats. For instance, one of the most radical opposition voices belonged to Boris Yeltsin, who would come to play a pivotal role in the eventual collapse of the Soviet regime. Gorbachev had the opportunity to sideline Yeltsin in 1987, when Yeltsin was forced to resign as the Moscow Party secretary. Yet, Gorbachev merely reassigned Yeltsin to a somewhat less prominent position. Later, Gorbachev failed to remove Yeltsin from electoral ballots, following the introduction of multi-candidate elections in 1988. Reddaway and Glinski (2001) (and Yeltsin himself) attribute Gorbachev's actions to his desire to use Yeltsin's radicalism as a threat against conservative opponents. Similarly, Reddaway and Glinski (2001, 160) contend that Gorbachev "covertly encouraged" organizers of a rally by Democratic Russia in 1990, as a means of threatening more conservative opponents and ensuring the passage of reforms creating a Soviet presidency. As Brown (2007, 128) argued of the rise of a popular opposition,

> "...Gorbachev and the progress of perestroika now have liberal as well as conservative critics. While in some ways this makes life even tougher for the Soviet leader, on balance it is to his political advantage."

The presence of a popular opposition and, more generally, the opening of the informational environment increased the mobilizational potential of the populace. As the public became better able to mobilize, members of the Soviet elite faced a greater risk in moving against Gorbachev. Any destabilization of the leadership posed an increased risk of giving rise to popular unrest. Despite these risks, conservative members of the elite did move to depose Gorbachev through a coup in August of 1991.[9] Consistent with the

---

[9]The proximate cause of the coup was the negotiation of a new union treaty giving increased indepen-

mechanisms of our theory, this coup gave rise to a counter-coup led by Boris Yeltsin in his capacity as the newly created president of the Russian Republic. This counter-coup led to the ouster of the putschists and gave rise to processes leading to the collapse of the Soviet regime and the dissolution of the Union.

2   Related Literature

Our theoretical analysis builds on a growing literature on autocratic government. The intuition that attempts by the elite to discipline autocratic leaders risk the continued stability of the autocratic regime, and hence officials' continued grasp their privileged positions, owes heavily to Bueno de Mesquita et al. (2003) and Besley and Kudamatsu (2007). Padró i Miquel (2007) also builds on this logic, and contends that this 'politics of fear' allows leaders to expropriate from their winning coalition.[10] Gehlbach and Keefer (2011) argue that because autocratic institutions help limit the risks faced by the regime from deposing the leader, institutionalization promotes the credibility of promises not to expropriate regime elites (see also Gehlbach and Keefer, 2012; Myerson, 2008). Geddes (1999) argues that institutionalization – and particularly the discrepancy between personalistic and party-based structures – explains the survival of autocratic regimes. Party-based regimes, she contends, rely less on the personal authority of the leader and incorporate groups that may act as a check on the leadership, facilitating regime-survival after a given leader dies or is removed from office. Gandhi (2008) also emphasizes the importance autocratic institutions, primarily focusing on the role of legislatures in co-opting potential opposition groups (see also Gandhi and Przeworski, 2006, 2007). She further codes the relationship between autocratic leaders and their elite inner sanctums (see Cheibub, Gandhi and Vreeland, 2010, 84-86), which is closer to our theoretical focus.

We also rely on insights from Svolik (2012, 2013b) regarding the interaction between the twin threats to authoritarian leaders – those from within the regime and those emerging from the populace. Bueno de Mesquita and Smith (2009) also build on the insight that autocratic leaders must meet these twin threats, and like this paper, they consider the role of public goods that enhance the ability of the populace to coordinate. Casper and Tyson (2014) analogously examine the relationship between protests and coups – and find feedback mechanisms between these two forms of threat. Specifically, they find that media freedoms increase the risk that one form of such threat sparks the other, in line with our primitive assumption that coups may give rise to protest, particularly in the presence of transparency.

Our paper also closely relates to a literature on various forms of transparency under autocratic rule. Typically, these studies focus on the role of the mass media. Accounts by Egorov, Guriev and Sonin (2009) and Lorentzen (2014) argue that autocratic regimes can effectively outsource the role of monitoring their lower level agents to the media. The benefits of such monitoring must be traded off against the risk that a free media may promote mass public opposition. King, Pan and Roberts (2013, 2014), in a series of

---

dent authority to the 15 constituent republics of the USSR.

[10]Di Lonardo and Tyson (2015) develop a related argument, in which an external threat diminishes group infighting, as applied to a terrorist group rather than an autocratic regime.

7

innovative studies of the PRC's 'Great Firewall', find supportive evidence for these accounts – online censors permit criticism of local government corruption and other forms of mis-governance, but delete calls for protest. In complementary theoretical explorations of media control under autocracy, Gehlbach and Sonin (2014) and Shadmehr and Bernhardt (2015) examine autocrats' incentive to engage in media censorship, given that citizens rationally update to discount good news about the regime or interpret no news as bad news, while Guriev and Treisman (2015) examine the trade-offs dictators face between employing censorship, propaganda and the co-optation of elites.[11]

While these papers share our focus on information dissemination under autocratic rule, the conception of transparency used in these papers differs fundamentally from that herein. Rather than focusing on the media, we emphasize the role of *government* disclosures of information to the public. Increasing transparency does little to enhance monitoring of lower-level public officials in our formulation, since there is no outsourcing of information collection to private organizations.

Boix and Svolik (2013) also examine the role of a form of transparency under autocratic rule, and – like this paper – they conclude that higher levels of transparency are associated with a reduced risk of coup. But, like studies that focus on the media, their conception of transparency differs fundamentally from that used in this paper. In their model, transparency relates to the ability of members of the regime to observe efforts by autocratic leaders to amass greater authority and power. Transparency is thus, in their formulation, the clarity of the 'rules of the game.' As the authority vested in an autocratic leader becomes more clearly defined, conflict is less likely to emerge between these leaders and members of their ruling coalition. While the clarity of the rules of the game may involve the availability of economic information – for instance information regarding the budget – the critical concern for Boix and Svolik (2013) is the role of information *available to the elite*. By contrast, our concern here is with information made available *to the wider public*. If a given piece of information is revealed to the public, it must perforce also be accessible to members of the regime elite – so, these two notions of transparency must be at least somewhat correlated. However, there is no reason to expect that the reverse holds – considerable amounts of information are likely circulated among autocratic elites and not disclosed to the broader public.

This paper most closely relates to recent work on the measurement and implications of government disclosure by Hollyer, Rosendorff and Vreeland (2014, 2015*a*,*b*). Our definition and – in empirical sections – our measure of transparency is derived from Hollyer, Rosendorff and Vreeland (2014), which focuses on the disclosure of economic information by the government to the broader public. We treat the findings of Hollyer, Rosendorff and Vreeland (2015*a*), that transparency increases the risk of mass public protest under autocratic rule, as a theoretical primitive. Hollyer, Rosendorff and Vreeland (2015*a*) also find, in an appendix, that transparency is associated with a reduced risk of autocratic regime collpase brought on by coups, a finding which is consistent with our argument. However, our interest here is in autocratic leader removal, which may or may not entail regime collapse. Finally, unlike these other pieces, this work focuses

---

[11]For an excellent review of the literature on information problems in non-democracies, see Wallace (2015).

on the conditions under which autocrats disclose.

3 Model

We present a model of autocratic disclosure, incorporating the threats posed to autocratic leaders from both members of the regime and the mass populace. Steps taken by regime-members to discipline their leaders threaten the stability of regime. In these assumptions, our model shares features of work by Besley and Kudamatsu (2007) and Padró i Miquel (2007). We, however, incorporate disclosure into the model, which increases the risk of mass mobilization. The regime's leaders, in order to forestall sanction by the elite, strategically choose a level of disclosure, threatening both the leader's and the rival elite's survival.

3.1 Model Primitives

Consider an interaction between an autocratic leader L, a group of regime elites R, and the mass of ciitizens denoted M. The leader L chooses whether to disclose $d \in \{0, 1\}$. Regime elites R observe the choice of d, and must determine whether to retain him in office.[12] We denote this decision $v \in \{0, 1\}$, where $v = 1$ denotes a decision to remove. v may thus be thought of as a decision to launch a coup.

Following the choices of d and v, a contest for power takes place between members of the regime and the populace. If R choses to keep L in place (sets $v = 0$), the probability that the regime falls is given by $p(d)$, where $1 > p(1) > p(0) > 0$. We thus make a primitive assumption that greater disclosure causes increased mobilizational capacity. This assumption is consistent with the findings of Hollyer, Rosendorff and Vreeland (2015a). If L is removed, the probability the regime falls to mass insurrection is given by $\omega p(d)$ where $\omega \in (1, \frac{1}{p(1)})$. The term ω reflects the tendency of internal strife to destabilize the regime. One can think of ω as reflecting the strength of L vis-à-vis R. To anticipate the empirical flavor of ω, following the approach of Bueno de Mesquita et al. (2003), ω is a function of L's time in power. During L's tenure, he entrenches himself and cultivates an R with increasing levels of dependency and loyalty (see chapter 3 in Svolik, 2012). Alternatively, Besley and Kudamatsu (2007) interpret this parameter as the the degree to which the autocracy is personalistic, in the sense of Geddes, Wright and Frantz (2014): In low-ω autocracies, autocratic succession is institutionalized, and R enjoys some power independent of L. The risk of regime collapse is higher in high-ω autocracies, where the regime's legitimacy depends on a cult of personality surrounding L. A third interpretation, which we use in our empirical appendix, conceives of ω as reflecting L-R relations along the lines of (Gandhi, 2008): in autocracies with a strict inner-circle hierarchy – i.e., military dictatorships and monarchies – L enjoys high-ω, whereas in autocracies where L power over the elites is more precarious – i.e., civilian autocracies – ω is low

Let national income (or, equivalently, the rents accruing to the regime) be y; if a leader survives in

---

[12]As is true throughout, we use the male pronoun to refer to autocratic leaders. We ascribe female pronouns to R and M for purposes of clarity.

office, he derives utility from the share of national income he consumes λy. We assume that $\lambda \in (\frac{1}{2}, 1)$ which captures a divergence of interest between the leader L and the elite R – a benefit to being the leader, and a motive for the elite to desire to remove the leader. R earns the residual $(1 - \lambda)y$, if the regime is not over thrown by the masses.

Of course, if mass insurrection takes down the regime, then L and R get nothing, regardless of the value of v. We allow for a further "congruence" of interest across L and R by another variable, r. One can think of r as the returns from regime cohesion. It represents the additional payoff earned by the leader and the elite in the instance that the elite chooses not to remove the leader and the regime (the leader and the elite together) survives potential mass insurrection. If both L and R survive in power together, then both get the added benefit r. If R sets $v = 1$, she looses r, and (the now-ousted) L receives a payoff normalized to 0. Note that overthrowing L involves, at a minimum, extremely careful planning by R and a single bullet, and more often involves an ugly stand-off between L and R, which can persist over long periods. We simplify the situation by assuming the R wins such stand-offs, but not without entailing losses. While the value of r is common knowledge, its value is drawn prior to the play of the game from a distribution with density G over domain $(0, \infty)$. That is, $G(0) = 0, G(\infty) = 1$ and $G'(.) > 0$

So, if $v = 0$, L has a $(1-p(d))$ chance of receiving $\lambda y(1+r)$, and a $p(d)$ chance that mass insurrection brings down the whole regime, and leaves L with 0. L also ends up with 0 if R sets $v = 1$, removing him from office.

In a complementary fashion, if $v = 0$, R has a $(1 - p(d))$ chance of receiving $(1 - \lambda)y(1 + r)$, and a $p(d)$ chance that mass insurrection brings down the whole regime, leaving R with 0. If the elite decides to oust L (setting $v = 1$), she gets the bigger piece of the y pie – the λ share – but the ouster is costly, such that she no longer gains r.[13] The ouster is costly in another way as well – it sends a signal to the masses. Sensing regime weakness, the masses may take to the streets and bring down the entire regime. The probability of successful mass insurrection is thus augmented by ω. So, if $v = 1$, R has only a $1 - \omega p(d)$ chance of getting λy (and no r) – and she faces a $\omega p(d)$ chance that mass insurrection brings down the whole regime, leaving R with 0.

Summarizing the expected payoffs of L and R, we have:

$$\text{Leader} \quad EU^L(d; v) = \begin{cases} (1 - p(d))\lambda y(1 + r) \text{ if } v = 0 \\ 0 \text{ if } v = 1 \end{cases}$$

$$\text{Elite} \quad EU^R(v; d) = \begin{cases} (1 - p(d))(1 - \lambda)y(1 + r) \text{ if } v = 0 \\ (1 - \omega p(d))\lambda y \text{ if } v = 1 \end{cases}$$

---

[13]Presumably, the rest of the y pie – the $(1 - \lambda)$ share – goes as a payoff to a newly emerging elite, but this new elite does not have a strategic move in this game, and is thus irrelevant to our model.

3.2 Three Preliminary Lemmas and Two Definitions

We begin by considering the elite's decision to remove the leader, conditional on having observed both the leader's action d and the realization of the variable r, capturing the benefits of loyalty. R sets $v = 1$ whenever $EU^R(1; d) \geq EU^R(0; d)$.

Lemma 1 establishes that this condition is satisfied when $r < \frac{2\lambda - 1 - p(d)(\lambda\omega - 1 + \lambda)}{(1-\lambda)(1-p(d))} \equiv r(d)$ for $d = 0, 1$. That is, the elite removes the leader whenever the benefits to cohesion, as reflected in the random variable r, are sufficiently low.

Lemma 2 formalizes the observation that since $p(1) > p(0)$ and $\omega > 1$, it must be that $r(1) < r(0)$. The threshold necessary to ensure elite quiescence ($r(d)$) falls when the leader chooses to disclose, relative to when he chooses not to do so.

It will also be convenient to define $\tilde{\omega} \equiv \frac{2\lambda - 1 + p(0)(1-\lambda)}{p(0)\lambda}$. Then Lemma 3 establishes that if $\omega < \tilde{\omega}$ then $r(0) > 0$. Proofs of these Lemmas, and all theoretical propositions, are presented in Appendix A.

3.3 Equilibrium

Recall that the elite and the leader are battling over shares of national income, and whoever holds the "leadership" gets the lion's share, $\lambda > 1/2$. Under the right parameter values, the leader discloses to prevent the elite from arranging his ouster.

If the masses win in their insurrection, then they become the leader, otherwise they get zero. So, they have a dominant strategy to always rise up. Hence we don't model their decision. When the masses win, both the elite and the leader are removed.[14]

Proposition 1 demonstrates that the subgame perfect equilibrium strategies for both R and L are conditional on the parameter ω. When ω is high – specifically $\omega > \tilde{\omega}$, the implications of infighting for regime (leader and elite) survival are sufficiently large that R will never challenge L – i.e., $v = 0$, $\forall\, r, d$. For lower values of ω, however, the elite may choose to remove the leader. They will only do so, however, if the benefits to cohesion (r) are sufficiently small. We can thus construct a subgame perfect equilibrium to this interaction as follows:

Proposition 1. *If $\omega < \tilde{\omega}$ then the Nash Equilibrium is:*
For any $r < r(1)$, $v = 1$ and $d = 1$;
For $r > r(0)$, $v = 0$ and $d = 0$;
For $r(1) \leq r \leq r(0)$, $d = 1$ and $v = 0$.

Note that if $\omega > \tilde{\omega}$, then in equilibrium, the leader never discloses ($d = 0$) since there is no credible coup threat.

---

[14]All results are robust to a more complicated decision-calculus for the masses, so long as they are more willing to arise given (1) infighting between L and R and (2) given disclosure (which we assume boosts the probability of mass success as a primitive).

Proposition 2. *If $\omega \geq \tilde{\omega}$ then the Nash Equilibrium is $v = 0$ and $d = 0$ for all r.*

Combining the insights of the two equilibria (which span the parameter space), we reach our central finding – that disclosure is less likely as the regime grows more personalistic. In personalistic regimes, the removal of the leader is particularly destabilizing, ω is high. From Proposition 1, we see that that disclosure occurs when $r \in [0, r(0)]$. Recall from the definition of $r(0)$ that this threshold is a function of ω. Our central result, Proposition 3 establishes that this disclosure region shrinks as ω rises, and since r is a random variable drawn from a distribution with a strictly monotonic density, we have:

Proposition 3. *Disclosure is less likely when ω rises.*

In Appendix B we offer a model with richer microfoundations that motivates this simple formulation of congruence or accountability between the leader and the elite. There, the primitive conflict between L and R is a matter of preferences – they are either congruent or divergent. In that model, in addition to choosing to disclose, the leader makes a policy choice. Congruent leaders share the policy preferences of the elite, divergent leaders do not. Since the leader's type is private information, we specify a perfect Bayesian equilibrium (which satisfies the intuitive criterion) to generate a unique equilibrium which has similar properties: as long as ω is not too large, there are parts of the parameter space in which disclosure is adopted (by a divergent type) to forestall a coup. Moreover in that specification, the leader chooses disclosure for larger subsets of the parameter space when ω is low, relative to when ω is high.

Both models reach two main conclusions that we can test empirically: (1) disclosure reduces the probability of regime disunity, as measured by coups, (2) high-ω leaders have little to fear from elites, and thus disclose less. We now turn to our empirical tests of these two results of our theory.

4 Empirics

4.1 Transparency and Coups

Proposition 1 forms the first basis of our empirical investigation. We empirically interpret this claim as holding that, conditional on autocratic institutions, transparency insulates leaders from coups.

Note that this is not a comparative static claim. Both disclosure and leader removal (i.e., coups) are endogenous in our model. However, it is clear from Proposition 1 that leaders disclose only to forestall the threat of coup.[15] Empirically, since leaders may be limited in their ability to boost disclosure – by, for instance, capacity constraints (Hollyer, Rosendorff and Vreeland, 2014; Stone, 2008) – or may imperfectly judge the level of disclosure necessary to forestall a coup, we expect that higher (lower) levels of transparency are associated with a reduced (enhanced) hazard of coup.

---

[15]One could trivially construct an analogous model, in which disclosure were not an option, and find that – in the interval $r(1) \leq r \leq r(0)$ – leaders no longer survive. We conduct an analogous exercise for a different variant of the model in Appendix B.

We estimate the relationship between the hazard of leader removal via coup and transparency using a series of Cox competing hazards specifications.[16] Competing hazards models are a means of estimating the hazard (the probability that an event takes place in time t, given that it has not already taken place) of one of several mutually exclusive events. In our case, the event of interest is leader removal via coup, and the competing hazards are alternate forms of leader removal.[17] The unit of observation is the autocratic leader-year, where leader identities and the methods and dates of leader removal are defined by Svolik (2012).

To adjust for the possibility that past successful coups predict future coups, we stratify the baseline hazard rate in our Cox estimates based on two measures of coup history.[18] The first is a simple indicator variable $\{0, 1\}$, equal to one if any past autocratic leader was removed via a coup. The second is an ordered variable $\{0, 1, 2, 3\}$, which – if equal to any element in $\{0, 1, 2\}$ denotes the number of past autocratic leaders removed via coup and, if equal to 3, denotes that more than two previous leaders have been removed via coup. We thus fit specifications of the following form:

$$h_l(t) = h_0(t, c_l) exp(\gamma transparency_{l,t-1} + \mathbf{X_{l,t-1}}\beta)$$ (1)

where l denotes autocratic leader, t denotes time, c_l is an indicator for coup history, and $\mathbf{X_{l,t-1}}\beta$ is a vector of time-varying controls and associated coefficients. $h_0(t, c_l)$ is estimated non-parametrically within each strata, based on the fraction of observations that experience a coup at time t as compared to the number of observations at risk (leaders who have not yet been removed for *any* reason). Duration dependence is thus factored out of the likelihood function, while a history of coups may shift both the shape and level of the baseline hazard function (Beck, Katz and Tucker, 1998; Box-Steffensmeier and Zorn, 2002). Our hypothesis holds that the coefficient on transparency, γ, should be negative.[19]

Our measure of transparency in these regressions is the HRV Index (Hollyer, Rosendorff and Vreeland, 2014). The HRV Index is a continuous measure which reflects the public disclosure, by governments, of credible economic information. It is constructed by relying on the presence or absence of data from the World Bank's World Development Indicators (WDI) data series. Since data disclosed to the World Bank

---

[16]These models are similar to results in the appendix to Hollyer, Rosendorff and Vreeland (2015*a*). However, those results pertain to regime, rather than leader, removal. Since their definition of regime collapse requires the replacement of one leader by another unaffiliated with the ruling clique, they estimate $Pr(leader\ removed)Pr(unaffiliated\ replacement|leader\ removed)$.

[17]See Goemans (2008) for an alternative application of the competing hazards model. This approach assumes that the hazard of one form of removal is conditionally independent of alternative forms of removal (Gordon, 2002).

[18]On this approach, see Box-Steffensmeier and Zorn (2002), who term this a 'conditional gap time' model.

[19]We test the proportional hazards assumption for all specifications and address violations as suggested by Box-Steffensmeier and Jones (2004) and Keele (2010).

are, in recent years, publicly available and, in the past, were widely available to researchers, the presence of such data are reflective of a general tendency to disclose economic information. Since the World Bank actively reviews these data for inaccuracies, and deletes reported observations it finds incredible, the availability of such observations reflects a tendency to disclose credible economic information. The HRV Index scales the presence/absence of 240 measures from the WDI using an item-response algorithm, producing a continuous measure of transparency, measured at the country-year level, for 125 countries from 1980-2010.

Our controls include several measures of autocratic institutions, drawn from the *Autocratic Regimes* dataset of Geddes, Wright and Frantz (2014) (henceforth the GWF dataset). This dataset partitions autocracies into four categories: personalistic dictatorships, single party regimes, military regimes, and monarchies. There are, however, only eight monarchical regimes in our sample, constituting less than 10 percent of the total. We thus treat non-personalistic military regimes and monarchies as a single reference category. We control for binary indicators of single party and personalistic regimes in all specifications. These controls are necessary given substantial evidence that autocratic regime survival is conditioned by institutional features (Boix and Svolik, 2013; Gandhi, 2008; Gandhi and Przeworski, 2007; Geddes, 1999).

We prefer the GWF dataset on autocratic institutions for our empirical specifications throughout, since we believe their definitions map more easily into our model parameter ω than common alternative measures of autocratic institutions. Personalistic regimes, we believe, correspond to $\omega > \tilde{\omega}$. However, we recognize that the mapping from this theoretical parameter into a set of empirical measures is somewhat fraught. To borrow a term from Pepinsky (2014*a*), ω reflects a 'logics' approach to the study of authoritarian regimes – it corresponds to a measure of the distribution of power between regime elites and dictators. This balance of power both gives rise to certain observable institutional configurations and is sustained by such configurations (on this point, see Pepinsky, 2014*b*; Svolik, 2013*a*). We emphasize that our empirical results pertaining to autocratic institutions should not be interpreted causally – institutions may play a causal role over both leader survival and disclosure, but institutions also are affected by unobserved intra-regime dynamics that may also have a causal effect on these terms.

We additionally control for economic covariates drawn from the Penn World Table version 7.1 (Heston, Summers and Aten, 2012). Specifically, we control for levels of economic growth in real PPP GDP, measured in percentage points. We additionally control for real GDP *per capita*, measured in thousands of constant 2005 PPP adjusted US dollars. The former control is necessary given substantial evidence that poor economic growth precipitates regime instability under autocratic rule (Haggard and Kaufman, 1995; Przeworski et al., 2000). The latter is included given the findings of a large literature on the role of economic development in destabilizing autocracies (Acemoglu et al., 2009; Boix, 2003; Boix and Stokes, 2003; Przeworski and Limongi, 1997; Przeworski et al., 2000).

Results from the model specified in Equation 1 are presented in Table 1. Results in which the baseline hazard is stratified based on whether there was a previous coup are presented in the leftmost column; those stratified based on the ordered indicator of coup history are presented in the center column; and results

that do not stratify the baseline hazard, but simply control for an indicator of past coups, are presented in the rightmost column.

Table 1: Hazard of Leader Removal via Coup

	Past Coup Strata	Coup Experience Strata	Past Coup Control
Transparency	-0.248	-0.282	-0.240
	[-0.480,-0.016]	[-0.531,-0.033]	[-0.461,-0.019]
Growth	-0.003	-0.005	-0.000
	[-0.031,0.026]	[-0.042,0.032]	[-0.029,0.029]
GDP *per capita*	-0.110	-0.094	-0.117
	[-0.208,-0.012]	[-0.175,-0.013]	[-0.229,-0.005]
Party	-1.793	-1.709	-1.735
	[-2.595,-0.991]	[-2.451,-0.967]	[-2.661,-0.810]
Party $\times$ t	0.113	0.112	0.109
	[0.045,0.181]	[0.049,0.175]	[0.037,0.182]
Personal	-0.807	-0.676	-0.809
	[-1.609,-0.004]	[-1.437,0.084]	[-1.592,-0.025]
Ever Past Coup			-0.047
			[-0.908,0.814]
# of Subjects	89	89	89
# of Failures	36	36	36

Results from Cox competing hazards regressions of leader removal via coup on transparency and controls. 95 percent confidence intervals are reported in brackets.

In all specifications, the coefficient on the transparency term is negative, with 95 percent confidence intervals bounded away from zero. Point estimates indicate that a one standard deviation increase in the HRV Index is associated with a 35-40% fall in the hazard of a coup. The associated 95 percent confidence interval runs from a decline of 3% to a decline of 58% in the hazard function.

To ease interpretation of these results, we present plots of smoothed estimates of the hazard function in Figure 2. The solid line depicts the smoothed hazard when the HRV Index is at its 10th percentile in the sample, while the dashed line presents the same when the HRV Index is at its 90th percentile. In the former instance, the hazard that a newly seated leader is ousted via a coup during the first year of his reign is roughly 3.25 percent. When transparency is at its 90th percentile, this falls to a hazard of roughly 1.75 percent.

Figure 2 also reveals that the hazard rate declines over time, particularly over the first 20 years of a

leader's rule. This is consistent with our claim that entrenched autocrats face fewer risks from within the ruling elite – a point to which we return below.

Figure 2: Smoothed Hazard Function of Leader Removal via Coup
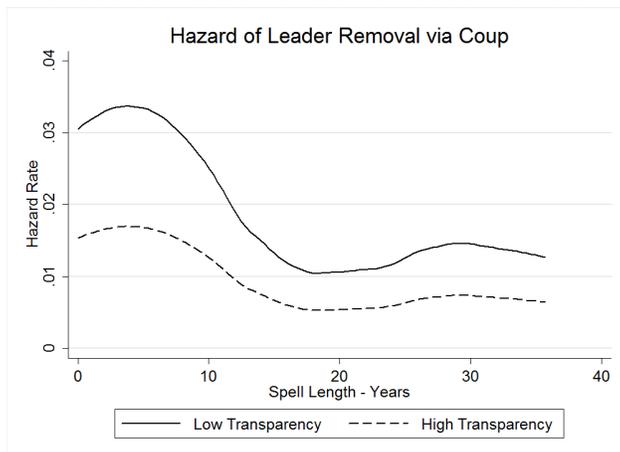
Estimates of the smoothed hazard function from the model in the third column of Table 1. The solid line depicts the hazard of coup when transparency is at its 10th percentile in the sample. The dashed line depicts the corresponding hazard when transparency is at its 90th percentile. Estimated hazards are for a non-personalistic military/monarchical regime that has not previously experienced leader removal via a coup.

In the appendix, we present robustness checks to our baseline specification. Specifically, we substitute alternative measures of autocratic institutions for the GWF measures. That model includes controls, drawn from the DD dataset (Cheibub, Gandhi and Vreeland, 2010). The DD dataset does not have a 'person-alistic' category. Instead, the dataset explicitly codes leader-elite relations as 'hierarchical' for the military dictatorships and monarchies (note that some of these are coded as "personalistic" in the GWF dataset), and "precarious' civilian autocracies (Cheibub, Gandhi and Vreeland, 2010, 84-86). In the former category the leader enjoys more control over elites through the hierarchy of the inner sanctum, while civilian rule is more 'precarious' (Cheibub, Gandhi and Vreeland, 2010, 84-86). Result are substantively similar to those presented here, if slightly less precisely estimated. The magnitude of the coefficient on the HRV Index is roughly the same as in the baseline model, but this estimate is only consistently significant at the 90 percent level.

4.2 Autocracy and Transparency

In addition to advancing claims about the relationship between transparency and the frequency of coups, our theory offers predictions about which autocratic regimes are likely to disclose information. Proposition 3 contends that the probability of disclosure rises as ω declines. We treat this proposition as advancing an

empirical claim that (1) personalistic regimes disclose less than other autocracies, and (2) leaders become less willing to disclose once entrenched in office (and are more willing to disclose when new to office).

In this section, we assess these two claims through a series of varying intercepts hierarchical regression models of the HRV Index on a series of institutional and time-varying characteristics. Specifically, we estimate models of the following form:

$$transparency_{i,t} = \rho transparency_{i,t-1} + \alpha_i + \mathbf{X_{i,t-1}}\beta + \epsilon_{i,t}$$
$$\alpha_i \sim N(\mathbf{Z_i}\gamma, \sigma_\alpha^2) \tag{2}$$

where i denotes a given autocratic regime, t denotes the year, $\mathbf{X_{i,t-1}}$ is a vector of time-varying controls, α_i is a regime-specific intercept term, and $\mathbf{Z_i}$ is a vector of regime-level (time-invariant) controls.[20]

Our definition of an autocratic regime is drawn from the dataset of Svolik (2012), in which an autocratic regime consists of the continuous succession of autocratic leaders affiliated with a given ruling clique. This dataset also provides our definition of a leader's time in office. We code an indicator $New\ Leader \in \{0,1\}$, which is set equal to one if a leader has been in power for 5 years or less.

As in the above, we draw our measure of transparency from the HRV Index and our measures of autocratic institutions from the GWF dataset. Our theory holds that personalistic regimes should disclose less than alternative regime classifications.

We draw a number of economic controls from the Penn World Table, version 7.1 (Heston, Summers and Aten, 2012). We include controls for both GDP *per capita* and chain-series GDP, both measured in PPP adjusted 2005 US dollars. These controls are included given that more developed states may simply be more able to disclose data than less developed ones. We additionally control for economic openness ($\frac{Exports+Imports}{GDP}$), given that more open economies may be subject to more pressure to disclose from international markets; economic growth (percent change in GDP), given the possibility that governments may be more tempted to disclose when times are good; and government consumption as a percent of GDP, given that statist economies may face less pressure to disclose information to the general public.

All regressions additionally include an indicator for fuel exporters, drawn from Easterly and Sewadeh (2001). We include this term given the findings of Egorov, Guriev and Sonin (2009), who find that natural resource exporters are more opaque than other autocracies.

The GWF regime classifications and the fuel exporter indicator constitute the time-invariant regime characteristics $\mathbf{Z_i}$, from equation 2. The $New\ Leader$ indicator and economic covariates constitute the time-varying terms $\mathbf{X_{i,t-1}}$. We additionally control for the lagged value of the HRV Index, which also varies over time.

This last term is included to capture model dynamics (Beck and Katz, 2011). Dynamics are of concern on both theoretical and empirical grounds. Empirically, transparency evolves according to a smooth (slow-

---

[20]The varying intercepts model is preferred over a fixed-effects estimator given that our institutional covariates of interest are largely time-invariant.

moving) process – indeed, the HRV Index is constructed in a manner that includes a non-parametric inter-temporal smoother (Hollyer, Rosendorff and Vreeland, 2014). Theoretically, it is implausible to assume that, for instance, a new leader aiming to increase levels of disclosure will be able to achieve this increase within the space of a single-year. Rather, one would expect this increase to play out over a period of several years. Where such dynamics are present, excluding the lagged dependent variable would result in model misspecification. To deal with potential bias resulting from the inclusion of panel-specific random effects and a lagged outcome variable, we fit variants of our model, based on an Anderson-Hsiao estimator, in Appendix C.2.1. Results are substantively similar, if somewhat less precisely estimated, to those presented here.

We estimate the model described in equation 2 via MCMC in JAGS 3.3.0. We place separate diffuse multivariate normal priors on, respectively, the coefficients on regime-level variables γ and the regime-year level variables β. We place an informative prior on $\rho \sim N(0,1)$.[21] All variables that are not either binary indicators nor time counters have been standadized, by subtracting the mean and dividing by the standard deviation.

The results from models based on equation 2 are presented in Table 2.

[21] Further details are available in Appendix C.2.1.

Table 2: Models of Disclosure: GWF Data

		Model 1	Model 2	Model 3
Regime Predictors	**Party**	0.002	0.002	0.002
		[-0.033, 0.038]	[-0.039, 0.031]	[-0.037, 0.036]
	Personal	-0.039	-0.038	-0.044
		[-0.083, -0.001]	[-0.085, -0.007]	[-0.087, -0.008]
	Fuel Exporter	-0.037	-0.036	-0.033
		[-0.082, 0.010]	[-0.073, 0.006]	[-0.070, 0.008]
Regime-Year Predictors	**Lag Transparency**	0.960	0.961	0.964
		[0.943, 0.978]	[0.943, 0.977]	[0.947, 0.980]
	GDP *per capita*	0.004	0.005	
		[-0.012, 0.021]	[-0.013, 0.020]	
	GDP	0.003	0.003	
		[-0.009, 0.017]	[-0.009, 0.016]	
	Ec. Openness	0.002		
		[-0.013, 0.017]		
	Growth	-4×10^{-4}		
		[-0.011, 0.010]		
	Gov't Consumption	-0.010	-0.010	
		[-0.02, 0.003]	[-0.023, 0.004]	
	New Leader	0.023	0.024	0.0248
		[-4×10^{-4}, 0.047]	[0.001, 0.048]	[0.002, 0.049]
	Cubic Time Polynomial	✓	✓	✓
	# Obs	1530	1530	1530
	# Regimes	119	119	119

Results from a hierarchical varying-intercepts linear regression of HRV transparency index scores on listed covariates. Covariates that shift the intercept term are described as 'Regime Predictors', while those that directly shift predicted transparency values are listed as 'Regime-Year Predictors.' All covariates that are neither indicators terms nor time counters have been standardized by subtracting the mean and dividing by the standard deviation. 95 percent credible intervals are presented in brackets.

In all specifications, the coefficient on the $New\ Leader$ indicator is positive with 95 percent credible intervals bounded away from zero. Recall further that the presence of the lagged dependent variable renders these models dynamic. We simulate the estimated marginal effect of a new leader over time, based on the results from our baseline specification, in Figure 3. All estimates are from Model 1 in Table 2.

Figure 3: Marginal Effect of a New Leader

Marginal Effect of New Leadership
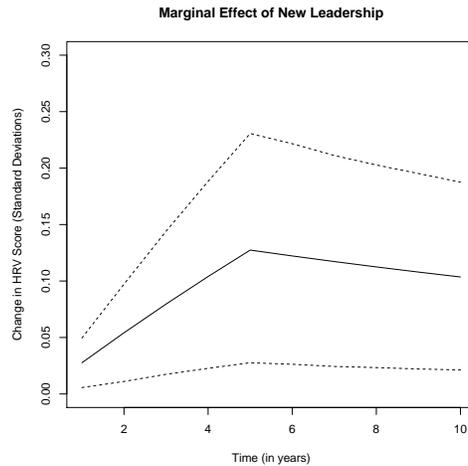
Estimates of the marginal effect of the introduction of a new leader in time $t = 1$
over a 10 year period. Standardized HRV transparency index scores are plotted
on the y-axis. Time, measured in years, is plotted on the x-axis. Solid lines depict
mean estimated marginal effects, dotted lines depict 95 percent credible intervals.

The estimated marginal effect of a new leader is small in absolute terms. The introduction of a new
leader in time $t = 0$ is anticipated to increase transparency scores by roughly 0.12 standard deviations by
time $t = 5$. However, our transparency scores tend to vary little within autocratic regimes over time. On
average, the standard deviation of the (standardized) transparency measure within an autocratic regime is
0.28. Thus, a new leader is predicted to increase transparency scores by roughly $\frac{1}{3}$ to $\frac{1}{2}$ their over-time
standard deviation by the conclusion of his first five years in office.

The coefficient on the indicator for personalist regimes is consistently negative – again in line with the-
oretical expectations. 95 percent credible intervals are bounded away from zero in all specifications. The
estimated contemporaneous marginal effect of changing from a non-personalistic military to a personalistic
regime is a decline of between 0.038 and 0.044 standard deviations in the HRV Index. (The analogous
contemporaneous marginal effect of a change from a party-based to a personalistic regime is a decline
of 0.04 standard deviations). Given the dynamics of the model, however, the long-term equilibrium asso-
ciation between regime classification and transparency is considerably larger. The long-term equilibrium
difference between personalistic and non-personalistic military regimes is estimated to be roughly 0.97
standard deviations.

Finally, the coefficient on the $fuel\ exporter$ indicator is negative in every model estimated, as is consis-
tent with Egorov, Guriev and Sonin (2009). Though, these results are not typically significant. Interestingly,
coefficients on economic covariates are consistently small and imprecisely estimated in all models.

In the appendix, we present robustness checks substituting measures of autocratic institutions from

the DD dataset. Results on leader entrenchment – i.e., that new leaders disclose more readily than those long established in office – continue to hold in these robustness checks. We similarly find that autocracies with hierarchical leader-elite relations (high-ω leaders) disclose at lower rates than precarious civilian autocracies (low-ω leaders).

5   Conclusion

Leaders in transparent autocracies enjoy a reduction in the risk of coup. Across a variety of statistical specifications, an increase of one standard deviation in HRV Index scores is associated with a roughly 35% reduction in the hazard of coup. While such leaders are more insulated from threats that emerge from within their regime, exiting results indicate that they suffer an increased risk from the masses (Hollyer, Rosendorff and Vreeland, 2015*a*). Autocrats are thus placed in a delicate position when deciding whether or not to disclose: they must trade off the threats they face from within their regime against those from without.

In this paper, we introduce a theoretical argument prefaced on such a balancing act. Transparency insulates leaders from internal threats, we argue, precisely *because* it exposes them to increased external threats. The common threat to both the leader and the autocratic elite posed by mass unrest ensures greater cohesion within the regime (Padró i Miquel, 2007; Di Lonardo and Tyson, 2015). Autocratic leaders, therefore, can manipulate this threat to ensure greater freedom of maneuver vis-à-vis their winning coalition. Transparency serves as one mechanism to gain this freedom.

This argument implies that autocrats should be most prone to disclose information when the internal balance of power tends to favor the autocratic elite over the leader. It is in these circumstances in which the leader stands to benefit most from any additional freedom of maneuver. Consistent with this argument, personalistic and entrenched autocrats are, empirically, less prone to transparency than their newly installed and institutionalized counterparts.

References

Acemoglu, Daron, Simon Johnson, James A. Robinson and Pierre Yared. 2009. "Reevaluating the Modernization Hypothesis." *Journal of Monetary Economics* 56(8):1043–1058.

Beck, Nathaniel and Jonathan N. Katz. 2011. "Modeling Dynamics in Time-Series-Cross-Section Political Economy Data." *Annual Review of Political Science* 14:331–352.

Beck, Nathaniel, Jonathan N. Katz and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42(4):1260–1288.

Besley, Timothy and Masayuki Kudamatsu. 2007. "Making Autocracy Work." Development Economics Discussion Paper Series 48, http://eprints.lse.ac.uk/3764/1/Making_Autocracy_Work.pdf.

Boix, Carles. 2003. *Democracy and Redistribution*. Cambridge University Press.

Boix, Carles and Milan W. Svolik. 2013. "The Foundations of Limited Authoritarian Government: Institutions, Commitment, and Power-Sharing in Dictatorships." *Journal of Politics* 75(2):300–316.

Boix, Carles and Susan C. Stokes. 2003. "Endogenous Democratization." *World Politics* 55(4):517–549.

Box-Steffensmeier, Janet M. and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. Cambridge University Press.

Box-Steffensmeier, Janet M. and Christopher Zorn. 2002. "Duration Models for Repeated Events." *Journal of Politics* 64(4):1069–1094.

Brown, Archie. 2007. *Seven Years that Changed the World: Perestroika in Perspective*. Oxford University Press.

Bueno de Mesquita, Bruce and Alastair Smith. 2009. "Political Survival and Endogenous Institutional Change." *Comparative Political Studies* 42(2):167–197.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: The MIT Press.

Bueno de Mesquita, Brue, James D. Morrow, Randolph M. Siverson and Alastair Smith. 2002. "Political Institutions, Policy Choice and the Survival of Leaders." *British Journal of Political Science* 32(4):59–590.

Bueno de Mesquita, Ethan. 2010. "Regime Change and Revolutionary Entrepreneurs." *American Political Science Review* 104(3):446–466.

Casper, Brett Allen and Scott A. Tyson. 2014. "Popular Protest and Elite Coordination in a Coup d'état." *Journal of Politics* 76(2):548–564.

Cheibub, José Antonio, Jennifer Gandhi and James Raymond Vreeland. 2010. "Democracy and Dictatorship Revisited." *Public Choice* 143:67–101.

Cho, In-Koo and David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *The Quarterly Journal of Economics* 102(2):179–222.

Dallin, Alexander. 1992. "Causes of the Collapse of the Soviet Union." *Post-Soviet Affairs* 8(4):279–302.

Di Lonardo, Livio and Scott A. Tyson. 2015. "Deterrence and Counterdeterrence in the Fight Against Global Terror." Presented at the Research Program in Political Economy Speaker Series, Princeton University, https://www.princeton.edu/rppe/speaker-series/TysonDET.pdf.

Easterly, William and Mirvat Sewadeh. 2001. "Global Development Network Growth Database." World Bank Economic and Development Research Group, Washington, D.C.

Edmond, Chris. 2013. "Information Manipulation, Coordination, and Regime Change." *Review of Economic Studies* 80(4):1422–1458.

Egorov, Georgy, Sergei Guriev and Konstantin Sonin. 2009. "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." *The American Political Science Review* 103(4):645–668.

Francois, Patrick, Ilia Rainer and Francesco Trebbi. 2014. "The Dictator's Inner Circle." NBER Working Paper 20216, http://www.nber.org/papers/w20216.

Fudenberg, Drew and Jean Tirole. 1991. *Game Theory*. The MIT Press.

Gandhi, Jennifer. 2008. *Political Institutions Under Dictatorship*. Cambridge University Press.

Gandhi, Jennifer and Adam Przeworski. 2006. "Cooperation, Cooptation, and Rebellion Under Dictatorships." *Economics and Politics* 18(1):1–26.

Gandhi, Jennifer and Adam Przeworski. 2007. "Authoritarian Institutions and the Survival of Autocrats." *Comparative Political Studies* 40:1279–1301.

Geddes, Barbara. 1999. "What Do We Know about Democratization after Twenty Years?" *Annual Review of Political Science* 2:115–144.

Geddes, Barbara, Joseph Wright and Erica Frantz. 2014. "Autocratic Breakdown and Regime Transitions: A New Data Set." *Perspectives on Politics* 12(2):313–331.

Gehlbach, Scott and Edmund Malesky. 2010. "The Contribution of Veto Players to Economic Reform." *Journal of Politics* 72(4):957–975.

Gehlbach, Scott and Konstantin Sonin. 2014. "Government Control of the Media." *Journal of Public Economics* 118:163–171.

Gehlbach, Scott and Philip Keefer. 2011. "Investment without Democracy: Ruling-Party Institutionalization and Credible Commitment in Autocracies." *Journal of Comparative Economics* 39(2):123–139.

Gehlbach, Scott and Philip Keefer. 2012. "Private Investment and the Institutionalization of Collective Action in Autocracies: Ruling Parties and Legislatures." *Journal of Politics* 74(2):621–635.

Goemans, Hein E. 2008. "Which Way Out? The Manner and Consequences of Losing Office." *Journal of Conflict Resolution* 52(6):771–794.

Gordon, Sanford C. 2002. "Stochastic Dependence in Competing Risks." *American Journal of Political Science* 46(1):200–217.

Guriev, Sergei and Daniel Treisman. 2015. "How Modern Dictators Survive: An Informational Theory of the New Authoritarianism." NBER Working Paper 21136, http://www.nber.org/papers/w21136.

Haggard, Stephan and Robert R. Kaufman. 1995. *The Political Economy of Democratic Transitions*. Princeton University Press.

Heston, Alan, Robert Summers and Bettina Aten. 2012. "Penn World Table Version 7.1." Center for International Comparisons of Production, Income and Prices at the Univeristy of Pennsylvania.

Hollyer, James R., B. Peter Rosendorff and James R. Vreeland. 2011. "Democracy and Transparency." *Journal of Politics* 73(4):1–15.

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2014. "Measuring Transparency." *Political Analysis* 22(4):413–434.

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2015*a*. "Transparency, Protest and Autocratic Instability." *American Political Science Review* 109(4):764–784.

Hollyer, James R., B. Peter Rosendorff and James Raymond Vreeland. 2015*b*. "Transparency, Protest and Democratic Stability." Presented at the 2013 American Political Science Association Conference, http://www.tc.umn.edu/~jhollyer/HRVDemStabilityDraft5.pdf.

Keele, Luke. 2010. "Proportionally Difficult: Testing for Nonproportional Hazards in Cox Models." *Political Analysis* 18(2):189–205.

King, Gary, Jennifer Pan and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107(2):326–343.

King, Gary, Jennifer Pan and Margaret E. Roberts. 2014. "Reverse-Engineering Censorship in China: Randomized Experimentation and Participant Observation." *Science* 345(6199):1–10.

Kotz, David M. and Fred Weir. 1997. *Revolution from Above: The Demise of the Soviet System*. Routledge.

Little, Andrew T., Tom LaGatta and Joshua A. Tucker. 2015. "Elections, Protest and Alternation of Power." *Journal of Politics* 77(4):1142–1156.

Lorentzen, Peter. 2014. "China's Strategic Censorship." *American Journal of Political Science* 58(2):402–414.

Morris, Stephen and Hyun Song Shin. 2002. "Social Value of Public Information." *American Econmic Review* 92(5):1521–1534.

Myerson, Roger B. 2008. "The Autocrat's Credibility Problem and Foundations of the Constitutional State." *The American Political Science Review* 102(1):125–139.

Padró i Miquel, Gerard. 2007. "The Control of Politicians in Divided Societies: The Politics of Fear." *Review of Economic Studies* 74:1259–1274.

Pepinsky, Thomas. 2014*a*. "Authoritarianism: Logics and Institutions." *APSA Comparative Politics Newsletter* 24(1):3–4.

Pepinsky, Thomas. 2014*b*. "The Institutional Turn in Comparative Authoritarianism." *British Journal of Political Science* 44(3):631–653.

Plümper, Thomas and Vera E. Troeger. 2007. "Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed-Effects." *Political Analysis* 15(2):124–139.

Przeworski, Adam and Fernando Limongi. 1997. "Modernization: Theories and Facts." *World Politics* 49.2:155–183.

Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. Cambridge University Press.

Reddaway, Peter and Dmitri Glinski. 2001. *The Tragedy of Russia's Reforms: Market Bolshevism Against Democracy*. United States Institute of Peace Press.

Shadmehr, Mehdi and Dan Bernhardt. 2014. "Vanguards in Revolution: Sacrifice, Radicalism and Coercion." http://moya.bus.miami.edu/~mshadmehr/VanguardsInRevolution.pdf.

Shadmehr, Mehdi and Dan Bernhardt. 2015. "State Censorship." *American Economic Journal: Microeconomics* 7(2):280–307.

Shih, Victor, Wei Whan and Mingxing Liu. 2010. *Contemporary Chinese Politics: New Sources, Methods, and Field Strategies*. Cambridge University Press chapter The Central Committee Past and Present: A method of Quantifying Elite Biographies, pp. 51–68.

Slater, Dan. 2009. "Revolutions, Crackdowns, and Quiessence: Communal Elites and Democratic Mobilization in Southeast Asia." *American Journal of Sociology* 115(1):203–254.

Stone, Randall W. 2008. "The Scope of IMF Conditionality." *International Organization* 62(4):589–620.

Svolik, Milan. 2013*a*. "Incentives, Institutions, and the Challenges to Research on Authoritarian Politics." *APSA Comparative Democratization Newsletter* 11(2):7–11.

Svolik, Milan W. 2012. *The Politics of Authoritarian Rule*. Cambridge University Press.

Svolik, Milan W. 2013*b*. "Contracting on Violence: The Moral Hazard in Authoritarian Repression and Military Intervention in Politics." *Journal of Conflict Resolution* 57(5):765–794.

Wallace, Jeremy L. 2015. "Information Politics in Dictatorships." *Emerging Trneds in the Social and Behaviorial Science: An Interdisciplinary, Searchable, and Linkable Resource* pp. 1–11. DOI: `10.1002/9781118900772.etrds0181`.

Wawro, Gregory. 2002. "Estimating Dynamic Panel Data Models in Political Science." *Political Analysis* 10(1):25–48.

Whitefield, Stephen. 1993. *Industrial Power and the Soviet State*. Oxford University Press.

A   Proofs

Lemma 1. *In any subgame perfect equilibrium, after elite has observed action d by leader, elite R removes the leader (sets $v = 1$) iff $r \leq r(d)$*

Proof.

$$EU^R(1; d) \geq EU^R(0; d) \qquad\qquad \Longleftrightarrow$$

$$(1 - \omega p(d))\lambda y \geq (1 - p(d))(1 - \lambda)y(1 + r) \qquad\qquad \Longleftrightarrow$$

$$\lambda - \lambda\omega p(d) \geq 1 - \lambda - p(d) + \lambda p(d) + r(1 - \lambda)(1 - p(d)) \qquad\qquad \Longleftrightarrow$$

$$2\lambda - 1 - p(d)\,(\lambda\omega - 1 + \lambda) \geq r(1 - \lambda)(1 - p(d)) \qquad\qquad \Longleftrightarrow$$

$$r \leq \frac{2\lambda - 1 - p(d)\,(\lambda\omega - 1 + \lambda)}{(1 - \lambda)(1 - p(d))} = r(d)$$

$\square$

Definition 1. *Define $r(d) \equiv \frac{2\lambda - 1 - p(d)(\lambda\omega - 1 + \lambda)}{(1 - \lambda)(1 - p(d))}$ for $d = 0, 1$.*

Lemma 2. $r(1) < r(0)$ *since $p(1) > p(0)$ and $\omega > 1$.*

Proof. Straightforward $\square$

Definition 2. *Define $\tilde{\omega} \equiv \frac{2\lambda - 1 + p(0)(1 - \lambda)}{p(0)\lambda}$.*

Lemma 3. *If $\omega < \tilde{\omega}$ then $r(0) > 0$.*

Proof.

$$r(0) \;=\; \frac{2\lambda - 1 - p(0)\,(\lambda\omega - 1 + \lambda)}{(1 - \lambda)(1 - p(0))} > 0 \Longleftrightarrow$$

$$\omega \;<\; \frac{2\lambda - 1 + p(0)(1 - \lambda)}{p(0)\lambda} = \tilde{\omega}$$

since $(1 - \lambda)(1 - p(0)) > 0$. $\square$

Proposition 1. *If $\omega < \tilde{\omega}$ then the Nash Equilibrium to the game is:*
For any $r < r(1)$, $v = 1$ and $d = 1$;
For $r > r(0)$, $v = 0$ and $d = 0$;
For $r(1) \leq r \leq r(0)$, $d = 1$ and $v = 0$.

Proof. Recall from Lemma 2 that $r(1) < r(0)$. Lemma 3 establishes that $r(0) > 0$. However $r(1)$ may be positive or negative. If $r(1) < 0$ then there does not exist any $r > 0$ such that $r < r(1)$.

Suppose $0 < r < r(1)$. Then $r < r(d)$ for $d = 0, 1$, and by Lemma 1, $v = 1$. Then payoff for L is zero irrespective of disclosure. We set $(d = 1)$ when L is indifferent.

Suppose $r > r(0)$. Then $r > r(d)$ for $d = 0, 1$, and by Lemma 1, $v = 0$. Then for L, non-disclosure $(d = 0)$ dominates disclosure $(d = 1)$ iff

$$EU^L(d = 1; v = 0) \leq EU^L(d = 0, v = 0) \qquad \Longleftrightarrow$$
$$(1 - p(1))\lambda y(1 + r) \leq (1 - p(0))\lambda y(1 + r) \qquad \Longleftrightarrow$$
$$p(1) \geq p(0)$$

which is true by assumption.

For $r(1) \leq r \leq r(0)$, Lemma 1 implies $v = 1$ for $d = 0$ and $v = 0$ for $d = 1$. Then for L, in equlibrium, L plays $d = 1$, and receives in equilibrium $EU^L(d = 1; v = 0) = (1 - p(1))\lambda y(1 + r)$. If instead, L played $d = 0$, then R plays $v = 1$. Then disclosure $(d = 1)$ dominates non-disclosure $(d = 0)$ iff

$$EU^L(d = 1; v = 0) \geq EU^L(d = 0, v = 1) \qquad \Longleftrightarrow$$
$$(1 - p(1))\lambda y(1 + r) \geq 0 \qquad \Longleftrightarrow$$
$$r \geq -1$$

Since r is bounded below by zero, this is always true. $\square$

Note that if $\frac{1}{p(0)} > \omega > \tilde{\omega}$, then there is no part of the parameter space in which, in equilibrium, the leader discloses with the intent to forestall a coup.

Proposition 2. *If $\omega \geq \tilde{\omega}$ then the Nash Equilibrium to the game is:*
For all r, $v = 0$ and $d = 0$.

Proof. If $\omega \geq \tilde{\omega}$ then by Lemma 3, $r(0) \leq 0$. Recall that $r \sim G$ over domain $(0, \infty)$, so $r > 0 \geq r(0)$. Then by Lemma 1, $v = 0$, and a best response to $v = 0$ is always $d = 0$. $\square$

Proposition 3. *Disclosure is less likely when w rises.*

Proof. Following from Proposition 1, $d = 1$ when $r \in [0, r(0)]$. Then $\frac{\partial r(0)}{\partial w} = \frac{-p(0)\lambda}{(1 - \lambda)(1 - p(0))} < 0$, and $\frac{\partial G(r(0))}{\partial \omega} = G' \frac{\partial r(0)}{\partial \omega} < 0$. $\square$

B The Model with Accountability Microfoundations

This version of the model builds on the theoretical framework introduced by Besley and Kudamatsu (2007), who construct a model of autocratic accountability involving both moral hazard and adverse selection.[22]

The game proceeds for two periods. In the first period, L must decide whether to disclose $d \in \{0, 1\}$, and in each period he must make a policy decision $e_t \in \{0, 1\}$, where t denotes the period of play $t \in \{1, 2\}$. Regime elites R, after witnessing the outcome of the leader's policy decision, must determine whether to retain him in office. We denote this decision $v \in \{0, 1\}$, where $v = 1$ denotes a decision to remove. Leaders may be of one of two types $\theta \in \{0, 1\}$. When $\theta = 1$, leaders are 'convergent' – they have a primitive preference for setting e_t equal to the value that maximizes the welfare of the group in power. When $\theta = 0$, leaders are 'divergent' – they have a primitive preference for setting e_t equal to a value viewed as suboptimal by the group in power. The prior probability that L is a convergent type is defined as π ($Pr(\theta = 1) = \pi$).

Following the conclusion of the first period of play, a contest for power takes place between members of the regime and the populace. If members of the regime chose to keep L in place, the probability that the regime falls is given by $p(d)$, where $1 > p(1) > p(0) > 0$. Define the effect of disclosure on mobilizational capacity as $\rho \equiv p(1) - p(0)$. If L was removed, either by members of the regime or by the victory of the populace, we assume a new leader is selected from the same distribution as described above (i.e., $Pr(\theta = 1) = \pi$). If the regime is overthrown, we assume M takes over as a new regime. We make a small change to the earlier setup: both members of the elite and leader enjoy λy if they survive in office and $\lambda \in (\frac{1}{2}, 1)$ reflects the disproportionate flow of resources to members of the regime.

Members of the regime also want L to set policy equal to a state variable $s_t \in \{0, 1\}$, which is randomly determined and where $Pr(s_t = 1) = \frac{1}{2} \ \forall \ t$, and s_t is independently drawn in each period. R's utility is thus given by:

$$u_{R,t}(e_t, s_t) = \begin{cases} I_t[\Delta + \lambda y] + (1 - I_t)(1 - \lambda)y \text{ if } e_t = s_t \\ I_t \lambda y + (1 - I_t)(1 - \lambda)y \text{ otherwise} \end{cases} \tag{3}$$

where $\Delta > 0$ and $I_t \in \{0, 1\}$ is an indicator variable equal to 1 if R is in power at time t and equal to zero if M is in power. Note that R always begins the game in power, hence $I_1 = 1$, whereas I_2 is determined by the contest for power that takes place at the end of the first period of play.

Those out of power have no ability to directly sanction the autocratic leader. Hence, their preferences over the policy variable e_t are irrelevant. We treat the utilities of those out of power as independent of e_t to emphasize that accountability, in the autocratic context, means accountability to the winning coalition of the regime.[23] If $I_2 = 0$, therefore, R has been removed from power, her share of national income declines

---

[22]This approach is also closely analogous to Padró i Miquel (2007).

[23]Our results would be substantively unchanged if those in and out of power both received this state-

to $(1 - \lambda)y$ and she is no longer concerned about the value of the policy variable.

Members of the populace derive utility only based on their share of national income $(1 - \lambda)y$ in the first period of play. Should they come to power in the second period, they will gain in their share of national income – which will rise to λy – and they will become concerned about the leader's choice of policy. The utility of the populace is thus given by:

$$u_{M,t}(e_t, s_t) = \begin{cases} I_t(1 - \lambda)y + (1 - I_t)[\Delta + \lambda y] \text{ if } e_t = s_t \\ I_t(1 - \lambda)y + (1 - I_t)\lambda y \text{ otherwise} \end{cases} \tag{4}$$

where I_t is as defined above. Note that this implies that the 'convergence' or 'divergence' of the leader is with reference to *whichever* group is in power at time t.

L's utility depends on his type. Convergent types ($\theta = 1$) share the policy preferences of the members of those elites in power (hence 'convergent'). Divergent types ($\theta = 0$), however, have primitive preferences opposed to sitting regime members. Specifically, we assume that divergent types derive a utility of r_t from setting $e_t \neq s_t$, where r_t is a random variable drawn from a CDF $G(\cdot)$. $G(\cdot)$ has support on $\mathbb{R}_+$ such that $G(\Delta) = 0$ (i.e., $r_t \in (\Delta, \infty)$) and has expected value μ. Leaders also derive utility from the share of national income flowing to the elite. Thus, leader utility is given by:

$$u_{L,t}(e_t, s_t; \theta) = \begin{cases} \Delta + \lambda y \text{ if } e_t = s_t \text{ and in power} \\ \lambda y \text{ if } e_t \neq s_t, \ \theta = 1 \text{ and in power} \\ r_t + \lambda y \text{ if } e_t \neq s_t, \ \theta = 0 \text{ and in power} \\ 0 \text{ if out of power.} \end{cases} \tag{5}$$

The order of play is as follows:

1. $Nature$ draws the the leader's type $\theta \in \{0, 1\}$, the state variable s_1 and the value of rents r_1, which are revealed to the leader but not to any citizen.

2. The leader chooses $d \in \{0, 1\}$ and the value of e_1

3. Members of the regime observe the choice of d and the realization of the policy outcome. They choose whether to unseat the leader $v \in \{0, 1\}$.

4. A contest for power between R and M takes place. M prevails with probability $p(d)$ if the leader was previously retained and with probability $\omega p(d)$ if the leader was previously removed.

5. (a) If M prevails, it is in power in round 2 and a new leader is chosen by $Nature$. This leader is of type $\theta = 1$ with probability π.

   (b) If R prevails after ousting the leader, a new leader is chosen by $Nature$. This leader is of type $\theta = 1$ with probability π.

---

contingent payoff.

(c) Otherwise, L remains in office.

6. $Nature$ chooses values of s_2 and r_2, which are revealed to the sitting leader, but not to any other player.

7. The sitting leader chooses e_2. All payoffs are realized and the game ends.

B.1 Equilibrium

We characterize a perfect Bayesian equilibrium (PBE) to this interaction (Fudenberg and Tirole, 1991). As is common in signaling games, this interaction can give rise to multiple such equilibria. We focus attention on a semi-separating PBE in pure strategies. We also restrict attention to equilibria in which L chooses $d = 1$ whenever indifferent over this choice. When we restrict attention to PBE in which L, when indifferent, chooses $d = 1$, this semi-separating equilibrium is the unique pure strategy PBE to satisfy an intuitive criterion refinement (Cho and Kreps, 1987).

A PBE in pure strategies consists of a strategy profile and a set of posterior beliefs for members of the regime. A strategy for L consists of a choice of an action pair (e_1, d) in the first period of play, where (e_1, d) is a mapping from his type θ and the realization of the value of rents r_1, $(e_1, d) : \{0, 1\} \times (\Delta, \infty) \to \{0, 1\} \times \{0, 1\}$. His strategy further consists of an action e_2 which is a mapping from his type, $e_2 : \{0, 1\} \to \{0, 1\}$. A strategy for R consists of a mapping from her posterior beliefs over L's type and the choice of disclosure d into her decision of whether to oust the leader $v : [0, 1] \times \{0, 1\} \to \{0, 1\}$. And R's posterior beliefs are updated according to Bayes' Rule and are consistent with the strategy profile.[24]

To characterize a semi-separating PBE, we first require several definitions. We first implicitly define two thresholds in ω:

Definition 3. *Define $\bar{\omega}$ and $\underline{\omega}$ such that:*

$$\pi\Delta = \frac{p(0)y(0)(\bar{\omega} - 1)(2\lambda - 1)}{1 - \bar{\omega}p(0)}$$
$$\pi\Delta = \frac{p(1)y(1)(\underline{\omega} - 1)(2\lambda - 1)}{1 - \underline{\omega}p(1)}.$$

Note that the right hand side of the equality $\pi\Delta = \frac{p(d)y(\omega-1)(2\lambda-1)}{1-\omega p(d)}$, which defines these two terms, is monotonic and increasing in ω and similarly monotonic and increasing in $p(d)$ over all admissible values of these terms. Note further that $\lim_{\omega \to 1} \frac{p(d)y(\omega-1)(2\lambda-1)}{1-\omega p(d)} = 0$, and $\lim_{\omega \to \frac{1}{p(d)}} \frac{p(d)y(\omega-1)(2\lambda-1)}{1-\omega p(d)} = \infty$. Thus, $\bar{\omega}$ and $\underline{\omega}$ are well defined. Moreover, since the expression is monotonically increasing and continuous in p, and $p(1) > p(0)$, $\bar{\omega} > \underline{\omega}$.

---

[24] M is non-strategic in this game, given that the interaction terminates after two periods. Even if M comes to power, members of the new regime cannot remove their sitting leader from power.

We further define a threshold in r_1, which we term $\bar{r}$. This threshold will be a function of the parameter ω.

Definition 4. *Define $\bar{r}(\omega)$ such that:*

$$\bar{r}(\omega) = \begin{cases} \Delta + [1 - p(0)][\mu + \lambda y] \text{ if } \omega < \underline{\omega} \\ \Delta + \rho[\mu + \lambda y] \text{ if } \omega \in [\underline{\omega}, \bar{\omega}] \\ \Delta \text{ if } \omega > \bar{\omega}. \end{cases}$$

We can now characterize the semi-separating PBE to this game in the following proposition.

Proposition 4. *A semi-separating PBE to this game consists of the following strategies and beliefs:*

1. *For L:*

$$(e_1, d) = \begin{cases} (\neg s_1, 1) \text{ if } r_1 \geq \bar{r}(\omega), \ \omega \leq \bar{\omega} \text{ and } \theta = 0 \\ (\neg s_1, 0) \text{ if } r_1 \geq \bar{r}(\omega), \ \omega > \bar{\omega} \text{ and } \theta = 0 \\ (s_1, 0) \text{ otherwise.} \end{cases}$$

$$e_2 = \begin{cases} \neg s_2 \text{ if } \theta = 0 \\ s_2 \text{ otherwise.} \end{cases}$$

2. *For R:*

$$v = \begin{cases} 0 \text{ if } \omega > \bar{\omega} \\ 0 \text{ if } \omega > \underline{\omega} \text{ and } d = 1 \\ 0 \text{ if } (e_1, d) = (s_1, 0) \\ 1 \text{ otherwise.} \end{cases}$$

3. *and R's beliefs are given (with some abuse of notation) by $Pr(\theta = 1 | e_1 = s_1, d = 0) > \pi$ and $Pr(\theta = 1 | e_1, d) = 0$ for all other realizations of (e_1, d).*

Proof. A PBE, in pure strategies, consists of (1) a strategy profile in which all actors adopt best responses given their beliefs, and (2) a set of beliefs that is weakly consistent with the strategy profile and updated according to Bayes rule, wherever possible (Fudenberg and Tirole, 1991). A semi-separating PBE in pure strategies to this game involves a first period strategy for L in which, for some range of realizations of r_1, different values of θ lead to different signals (e_1, d).

We begin via backward induction. In the final period of play, L has a dominant strategy, conditional on his type θ. $e_2 = s_2$ if $\theta = 1$ and $e_2 \neq s_2$ if $\theta = 0$.

Given this strategy for all types of L, we can now characterize R's decision over $v \in \{0, 1\}$. If L is retained, and the regime survives, R receives an expected utility of $Pr(\theta = 1 | e_1, d)\Delta + \lambda y$. However, the regime only survives with probability $1 - p(d)$. With probability $p(d)$ the regime collapses, in which case R is

31

guaranteed a utility of $(1-\lambda)y$. Hence, the expected utility of retention is $p(d)(1-\lambda)y + [1-p(d)][Pr(\theta = 1|e_1, d)\Delta + \lambda y]$.

If R sets $v = 1$ and the regime survives, she receives expected utility of $\pi\Delta + \lambda y$, where π denotes the probability with which Nature chooses $\theta = 1$. However, following $v = 1$, the regime survives with probability $\omega p(d)$. Hence, the expected utility of removal is $\omega p(d)(1-\lambda)y + [1-\omega p(d)][\pi\Delta + \lambda y]$.

Comparing these two utilities yields expression 6 in the text. $v = 1$ iff:

$$\omega p(d)(1-\lambda)y + [1-\omega p(d)][\pi\Delta + \lambda y] > p(d)(1-\lambda)y + [1-p(d)][Pr(\theta = 1|e_1, d)\Delta + \lambda y]$$

Notice that the RHS of expression 6 is monotonic and increasing in $Pr(\theta = 1|e_1, d)$, while the LHS is invariant in this term. Therefore, if the expression fails to hold when $Pr(\theta = 1|e_1, d) = 0$, it will never hold.

Substituting $Pr(\theta = 1|e_1, d) = 0$ and rearranging yields:

$$\pi\Delta > \frac{p(d)y(\omega - 1)(2\lambda - 1)}{1 - \omega p(d)}$$

which defines the two thresholds $\bar{\omega}$ and $\underline{\omega}$ in Definition 3. If $\omega > \bar{\omega}$ this expression can never hold for any value of d, so R has a strictly dominant strategy of setting $v = 0$. If $\omega > \underline{\omega}$, the expression can never hold so long as $d = 1$. So, R must respond to $d = 1$ by setting $v = 0$.

Notice further that, given $\omega > 1$ and $\lambda > \frac{1}{2}$, the inequality in expression 6 can never hold for $Pr(\theta = 1|e_1, d) \geq \pi$. Hence, for any $Pr(\theta = 1|e_1, d) \geq \pi$, R must set $v = 0$.

Given these preliminaries, we can now consider L's strategy in the first period of play. Clearly, for $\omega > \bar{\omega}$, L has a dominant strategy, conditional on type. For $\theta = 1$, it must be the case that $(e_1, d) = (s_1, 0)$. Analogously, for $\theta = 0$, it must be the case that $(e_1, d) = (\neg s_1, 0)$.

Consider, now, values of $\omega < \bar{\omega}$. Let L play a strategy of $(e_1, d) = (s_1, 0)$ when $\theta = 1$ – as is consistent with his primitive preference. Given this strategy when $\theta = 1$, R must hold beliefs such that $Pr(\theta = 1|e_1 = s_1, d = 0) \geq \pi$, hence sending this signal pair guarantee's $v = 0$ for all types of L.

We must now consider the strategy for L when $\theta = 0$. Let us first consider the case when $\omega \in [\underline{\omega}, \bar{\omega}]$. He can guarantee retention by adopting the action pair $(s_1, 0)$, in which case his expected utility is given by $\Delta + \lambda y + [1 - p(0)][\mu + \lambda y]$. Notice that, for $\omega \in [\underline{\omega}, \bar{\omega}]$, a divergent L may also guarantee retention by setting $d = 1$. If he does this, he also strictly prefers to follow his primitive preference over policy by setting $e_1 \neq s_1$. (Analogously, if he chooses to separate by setting $e_1 \neq s_1$, he strictly prefers to set $d = 1$ and guarantee $v = 0$ rather than $v = 1$.) By setting $(e_1, d) = (\neg s_1, 1)$, L obtains $r_1 + \lambda y + [1 - p(1)][\mu + \lambda y]$. Hence, for $\theta = 0$ and $\omega \in [\underline{\omega}, \bar{\omega}]$, L chooses $(\neg s_1, 1)$ if $r_1 > \Delta + \rho[\mu + \lambda y]$, where $\rho \equiv p(1) - p(0)$, and chooses $(s_1, 0)$ otherwise.

Let us now consider the case when $\omega < \underline{\omega}$. Here, L can guarantee retention by adopting $(s_1, 0)$, but faces certain removal for any other signal pair. His expected utility from $(s_1, 0)$ is identical to that given above. His utility from any alternative in which $e_1 \neq s_1$ is given by $r_1 + \lambda y$. (Clearly, the signal $(s_1, 1)$ is dominated.) L thus prefers to choose $e_1 \neq s_1$ if $r_1 > \Delta + [1 - p(0)][\mu + \lambda y]$, and to choose $(s_1, 0)$

32

otherwise. L is indifferent between $(\neg s_1, 1)$ and $(\neg s_1, 0)$. As noted above, we assume $d = 1$ whenever L is indifferent, such that L adopts the signal $(\neg s_1, 1)$ when $r_1 > \Delta + [1 - p(0)][\mu + \lambda y]$, $\theta = 0$ and $\omega < \underline{\omega}$.

We can summarize these cut-points in r_1 by $\bar{r}(\omega)$, as defined in Definition 4. Given this strategy by L, R must believe $Pr(\theta = 1 | e_1 = s_1, d = 0) > \pi$ and, when $\omega > \bar{\omega}$, must believe $Pr(\theta = 1 | e_1 \neq s_1, d = 0) = 0$, when $\omega < \bar{\omega}$, must believe $Pr(\theta = 1 | e_1 \neq s_1, d = 1) = 0$. Other messages remain off the path of play and beliefs are unrestricted. We set these beliefs to be $Pr(\theta = 1 | e_1, d) = 0$ for all off-path messages (e_1, d). Given these beliefs, R's best response is to set $v = 0$ when $(e_1, d) = (s_1, 0)$. As noted above, her best response is also to set $v = 0$ whenever $\omega > \bar{\omega}$ or $\omega \geq \underline{\omega}$ and $d = 1$. R's best response is to set $v = 1$ in all other cases. This completes the characterization of the semi-separating equilibrium. □

Proof that the Semi-Separating Equilibrium Uniquely Satisfies the Intuitive Criterion. This proof consists of two requirements. First, we must demonstrate that the semi-separating equilibrium satisfies the intuitive criterion refinement. Second, we must demonstrate that no alternative pure strategy PBE, in which L discloses whenever indifferent, satisfies the intuitive criterion refinement. We take each of these steps in turn.

We begin by demonstrating that the PBE characterized by Proposition 4 satisfies the intuitive criterion refinement. There are information sets that are never hit in equilibrium, though which information sets are not hit depends on the value of the parameter ω. For $\omega > \bar{\omega}$, the messages $(s_1, 0)$ and $(\neg s_1, 0)$ are observed in equilibrium, while $(s_1, 1)$ and $(\neg s_1, 1)$ are never observed. Trivially, however, both off-path messages are equilibrium dominated for both types of L – since, when $\omega > \bar{\omega}$, both types of L have a dominant strategy. Hence, the intuitive criterion does not limit off equilibrium path beliefs, and the equilibrium survives.

For $\omega < \bar{\omega}$, the messages $(s_1, 0)$ and $(\neg s_1, 1)$ are observed in equilibrium, while the messages $(s_1, 1)$ and $(\neg s_1, 0)$ are not. The message $(s_1, 0)$ equilibrium dominates both off path messages for convergent types of L, since $(s_1, 0)$ returns the highest possible utility for L when $\theta = 1$. Given that this is the case, either the intuitive criterion (1) does not limit beliefs for these off-path messages (the off-path message is equilibrium dominated for all types of L) or (2) the intuitive criterion requires that these off-path beliefs are such that $Pr(\theta = 1 | e_1, d) = 0$ (the off-path message is equilibrium dominated when $\theta = 1$, but not when $\theta = 0$). Given that we specify $Pr(\theta = 1 | e_1, d) = 0$ for $(e_1, d) = \{(s_1, 0), (\neg s_1 1)\}$ in the semi-separating equilibrium, these beliefs are consistent with the intuitive criterion refinement.

We now must demonstrate that alternative pure strategy PBE, in which L discloses whenever indifferent, do not satisfy an intuitive criterion refinement. To accomplish this task, we must first characterize the alternative pure strategy PBE to this game.

First, notice that there cannot exist any PBE in which divergent types adopt a pure strategy of setting $e_1 = s_1$. No such equilibrium can exist because, for a sufficiently high realization of the rents term r_1, divergent types have a dominant strategy of setting $e_1 \neq s_1$ – i.e., they will do so regardless of R's beliefs.

Notice further that, for $\omega > \bar{\omega}$, both types of L have a dominant strategy. In any PBE, it must be the case that for $\omega > \bar{\omega}$, L chooses $(s_1, 0)$ when $\theta = 1$ and $(\neg s_1, 0)$ when $\theta = 0$. We will therefore confine

our attention to equilibrium strategies when $\omega < \bar{\omega}$.

These restrictions rule out any equilibrium in which all types of L pool on the action $e_1 = s_1$ when $\omega < \bar{\omega}$. The intuitive criterion rules out any equilibrium in which, for some range of values of ω, both types of leader pool on the either the message $(\neg s_1, 1)$ or $(\neg s_1, 0)$. For any range of parameter values such that both types of L pool on $(e_1, d) \in \{(\neg s_1, 1), (\neg s_1, 0)\}$, R's posterior belief on witnessing this signal must be such that $Pr(\theta = 1|e_1, d) = \pi$, which, in turn, guarantees that her best response to this message is $v = 0$. Since either equilibrium message then satisfies L's preferences over both policy and retention when $\theta = 0$, either message will equilibrium dominate any off path message $(e_1 = s_1, d = \cdot)$ when $\theta = 0$. However, the eqiulibrium message cannot equilibrium dominate any off path message $(e_1 = s_1, d = \cdot)$ for convergent types, since $e_1 = s_1$ satisifies L's policy preference when $\theta = 1$. Thus, the intuitive criterion requires that R hold off-path beliefs $Pr(\theta = 1|e_1 = s_1, d = \cdot) = 1$. Given these off-path beliefs, convergent types prefer to deviate from the equilibrium – i.e., any such equilibrium must fail the intuitive criterion refinement.

There remains one additional scenario to consider – an alternative semi-separating PBE in which convergent types adopt the strategy of sending the message $(s_1, 1)$. Divergent types pool with this message for values of $r_1 < \bar{r}(\omega)$ when $\omega < \underline{\omega}$, and will deviate to $(\neg s_1, 1)$ otherwise, where $\bar{r}(\omega)$ is as defined in Definition 4. In this equilibrium, R's beliefs are defined such that $Pr(\theta = 1|e_1 = s_1, d = 1) > \pi$ and $Pr(\theta = 1|e_1, d) = 0$ for any other message pair. Given this, R's strategy is to set $v = 0$ if $(e_1, d) = (s_1, 1)$; $v = 0$ if $d = 1$ and $\omega > \underline{\omega}$; and $v = 1$ otherwise.

In this equilibrium, the off-path message $(s_1, 0)$ is equilibrium dominated for divergent types for all $\omega > \underline{\omega}$ and when $\omega < \underline{\omega}$ and $r_1 \geq \bar{r}(\omega)$. This off-path message is not, however, equilibrium dominated for convergent types. Hence, R's beliefs on witnessing $(s_1, 0)$ must be such that $Pr(\theta = 1|s_1, 0) > \pi$ – and R must respond to these belief by setting $v = 0$ on witnessing $(s_1, 0)$. However, if R adopts this strategy, convergent types of L strictly prefer to deviate from the equilibrium, and send the message $(s_1, 0)$. The equilibrium does not survive an intuitive criterion refinement.

We have thus considered all alternative pure strategy PBE, in which L discloses where indifferent, to the game. The semi-separating equilibrium characterized by Proposition 4 survives the intuitive criterion refinement. Alternative pure strategy PBE do not. Hence, the equilibrium characterized by Proposition 4 is the unique pure strategy PBE to satisfy the intuitive criterion. $\qquad \square$

B.2 Intuitions

In the final period of play, the sitting L sets his policy decision e_2 according to type. That is $e_2 = s_2$ if $\theta = 1$ and $e_2 \neq s_2$ if $\theta = 0$. This is a dominant strategy.

Given this strategy by the leader, R must make her decision regarding retention weighing her expectations about L's future policy choice and the consequences of removal for regime survival. R can update her beliefs based on the leader's previous decisions over policy and disclosure (e_1, d), and we denote these beliefs, with some abuse of notation, as $Pr(\theta = 1|e_1, d)$. Thus, R can expect policy returns of $Pr(\theta = 1|e_1, d)\Delta$ from retaining the leader, and returns of $\pi\Delta$ from replacing the leader with an alterna-

tive. Removal, however, also has consequences for regime survival. If the leader is retained, the regime is removed with probability $p(d)$ – if he is removed, the regime is ousted with probability $\omega p(d)$. Weighing these various concerns, we can state that R sets $v = 1$ iff:

$$\omega p(d)(1-\lambda)y + [1 - \omega p(d)][\pi\Delta + \lambda y] > p(d)(1-\lambda)y + [1 - p(d)][Pr(\theta = 1|e_1, d)\Delta + \lambda y]. \quad (6)$$

Because of the adverse consequences of leader removal for the survival of the regime, circumstances can arise under which R chooses to retain their leader even if certain that the leader is a divergent type. That is, $v = 0$ even if $Pr(\theta = 1|e_1, d) = 0$. The risks incurred by removing the leader simply aren't worth the expected benefits $(\pi\Delta)$ from securing a more desirable dictator. Substituting $Pr(\theta = 1|e_1, d) = 0$ into expression 6 and simplifying yields the values of $\bar{\omega}$ and $\underline{\omega}$ as defined in Definition 3.

So, for $\omega > \bar{\omega}$, regardless of the actions of L in the first period, regime elites choose $v = 0$. The risks of removing the leader simply aren't worth any possible policy gains that might be derived from doing so. We refer to autocracies with values of $\omega > \bar{\omega}$ as either *personalistic* or *entrenched*.

For $\omega \in [\underline{\omega}, \bar{\omega}]$, leaders can guarantee their retention if they disclose, but are not offered any such guarantee otherwise. Because disclosure heightens the risk of public mobilization, it also renders ousting the leader more risky. Elites are thus less willing to remove when $d = 1$ than when $d = 0$.

We can now consider L's decision over policy and disclosure (e_1, d) in the first period of play. When $\omega > \bar{\omega}$, L can make this decision free of any consideration of the regime's response. Each type of L thus acts on his primitive preference – $e_1 = s_1$ if $\theta = 1$ and not otherwise. Similarly, the leader can act on his primitive preference over disclosure, which is always not to disclose $(d = 0)$. In the model, disclosure acts only to increase the mobilizational capacity of the populace. For leaders who are already secure from elite accountability, disclosure offers only a cost with no benefit.

In a semi-separating equilibrium, a convergent type of leader also acts on his primitive preferences – setting $e_1 = s_1$ and $d = 0$ – for all values of $\omega < \bar{\omega}$. So long as elites correctly interpret this behavior as a positive signal of L's type, this maximizes a convergent leader's utility. The question, then, is whether a divergent type would choose to pool with convergent and set $e_1 = s_1$ and $d = 0$, or choose to separate. A divergent leader will base this decision on the realization of the rents he receives from deviating from the preferences of the elite, i.e., r_1.

Recall that a divergent type earns rents r_1 from setting $e_1 \neq s_1$, where r_1 is a random variable and $r_1 > \Delta$. Hence, divergent types always have some incentive to separate from convergent, and – for a sufficiently high realization of r_1 – this will be a dominant strategy. Since convergent types set $(e_1, d) = (s_1, 0)$, whereas divergent types may choose not to do so, $Pr(\theta = 1|e_1 = s_1, d = 0) > \pi$. Substituting these beliefs into expression 6 reveals that R always strictly prefers to retain given the action pair $(s_1, 0)$. A divergent type thus knows he will be retained with certainty by mimicking convergent types.

The value of rents necessary to induce a divergent type to separate is given by $\bar{r}(\omega)$. For $r_1 > \bar{r}(\omega)$, the leader's policy gains from defying his winning coalition outweigh the benefits from certain retention by the elite. When a divergent leader chooses to so defy his elite, he also chooses to disclose. This is

because, for $\omega \in [\underline{\omega}, \bar{\omega}]$ disclosure guarantees the elite's quiescence. The leader is able to free himself from elite accountability via disclosure. For $\omega < \underline{\omega}$, disclosure does not insulate the leader from regime backlash. But, since the leader anticipates that the elite will mobilize for his removal, the consequences of disclosure will only be felt by his successor. The leader is indifferent between disclosing and not, and – in an act of schadenfreude – punishes his winning coalition for their disloyalty.

Combining these strategies and beliefs, we are left with PBE characterized by Proposition 4. Convergent types always act according to their primitive preferences over both policy and disclosure, which are also the preferences of the elite. Divergent types only act according to their primitive preferences over both disclosure and policy if they are personalistic and entrenched. If not, they may choose to act on their preference over policy if the rents from doing so are sufficiently high. However, if they so separate themselves from convergent types, they must also choose to disclose. Members of the elite, responding to these equilibrium strategies, oust the leader when $(e_1, d) \neq (s_1, 0)$ and $\omega < \underline{\omega}$, and choose to retain otherwise.

B.3 Comparative Statics

We advance two empirical claims based on our theory: The first pertains to leader removal – leaders who disclose more readily should be at a reduced threat of removal via coup. The second examines the circumstances under which autocratic leaders choose to disclose – we should witness less disclosure when autocratic institutions are personalistic and when the leadership is long entrenched, and greater disclosure when the regime is institutionalized (not personalistic) and the leadership is new to office.

Our claims regarding the insulating effects of disclosure, however, cannot be examined via a conventional comparative static. The decision to disclose is endogenous, and hence cannot be manipulated like an exogenous parameter. This is true despite the fact that, in a partial equilibrium analysis, it is clear that R's decision over removal is influenced by disclosure, such that she is less likely to set $v = 1$ when $d = 1$.

We instead, therefore, compare the equilibrium defined in Proposition 4 to an analogous semi-separating equilibrium in a game isomorphic to that above, save only that L does not possess the option to disclose. Practically, autocrats face a number of obstacles to increasing disclosure, particularly within a short time-frame. The disclosure of economic information requires non-trivial amounts of bureaucratic expertise, which must be amassed over time, in addition to large-scale data collection efforts. In short, we might imagine that autocratic leaders vary in their capacity to disclose – or to increase levels of disclosure (for an extensive discussion of the relationship between capacity and transparency, see Hollyer, Rosendorff and Vreeland, 2014). In the following proposition, we demonstrate that, when autocratic leaders possess a greater capacity to disclose – and hence disclose more frequently – they survive for a grater range of parameter values than would be the case where they cannot disclose.

Proposition 5. *In a semi-separating equilibrium to a model without disclosure, when* $\omega \in [\underline{\omega}, \bar{\omega}]$ *and* $r_1 > \Delta + [1 - p(0)][\mu + \lambda y]$, *divergent types of* L *are removed by the elite with certainty. For the same*

set of parameter values, in a semi-separating equilibrium where disclosure is possible, divergent types of
L are retained with certainty and choose $d = 1$.

Proof. To complete this proof, we first must construct an analogous semi-separating equilibrium to that specified in Proposition 4 to a game isomorphic to that above, save that disclosure is not an option. That is, the game is identical, except that d is fixed and equal to zero.

As in the original game, L has a dominant strategy of playing according to type in the final period of play. Given this, R's retention decision is identical to the original model, and her best response is given by expression 6. As in the original game, we can therefore characterize a value of $\bar{\omega}$ to the alternative game, and this value is given by Definition 3. There is no analogue to the value $\underline{\omega}$, however.

We can now consider L's strategy regarding his policy decision in the firs period of play e_1. This is now his only message. Clearly, if $\omega > \bar{\omega}$, because R has a dominant strategy of setting $v = 0$, L also has a dominant strategy of playing according to type – i.e., $e_1 = s_1$ if $\theta = 1$ and $e_1 \neq s_1$ if $\theta = 0$.

As in the original semi-separating equilibrium, we consider an equilibrium in which, when L is a convergent type, $e_1 = s_1$ when $\omega < \bar{\omega}$ as well. Given that this pure strategy by convergent types of L, R's belief on witnessing $e_1 = s_1$ must be such that $Pr(\theta = 1 | e_1 = s_1) \geq \pi$ – in which case R's best response is to set $v = 0$.

Divergent types of L thus have a choice of whether to pool with convergent types when $\omega < \bar{\omega}$ – and enjoy guaranteed retention – or to separate and gain rents at the expense of removal. A divergent type gains expected utility $\Delta + \lambda y + [1 - p(0)][\mu + \lambda y]$ from setting $e_1 = s_1$. He gains expected utility $r_1 + \lambda y$ from setting $e_1 \neq s_1$. Hence, divergent types will set $e_1 \neq s_1$ if $r_1 > \Delta + [1 - p(0)][\mu + \lambda y]$ and will set $e_1 = s_1$ otherwise.

We can now characterize a semi-separating equilibrium to the alternative game:

1. For L:

$$
e_1 = \begin{cases} \neg s_1 \text{ if } \theta = 0 \text{ and } \omega > \bar{\omega} \\ \neg s_1 \text{ if } \theta = 0 \text{ and } r_1 > \Delta + [1 - p(0)][\mu + \lambda y] \text{ and } \omega < \bar{\omega} \\ s_1 \text{ otherwise.} \end{cases} \quad e_2 = \begin{cases} \neg s_1 \text{ if } \theta = 0 \\ s_1 \text{ otherwise.} \end{cases}
$$

2. For R:

$$
v = \begin{cases} 0 \text{ if } \omega > \bar{\omega} \\ 0 \text{ if } e_1 = s_1 \\ 1 \text{ otherwise.} \end{cases}
$$

3. and R's beiefs are given by $Pr(\theta = 1 | e_1 \neq s_1) = 0$ and $Pr(\theta = 1 | e_1 = s_1) > \pi$

Comparing these two equilibria, it is apparent that, in the alternative model, for $\omega < \bar{\omega}$, any decision to set $e_1 \neq s_1$ results in removal. This is not true in the original model, where the message $(\neg s_1, 1)$ results

37

in retention for $\omega \in [\underline{\omega}, \bar{\omega}]$. In the alternative model, divergent types choose to set $e_1 \neq s_1$ when $\omega < \bar{\omega}$ iff $r_1 > \Delta + [1 - p(0)][\mu + \lambda y]$. In the original model, divergent types choose to send the message $(\neg s_1, 1)$ for $\omega \in [\underline{\omega}, \bar{\omega}]$ for $r_1 > \Delta + \rho[\mu + \lambda y]$ where $1 - p(0) > \rho$.

Hence, we can say, that in the alternative model, leaders are removed for $\omega \in [\underline{\omega}, \bar{\omega}]$ when $r_1 > \Delta + [1 - p(0)][\mu + \lambda y]$; whereas, in the original model, such leaders would choose to disclose and would be retained. $\qquad \square$

Empirically, we interpret Proposition 5 as indicating that – conditional on institutional covariates – disclosure should be associated with a reduced risk of leader removal via coup. For this range of parameter values, leaders who are able to disclose survive – and their survival results directly from their disclosure decision. The increased risk of popular mobilization cows the elite.

Proposition 6. *L chooses $d = 1$ for a wider range of realizations of r_1 and θ when $\omega \leq \bar{\omega}$ than when $\omega > \bar{\omega}$.*

Proof. This proposition follows trivially from the equilibrium definition in Proposition 4. $\qquad \square$

Proposition 6 tells us which autocrats should be most prone to disclose. Disclosure should be greater when $\omega \leq \bar{\omega}$ – i.e., when the leader is neither personalistic nor entrenched – than otherwise. ω here represents the balance of power within the regime – specifically the risk leader removal poses for regime survival. This risk is affected by political institutions – where these are more strongly codified ω shrinks and disclosure is more likely – and by the leader's time in office – $\omega \leq \bar{\omega}$ when a leader is newly installed in office.

C   Empirical Appendix

C.1   Transparency and Coups: Robustness Checks

In body of the paper, we present a series of Cox competing hazards models of the risk of leader removal via coup in autocracies as a function of transparency, controlling for regime characteristics (Single Party, and Personalistic regimes) as defined by the GWF dataset. Here we present analogous results controlling for measures of autocratic institutions as defined by the DD dataset (Cheibub, Gandhi and Vreeland, 2010).

Specifically, we control for a binary indicator for 'hierarchical' versus 'precarious' autocracies, as described in the main text.

We present coefficient estimates from these models in Table 3. As in the main body of the paper, we stratify the baseline hazard function based on (1) whether there was a prior leader removal via coup (results in the leftmost column in Table 3) and (2) based on the number of past leader removals via coup (the center column). In the rightmost column, we present estimates in which we do not stratify the baseline hazard, and rather simply control for an indicator of whether a previous leader was removed via coup.

Table 3: Hazard of Removal via Coup (DD Controls)

	Past Coup Strata	Coup Experience Strata	Past Coup Control
Transparency	-0.202	-0.228	-0.217
	[-0.401,-0.002]	[-0.450,-0.006]	[-0.420,-0.014]
Growth	-0.008	-0.006	-0.006
	[-0.035,0.019]	[-0.040,0.028]	[-0.033,0.021]
GDP *per capita*	-0.073	-0.071	-0.078
	[-0.145,-0.001]	[-0.138,-0.004]	[-0.154,-0.002]
Hierarchical	0.410	0.280	0.430
	[-0.196,1.017]	[-0.311,0.871]	[-0.182,1.041]
Ever Past Coup			-0.126
			[-0.969,0.717]
# of Subjects	94	94	94
# of Failures	37	37	37

Results from Cox competing hazards regressions of leader removal via coup on transparency and controls. Results in the leftmost column are from a model in which the baseline hazard rate is stratified based on whether there was a previous leader removal via coup in the sample; results in the center column are from a model in which the baseline hazard is stratified based on an ordered variable based on the frequency of past coups; results in the rightmost column do not stratify the baseline hazard, and simply control for an indicator for past coups. 95 percent confidence intervals are reported in brackets.

The coefficient estimates on the HRV Index are consistently negative and are of substantively similar magnitude to those in Table 1. These estimates are significant at the 90 percent level or above in all specifications, and at the 95 percent level when we simply control for a past history of coups.

Point estimates on the hierarchical indicator is not significant in any specification.

C.2 Autocracy and Transparecy Results

C.2.1 Time-Series Cross-Section: GWF Controls

We assess the relationship between transparency, autocratic institutionalization (as measured by the GWF indicators for autocratic regime type), and the presence of a new leader through two sets of regressions. The first of these, based on the model described in Equation 2, consist of Bayesian hierarchical regressions of transparency on covariates. The unit of observation is the regime-year, and we estimate a varying intercepts model, in which each regime receives its own intercept, which is a function of time-invariant

covariates. We estimate this model using an MCMC estimator run in JAGS 3.3.0. We run two chains of 20,000 iterations each, with the first 10,000 iterations used as a burn-in period. Gelman-Rubin diagnostics on all model parameters – including regime specific intercepts – indicate that the MCMC estimator has converged.

While the lagged dependent variable is necessary to incorporate dynamics into the model, its presence, when coupled with regime-specific random effects, raises a potential problem with this specification. Much as with fixed-effects estimators, the presence of variable intercepts induces a correlation between the lagged dependent variable and the error process, resulting in bias. This bias is inversely proportional to the number of time periods observed (Wawro, 2002). While our dataset covers a relatively long period, autocratic regimes are highly heterogeneous in the time over which they survive, and thus over which they are observed. On average, autocratic regimes are observed for 19 years in our dataset.

To correct for this bias, we employ the following procedure: We obtain unbiased estimates of time-varying coefficients (including ρ, the coefficient on the lagged dependent variable) through the use of the Anderson-Hsiao estimator. That is, we first remove unit specific effects by first-differencing equation 2. We then estimate the model in first-differences, instrumenting for the lagged dependent variable by using the twice-lagged level of transparency (i.e., $transaprency_{i,t-2}$), which is independent of the first-differenced error process. We then reestimate equation 2, constraining all coefficients on time-varying variables (ρ and β) to equal the estimates from the (unbiased) first-differenced model. This process then provides our estimates on regime-level parameters γ.[25] We thus estimate a series of equations of the following form:

$$\Delta transparency_{i,t-1} = \mu + \zeta transparency_{i,t-2} + \mathbf{\Delta X_{i,t-1}}\psi + \nu_{i,t-1} \tag{7}$$

$$\Delta transparency_{i,t} = \hat{\rho}\Delta \hat{transparency}_{i,t-1} + \mathbf{\Delta X_{i,t-1}}\hat{\beta} + \eta_{i,t} \tag{8}$$

$$transparency_{i,t} = \alpha_i + \hat{\rho} transparency_{i,t-1} + \mathbf{X_{i,t-1}}\hat{\beta} + \epsilon_{i,t} \tag{9}$$

$$\alpha_i \sim N(\mathbf{Z}\gamma, \sigma_\alpha)$$

where Δ is the first-difference operator, $\Delta \hat{transparency}_{i,t-1}$ is the predicted value of the lagged first-difference of transparency obtained from equation 7, and $\hat{\rho}$ and $\hat{\beta}$ are obtained from equation 8. We estimate all four equations in the same MCMC algorithm run from JAGS 3.3.0. Estimates from this model are reported in the three rightmost columns of Tables 4. We run two chains of 30,000 iterations each, where the first 20,000 iterations are used as a burn-in period. Gelman-Rubin diagnostics on all model parameters indicate that the model has converged.

---

[25]This process bears some semblance to the fixed-effects vector decomposition method of Plümper and Troeger (2007), in that estimates of coefficients on time-varying coefficients are obtained from one unbiased model and substituted into a (biased) model to obtain estimates for time-invariant characteristics. Here, however, this process is meant to address Nickell bias rather than discrepancies between fixed and random effects estimators.

Table 4: Models of Disclosure: GWF Data, Andersion-Hsiao Estimation

	Model 1	Model 2	Model 3
Regime Predictors	**Party**	0.002	-0.002	-0.003
		[-0.031, 0.032]	[-0.037, 0.028]	[-0.034, 0.024]
	Personal	-0.037	-0.037	-0.040
		[-0.073, $4{\times}10^{-4}$]	[-0.070, -0.001]	[-0.072, -0.007]
	Fuel Exporter	-0.029	-0.027	-0.027
		[-0.065, 0.008]	[-0.061, 0.008]	[-0.059, 0.005]
Regime-Year Predictors	**Lag Transparency**	0.645	0.647	0.648
		[0.634, 0.656]	[0.636, 0.657]	[0.637, 0.657]
	GDP *per capita*	$7{\times}10^{-4}$	0.001	
		[-0.013, 0.015]	[-0.012, 0.014]	
	GDP	0.004	0.003	
		[-0.007, 0.014]	[-0.008, 0.013]	
	Ec. Openness	0.004		
		[-0.008, 0.017]		
	Growth	-0.006		
		[-0.013, $2{\times}10^{-4}$]		
	Gov't Consumption	-0.008	-0.007	
		[-0.019, 0.003]	[-0.018, 0.003]	
	New Leader	0.007	0.007	0.008
		[-0.011, 0.026]	[-0.010, 0.027]	[-0.011, 0.026]
	Cubic Time Polynomial	✓	✓	✓
	# Obs	1411	1411	1411
	# Regimes	111	111	111

Results from a hierarchical varying-intercepts linear regression of HRV transparency index scores on listed covariates. Covariates that shift the intercept term are described as 'Regime Predictors', while those that directly shift predicted transparency values are listed as 'Regime-Year Predictors.' All covariates that are neither indicators terms nor time counters have been standardized by subtracting the mean and dividing by the standard deviation. 95 percent credible intervals are presented in brackets.

Table 4 provides coefficient estimates from the second- and third-stage of this system of equations. Table 5, below, presents estimates from the first stage.

Table 5: First Stage Anderson-Hsiao Estimates, GWF Controls

Variable Name	Model 1	Model 2	Model 3
Constant	-0.386	0.146	-0.507
	(22.374)	(22.835)	(22.451)
GDP per capita	-0.007	0.002	
	(0.052)	(0.05)	
GDP	0.032	0.033	
	(0.03)	(0.03)	
Ec. Openness	0.018		
	(0.019)		
Growth	0.004		
	(0.004)		
Gov't Consumption	0.001	0	
	(0.016)	(0.015)	
New Leader	0.01	0.01	0.01
	(0.015)	(0.014)	(0.014)
Time	0.419	-0.114	0.54
	(22.375)	(22.834)	(22.451)
$Time^2$	0.023	0.028	0.02
	(0.132)	(0.136)	(0.135)
$Time^3$	-0.018	-0.019	-0.016
	(0.028)	(0.029)	(0.028)
$Transparency_{t-2}$	-0.009	-0.009	-0.008
	(0.005)	(0.005)	(0.005)

Coefficient estimates from the first stage regression of the lagged first-difference of transparency on the twice lagged level of transparency and time-varying covariates. Standard errors are in parentheses.

C.2.2 Time-Series Cross-Section: DD Robustness Checks

We additionally fit models analogous to those described above, using the DD measures of autocratic institutionalization in place of the GWF measures. First, we describe the results from the hierarchical model using these covariates, as described by Equation 2. We estimates this model via MCMC from JAGS 3.3.0, running 2 chains of 20,000 iterations each, with the first 10,000 iterations used as a burn-in period. Gelman-Rubin diagnostics on all parameters indicate that the model has converged. The results of this estimation are presented in Table 6.

We include the following controls for autocratic institutional features, drawn from the DD dataset: In all

specifications we a binary indicator hierarchical (versus precarious) regimes. In some specifications, we additional includes a binary indicator equal to 1 if the regime has an elected legislature, for indicators equal to 1 if the regime allows for the existence of a single political party, and another indicator equal to one if it allows multiple parties (non-party dictatorships are the reference category), and an another indicator equal to 1 for communist governments.

As in the baseline model, estimates on the $New\ Leader$ term are positive in all specifications with the DD controls. 95 percent credible intervals are bounded away from zero in models based on the equation 2. These coefficients are estimated with somewhat greater error in the Anderson-Hsiao based specifications.

Coefficients on the institutional controls are typically imprecisely estimated. The exception is the coefficient on our variable of interest, hierarchical regimes, which is negative – with 95 percent credible intervals bounded away from zero in four of six specifications.

As with the GWF data, we estimate an Anderson-Hsiao correction using the DD data. We present the first-stage estimates from this series of equations, in which the lagged first-difference of transparency is regressed on its twice lagged level and the first-differences of time-varying covariates, in Table 7.

Table 6: Models of Disclosure: DD Data

	LDV Models			Instrumented LDV Models		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Regime Predictors						
Single Party	-0.018 [-0.088, 0.039]			-0.021 [-0.072, 0.041]		
Multi-Party	0.015 [-0.041, 0.065]			0.009 [-0.038, 0.058]		
Legislature	0.027 [-0.012, 0.066]			0.017 [-0.02, 0.059]		
Hierarchical	-0.027 [-0.06, 0.002]	-0.036 [-0.062, -0.011]	-0.033 [-0.057, -0.005]	-0.023 [-0.051, 0.004]	-0.029 [-0.053, -0.007]	-0.026 [-0.052, -0.005]
Communist	0.024 [-0.035, 0.085]			0.027 [-0.029, 0.079]		
Fuel Exporter	-0.026 [-0.073, 0.016]	-0.033 [-0.07, 0.008]	-0.028 [-0.067, 0.009]	-0.02 [-0.059, 0.018]	-0.025 [-0.06, 0.01]	-0.022 [-0.057, 0.01]
Lag Transparency	0.959 [0.941, 0.977]	0.964 [0.948, 0.978]	0.967 [0.952, 0.982]	0.644 [0.632, 0.656]	0.648 [0.638, 0.657]	0.65 [0.64, 0.66]
GDP per capita	0.011 [-0.005, 0.029]	0.008 [-0.009, 0.02]		0.006 [-0.007, 0.022]	0.004 [-0.007, 0.017]	
GDP	0 [-0.014, 0.012]	0.002 [-0.01, 0.014]		0.002 [-0.01, 0.013]	0.003 [-0.008, 0.013]	
Regime-Year Predictors						
Ec. Openness	-0.003 [-0.019, 0.011]			0.001 [-0.011, 0.014]		
Growth	-0.004 [-0.014, 0.006]			-0.008 [-0.015, -0.002]		
Gov't Consumption	-0.015 [-0.028, -0.001]	-0.016 [-0.028, -0.003]		-0.012 [-0.024, 0]	-0.012 [-0.023, -0.002]	
New Leader	0.031 [0.008, 0.053]	0.03 [0.009, 0.051]	0.03 [0.009, 0.053]	0.001 [-0.016, 0.02]	0.002 [-0.015, 0.02]	0.002 [-0.017, 0.02]
Cubic Time Polynomial	✓	✓	✓	✓	✓	✓
# Obs	1623	1623	1623	1486	1486	1486
# Regimes	135	135	135	121	121	121

Results from a hierarchical varying-intercepts linear regression of HRV transparency index scores on listed covariates. Covariates that shift the intercept term are described as 'Regime Predictors,' while those that directly shift predicted transparency values are listed as 'Regime-Year Predictors.' Results in the left three columns are from the linear model described by equation 2, while those in the right three columns are from the system of equations 7-9. All covariates that are neither indicators terms nor time counters have been standardized by subtracting the mean and dividing by the standard deviation. 95 percent credible intervals are presented in brackets.

Table 7: First Stage Anderson-Hsiao Estimates, DD Controls

Variable Name	Model 1	Model 2	Model 3
Constant	0.225	0.291	0.123
	(22.345)	(22.867)	(22.245)
GDP per capita	-0.013	-0.005	
	(0.051)	(0.05)	
GDP	0.03	0.031	
	(0.03)	(0.03)	
Ec. Openness	0.011		
	(0.019)		
Growth	0.006		
	(0.004)		
Gov't Consumption	0.006	0.003	
	(0.016)	(0.015)	
New Leader	0.007	0.007	0.002
	(0.013)	(0.013)	(0.014)
Time	-0.197	-0.264	-0.094
	(22.344)	(22.867)	(22.246)
Time^2	0.055	0.055	0.047
	(0.133)	(0.129)	(0.132)
Time^3	-0.023	-0.023	-0.02
	(0.028)	(0.027)	(0.028)
$\text{Transparency}_{t-2}$	-0.009	-0.009	-0.008
	(0.005)	(0.005)	(0.005)

Coefficients from the regression of the lagged first difference of transparency on its twice lagged level and the first-differences of time-varying covariates. Standard errors are presented in parentheses.