The Weight of Precedent: Parties, Institutions, and Executive Norms

Daniel A. N. Goldstein* Collin T. Schumock†

April 2024

Abstract

Political executives often adhere to informal traditions established by their predecessors. Without the backing of formal laws, elites have incentives to violate norms, particularly if doing so yields a political advantage. When do constraining *executive norms* carry weight and when do they falter? We examine an infinite horizon principal-agent model to analyze the maintenance of executive norms. We first consider a model which is played only between the executive and their party. This model demonstrates the importance of intra-party accountability in the maintenance of norms, as well as the role that differences in patience can play in willingness to violate norms. Next, we consider an expanded model with two parties and two executives. This shows how expectations over the actions of other political parties shape the willingness to violate norms when in-office. The insights from the models are used to categorize types of executive norms and their relative fragility. We also chart the trajectory of one executive norm in-depth: the two term tradition of the American presidency. Overall, the study holds implications for how informal institutions regulate executive behavior and for understanding the interplay between informal and formal institutions.

Word Count: 11,661

*University of Oslo, daniel.goldstein@stv.uio.no.

†Yale University, collin.schumock@yale.edu.

We would like to thank Peter Bils, Marco Colleoni, Maria Carreri, Gregory Huber, Clement Minaudier, Ian Turner, Dan Alexander, and audiences from Yale, European University Institute, APSA, MPSA, and the Yale-Universitat de Barcelona Historical Political Economy Workshop.

"General Washington set the example of voluntary retirement after 8 years. I shall follow it, and a few more precedents will oppose the obstacle of habit to anyone after a while who shall endeavor to extend his term."

Thomas Jefferson to John Taylor, January 6, 1805[1]

Norms can serve as an informal means of keeping political executives in check. In fact, informal restrictions have become increasingly important in light of the expansion of presidential power, particularly in the United States. Executive aggrandizement has emerged as a pervasive concern for democracies (Bermeo 2016; Levitsky and Ziblatt 2019). Moreover, norms have played an important role in the overall growth of the presidency: "Presidential power is both augmented and constrained" by norms (Renan 2018, p. 2189). The political norms that govern executive behavior operate in a variety of ways and settings, including the context of elections (Azari and Smith 2012). Given that politicians are most often office seekers and want to stay in power (Downs 1957), how can informal norms that constrain politicians' ambitions survive and prove effective?

We focus on a set of norms that regulate the behavior of political elites, which we term "executive norms." These norms are distinct among political norms in that they focus specifically on political executives who have been granted a great deal of unilateral authority. Such political actors are particularly relevant because it is challenging to formally circumscribe all of their possible authority, and they are likely to test the bounds on their power. Thus, norms have played a key role in limiting the expansion of executive power, as they have done with the American presidency (Whittington and Carpenter 2003; Huq 2012). But oversight is only credible if there are actors to enforce it. Although separation of powers serves this purpose to some extent (Skowronek 1997), we focus specifically on an executive's political party as the primary pillar of accountability given its critical role in supporting their candidacy and legislative agenda.

In contrast to other studies of political norms among political parties (Helmke, Kroeger and Paine 2022) or even horizontal accountability among legislative colleagues (Matthews 1959), we examine a principal-agent relationship in which a party can check its executive. Although the party and the executive have aligned interests in the success of an executive that comes from their

[1] Jefferson (1805) on the intention to initiate a norm of the U.S. president serving only two terms.

political party, the party's long-term perspective may serve as a check on the ambitions of an executive bound mostly by informal precedent. Of course, by limiting the power of an aligned executive, a political party or its supporters face a trade-off between, for example, increasing the likelihood of political victories and maintaining a norm that may be beneficial only in the long run. Such trade-offs have been shown to characterize and complicate the motivation of citizens to hold corrupt political executives accountable (Boas, Hidalgo and Melo 2019) as well as trade off elements of democratic principles (Graham and Svolik 2020). In addition, parties can play an important role as "gatekeepers" who can sideline politicians who have the ambition to abandon democratic principles (Levitsky and Ziblatt 2019). This has become a growing concern with the rise of populist politicians who appear eager and willing to transgress both formal and informal institutions (Siegel 2018).

We present a number of executive norms in Table 1. Such norms can range from informal or ad hoc rules of behavior, such as the typical discourse expected from political leaders (Jamieson and Taussig 2017), informational transfers, or ceremonial roles, to regularities that could easily be formalized into law, for example, the appointment of state judges. A prominent example of the latter case is the "two-term tradition" surrounding the presidency in the United States. The long-standing tradition was only institutionalized with the Twenty-second Amendment in 1951. For each example in Table 1, we try to classify its status in the United States with "violated" indicating a clear violation of the norm, "transgressed" indicating that a transgression of a norm has occurred, but has not necessarily been successful in destroying the norm, and "adherence" indicating that the norm persists without serious attempts to violate it. For example, the tradition of presidents limiting themselves to two terms was institutionalized after its violation. For other norms that have been transgressed, such as candidates releasing private information and executives not abusing their pardon powers, it remains to be seen whether political elites can maintain them informally or if formal change will be necessary.[2]

[2]Indeed, legislation has been introduced to constrain presidents' pardon power ("H.R.1627 - 116th Congress (2019-2020): Abuse of the Pardon Prevention Act," `https://www.congress.gov/bill/116th-congress/house-bill/1627`) and to require presidents to release their tax returns ("H.R.347 - 117th Congress (2021-2022): Presidential Tax Transparency Act," `https://www.congress.gov/bill/`

Table 1: Sample of Executive Norms

	Types	Examples	Status in US
1	Informal term limits	American Presidential two-term tradition; Italian Presidency one term tradition	Violated
2	Release of private candidate information	US Presidential candidates tax and medical records	Transgressed
3	Conceding of electoral results	Peaceful transition of power; Assist transition of next administration	Transgressed
4	Direct policy debate with opponents	Participation in Presidential debates	Adherence
5	Direct engagement with critical media	Presidential press conferences	Adherence
6	Pardon powers	Not used for self-gain or to benefit close associates	Transgressed
7	Information sharing among government branches	Executive branch sharing information with Congress by courtesy	Transgressed

The model we present is general enough to extend to most settings where informal norms dictate and constrain the behavior of leaders and prominent elites. For example, norms may inform the use of executive powers, such as the American president's legislative veto (Spitzer 1988), or informal precedents may shape the retirement of positions with lifetime appointments, such as judges, or with open-ended contracts, such as organizational leaders. It may also speak to related norms against undesirable behavior, such as the reputational stigma associated with corruption.

We present two dynamic formal models that outline the elements at play in the maintenance of an executive norm. In the baseline model, which we refer to as the Intra-Party Model, we examine a coordination game between a party and their executive. These two actors interact repeatedly. They may coordinate on jointly beneficial actions, or coordination can fail, with the actors focusing on more individualistic choices. The latter scenario could capture an executive prioritizing their election over party needs, or a party undermining an executive's agenda or campaign (which could occur if it is not electorally or politically expedient for party members). This approach is consistent with how norms have often been discussed and modeled in the literature as expectations of a given joint outcome in repeated interactions (Axelrod 1986; Mailath and Samuelson 2006).

However, we go a step beyond conceiving of norms as simply a coordination game and also

117th-congress/house-bill/347.)

integrate a higher level *precedent game* that captures the possibility of a fundamental change in the underlying executive norms game. The executive and the party may gain a short-term advantage by violating an executive norm, but it comes at the cost of shifting the value of the coordination outcome. That is, by disrupting a long-standing norm, we argue that an executive not only shifts expectations (or reputations) in repeated interactions (which could be recovered), but fundamentally shifts the payoff to coordination. We use this approach to distinguish executive norms as precedents, which are durable norms that establish the informal "rules of the game" (North and Weingast 1989), whereas coordinated outcomes are simply conventions or descriptive norms (Lewis 2008; Cialdini, Reno and Kallgren 1990), which are more transient reputations or simply expectations tied to individual players. Our more complex modeling of norms allows for a richer conceptualization of norm adherence, transgression, and discipline.

The model details the importance of an executive's party in ensuring norm compliance. It outlines the short-term rewards to a party of indulging a norm-violating executive, as well as the long-term costs. We uncover equilibria that allow for both the persistence of an executive norm and the collapse of the norm, which is critically altered by the party's willingness to discipline norm transgressions and to enforce such discipline with sufficient punishment through suboptimal subsequent equilibria play. This underscores that the party is a legitimate gatekeeper of informal checks on executives only if it is credibly willing to implement a sufficiently harsh punishment strategy to deter executive transgressions.

We then build on the baseline setup in a second game, which we call the Inter-party Model. We now introduce a second political party and an executive. In each period, one executive and one party are in office and have the opportunity to transgress the norm. After each period, control of the government can switch to the out-party, which then has the opportunity to transgress the norm. This setup makes the analysis significantly more complex, as the in-office executives and parties must now account for the possibility that the out-party will take office and end the executive norm.

The addition of a second party has two primary effects, depending on the strategies of the parties. First, if the out-party does not discipline and, therefore, facilitates norm transgressions by its executive, then the in-party and its executive have greater incentives to violate the norm immediately in order to reap the benefits. In this way, party competition can undermine the survival of executive norms. However, if the out-of-office party executive respects the norm, this can empower the in-office party to credibly and effectively discipline its executive. This positive spillover from party competition resonates with the literature on mutual forbearance through which parties serve as effective gatekeepers of democratic governance and public sectors insulated from excessive political influence (de Figueiredo, Rui J. P. 2002; Helmke, Kroeger and Paine 2022).

Following the analysis of the model, we detail a case study of the long-standing archetypal executive norm of the "two-term tradition" within American democracy. This norm held that the president would serve for no more than two terms in office. Seen as a critical check on preventing the powerful American president from becoming a monarch, this tradition has been upheld within American political development as an enduring legacy of George Washington.[3] This tradition eventually ended when Franklin D. Roosevelt was elected to a third term in 1940. However, this was not the only attempt to end the tradition. Earlier attempts were made by Ulysses S. Grant and FDR's cousin, Theodore Roosevelt, when he ran for a third term as a member of the Progressive Bull Moose Party in 1912 after failing to receive the Republican nomination. The informal tradition of a two-term limit was eventually legally formalized by the Twenty-Second Amendment to the U.S. Constitution, which was ratified in 1951 (Azari and Smith 2012). How could the informal norm of presidential term limits in the United States survive for 144 years without a legal mandate? And what factors led this tradition to end? The history of the two-term tradition serves to highlight the trajectory of an executive norm.

The study contributes to several strands of literature. Political norms have become an area of increasing focus, particularly since "democratic norms" have been violated in recent years (Helmke, Kroeger and Paine 2022; Grillo and Prato 2023). In particular, the model speaks to cases where informal norms support term limits and combat democratic backsliding in the form of executive

[3]Though it was Thomas Jefferson who initially formulated two terms as a norm: "The danger is that the indulgence & attachments of the people will keep a man in the chair after he becomes a dotard, that reelection through life shall become habitual, & election for life follow that (Jefferson 1805)."

aggrandizement, which aims to undermine or eliminate formal term limits and has been shown to be one of the most common tactics of would-be authoritarian leaders (Meng 2020; Versteeg et al. 2020). In the United States, this issue continues to be a topic of discussion, especially since former President Donald Trump alluded to possibly serving more than two terms while in power (Montanaro 2022).

The study of political norms also falls within a larger field that examines the role of informal institutions in producing political outcomes (North and Weingast 1989; Greif 1993) as well as endogenous institutional change (Greif and Laitin 2004; Mahoney and Thelen 2009; Roland 2004). Moreover, our study contributes to understanding of how informal factors, such as norms, contribute to the maintenance and support of formal political institutions (Azari and Smith 2012; Helmke and Levitsky 2004; Acemoglu and Jackson 2017; Hart 2012; Tyler 2006).

In addition, the model joins research on the role of intra-party competition and party control of candidates (Snyder and Ting 2011). Lastly, our study contributes to a developing field of formal models of historical political development (Gailmard 2017, 2021). Such studies serve to elucidate the salient motives of actors during critical junctures, particularly as formal theory offers researchers an exciting avenue to further explore institutional shifts.

Next, we discuss our definition of executive norms in more detail. We then describe the one-party, intra-party model before turning to the two-party, inter-party model. Then, we consider the implications for several types of executive norms and provide an in-depth case study of the American two-term tradition before concluding.

Executive Norms

Executive norms raise an important distinction between "conventions" and "norms." While conventions are simply empirical regularities that serve to establish expectations (Lewis 2008), norms are enforced by some mechanism, which may be social sanctioning by conforming members (Bicchieri 2005; Bicchieri and Muldoon 2011).[4] What complicates the executive norm we consider

[4]In a sense, conventions refer to "descriptive norms" that detail expected behavior (Cialdini, Reno and Kallgren 1990).

is that it applies to political elites and changes the behavior of only one individual at a time. In this sense, while political norms can be subdivided into "elite norms" that govern the behavior of politicians, executive norms are a further subset that focus specifically on prominent political leaders. Thus, while social norms generally apply to a large population of individuals (e.g., Bénabou and Tirole (2006)), such executive norms have the property that they may be enforced by a large group of people (e.g., party elites and voters) but apply to a single incumbent executive. In this case, while social sanctions are an appropriate conceptual punishment to ensure norm compliance among a large group, elite norms require a different form of punishment.

We focus on the case of party discipline of political executives. In particular, we consider the various coordination problems that parties and executives must overcome to support the electoral and policy success of a political executive. Such coordination problems may include organizing party members to fundraise for the executive, rallying support among party elites and members for an executive at a convention, or supporting the prioritization of the executive's policy agenda. These are largely informally organized activities that require coordination, but which involve disparate party members and elites to support the success of the executive. As we detail in the model, a party could choose to undermine a norm-breaking executive by deviating from coordinated activities. In addition, we consider how norm violations may actually alter the returns to such coordinated outcomes.

The relative ability of parties to discipline their executive speaks to an important scope condition of our analysis, which is the type of political system. Executive norms are particularly relevant in presidential systems, where the executive is separate from the legislature. In parliamentary systems, the head of state is a member of the majority party (or coalition), which ensures a greater degree of authority from other party elites. In presidential systems, the executive can more easily be a political outsider who is less beholden to party elites (Levitsky and Ziblatt 2019). This can pit party elites against executives who may not be drawn directly from their rank and file. These factors ultimately amount to relative party strength, where weak parties are less able to control their members and platforms (Rosenbluth and Shapiro 2018). Nevertheless, in both presidential and parliamentary systems, executive norms still apply to fill in unwritten informal institutions (Helmke and Levitsky 2004).

As discussed below, we distinguish between attempted violations of an executive norm, which we refer to as "transgressions" by an executive, and successful norm violations, where the expectation of the informal action has disappeared. For example, with respect to the two-term tradition, one could transgress the norm by attempting to run for a third term, but the norm is violated only if reelection actually occurred.

Furthermore, while it is largely beyond our scope to consider the creation of executive norms, note that in the case of the term limit norm, while George Washington established the convention of a two-term limit, it was the actions of Thomas Jefferson and subsequent politicians that established a legacy cost for running for a third term. As we detail in the case study, Jefferson was largely motivated by ideological concerns about limiting potential monarchs. This tradition essentially made it "inappropriate" for the executive to be reelected to a third term. In this sense, executive norms can also be related to the literature on "injunctive norms," which are beliefs about the type of behavior that others consider appropriate (Cialdini, Reno and Kallgren 1990). Thus, in this case, underlying political values shape appropriateness, such as hearkening back to the Whig ideology that the possibility of indefinite reelection gives the executive too much authority.

Overview of the Models

We consider two models that build on each other. First, we consider an Intra-Party Model. Here, we consider the dynamics between a single party and its executive in a principal-agent relationship. This model shows how intra-party accountability can sustain a norm over time and when party control falters. A key tension we highlight is that the executive and the party may have different time horizons. In particular, because parties are long-lived entities, while executives have finite lives, they are likely to have more myopic preferences. While this model introduces intra-party dynamics, it has the weaknesses of focusing only on a single party that may violate the norm. We weaken this assumption in the second model.

Second, we examine two executive-party relationships in an inter-party model. There are now two parties, each with an executive who may (potentially) hold office. One party and one executive can hold office in each period. A tension arises because either party-executive dyad can break the precedent, taking all the short-term gains while shifting the norms game for both parties going forward. This model examines how parties can implicitly coordinate to both manage their executives and maintain a mutually beneficial norm. This builds on related models of political norms that analyze so-called "mutual forbearance" among parties (Helmke, Kroeger and Paine 2022), although previous studies lacked the intra-party dynamic.

Intra-Party Model

We first present the simpler model, with one party and their executive, to establish the baseline intra-party results.

Setup

We consider an infinite-horizon, discrete-time stochastic game. The players consist of an (E)xecutive and their (P)arty.[5] In each period, the executive and the party engage in a stage game that may consist of an overarching sequential precedent game and one of two types of substages, which are detailed in Figure 1.[6] In the overarching *precedent stage game*, the executive decides whether to transgress or adhere to an existing norm.[7] If the executive chooses to transgress the norm, the party may or may not discipline the executive. If the party chooses not to discipline, then we say that the norm is *violated* and ceases to exist in the future. If the party disciplines the executive, there is a lottery on whether the norm transgression successfully violates the norm. Party discipline counters the transgression—and preserves the norm—with probability $q \in (0,1)$ and fails (leading to a norm

[5]When referring to the executive we use female pronouns and when referring to the party we use third-person pronouns.

[6]We refer to these normal form games as "substages" rather than subgames because they do not encapsulate all subsequent nodes and, therefore, are not properly defined as subgames. We use the nomenclature of substage, as a convenient shorthand to refer to the simultaneous (coordination) game portions of the overall t period stage game.

[7]The precedent game aligns with the theory of Axelrod (1986) that there can be metanorms that exist to support lower-level norms. In our game, the adhere/transgress and discipline/no discipline sequences are more closely aligned with the metanorm, and the equilibrium of the coordination substage is a lower level norm.

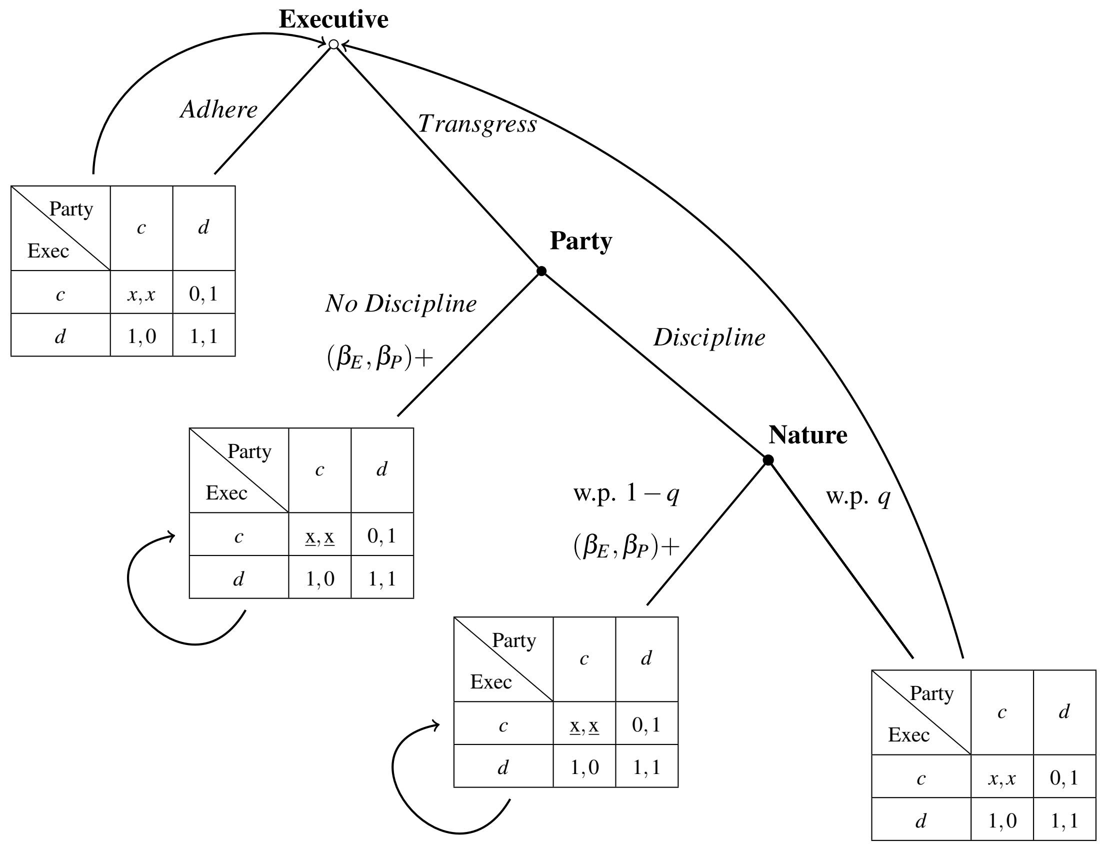
violation) with complementary probability, $1-q$.

Following the precedent game, the players play one of two possible coordination substages. The *norm adherence substage* occurs after the executive adheres to the norm at the beginning of the precedent game, or after the party successfully disciplines the executive following a transgression. After the adherence substage, the game transitions back to the precedent game in the following period with certainty. The process is then repeated as shown in Figure 1.

Conversely, the *norm violation substage* occurs when the executive transgresses the norm and the party does not discipline, or the party's attempt to discipline the executive fails and the norm is violated. The norm violation substage is an absorbing state, in part representing that the norm has ceased to exist, and the players remain in this substage for every subsequent period.

Within each substage, the executive and the party play a coordination game. This is modeled after the stag hunt game, which is based on an account due to Jean-Jacques Rousseau. In the original game, players *c*oordinate to hunt a stag (a larger prize) or *d*efect and hunt a hare (a smaller prize), $a_t^i \in \{c, d\}$. The hare (defection) is the risk-averse action because it has a guaranteed payoff. We normalize the risk dominant payoff to 1. The stag (coordination) payoff (x in our model) is assumed to be larger, $x > 1$, but it is a potentially risky action because if the second player decides to hunt the hare instead, the stag hunter receives nothing. The stag hunt has often been proposed as a model of norms because it is characterized by two pure strategy Nash equilibria (Skyrms 2004).

Figure 1: Intra-Party Stage Game



Following this insight, we use the stag hunt setup to consider possible coordination between the executive and the party. For example, the executive and the party may coordinate to support the executive's electoral or legislative agenda. Alternatively, the executive and the party could act more independently, reducing their payoffs. This could capture the reduced likelihood of reelection or successful passage of partisan legislation. Most critically, the multiple equilibria of the coordination game gives the party a credible discipline mechanism: the coordination equilibrium (and subsequent payoff) represents a viable mechanism for the party to use as a reward for adherence to the existing norm, while the non-cooperative (mutual defection) equilibrium and payoff represents a credible punishment tool for the party following an executive transgression. Note that we refer to "discipline" to indicate the singular action of the party, whereas "punishment" refers to the τ periods in which the party plays defect. Thus, an effective punishment mechanism can make the threat of discipline credible.[8]

Breaking a precedent in political systems is a risky endeavor that can lead to unpredictable new rules and norms (Shepsle 2017). We assume that mutual defection always exists as a risk-averse norm. Further, if the norm is violated, the coordination outcome becomes a new, worse "norm," $\underline{x} < 1$, that gives a payoff lower than the risk-adverse defection outcome. Conceptually, this represents that adhering to a norm that "no longer exists" is suboptimal.[9] Analytically, this results in the executive and the party having strictly dominant strategies of d in the norm violation substages.

Strategies. The set of actions available to the executive in a stage game is history dependent. If the norm has not been violated the set of actions is:

$$A_t^E = \{a_t^E, a_t^E(A), a_t^E(TN), a_t^E(TDV), a_t^E(TDA)\}$$

where $a_t^E \in \{A, T\}$ is the executive's adherence or transgression action and the remaining actions $a_t^E(A)$, $a_t^E(TN)$, $a_t^E(TDV)$, $a_t^E(TDA) \in \{c, d\}$ are the executive's coordinate or defect action in each of the substage games (following *Adhere*, *Transgress* and *No Discipline*, *Transgress*, *Discipline* and successful norm violation, and *Transgress*, *Discipline* and successful discipline, respectively). If the norm has been violated, then the set of actions available to the executive is:

$$A_t^E = \{a_t^E(violation)\}$$

where $a_t^E(violation)$ are the executive's coordinate or defect actions in a violation substage game.

[8]This two-step punishment mechanism is necessary to avoid the classical result that punishment is perfectly effective (preventing transgressions) or completely ineffective (allowing all norm violations). These results can be obtained by taking the limit of τ towards its extreme points. Thus, this mechanism allows us to study the intermediate case, consistent with reality, in which punishment is probabilistically effective.

[9]Consider norms that have at least nominal costs associated with adherence. Once the norm has been violated, whatever benefits the player gained from adhering to the norm no longer exist, only the costs. For example, there is no benefit to limiting a candidate to two terms in office if there is no longer a tradition of two terms expected by voters, or to disclosing private information about candidates if there is no expectation to do so. Moreover, the new norm may actually be worse. For instance, politicians might hide damaging information that could come out later. Faced with a worse norm, a party may be better off distancing itself from a candidate than condoning a new, worse norm.

Similarly, the set of actions available to the party in a stage game where no violation has occurred is:

$$A_t^P = \{a_t^P, a_t^P(A), a_t^P(TN), a_t^P(TD), a_t^P(TD)\}$$

where $a_t^P \in \{N, D\}$ is the party's discipline or no discipline action and the remaining actions $a_t^P(A)$, $a_t^P(TN)$, $a_t^P(TDV)$, $a_t^P(TDA) \in \{c, d\}$ are the party's coordinate/defect action in each of the substage games (following *Adhere*, *Transgress* and *No Discipline*, *Transgress*, *Discipline* and successful norm violation, and *Transgress*, *Discipline* and successful discipline, respectively). If the norm has been violated, then the set of actions available to the party is:

$$A_t^E = \{a_t^P(violation)\}$$

where $a_t^P(violation)$ are the party's coordinate or defect actions in a violation substage game.

In the appendix, we fully detail the history-dependent strategies, see Appendix A. We focus on two particular strategies each for the executive and the party. For clarity and simplicity we label these strategies verbally (instead of displaying them mathematically in the main text) as "Adhere" and "Transgress,"—though note that these strategies are complete and do list actions at every decision node, not only the first node. We focus on two strategies for the party that we will call "Punish" and "Not (Punish)." We use this naming scheme for the party to make clear that the party's punishment strategy relies on more than the party simply playing "discipline" following a transgression. The strategies are detailed in greater length in the following analysis section.

Payoffs. Each player receives the payoffs specified in the appropriate substage for a period t if it is on the equilibrium path. In addition, if the executive violates the norm in period t, both players receive the norm violation payoff, which we assume is positive for both players, $\beta_i > 0$. However, again note that a norm violation is not equivalent to a norm transgression. A violation occurs if the executive transgresses and the party does not discipline, or if the party disciplines but the transgression is still successful, which occurs with probability $1 - q$ (i.e., the norm was successfully violated and the party benefits in the short run regardless of the opposition). The payoffs of the players are discounted in each period after the first by $\delta_i \in (0,1)$. Note that, with the discount factor, two interpretations of the model are valid. The straightforward interpretation is that the game is played between the party and the executive for an infinite number of periods. An equally valid interpretation of the discount factor is that it represents the probability that the party and executive will be matched to play the game again in the next period. In our context, "not being matched to play again" could mean, for example, that the executive has retired or that the party has dissolved.

Sequence of play. The timing of the stage game is summarized below:

1. At the beginning of period t, if we are not stuck in the norm violation substage, the executive can either adhere to the norm or violate it. If the executive adheres to the norm, the executive and the party play a norm adhering substage. We then move to period $t+1$ and return to this first step of the stage game.

2. If the executive transgresses the norm, the party either disciplines the executive or does not (accepting transgression, resulting in a norm violation). In the latter case, we move to the norm violation substage with certainty and remain there *ad infinitum*.

3. If the party disciplines the executive, the game moves to the norm violation substage with probability $1-q$ and remains there *ad infinitum*. With probability q the norm survives and we move to a norm adhering substage. Then we move to period $t+1$ and the precedent game repeats.

Solution concept. There is complete information in the game, and we search for pure-strategy subgame perfect Nash equilibria (by only considering two specific strategies for each player, we focus on a subset of these equilibria, which we will discuss shortly). A stage game strategy profile consists of the players' choices in the precedent game and then the four norm substages.

Analysis and Equilibrium Selection

To uncover equilibria, we apply the one-stage-deviation principle in infinite horizon games (Fudenberg and Tirole 1991). This allows us to verify that a strategy profile is subgame perfect if no player has a profitable one-shot deviation (on or off the equilibrium path). That is, holding the strategy of player $\neg i$ fixed, we can examine a deviation by player i in period t, which then returns to the specified equilibrium play (which may include the prescribed punishment). Thus, for example, if a strategy profile specifies adherence for the executive, it is sufficient to examine one transgression in period $t = 1$ and then a return to adherence, rather than a series of transgressions.

We first solve the substages, since they are invariant to the actions in the precedent game. In the adherence substages, the party and executive either jointly take the coordinated action (c, c), or jointly take the non-cooperative/defect action (d, d).[10] This yields an expected utility of x in the cooperate outcome and 1 in the non-cooperative outcome. Furthermore, in the norm violation substages, there is a dominant strategy to play defect as $1 > \underline{x}$.

We use the multiplicity of equilibria in the adherence substage to draw substantive insights about norms. First, we assume that a benefit of the norm's existence is that the players can (and do) coordinate on the payoff dominant (cooperate) equilibrium in the adherence substage. The party is also able to use the multiplicity of equilibria in the adherence substage as part of their punishment strategy (which we will detail below). In short, the party can credibly threaten to play the defect (or non-cooperative) strategy in the adherence substage for a number of periods following an attempted transgression by the executive.

Turning to the precedent game, we first consider when the executive adheres to the precedent and when they are willing to transgress. Note that if the executive violates the norm in t and the party disciplines, then with probability $1 - q$ the game permanently goes to the norm violation substage. But with probability q, the game enters the norm adherence substage and the precedent repeats in period $t + 1$. We focus on candidate equilibria where, if the executive transgresses and the party discipline is successful (preserving the norm), the executive subsequently adheres in the

[10] There also exists a mixed strategy equilibrium, but we focus on pure strategies in the analysis.

ensuing periods.[11]

Punishment Strategy

The punishment action of the party involves two steps. First, the party chooses whether or not to discipline the executive in the sequential *precedent game*. If the party does not discipline the executive when the executive transgresses, the norm is violated with certainty. (Effectively, the party is permitting, or condoning, the transgression.) However, if the party does discipline, there is a positive probability that the norm transgression is ineffective, and the precedent remains in place. The discipline action, thus, is one punishment lever of the party. By exercising discipline, the party keeps the payoff-dominant norm in play in later periods with positive probability (q). That being said, discipline in the sequential portion of the game alone is not enough to deter an executive determined to violate the norm. But, the party can leverage the multiplicity of equilibria in the adherence substage to build an effective punishment. To deter transgressions, the party bundles the action of discipline with the commitment to play the defect strategy in the adherence substage game for τ periods. This reduces the payoff of the executive into later periods if her transgression fails to violate the norm. Note that this punishment of playing the defect equilibria in the adherence substage is also costly to the party, as they forego the additional payoff that is attainable through the cooperative equilibrium.

The party punishment strategy we consider is as follows. First, the two key elements are these: the party always plays discipline at their first decision node, regardless of the executive's prior action, and the party plays defect for τ periods following a failed transgression by the executive. The history dependent nature of these actions can successfully deter the executive from transgressing. After τ periods of punishment the party returns to cooperating in the adherence substage. Additionally, if the executive has never transgressed the norm, the party plays cooperate in the adherence

[11]We focus on a one-period transgression following the one-stage deviation principle. However, we could also consider equilibrium play in which the executive always transgresses or transgresses for a finite period greater than one period. We limit consideration to a single transgression in equilibrium because this accounts for the finite (political and real-world) lifespans of political executives. Should the executive fail in the norm transgression, a party is likely to remove that politician from the role going forward. Alternatively, the executive may simply have learned that they must also play by the "rules of the game."

substage, and if the norm is violated the party always plays defect in the violation substages. (For a formal presentation refer to Appendix A.)

The other party's strategy "no punishment" is identical to the punishment strategy, except the party always plays no discipline, regardless of the executive's action. The version of the "no punishment" strategy we have specified in the appendix keeps the τ period punishment. We do this to highlight that the τ-defection punishment is not enough to prevent the executive from transgressing, just as the party playing discipline without the τ-defection punishment is insufficient.

The executive transgression strategy we consider consists of the executive attempting a norm violation for one period only. Her transgression strategy says to transgress in any period where the game's history does not contain a transgression. Following a failed transgression, the the executive reverts to adhering, and plays defect for τ periods in the adherence subgame. If her transgression is successful, she plays defect in every violation subgame. If she finds herself in an adherence subgame without having transgressed, or after the τ punishment, she reverts to playing cooperate. (Again, see Appendix A for the formal presentation.)

Finally, the executive's adherence strategy is similar to her transgression strategy, except she always plays adhere. On the equilibrium path, this ensures that every period goes to the adherence substage. Off the equilibrium path, she plays the same actions as she does in her transgression strategy.

Simultaneously, the executive transgresses for one period and then adheres for the rest of the game, plays defect for τ periods before returning to cooperate (indefinitely) in adherence substages following a failed transgression attempt, and plays defect in any violation substages: Transgress.[12]

The party's expected utility when playing Punish with the executive playing Transgress is:

$$Eu_P(\text{Punish}|\text{Transgress}) = (1-q)\left(\beta_P + \frac{1}{1-\delta_P}\right) + q\left(\frac{1-\delta_P^\tau}{1-\delta_P} + \delta_P^\tau\left(\frac{x}{1-\delta_P}\right)\right). \qquad (1)$$

The party's expected utility when playing "Not Punish" with the executive playing "Trangress"

[12]But, again, if the executive adheres in every period then the party and executive play the (payoff dominant) cooperative equilibrium in every period.

is:

$$Eu_P\left(\text{Not}|\text{Transgress}\right) = \beta_P + \frac{1}{1-\delta_P}. \tag{2}$$

Comparing these utilities lead to Lemma 1.

Lemma 1. *The party will punish a norm transgression in any period as long as:*

$$\delta_P^{\tau}\left(\frac{x-1}{1-\delta_P}\right) \geq \beta_P.$$

Given a sufficient number of punishment periods, even the party would be unwilling to engage in such punishment. For example, if we take the limit of τ, this condition will never hold given a positive benefit to violating the norm to the party. However, given a finite τ, we can see that the party is willing to discipline norm transgressions as long as they are sufficiently patient and the value of a norm violation to the party is relatively low. Moreover, this requirement is easier to satisfy as norm adherence becomes more valuable (x is greater).

Note that, as we consider discount factors less than one, by the one-shot deviation principle (Mailath and Samuelson 2006), it is sufficient to examine whether the party would deviate to not punishing in one period to check for any number of deviations.

Transgression Strategy

Now we consider the executive. The executive's expected utility when playing "Adhere" with the executive playing "Punish" is:[13]

$$Eu_E(\text{Adhere}|\text{Punish}) = \frac{x}{1-\delta_E}. \tag{3}$$

Now consider the executive's expected utility from playing a one-shot transgression when the party does not discipline transgressions. The executive will surely achieve a violation of the norm, and then the executive and party will play the violation substage game going forward, where both

[13]Note that when the executive adheres, her expected utility is independent of the party's discipline action in the transgression game.

actors will play the non-cooperative action. The executive's expected utility from this is:

$$Eu_E(\text{Transgress}|\text{Not}) = \beta_E + \frac{1}{1-\delta_E}. \tag{4}$$

And finally consider the executive's expected utility from transgressing for one period when the party does punish transgressions. The executive achieves a violation of the norm with probability q (in which case the players then play the violation substage game in perpetuity) and fails to violate with probability $1-q$. The executive's expected utility from this strategy is:

$$Eu_E(\text{Transgress}|\text{Punish}) = (1-q)\left(\beta_E + \frac{1}{1-\delta_E}\right) + q\left(\frac{1-\delta_E^\tau}{1-\delta_E} + \delta_E^\tau\left(\frac{x}{1-\delta_E}\right)\right). \tag{5}$$

Lemma 2 details the optimal transgression strategy by the executive given both potential strategies by the party.

Lemma 2. *When the party does not punish transgressions, the executive transgresses if:*

$$\frac{x-1}{1-\delta_E} \leq \beta_E.$$

When the party does punish transgressions, the executive transgresses if:

$$\frac{x-1}{1-q}\left(\frac{1-\delta_E^\tau}{1-\delta_E}\right) + \delta_E^\tau\left(\frac{x-1}{1-\delta_E}\right) \leq \beta_E.$$

Expectedly, the condition on β_E is stricter when the party punishes trangressions. The likelihood that the transgression will be successfully disciplined (q) only enters into consideration when the party punishes. At one limit ($q \to 0$), party discipline does not affect the choice of the executive. At the other limit (large τ and perfect discipline, $q \to 1$) the party can always prevent a transgression, given their willingness to do so.

Intra-Party Equilibria

Before stating the propositions we define "voluntary adherence" and "party-induced adherence" which simplifies the proposition statements.

Definition 1 (Voluntary Executive Adherence)**.** An executive voluntarily adheres to a norm whenever they choose to adhere regardless of their party's action following a transgression by the executive.

Definition 2 (Party-Induced Adherence)**.** An executive is induced by their party to adhere to a norm any time that they would choose to transgress the norm if the party did not discipline, but choose instead to adhere to the norm when the party does discipline following a transgression.

Thus, we uncover two set of equilibria, as detailed in Proposition 1.

Proposition 1. *With our restriction of attention to one-stage transgressions by the executive, and finite period punishments by the party in response to transgressions, we uncover two sets of equilibria in the intra-party game.*

1. *Adherence equilibria, where the executive adheres to the norm in every period, of which there are two types:*
    - *Voluntary executive adherence, which occurs when the executive adheres to the norm regardless of the party's discipline strategy.*
    - *Party-induced adherence, which occurs when the executive adheres to the norm in order to avoid a punishment by her party, but she would otherwise prefer to violate the norm.*
2. *Transgression equilibria, where the executive transgresses the norm in the first period, of which there are also two types:*
    - *Transgression and punishment, where the executive transgresses the norm despite knowing that the party will discipline her for doing so.*

- *Transgression and no punishment, where the executive transgresses the norm and the party effectively condones the transgression and does not punish the executive.*

Among the equilibria, a notable result is the transgression equilibrium when the party punishes. As we have argued, it is natural to assume that party will value the future more than the executive. Moreover, since the executive is likely to benefit most directly from the norm violation, it is reasonable to assume that the benefits of the norm violation will most directly aid the executive. Therefore, this outcome may occur when the executive transgresses the precedent but the party moves to block the violation. This could apply to examples of executive norms that we have highlighted—for instance, if the president wishes to run for a third term, breaking the two-term tradition, but the party does not believe that the short-term benefit of having its candidate remain in office is worth violating the long-term benefit of preserving the norm.

Discussion

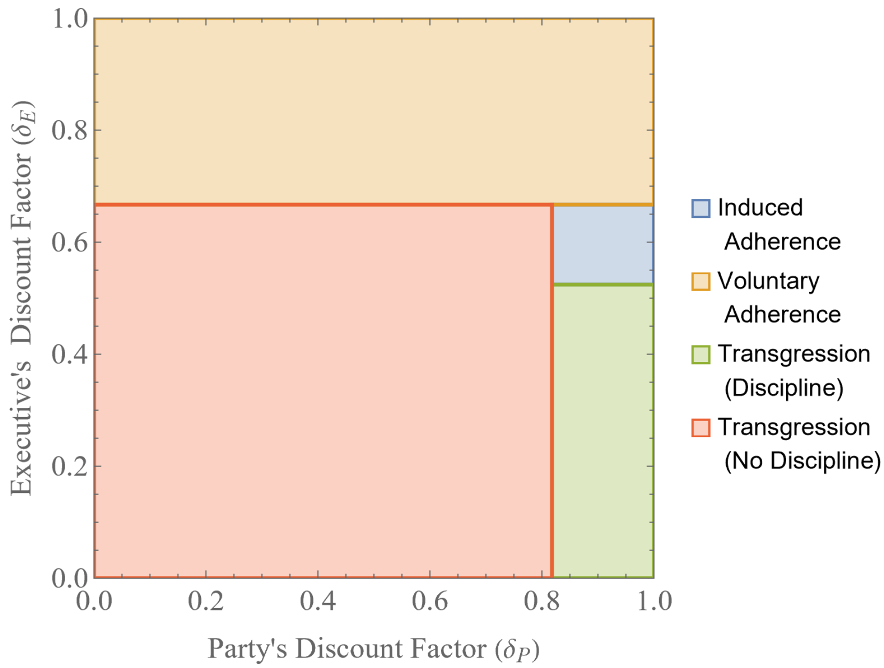
First, Figure 2 plots the four equilibrium results given the party's discount factor on the horizontal axis and the executive's discount factor on the vertical axis. There is reason to believe that the party is likely to discount the future less than the executive. This may simply be because executives are human beings with a finite life span, whereas parties can theoretically exist indefinitely (of course, some parties die or evolve into new parties, such as the Whigs). Moreover, while we model the party as a unitary actor, they are composed of members who may later become the executive. Therefore, they may well prefer to preserve a norm in order to have the benefit of violating it themselves in the future (should they ever come to power). A challenge to this assumption is that it is a necessary condition for the Adherence Equilibrium, with or without party discipline, that the executive has a relatively high discount factor. Moreover, the party must also value the future highly in order to be credibly willing to impose discipline. Only in the latter case is it possible for the executive to adhere to the norm when the discount level is relatively moderate (in this case, close to half). However, if the executive has a very low discount factor, which could reflect, for example, that they are term limited, then party discipline becomes a useless tool and the executive will violate the norm anyway (the lower two rectangles of Figure 2). Here, the party is willing to use executive discipline anyway because it may successfully block the norm violation, making cooperation feasible in the future.

Next, Figure 3 details the interplay between the severity of the party's finite defection penalty and the likelihood that the norm survives transgression.[14] Note that the other parameters are slightly altered so that the party has a relatively low value for violating the norm and a relatively high discount factor; otherwise, the party would be unwilling to discipline the executive. Nevertheless, too severe a punishment, i.e., too many periods of deviation play, is not a credible deterrent to norm transgression because the party is unwilling to bear this cost. At more moderate levels of punishment, this form of discipline is an effective deterrent to norm violations. However, there is a subtle non-linearity as the punishment becomes weaker. This is because even as the transgression becomes less likely to be successful, the severity of the punishment is markedly reduced.

Finally, Figure 4 compares the respective benefits to the executive and the party for violating the norm. The norms we consider most directly benefit the electoral or political ambitions of the executive. Nevertheless, the party is likely to benefit, at least in the short run, if its executive either remains in office or wins policy victories. Again, we adjust select parameters to focus on the case where the party is more patient than the executive. For relatively modest returns to violating the norm, the party is credibly willing to discipline the executive. In this case, the returns to the executive successfully violating the norm must be relatively high to warrant accepting the risk of punishment. However, if the party benefits greatly from violating the norm, the case for the executive's willingness to adhere to the norm is greatly diminished. Thus, the risk of an executive norm falling is largest when the interests of the executive and the party are aligned in wishing to end it.

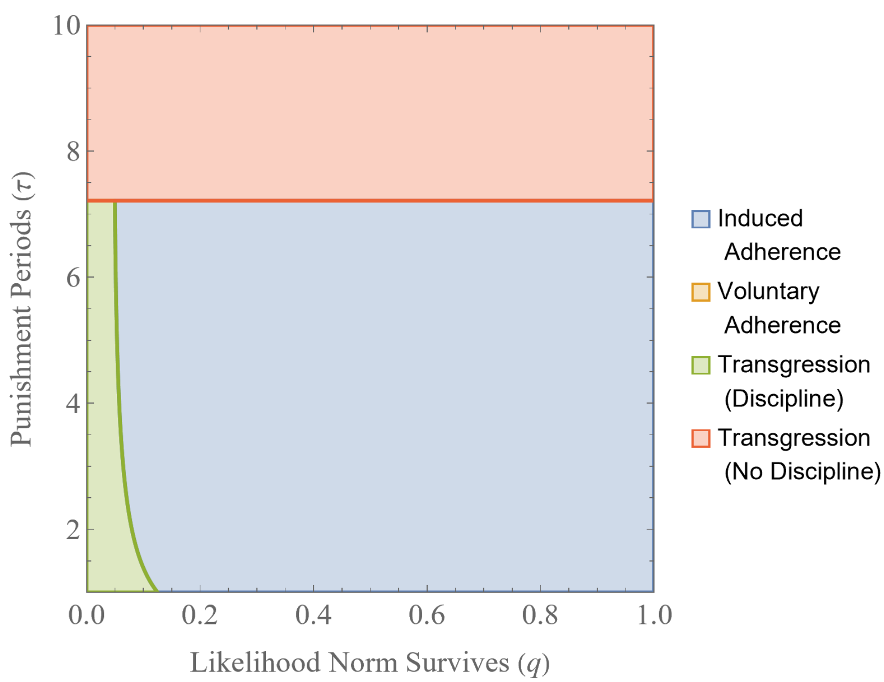
[14]Note that the number of penalty periods τ is a discrete variable. We treat it as continuous here for the purpose of constructing the figure.

Figure 2: Party versus Executive Discount Factors (Intra-Party Model)



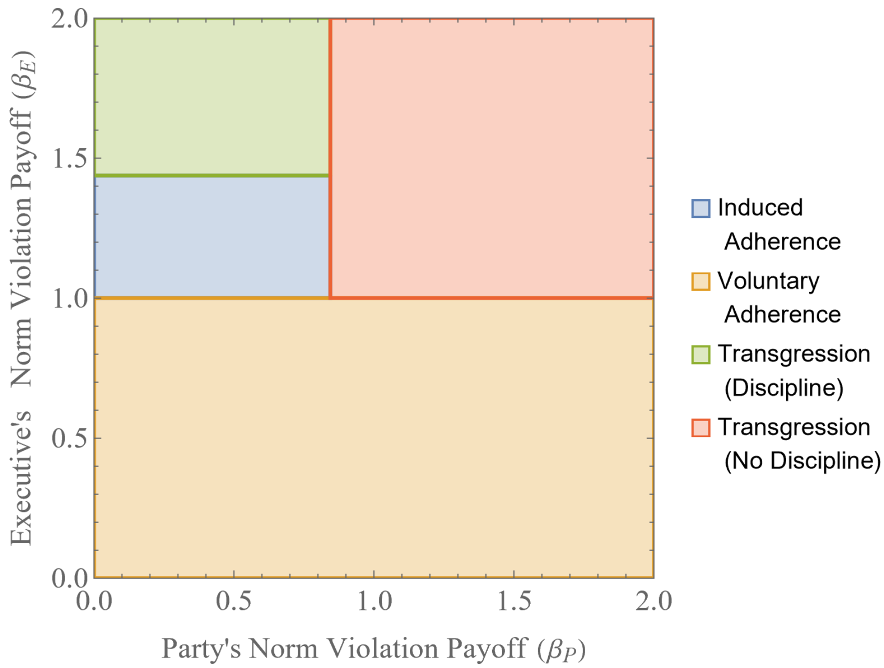
Specification: $x = 1.5, q = \frac{1}{3}, \tau = 3, \beta_E = 1.5, \beta_P = 1.5$

Figure 3: Punishment Periods & Norm Survival (Intra-Party Model)



Specification: $x = 1.5, \beta_E = 1.5, \beta_P = 0.5, \delta_E = .65, \delta_P = 0.8$

Figure 4: Party versus Executive Norm Violation Payoffs Factors (Intra-Party Model)



Specification: $x = 1.5, q = \frac{1}{3}, \tau = 3, \delta_E = 0.5, \delta_P = 0.75$

Minimal Credible Punishment

Our model takes the number of punishment periods as an exogenous parameter. As an extension, we determine for which values of τ punishment will be effective and credible.[15] For the Intra-Party Model the minimum effective punishment length is:[16]

$$\hat{\tau} = \left\lceil \frac{\ln\left(\frac{x-1-\beta_E+q\beta_E+\beta_E\delta_E-q\beta_E\delta_E}{q(x-1)}\right)}{\ln(\delta_E)} \right\rceil.$$

This value represents the minimum τ that will deter the executive from transgressing.

Similarly, we can calculate the maximum τ that the party will be willing to endure to stop the executive from transgressing:

$$\hat{\hat{\tau}} = \left\lfloor \frac{\ln\left(\beta_P\left(\frac{1-\delta_P}{x-1}\right)\right)}{\ln(\delta_P)} \right\rfloor.$$

[15] The model could be extended to treat the duration of punishment, τ, as a choice variable of the party.

[16] Please refer to Appendix D.1 for the derivation of this result.

If there is a τ^* such that $\hat{\tau} \leq \tau^* \leq \hat{\hat{\tau}}$ then that τ^* is an effective punishment (it deters the executive from transgressing) and it is credible (the party is willing to bear it).

Regarding the benefit of violating the norm for the executive (β_E), the partial derivative of the minimum effective punishment length $\hat{\tau}$ is given by:

$$\frac{\delta_E - \delta_E q + q - 1}{(x - 1 + \beta_E(\delta_E - \delta_E q + q - 1))\ln(\delta_E)}.$$

Since $\delta_E - \delta_E q + q - 1$ is always negative within our parameter bounds, the minimum effective punishment increases as long as $x < 1 - \beta_E(q + \delta_E - 1 - q\delta_E)$. That is, as long as the benefit of violating the norm sufficiently outweighs the reward of coordinating, the minimum punishment length sufficient to deter the executive will increase as the benefit increases.

Inter-Party Model

We now extend the intra-party model to consider the role that inter-party competition, and potentially coordination, plays in maintaining executive norms.

Setup

We again consider an infinite horizon repeated stochastic game in discrete time. We now introduce a second party and their executive, which we call Party 2 and Executive 2. In this larger inter-party game, both parties (potentially) play the precedent game of the intra-party model. Layered above the sequential game, Nature determines which party and executive play the precedent game before each period. Nature chooses Party 1 with probability $p \in (0,1)$ and Party 2 with complementary probability.[17]

After Nature chooses a party to play the precedent game, Model 2 proceeds identically to Model 1. As before, when an executive (of either party) chooses to transgress the norm, their party leadership can choose to either discipline the executive for that transgression, or not. If no

[17] A natural interpretation of Nature is as an abstraction of the electorate. Hence, this parameter capture the relative popularity of the incumbent executive. A future extension could formalize the role of the voter.

discipline occurs, the norm is transgressed and subsequently ceases to exist *for both parties*. Party discipline is effective with probability q, and when the party successfully disciplines their executive the norm survives.[18] And further, as in Model 1, following a violation of the norm (either through a lack of party discipline or failed party discipline), parties and their executive play the norm violation substage game. Following maintenance of the norm (through adherence or successful discipline) the party and their executive play the *norm adherence* substage game.

The addition of a second party complicates the strategic calculation of the party.[19] We assume that a successful transgression of the norm by an executive of either party leads the norm to end. Unlike Model 1, in this case, the norm is not a *partisan* norm unique only to one party, but is supra-partisan and exists at the office or election level. Hence, Party 1 needs to consider not only the possibility that their executive will attempt to transgress the norm, but also the potential of Party 2's executive violating the norm, and the likelihood that Party 2 will attempt to discipline their executive. Norm preservation will be more difficult in this environment due to inter-party considerations: if Party 1 holds office and anticipates that Party 2 will not attempt to discipline their executive, Party 1 may now allow their executive to transgress the norm, even if, absent inter-party considerations, they would prefer the norm to be preserved. Nonetheless, tacit coordination between the parties to preserve the norm is possible through mutual discipline of attempted norm transgressions.

The set of actions available to each party and their executive is identical to the set of actions in Model 1.

The payoffs received in the substages are essentially identical to those laid out in Model 1. The difference is that now we assume the out-party and their executive (i.e., the party not chosen by nature) receive payoffs of zero whenever they are not in office. The norm violation payoff β_i continues to only be received once, and only by the party and executive in office when the norm is successfully violated. Although the model allows for β_P and β_E to differ across the parties and

[18]We assume that parties are equally effective in disciplining their executives. A future extension will explore q values that are unique to each party.

[19]We generically focus on Party 1 without loss of generality.

executives, so that $\beta_{P1} \neq \beta_{P2}$ and $\beta_{E1} \neq \beta_{E2}$, we generally consider symmetric norm violation benefits across parties. Discounting is the same as in Model 1, allowing β_i, δ_i to vary across party. We are primarily interested in the differences within parties (δ_{P1} vs. δ_{E1}) as we see the time incongruency between parties and their executives as a particularly important aspect of norm preservation. Comparison across parties (e.g. β_{P1} vs. β_{P2}) allows us to study other substantively interesting questions like differences across parties in their values of the norm, or their gain from violating it.

Information is complete and we again use the subgame perfect Nash equilibrium solution concept in pure strategies. A strategy profile for Game 2 consists of both parties' actions and their executives' actions at each decision node. The timing of one complete period of the game (see Figure 5) is:

1. At the beginning of period t, Nature determines which party's executive holds office: Party 1 with probability p and Party 2 with probability $1-p$.

2. Next, one of two things may occur:

   (a) If the game has not reached the (absorbing) norm violation substage, the executive can either adhere to the norm or transgress. If the executive adheres to the norm, the executive and the party play a norm adhering substage. We then move to period $t+1$.

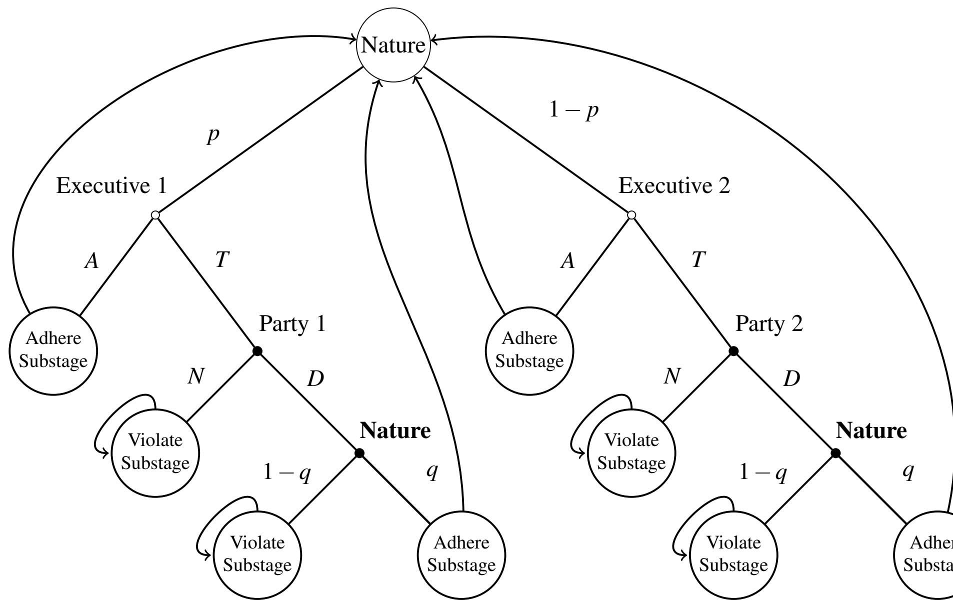
   (b) Alternatively, if the game has reached the norm violation substage, then the executive and party placed in office by Nature in preceding period play the norm violation substage.

3. If the in-office executive transgresses the norm at step 2(a) above, the party either disciplines the executive or does not (accepting the norm violation). In the latter case, we move to the norm violation with certainty and remain there *ad infinitum*. That is, the game cycles through steps 1 and 2(b) repeatedly for the rest of time.

4. If the party disciplines the executive at step 3, the game moves to the norm violation substage

with probability $1-q$ and remains there *ad infinitum*. With probability q the norm survives and we move to the norm adhering substage. Then we move to period $t+1$ and the precedent game repeats.

Figure 5: Inter-Party Stage Game



Analysis

We focus on the same types of equilibria as in Model 1: those where the executives consider a one-stage deviation from adherence to transgression, those where the parties and their executives play the cooperative strategy in the adherence substage if the executive does not attempt a violation, and those where the parties punish a transgression by disciplining their executives and playing the risk dominant strategy in the adherence substages for a finite τ number of periods. The analytical complexity added in Model 2 is that executives make their transgression decisions knowing that they may lose the opportunity to violate the norm if the out-party takes office and successfully violates, and the parties expect to spend at some periods of the game out of power, during which time the norm can be violated. This means that in-party actors must evaluate the out-party actors' resolve to maintain the norm.

Discipline Strategy

As in Model 1, the discipline actions of the parties are comprised of two components—discipline in the precedent game as well as playing the risk dominant action for some number of periods in the adherence substage. By combining these disciplinary actions, the party may be able induce the executive to adhere to the norm. We again consider a discipline strategy where, following a transgression by the executive, the party disciplines and immediately begins playing defect in the adherence substage for a total of τ periods.

We consider one party's decision to discipline or not from the perspective that they know that their executive holds office (we call this the "in-party"). This applies to the case where either party first holds office as the calculations are symmetric (aside from changing idiosyncratic variables and the probability p). To simplify the exposition, we analyze Model 2 *by cases* with the first case being that where the in-party (correctly) anticipates that the out-party executive adheres to the norm. This simplifies the exposition by allowing us to drop the out-party actors' strategies from the in-party actors' utility statements. Effectively, the utility statements that follow assume the out-party and out-party executive play strategies that lead to the outcome described by the case.[20] Those three outcomes are: out-party executive adheres, out-party executive transgresses and is not punished, and out-party executive transgresses and is punished. Note that, for example, the outcome of "out-party executive adheres" can come about in two ways, namely, from the out-party inducing its executive to comply or from the out-party executive voluntarily complying. But *how* this outcome comes about if the out-party takes office is of no consequence to the in-party's expected utility.

Case 1: *out-party executive adheres*. We begin by analyzing the case where the out-party executive adheres to the norm, either because the out-party executive unilaterally prefers to adhere to the norm, or because the out-party's threat of discipline is effective in deterring the out-party executive from transgressing.

First, we derive the party's utility from not disciplining their executive's transgression (mean-

[20]The utility derivations in Appendix B include the out-party strategies.

ing the norm will surely be violated). The party plays no discipline following a transgression, cooperate in any adherence substage, and defect in any violation substage. Consider the party's utility from playing this strategy when their executive plays a one-shot transgression, cooperate in any adherence substage, and defect in any violation substage:[21]

$$Eu_P(\text{Not}|\text{Transgress}) = \beta_P + 1 + p\frac{\delta_P}{1-\delta_P}. \tag{6}$$

Note that the primary difference from Model 1 is that, even in the norm violation case, a party only receives a payoff when their executive holds office, which occurs with probability p from the perspective of Party 1.

Next, we present the in-party's utility from disciplining their executive's transgression, using the same τ-period discipline strategy discussed in Model 1 and with their executive playing the same transgression strategy as above:

$$\begin{aligned} &Eu_P(\text{Punish}|\text{Transgress}) \\ &= (1-q)\left(\beta_P + 1 + \frac{\delta_P \cdot p}{1-\delta_P}\right) + q\left(1 + \left(p\frac{\delta_P - \delta_P^\tau}{1-\delta_P}\right) + \left(p\frac{\delta_P^\tau x}{1-\delta_P}\right)\right). \end{aligned} \tag{7}$$

The party prefers to play a τ-period discipline strategy to one of not disciplining transgressions if $Eu_P(\text{Punish}|\text{Transgress}) \geq Eu_P(\text{Not}|\text{Transgress})$ which we present as Lemma 3.

Lemma 3. *The party will punish a norm transgression in any period when the out-party executive adheres to the norm as long as*

$$p\delta_P^\tau\left(\frac{x-1}{1-\delta_P}\right) \geq \beta_P.$$

Given that $p \in (0,1)$ by assumption, the condition for the party to discipline their executive when considering inter-party dynamics is strictly stronger than the one found in Model 1, but this is solely due to the fact that the party expects to sometimes not hold office. Note that as $p \to 1$ we collapse back to the condition in Model 1 and as $p \to 0$ the party never chooses to discipline. This

[21]With the out-party strategies: $Eu_P(\text{Not}|\text{Transgress}, (\text{Adhere}, \text{Punish}))$, or $Eu_P(\text{Not}|\text{Transgress}, (\text{Adhere}, \text{Not}))$.

is because, in the extreme $p = 0$ the party never expects to obtain office again and thus will not discipline in order to ensure the capture of the β_P violation reward.

The same comparative statics hold for this result as the equivalent in Model 1: with a finite τ the party (may be) willing to discipline, the condition is easier to satisfy as x increases, and the condition is easier to satisfy as δ_P goes to one.

Case 2: *out-party executive transgresses and is not disciplined.* Analysis of the case where the in-party expects the out-party's executive to violate the norm with certainty after taking office (so the out-party is playing a "no punishment" strategy) is slightly more complicated. Now, the party must consider how long they will remain in office, as they rationally expect the norm to fall as soon as they lose power. The role that this party switching plays is clear—if the party anticipates remaining in office for multiple periods then it may be worth preserving the norm, but if the party anticipates losing office relatively quickly it may well be worthwhile to allow their executive to violate the norm in order to capture the one period advantage and to prevent the out-party from capturing the same.

The in-party's utility from not disciplining the norm violation is identical regardless of whether the in-party expects the out party to violate the norm or not, because the in-party executive will violate the norm first. So the in-party's utility from not disciplining is the same as Equation (6).

The in-party's expected utility stream from disciplining an attempted transgression when the in-party anticipates that the out-party executive will violate the norm is a complicated object because there is always the (ever diminishing) possibility that the in-party will remain in power in each period and the norm will not be violated. To compute this infinite regress, we consider the probability that the in-party remains in power through the τ punishment periods and then remains in office indefinitely under the cooperation outcome, while also considering the possibility in each period that the out-party gains office and the norm ends. We again consider the in-party's executive adopting a strategy of a one-shot transgression with the attendant τ-period deviation followed by continuous cooperation in any adherence substage and deviation in any violation substage

$$
\begin{aligned}
&Eu_P(\text{Punish}|\text{Transgress})\\
&= (1-q)\left(\beta_P + 1 + \frac{\delta_P p}{1-\delta_P}\right) + q\left(\frac{p^\tau \delta_P^\tau \left(x + \frac{(1-p)p\delta_P^2}{1-\delta_P}\right)}{1-\delta_P p} + \frac{\left(1 + \frac{(1-p)p\delta_P^2}{1-\delta_P}\right)(p^\tau \delta_P^\tau - 1)}{\delta_P p - 1}\right). \quad (8)
\end{aligned}
$$

When the out-party is expected to violate the norm, the party prefers to play a τ-period discipline strategy to one of not disciplining transgressions when $Eu_P(\text{Punish}|\text{Transgress}) \geq Eu_P(\text{Not}|\text{Transgress})$ which we present as Lemma 4.

Lemma 4. *The party will punish a norm transgression in any period when the out-party executive violates the norm (the out-party executive transgresses and the party does not punish) if:*

$$
\frac{p^\tau \delta_P^\tau \left(x + \frac{(1-p)p\delta_P^2}{1-\delta_P}\right)}{1-\delta_P p} + \frac{\left(1 + \frac{(1-p)p\delta_P^2}{1-\delta_P}\right)(p^\tau \delta_P^\tau - 1)}{\delta_P p - 1} - 1 - \frac{\delta_P p}{1-\delta_P} \geq \beta_P.
$$

Comparing this to the result from Lemma 3 when the out-party executive is expected to adhere to the norm, clearly the condition for the in-party to discipline their executive when the out-party executive is expected to violate the norm is stricter than that when the out-party respects the norm.

Case 3: *out-party executive transgresses and is disciplined.* This case is distinguished from case 2 in that we are considering the out-party executive transgressing and (potentially) being punished by their party.

The in-party's utility from not punishing is again the same as the other cases, refer to Equation (6).

The in-party's utility from punishing their executive's transgression when the out party will attempt a norm violation (out-party executive transgress and out-party punish) is:[22]

[22]The conditions depend on if the punishment is one period or greater than one period.

$$
Eu_P(\text{Punish}|\text{Transgress}) =
$$

$$
\begin{cases}
(1-q)\left(\beta_P - \frac{\delta p}{\delta_P - 1} + 1\right) \\
+q\left(\frac{\delta_P^\tau p^\tau \left(\delta_P(1-p)\left(\frac{\delta_P p(1-q)}{1-\delta_P} + \frac{\delta_P pqx}{1-\delta_P}\right) + x\right)}{1-\delta_P p}\right. \\
\left. + \frac{\left(\delta_P^\tau p^\tau - 1\right)\left(\delta_P(1-p)\left(q\left(\frac{px\delta_P^\tau}{1-\delta_P} - \frac{p\left(\delta_P^2 - \delta_P^\tau\right)}{\delta_P - 1}\right) - \frac{\delta_P p(1-q)}{\delta_P - 1}\right) + 1\right)}{\delta_P p - 1}\right), & \text{if } \tau \geq 2 \\
(1-q)\left(\beta_P - \frac{\delta_P p}{\delta_P - 1} + 1\right) \\
+q\left(\frac{\delta_P^\tau p^\tau \left(\delta_P(1-p)\left(\frac{\delta_P p(1-q)}{1-\delta_P} + \frac{\delta_P pqx}{1-\delta_P}\right) + x\right)}{1-\delta_P p} + \frac{\left(\delta_P^\tau p^\tau - 1\right)\left(\delta_P(1-p)\left(-\frac{\delta_P p(1-q)}{\delta_P - 1} - \frac{\delta_P^2 pqx}{\delta_P - 1}\right) + 1\right)}{\delta_P p - 1}\right), & \text{else.}
\end{cases}
\tag{9}
$$

When the out-party executive is expected to transgress the norm and be disciplined by their party, the in-party prefers to play a τ-period discipline strategy to one of not disciplining transgressions when $Eu_P(\text{Punish}|\text{Transgress}) \geq Eu_P(\text{Not}|\text{Transgress})$.

Lemma 5. *The party will punish their executive's norm transgression in any period when the out-party executive is expected to transgress the norm and be punished by their party (the out-party) if Condition 1 holds (see appendix).*

Transgression Strategy

Now we consider the decision of the in-party executive to attempt a transgression of the norm. We adapt the same analytical approach, considering the choice made by the executive after they have been chosen for office by Nature. We again split the analysis into cases and drop the out-party actors' strategies from the left-hand side of the utility statements.

Case 1: *out-party executive adheres*. The executive's utility for playing the adherence strategy when she expects the out-party to respect the norm, is:

$$Eu_E(\text{Adhere}|\text{Punish}) = Eu_E(\text{Adhere}|\text{Not}) = x + p\frac{\delta_E x}{1-\delta_E}. \tag{10}$$

The in-party executive's utility for a one-shot transgression when the out-party respects the norm and when the in-party does not discipline transgressions is:

$$Eu_E(\text{Transgress}|\text{Not}) = \beta_E + 1 + p\frac{\delta_E}{1-\delta_E}. \tag{11}$$

Whereas the in-party executive's utility for transgressing when her party disciplines her for τ periods using the previously prescribed strategy and when the out-party respects the norm is:

$$\begin{aligned} &Eu_E(\text{Transgress}|\text{Punish}) \\ &= (1-q)\left(\beta_E + 1 + p\frac{\delta_E}{1-\delta_E}\right) + q\left(1 + \left(p\frac{\delta_E - \delta_E^\tau}{1-\delta_E}\right) + \left(p\frac{\delta_E^\tau x}{1-\delta_E}\right)\right). \end{aligned} \tag{12}$$

Now we can determine when the executive chooses to transgress the norm instead of adhering to it, which is detailed by Lemma 3.

Lemma 6. *When the out-party executive adheres to the norm, the in-party executive transgresses the norm when*

- *The party does not discipline, if*

  $\beta_E \geq x - 1 + p\delta_E\left(\frac{x-1}{1-\delta_E}\right)$,

- *Or, when the party disciplines, if*

  $\beta_E \geq \frac{x-q}{1-q} + p\frac{\delta_E x + q\delta_E^\tau - q\delta_E - q\delta_E^\tau x}{(1-\delta_E)(1-q)} - p\frac{\delta_E}{1-\delta_E} - 1.$

Case 2: *out-party executive transgresses and is not disciplined*. Now we consider the executive transgression decision when the executive expects her out-party counterpart to transgress the norm once she takes office. First, the in-party executive's utility for adhering (regardless of the in-party's discipline decision) when the out-party transgresses is:

$$Eu_E(\text{Adhere}|\text{Punish}) = \frac{x + \frac{\delta_E^2 p(1-p)}{1-\delta_E}}{1-\delta_E p}. \tag{13}$$

The executive's utility for transgressing when she expects that her party will not discipline her, and when the out-party will transgress, is the same in the case where the out-party respects the norm, see Equation (11).

$$Eu_E(\text{Transgress}|\text{Not}) = \beta_E + 1 + p\frac{\delta_E}{1-\delta_E}.$$

And the executive's expected utility when transgressing while the Party disciplines for τ periods, and the out-party executive is expected to transgress when in office is:

$$\begin{aligned} &Eu_E(\text{Transgress}|\text{Punish}) \\ &= (1-q)\left(\beta_E + 1 + \frac{\delta_E p}{1-\delta_E}\right) + q\left(\frac{p^\tau \delta_E^\tau \left(x + \frac{(1-p)p\delta_E^2}{1-\delta_E}\right)}{1-\delta_E p} + \frac{\left(1 + \frac{(1-p)p\delta_E^2}{1-\delta_E}\right)(p^\tau \delta_E^\tau - 1)}{\delta_E p - 1}\right). \end{aligned} \tag{14}$$

Now we can detail the conditions for which the executive chooses to transgress the norm instead of adhering to it and when they expect that the out-party will transgress the norm when taking office, which is detailed by Lemma 7.

Lemma 7. *When the out-party executive will successfully violate the norm upon taking office, the in-party executive transgresses the norm when*

- *The party does not discipline, if*

  $\beta_E \geq \frac{x-1}{1-\delta_E p}$,

- *or, when the party disciplines, if,*

  $$\beta_E \geq \frac{x + \frac{\delta_E^2 p(1-p)}{1-\delta_E}}{(1-\delta_E p)(1-q)} - \frac{q}{1-q}\left(\frac{p^\tau \delta_E^\tau \left(x + \frac{(1-p)p\delta_E^2}{1-\delta_E}\right)}{1-\delta_E p} + \frac{\left(1 + \frac{(1-p)p\delta_E^2}{1-\delta_E}\right)(p^\tau \delta_E^\tau - 1)}{\delta_E p - 1}\right) - 1 - \frac{\delta_E p}{1-\delta_E}.$$

Case 3: *out-party executive transgresses and is disciplined.* Now, when the out-party executive transgresses but the out-party attempts to discipline, the in-party executive's utility from adhering:

$$Eu_E(\text{Adhere}|\text{Punish}) = \frac{x + \delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E - 1} - \frac{\delta_E pqx}{\delta_E - 1}\right)}{1 - \delta_E p}. \tag{15}$$

The executive's utility for transgressing when she expects that her party will not discipline her, is the same as the previous two cases, refer to Equation (11).

And the executive's expected utility for transgressing when her party will punish her, and the out-party executive is expected to transgress but be punished when in office is:

$$Eu_E(\text{Punish}|\text{Transgress}) =$$

$$\begin{cases} (1-q)\left(\beta_E - \frac{\delta p}{\delta_E - 1} + 1\right) & \text{if } \tau \geq 2 \\ +q\left(\frac{\delta_E^\tau p^\tau\left(\delta_E(1-p)\left(\frac{\delta_E p(1-q)}{1-\delta_E} + \frac{\delta_E pqx}{1-\delta_E}\right) + x\right)}{1-\delta_E p}\right. & \\ \left. + \frac{\left(\delta_E^\tau p^\tau - 1\right)\left(\delta_E(1-p)\left(q\left(\frac{px\delta_E^\tau}{1-\delta_E} - \frac{p\left(\delta_E^2 - \delta_E^\tau\right)}{\delta_E - 1}\right) - \frac{\delta_E p(1-q)}{\delta_E - 1}\right) + 1\right)}{\delta_E p - 1}\right), & \\ (1-q)\left(\beta_E - \frac{\delta_E p}{\delta_E - 1} + 1\right) & \text{else.} \\ +q\left(\frac{\delta_E^\tau p^\tau\left(\delta_E(1-p)\left(\frac{\delta_E p(1-q)}{1-\delta_E} + \frac{\delta_E pqx}{1-\delta_E}\right) + x\right)}{1-\delta_E p} + \frac{\left(\delta_E^\tau p^\tau - 1\right)\left(\delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E - 1} - \frac{\delta_E^2 pqx}{\delta_E - 1}\right) + 1\right)}{\delta_E p - 1}\right), & \end{cases} \tag{16}$$

Lemma 8. *When the out-party executive will attempt to violate the norm upon taking office (out-party executive transgresses and is disciplined by their party), the in-party executive transgresses the norm when their party does not discipline, if*

$$\beta_E \geq \frac{x + \delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E - 1} - \frac{\delta_E pqx}{\delta_E - 1}\right)}{1 - \delta_E p} - 1 - \frac{\delta_E p}{1 - \delta_E},$$

or, when their party disciplines according to Condition 2 (see appendix).

Inter-Party Equilibria

We can now state three sets of equilibria for the inter-party model. Proposition 2 covers a range of party-party and party-executive dyadic behavior—mutual norm adherence, one-sided adherence, and mutual transgression—and characterizes the conditions for them to occur.[23]

Proposition 2. *With restriction of attention to one-stage transgressions by the in office executive and finite punishments by their party in response to transgressions, there are three sets of equilibria in the inter-party game (categorized by the out-party behavior), each with different forms that may include multiple unique equilibria.*

1. ***Out-party adherence:*** *When the out-party executive voluntarily adheres, the in-party executive will voluntarily adhere under some conditions, be induced to adhere under others, and transgress in others. The party will punish their executive for some parameter conditions and not for others.*

2. ***Out-party violation:*** *Similarly, when the out-party executive will successfully violate the norm, there are conditions that support the in-party executive voluntarily adhering to the norm, being induced to adhere to the norm, and transgressing the norm. Likewise, the party will at times punish transgressions and not in others.*

3. ***Out-party transgression and disciplined:*** *Finally, there is an intermediate region in which the out-party disciplines an out-party executive who transgresses upon taking office. Again, there are conditions that support the in-party executive voluntarily adhering to the norm, being induced to adhere to the norm, and transgressing the norm. Likewise, the party will sometimes punish transgressions and other times not.*

[23] See the proof of Proposition 2 in the appendix for detailed statements of the conditions necessary for each type of equilibrium.

Discussion

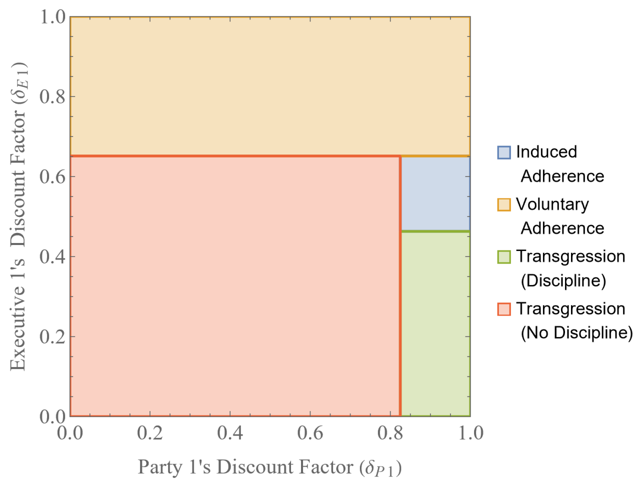
There is a considerable complexity that arises from the inclusion of inter-party competition. In particular, the conditions for adherence and punishment for the in-party are stricter when the out-party violates than when it adheres. We can observe this in Figure 6, which includes when the out-party is certain to adhere (Figure 6a) and when the out-party is certain to violate (Figure 6b). In the case where the out-party adheres, the equilibria are essentially the same as Figure 2 from the intra-party model. However, given that the in-party faces an imminent violation should the out-party gain office, Figure 6b shows that either the executive violates without being disciplined (successfully ending the executive norm) or voluntarily adheres. Thus, in this example, the out-party violating the norm essentially nullifies the effectiveness of the party's disciplinary tool. This further highlights the importance of political executives in maintaining executive norms, as parties require behavior outside of their control (i.e., the out-party and its executive's adherence) in order to even have the possibility of effectively threatening their own executive.

The inclusion of party competition also yields a number of insights. Figure 7 and Figure 8 detail the effect of the probability that the in-party retains office (p) for different values of the in-party executive's discount factor (Figures 7a and 8a) and the in-party's own discount factor (Figures 7b and 8b). For the first specification, the out-party adheres, which requires that we restrict the value of p from above, which essentially means that the out-party would adhere to the norm only if there is a sufficiently high probability that they will win (and retain) office (i.e., p is not too high). For the second specification, we assume that the out-party violates, which requires that the likelihood of retention must exceed a minimum value, as the out-party would deviate to adherence below this value (i.e., p is not too low).

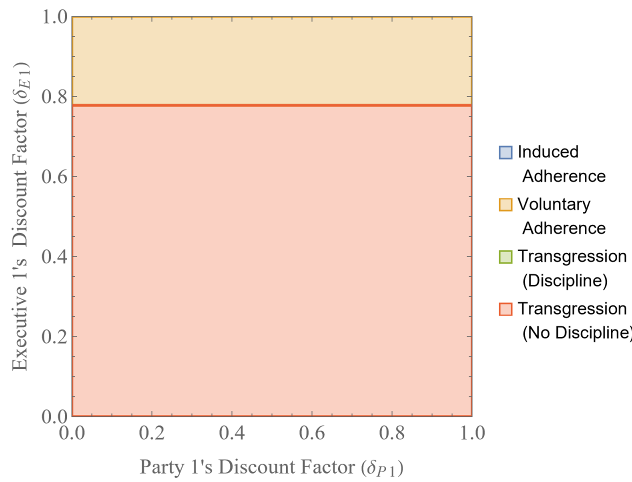
Figure 7a shows that increasing the probability of remaining in office can lead to forbearance in maintaining an executive norm. However, this is only a probabilistic result. It must be the case that the in-party executive sufficiently values the future, otherwise they would attempt a violation of the norm. Moreover, in this region where reelection is likely (but, critically, not too likely), the party gains sufficient leverage to successfully induce adherence through the threat of punishment.

Figure 6: Party versus Executive Discount Factors (Inter-Party Model)

(a) Out-party Adheres ($\delta_{E_2} = 0.9, \delta_{P_2} = 0.9$)



(b) Out-party Violates ($\delta_{E_2} = 0.7, \delta_{P_2} = 0.7$)



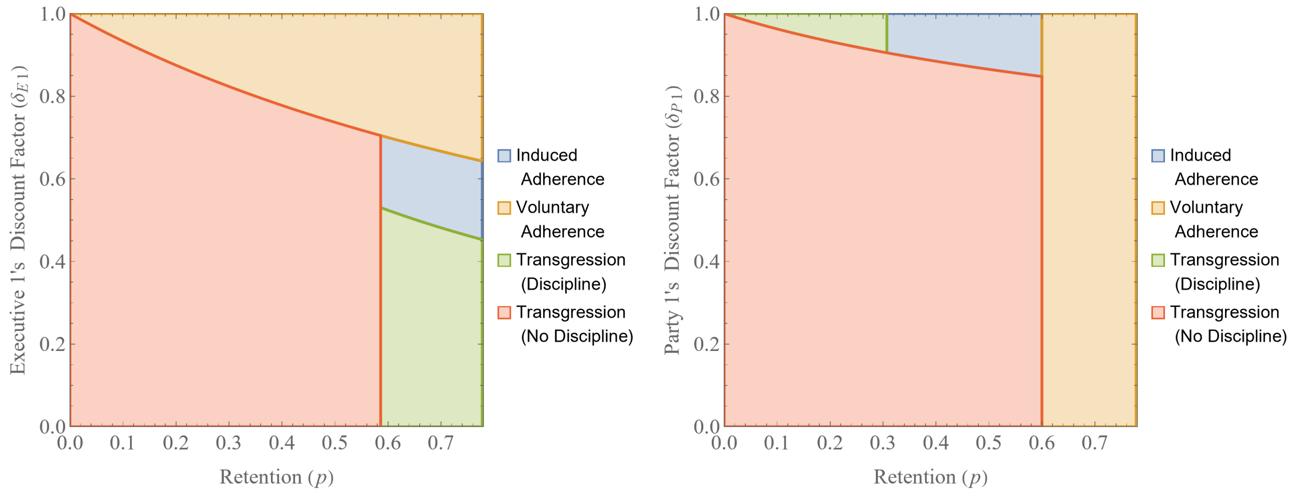
Specification: $x = 1.5, q = \frac{1}{3}, p = \frac{3}{4}, \tau = 3, \beta_{E_1} = 1.2, \beta_{P_1} = 1.2, \beta_{E_2} = 1.5, \beta_{P_2} = 1.5.$

Similarly, Figure 7b shows that the in-party may well allow the norm to be violated unless there is a sufficiently high probability that they will retain power or they are very forward-looking. These results suggest that executive norms are most likely to be violated when a party has a tenuous grip on office. Finally, Figure 8a and Figure 8b show that the out-party violating the norm shifts the dynamic such that the in-party must be extremely likely to retain office in order for the norm to survive. Conversely, a lower retention probability increases the incentives for the in-party and its executive to capture the benefits of violating the norm while it is still possible to do so.

Finally, while we generally assume symmetric norm violation benefits across parties and executives, it is possible that parties and executives value norm violations differently. For example, as we discuss in the next section, with respect to the two-term tradition, it is reasonable to think that the norm violation benefit was lower for the Republican Party than for the Democratic Party. The concept of presidential term limits can be traced from Jefferson and the Democratic-Republicans to the Whigs and the Republican Party. This long history and commitment to the norm would reasonably have made the Republican Party leadership more hesitant to support a third term for one of its presidents.

Figure 7: Retention Likelihood (Inter-Party Model)

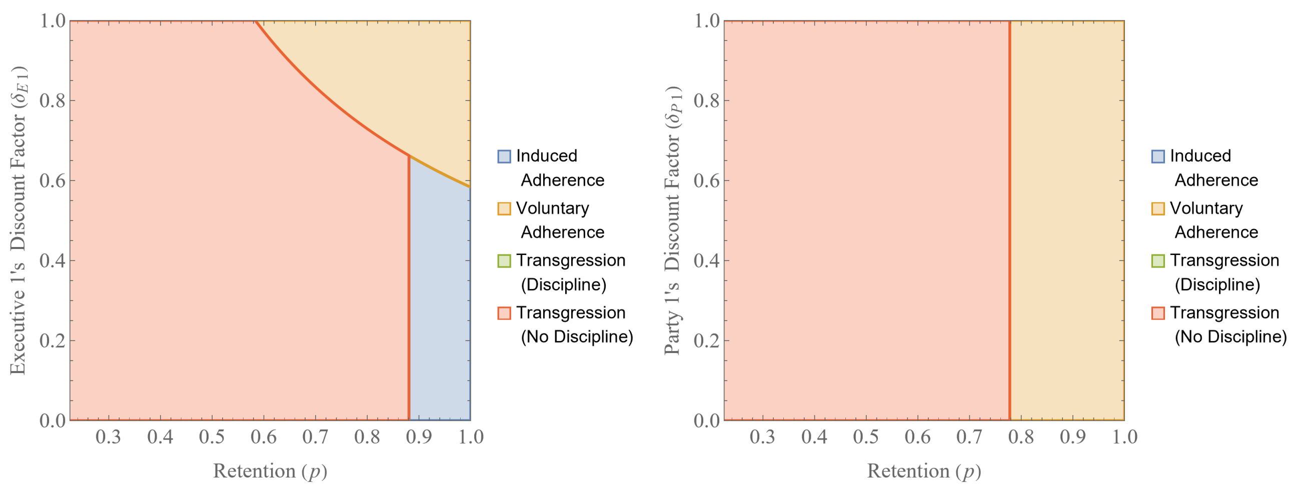
(a) Executive 1's Discount Factor ($\delta_{P_1} = 0.85$) (b) Party 1's Discount Factor ($\delta_{E_1} = 0.7$)



Specification: $x = 1.5, q = \frac{1}{3}, \tau = 3, \beta_{E_1} = 1.2, \beta_{P_1} = 1.2$ and out-party adheres ($\beta_{E_2} = 1.5, \beta_{P_2} = 1.5, \delta_{E_2} = 0.9, \delta_{P_2} = 0.9$).

Figure 8: Retention Likelihood (Inter-Party Model)

(a) Executive 1's Discount Factor ($\delta_{P_1} = 0.85$) (b) Party 1's Discount Factor ($\delta_{E_1} = 0.7$)



Specification: $x = 1.5, q = \frac{1}{3}, \tau = 3, \beta_{E_1} = 1.2, \beta_{P_1} = 1.2$ and out-party violates ($\beta_{E_2} = 1.5, \beta_{P_2} = 1.5, \delta_{E_2} = 0.7, \delta_{P_2} = 0.7$).

Empirical Implications

The models demonstrate the critical interplay between political executives and political parties in maintaining informal institutions such as executive norms. Although we highlight several examples of executive norms in Table 1, not all of these norms are alike, both in their consequences for government, citizen welfare, and in their relative fragility. There are several relatively frivolous traditions that have developed with the American presidency, such as pardoning a turkey at Thanksgiving or giving an interview before the Super Bowl. Such norms have relatively limited benefits, so parties are unlikely to credibly punish a political executive to maintain them. At the same time, these norms have relatively little upside in violating them and thus may persist for long periods of time.

Then there are coordination-based executive norms, such as the executive sharing information with members of Congress, that allows the legislature to operate in a richer information environment. While not necessarily legally required, such informal arrangements provide welfare benefits through improved policy-making. These should be relatively stable because they hold tangible coordination benefits, and there are fewer benefits to violating such norms (although this can occur if the executive wants to mislead the legislature, which can serve as an interbranch check on the executive).

Of particular importance are accountability mechanisms and limits on executive power. Constitutions often leave gaps that can be exploited by politicians with ill intentions for personal gain or the accumulation of power (Helmke and Levitsky 2004). Thus, should a political executive, for example, avoid interactions with critical media, it could undermine citizens' understanding of their government and facilitate hidden corruption or negative policy outcomes. Furthermore, withholding information about candidates could lead to later crises if a political executive falls ill or is implicated in a scandal while in office.

Perhaps most consequential is the crossing of democratic red lines, such as conceding electoral results or leaving office by popular vote. Democratic backsliding often occurs on constitutional grounds (Levitsky and Ziblatt 2019; Przeworski 2019), and, therefore, citizens and parties must serve to credibly enforce constraining executive norms.

We now turn to a case study of an executive norm that falls closer to this critical category in the form of informal term limits. Although weakening term limits can signal a move toward executive aggrandizement and democratic backsliding in both unstable and seemingly consolidated democracies (Meng 2020; Versteeg et al. 2020), there are still a number of trade-offs with term limits (Smart and Sturm 2013; Gersbach, Jackson and Tejada 2020). Much of the historical content of the case study that follows draws on the excellent work, *Presidential Term Limits in American History: Power, Principles, and Politics*, by Michael Korzi (2011).

Finally, note that it may be observationally (or empirically) impossible to distinguish between, for example, mutual voluntary adherence and mixed adherence. In equilibrium, voters will observe leaders of both parties adhering to the norm, but they may not be able to tell whether the executives are adhering voluntarily or under threat of punishment by the party. Executives may have incentives to appear to adhere voluntarily when in fact they are being coerced by their party. Nevertheless, given historical accounts of internal party behavior, it may be possible to infer which equilibria fit a given scenario.

The American Two-Term Tradition

George Washington's decision to only serve two terms as President initiated a tradition, which lasted well over a century, that no executive would exceed two terms in office. However, this informal norm was tested at various points and eventually ended in 1940 when FDR was reelected to a third term. The model speaks to when challenges to the executive norm failed in the 19th and early 20th centuries and when the norm ultimately faltered.

Debates over executive term limits date back to the founding of the United States (Peabody 2001). Skepticism of executive power owing to the colonialists' experiences with England, the states wrote constitutions that established strong legislatures and weak governors who were restricted by term limits or rotation requirements (Korzi 2011). The Articles of Confederation went as far as not establishing an executive branch of the (con)federal government. As problems with a weak national government came into focus, views on executive power shifted.

Ultimately, the Framers would settle on a presidency without any reeligibility restrictions. Although no formal eligibility restrictions were put in place, informal restrictions would quickly arise. Thomas Jefferson supported term limits and used Washington's retirement after two terms as a way to bolster the nascent two-term norm, which became a "virtual postscript" to the Constitution (Korzi 2011, p. 41).

By the time Ulysses S. Grant attempted a third run for president, the two-term tradition was a well-established executive norm, "It is fully imbedded in the minds of the people as if it were written in the Constitution that no man shall serve more than two terms. . . . Even in the case of war, it would be better for Grant to be in the field than in the Presidency," (Joseph Hawley, as cited in Korzi (2011, p. 60)). Grant's "rotation" out of office before his third term attempt in 1880 softened some concerns, but ultimately the Republican Convention chose James Garfield to be their party's candidate.[24]

The next threat to the two-term tradition came from Theodore Roosevelt. While still in office, Roosevelt explicitly denied that he would run for a third term. However, in 1912, once out of office, Roosevelt did not suppress calls for him to run a third time. With President Taft and (some of) the Republican Party opposed to a third Roosevelt term, Roosevelt ran under the banner of the Progressive ("Bull Moose") Party. This would essentially split the Republican vote and ensure victory for Democratic nominee Woodrow Wilson (Korzi 2011).[25]

The penultimate threat to the two-term norm came from Calvin Coolidge, who took office after the death of President Harding. The relatively short portion of Harding's term that Coolidge served (less than two years) led some in his party to argue and advocate for a third term, saying that the end of Harding's term should not count. Although not a violation in the sense of winning a third term, concerns about violating the spirit of the two-term tradition motivated Coolidge, a Republican, not

[24]Of course, given the convention nominating process in use at the time, by "attempt at a third run" we mean that Grant was offered as a potential candidate for the general election at the convention.

[25]In comparison to Grant, Roosevelt's 1912 run is perhaps a more significant "attempt" at third term. Like Grant, it is worth noting that Roosevelt spent a term out of office.

to seek a third term. As he reportedly explained: "If I take another term, I will be in the White House till 1933 . . . Ten years in Washington is longer than any other man has had it — too long!" (White 2018, p. 361).

Franklin Roosevelt was the first (and the last) president to violate the tradition of two terms. Continuing the Roosevelt presidency was seen as the best way to maintain the New Deal programs that were implemented to address economic collapse of the late 1920s and early 1930s. As the 1940 election approached and the war in Europe came into clearer focus, support for a third term grew within the Democratic party and the broader public. Once FDR was clearly the Democratic Party's candidate, the Republican candidate Wendell Willkie made the issue of the two-term tradition (and FDR's violation of it should he be reelected) a centerpiece of the campaign against FDR. The supporters of Roosevelt and the campaign invoked Hamilton and noted the importance of stable leadership in the face of a serious crisis, and FDR noted the successes of his administration and the importance of defending them (Korzi 2011). Ultimately, the two-term norm was not enough to prevent FDR's reelection to a third and fourth term.

Table 2 details how the model equilibria relate to historical cases (italics indicate when the norm was violated). The rows indicate when the executive attempted to violate the norm (transgressed) by running for office, while the columns detail when their party supported that attempted violation (no discipline). This serves to classify four types of equilibrium outcomes.

First, there is the case where the executive transgresses the norm, but the party disciplines the attempt. The "T. Roosevelt 1912 equilibrium" refers to when T. Roosevelt ran as the Progressive Party candidate, and the Republican Party did not support him. Second, we have the case where there was discussion of a third term for an executive, but they did not transgress, and the party would have disciplined them (such as withholding support). This is consistent, for example, with a "Grant 1880 equilibrium." Grant was debated at the nominating convention even though he did not formally seek the nomination and was not selected. Third, there is the case where the executive adheres to the norm, but his party likely would not have disciplined a transgression. We call this outcome the "Jefferson 1808 equilibrium." In this case, the incumbent decided not to enter the race

Table 2: Equilibrium Empirical Cases

		Party: No Discipline	Party: Discipline
Executive	Transgress	*FD Roosevelt 1940*	T. Roosevelt 1912
	Adhere	Washington 1796; Jefferson 1808; Coolidge 1928	Grant 1880; Truman 1952

(and, in fact, initiated the norm) for ideological reasons (i.e., low returns to violating the norm). Finally, fourth, there is the case where the incumbent transgresses the norm and the party does not discipline this action. We call this the "F. Roosevelt 1940 equilibrium," which is when the norm is violated with certainty and the tradition ends.

Why was Franklin Roosevelt able to violate the norm, but not Grant or Theodore Roosevelt? The crisis facing the nation leading up to the 1940 election was clearly a uniquely difficult time for the country. The recovery from the depression was still not fully complete and the situation in Europe was dire. The need for experienced and competent leadership was distinctly high, and the continuation of the office provided an additional level of stability. Indeed, the times more or less precisely matched the scenario in which the Framers envisioned a need for extended presidential tenure. The same cannot be said for the elections of 1880, 1912, or 1928. This could relate to the party's return to the executive violating the norm in the models, indicating that there were significant gains for the party (and the nation) to maintain competent leadership during this turbulent time.

Beyond the unparalleled situation the nation found itself in 1940, there is also the fact that FDR was a Democrat while Grant, Theodore Roosevelt and Coolidge were all Republicans. Simply by tracing the roots of the Republican Party, it is clear that of the two major parties, the Republicans had a greater attachment to and belief in Whiggish philosophy, which from the founding of the country had supported restraints on presidential tenure and term limits. Contrast this with the Democratic Party, which by the time of Franklin Roosevelt's presidency had come to embrace a more “plebiscitary” view of presidential leadership, which argued that the people should not be prevented from electing a president to a third term (Korzi 2011, p. 99). The strong support of FDR by the Democratic Party served as a signal of the party’s confidence in FDR and his administration’s ability to handle the crisis and that handling the crisis effectively (i.e., the benefit of violating the norm) was more important than the two-term tradition.

Conclusion

Governments are rife with informal institutions that shape politicians’ actions. We examine in detail the norms that constrain and establish the behavior of political executives, who hold singular roles as leaders able to shape the trajectory of nations. Using a pair of formal models, we elucidate how such norms shape behavior, the forces that allow them to persist, and when they falter.

When norms are weakened or disregarded, this can create openings for toppling formal institutions that are supported by norms. Maintaining norms that constrain leaders may be critical to preventing power grabs by politicians. The models may apply not only to the history of the United States, but to a variety of settings and party systems in which informal checks on executive power exist. Given that parties gain and lose influence over time, and that the relative strength of parties varies across countries, it is not only up to party elites but also politicians themselves to maintain informal limits on executive power. However, given the centrality of parties in checking executive power, norm-breaking executives may well attempt to co-opt party leadership that could potentially check executive power. This outcome appears to have occurred recently with former President Donald Trump and the leadership of the Republican National Committee, where his daughter-in-law now holds a leadership role (Gold 2024). Thus, it is critical that the party remain strong and relatively independent of unilateral executive influence in order to maintain its gate keeping role of protecting executive norms.

Finally, in order to outline parsimonious motivations, we have abstracted a number of important factors that future research and extension of this work may address. For example, while we have considered the role of party competition in maintaining norms, presidential systems also contend with acute interbranch conflict. As the power of an executive grows over time, there may be more consensus among other branches, such as the legislature, that institutionalization of norms limiting executive power could serve to restrain a potentially unmanageable executive. Taking these dynamics into account may be a fruitful avenue for future research.

Additionally, an extension of this model could replace Nature with a voter who reacts strategically, instead of probabilistically, to norm violations. One could imagine a model in which norm violations help an executive's short run electoral or policy goals, but harm the party in the long run if the voter shifts their support in response (akin to metanorms discussed in Axelrod (1986)). In such a model, party discipline could signal to voters that the party is resolved to maintain norms. Alternatively, a party condoning a violation may signal to the voter that the norm is unimportant.

Another refinement of the model would be to endogenize the length of punishment as a party choice variable. Finally, further research could examine when institutionalized rules on executives are overturned and what role norms play in actually preventing the erosion of such formalized constraints in the first place.

References

Acemoglu, Daron and Matthew O. Jackson. 2017. "Social Norms and the Enforcement of Laws." *Journal of the European Economic Association* 15(2):245–295.

Axelrod, Robert. 1986. "An Evolutionary Approach to Norms." *The American Political Science Review* 80(4):1095–1111.

Azari, Julia R. and Jennifer K. Smith. 2012. "Unwritten Rules: Informal Institutions in Established Democracies." *Perspectives on Politics* 10(1):37–55.

Bénabou, Roland and Jean Tirole. 2006. "Incentives and prosocial behavior." *American economic review* 96(5):1652–1678.

Bermeo, Nancy. 2016. "On Democratic Backsliding." *Journal of Democracy* 27(1):5–19.

Bicchieri, Cristina. 2005. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press.

Bicchieri, Cristina and Ryan Muldoon. 2011. Social Norms. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta and Uri Nodelman. Spring 2011 ed. Metaphysics Research Lab, Stanford University.

Boas, Taylor C., F. Daniel Hidalgo and Marcus André Melo. 2019. "Norms versus Action: Why Voters Fail to Sanction Malfeasance in Brazil." *American Journal of Political Science* 63(2):385–400.

Cialdini, Robert B, Raymond R Reno and Carl A Kallgren. 1990. "A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places." *Journal of Personality and Social Psychology* 58(6):1015–1026.

de Figueiredo, Rui J. P., Jr. 2002. "Electoral Competition, Political Uncertainty, and Policy Insulation." *The American Political Science Review* 96(2):321–333.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.

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

Gailmard, Sean. 2017. "Building a New Imperial State: The Strategic Foundations of Separation of Powers in America." *The American Political Science Review* 111(4):pp. 668–685.

Gailmard, Sean. 2021. "Imperial Governance and the Growth of Legislative Power in America." *American Journal of Political Science* 65(4):912–925.

Gersbach, Hans, Matthew O Jackson and Oriol Tejada. 2020. "Optimal Term-Length and the Structure of Democracy.". Available at SSRN: `https://ssrn.com/abstract=3615407` or http://dx.doi.org/10.2139/ssrn.3615407.

Gold, Michael. 2024. "Trump Ally and Daughter-in-Law Officially Take Over R.N.C. Leadership." *The New York Times* .
URL: *https://www.nytimes.com/2024/03/08/us/politics/trump-rnc-lara-michael-whatley.html*

Graham, Matthew H and Milan W Svolik. 2020. "Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States." *American Political Science Review* 114(2):392–409.

Greif, Avner. 1993. "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition." *The American Economic Review* 83(3):525–548.

Greif, Avner and David D. Laitin. 2004. "A Theory of Endogenous Institutional Change." *The American Political Science Review* 98(4):633–652.

Grillo, Edoardo and Carlo Prato. 2023. "Reference Points and Democratic Backsliding." *American Journal of Political Science* 67(1):71–88.

Hart, Herbert Lionel Adolphus. 2012. *The Concept of lLaw*. Oxford University Press.

Helmke, Gretchen, Mary Kroeger and Jack Paine. 2022. "Democracy by Deterrence: Norms, Constitutions, and Electoral Tilting." *American Journal of Political Science* 66(2):434–450.

Helmke, Gretchen and Steven Levitsky. 2004. "Informal Institutions and Comparative Politics: A Research Agenda." *Perspectives on Politics* 2(4):725–740.

Huq, Aziz Z. 2012. "Binding the Executive (by Law or by Politics) Reviews." *University of Chicago Law Review* 79:777–836.

Jamieson, Kathleen Hall and Doron Taussig. 2017. "Disruption, Demonization, Deliverance, and Norm Destruction: The Rhetorical Signature of Donald J. Trump." *Political Science Quarterly* 132(4):619–650.

Jefferson, Thomas. 1805. "Presidential Term Limits in American History: Power, Principles & Politics.".
URL: `https://www.loc.gov/resource/mtj1.032_0169_0169/?st=text`

Korzi, Michael J. 2011. *Presidential Term Limits in American History: Power, Principles & Politics*. Texas A&M University Press.

Levitsky, Steven and Daniel Ziblatt. 2019. *How Democracies Die*. Crown.

Lewis, David. 2008. *Convention: A Philosophical Study*. John Wiley & Sons.

Mahoney, James and Kathleen Thelen, eds. 2009. *A Theory of Gradual Institutional Change*. Cambridge University Press pp. 1–37.

Mailath, George J and Larry Samuelson. 2006. *Repeated Games and Reputations: Long-Run Relationships*. Oxford university press.

Matthews, Donald R. 1959. "The Folkways of the United States Senate: Conformity to Group Norms and Legislative Effectiveness." *The American Political Science Review* 53(4):1064–1089.

Meng, Anne. 2020. *Constraining Dictatorship: From Personalized Rule to Institutionalized Regimes*. Political Economy of Institutions and Decisions Cambridge University Press.

Montanaro, Domenico. 2022. "Trump Can Only Serve 4 More Years. The Reason Why Has a Long and Sordid History." *NPR* . Available at: https://www.npr.org/2022/09/06/1120383809/22nd-amendment-explainer.

North, Douglass C. and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *The Journal of Economic History* 49(4):803–832.

Peabody, Bruce G. 2001. "George Washington, Presidential Term Limits, and the Problem of Reluctant Political Leadership." *Presidential Studies Quarterly* 31(3):439–453.

Przeworski, Adam. 2019. *Crises of Democracy*. Cambridge University Press.

Renan, Daphna. 2018. "Presidential Norms and Article II." *Harvard Law Review* 131:2187–2282.

Roland, Gérard. 2004. "Understanding Institutional Change: Fast-Moving and Slow-Moving Institutions." *Studies in Comparative International Development* 38:109–131.

Rosenbluth, Frances and Ian Shapiro. 2018. *Responsible Parties: Saving Democracy from Itself*. Yale University Press.

Shepsle, Kenneth A. 2017. *Rule Breaking and Political Imagination*. University of Chicago Press.

Siegel, Neil S. 2018. "Political Norms, Constitutional Conventions, and President Donald Trump Symposium: The Future of the U.S. Constitution." *Indiana Law Journal* 93:177–206.

Skowronek, Stephen. 1997. *The Politics Presidents Make: Leadership from John Adams to Bill Clinton*. Harvard University Press.

Skyrms, Brian. 2004. *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press.

Smart, Michael and Daniel M. Sturm. 2013. "Term Limits and Electoral Accountability." *Journal of Public Economics* 107:93–102.

Snyder, Jr, James M. and Michael M. Ting. 2011. "Electoral Selection with Parties and Primaries." *American Journal of Political Science* 55(4):782–796.

Spitzer, Robert J. 1988. *The Presidential Veto*. SUNY Press.

Tyler, Tom R. 2006. *Why People Obey the Law*. Princeton University Press.

Versteeg, Mila, Timothy Horley, Anne Meng, Mauricio Guim and Marilyn Guirguis. 2020. "The Law and Politics of Presidential Term Limit Evasion." *Columbia Law Review* 120(1):173–248.

White, William Allen. 2018. *A Puritan in Babylon: The Story of Calvin Coolidge*. Pickle Partners Publishing.

Whittington, Keith E and Daniel P Carpenter. 2003. "Executive Power in American Institutional Development." *Perspectives on Politics* 1(3):495–513.

Appendix

Contents

A Strategies

Here we formally present the strategies we consider for the players in each model.

A.1 Model 1

We define a history H_t as the set of all feasible histories where a given history is defined as $h_t \in H_t$. Therefore, the set of all histories is defined as $H = \bigcup_{t=1}^{\infty} H_t$. Hence, we define a strategy for a player i as a mapping of histories to strategies, $H \to s_i$. Thus, a strategy for a player is set $S_i = \{s_{i,1}(h_1), s_{i,2}(h_2), \ldots\}$, which specifies a complete strategy profile for all contingencies and where each stage game is denoted by A_t^i.

For convenience of presentation we define $x_i = \{c, d\}$ where x_i can take either value for either player. This allows us to express multiple histories with one entry.

A.1.1 Executive

The adherence strategy we consider for the executive is as follows.

Stage 1:

- Play $Acddd$.

Stage 2:

- Play $Acddd$ if the outcome of stage 1 is (Ax, x) for any x_i.
- Play d if the outcome of stage 1 is (Tx, Nx) or $(Tx, Dx, (1-q))$ for any x_i.
- Play $Adddd$ if the outcome of stage 1 is (Tx, Dx, q) and $\tau \geq 2$ for any x_i.
- Play $Acddd$ if the outcome of stage 1 is (Tx, Dx, q) and $\tau = 1$ for any x_i.

Stage 3–∞:

- Play $Acddd$ if h_{t-1} is a sequence consisting of only combinations of (Ax, x).
- Play d if (Tx, Nx) or $(Tx, Dx, (1-q))$ occurs anywhere in h_{t-1} for any x_i.

- Play $Adddd$ if h_{t-1} is a sequence that does not contain (Tx, Nx) or $(Tx, Dx, (1-q))$ anywhere and contains a subsequence of histories where (Tx, Dx, q) is followed by $n < \tau$ periods of (Ax, x) that is not itself a subsequence of histories (Tx, Dx, q) followed by $n \geq \tau$ periods of (Ax, x).
- Play $Acddd$ if h_{t-1} is a sequence that does not contain (Tx, Nx) or $(Tx, Dx, (1-q))$ and contains a subsequence of histories (Tx, Dx, q) followed by $n \geq \tau$ periods of (Ax, x) and no other unique subsequence of histories where (Tx, Dx, q) is followed by $n < \tau$ periods of (Ax, x).

In words, this strategy says that the executive adheres in the first period. Then, in period two, if the outcome of the first period is that she adhered, regardless of what the players play in the adherence substage game, she will adhere again. Then, in periods three through infinity, if she adhered in every prior period, she will adhere in the current period (again, regardless of the outcome of the adherence substage game). Given that the executive is the first mover and she and the party immediately play an adherence substage game following the executive playing adhere, we can see that this will be the relevant part of the executive's strategy along the equilibrium path. But her strategy must specify an action for all possible contingencies. The second bullet points of stages two and three say that the executive will always play d in a violation substage. The third and fourth bullet points of stage two prescribe actions that differ only at the first adherence substage game. The third bullet point says that if the executive trangsresses and is successfully disciplined by the party she will play defect at the adherence substage game following playing adhere if the length of punishment, τ, is two periods or longer. Alternatively, if the party plays a punishment that only lasts one period, the executive immediately returns to playing cooperate at the first adherence substage. In stages three through infinity the third and fourth bullet points say the following: if the history is such that the executive transgressed and was sucessfully disciplined by the party, and the number of periods since that discipline have not yet reached the punishment length, play defect at the first adherence substage; if the number of periods that have passed are greater than or equal to the punishment length, play cooperate in the first adherence subgame.

Now we present the executive's one period transgression strategy.

Stage 1:

- Play $Tcddd$.

Stage 2:

- Play $Tcddd$ if the outcome of stage 1 is (Ax, x) for any x_i.
- Play d if the outcome of stage 1 is (Tx, Nx) or $(Tx, Dx, (1-q))$ for any x_i.
- Play $Adddd$ if the outcome of stage 1 is (Tx, Dx, q) and $\tau \geq 2$ for any x_i.
- Play $Acddd$ if the outcome of stage 1 is (Tx, Dx, q) and $\tau = 1$ for any x_i.

Stage 3–∞:

- Play $Tcddd$ if h_{t-1} is a sequence consisting of only combinations of (Ax, x).
- Play d if (Tx, Nx) or $(Tx, Dx, (1-q))$ occurs anywhere in h_{t-1} for any x_i.
- Play $Adddd$ if h_{t-1} is a sequence that does not contain (Tx, Nx) or $(Tx, Dx, (1-q))$ anywhere and contains a subsequence of histories where (Tx, Dx, q) is followed by $n < \tau$ periods of (Ax, x) that is not itself a subsequence of histories (Tx, Dx, q) followed by $n \geq \tau$ periods of (Ax, x).
- Play $Acddd$ if h_{t-1} is a sequence that does not contain (Tx, Nx) or $(Tx, Dx, (1-q))$ and contains a subsequence of histories (Tx, Dx, q) followed by $n \geq \tau$ periods of (Ax, x) and no other unique subsequence of histories where (Tx, Dx, q) is followed by $n < \tau$ periods of (Ax, x).

This strategy is nearly identical to the executive's adherence strategy. The key difference is that, under the transgression strategy, the executive transgresses in any period when she has not transgressed in a prior period. On the equilibrium path, this will mean that the executive transgresses in period 1.

A.1.2 Party

The punishment strategy we consider for the party is as follows.

Stage 1:

- Play *Dcddd* regardless of whether the executive plays A or T.

Stage 2:

- Play *Dcddd** if the outcome of stage 1 is (Ax,x) for any x_i.
- Play d if the outcome of stage 1 is (Tx,Nx) or $(Tx,Dx,(1-q))$ for any x_i.
- Play *Ddddd** if the outcome of stage 1 is (Tx,Dx,q) and $\tau \geq 2$ for any x_i.
- Play *Dcddd** if the outcome of stage 1 is (Tx,Dx,q) and $\tau = 1$ for any x_i.

Stage 3–∞:

- Play *Dcddd** if h_{t-1} is a sequence consisting of only combinations of (Ax,x).
- Play d if (Tx,Nx) or $(Tx,Dx,(1-q))$ occurs anywhere in h_{t-1} for any x_i.
- Play *Ddddd** if h_{t-1} is a sequence that does not contain (Tx,Nx) or $(Tx,Dx,(1-q))$ anywhere and contains a subsequence of histories where (Tx,Dx,q) is followed by $n < \tau$ periods of (Ax,x) that is not itself a subsequence of histories (Tx,Dx,q) followed by $n \geq \tau$ periods of (Ax,x).
- Play *Dcddd** if h_{t-1} is a sequence that does not contain (Tx,Nx) or $(Tx,Dx,(1-q))$ and contains a subsequence of histories (Tx,Dx,q) followed by $n \geq \tau$ periods of (Ax,x) and no other unique subsequence of histories where (Tx,Dx,q) is followed by $n < \tau$ periods of (Ax,x).

*In each case the party plays this set of actions regardless of whether the executive played A or T in the move immediately prior.

Again, this punishment strategy contains a structure very similar to that of the executive's strategies. It says that the party always plays D at their discipline action node, regardless of the executive's choice of T or A. This can serve to deter the executive from transgressing, as we will demonstrate. The strategy also says that the party always plays defect when it finds itself in a violation subgame. Finally, when the executive has transgressed the norm and the executive's discipline has successfully blocked the violation, the strategy says that the party will play defect in the first adherence substage until τ periods have passed since the transgression attempt.

Finally we present the no punishment strategy we consider for the party.

Stage 1:

- Play $Ncddd$ regardless of whether the executive plays A or T.

Stage 2:

- Play $Ncddd^*$ if the outcome of stage 1 is (Ax,x) for any x_i.
- Play d if the outcome of stage 1 is (Tx,Nx) or $(Tx,Dx,(1-q))$ for any x_i.
- Play $Ndddd^*$ if the outcome of stage 1 is (Tx,Dx,q) and $\tau \geq 2$ for any x_i.
- Play $Ncddd^*$ if the outcome of stage 1 is (Tx,Dx,q) and $\tau = 1$ for any x_i.

Stage 3–∞:

- Play $Ncddd^*$ if h_{t-1} is a sequence consisting of only combinations of (Ax,x).
- Play d if (Tx,Nx) or $(Tx,Dx,(1-q))$ occurs anywhere in h_{t-1} for any x_i.
- Play $Ndddd^*$ if h_{t-1} is a sequence that does not contain (Tx,Nx) or $(Tx,Dx,(1-q))$ anywhere and contains a subsequence of histories where (Tx,Dx,q) is followed by $n < \tau$ periods of (Ax,x) that is not itself a subsequence of histories (Tx,Dx,q) followed by $n \geq \tau$ periods of (Ax,x).
- Play $Ncddd^*$ if h_{t-1} is a sequence that does not contain (Tx,Nx) or $(Tx,Dx,(1-q))$ and contains a subsequence of histories (Tx,Dx,q) followed by $n \geq \tau$ peri-

ods of (Ax, x) and no other unique subsequence of histories where (Tx, Dx, q) is followed by $n < \tau$ periods of (Ax, x).

*In each case the party plays this set of actions regardless of whether the executive played A or T in the move immediately prior.

A.2 Model 2

The strategies for the actors are nearly identical for Model 2. The main difference is that nature's choice of the party in office and the out-party's actions enter into the histories. Nature's choice of the party in office is only consequential insofar as the executive and party do not take an action when not in office. The out-party actors' actions are also inconsequential, except when the executive (regardless of party) successfully violates the norm that also erodes the norm for the other party. Instead of entirely re-writing the strategies we simply present how the strategies differ in Model 2.

A.2.1 Executive

The adherence strategy for the executive in Model 2 is modified in the following way.

Stage 2–∞:

- Play d if h_{t-1} contains a period where the other party held office and the outcome was (Tx, Nx) or $(Tx, Dx, (1-q))$ for any x_i.

Any stage where Nature selects the other party:

- The action set is empty.
- Count the period as one of the τ punishment periods.

A.2.2 Party

The adherence strategy for the party in Model 2 is modified in the same way.

Stage 2–∞:

- Play d if h_{t-1} contains a period where the other party held office and the outcome was (Tx, Nx) or $(Tx, Dx, (1-q))$ for any x_i.

Any stage where Nature selects the other party:

- The action set is empty.
- Count the period as one of the τ punishment periods.

In Model 2, the executive and party strategies only change in minor ways. If the out party violates the norm (sending the game to an infinite sequence of violation substages) the actors always play defect. When Nature selects the other party to hold office, the out-party executive and party take no actions. Additionally, the players count the out-periods as part of the τ-punishment sequence.

B Derivation of Expected Utilities

Here we present the derivation of the various expected utilities presented in the main text.

Equation (1):

$$
\begin{aligned}
&Eu_P(\text{Punish}|\text{Transgress}) \\
&= (1-q)\left(\beta_P + \sum_{t=1}^{\infty} \delta_P^{t-1} \cdot 1\right) + q\left(\sum_{t=1}^{\tau} \delta_P^{t-1} \cdot 1 + \sum_{t=\tau+1}^{\infty} \delta_P^{t-1} \cdot x\right) \\
&= (1-q)\left(\beta_P + \frac{1}{1-\delta_P}\right) + q\left(\frac{1-\delta_P^{\tau}}{1-\delta_P} + \delta_P^{\tau}\left(\frac{x}{1-\delta_P}\right)\right)
\end{aligned}
$$

Equation (2):

$$
\begin{aligned}
Eu_P(\text{Not}|\text{Transgress}) &= \beta_P + \sum_{t=1}^{\infty} \delta_P^{t-1} \cdot 1 \\
&= \beta_P + \frac{1}{1-\delta_P}
\end{aligned}
$$

Equation (3):

$$
\begin{aligned}
Eu_E(\text{Adhere}|\text{Punish}) &= \sum_{t=1}^{\infty} \delta_E^{t-1} \cdot x \\
&= \frac{x}{1-\delta_E}
\end{aligned}
$$

Equation (4):

$$
\begin{aligned}
Eu_E(\text{Transgress}|\text{Not}) &= \beta_E + \sum_{t=1}^{\infty} \delta_E^{t-1} \cdot 1 \\
&= \beta_E + \frac{1}{1-\delta_E}
\end{aligned}
$$

Equation (5):

$$
\begin{aligned}
&Eu_E(\text{Transgress}|\text{Punish}) \\
&= (1-q)\left(\beta_E + \sum_{t=1}^{\infty} \delta_E^{t-1} \cdot 1\right) + q\left(\sum_{t=1}^{\tau} \delta_E^{t-1} \cdot 1 + \sum_{t=\tau+1}^{\infty} \delta_E^{t-1} \cdot x\right) \\
&= (1-q)\left(\beta_E + \frac{1}{1-\delta_E}\right) + q\left(\frac{1-\delta_E^{\tau}}{1-\delta_E} + \delta_E^{\tau}\left(\frac{x}{1-\delta_E}\right)\right)
\end{aligned}
$$

Equation (6):

$$
\begin{aligned}
&Eu_P(\text{Not}|\text{Transgress},(\text{Adhere},\text{Punish})) = Eu_P(\text{Not}|\text{Transgress},(\text{Adhere},\text{Not})) \\
&= \beta_P + 1 + \sum_{t=2}^{\infty} \left(\delta_P^{t-1}\left(p \cdot 1 + (1-p) \cdot 0\right)\right) = \beta_P + 1 + p\frac{\delta_P}{1-\delta_P}
\end{aligned}
$$

Equation (7):

$$
\begin{aligned}
&Eu_P(\text{Punish}|\text{Transgress},(\text{Adhere},\text{Punish})) = Eu_P(\text{Punish}|\text{Transgress},(\text{Adhere},\text{Not}))\\
&= (1-q)\left(\beta_P + 1 + \sum_{t=2}^{\infty}\left(\delta_P^{t-1}\left(p\cdot 1 + (1-p)\cdot 0\right)\right)\right)\\
&+ q\left(1 + \sum_{t=2}^{\tau}\left(\delta_P^{t-1}\left(p\cdot 1 + (1-p)\cdot 0\right)\right) + \sum_{t=\tau+1}^{\infty}\left(\delta_P^{t-1}\left(p\cdot x + (1-p)\cdot 0\right)\right)\right)\\
&= (1-q)\left(\beta_P + 1 + \frac{\delta_P\cdot p}{1-\delta_P}\right) + q\left(1 + \left(p\frac{\delta_P - \delta_P^{\tau}}{1-\delta_P}\right) + \left(p\frac{\delta_P^{\tau}x}{1-\delta_P}\right)\right)
\end{aligned}
$$

Equation (8):

$$
\begin{aligned}
&Eu_P(\text{Punish}|\text{Transgress},(\text{Transgress},\text{Not}))\\
&= (1-q)\left(\beta_P + 1 + \sum_{t=2}^{\infty}\left(\delta_P^{t-1}\left(p\cdot 1 + (1-p)\cdot 0\right)\right)\right)\\
&+ q\Bigg[1 + (1-p)\left(\delta_P\cdot 0 + \sum_{t=2}^{\infty}\left(\delta_P^{t}\left(p\cdot 1 + (1-p)\cdot 0\right)\right)\right)\\
&+ p\Bigg(\delta_P\cdot 1 + (1-p)\left(\delta_P^{2}\cdot 0 + \sum_{t=3}^{\infty}\left(\delta_P^{t}\left(p\cdot 1 + (1-p)\cdot 0\right)\right)\right) + p\Bigg((\dots)\\
&+ p\Bigg(\delta_P^{\tau-1}\cdot 1 + (1-p)\left(\delta_P^{\tau}\cdot 0 + \sum_{t=\tau+1}^{\infty}\left(\delta_P^{t}\left(p\cdot 1 + (1-p)\cdot 0\right)\right)\right)\\
&+ p\Bigg(\delta_P^{\tau}x + (1-p)\left(\delta_P^{\tau+1}\cdot 0 + \sum_{t=\tau+2}^{\infty}\left(\delta_P^{t}\left(p\cdot 1 + (1-p)\cdot 0\right)\right)\right) + p\Big((\dots)\Big)\Bigg)\Bigg)\Bigg)\Bigg)\Bigg]\\
&= (1-q)\left(\beta_P + 1 + \sum_{t=2}^{\infty}\delta_P^{t-1}p\right)\\
&+ q\Bigg[\left(1 + \delta_P(1-p)\left(\sum_{t=2}^{\infty}\delta_P^{t-1}p\right)\right)\sum_{t=1}^{\tau}\delta_P^{t-1}p^{t-1}\\
&+ \left(x + \delta_P(1-p)\left(\sum_{t=\tau+1}^{\infty}\delta_P^{t-\tau}p\right)\right)\sum_{t=\tau+1}^{\infty}\delta_P^{t-1}p^{t-1}\Bigg]\\
&= (1-q)\left(\beta_P + 1 + \frac{\delta_P p}{1-\delta_P}\right) + q\left(\frac{p^{\tau}\delta_P^{\tau}\left(x + \frac{(1-p)p\delta_P^{2}}{1-\delta_P}\right)}{1-\delta_P p} + \frac{\left(1 + \frac{(1-p)p\delta_P^{2}}{1-\delta_P}\right)(p^{\tau}\delta_P^{\tau} - 1)}{\delta_P p - 1}\right)
\end{aligned}
$$

Equation (9):

$$Eu_P(\text{Punish}|\text{Transgress})$$

$$= (1-q)\left(\beta_P + 1 + \sum_{t=2}^{\infty}\left(\delta_P^{t-1}\left(p\cdot 1 + (1-p)\cdot 0\right)\right)\right)$$

$$+ q\Bigg[1+(1-p)\Bigg(\delta_P\cdot 0 + q\left(\begin{cases}\sum_{t=2}^{\tau-1}\left(\delta_P^t\cdot p\cdot 1\right)+\sum_{t=\tau}^{\infty}\left(\delta_P^t\cdot p\cdot x\right), & \text{if } \tau\ge 2\\ \sum_{t=2}^{\infty}\left(\delta_P^t\cdot p\cdot x\right), & \text{else}\end{cases}\right) + (1-q)\sum_{t=2}^{\infty}\left(\delta_P^t\cdot p\cdot 1\right)\Bigg)$$

$$+ p\Bigg(\delta_P^{\tau-1}\cdot 1 + (1-p)\Bigg(\delta_P^{\tau}\cdot 0 + q\sum_{t=\tau+1}^{\infty}\left(\delta_P^t\cdot p\cdot x\right) + (1-q)\sum_{t=\tau+1}^{\infty}\left(\delta_P^t\cdot p\cdot 1\right)\Bigg)$$

$$+ p\Bigg(\delta_P^{\tau}x + (1-p)\Bigg(\delta_P^{\tau+1}\cdot 0 + q\sum_{t=\tau+2}^{\infty}\left(\delta_P^t\cdot p\cdot x\right) + (1-q)\sum_{t=\tau+2}^{\infty}\left(\delta_P^t\cdot p\cdot 1\right)\Bigg) + p\Bigg((\ldots)\Bigg)\Bigg)\Bigg)\Bigg)\Bigg)\Bigg]$$

$$= (1-q)\left(\beta_P + 1 + \sum_{t=2}^{\infty}\delta_P^{t-1}p\right)$$

$$+ q\Bigg[\Bigg(1+\delta_P(1-p)\Bigg(q\Bigg(\begin{cases}\sum_{t=2}^{\tau-1}\left(\delta_P^t\cdot p\cdot 1\right), & \text{if } \tau\ge 2\\ +\sum_{t=\tau}^{\infty}\left(\delta_P^t\cdot p\cdot x\right) & \\ \sum_{t=2}^{\infty}\left(\delta_P^t\cdot p\cdot x\right), & \text{else}\end{cases}\Bigg) + (1-q)\left(\sum_{t=2}^{\infty}\delta_P^{t-1}p\right)\Bigg)\Bigg)\sum_{t=1}^{\tau}\delta_P^{t-1}p^{t-1}$$

$$+ \Bigg(x+\delta_P(1-p)\Bigg(q\left(\sum_{t=\tau+1}^{\infty}\delta_P^{t-\tau}px\right) + (1-q)\left(\sum_{t=\tau+1}^{\infty}\delta_P^{t-\tau}p\right)\Bigg)\Bigg)\sum_{t=\tau+1}^{\infty}\delta_P^{t-1}p^{t-1}\Bigg]$$

$$\Rightarrow \text{If } \tau\ge 2 := (1-q)\left(\beta_P - \frac{\delta p}{\delta_P - 1} + 1\right)$$

$$+ q\Bigg(\frac{\delta_P^{\tau}p^{\tau}\left(\delta_P(1-p)\left(\frac{\delta_P p(1-q)}{1-\delta_P} + \frac{\delta_P pqx}{1-\delta_P}\right) + x\right)}{1-\delta_P p}$$

$$+ \frac{\left(\delta_P^{\tau}p^{\tau}-1\right)\left(\delta_P(1-p)\left(q\left(\frac{px\delta_P^{\tau}}{1-\delta_P} - \frac{p\left(\delta_P^2-\delta_P^{\tau}\right)}{\delta_P-1}\right) - \frac{\delta_P p(1-q)}{\delta_P-1}\right)+1\right)}{\delta_P p - 1}\Bigg)$$

$$\Rightarrow \text{If } \tau = 1 := (1-q)\left(\beta_P - \frac{\delta_P p}{\delta_P - 1} + 1\right)$$

$$+ q\left(\frac{\delta_P^{\tau}p^{\tau}\left(\delta_P(1-p)\left(\frac{\delta_P p(1-q)}{1-\delta_P} + \frac{\delta_P pqx}{1-\delta_P}\right) + x\right)}{1-\delta_P p} + \frac{\left(\delta_P^{\tau}p^{\tau}-1\right)\left(\delta_P(1-p)\left(-\frac{\delta_P p(1-q)}{\delta_P-1} - \frac{\delta_P^2 pqx}{\delta_P-1}\right)+1\right)}{\delta_P p - 1}\right)$$

Equation (10):

$$
\begin{aligned}
&Eu_E(\text{Adhere}|\text{Punish},(\text{Adhere},\text{Punish})) = Eu_E(\text{Adhere}|\text{Not},(\text{Adhere},\text{Punish}))\\
&Eu_E(\text{Adhere}|\text{Punish},(\text{Adhere},\text{Not})) = Eu_E(\text{Adhere}|\text{Not},(\text{Adhere},\text{Not}))\\
&= x + \sum_{t=2}^{\infty}\left(\delta_E^{t-1}\left(p\cdot x + (1-p)\cdot 0\right)\right)\\
&= x + p\frac{\delta_E x}{1-\delta_E}.
\end{aligned}
$$

Equation (11):

$$
\begin{aligned}
&Eu_E(\text{Transgress}|\text{Not},(\text{Adhere},\text{Punish})) = Eu_E(\text{Transgress}|\text{Not},(\text{Adhere},\text{Not}))\\
&= \beta_E + 1 + \sum_{t=2}^{\infty}\left(\delta_E^{t-1}\left(p\cdot 1 + (1-p)\cdot 0\right)\right)\\
&= \beta_E + 1 + p\frac{\delta_E}{1-\delta_E}.
\end{aligned}
$$

Equation (12):

$$
\begin{aligned}
&Eu_E(\text{Transgress}|\text{Punish},(\text{Adhere},\text{Punish})) = Eu_E(\text{Transgress}|\text{Punish},(\text{Adhere},\text{Not}))\\
&= (1-q)\left(\beta_E + 1 + \sum_{t=2}^{\infty}\left(\delta_E^{t-1}\left(p\cdot 1 + (1-p)\cdot 0\right)\right)\right)\\
&+ q\left(1 + \sum_{t=2}^{\tau}\left(\delta_E^{t-1}\left(p\cdot 1 + (1-p)\cdot 0\right)\right) + \sum_{t=\tau+1}^{\infty}\left(\delta_E^{t-1}\left(p\cdot x + (1-p)\cdot 0\right)\right)\right)\\
&= (1-q)\left(\beta_E + 1 + p\frac{\delta_E}{1-\delta_E}\right) + q\left(1 + \left(p\frac{\delta_E - \delta_E^{\tau}}{1-\delta_E}\right) + \left(p\frac{\delta_E^{\tau} x}{1-\delta_E}\right)\right)
\end{aligned}
$$

Equation (13):

$$
\begin{aligned}
&Eu_E(\text{Adhere}|\text{Punish}, (\text{Transgress}, \text{Not})) \\
&= \left(x + \delta_E(1-p)\left(\sum_{t=2}^{\infty} \delta_E^{t-1} p\right)\right) \sum_{t=1}^{\infty} \delta_E^{t-1} p^{t-1} \\
&= \frac{x + \frac{\delta_E^2 p(1-p)}{1-\delta_E}}{1-\delta_E p}
\end{aligned}
$$

Equation (14):

$$
\begin{aligned}
&Eu_E(\text{Transgress}|\text{Punish}, (\text{Transgress}, \text{Not})) \\
&= (1-q)\left(\beta_E + 1 + \sum_{t=2}^{\infty} \left(\delta_E^{t-1}\left(p \cdot 1 + (1-p) \cdot 0\right)\right)\right) \\
&+ q\Bigg[1 + (1-p)\left(\delta_E \cdot 0 + \sum_{t=2}^{\infty} \left(\delta_E^{t}\left(p \cdot 1 + (1-p) \cdot 0\right)\right)\right) \\
&+ p\Bigg(\delta_E \cdot 1 + (1-p)\left(\delta_E^2 \cdot 0 + \sum_{t=3}^{\infty} \left(\delta_E^{t}\left(p \cdot 1 + (1-p) \cdot 0\right)\right)\right) + p\Bigg((\ldots) \\
&+ p\Bigg(\delta_E^{\tau-1} \cdot 1 + (1-p)\left(\delta_E^{\tau} \cdot 0 + \sum_{t=\tau+1}^{\infty} \left(\delta_E^{t}\left(p \cdot 1 + (1-p) \cdot 0\right)\right)\right) \\
&+ p\Bigg(\delta_E^{\tau} \cdot x + (1-p)\left(\delta_E^{\tau+1} \cdot 0 + \sum_{t=\tau+2}^{\infty} \left(\delta_E^{t}\left(p \cdot 1 + (1-p) \cdot 0\right)\right)\right) + p\Bigg((\ldots)\Bigg)\Bigg)\Bigg)\Bigg)\Bigg)\Bigg] \\
&= (1-q)\left(\beta_E + 1 + \sum_{t=2}^{\infty} \delta_E^{t-1} p\right) \\
&+ q\Bigg[\left(1 + \delta_E(1-p)\left(\sum_{t=2}^{\infty} \delta_E^{t-1} p\right)\right) \sum_{t=1}^{\tau} \delta_E^{t-1} p^{t-1} \\
&+ \left(x + \delta_E(1-p)\left(\sum_{t=\tau+1}^{\infty} \delta_E^{t-\tau} p\right)\right) \sum_{t=\tau+1}^{\infty} \delta_E^{t-1} p^{t-1}\Bigg] \\
&= (1-q)\left(\beta_E + 1 + \frac{\delta_E p}{1-\delta_E}\right) + q\left(\frac{p^{\tau}\delta_E^{\tau}\left(x + \frac{(1-p)p\delta_E^2}{1-\delta_E}\right)}{1-\delta_E p} + \frac{\left(1 + \frac{(1-p)p\delta_E^2}{1-\delta_E}\right)\left(p^{\tau}\delta_E^{\tau} - 1\right)}{\delta_E p - 1}\right)
\end{aligned}
$$

Equation (15):

$$
\begin{aligned}
&Eu_E(\text{Adhere}|\text{Punish}) \\
&= \left(x + \delta_E(1-p)\left((1-q)\left(\sum_{t=2}^{\infty} \delta_E^{t-1} p\right) + q\left(\sum_{t=2}^{\infty} \delta_E^{t-1} px\right)\right)\right) \sum_{t=1}^{\infty} \delta_E^{t-1} p^{t-1} \\
&= \frac{x + \delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E - 1} - \frac{\delta_E pqx}{\delta_E - 1}\right)}{1 - \delta_E p}
\end{aligned}
$$

Equation (16):

$Eu_E(\text{Transgress}|\text{Punish})$

$$= (1-q)\left(\beta_E + 1 + \sum_{t=2}^{\infty}\left(\delta_E^{t-1}\left(p\cdot 1 + (1-p)\cdot 0\right)\right)\right)$$

$$+q\Bigg[1+(1-p)\left(\delta_E\cdot 0 + q\left(\begin{cases}\sum_{t=2}^{\tau-1}\left(\delta_E^t\cdot p\cdot 1\right)+\sum_{t=\tau}^{\infty}\left(\delta_E^t\cdot p\cdot x\right), & \text{if } \tau\geq 2\\ \sum_{t=2}^{\infty}\left(\delta_E^t\cdot p\cdot x\right), & \text{else}\end{cases}\right)+(1-q)\sum_{t=2}^{\infty}\left(\delta_E^t\cdot p\cdot 1\right)\right)$$

$$+p\left(\delta_E^{\tau-1}\cdot 1+(1-p)\left(\delta_E^{\tau}\cdot 0+q\sum_{t=\tau+1}^{\infty}\left(\delta_E^t\cdot p\cdot x\right)+(1-q)\sum_{t=\tau+1}^{\infty}\left(\delta_E^t\cdot p\cdot 1\right)\right)\right.$$

$$\left.+p\left(\delta_E^{\tau}x+(1-p)\left(\delta_E^{\tau+1}\cdot 0+q\sum_{t=\tau+2}^{\infty}\left(\delta_E^t\cdot p\cdot x\right)+(1-q)\sum_{t=\tau+2}^{\infty}\left(\delta_E^t\cdot p\cdot 1\right)\right)+p\Big((\ldots)\Big)\right)\right)\Bigg]$$

$$= (1-q)\left(\beta_E+1+\sum_{t=2}^{\infty}\delta_E^{t-1}p\right)$$

$$+q\left[\left(1+\delta_E(1-p)\left(q\left(\begin{cases}\sum_{t=2}^{\tau-1}\left(\delta_E^t\cdot p\cdot 1\right), & \text{if } \tau\geq 2\\ +\sum_{t=\tau}^{\infty}\left(\delta_E^t\cdot p\cdot x\right) & \\ \sum_{t=2}^{\infty}\left(\delta_E^t\cdot p\cdot x\right), & \text{else}\end{cases}\right)+(1-q)\left(\sum_{t=2}^{\infty}\delta_E^{t-1}p\right)\right)\right)\sum_{t=1}^{\tau}\delta_E^{t-1}p^{t-1}\right.$$

$$\left.+\left(x+\delta_E(1-p)\left(q\left(\sum_{t=\tau+1}^{\infty}\delta_E^{t-\tau}px\right)+(1-q)\left(\sum_{t=\tau+1}^{\infty}\delta_E^{t-\tau}p\right)\right)\right)\sum_{t=\tau+1}^{\infty}\delta_E^{t-1}p^{t-1}\right]$$

$$\Rightarrow \text{If } \tau\geq 2 := (1-q)\left(\beta_E-\frac{\delta p}{\delta_E-1}+1\right)$$

$$+q\left(\frac{\delta_E^{\tau}p^{\tau}\left(\delta_E(1-p)\left(\frac{\delta_E p(1-q)}{1-\delta_E}+\frac{\delta_E pqx}{1-\delta_E}\right)+x\right)}{1-\delta_E p}\right.$$

$$\left.+\frac{\left(\delta_E^{\tau}p^{\tau}-1\right)\left(\delta_E(1-p)\left(q\left(\frac{px\delta_E^{\tau}}{1-\delta_E}-\frac{p\left(\delta_E^2-\delta_E^{\tau}\right)}{\delta_E-1}\right)-\frac{\delta_E p(1-q)}{\delta_E-1}\right)+1\right)}{\delta_E p-1}\right)$$

$$\Rightarrow \text{If } \tau=1 := (1-q)\left(\beta_E-\frac{\delta_E p}{\delta_E-1}+1\right)$$

$$+q\left(\frac{\delta_E^{\tau}p^{\tau}\left(\delta_E(1-p)\left(\frac{\delta_E p(1-q)}{1-\delta_E}+\frac{\delta_E pqx}{1-\delta_E}\right)+x\right)}{1-\delta_E p}+\frac{\left(\delta_E^{\tau}p^{\tau}-1\right)\left(\delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E-1}-\frac{\delta_E^2 pqx}{\delta_E-1}\right)+1\right)}{\delta_E p-1}\right)$$

C Proofs

Lemma 1. *The party will punish a norm transgression in any period as long as:*

$$\delta_P^\tau \left(\frac{x-1}{1-\delta_P} \right) \geq \beta_P.$$

Proof of Lemma 1. The party prefers to punish a norm transgression whenever their expected utility from doing so is (weakly) greater than that from not punishing.

$$\begin{aligned} Eu_P(\text{Punish}|\text{Transgress}) &\geq Eu_P(\text{Not}|\text{Transgress}) \\ \Rightarrow (1-q)\left(\beta_P + \frac{1}{1-\delta_P}\right) + q\left(\frac{1-\delta_P^\tau}{1-\delta_P} + \delta_P^\tau\left(\frac{x}{1-\delta_P}\right)\right) &\geq \beta_P + \frac{1}{1-\delta_P} \\ \Rightarrow \delta_P^\tau\left(\frac{x-1}{1-\delta_P}\right) &\geq \beta_P. \end{aligned}$$

This is sufficient to establish the condition necessary for the party to prefer punishment over non-punishment. When the above condition is not met, the party prefers to not punish (when $Eu_P(\text{Punish}|\text{Transgress}) < Eu_P(\text{Not}|\text{Transgress})$).

■

Lemma 2. *When the party does not punish transgressions, the executive transgresses if:*

$$\frac{x-1}{1-\delta_E} \leq \beta_E.$$

When the party does punish transgressions, the executive transgresses if:

$$\frac{x-1}{1-q}\left(\frac{1-\delta_E^\tau}{1-\delta_E}\right) + \delta_E^\tau\left(\frac{x-1}{1-\delta_E}\right) \leq \beta_E.$$

Proof of Lemma 2. First consider the case where the party is playing a strategy of not punishing transgressions. In this case, the executive prefers to transgress if her expected utility from a norm violation (which will happen with probability one if the party does not discipline) is greater than her expected utility from adhering to the norm:

$$\begin{aligned}
& Eu_E(\text{Adhere}|\text{Punish}) \leq Eu_E(\text{Transgress}|\text{Not}) \\
& \Rightarrow \frac{x}{1-\delta_E} \leq \beta_E + \frac{1}{1-\delta_E} \\
& \Rightarrow \frac{x-1}{1-\delta_E} \leq \beta_E.
\end{aligned}$$

Next consider the case where the party is playing a strategy of punishing transgressions. Now the executive only prefers to attempt a transgression if her expected utility following a transgression and party discipline (which follows from the party's strategy) is greater than that of adhering:

$$\begin{aligned}
& Eu_E(\text{Adhere}|\text{Punish}) \leq Eu_E(\text{Transgress}|\text{Punish}) \\
& \Rightarrow \frac{x}{1-\delta_E} \leq (1-q)\left(\beta_E + \frac{1}{1-\delta_E}\right) + q\left(\frac{1-\delta_E^\tau}{1-\delta_E} + \delta_E^\tau\left(\frac{x}{1-\delta_E}\right)\right) \\
& \Rightarrow \frac{x-1}{1-q}\left(\frac{1-\delta_E^\tau}{1-\delta_E}\right) + \delta_E^\tau\left(\frac{x-1}{1-\delta_E}\right) \leq \beta_E.
\end{aligned}$$

Note that for each condition, replacing the inequality with an equality produces the condition for which the executive is indifferent between her strategies, and reversing the inequality gives the case where the executive prefers to adhere. ■

Proposition 1: *With our restriction of attention to one-stage transgressions by the executive, and finite period punishments by the party in response to transgressions, we uncover two sets of equilibria in the intra-party game.*

- *Adherence equilibria, where the executive adheres to the norm in every period, of which there are two types:*
  - *Voluntary executive adherence, which occurs when the executive adheres to the norm regardless of the party's discipline strategy.*
  - *Party-induced adherence, which occurs when the executive adheres to the norm in order to avoid a punishment by her party, but she would otherwise prefer to violate the*

norm.

- *Transgression equilibria, where the executive transgresses the norm in the first period, of which there are also two types:*
  - *Transgression and punishment, where the executive transgresses the norm despite knowing that the party will discipline her for doing so.*
  - *Transgression and no punishment, where the executive transgresses the norm and the party effectively condones the transgression and does not punish the executive.*

Proof of Proposition 1. To uncover equilibria, we use the one-stage-deviation principle in infinite horizon games. We may apply this theorem in the case where our game is "continuous at infinity," which holds because the per-period payoffs are discounted ($\delta_i < 1$) and each per-period payoff is uniformly bounded (by either x or $\beta_i + 1$, depending on the parameter values). In particular, given the one-stage-deviation principle, it is sufficient to examine that there is no strategy s_i given the history h_{t-1} where $s_i'(h_{t-1}) > s_i(h_{t-1})$. Following from the proof of the one-stage-deviation principle by Fudenberg and Tirole (1991), essentially should there be a strategy s_i' that is a profitable deviation over s_i at the history h_t, it must be the case that there is another strategy s_i'' that is the same as s_i' except for a finite number of periods after h_t where this strategy aligns with s_i'. Then by induction, we may establish that a strategy is optimal given that it is one stage unimprovable.

Given our substantive focus, we only consider one form of an executive strategies comprised of either adherence to the norm in every period or a one-period attempt at transgression and party strategies either no punishment or a τ-period punishment. More specifically, the executive adherence strategy we consider consists of adherence in every period, as well as cooperation in every adherence substage and defection in every violation substage. The party's no punishment strategy consists of no discipline regardless of the executive's preceeding action, cooperation in every adherence substage, and defection in every violation substage. The party's punishment strategy consists of playing discipline, defection in τ adherence substages, followed by cooperation in all remaining adherence substages, and defection in every violation substage in any history where the executive transgressed and consists of playing no discipline, cooperation in all adherence substages, and defection in every violation substage in any history where the executive has not transgressed. Finally, the executive's transgression strategy consists of transgressing once (followed by adherence in any potential future periods where the executive must choose between between adherence and transgression), defection in τ adherence substages, followed by cooperation in all remaining adherence substages, and defection in every violation substage. Given the form of executive strategies we consider, the one-stage deviation principle can clearly be applied.

We proceed by conjecturing equilibria made up of combinations of these strategies and then derive the conditions under which they exist. There are many other conceivable punishment, no punishment, transgression, etc. strategies. Hence, we do not prove uniqueness of the following equilibria (as they are indeed *not* unique).

First, it is useful to solve for the Nash equilibria of the substage games. In any adherence substage, we have that $BR_i(c) = c$ and $BR_i(d) = d$ and so the set of pure strategy Nash equilibria are:

$$\{(c,c),\ (d,d)\}.$$

In any violation substage we have that $BR_i(c) = BR_i(d) = d$ (because $\underline{x} < 1$) and so the pure strategy Nash equilbrium is:

$$\{(d,d)\}.$$

Thus, any equilibrium must have the players playing either mutual cooperation or defection in adherence substages and mutual defection in violation substages.

We now move through the four equilibrium presented in the proposition.

The voluntary executive adherence equilibria requires that the executive adheres in every period, and moreover that they do so without threat of punishment from the party. The condition for the executive to adhere (Lemma 2) is stricter when the party does not punish transgressions, so this is the binding condition on the executive. When the executive plays adhere in response to the party not punishing, the party is indifferent between punishing and not and therefore these equilibria consist of the executive playing their adhere strategy and the party playing either the punishment or no punishment strategies. We thus have the following **voluntary executive adherence equilibria**

- Executive adhere and party punish: occurs when $\frac{x-1}{1-\delta_E} \geq \beta_E$
- Executive adhere and party not punish: occurs when $\frac{x-1}{1-\delta_E} \geq \beta_E$.

From Lemma 2 the executive wants to adhere regardless of the party's strategy when $\frac{x-1}{1-\delta_E} \geq \beta_E$, and thus does not have a profitable deviation. The party is always indifferent between playing the punishment and no punishment strategies when the executive adheres regardless. Moreover note that every substage (adherence in this case) consists of mutual best responses of cooperate.

The party-induced adherence equilibrium requires that the executive adheres, but only when there is a threat of punishment—with no punishment the executive would transgress. This also means that the party must want to punish a transgressing executive. For the executive, it follows from Lemma 2 that

$$\frac{x-1}{1-q}\left(\frac{1-\delta_E^\tau}{1-\delta_E}\right)+\delta_E^\tau\left(\frac{x-1}{1-\delta_E}\right) \geq \beta_E \geq \frac{x-1}{1-\delta_E},$$

must be true for the executive to have no profitable deviation when the party is playing their punishment strategy, and for the executive to want to play transgress if the party were not playing punish. Similarly, it follows from Lemma 1 that

$$\delta_P^\tau\left(\frac{x-1}{1-\delta_P}\right) \geq \beta_P$$

must be true for the party to have no profitable deviation. The players are also mutually best responding in the substages. These conditions on β_E and β_P together produce the party-induced adherence equilibrium.

The transgression and punishment equilibrium requires that the executive transgresses the norm and the party punishes that transgression. From Lemma 2 if,

$$\frac{x-1}{1-q}\left(\frac{1-\delta_E^\tau}{1-\delta_E}\right)+\delta_E^\tau\left(\frac{x-1}{1-\delta_E}\right)\leq\beta_E$$

the executive's utility is maximized by transgressing, even with the party punishing. From Lemma 1,

$$\delta_P^\tau\left(\frac{x-1}{1-\delta_P}\right)\geq\beta_P$$

the party's utility is maximized by punishing the executive following a transgression. The actors are again best responding in the substages, and when these conditions on β_E and βP hold, neither the executive nor the party have profitable deviations from the prescribed strategies and the transgression and punishment equilibrium is obtained.

Finally the transgression and no punishment equilibrium similarly requires that the executive transgresses the norm, but now that the party prefers not to punish. From Lemma 2 if,

$$\frac{x-1}{1-\delta_E}\geq\beta_E$$

the executive transgresses when the party does not punish. From Lemma 1, if,

$$\delta_P^\tau\left(\frac{x-1}{1-\delta_P}\right)\leq\beta_P$$

then the party does not punish norm transgressions. When these conditions exist together, neither actor has a profitable deviation from transgression (in the case of the executive) or no punishment (in the case of the party) and hence we are in the transgression and no punishment equilibrium. ■

Lemma 3. *The party will punish a norm transgression in any period when the out-party executive adheres to the norm as long as*

$$p\delta_P^\tau\left(\frac{x-1}{1-\delta_P}\right)\geq\beta_P.$$

Proof of Lemma 3. The party prefers to punish a norm transgression whenever their expected utility from doing so is greater than that from not punishing. In Model 2, when the out-party executive adheres, we have:[26]

$$
\begin{aligned}
& Eu_P(\text{Punish}|\text{Transgress}, (\text{Adhere}, \text{Punish})) \geq Eu_P(\text{Not}|\text{Transgress}, (\text{Adhere}, \text{Punish})) \\
& \Rightarrow (1-q)\left(\beta_P + 1 + p\frac{\delta_P}{1-\delta_P}\right) + q\left(1 + \left(p\frac{\delta_P - \delta_P^\tau}{1-\delta_P}\right) + \left(p\frac{\delta_P^\tau x}{1-\delta_P}\right)\right) \\
& \geq \beta_P + 1 + p\frac{\delta_P}{1-\delta_P} \\
& \Rightarrow q\left(1 + \left(p\frac{\delta_P - \delta_P^\tau}{1-\delta_P}\right) + \left(p\frac{\delta_P^\tau x}{1-\delta_P}\right)\right) \geq q\left(\beta_P + 1 + p\frac{\delta_P}{1-\delta_P}\right) \\
& \Rightarrow 1 + \left(p\frac{\delta_P - \delta_P^\tau}{1-\delta_P}\right) + \left(p\frac{\delta_P^\tau x}{1-\delta_P}\right) \geq \beta_P + 1 + p\frac{\delta_P}{1-\delta_P} \\
& \Rightarrow p\frac{\delta_P^\tau x - \delta_P^\tau}{1-\delta_P} \geq \beta_P \\
& \Rightarrow p\delta_P^\tau\left(\frac{x-1}{1-\delta_P}\right) \geq \beta_P.
\end{aligned}
$$

■

Lemma 4. *The party will punish a norm transgression in any period when the out-party executive violates the norm (the out-party executive transgresses and the party does not punish) if:*

$$
\frac{p^\tau \delta_P^\tau \left(x + \frac{(1-p)p\delta_P^2}{1-\delta_P}\right)}{1-\delta_P p} + \frac{\left(1 + \frac{(1-p)p\delta_P^2}{1-\delta_P}\right)(p^\tau \delta_P^\tau - 1)}{\delta_P p - 1} - 1 - \frac{\delta_P p}{1-\delta_P} \geq \beta_P.
$$

Proof of Lemma 4. The party's expected utility from disciplining their executive following a transgression when the out-party will successfully violate the norm once in office is greater than the

[26]Note that this is equivalent with $Eu_P(\text{Punish}|\text{Transgress}, (\text{Adhere}, \text{Not})) \geq Eu_P(\text{Not}|\text{Transgress}, (\text{Adhere}, \text{Not}))$ instead.

party's expected utility from not disciplining their executive if:

$$
\begin{aligned}
&Eu_P(\text{Punish}|\text{Transgress},(\text{Transgress},\text{Not})) \geq Eu_P(\text{Not}|\text{Transgress},(\text{Transgress},\text{Not}))\\
&(1-q)\left(\beta_P+1+\frac{\delta_P p}{1-\delta_P}\right)+q\left(\frac{p^\tau\delta_P^\tau\left(x+\frac{(1-p)p\delta_P^2}{1-\delta_P}\right)}{1-\delta_P p}+\frac{\left(1+\frac{(1-p)p\delta_P^2}{1-\delta_P}\right)(p^\tau\delta_P^\tau-1)}{\delta_P p-1}\right)\\
&\geq \beta_P+1+p\frac{\delta_P}{1-\delta_P}\\
&\Rightarrow \frac{p^\tau\delta_P^\tau\left(x+\frac{(1-p)p\delta_P^2}{1-\delta_P}\right)}{1-\delta_P p}+\frac{\left(1+\frac{(1-p)p\delta_P^2}{1-\delta_P}\right)(p^\tau\delta_P^\tau-1)}{\delta_P p-1}-1-\frac{\delta_P p}{1-\delta_P}\geq \beta_P.
\end{aligned}
$$

■

Lemma 5. *The party will punish their executive's norm transgression in any period when the out-party executive is expected to transgress the norm and be punished by their party (the out-party) if Condition 1 holds (see appendix).*

Proof of Lemma 5. The party's expected utility from disciplining their executive following a transgression when the out-party executive will attempt a norm violation once in office (out-party's executive transgress and out-party punish) is greater than the party's expected utility from not disciplining their executive if condition 1 (below) holds:

$$Eu_P(\text{Punish}|\text{Transgress}) \geq Eu_P(\text{Not}|\text{Transgress})$$

If $\tau \geq 2$,

$$
\begin{aligned}
&(1-q)\left(\beta_P - \frac{\delta p}{\delta_P - 1} + 1\right) \\
&+q\left(\frac{\delta_P^\tau p^\tau \left(\delta_P(1-p)\left(\frac{\delta_P p(1-q)}{1-\delta_P} + \frac{\delta_P pqx}{1-\delta_P}\right) + x\right)}{1-\delta_P p}\right. \\
&\left.+\frac{(\delta_P^\tau p^\tau - 1)\left(\delta_P(1-p)\left(q\left(\frac{px\delta_P^\tau}{1-\delta_P} - \frac{p\left(\delta_P^2 - \delta_P^\tau\right)}{\delta_P - 1}\right) - \frac{\delta_P p(1-q)}{\delta_P - 1}\right) + 1\right)}{\delta_P p - 1}\right) \\
&\geq \beta_P + 1 + p\frac{\delta_P}{1-\delta_P}
\end{aligned}
$$

If $\tau = 1$,

$$
\begin{aligned}
&(1-q)\left(\beta_P - \frac{\delta_P p}{\delta_P - 1} + 1\right) \\
&+q\left(\frac{\delta_P^\tau p^\tau \left(\delta_P(1-p)\left(\frac{\delta_P p(1-q)}{1-\delta_P} + \frac{\delta_P pqx}{1-\delta_P}\right) + x\right)}{1-\delta_P p} + \frac{(\delta_P^\tau p^\tau - 1)\left(\delta_P(1-p)\left(-\frac{\delta_P p(1-q)}{\delta_P - 1} - \frac{\delta_P^2 pqx}{\delta_P - 1}\right) + 1\right)}{\delta_P p - 1}\right) \\
&\geq \beta_P + 1 + p\frac{\delta_P}{1-\delta_P}
\end{aligned}
$$

■

Lemma 6. *When the out-party executive adheres to the norm, the in-party executive transgresses the norm when*

- *The party does not discipline, if*

$$\beta_E \geq x - 1 + p\delta_E\left(\frac{x-1}{1-\delta_E}\right),$$

- *Or, when the party disciplines, if*

$$\beta_E \geq \frac{x-q}{1-q} + p\frac{\delta_E x + q\delta_E^\tau - q\delta_E - q\delta_E^\tau x}{(1-\delta_E)(1-q)} - p\frac{\delta_E}{1-\delta_E} - 1.$$

Proof of Lemma 6. First, when the executive does not expect her party to punish a transgression, her expected utility from doing so is greater than adhering if (with the out-party executive adhering to norm):[27]

$$
\begin{aligned}
&Eu_E(\text{Transgress}|\text{Not}) \geq Eu_E(\text{Adhere}|\text{Not}) \\
&\beta_E + 1 + p\frac{\delta_E}{1-\delta_E} \geq x + p\frac{\delta_E x}{1-\delta_E} \\
&\Rightarrow \beta_E \geq x - 1 + p\delta_E\left(\frac{x-1}{1-\delta_E}\right).
\end{aligned}
$$

When the executive expects her party to punish her transgression, her expected utility from doing so is greater than her expected utility from adhering if (again, with the out-party executive expected to adhere to the norm):

$$
\begin{aligned}
&Eu_E(\text{Transgress}|\text{Punish}) \geq Eu_E(\text{Adhere}|\text{Punish}) \\
&(1-q)\left(\beta_E + 1 + p\frac{\delta_E}{1-\delta_E}\right) + q\left(1 + \left(p\frac{\delta_E - \delta_E^\tau}{1-\delta_E}\right) + \left(p\frac{\delta_E^\tau x}{1-\delta_E}\right)\right) \geq x + p\frac{\delta_E x}{1-\delta_E} \\
&\Rightarrow \beta_E \geq \frac{x-q}{1-q} + p\frac{\delta_E x + q\delta_E^\tau - q\delta_E - q\delta_E^\tau x}{(1-\delta_E)(1-q)} - p\frac{\delta_E}{1-\delta_E} - 1.
\end{aligned}
$$

■

Lemma 7. *When the out-party executive will successfully violate the norm upon taking office, the in-party executive transgresses the norm when*

- *The party does not discipline, if*

  $\beta_E \geq \frac{x-1}{1-\delta_E p}$,

- *or, when the party disciplines, if*

  $$\beta_E \geq \frac{x+\frac{\delta_E^2 p(1-p)}{1-\delta_E}}{(1-\delta_E p)(1-q)} - \frac{q}{1-q}\left(\frac{p^\tau \delta_E^\tau\left(x+\frac{(1-p)p\delta_E^2}{1-\delta_E}\right)}{1-\delta_E p} + \frac{\left(1+\frac{(1-p)p\delta_E^2}{1-\delta_E}\right)(p^\tau \delta_E^\tau - 1)}{\delta_E p - 1}\right) - 1 - \frac{\delta_E p}{1-\delta_E}.$$

[27]Omitting the out-party's strategy for reasons of conciseness, but they are either $(\text{Adhere}, \text{Punish})$ or $(\text{Adhere}, \text{Not})$.

Proof of Lemma 7. First, when the executive does not expect her party to punish a transgression, her expected utility from transgressing is greater than adhering if (with the out-party executive expected to successfully violate the norm once in office):

$$\begin{aligned}
&Eu_E(\text{Transgress}|(N^{\infty},c^{\infty},d^{\infty}) \geq Eu_E(\text{Adhere}|\text{Punish})\\
&\beta_E + 1 + p\frac{\delta_E}{1-\delta_E} \geq \frac{x+\frac{\delta_E^2 p(1-p)}{1-\delta_E}}{1-\delta_E p}\\
&\Rightarrow \beta_E \geq \frac{x-1}{1-\delta_E p}.
\end{aligned}$$

When the executive expects her party to punish her transgression, her expected utility from doing so is greater than her expected utility from adhering if (again, with the out-party executive expected to successfully violate the norm once in office):

$$\begin{aligned}
&Eu_E(\text{Transgress}|\text{Punish} \geq Eu_E(\text{Adhere}|\text{Punish})\\
&(1-q)\left(\beta_E+1+\frac{\delta_E p}{1-\delta_E}\right)+q\left(\frac{p^{\tau}\delta_E^{\tau}\left(x+\frac{(1-p)p\delta_E^2}{1-\delta_E}\right)}{1-\delta_E p}+\frac{\left(1+\frac{(1-p)p\delta_E^2}{1-\delta_E}\right)(p^{\tau}\delta_E^{\tau}-1)}{\delta_E p-1}\right)\\
&\geq \frac{x+\frac{\delta_E^2 p(1-p)}{1-\delta_E}}{1-\delta_E p}\\
&\Rightarrow \beta_E \geq \frac{x+\frac{\delta_E^2 p(1-p)}{1-\delta_E}}{(1-\delta_E p)(1-q)}\\
&-\frac{q}{1-q}\left(\frac{p^{\tau}\delta_E^{\tau}\left(x+\frac{(1-p)p\delta_E^2}{1-\delta_E}\right)}{1-\delta_E p}+\frac{\left(1+\frac{(1-p)p\delta_E^2}{1-\delta_E}\right)(p^{\tau}\delta_E^{\tau}-1)}{\delta_E p-1}\right)-1-\frac{\delta_E p}{1-\delta_E}.
\end{aligned}$$

■

Lemma 8. *When the out-party executive will attempt to violate the norm upon taking office (out-party executive transgresses and is disciplined by their party), the in-party executive transgresses the norm when their party does not discipline, if*

$$\beta_E \geq \frac{x + \delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E - 1} - \frac{\delta_E pqx}{\delta_E - 1}\right)}{1 - \delta_E p} - 1 - \frac{\delta_E p}{1 - \delta_E}$$

or, when their party disciplines according to Condition 2 (see appendix).

Proof of Lemma 8. First, when the executive does not expect her party to punish a transgression, her expected utility from transgressing is greater than adhering if (with the out-party executive expected to transgress the norm and be punished by their party once in office):

$$\begin{aligned}
&Eu_E(\text{Transgress}|\text{Not}) \geq Eu_E(\text{Adhere}|\text{Not}) \\
&\beta_E + 1 + p\frac{\delta_E}{1-\delta_E} \geq \frac{x + \delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E - 1} - \frac{\delta_E pqx}{\delta_E - 1}\right)}{1 - \delta_E p} \\
&\Rightarrow \beta_E \geq \frac{x + \delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E - 1} - \frac{\delta_E pqx}{\delta_E - 1}\right)}{1 - \delta_E p} - 1 - \frac{\delta_E p}{1 - \delta_E}.
\end{aligned}$$

When the executive expects her party to punish her transgression, her expected utility from doing so is greater than her expected utility from adhering if (again, with the out-party executive expected to successfully violate the norm once in office) according to Condition 2:

$Eu_E(\text{Transgress}|\text{Punish} \geq Eu_E(\text{Adhere}|\text{Punish})$

$$\Rightarrow \text{If } \tau \geq 2 := (1-q)\left(\beta_E - \frac{\delta p}{\delta_E - 1} + 1\right)$$

$$+q\left(\frac{\delta_E^\tau p^\tau \left(\delta_E(1-p)\left(\frac{\delta_E p(1-q)}{1-\delta_E} + \frac{\delta_E pqx}{1-\delta_E}\right) + x\right)}{1-\delta_E p}\right.$$

$$\left.+\frac{(\delta_E^\tau p^\tau - 1)\left(\delta_E(1-p)\left(q\left(\frac{px\delta_E^\tau}{1-\delta_E} - \frac{p\left(\delta_E^2 - \delta_E^\tau\right)}{\delta_E - 1}\right) - \frac{\delta_E p(1-q)}{\delta_E - 1}\right) + 1\right)}{\delta_E p - 1}\right)$$

$$\geq \frac{x + \delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E - 1} - \frac{\delta_E pqx}{\delta_E - 1}\right)}{1-\delta_E p}$$

$$\Rightarrow \text{If } \tau = 1 := (1-q)\left(\beta_E - \frac{\delta_E p}{\delta_E - 1} + 1\right)$$

$$+q\left(\frac{\delta_E^\tau p^\tau \left(\delta_E(1-p)\left(\frac{\delta_E p(1-q)}{1-\delta_E} + \frac{\delta_E pqx}{1-\delta_E}\right) + x\right)}{1-\delta_E p} + \frac{(\delta_E^\tau p^\tau - 1)\left(\delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E - 1} - \frac{\delta_E^2 pqx}{\delta_E - 1}\right) + 1\right)}{\delta_E p - 1}\right)$$

$$\geq \frac{x + \delta_E(1-p)\left(-\frac{\delta_E p(1-q)}{\delta_E - 1} - \frac{\delta_E pqx}{\delta_E - 1}\right)}{1-\delta_E p}.$$

■

Proposition 2: *With restriction of attention to one-stage transgressions by the in office executive and finite punishments by their party in response to transgressions, there are three sets of equilibria in the inter-party game (categorized by the out-party behavior), each with different forms that may include multiple unique equilibria.*

1. ***Out-party adherence:*** *When the out-party executive voluntarily adheres, the in-party executive will voluntarily adhere under some conditions, be induced to adhere under others, and transgress in others. The party will punish their executive for some parameter conditions and not for others.*

2. ***Out-party violation:*** *Similarly, when the out-party executive will successfully violate the norm, there are conditions that support the in-party executive voluntarily adhering to the norm, being induced to adhere to the norm, and transgressing the norm. Likewise, the party will at times punish transgressions and not in others. The conditions for adherence and punishment are stricter when the out-party violates than when they adhere.*

3. ***Out-party transgression and disciplined:*** *Finally, there is an intermediate region where the out-party disciplines an out-party executive transgressing when they enter office. Again, in this case, there are conditions that support the in-party executive voluntarily adhering to the norm, being induced to adhere to the norm, and transgressing the norm. Likewise, the party will at times punish transgressions and not in others.*

Proof of Proposition 2. For the proof, we consider the case where Party 1 and their executive are the in-party. The proof for Model 2's proposition largely follows the structure of Proposition 1. We again use the one-stage deviation principle to solve for equilibria—per-period payoffs are discounted δ_i, with $0 < \delta_i < 1$ $\forall i$ and uniformly bounded (again by either x or $\beta_i + 1$ depending on parameter values). Additionally we only consider the strategies outlined throughout the paper—"adherence" and "transgression" strategies for the executive and "punish" or "do not punish" strategies for the parties.

Out-party adherence: Now we present the necessary conditions for the four combinations of in-party strategy when the out-party adheres in equilibrium. These follow directly from the results derived in Lemmas 3 and 6. For the in-party,

- Voluntary Adherence
  - Executive: $\beta_{E1} < x - 1 + p\delta_{E1}\left(\frac{x-1}{1-\delta_{E1}}\right)$
  - Party: No restriction
- Induced Adherence
  - Executive: $x - 1 + p\delta_{E1}\left(\frac{x-1}{1-\delta_{E1}}\right) \leq \beta_{E1} < \frac{x-q}{1-q} + p\frac{\delta_{E1}x + q\delta_{E1}^{\tau} - q\delta_{E1} - q\delta_{E1}^{\tau}x}{(1-\delta_{E1})(1-q)} - p\frac{\delta_{E1}}{1-\delta_{E1}} - 1$

- Party: $p\delta_{P1}^{\tau}\left(\frac{x-1}{1-\delta_{P1}}\right) \geq \beta_{P1}$

- Transgression with Punishment
  - Executive: $\beta_{E1} \geq \frac{x-q}{1-q} + p\frac{\delta_{E1}x+q\delta_{E1}^{\tau}-q\delta_{E1}-q\delta_{E1}^{\tau}x}{(1-\delta_{E1})(1-q)} - p\frac{\delta_{E1}}{1-\delta_{E1}} - 1$
  - Party: $p\delta_{P1}^{\tau}\left(\frac{x-1}{1-\delta_{P1}}\right) \geq \beta_{P1}$
- Transgression without Punishment
  - Executive: $\beta_{E1} \geq x-1+p\delta_{E1}\left(\frac{x-1}{1-\delta_{E1}}\right)$
  - Party: $p\delta_{P1}^{\tau}\left(\frac{x-1}{1-\delta_{P1}}\right) < \beta_{P1}$

Turning to the out-party, as we only consider one transgression, the out-party knows that there will be adherence should it return to the (original) in-party. Hence, the out-party adheres voluntarily given: $\beta_{E2} < x-1+(1-p)\delta_{E2}\left(\frac{x-1}{1-\delta_{E2}}\right)$. There is induced adherence given: $x-1+(1-p)\delta_{E2}\left(\frac{x-1}{1-\delta_{E2}}\right) \leq \beta_{E2} < \frac{x-q}{1-q}+(1-p)\frac{\delta_{E2}x+q\delta_{E2}^{\tau}-q\delta_{E2}-q\delta_{E2}^{\tau}x}{(1-\delta_{E2})(1-q)} - (1-p)\frac{\delta_{E2}}{1-\delta_{E2}} - 1$ & $(1-p)\delta_{P2}^{\tau}\left(\frac{x-1}{1-\delta_{P2}}\right) \geq \beta_{P2}$.

Out-party violation: Now we present the necessary conditions for the four combinations of in-party strategy when the out-party violates the norm when in office (transgresses and is not punished) in equilibrium. These follow directly from the results derived in Lemmas 4 and 7. For the in-party,

- Voluntary Adherence
  - Executive: $\beta_{E1} < \frac{x-1}{1-\delta_{E1}p}$
  - Party: No restriction
- Induced Adherence

– Executive:

$$\frac{x-1}{1-\delta_{E1}p} \leq \beta_{E1} < \frac{x+\frac{\delta_{E1}^2 p(1-p)}{1-\delta_{E1}}}{(1-\delta_{E1}p)(1-q)} - \frac{q}{1-q}\left(\frac{p^\tau \delta_{E1}^\tau \left(x+\frac{(1-p)p\delta_{E1}^2}{1-\delta_{E1}}\right)}{1-\delta_{E1}p} + \frac{\left(1+\frac{(1-p)p\delta_{E1}^2}{1-\delta_{E1}}\right)(p^\tau \delta_{E1}^\tau - 1)}{\delta_{E1}p-1}\right) - 1 - \frac{\delta_{E1}p}{1-\delta_{E1}}$$

– Party: $\frac{p^\tau \delta_{P1}^\tau \left(x+\frac{(1-p)p\delta_{P1}^2}{1-\delta_{P1}}\right)}{1-\delta_{P1}p} + \frac{\left(1+\frac{(1-p)p\delta_{P1}^2}{1-\delta_{P1}}\right)(p^\tau \delta_{P1}^\tau - 1)}{\delta_{P1}p-1} - 1 - \frac{\delta_{P1}p}{1-\delta_{P1}} \geq \beta_{P1}$

- Transgression with Punishment

– Executive:

$$\beta_{E1} \geq \frac{x+\frac{\delta_{E1}^2 p(1-p)}{1-\delta_{E1}}}{(1-\delta_{E1}p)(1-q)} - \frac{q}{1-q}\left(\frac{p^\tau \delta_{E1}^\tau \left(x+\frac{(1-p)p\delta_{E1}^2}{1-\delta_{E1}}\right)}{1-\delta_{E1}p} + \frac{\left(1+\frac{(1-p)p\delta_{E1}^2}{1-\delta_{E1}}\right)(p^\tau \delta_{E1}^\tau - 1)}{\delta_{E1}p-1}\right) - 1 - \frac{\delta_{E1}p}{1-\delta_{E1}}$$

– Party: $\frac{p^\tau \delta_{P1}^\tau \left(x+\frac{(1-p)p\delta_{P1}^2}{1-\delta_{P1}}\right)}{1-\delta_{P1}p} + \frac{\left(1+\frac{(1-p)p\delta_{P1}^2}{1-\delta_{P1}}\right)(p^\tau \delta_{P1}^\tau - 1)}{\delta_{P1}p-1} - 1 - \frac{\delta_{P1}p}{1-\delta_{P1}} \geq \beta_{P1}$

- Transgression without Punishment

– Executive: $\beta_{E1} \geq \frac{x-1}{1-\delta_{E1}p}$

– Party: $\frac{p^\tau \delta_{P1}^\tau \left(x+\frac{(1-p)p\delta_{P1}^2}{1-\delta_{P1}}\right)}{1-\delta_{P1}p} + \frac{\left(1+\frac{(1-p)p\delta_{P1}^2}{1-\delta_{P1}}\right)(p^\tau \delta_{P1}^\tau - 1)}{\delta_{P1}p-1} - 1 - \frac{\delta_{P1}p}{1-\delta_{P1}} < \beta_{P1}$

Turning to the out-party, as we only consider one transgression, the out-party knows that there will be adherence should it return to the (original) in-party. Hence, the out-party transgress without punishment given: $\beta_{E2} \geq x-1+(1-p)\delta_{E2}\left(\frac{x-1}{1-\delta_{E2}}\right)$ & $(1-p)\delta_{P2}^\tau\left(\frac{x-1}{1-\delta_{P2}}\right) < \beta_{P2}$

Out-party transgression and disciplined: Now we present the necessary conditions for the four combinations of in-party strategy when the out-party executive transgresses and the out-party disciplines. These follow directly from the results derived in Lemmas 5 and 8.

- Voluntary Adherence
  - Executive: $\beta_{E1} < \frac{x+\delta_{E1}(1-p)\left(-\frac{\delta_{E1}p(1-q)}{\delta_{E1}-1}-\frac{\delta_{E1}pqx}{\delta_{E1}-1}\right)}{1-\delta_{E1}p} - 1 - \frac{\delta_{E1}p}{1-\delta_{E1}}$
  - Party: No restriction
- Induced Adherence
  - Executive: Opposite inequality sign on Condition 2 from proof of Lemma 8
  - Party: Condition 1 from proof of Lemma 5
- Transgression with Punishment
  - Executive: Condition 2 from proof of Lemma 8
  - Party: Condition 1 from proof of Lemma 5
- Transgression without Punishment
  - Executive: $\beta_{E1} \geq \frac{x+\delta_{E1}(1-p)\left(-\frac{\delta_{E1}p(1-q)}{\delta_{E1}-1}-\frac{\delta_{E1}pqx}{\delta_{E1}-1}\right)}{1-\delta_{E1}p} - 1 - \frac{\delta_{E1}p}{1-\delta_{E1}}$
  - Party: Opposite inequality sign on Condition 1 from proof of Lemma 5

Turning to the out-party, as we only consider one transgression, the out-party knows that there will be adherence should it return to the (original) in-party. Hence, the out-party transgress without punishment adherence given: $\beta_{E2} \geq x - 1 + (1-p)\delta_{E2}\left(\frac{x-1}{1-\delta_{E2}}\right)$ & $(1-p)\delta_{P2}^{\tau}\left(\frac{x-1}{1-\delta_{P2}}\right) < \beta_{P2}$

This completes the proof for the combinations of strategies under consideration. ■

D Derivation of Minimal Credible Punishment

Here, we demonstrate how to solve for the effective and credible τ punishment length.

D.1 Model 1

The executive is indifferent between transgressing and adhering when there is a threat of discipline for τ periods, $Eu_E(\text{Adhere}|\text{Punish}) = Eu_E(\text{Transgress}|\text{Punish})$ if:

$$\frac{x-1}{1-q}\left(\frac{1-\delta_E^{\tau}}{1-\delta_E}\right)+\delta_E^{\tau}\left(\frac{x-1}{1-\delta_E}\right)=\beta_E.$$

Solving for τ:

$$\tau = \frac{\ln\left(\frac{x-1-\beta_E+q\beta_E+\beta_E\delta_E-q\beta_E\delta_E}{q(x-1)}\right)}{\ln(\delta_E)}.$$

But recall that $\tau \in \mathbb{Z}^+$. If the above τ is between two integers, τ^{n-1} and τ^n and the party were to punish for τ^{n-1} periods, the executive would not be deterred from transgressing. Thus, take the ceiling function of the above equality:

$$\hat{\tau} = \left\lceil \frac{\ln\left(\frac{x-1-\beta_E+q\beta_E+\beta_E\delta_E-q\beta_E\delta_E}{q(x-1)}\right)}{\ln(\delta_E)} \right\rceil.$$

Similarly, we can find the maximum punishment length that the party is willing to endure in order to deter their executive. The party is indifferent between punishing for τ periods and not disciplining when:

$$\delta_P^{\tau}\left(\frac{x-1}{1-\delta_P}\right)=\beta_P.$$

Solving for τ:

$$\tau = \frac{\ln\left(\beta_P\left(\frac{1-\delta_P}{x-1}\right)\right)}{\ln(\delta_P)}.$$

But if τ is between two integers the party will only punish if at the lower value, so we take the floor function of the above value:

$$\hat{\hat{\tau}} = \left\lfloor \frac{\ln\left(\beta_P\left(\frac{1-\delta_P}{x-1}\right)\right)}{\ln(\delta_P)} \right\rfloor.$$