Embracing the Enemy*

Álvaro Delgado-Vega† Johannes Schneider‡

June 2024

Abstract

We study an organization with a principal and two agents. All three have a long-run agenda which drives their repeated interactions. The principal can influence the competition for agency by endorsing an agent. Her agenda is more aligned with her "friend" than with her "enemy." Even when fully aligned with the friend, the principal embraces the enemy by persistently endorsing him once an initial "cordon sanitaire" to exclude the enemy breaks exogenously. A dynamically optimizing principal with extreme agenda either implements the commitment solution or reverts to static Nash. For less extreme principals, losing commitment power has more gradual effects.

JEL Codes: D23, D74, D86

Keywords: Dynamic Games, Relational Contracts, Horizontal Moral Hazard, Organizational Economics, Coalition Formation, Optimal Endorsement

Politics is the art of the possible, the attainable—the art of the next best.

Otto von Bismarck

1 Introduction

Organizations grapple with disagreements. In a coalitional government, political parties with conflicting policy preferences have to work together. In firms, divisions disagree about project prioritization. At times, this disagreement is exacerbated when a particular agent with vastly different preferences enters the arena. They may come as non-establishment political groups, newly acquired divisions, or revolutionary managers. Regardless of their form, these agents compete for influence with the establishment, thereby challenging the organization's decision-making process.

*We are grateful to Antoine Loeper for impactful discussions at the beginning of this project. We thank Paweł Doligalski, Dana Foarta, Gabriele Gratton, Elia Sartori, Christoph Schottmüller, Jonathan B. Slapin, Takuo Sugaya, Heidi Thysen, Jan Zapal, and audiences at various places.

†ETH Zürich; alvaro.delgado@mtec.ethz.ch

‡Universidad Carlos III de Madrid; jschneid@eco.uc3m.es

When repeatedly dealing with such fundamental differences, *who* takes the lead is not only important for its own sake, but also because it determines *how* the organization moves forward. In these struggles, we observe "power brokers" who mediate between competitors—for example, a centrist party, or a conciliatory CEO. In the situations we imagine, such a power broker's opinion is often situated between the most extreme agents. What strategy would a power broker employ? Is it prudent to establish a *cordon sanitaire* around a discordant agent, akin to the one sometimes created by mainstream parties when dealing with radical entrants? Or should she, on the contrary, back the extremist to incentivize his moderation?

To address these questions, we consider an infinite-horizon setting where two agents repeatedly compete for the right to determine policy. Both are power-hungry and agenda-driven. A third player, the principal, has an agenda too and solely cares about that. The principal's ideal policy is between the ideals of the agents, closer to one (her "friend") than to the other (her "enemy").

The principal cannot choose policies herself, but influences who gets the right to act: she can increase an agent's chances of taking the lead, yet cannot fully determine the allocation of power. The notion we aim to capture is that sometimes a centrist party or a CEO can determine who is in charge. In other occasions, factors such as parliamentary numbers or the available talent leave no choice whom to grant agency.

We study the dynamics of the principal's optimal contract. Our results deliver three distinct predictions: (i) if an enemy appears, at first maintaining a cordon sanitaire against him is optimal; (ii) eventually, that cordon sanitaire will fall, and (iii) once it fell, it is not optimal to restore it. From then on, the principal will "embraces the enemy".

This analysis unveils an inherently dynamic logic. If the game was played only once, whichever agent was given the right to act (the "leading agent"), would choose his preferred policy leading to extreme policies and polarized behavior. Anticipating this polarization, the principal would want to maximize the likelihood that the agent closer to her preferences takes the lead: she would fully endorse the friend. However, if the same interaction occurs repeatedly, a permanent cordon sanitaire around the enemy would severely limit the principal's influence. To see why, suppose the enemy becomes the leading agent today. Then, the principal can offer him the following contract: "If you choose a policy close to my bliss point today, I will endorse you the next period."[1]

We show that this exchange of policy moderation for endorsement is not only effective in disciplining the enemy, but furthermore, it becomes most relevant when the friend's agenda fully aligns with the principal's. In this case, the principal's behavior radically changes when the enemy takes the lead for the first time. As long as only the friend leads, the principal optimally endorses her friend keeping a cordon sanitaire around her enemy. However, by design, the principal cannot permanently exclude the enemy from taking

[1]For the sake of the argument only, assume the principal (but no agent) has commitment power.

the lead. Once the enemy leads for the first time, the principal immediately shifts her strategy. From that moment onward, the principal endorses the enemy persistently in exchange for the enemy's moderation. Since the friend fully agrees with the principal, the friend will still choos the principal's preferred policy.

If the principal and the friend are not fully aligned, the principal faces an additional challenge: she also needs to incentivize the friend to pick policies closer to her bliss point. The principal can make use of two "carrots" to achieve that goal. She can either (i) promise endorsement to the friend (which benefits the friend power- and policy-wise), or (ii) promise to use her endorsement to moderate the enemy (which benefits the friend policy-wise). If the policy disagreement between the principal and the friend is sufficiently large, the first carrot is needed as well so the friend also receives some endorsement. In terms of payoffs, a principal who has some disagreement with the friend performs better than a principal who fully aligns with him or disagrees equally strong with both agents.

Our predictions relate to the well-known *Ally Principle*, by which a principal always prefers to delegate to her closest friend (Bendor, Glazer, and Hammond, 2001; Bendor and Meirowitz, 2004; Callander et al., 2008). We offer a nuanced message: The ally principle is not invalidated in general, but it is suboptimal to uphold it forever. Indeed, the behavior we outlined resembles well with the real-world. Consider, for example, the way parliamentary democracies often deal with an extremist new party. At first, mainstream parties try to exclude it. However, once the newcomer attains some power, they switch to an embracing strategy to moderate it. Often, the extremist then becomes part of the political establishment.[2] Similarly, as Xuan (2009) shows that CEOs appear as persistent bridge-builders by favoring divisions distant from their background.

On a conceptual level, we depart from a canonical repeated moral hazard settings in various ways. The crucial assumption is that we consider a setting in which multiple agents have to work together in a community of fate. This assumption implies that we must be more precise in our notion of moral hazard. We assume that moral hazard in our setting does not come from the agent's desire to shirk, but rather from the agent's desire to push his agenda. Moreover, we assume that agents attach an intrinsic value to holding decision rights, and that the principal has limited power. Neither of these departures make a difference in a single-agent organization, but are crucial with multiple agents. To fit our applications, we also assume a setting in which utility is non-transferable. In sum, any attempt by any pair of players to reward or punish each other impacts the incentives and payoffs of the third player.

Naturally, one may wonder whether, if given the choice, the principal would prefer multiple agents in our setting. We show that a principal (unless fully aligned with the friend) prefers to "bring in the enemy" if interactions are frequent. One the one hand,

[2]When discussing our results, we provide a case study from the German Weimar Republic politics on how cordon sanitaires are sustained and broken.

bringing in the enemy inevitably involves the enemy in policy-making. On the other hand, it enables the principal to threaten the friend thereby influencing his policy choices.

In the second part of the paper, we extend our analysis to the case in which the principal lacks *commitment power*—she offers a relational contract. The absence of commitment is important for two reasons. First, a committed principal *can credibly promise* the enemy to endorse him in future periods, even if she may later be tempted to renege on that promise and favor her friend instead. That allows her to offer endorsement to the enemy and, in return, ask for concessions. Second, a committed principal *can credibly threaten* to punish a deviating friend. This threat is curtailed if the principal lacks commitment because she genuinely aligns more with the friend than with the enemy.

In some parameter regions, both problems can be circumvented. We can replicate the commitment outcome if the agents attach sufficient value to having the lead or if the principal is sufficiently balanced. Does cooperation unravel completely otherwise? We show that this depends on the level of alignment between the principal and the friend. If they are fully aligned, the optimal relational contract follows a bang-bang logic: either the value for having the lead is sufficiently large to sustain the commitment contract, or any form of cooperation unravels. On the contrary, if the principal is relatively balanced, cooperation gradually unravels when the intrinsic value of agency decreases. Once the commitment contract ceases to exist, the principal can still uphold some cooperation through some embracing strategy.

To summarize, this paper has three main contributions. First, we provide a novel perspective on the incentives in a community of fate when power needs to be allocated between different characters. Second, we provide a dynamic theory of why power-brokers might start to embrace an enemy, once that enemy has gained agency (but not before). Third, we shed some light on the role of the commitment in such a relationship. In particular, whether cooperation unravels depends crucially on the disagreement between principal and friend.

Related Literature. The allocation of power within organization is at the heart of organizational economics and political economy (see Bolton and Dewatripont, 2013, for a survey). Different from large parts of the early literature we follow more recent approaches and consider repeated interactions. The main approach in the literature is to consider either situations in which the principal allocates power with probability one (e.g., Li, Matouschek, and Powell, 2017; Lipnowski and Ramos, 2020; Acharya, Lipnowski, and Ramos, 2024), or the allocation decision is stochastic and exogenous (e.g., Dixit, Grossman, and Gul, 2000). We take a middle ground in which the principal can influence the allocation of power to some extent, but not fully. To the best of our knowledge, Delgado-Vega (2023) is the only paper with a similar approach. However, different from this paper, Delgado-Vega (2023) considers a setting of classical, vertical, moral hazard. Our setting, instead, is about horizontal moral hazard: shirking is not the main incentive problem but the direction of the action is. As a consequence, the differentiation between "friend" and "enemy" becomes relevant, while it is absent in vertical moral hazard problems.

Conceptually, we approach our problem from the perspective of relational contracts (see, e.g., the surveys of Malcomson, 2012; MacLeod, 2014; Watson, 2021), that is, different from the classical repeated games approach, we take certain parameter configuration as given and aim to (behaviorally) characterize the dynamic interaction. Some other papers within that literature consider relationships with multiple competing agents at the same time. Examples include Levin (2002), Rayo (2007), Board (2011), Andrews and Barron (2016), Barron and Powell (2019), Calzolari and Spagnolo (2020), and Nocke and Strausz (2023). Although these papers address diverse problems, they have one feature in common: deviating agents are punished by exclusion. In our setting, such exclusion is ruled out. The principal only has limited power, and our organization is a "community of fate:" both agents and the principal are directly affected by the organization's policy. Therefore, the principal cannot credibly commit to excluding an agent from taking actions in future periods, nor can unilaterally invoke an outside option. Instead, punishments themselves are non-trivial relational contracts, in which the compliant parties take incentive-compatible actions to minimize a deviator's payoff. The principal's temptation to renege on the contract herself is then relevant both on and off the equilibrium path. We share with Board (2011), Barron and Powell (2019), and Calzolari and Spagnolo (2020) that direct and persistent competition between agents is at the heart of our analysis. However, unlike their approaches, we study a setting with non-transferable utility—thus precluding the usual front-loading of payments. Instead, in our model the principal needs to provide a more sophisticated reward scheme to ensure agents' cooperation.

We relate to the growing literature on relational incentives in the political arena (see, e.g., Yared, 2010; Padró i Miquel and Yared, 2012; Anesi and Buisseret, 2022; Callander, Foarta, and Sugaya, 2022; Acharya, Lipnowski, and Ramos, 2024). Cooperation between, for example, political parties or between parties and interest groups, often lacks third-party enforcement, and hence their agreements take the form of relational contracts. Our main contribution is to study the role of these relational incentives in a setting where there is a persistent partisan disagreement and authority swings between the players. Our results apply both to the formation of coalition governments and to lobbying with campaign contributions. These two bodies of literature share a common delegation problem: giving office to a politician entails transferring control rights over certain policies. Hence, in our setting, both the office rent and the control over policy go together, and therefore, the problem of keeping promises gains relevance.[3]

On a technical level the non-stationarity connects our paper to an old literature on punishment in repeated games. In our setting, a principal without commitment,

[3] In previous works in coalition formation like Baron and Diermeier (2001) office rents and policy could be exchanged freely.

differentiates in her punishment threats between the friend and the enemy. In fact, while the enemy and the principal are best punished using traditional trigger punishment, the friend is punished via a stick-and-carrot approach. During the stick period, the principal provides strong support to the enemy without asking much in return. That punishment is sustainable on the principal's side because the principal expects the carrot of eventually returning to the on-path (optimal) behavior. There is a subtle difference, however, to Abreu (1986)'s seminal stick-and-carrot idea. In Abreu (1986) the stick ensures a fixed and given period of low payoffs to *everyone*. In our case, the length of the stick depends on nature's choices resembling ideas from Green and Porter (1984).

2 Model

Time is discrete and indexed by $t = 1, 2 \ldots$ There are three risk-neutral players, one Principal, P, and two Agents, the friend, L, and the enemy, R. Players discount the future exponentially with discount factor $\beta \in (0, 1)$. Information is perfect and, in each period, t, players play the following stage game.[4]

Stage game. At the beginning of a period, P selects the level of endorsement $s_t \in [-m, m]$, where $s_t = -m$ implies full endorsement of L and $s_t = m$, full endorsement of R. Then nature tosses a biased coin which realizes as $k_t \in \{L, R\}$. Endorsement s affects the probabilities of this biased coin. The probability of $k_t = R$ is

$$p(s_t) = \frac{1}{2} + s_t.$$

and k_t=L realizes with the complementary probability, $1 - p(s_t)$. The principal's power is limited, i.e., $m \in (0, 1/2)$. If k realizes, we say agent k is the *leading agent.* The leading agent chooses a policy $y_t \in [0, 1]$. The stage game ends, and all players collect their within-period payoffs.[5]

Payoffs. If the leader's policy is y, player $i \in \{L, R, P\}$ receives a stage payoff from that policy given by

$$u_{i,t}(y) := -|y - \theta_i|,$$

where θ_i denotes player i's persistent bliss point in the policy space,

$$\theta_L \equiv 0 \qquad \theta_P \equiv \theta \in [0, 1/2) \qquad \theta_R \equiv 1.$$

Only the agent that leads, k, in that current period receives a rent for leading, $b > 0$, in addition to $u_{k,t}$. Notationally, it is useful to distinguish between a players' continuation payoff at the beginning of a period, $w_i(\cdot)$, and that at the interim *after selection*, $v_i(\cdot)$.

[4]We assume that the principal can make use of a public randomization device which—although not strictly necessary in our perfect information game—simplifies some of our proofs.

[5]Our model is isomorphic to one in which nature first determines if the principal can delegate (which happens with probability $2m$) or agency is determined by an unbiased coin flip.

Timing and Solution Concept. We are looking for an ex-ante principal-preferred subgame perfect Nash equilibrium (SPNE) of the repeated stage game.

A strategy of a player describes her action choices at all nodes of the game in which the player takes an action. We describe it by the function $s(\cdot)$ for the principal, $y_L(\cdot)$ for agent L and $y_R(\cdot)$ for agent R. We let $\sigma := (s(\cdot), y_l(\cdot), y_R(\cdot))$ describe a strategy profile, and we call strategy profiles that constitute an SPNE a *contract* using C for a generic one. The *optimal contract*, C^*, is the principal's ex-ante preferred contract.

Agents dynamically optimize. In the paper, we consider two scenarios for the principal. First, the *commitment case* in which the principal ex-ante commits to a mapping from histories into continuation strategies. Second, the *no commitment case* in which C must be a proper relational contract between three, dynamically optimizing players.

Discussion. Before moving on, we briefly discuss two key features of our model. We revisit these assumptions again after we present our main results.

Externalities. A defining feature of our model is that players form a community of fate. Policy decisions are payoff relevant to all players in all periods, regardless of who takes them. This is distinct from classical delegation problems, where sidelined agents receive an exogenous outside option. The community of fate aspect implies that all players hold stakes in the organization and therefore views about *how* things should be done.

Principal's Power. Contrary to standard delegation problems, we assume a principal with limited power. The principal influences the selection process but cannot fully control it. This assumption guarantees that irrespective of the other's choices, any agent expects to return to the lead within finite time—no agent can be sidelined forever.

3 Analysis

We analyze the model in steps. We begin with some preliminary observations about implementability. Then, we turn to a special case of the model in which the principal fully aligns with one agent in policy preferences. This example is useful to form a basic intuition about the key mechanics. Yet, as will become apparent in the analysis of the full model thereafter, it only tells part of the story.

3.1 Preliminaries

Implementable contracts. Because a contract needs to be incentive compatible by definition, each player must prefer to choose the policy the contract prescribes over her best deviation. If agent k complies with contract C's recommendation, $y_k(h)$, at a given history h, he will choose policy $y_k(h)$ and if he chooses indeed $y_k(h)$, will expect an on-path continuation value $w_k(C, h')$. If, on the contrary, he deviates, e.g., by choosing his preferred policy θ_k, he is punished with the worst possible continuation value, which we denote by α_k.[6] Formally, a policy y is enforceable for k within contract C at history h if and only if

$$v_k(h) = -|y - \theta_k| + \beta w_k(C, h') \geq \beta \alpha_k. \tag{DEC}$$

Following Levin (2003), we refer to (DEC) as k's *dynamic enforcement constraint*. If it holds with equality, we say that k's (DEC) binds. A necessary condition for a a strategy profile to be a contract is that (DEC) holds at all histories. If the principal has commitment power, the condition is also sufficient.

In contrast to the classical contracting literature, termination is not a possible punishment because (i) agents cannot leave the organization and (ii) the principal has limited commitment power. Instead, the response to a deviation is a contract itself and thus α_k is a non-trivial object. For example, in the commitment case, any agent is maximally punished by P promising unconditional full-endorsement to the non-deviator, which results in complete polarization in the agents' choices, $y_L = 0$ and $y_R = 1$.[7] In the non-commitment case, punishments become more sophisticated.

Agent's disagreement. Agents disagree in two dimensions. First, they disagree along the horizontal policy dimension. There, they distribute a constant utility of -1 between them. Second, they compete over the per-period rent of leading, b. Taken together, agents utility in every period sums to

$$b + u_{L,t}(y) + u_{R,t}(y) \equiv b - 1.$$

An immediate consequence of this constant-sum property *between agents* is that the optimal contract which maximizes the principal's payoff, maximizes utilitarian welfare and is thus Pareto efficient.

The constant-sum property also provides a natural benchmark, which is simply to assume the principal has no authority, that is, $m = 0$.

Observation 1 (**Complete Polarization**)**.** If $m = 0$, the game has an (essentially) unique equilibrium. Whenever agent k leads, he chooses his bliss point $y_k = \theta_k$.

The result follows immediately from the constant-sum property. Any concession by one agent needs to be repaid by a future concessions of the other agent. Because players are impatient, future concession need to be larger than present ones. But then, they require larger concessions thereafter and so on. Thus, absent the principal, agents can only agree to polarize.

[6] Simple penal codes á la Abreu (1988) suffice in our setting because information is perfect.

[7] To ease exposition, we drop the history notation when we mean "for all on-path histories" that satisfy a certain condition usually indicated by sub-/superscript. For example, by y_k we mean "for all on-path histories (within the punishment contract) in which k is the leading agent."

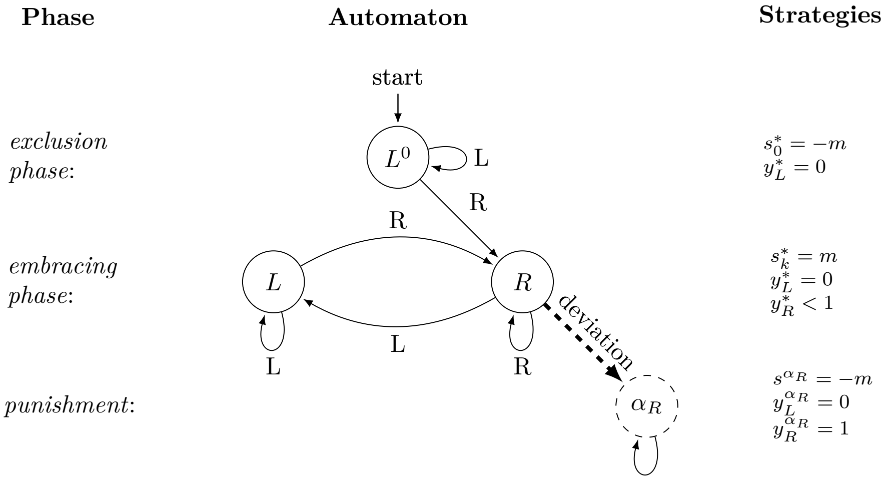

Figure 1: **The Optimal Commitment Contract for $\theta = 0$.** On-path transitions through nature's draws of $k \in \{L, R\}$ are represented by solid edges, transitions induced through a deviating agent by dashed edges. In the exclusion phase, the principal fully endorses L. In the embracing phase, she fully endorses R. If R deviates, he is punished by unconditional endorsement of L until eternity.

3.2 A Simple Example—Commitment

We now introduce the principal and solve the special case in which she fully aligns with agent L, $\theta = \theta_L = 0$.

A Candidate Contract. A natural candidate would be for the principal to always fully endorse the friend L, i.e., $s = -m$. This strategy maximizes the probability of the friend leading and, once leading, L chooses the principal's bliss point, $y_L = 0$. However, whenever the enemy, R, leads, the enemy has no incentives to choose anything but his bliss, $y_R = 1$. Thus, the principal enjoys, in the majority of cases, her preferred policy $y_L = 0$, but suffers from her worst policy $y_R = 1$ otherwise. Such behavior trivially describes a contract since it repeats the stage Nash equilibrium (referred to simply as Stage Nash henceforth). However, we may ask: Is this contract ever optimal?

The Optimal Contract. It turns out that this intuitive candidate is not optimal for any parameter combination (β, m, b). Instead, (almost) the opposite is true: the principal embraces the enemy, once an initial attempt to exclude him failed.

Proposition 1. *Suppose $\theta = 0$ and a principal with full commitment. For any parameter setting, (m, β, b), the optimal contract consists of two phases on the equilibrium path:*

(i) an exclusion phase *at the beginning, in which the principal fully endorses L, and*

(ii) an embracing phase *later, in which the principal fully endorses R.*

The contract switches to the embracing phase once R leads for the first time. *In all phases, L chooses $y_L^* = 0$, and R moderates to*

$$y_R^* = \max\left\{1 - \frac{4\beta m(b+1)}{2 - \beta(1-2m)}, 0\right\}.$$

Proposition 1 shows that the principal switches on-path strategies mid-game. Figure 1 provides the corresponding automaton representation of Proposition 1.[8] The game begins with an exclusion phase (node L^0) in which we remain (edge L) until nature selects R to lead for the first time (edge R). Now, we transition irreversibly to the embracing phase (nodes L and R). The principal fully endorses R and nature determines the state each period (edges L and R). If agent R deviates, we enter a punishment phase (dashed edge) in which the principal unconditionally endorses L, leading to polarization ($y_L = 0, y_R = 1$) for eternity. Since L has no incentive to deviate in the first place, his deviations can remain unpunished.

The principal's endorsement has two effects. It influences who gets selected to lead and serves as a reward for the agent's behavior. The principal wants agent L to lead, but endorsing L does not change L's behavior when leading: L already chooses the principal's ideal policy. However, if R leads, the principal wants to discourage R from choosing the principal's worst policy. She promises endorsment in exchange for moderating policies.[9]

Endorsement implies that an agent takes the lead more often. He profits from leading in two ways: he values holding power per se, and also using power to choose the policy. The principal, in turn, only cares about the use of power. The principal therefore leverages her indifference over *who* leads to influence *how* they lead. By promising the enemy the lead more often in the future, the principal can ask him to choose a policy closer to her bliss point. The enemy complies because he is compensated for his concession by enjoying the value of holding power more often.

These considerations, become only relevant once the enemy leads for the first time. The reason is simply that past endorsements are sunk when it comes to incentives for the leading agent. Thus, before the enemy leads for the first time, it only matters *who* leads next. The principal prefers her friend to lead and therefore, she tries to exclude the enemy. Once the enemy leads, however, the principal's main concern is to limit her losses from the enemy's choices. Promising indefinite full endorsement maximizes her net present value of future policy choices by both agents.

3.3 A Simple Example—No Commitment

The previous results rely on the principal promising future endorsement to the enemy if he moderates today. Thus, they appear to be tightly connected to the principal's commitment power. After all, the principal always prefers L to lead. Here, we discuss the role of the principal's commitment.

[8] An automaton representation common in the representation of repeated games groups "relevant" histories into (automaton) states. It avoids keeping track of complicated game trees in cases in which we are interested in more than Markov Perfect Equilibria (see Mailath and Samuelson, 2006, for a detailed discussion).

[9] Note that due to our linearity (and thus constant marginal utility) assumption we avoid the usual moderation through exploitation of decreasing returns as in, e.g., Acemoglu, Golosov, and Tsyvinski (2011). Assuming a concave loss function as in that literature would only amplify our effect.

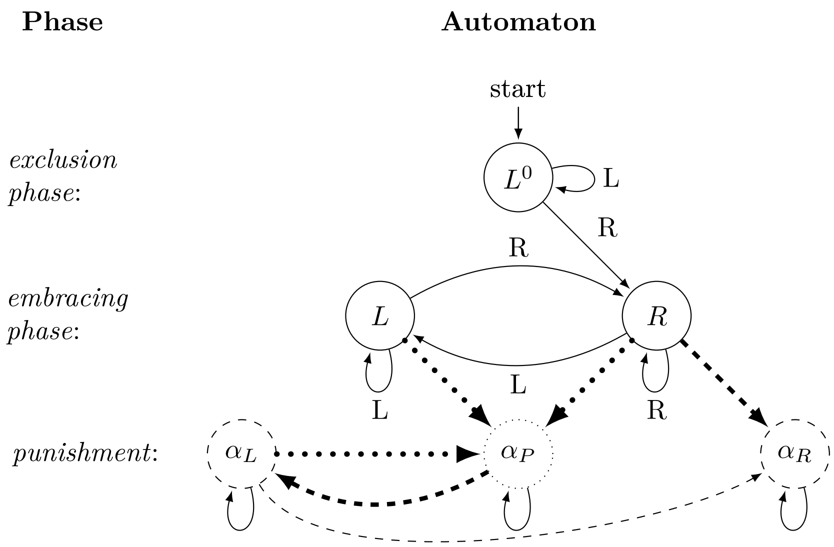

Figure 2: **The Optimal No-Commitment Contract if $\theta = 0$.** On-path transitions through nature's draws of $k \in \{L, R\}$ are represented by solid edges, transitions induced through a deviating agent by dashed edges, transitions induced through a deviating principal by dotted edges. Thick dashed edges represent agent's binding (DEC) in the optimal contract. Thick dotted edges represent the principal's binding (DEC)'s if $b = \hat{b}$. L's deviations on-path and during R's punishment can be left unpunished and so can R's deviations during P's punishment.

Without commitment, also the principal optimizes dynamically. She, too, must find it (interim) optimal to honor her promises. We capture that requirement through the principal's dynamic enforcement constraint: any contract C at any history h must satisfy

$$w_P(C, h) \geq \alpha_P, \tag{DEC'}$$

where α_P is the principal's punishment, given by the ex-ante payoff of her worst contract.

Constructing this punishment payoff, α_P, is a non-trivial exercise. In the principal's worst contract, the friend and the enemy collude. They both choose a policy to the right of the principal's bliss point. Punishing the principal is particularly demanding on L. How far to the right is he is willing to move depends on what his unishment. Thus, also L's deviations become relevant, albeit only off the equilibrium path. Figure 2 provides the updated automaton representation. Instead of the single binding (DEC) in the commitment contract, we now need to determine a more sophisticated punishment. When punishing P, L's (DEC) binds, whereas P's (DEC') binds when punishing L and (potentially) on-path. We provide a detailed construction in appendices D and G

Proposition 2. *Assume a dynamically optimizing principal, and suppose $\theta = 0$. The optimal commitment contract can be implemented if and only if*

$$b \geq \bar{b}_0 := \frac{(1-\beta)^2}{\beta(2-\beta(1+1/2-m))(1/2+m)}.$$

Otherwise, Stage Nash is the only implementable on-path behavior.

Proposition 2 states that the principal can keep her promises if b is sufficiently high.

This result is intuitive. The larger b, the more R is willing to moderate in exchange for endorsement. Hence, the larger b, the greater the principal's incentives to stick to her promise to endorse R.

But, perhaps surprisingly, if the optimal commitment contract cannot be upheld, any chance of cooperation with the enemy unravels. That is, if b is too low for the principal to keep her commitment-contract promises, the only remaining contract is our initial candidate: full polarization and endorsement of the friend.

To see why cooperation with the enemy follows this bang-bang pattern, consider a class of (commitment) contracts in which the principal offers some intermediate endorsement $s \in (-m, m]$ and R moderates as much as his (DEC) permits in return. We may conjecture that even if $s = m$ is not sustainable without commitment, some other contract in this class could fix the principal's commitment problem. The principal always prefers L over R to lead, and decreasing endorsement for R should thus increase the principal's payoffs. However, this marginal effect is only part of the story. There is an inframarginal effect to be considered too: if P decreases endorsement for R, R will respond by reducing her moderation. That, in turn, will decrease the principal's utility *conditional* on R leading.

It turns out, that the joint effect can never fix the principal's commitment problem, if the contract from Proposition 1 cannot do it. To build intuition, it is best to start at the other extreme. That is, assume players play Stage Nash. This contract trivially satisfies P's dynamic enforcement constraint. Let us now consider a marginal increase of s in return for R's moderation. Increasing s slightly has a negative, marginal, effect: the friend leads less often. But it has a positive inframarginal effect: the enemy moderates.

At $s = -m$, the marginal effect is large because R, when leading, chooses polarizes a very distant policy. Hence, a leading R is much worse for the principal than a leading L. As R moderates more, that marginal effect diminishes, so selecting R becomes less costly for P. The inframarginal effect remains—due to linearity of the preferences—(roughly) constant.[10] Thus, eventually, the inframarginal effect becomes relatively more important, leading to a conditional value function for P that is convex in s. Eventually, the inframarginal effect becomes dominant and this value function increases. If b is large enough, P's on-path value becomes large enough for $s < m$ to credibly promise that endorsement s to R. But then, this conditional value function keeps increasing in s implying that the principal can promise $s = m$ and thus the optimal commitment contract is implementable.

3.4 General Model—Commitment

We now turn to the general case, $\theta \in [0, 1/2]$, assuming the principal can commit. In the main text, we focus on the economic discussion, the intuition, and the implications. The

[10]There is some non-linearity due to discounting involved, but because it affects marginal and inframarginal effect alike, it is irrelevant for this intuition.

Results we present here are corollaries of the optimal contract which we fully characterize in appendix E for arbitrary (β, m, b, θ).

First Best Contracts. The principal obtains her first best if all agents choose $y_k = \theta$. The fact that the principal attains her first best if the vector (β, m, b) becomes large, is not surprising. Fist, when β increases, the equilibrium set expands through the folk-theorem logic. Then, we can attain first best provided (m, b) are large enough for the principal's future behavior to matter. Second, when $m \to 1/2$ the principal obtains complete control over selection and can promise to exclude a deviator. For agents who care about the future that alone incentivizes $y_k = \theta$. Finally, when $b \to \infty$ agents compete mainly for taking the lead. Any policy that improves their leading chances is enough for incentivizing then to choose $y_k = \theta$. How large the vector (β, m, b) needs to be depends on the principal's position, θ. Formally, we define two values

$$\overline{\theta} := 2\frac{\beta}{1-\beta}\left(1-\beta\left(\frac{1}{2}+m\right)\right)(b+1)m,$$
$$\breve{\theta} := 1-\beta\left(\frac{1}{2}+m(1+2b)\right).$$

Result 1. *The principal achieves her first best if and only if* $\theta > \max\{1-\overline{\theta}, \breve{\theta}\}$.

Observe that when $b \to \infty$, the cutoff converges to $-\infty$ and thus first best can always be attained. Moreover, as $m \to 1/2$ the cutoff converges to $1-\beta(1+b)$ which implies that first best can be attained for any θ provided that $\beta > 1/(1+b)$.

Second-Best Contracts. Our problem becomes interesting if interaction is somewhat infrequent ($\beta << 1$), the principal's power is limited ($m << 1/2$) and policy matters to the agents (b not too large). From now on, we will assume that we are in such a second-best world, i.e., $\theta < \max\{1-\overline{\theta}, \breve{\theta}\}$. As in the example, we distinguish between an exclusion phase and an embracing phase.

Exclusion phase. First, focus on the periods before R leads for the first time. Recall from the example that then, the principal wants to keep the enemy from leading. The reasoning was this: until R leads for the first time, there are no incentives to provide him yet. This logic is unchanged as we introduce some disagreement between the principal and the friend ($\theta > 0$). In fact, by threatening to remove her endorsement, the principal provides incentives for the friend to moderate toward the principal's bliss point. That threat is effective as long as the disagreement between the principal and the friend is not too large. The following result formalizes this idea using the cutoff $\overline{\theta}$ from above.

Result 2 (Exclusion phase). *There exists an exclusion phase until the enemy leads for the first time. On the equilibrium path, the exclusion phase is characterized by full endorsement of the friend,* $s_0^* = -m$*, who, in turn, chooses*

(i) the principal's bliss point, $y_L^* = \theta$*, if and only if* $\theta \leq \overline{\theta}$*, and*

(ii) the largest incentive compatible policy, $y_L^* = \overline{\theta}$*, if and only if* $\theta \geq \overline{\theta}$*.*

Regardless of θ, the principal can push L at most to $y_L = \overline{\theta}$. However, the principal has no interest to "overshoot" and thus only pushes L to $y_L = \theta$ if $\theta < \overline{\theta}$.

Embracing phase. We now move to the second phase of our contract. Once R receives the power to choose the policy, R's incentives become relevant. In the example, the principal would not need to incentivize a leading L. She fully shifted endorsement to the enemy in exchange for his moderation. A less extreme principal also aims to moderate his friends. The friend's incentives to moderate come (i) from the principal's direct endorsement and (ii) her success at moderating R. The further R moderates, the better for L, providing him larger incentives to respect the contract.

It turns out that embracing the enemy fully as in Proposition 1 remains optimal for a range of $\theta > 0$. If θ is sufficiently close to 0, L can be incentivized solely by the principal's promise to push for R's moderation and by threatening to withdraw that push should L refuse to comply. Specifically, when the principal's bliss point is to the left of threshold

$$\underline{\theta} := 2\frac{\beta}{1-\beta}(b+1)m\beta\left(\frac{1}{2}+m\right),$$

the friend's (DEC) has slack on the equilibrium path which implies that endorsing the enemy fully during the embracing phase is optimal.

Result 3 (**Full Embracement**)**.** *Consider the embracing phase of the optimal contract. If and only if $\theta \leq \underline{\theta}$, the principal fully endorses the enemy irrespective of who has led last.*

When leading, the friend chooses the principal's bliss point, $y_L^* = \theta$. The enemy, in turn, moderates from her bliss point to $y_R^* \in (\theta, 1)$. The enemy's (DEC) *binds and moderation increases in the parameter θ.*

Result 3 already indicates how the principal can incentivize the friend when the enemy's moderation is insufficient zo make him choose her bliss point: the principal can offer some endorsement also to L. Hence, for $\theta > \underline{\theta}$, the principal shifts some endorsement to the friend. This leads to the following phenomenon: The more the principal and the friend disagree, the more endorsement the friend receives.

The principal's endorsement strategy tilts. She conditions her endorsement on who has led last. In particular, when R has led last, she continues to fully endorse R, $s_R^* = m$. But when L has led last, the principal takes a more L-leaning endorsement strategy. The more balanced the principal, the more she endorses L.

To see why such behavior is optimal, recall the constant sum property between agents. Moreover, note that R's (DEC) binds since he is asked to moderate as much as possible. Combining these two facts implies that L's payoff conditional on *not leading* is already maximized even under full endorsement of R. Thus, the only way to relax L's constraint is to increase the endorsement she receives after having led last. Finally, any situation in which L's dynamic enforcement constraint binds, has the same value to R but not to the principal. The principal prefers, if possible, the case in which $y_L^* = \theta$. Thus, she adjust

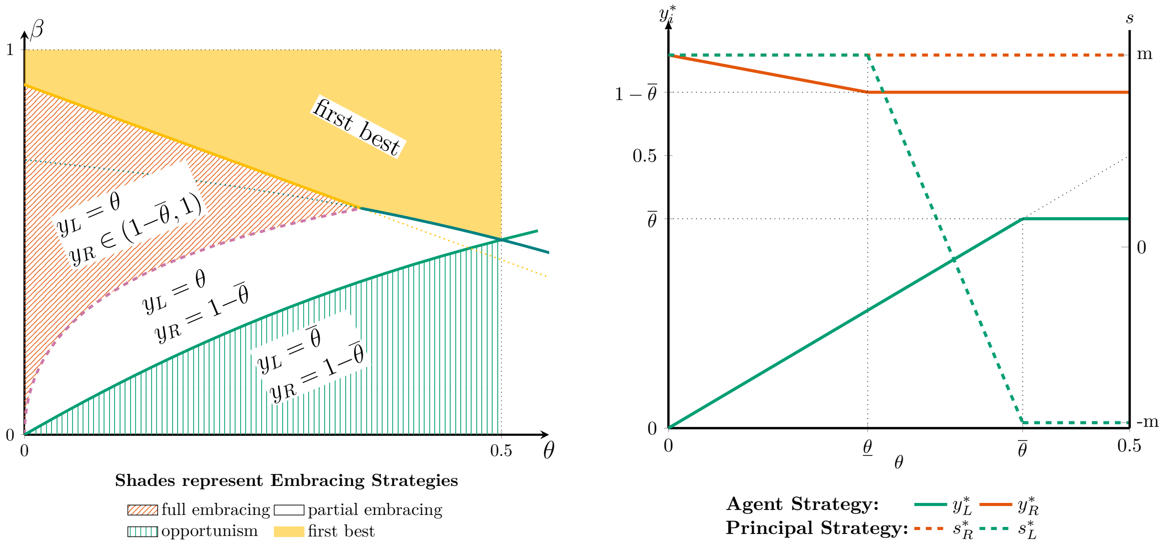

Figure 3: *Embracing Phase under Commitment.* The **left panel** shows the on-path strategies in the embracing phase in the (θ, β) space. Outside the first best, full embracing is optimal for low θ, partial embracement for intermediate θ, and opportunism for large θ ($b = 2.5, m = 0.2$). The **right panel** depicts strategies as function of θ for a given (b, m, β). The left scale is for agent's decisions, the right scale for principal's.

strategies to ensure $y_L^* = \theta$ if possible. The next result summarizes these findings, and Figure 3 provides an illustration of Results 1, 3 and 4.

Result 4 (Opportunistic Embracement). *Suppose $\theta > \underline{\theta}$. Consider the embracing phase of the optimal contract.*

The principal plays a distinct strategy s_k^* when k has led last. When the enemy has led last, the principal fully endorses him, $s_R^* = m$. When the friend has led last, endorsement is as follows.

(i) The principal endorses at an intermediate level, $s_L^* \in (-m, m)$, if and only if $\underline{\theta} < \theta < \overline{\theta}$. The friend receives larger endorsement (lower s_L^*), the stronger her disagreement with the principal (larger θ).

(ii) The principal fully endorses the friend, $s_L^* = -m$, if and only if $\theta \geq \overline{\theta}$.

The agents' policy choices are such that both (DEC) *bind. The enemy chooses $y_R^* = 1 - \overline{\theta}$ and the friend chooses:*

(i) the principal's bliss point, $y_L^* = \theta$ if and only if $\theta \leq \overline{\theta}$.

(ii) the largest incentive compatible policy, $y_L = \overline{\theta}$, if and only if $\theta \geq \overline{\theta}$.

Result 4 shows that as the principal becomes more balanced, the friend receives more endorsement. For $\theta \geq \overline{\theta}$ the optimal contract implies a fully opportunistic principal. That principal simply endorses whoever has led last.

Result 4 shows that as the principal becomes more balanced, the friend receives more endorsement. For $\theta \geq \overline{\theta}$ the optimal contract implies a fully opportunistic principal. That principal simply endorses whoever has led last. Note that along that process, policies are only affected by the endorsement choice immediately after the agent's move. Hence, R's policy is fixed at $y_R = 1 - \overline{\theta}$, while L's policy moves as the principal's endorsement for L increases, until $s_L = -m$ and $y_L = \overline{\theta}$.

The key to understanding our results is understanding how the principal's disagreement with the agents is distributed. A balanced principal holds large disagreements with both agents and needs to incentivize them both to moderate. If her bliss point is to the right of $\overline{\theta}$, she needs to fully endorse whoever led last to uphold that period's incentives for moderation, but none of the agents ever chooses the principal's bliss point. Thus the optimal contract is identical for all $\theta \in [\overline{\theta}, 1/2]$.

3.5 Implications

Here we formulate the economic lessons from the optimal contract.

Polarization. Our results predict that the level of polarization in policies is a function of the principal's disagreement with the agents.

Result 5 (**Polarization**). *The distance between the agent's equilibrium policies, i.e., $y_R^* - y_L^*$ decreases as the principal's preferences become more balanced.*

This result contrasts with the full polarization that occurs in the absence of the principal (Observation 1). It adds a new channel to the political economy literature that explores moderation as a form of repeated favor trading (e.g., Dixit, Grossman, and Gul, 2000; Acemoglu, Golosov, and Tsyvinski, 2011; Invernizzi and Ting, 2023; McClellan and Liao, 2023). These papers consider settings without a principal, but assume curvature in the agent's utilities. By contrast, our setting is constant-sum between agents by assumption thereby ruling out that agents everage on differences in marginal utilities. Yet the two effects (theirs and ours) amplify each other such that with curvature in the utility, balanced principals will lead to even more moderation.

Principal's welfare. A priori, it is not obvious whether a principal fully aligned with an agent is better or worse off than a completely balanced principal. Our next result shows, that a principal's payoff is highest when he has some bias toward an agent, but that bias is not extreme.

Result 6 (**Principal's Payoff**). *Consider (β, m, b) such that the first best is unattainable for any θ. Then the principal's ex-ante payoff has a unique maximum in θ at $\theta = \overline{\theta} < 1/2$.*

The reason behind this result can be seen when starting with a fully extreme principal, $\theta = 0$. As the principal becomes more biased, first, she is able to ensure that agent L still selects her preferred policy θ. Second, agent R gets mechanically closer and due to L's movement is willing to moderate more. This effect stops once $\theta = \overline{\theta}$ after which all strategies remain fix as θ increases. Now, the principal payoff decreases when L leads and increases when R leads. Due the exclusion phase, however, alignment with L is more important for the principal's payoff and hence her payoff decreases in θ for $\theta > \overline{\theta}$.

An implication of Result 6 is on the selection of the principal. If we suppose that becoming a principal implies incurring some initial cost, moderately biased principals have the largest incentives to take on such cost.

Competition. Our third implication compares our model with one of a single agent L. A priori it is not clear whether the principal prefers to lose her disciplining power by fully delegating authority to L ex-ante who then has no incentives to moderate or prefers to deal with a (potentially distant) enemy. The following proposition shows that a cutoff exists such that a principal, which is more balanced than this cutoff, prefers competition over total exclusion of the enemy.

Result 7 (**Competition**). *Suppose parameters are such that the first best is not attainable. There exists a threshold $\tilde{\theta} \in [0, 1/2)$ such that for all $\theta > \tilde{\theta}$ the principal prefers to operate with two agents, L and R. Otherwise, she prefers to eliminate R ex ante.*

Moreover, if players are sufficiently patient, any principal with $\theta > 0$ strictly prefers to operate with two agents L and R.

Intuitively, as the principal becomes more balanced, her disagreement with the friend increases. The presence of the enemy allows her to discipline both by threatening each to support their rival. The cost, naturally, is that, sometimes, the agent choosing the policy has vastly different preferences. However, as the principal moves to the center, this cost declines. If players are patient—which can be interpreted as frequent opportunities to act—the principal's power to moderate strong. Therefore, she prefers a competitive setting with the enemy even when she is almost aligned with her friend's preferences.

3.6 Discussion: Commitment

Before removing the commitment assumption, it is worthwhile to pause and reflect on the touch-points between our model and the real world.

A Community of Fate. We chose a highly stylized model based on a central, conceptual assumption: our organization resembles a community of fate.

A community of fate implies that all members hold stakes in the organization. Moreover, although they may disagree on the direction the organization should take, it is too costly to break up the organization or to permanently bar some opinions from decision-making. That is, although disagreement may be a burden, there is—at least in the medium run—no solution to enforce conformity of the members.

The prohibitively high cost may have different reasons. In a political context removing a discordant faction of the society may only be done through breaking up the polity entirely or, worse, through violence. In a company firing a discordant specialist requires that a concordant replacement with equal skills can be found on the market. Often such types simply do not exist.

A second, connected issue is that because members cannot be excluded, they often hold bargaining power. This is evident in the political arena, but also applies to innovative companies where traditional, rigid hierarchies threaten the retainment of high-skilled workers. In all these settings, managing the disagreements about the organization's policies becomes a crucial task.

Modeling Choices. Our model aims to parsimoniously capture the idea of a community of fate. We make three key modeling choices: the organization consists of the principal and two competing agents, disagreements are horizontal, and the principal is weak. Although the principal has influence in the power allocation, she cannot sideline either agent permanently. Instead, she acts as a power broker between the agents.[11]

Although highly stylized, our setting captures, for example, the crucial features of parliamentary coalition formation, or executive committees of companies. The delegation of decision-making power is at the center of these applications: a firm's CEO assigns projects to different divisions or a centrist party chooses its coalition partners. Yet, exogenous forces curb their power: sometimes expertise implies directly which division should take the lead, sometimes exogenous events grant extremist parties a majority, which renders the centrist player temporarily powerless.

Horizontal disagreement is traditional to political economy settings, but disagreements about *how* things should be done are also common among the upper echelons of corporations, overshadowing the classic moral hazard problem of exerting effort.[12] Even when institutions do not exogenously grant power to a certain principal, our setting arises naturally. For example, a centrist party with the ability to "play both sides" endogenously becomes the principal of our setting: she can occasionally break ties by siding with one of the more polarized blocks. Throughout history, European parliamentary systems—for instance in Italy, France or Germany—had liberal or Christian-democratic parties that exploited their centrality by taking a role in line with that of our principal.

Embracing the Enemy vs the Ally Principle. We now compare our Embracing the Enemy result with a well known concept in the politics of delegation: The ally principle (see e.g., Gehlbach, 2021). The ally principle posits that a principal should delegate to her closest ally. In our model, adhering to the ally principle is a sound strategy as long as the enemy lacks agency. However, once the enemy holds agency, embracing him is optimal. The reason behind that shift lies in the community of fate assumption. No agent can be permanently excluded from power and hence the principal needs to envision a relationship that accommodates both friends and foes.

In the real world, we frequently observe deviations from the Ally Principle. Prime

[11]Note that each deviation from the classical principal-agent framework in isolation does not imply changes to its predictions. Yet, as we have seen, combining all three implies a significant departure.

[12]Often leading figures have enough "skin in the game" to work hard, but may disagree on the road to success for various reasons: signaling motives for their own career (Holmstörm, 1999), competing narratives over the right course of action (Eliaz and Spiegler, 2020), or identity concerns (Akerlof and Kranton, 2005)

Ministers incorporate their most vocal critics into their cabinets, and CEOs distribute company projects across various divisions. . . Several (static) ideas have been put forward to explain such deviations. They include "face-washing," "power sharing," or the principal's need for commitment (see, e.g., Rogoff (1985) on why lavish politicians might delegate monetary policy to a conservative central banker). More formally, Bendor, Glazer, and Hammond (2001) and Callander et al. (2008) demonstrate how the Ally Principle can be invalidated if information acquisition is costly for the agents. We present a different, inherently dynamic explanation in which information plays no role: the principal delegates to an enemy because her power is limited.

Cordon sanitaires—An application. We now discuss a particular application of our setting: the entry of radical newcomers in parliament. Our results provide a theory of the dynamics of so-called *cordon sanitaires* that mainstream parties sometimes attempt to establish around radical newcomers: ostracizing them in the political process. According to our theory, a *cordon sanitaire* is initially optimal but once it is broken, it remains broken. The first time the newcomer allies with the establishment, the pattern changes: ostracized parties become regular coalition partners and moderate their positions in return.

Our characterization offers a clear-cut prediction on *when cordon sanitaires* are broken, namely, when agency of the enemy is unavoidable. A historical example is Weimar Germany and the politics around the participation of the German National People's Party (DNVP) in government. The DNVP rejected the republican Constitution of 1919 and the Treaty of Versailles, and was accordingly excluded by democratic parties from coalitions. However, in 1924, passing the constitutional amendments to ratify the Dawes Plan made the DNVP's votes necessary. This necessity led to a bridge-building attitude by centrist parties—the Zentrum and German People's Party (DVP)—toward the DNVP. After the Dawes Plan was approved, the DVP's leader Gustav Stresseman

> was conciliatory towards the DNVP, arguing that having been bought to accept the realities it should be allowed to share the government [. . .] 'National realpolitik' was an attempt to build a bridge to the DNVP. [. . .] He argued that bringing the DNVP into government would be an act of statesmanship. No party, he said, had more reason than they to dislike the DNVP: 'But one cannot make domestic policy with sentimentality.' (Wright, 2002)

The passing of the Dawes plan started a period of DNVP moderation, which coincided with Weimar republic's so-called "golden age". The DNVP joined a republican cabinet and a liberal-conservative coalition. To maintain its government position, the DNVP supported laws which led to an implicit recognition of the Weimar constitution.

Preferences Reports: Signalling Extremism. We close our discussion with addressing a question somewhat outside our model. Would an agent benefit from signalling extremism. To this end, suppose an agent could announce his bliss point before the game begins, and assume the others naively believe his claim. Would the agent report his true bliss point? He would not. The logic can be most easily seen for the friend. As the principal believes that L's bliss point is further from her own bliss point, she gradually begins to tilt endorsement in L's favor but never requests a policy further to the right than the principal's bliss point. Furthermore, a second-order effect appears: since L reports more extreme preferences, R is more afraid of polarization and willing to concede more, thereby benefiting L. Yet, in equilibrium, of course, players would not believe such reports and "signalling extremism" becomes complicated. We consider it beyond scope here.

3.7 General Model—No Commitment

We now remove the principal's commitment. Recall the principal's (DEC') is given by

$$w_P(C, h) \geq \alpha_P, \tag{DEC'}$$

which implies that at any history h the principal prefers continuing with contract C over switching to her worst continuation contract.

The principal's commitment problem. On the equilibrium path, (DEC') becomes particularly relevant when the contract asks the principal to endorse the enemy, as she is naturally tempted to endorse the friend instead. Importantly, this temptation is also relevant *when trying to punish the friend.* The reason is straightforward. Punishing the friend involves endorsing the enemy, an action the principal wishes to reneg on.

When the value of leading, b, is high, however, the principal's commitment problem is moot. As shown in Result 1, the principal's first best is easily achieved for $b \to \infty$. Thus, she is indifferent whom to endorse and plays both sides, with or without commitment.

Proposition 3. *For any (β, m, θ), there exists an $\bar{b}$ such that for $b \geq \bar{b}$, the optimal commitment contract can be implemented.*

Balance is commitment. A fully balanced principle, $\theta = 1/2$, has no issues to implement the commitment solution. To her, any agent is friend and foe alike. Switching allegiance comes at no cost. Our next result shows that there is a non-empty neighborhood of $\theta = 1/2$, in which the commitment solution can be implemented, too. That is, although the notions of friend and an enemy start to have meaning, the commitment assumption still plays no role.

Proposition 4. *Take any (β, m, b) such that $\overline{\theta} < 1/2$. There exists a threshold type $\hat{\theta} \in [\overline{\theta}, \frac{1}{2})$ such that if the principal's type is $\theta \geq \hat{\theta}$, her maximum payoff without commitment is identical to that with commitment. Moreover, that threshold decreases in b.*

To understand the intuition behind Proposition 4, recall first that we are in an environment where—even with principal commitment—*both* agents' (DEC) bind. Second, also recall that although the principal's commitment is irrelevant for punishing the enemy, it plays a role in punishing the friend. Thus, the key to upholding the commitment solution, is to uphold the friend's punishment value.

The following two-phase contract provides the same payoff to the friend as the grim trigger punishment yet is easier to sustain for the principal:

(i) *"punishment" phase.* Initially, the contract is identical to the punishment contract under commitment: the principal endorses R, and R polarizes choosing $y_R = 1$. This phase continues until L is (re-)selected to lead for the first time after deviation.

(ii) Then we enter the *"back-to-business" phase*: play returns to the behavior of the optimal contract.

This type of punishment is reminiscent of Abreu (1988)'s "sticks and carrots." The punishment phase is painful, not only—as intended—for the friend, but—as a side product—also for the principal. To make up for it, the principal is promised a carrot: going back to the optimal contract. This back-to-business carrot is effective because it is both, the best continuation value the principal can be promised, and, undesirable for the friend. For the friend, returning to the principal's best contract, implies that he receives the *lowest continuation payoff conditional on leading* since his (DEC) binds.[13]

To close the argument, observe that the principal is willing to go through the punishment phase only if *she* fears *her* worst punishment enough. In fact, during her attempt to punish the friend, the principal hopes for the punishment phase to end by having the friend return to lead. Yet, she may be willing to endorse the enemy if non-compliance is responded by playing, e.g., Stage Nash from then on. Recall, that Stage Nash implies full polarization, a situation a balanced principal fears most. Therefore, Stage Nash implies enough incentives for the principal to follow the friend's punishment contract.

Gradual unraveling. We now turn to parameter specifications such that $\overline{\theta} < \theta < \hat{\theta}$.

Before we proceed, it is instructive to recall how in our special case $\theta = 0$, when b falls below threshold $\bar{b}_0$, (i) the commitment outcome is no longer available, (ii) the principal's (DEC') binds on the equilibrium path, and (iii) the problem fully unravels allowing only the repeated stage Nash as an equilibrium.

As we will show now, for a large θ these three effects become relevant at different stages and hence, the optimal contract does not unravel completely. First, we start with a θ immediately to the left of $\hat{\theta}$. We show that for these cases, neither effects (ii) or (iii) apply. Hence, although the principal cannot attain the commitment outcome, she still moderates the agents.

The principal's main problem in this region is that she cannot credibly threaten to punish the friend as harsh as she would want to ex-ante. Instead she needs to promise a punishment she will be willing to sustain. For the same reasons as above, the optimal punishment will have the back-to-business property. However, even with that carrot, the

[13]There are some notable differences to Abreu (1986). First, our game is not symmetric, which he requires. Second, while in his case the carrot is needed to keep all non-deviators on-board in the stick phase in our case relaxes P's constraint, but tightens R's. Third, due to the constant-sum logic between the agents, R profits from punishment, while in Abreu (1986) all players suffer.

principal cannot promise a punishment phase with full and unconditional endorsment of the enemy. Instead, the principal needs to ask for some moderation by the enemy in exchange for her full endorsement to be willing to go through with the endorsement. A side effect of this exchange is that the friend's punishment gets reduced.

In the following proposition we formulate this result in terms of the value of leading, b, which has a one-to-one mapping to $\hat{\theta}$ in that region as established in Proposition 4.

Proposition 5. *Take some set of parameters (β, m, θ) such that $\theta \in (\overline{\theta}, \hat{\theta})$ for some b.*

Then, there exists some $\hat{b} < \overline{b}$ such that for any $b \geq \hat{b}$, the optimal contract is qualitatively similar to the commitment one in the following sense:

(i) the principal plays the opportunistic strategy $s_L = -m$ and $s_R = m$, and,

(ii) the agents choose on-path policies y_i^* such that their (DEC) *binds.*

As b decreases, L's punishment is less severe resulting in more polarization in the sense that y_L^* decreases and y_R^* increases. The principal's ex-ante payoff therefore decreases when b decreases in this region.

In this proposition we see that, as long as b is above $\hat{b}$, there is a crucial element of continuity with respect to the commitment optimal contract: the principal's (on-path) strategy is opportunistic. The principal's endorsement is unchanged because her (DEC') has slack on the equilibrium path.

Of course, there is an on-path effect of the reduced punishment because the friend's (DEC) changes even if it is an off-path object. Since the friend's punishment becomes milder, he is not willing to choose the same policy as under commitment. Instead, he choose a more polarized option. Moreover, since the friend moderates less, the enemy's willingness to moderate declines. Both effects reduce the principal's on-path value.

If the value of taking the lead is below $\hat{b}$, the principal's (DEC') not only binds (off-path) when punishing the friend. Instead, (DEC') becomes an on-path constraint. This has a two effects. First, and similarly to before, the friend's punishment becomes milder. Second, now the principal's endorsement strategy needs to change on-path. She has to promise partial endorsement to the friend even when the enemy is in the lead, as otherwise she would not appear credible. Through that partial endorsement, however, she is able to save at least some of the positive embracing-the-enemy effect. If the value of punishment is even lower and belwo $\check{b}$, the Stage Nash is the only remaining option. Due to the marginal adjustments throught, the optimal contract smoothly converges to the the Stage Nash as b decreases towards $\check{b}$.

Proposition 6. *Take a set of parameters (β, m, b, θ) such that $\theta \in (\overline{\theta}, \hat{\theta})$.*

Then, there exists a cutoff $\check{b}$ such that if $b \geq \check{b}$, the optimal contract takes the following form:

(i) whenever L led last, the principal fully endorses L, $s_L = -m$,

(ii) whenever R led last, the principal partially endorses R , $s_R \in (-m, m)$, and,

(iii) the policies are such that both agents' (DEC) *bind on the equilibrium path.*

If, starting from $b > \check{b}$, b decreases, then s_R decreases, on-path polarization increases in the sense that that y_L decreases and y_R increases, and the principal's optimal payoff decreases.

If $b < \check{b}$, the repeated stage Nash equilibrium is the only implementable on-path behavior.

Together, Proposition 3, 5 and 6 give a clear intuition: the principal's commitment assumption, although relevant, has less extreme effects if the principal is more balanced. In the case of an extreme principal, small changes in the parameter values have drastic consequences: Instead of the full commitment solution, nothing but the Stage Nash becomes available to the principal. With a more balanced principal, these changes imply less severe consequences. Even absent commitment, small changes only produce gradual changes in the implementable contracts.

3.8 Discussion: No Commitment

So far we have considered both an extreme principal (Proposition 2) and a sufficiently balanced principal (Proposition 4 to 6). One may wonder about the optimal no-commitment contract between these polar cases. While a (behavioral) characterization is possible, we need to consider a several issues simultaneously, making it somewhat cumbersome. To give an idea of the construction we discuss informally some of the complexities that arise and relate them to our results.

First, the cases we have discussed imply a simple punishment contract for the principal. A balanced principal is best punished by polarization (repeated Stage Nash). An extreme principal, on the other hand, is best punished by a the friend and the enemy colluding to select an action to the right of the principal's bliss point. In the general case, the optimal punishment for the principal depends on the parametric configuration.[14]

Second, determining which punishment is best for a deviating friend becomes complicated for intermediate principals. On the on hand, we need a carrot for the principal. On the other hand, we need a stick for the friend. Because the optimal contract for the principal involves an exclusion phase in which the friend's (DEC) has slack, there is a tradeoff. Should we reduce the friend's punishment in the back-to-business phase to strengthen the principal's incentives to endorse the enemy in the punishment phase? It turns out that, if possible, the back-to-business phase returns to the part of the optimal contract in which the friend's (DEC) binds. That scheme balances the incentives for the principal with the punishment for the friend.

Third, combining these two effects, two more issues arise. One issue arises if the principal's punishment does involves collusion between friend and enemy. Then the friend's (DEC) binds on-path and when punishing the principal. The principal's (DEC') binds when punishing the friend and potentially on-path. Thus gradual adjustments on-path of the sort discussed in Proposition 6 imply changes in the punishment possibilities which

[14]We provide the technical details in appendix G.

feeds back into the principal's on-path options. We know from Proposition 2 that at the extreme that problem is severe and the principal can either sustain the commitment solution or nothing at all.

Another issue, arises for contracts for which, under commitment, the friend's (DEC) has slack throughout. Then sustaining the commitment solution is possible even if we reduce the ability to punish the friend. That effect may imply a non-monotonicity of implementing the commitment solution when letting $\theta \to 0$. On the one hand, sustaining the friends punishment gets more difficult because it gets harder to punish the principal. On the other hand, sustaining punishment for the friend becomes less relevant as her desire to deviate vanishes. Although an approach toward such a construction is straightforward, the case distinctions make it cumbersome to execute. We consider it beyond scope.

4 Conclusion

We provide a simple model of endorsement that shows that a principal may find it beneficial to fully endorse an agent with less aligned preferences for a task with externalities on all players in the organization. The reason is that there is greater potential in embracing the enemy, thereby moderating her policy choices, than ensuring that a close ally is chosen more often.

We show that for a given conflict between agents, moderately biased principals enjoy larger utility than those at one extreme or those fully balanced, suggesting this is the type of principals we are more likely to see. Moreover, the principal dealing with an all-powerful ally that is not fully aligned finds it beneficial to add a distant agent to the relationship to induce competition, especially if agents take action frequently. Agents, in turn, find it beneficial to appear more extremist than they are to extract larger endorsements from the principal.

While our model highlights a novel channel in dynamic relationships that leads a power broker to endorse her enemy, there are several aspects our model ignores that provide scope for future research. Among these are the possibility of evolving preferences, the question of perfect monitoring of the principal's endorsement, and the potential role of asymmetric information.

A Notation

For later reference, we introduce here some notation that we use in the proofs.

Histories. Our game is a repeated extensive form game. A history h is a vector that fully describes the set of past actions and realizations. The initial note is given by $h_0 \equiv \emptyset$. A history h is said to be included in history h', i.e., $h \subseteq h'$, if history h' is reached in a continuation game of h. If $h \neq h'$ we say $h \subset h'$. To ensure readability, we abuse notation and suppress within stage game histories. For example, we write $y_k(h)$ meaning agent k's policy when selected in the stage game starting at history h following "on-path play" by the principal in her choice $s(h)$. This allows us, for example, to compare the mutually exclusive $y_L(h)$ and $y_R(h)$ in a compact manner.

Moreover, we will use y_k without a particular history when we describe behavior that is on-path stationary in the following sense: *In all events that occur with positive proabability ex-ante in which k gets to act, she chooses action y_k.*

Continuation Values. We denote i's greatest (smallest) continuation value sustainable as a contract conditional on k leading by $\overline{v}_i(k)$ ($\underline{v}_i(k)$); analogously, we use $\overline{w}_i$ and $\underline{w}_i$ for the beginning-of-a-period continuation values. The existence of a greatest and smallest value is guaranteed by the compactness of the equilibrium payoff set and the self-generating property which follows from standard arguments.

Effective Discounting. Our game is a stochastic game, such that players weight future paths of play with their probability of occurring. Because the evolution of the state takes a particular form, we sometimes need to describe (conditional) probabilities of future events occurring. To avoid long expressions, we use $\delta(h'|h)$ to describe the history-h values of events expected to occur at a future history $h' \supset h$. That is, if history h occurs at time t and history $h' \supseteq h$ occurs at t', we have an effective discount factor of

$$\delta(h'|h) := \beta^{t'-t} Pr(h'|h)$$

where $Pr(h'|h)$ is the probability that history h' occurs conditional on having reached history h. For any $h' \not\supseteq h$ we set $\delta(h'|h) = 0$.

B Useful Lemmas

Here we state and prove 4 Lemmas that reduce the relevant contract space significantly. Lemma 1 establishes a natural order in the on-path action choices in *all relevant* contracts, Lemma 2 establishes 3 key properties of the *optimal* contract. Lemma 3 and 4 provide further structure on punishment and optimality in the commitment case.

Lemma 1. *Take any contract with ex-ante value $w_P(h_0)$ to the principal. Then there exists a contract with value $w_P(h_0)$ such that $y_L(\cdot) \leq y_R(\cdot)$.*

Proof. We prove this lemma by constructing a contract that replaces policies at a particular history h such that at this history $y_L(h) \leq y_R(h)$ without changing anything else, in particular not the ex-ante payoff to the principal. To achieve this, consider a contract C where at some history h, $y_L(h) > y_R(h)$. At this history, the principal's selection decision is $p(s(h))$, and the principal's continuation value before selection is $w_P(h)$.

Now, suppose that, in addition, either $y_L(h) < \theta$ or $y_R(h) > \theta$. Consider a contract $\widetilde{C}$ that is identical to C except that at history h it specifies $\tilde{y}_L(h) = y_L(h) - \varepsilon$ and $\tilde{y}_R(h) = y_R(h) + \frac{1-p(s(h))}{p(s(h))}\varepsilon$. Because $|\tilde{y}_k - \theta_k| \leq |y_k - \theta_k|$ at any history, $\tilde{C}$ is indeed a contract. Moreover, notice that at history h the principal's value is identical because

$$\begin{aligned}\tilde{w}_P(h) - \beta E_{h'}[w_P(h')|h] &= p(s(h)\tilde{y}_R(h) + (1 - p(s(h))\tilde{y}_L(h) \\ &= p(s(h)y_R(h) + (1 - p(s(h))\varepsilon + (1 - p(s(h))y_L(h) - (1 - p(s(h)))\varepsilon \\ &= w_P(h) - \beta E_{h'}[w_P(h')|h],\end{aligned}$$

where $\beta E_{h'}[w_P(h')|h]$ describes the net-present value to the principal of the game *after* completing this period's stage game where the "present" is node h. This object is constant across the two contracts by construction.

Now consider the case $y_R(h) \leq \theta \leq y_L(h)$. As before, consider $\widetilde{C}$ that is identical to C but with

$$\tilde{y}_L(h) = \theta - \frac{p(s(h))}{1 - p(s(h))}(\theta - y_R(h)) \qquad \text{and} \qquad \tilde{y}_R(h) = \theta + \frac{1 - p(s(h))}{p(s(h))}(y_L(h) - \theta).$$

Thus, $\tilde{C}$ is a contract with $y_L(h) \leq y_R(h)$, and P's value is unchanged because

$$\begin{aligned}\tilde{w}_P(h) - \beta E_{h'}[w_P(h')|h] =& p(s(h)) \left(\frac{(1 - p(s(h)))}{p(s(h))}(y_L(h) - \theta) \right) \\ &+ (1 - p(s(h)) \left(\frac{p(s(h))}{1 - p(s(h))}(\theta - y_R(h)) \right) \\ =&(1 - p(s(h))(y_L(h) - \theta) + p(s(h))(\theta - y_R(h)) \\ =&w_P(h) - \beta E_{h'}[w_P(h')|h].\end{aligned}$$

□

Lemma 2. *The optimal on-path policies satisfy $y_R(h) \geq \theta \geq y_L(h)$ for any history h. It is without loss of generality to focus on contracts in which whenever an agent gets to choose y on the equilibrium path, (i) either he chooses the principal's bliss point $y_k = \theta$, or (ii)* (DEC) *binds.*

Proof. We prove both statements constructive and separately.

First Statment: $y_R \geq \theta \geq y_L$. We show that for any contract violating $y_R \geq \theta \geq y_L$, there exists another contract that does not violate the inequalities and is weakly preferred by P.

Fix a contract C such that $y_R(h) < \theta$ for some on-path history h. We will augment the contract in (at most) two steps to construct an alternative that the principal prefers. First, consider an identical contract apart from increasing $y_R(h)$ marginally. That contract improves the principal's payoff at node h by dy_R, relaxes R's (DEC) in all histories $h' \subseteq h$ and leaves them unchanged in all period $h'' \supset h$. It is thus incentive compatible for R. Now consider the last history $\hat{h} \subset h$ in the sequence leading up to h in which it is L's turn to act if such a history exists. The increase of $y_R(h)$ has tightened L's (DEC) by $\delta(h|\hat{h})dy_{R,t}$. Thus, augmenting the contract for a second time by reducing the prescribed $y_L(\hat{h})$ by at most $\delta(h|h')dy_{R,t}$, restores (DEC) of L, without violating R's (DEC) anywhere. Now, observe that the first and second augmentation jointly are at worst ex-ante payoff neutral to P because at node $\hat{h}$ her gains and losses equate. Hence, assuming $y_R(h) \geq \theta$ is without loss. The case for $\theta \geq y_L$ is analogous.

Second Statement: Either (DEC) **binds or** $y_k = \theta$**.**

Definition 1. A history h with property Π is said to be a *first history with property* Π if there is no history $h' \subset h$ that has property Π. It is a *last history with property* Π if no history following h has property Π.

Step 1. k's dynamic enforcement constraint upon first selection. Take an arbitrary contract C with $y_R(\cdot) \geq y_L(\cdot)$ and assume there is a first history in which k leads, $h \in \mathcal{H}_k^*$, and k's (DEC) has slack. Then consider a contract $\widetilde{C}$ identical to C apart from $\tilde{y}_k(h)$ which is chosen such that $|\tilde{y}_k(h) - \theta| < |y_k(h) - \theta|$ and k's (DEC) at h is not violated. Since $\tilde{y}_k$ implies greater moderation, $\tilde{C}$ is strictly preferred by P and $-k$. Now, either $\tilde{C}$ exists for $\tilde{y}_k(h) = \theta$, or there is a $\tilde{y}_k(h) \neq \theta$ such that k's dynamic enforcement constraint binds. We iterate this procedure for any $h \in \mathcal{H}_k^*$ and obtain that an optimal contract exists in the class $\mathcal{C}^1$ in which the second statement holds for any k and $h \in \mathcal{H}_K^*$.

Step 2. k**'s dynamic enforcement constraint upon re-selection.** Consider a contract C^1 in the class of $\mathcal{C}^1$ and assume that there is at least one history h such that k's dynamic enforcement constraint has slack. Take the first on-path history h with that property.

Consider then a candidate contract C^2 which is identical to C^1 with the exception that at history h, we choose the associated $\tilde{y}_k(h)$ such that either $\tilde{y}_k(h) = \theta$ or k's dynamic enforcement constraint binds. Using the same arguments as in step 1, C^2 is incentive compatible for P and $-k$. However, it may not be incentive-compatible for k at a previous selection.

To construct a contract incentive compatible for k, take the last on-path history $h' \subset h$ at which k led before h. Consider now a contract C^3 identical to C^2, but with the difference that at history h', agent k chooses $\tilde{y}_k(h')$ such that $|\tilde{y}_k(h') - y_k(h')| = \delta(h|h')|\tilde{y}_t(h) - y_t(h)|$ and $|\tilde{y}_k(h') - \theta_k| < |y_k(h) - \theta_k|$. Such an $\tilde{y}_k(h')$ exists by construction and implies that k decreases his moderation at history h' by exactly the net-present value of his own increase in moderation at history h. Thus, k is indifferent at history h' between contracts C^3 and C^1 and, by construction, k's dynamic enforcement constraint at node $h' \subset h$ holds in contract C^3. Further, since these concessions are equivalent in history-h' net-present values, they are also equivalent at any $h'' \subset h'$. Two implications follow from this feature: first, C^3 is ex-ante payoff equivalent to C^1 for P and second, C^3 is incentive compatible if the initial C^1 is.

Using the procedure from Step 2 iteratively at any history in which agent k leads, the statement follows. □

Definition 2 (Grim Trigger punishment)**.** The punishment contract of agent k is said to be *grim trigger punishment* if it prescribes that the principal fully endorses $j \neq k$ in any continuation game and both agents choose their respective bliss points whenever leading, $y_L = 0$ and $y_R = 1$.

Lemma 3. *An agent's worst sustainable payoff conditional on leading is the one obtained from grim trigger punishment if and only if* $(1-\beta) \geq 2m\beta$. *It is given by* $b + \beta\alpha^C$*,*

$$\alpha^C := \frac{(1/2-m)b - (1/2+m)}{1-\beta}.$$

Proof. Suppose an agent prefers to lead. We then construct the deviator agent's worst contract. First, P has commitment power and thus can promise to support the non-deviator unconditionally. Moreover, playing one's bliss point is trivially incentive compatible and, by the zero-sum property, the worst choice for the non-leading agent. The resulting payoff is α^C.

To close the argument, we need to show that an agent prefers to lead. For this step, we use the zero-sum property. Winning is desired if the worst payoff conditional on leading, $\underline{v}_L(L) = b + \beta\alpha^C$, is larger than the best payoff conditional on not leading,

$$\overline{v}_L(R) = \frac{b-1}{1-\beta} - \underline{v}_R(R) = \frac{b-1}{1-\beta} - (b + \beta\alpha^C).$$

Thus,

$$\begin{aligned} \frac{(b-1)}{(1-\beta)} &\leq 2(b+\beta\alpha^C) = 2b + \frac{\beta(1-2m)(b-1)}{(1-\beta)} - \frac{\beta(1-m)}{1-\beta} \\ \Leftrightarrow \quad \frac{(1+b)(1-\beta-2m\beta)}{1-\beta} &\geq 0 \qquad \Leftrightarrow (1-\beta) \geq 2m\beta. \end{aligned}$$

□

Lemma 4. *Assume commitment of the principal. Suppose there is an optimal contract such that* $y_L(\cdot) = \theta$*, and* $y_R(h) \neq \theta$ *for some* h*. Then in such optimal contract, the principal either fully endorses* R *after the first time* R *leads, or* L*'s* (DEC) *binds.*

Proof. Consider a history h and assume R leads. We can write R's on-path value at h as

$$\begin{aligned} v_R(h) &= b - (1 - y_R(h)) \\ &+ \beta\left(\sum_{\eta \supset h} \delta(\eta|h)\left(\left(\frac{1}{2} + s(\eta)\right)(b - (1 - y_R(\eta))) + \left(\frac{1}{2} - s(\eta)\right)(y_L(\eta) - 1)\right)\right) \end{aligned}$$

Similarly, we can write P's on path value at h conditional on R leading as

$$\begin{aligned} v_P(h) &= \theta - y_R(h) \\ &+ \beta\left(\sum_{\eta \supset h} \delta(\eta|h)\left(\left(\frac{1}{2} + s(\eta)\right)\left(\theta - y_R(\eta)\right)) + \left(\frac{1}{2} - s(\eta)\right)(y_L(\eta) - \theta)\right)\right). \end{aligned}$$

Summing the two yields

$v_R(h) + v_P(h) =$

$$b - (1-\theta) + \beta \left(\sum_{\eta \supset h} \delta(\eta|h) \left(\left(\frac{1}{2} + s(\eta) \right) (b + 2\theta - 2y_L(\eta)) + (2y_L(\eta) - (1+\theta)) \right) \right).$$

We now substitute $y_L = \theta$ everywhere, and replace $v_R(h) = b + \beta\alpha^C$ because (DEC) binds for R by Lemma 2 and the fact that $y_R(h) \neq \theta$. We obtain

$$v_P(h) = (\theta - 1)(1+\beta) + \beta \left(\sum_{\eta \supset h} \delta(\eta|h) b(\frac{1}{2} + s(\eta)) \right) - \beta\alpha^C \tag{1}$$

and thus, the value of the contract increases in $s(\eta)$ for any $\eta \supset h$. Thus, it is optimal to set $s(\eta) = m$, which proves the claim if $(s = m, y_L = \theta)$ satisfies L's (DEC). □

C Proof of Proposition 1

Proof. Since $\theta = 0$, in any optimal contract $y_L = \theta$ with L's (DEC) trivially satisfied.

Before R leads for the first time, i.e., for $h \in \mathcal{H}_L$, P's endorsement has no effect on R's incentives, thus $s(h) = -m$.

After the first time R leads, i.e., for $h \notin \mathcal{H}_L$, by Lemma 2, either R chooses P's preferred policy θ or R's dynamic enforcement constraint binds. Applying Lemma 4 this implies $s(h) = m$. Solving for $y_R(h)$ using inequality (DEC) yields y_R^*. □

D Proof of Proposition 2

Proof. When the commitment contract yields the first best, the result holds trivially. For the remainder we prove the non-first-best case.

First Statement, "⇐" Direction. We begin with the *if* part. That is, we assume $b > \bar{b}_0$ and show that the commitment solution is implementable.

Take a contract that, on the equilibrium path, is identical to the commitment contract. Further, assume the following off-path punishment contracts, which are all (on-path) stationary and are fully described by the tuple (y_L, y_R, s):

If R **deviated.** The continuation contract is $(y_L = 0, y_R = 1, s = -m)$

If P **deviated.** The continuation contract is $(\hat{y}_L, y_R = 1, s = -m)$; $\hat{y}_L > 0$ and makes L's (DEC) bind.

If L **deviated.** The continuation contract is $(y_L = 0, y_R = \hat{y}_R, s = m)$, where $\hat{y}_R$ makes P's dynamic enforcement constraint bind.

The commitment outcome is a contract depends on if $\hat{y}_L$ is such that P prefers to stay on-path. However, to determine $\hat{y}_L$ we must determine $\hat{y}_R$ which, by construction, needs to satisfy $\hat{y}_R \geq y_R^*$.

In the payoff space, P's and L's punishment payoffs α_P and α_L are related as follows:

$$\alpha_P = -\frac{1}{1-\beta}(p(m)\hat{y}_L + (1-p(m))))$$

where $\hat{y}_L$ is given by L's (DEC) in P's punishment contract:

$$b - \hat{y}_L + \beta \underbrace{(\alpha_P + \frac{p(m)b}{1-\beta})}_{w_L} = b + \beta\alpha_L,$$

which depends on L's punishment payoff

$$\alpha_L = \frac{(1-p(m))b - p(m)\hat{y}_R}{1-\beta},$$

where $\hat{y}_R$ is given by P's (DEC') in L's punishment contract, $-p(m)\hat{y}_R + \beta\alpha_P = \alpha_P$. Solving these four (linear) equations for the unknowns $\hat{y}_L$, $\hat{y}_R, \alpha_L$, yields in particular

$$\alpha_P = -\frac{1-\beta-2m+2(1+b)\beta m + 4b\beta m^2}{2(1-\beta)^2}.$$

Next, note that P's (DEC') in the commitment contract becomes relevant, *after* R led for the first time. The continuation contract from then onwards is stationary with a continuation payoff following a history $h \in \mathcal{H}_R^*$,

$$w_P(h) = -p(m)y_R^* = -\left(\frac{1}{2}+m\right)\left(1 - \frac{4(1+b)\beta m}{2(1-\beta(1-p(m)))}\right), \tag{2}$$

where y_R^* is given by Proposition 1. Pluggin into $w_P(h) \geq \alpha_P$ and solving for b, we obtain the sufficient condition:

$$b \geq \frac{(1-\beta)^2}{\beta(2-\beta(1+1/2-m))\,(1/2+m)} =: \bar{b}_0.$$

First Statement, "⇒" Direction. To prove the *only if* part, we need to show that the α_P constructed is an optimal punishment payoff for P.

By Lemma 14 in appendix G, the constructed α_P is P's optimal punishment payoff if α_L is indeed L's optimal punishment payoff.

To see that L's optimal punishment implies α_L, observe first that if $\hat{y}_R = 1$ in the description of L's punishment, we're back to the commitment punishment. If implementable, this is L's punishment. However, if not, then P's dynamic enforcement constraint (DEC') holds with equality. Now fix a contract with $(s < m, \tilde{y}_R)$ and assume (DEC') holds with equality. Then we can increase s and decrease $\tilde{y}_R$ such that P's value remains the same and R's (DEC) continuoues to hold. That operation makes L worse off since P and L have identical preferences over policies, but L, in addition, suffers from leading less often. Thus, any $(s < m)$ cannot be L's worst punishment.

Second Statement. Finally, we need to show that for $b < \bar{b}_0$ only the repeated stage Nash can be implemented.

To prove this statement, take a stationary contract $(s, y_L, \hat{y}_R)$ with $y_L = 0$ and $\hat{y}_R$

such that R's (DEC) binds. Now consider P's on-path continuation payoff[15]

$$w_P(s) = -p(s)\underbrace{\frac{1-\beta p(m) - b\beta(m+s)}{1-(1-p(s))\beta}}_{=\hat{y}_R}.$$

Based on this term, we make three observations

1. at $s = -m$, P's (DEC') holds
2. at $s = -m$, $w_P(s)$ decreases in s if $b < \bar{b}_0$
3. throughout, $w_P(s)$ is convex in s

The first observation holds trivially. For the second observation, note that

$$\left.\frac{\partial w_P(s)}{\partial s}\right|_{s=-m} \leq 0 \Leftrightarrow b \leq \frac{1}{1-p(m)}\frac{1-\beta}{\beta} =: \hat{b}.$$

Now, for $\hat{b} < \bar{b}_0$ we need that (dividing both sides by the common terms $(1-\beta)/\beta$ and multiplying by $p(s)$)

$$\frac{1/2+m}{1/2-m} < \frac{1-\beta}{2-\beta(3/2-m)}.$$

Since the left-hand side is bounded below by two and the right-hand side is bounded above by $1/2$, such an order never occurs. Thus, whenever $b < \bar{b}_0$, then $b < \hat{b}$ and the inequality is never satisfied. The last observation follows because $\hat{y}_R$ strictly decreases in s at an increasing rate and the first term, $-p(s)$, linearly decreases in s.

Because α_P is constant in the on-path s and $w_P(-m)$ is implementable, we get the following result: If the principal can implement any $s > -m$ in a stationary contract, she can implement $s = m$ which is strictly preferred ex-ante.

To close the proof, what remains to be shown is that no contract outside the class of the stationary contracts described above is implementable and delivers a larger payoff than repeated stage Nash when $b < \bar{b}_0$.

To see this, observe first, that if such an optimal contract exists, an optimal contract exists in which $y_L = 0$ whenever L leads by Lemma 2. Second, because $y_R \neq \theta$, again by Lemma 2, R's dynamic enforcement constraint has to hold with equality. That property implies that R's continuation payoffs are stationary—they depend solely on who got is leading.

But then, a version of Lemma 4 applies. Using the arguments from that proof, it is optimal to increase s up to the point at which P's dynamic enforcement constraint binds. Iterating backward, that means, P's dynamic enforcement constraint binds at *all* instances after R has led for the first time. Thus, the stationary class of contracts considered above is sufficient. □

[15]we abuse notation throughout this prove conditioning w_P on the stationary strategy s rather than a specific history.

E The Commitment Contract and its Implications

Here, we fully characterize the optimal commitment contract for an arbitrary (β, m, b, θ) through a series of Lemmas.

Proposition 7. *The optimal contract implies Results 1 to 7.*

Proof. The proof is organized through the following steps.

1. Lemma 5 to 7 characterize a first-best region providing the if-part of Result 1.
2. Lemma 8 and 9 characterize the full embracing region and the associated contract implying Results 2 and 3 for $\theta \leq \underline{\theta}$.
3. Lemma 10 to 12 characterize the optimal contract for $\theta \in (\underline{\theta}, 1/2]$ implying Result 4, Result 2 for $\theta > \underline{\theta}$, and, in combination with the previous optimality results, the only-if part of Result 1.
4. Lemma 13 provides the implications Results 5 to 7.

To simplify notation, we make frequent use of the following shorthand expressions

$$\hat{\theta} := p(m)\beta \qquad\qquad \psi := -\beta\left(\frac{b+1-2\beta b}{1-\beta} + 2\beta\alpha^C\right).$$

Lemma 5. *Suppose $\breve{\theta} \leq \theta \leq \hat{\theta}$. Then there is an optimal contract $(s, y_L, y_R) = (m, \theta, \theta)$.*

Proof.

Step 1. A candidate. Take the candidate $(s, y_L, y_R) = (m, \theta, \theta)$ with deviations punished via grim trigger.

Step 2. The candidate is a contract. Optimality follows because $y_k{=}\theta$. R's (DEC) is

$$b - (1 - y_R) + \frac{\beta}{1-\beta}\Big(p(m)(b - (1 - y_R)) - (1 - p(m))(1 - \theta)\Big) \geq b + \beta\alpha^C$$

which is equivalent to

$$y_R \geq 1 - \beta\frac{4m(b+1) + \theta(1-2m)}{2 - \beta(1-2m)}. \tag{3}$$

Hence, $y_R = \theta$ satisfies (DEC) if

$$\theta \geq 1 - \beta\frac{4m(b+1) + \theta(1-2m)}{2 - \beta(1-2m)} \qquad \Leftrightarrow \qquad \theta \geq 1 - \beta\left(\frac{1}{2} + m(1+2b)\right) = \breve{\theta}.$$

Analogously, L's (DEC) is

$$b - y_L + \frac{\beta}{1-\beta}\Big((1 - p(m))(b - y_L) - p(m)\theta\Big) \geq b + \beta\alpha^C$$

which implies

$$y_L \leq \frac{p(m)(1-\theta)\beta}{1 - \beta p(m)}.$$

Hence, $y_L = \theta$ satisfies (DEC) if

$$\theta \leq \frac{p(m)(1-\theta)\beta}{1-\beta p(m)} \qquad \Leftrightarrow \qquad \theta \leq p(m)\beta = \hat{\theta}. \qquad \square$$

Lemma 6. *Suppose* $\theta \geq \max\{1-\overline{\theta}, \underline{\theta}\}$. *Then, the optimal contract implies* $y_k = \theta$.

Proof.

Step 1. A candidate. Consider the following strategy profile with $y_k = \theta$ and on-path strategies, (s_L, s_R), where s_L solves

$$\theta = \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) \underbrace{-\beta\left(\frac{b+1-2\beta b}{1-\beta} + 2\beta\alpha^C\right)}_{=\psi} s_L,$$

and s_R solves

$$\theta = 1 - \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_R.$$

The first-period endorsement strategy is arbitrary, with deviations punished via grim trigger.

Step 2. The candidate is a contract. The joint per-period utility of L and R is $u_L + u_R = b - 1$. Assume, for now, that we are in a setting in which both players' (DEC) binds when leading. THus, the leading agent's continuation payoff is identical to deviating and entering the worst punishment, i.e., $b + \beta\alpha^C$. That implies that the non-leading agent receives the residual $(b-1)/(1-\beta) - (b+\beta\alpha^C)$. But then, because both (DEC) bind,

$$-y_L + \beta\left((1-p(s_L))\left(b+\beta\alpha^C\right) + p(s_L)\left(\frac{b-1}{(1-\beta)} - (b+\beta\alpha^C)\right)\right) = \beta\alpha^C$$

$$(y_R - 1) + \beta\left(p(s_R)\left(b+\beta\alpha^C\right) + (1-p(s_R))\left(\frac{b-1}{(1-\beta)} - (b+\beta\alpha^C)\right)\right) = \beta\alpha^C,$$

we can solve for y_L and y_R as

$$y_L(s_L) = \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_L; \text{ and } y_R(s_R) = 1 - \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_R. \quad (4)$$

Note that since both equations in (4) are expected to hold in all future periods, the leading agent's choice is only a function of the next expected endorsement by P. Thus, $y_k = \theta$ satisfies both agents' (DEC), if and only if there are $s_k \in [-m, m]$ such that each RHS in equation (4) is equal to θ.

Assume for now that $\psi < 0$, which is a property that is necessary and sufficient for the interval we construct to be non-empty, and that we will verify later. Then, an s_L that ensures $y_L = \theta$ exists if and only if

$$\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi m \leq \theta \leq \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) - \psi m,$$

and rearranging it,

$$\underline{\theta} = \frac{\beta}{1-\beta}(b+1)mp(m)\beta \leq \theta \leq \frac{\beta}{1-\beta}(b+1)m\left(2-\beta p(m)\right) = \overline{\theta},$$

which is non-empty (and hence indeed $\psi < 0$) if and only if $1-\beta > 2m\beta$. Similarly, an s_R that ensures $y_R = \theta$ exists if and only if

$$1-\beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi m \leq \theta \leq 1-\beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) - \psi m,$$

which is equivalent to $1-\overline{\theta} \leq \theta \leq 1-\underline{\theta}$. Using that $\theta \leq 1/2$ we can conclude

$$\theta > 1-\overline{\theta} \Rightarrow \theta < \overline{\theta}, \quad \text{and} \quad \theta > \underline{\theta} \Rightarrow \theta < 1-\underline{\theta}.$$

Thus, a first best contract exists $\theta > \max\left\{1-\overline{\theta}, \underline{\theta}\right\}$ conditional on $\psi < 0$, which remains to be shown.

Recall from above that $\psi < 0$ if and only if $1-\beta > 2m\beta$. Now observe that if instead $1-\beta \leq 2m\beta$, then

$$\underline{\theta} = \frac{1}{1-\beta}\underbrace{m\beta}_{\geq\frac{1-\beta}{2}}\underbrace{(b+1)}_{>1}\underbrace{(1+2m)\beta}_{\geq 1} > 1/2$$

which implies that $\psi \geq 0 \Rightarrow \theta < \underline{\theta}$ making Lemma 6 redundant.

Step 3. Exclusion Phase. Finally, since $y_i = \theta$ in any continuation game, P is indifferent between any initial endorsement, which concludes the proof. □

Lemma 7. *Suppose $\underline{\theta} > \theta > \hat{\theta}$. Then the optimal contract implies $y_k = \theta$.*

Proof.

Step 1: A candidate. Consider the following strategy profile with $y_k = \theta$ and on-path strategies, $s_L = m$ and,

$$s_R = m - \frac{\theta - \beta p(m)}{\beta\left(\beta p(m)(b+1) - \theta\right)} =: \overline{s}_R.$$

The first-period endorsement strategy is arbitrary, deviations are punished via grim trigger.

Step 2: The candidate is a contract. We first show that the candidate satisfies L's (DEC). L's value conditional on leading is

$$v_L(L) = b - \theta + \beta\left(p(m)v_L(R) + (1-p(m))v_L(L)\right),$$

and L's value conditional on *not* leading is

$$v_L(R) = -\theta + \beta\left(p(s_R)v_L(R) + (1-p(s_R))v_L(L)\right).$$

Solving this system for $v_L(L)$ and $v_L(R)$, we obtain

$$v_L(L) = \frac{(1-\beta p(s_R))b}{(1-\beta)(1+\beta(m-s_R))} - \frac{\theta}{(1-\beta)}, \text{ and } v_R(L) = \frac{(1-p(s_R))\beta b}{(1-\beta)(1+\beta(m-s_R))} - \frac{\theta}{1-\beta}.$$

Hence, L's (DEC) holds if and only if $v_L(L) \geq b + \beta\alpha^C \Leftrightarrow s_R \leq \overline{s}_R$, which, in turn, holds by assumption. Observe that $\overline{s}_R < m$ if $\underline{\theta} > \theta > \hat{\theta}$ and $\overline{s}_R = m$ if $\theta = \hat{\theta}$.

Consistency with R's (DEC) follows from recalling that $u_L + u_B = b - 1$ in every period, which implies that

$$v_R(R) = \frac{b-1}{1-\beta} - v_L(R) = \frac{b+\theta-1}{1-\beta} - \frac{(1-p(s_R))\beta b}{(1-\beta)(1+\beta(m-s_R))}.$$

Our candidate satisfies R's (DEC) if

$$v_R(R) \geq b + \beta\alpha^C \qquad \Leftrightarrow \qquad s_R \geq m - \frac{\theta - (1-\beta p(m)) + 2bm\beta}{\beta\left((b+1)(1-\beta p(m)) - \theta\right)} =: \underline{s}_R.$$

What remains, is to show that $\overline{s}_R \geq \underline{s}_R$ for all $\theta \in (\hat{\theta}, \underline{\theta})$ which is equivalent to showing

$$\frac{4b\left((b+1)m(1+2m)\beta^2 - (1-\beta)\theta\right)}{\beta\Big((1+b)^2(1+2m)\beta(2-\beta(1+2m)) - 4\theta\left((1+b)-\theta\right)\Big)} > 0.$$

Because $\theta < \underline{\theta}$ by assumption, the numerator is positive. So what remains, is to show that the denominator is positive, too. We do this by sequentially deriving bounds. First notice that the relevant term (within the larger brackets) is decreasing in θ because $b > 0$ and $\theta < 1/2$. Thus, that term is positive for all $\theta \in (\hat{\theta}, \underline{\theta})$ if and only if it is non-negative for $\theta = \underline{\theta}$ that is if

$$\frac{(1+b)^2(1+2m)\beta}{(1-\beta)^2}\left(2 + \beta\left(2m + 4\beta - (1-2m)\beta^2(1+2m)^2 - 5\right)\right) \geq 0. \tag{5}$$

Once again, the term in large brackets is the relevant one and needs to be non-negative. We will first show that that term is increasing in m. The derivative of the term in big brackets is

$$\beta(2 - 4(1-4m^2)\beta^2 + 2(1+2m)^2\beta^2).$$

The second derivative is $(1+6m)\beta^3 \geq 0$, which implies an increasing first derivative for $m \in (0, 1/2)$. Now observe that the first derivative at $m = 0$ is $2\beta(1-\beta^2) > 0$ which implies the term in brackets of (5) is increasing on $m \in (0, 1/2)$.

Finally, the term itself at $m = 0$ collapses to $(2-\beta)(1-\beta)^2 > 0$, which proves that it is positive for any $m \in (0, 1/2)$.

Step 3. Exclusion Phase. Finally, since $y_k = \theta$ in any continuation game, P is indifferent between any initial endorsement which concludes the proof. □

Lemma 8. *If $\theta \leq \underline{\theta}$, there is an incentive-compatible contract with $y_L = \theta$ and $y_R \geq \theta$. If, in addition, $\theta < \hat{\theta}$, agent L's dynamic enforcement constraint has slack. In that contract*

$s = -m$ until the first time R leads. Thereafter, $s = m$.

Proof. The claim holds for all $\theta \geq \hat{\theta}$ by Lemma 7. Observe that $\hat{\theta}$ and $\underline{\theta}$ intersect at most once for $\beta > 0$ because both are increasing in β with $\underline{\theta}$ convex, $\hat{\theta}$ linear and $\underline{\theta} \to \hat{\theta}$ as $\beta \to 0$. This intersection occurs at $\beta = \frac{1}{1+2m(1+b)} =: \hat{\beta}$, and $\hat{\theta} < \underline{\theta}$ if and only if $\beta > \hat{\beta}$. Thus, what remains is to show the statement for $\underline{\theta} < \theta < \hat{\theta}$.

Step 1. A candidate. Consider the following candidate with an on-path strategy $y_L = \theta$ whenever L leads, and an on-path strategy

$$\tilde{y}_R = 1 - \frac{2m(b+1) + \theta(1/2 - m)}{1 - \beta(1/2 - m)}\beta.$$

whenever R leads. Until R leads for the first time, $s_0 = -m$, and thereafter, $s = m$ in any on-path node that follows. Deviations are punished via grim trigger.

Step 2. The candidate is a contract. Take R's (DEC) under the candidate. Observe that

$$\tilde{y}_R = 1 - \frac{2m(b+1) + \theta(1/2 - m)}{1 - \beta(1/2 - m)}\beta.$$

solves R's (DEC) with equality. Because $\theta < \hat{\theta}$ the resulting $\tilde{y}_R > \theta$ (see the proof of Lemma 5 for details). Now consider L's (DEC) which is

$$(b - \theta) + \frac{\beta}{1-\beta}\Big(1/2 - m)(b - \theta) - (1/2 + m)\tilde{y}_R\Big) \geq b + \beta\alpha^C,$$

and substituting for $\tilde{y}_R$ and α^C implies

$$\frac{2m(b+1) + \theta(1/2 - m)}{1 - \beta(1/2 - m)}\beta \geq \frac{1 - \beta(1/2 + m)}{\beta(1/2 + m)}\theta,$$

which, in turn, is equivalent to

$$\theta \leq \frac{\beta^2 m(2m+1)}{1-\beta}(b+1) = \underline{\theta}.$$

Step 3. Initial periods. L's (DEC) has slack when playing the principal's bliss point, $y_L = \theta$, and receiving no endorsement. Thus, more endorsement, $s < m$, leaves L's incentives undistorted. Because R optimizes dynamically, endorsement decisions prior to R's first lead are inconsequential. □

Lemma 9. *The contract constructed in the proof of Lemma 8 is optimal in its region.*

Proof. In steps 1 and 2 we focus on histories $h \notin \mathcal{H}_L$ and show that no better contract for P exists, in step 3 we close by addressing $h \in \mathcal{H}_L$.

Step 1. The relevant class of contracts. Invoking Lemma 2, we, without loss, restrict attention to contracts such that at any history the leading agent either chooses $y_i = \theta$ or that agent's (DEC) binds. We can restrict the contract space further by focusing on $y_L \leq \theta \leq y_R$. Lemma 4 implies that within the class of contracts such that $y_L = \theta$ and

$y_R \neq \theta$ we only need to consider those in which $s = m$ unless either L's (DEC) binds or R plays $y_R(h) = \theta$ at least at some on-path history h.

Step 2. No other candidate in that class is optimal.

Step 2a. If L's dynamic enforcement constraint binds, R's dynamic enforcement constraint must have slack thereafter. Consider a history in which a leading L's (DEC) holds with equality. Then, a leading R's (DEC) must have slack if R leads the period immediately after (history h'). To see this, recall first that $u_L + u_R = b - 1$ in each period. Thus, if L's (DEC) binds in h and R's (DEC) binds in h', then L's value of the game when in the lead at h is $b + \beta\alpha^C$ while L's value of the game when R leads in the next period is $(b-1)/(1-\beta) - b - \alpha^C$. For now, assume in addition, that if L leads in h', his (DEC) also binds. Then, L's (DEC) becomes

$$b - y_L + \beta(1/2 - s_L)(b + \beta\alpha^C) + (1/2 + s_L)((b-1)/(1-\beta) - b - \alpha^C) = b + \beta\alpha^C$$
$$\Leftrightarrow y_L = \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_L.$$

If $\theta < \underline{\theta}$, there is no $s_L \in [-m, m]$ such that $y_L \leq \theta$ solves (4) and hence, if R leads in h', R's (DEC) has to have slack. Moreover, notice that relaxing L's (DEC) in h' only makes matters worse. Invoking Lemma 2, that argument implies that $y_R(h') = \theta$ in h' if L's (DEC) binds in h.

Step 2b. No optimal contract exists in which $y_R(h) = \theta$. Finally, we show that there is no optimal contract in which, at h', $y_R(h) = \theta$. To see this, recall that under the candidate from Lemma 8 when R leads she is promised the best possible continuation payoff within the relevant class of contracts in all nodes in which R is not playing: P fully endorses R, and L chooses the highest policy, $y_L = \theta$, available within the relevant class of contracts. Leaving these actions unchanged to derive an upper bound of what is possible, it is only possible to let R choose $y_R(h) = \theta$ if R would choose $y_R(h') > \theta$ if leading at some future history h'. Once again by Lemma 2, no such contract exists in the relevant class of contracts that is such that R's (DEC) holds with slack in h'. But then, from R's perspective such promise is no different from promising to return to the candidate after h ruling out that $y_R(h) = \theta$ satisfies R's (DEC) if $\theta < 1 - \frac{\beta}{2} - (1+2b)m\beta$ and hence the principal's best response to $(s = m, y_L = \theta)$ is $y_R > \theta$.

Step 3. Exclusion Phase. The contract delivers P's first best until R leads for the first time and $s = -m$ minimizes the chances of this event. That completes the proof. □

Lemma 10. *If $\theta \in (\underline{\theta}, \overline{\theta}]$, there exists a contract in which* (DEC) *binds for both agents and $y_L = \theta$. In that contract, $s = -m$ until the first time R leads. Thereafter, $s_R = m$ when R led last and $s_L < m$ when L led last.*

Proof. We restrict attention to $\theta \leq 1 - \overline{\theta}$. Lemma 6 covers the complementary case.

Step 1. A candidate. When L led last, the principal's on-path endorsement strategy is

$$s_L = \frac{\alpha^C\beta + \theta - \frac{b-1}{2}\frac{\beta}{(1-\beta)}}{\psi} < m.$$

When R led last, it is $s_R = m$. On path, $y_L = \theta$ and $y_R = 1 - \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi m$.

Until R leads for the first time, $s_0 = -m$; deviations are punished via grim trigger.

Step 2. The candidate is a contract. Suppose that both (DEC) bind. That is,

$$y_R = 1 - \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi m \qquad y_L = \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_L. \tag{4}$$

The proposed s_L solves the second equation by construction for $y_L = \theta$. However, analogously to the proof Lemma 6, $y_L = \theta$ can only be sustained if $\underline{\theta} \leq \theta \leq \overline{\theta}$ conditional on $\psi \leq 0$. To see that $\psi < 0$ assume for a contradiction that $\psi > 0$ and recall that this implies $\beta(1+2m) > 1$ or $1 - \beta < 2\beta m$. Moreover, recall that, for this case to be relevant, we require

$$1/2 \geq \underline{\theta} \quad \Leftrightarrow \quad \frac{\beta}{1-\beta}\beta m(2m+1)(b+1) \leq 1/2 \quad \Leftrightarrow \quad 2\beta^2 m(2m+1)(b+1) \leq 1-\beta.$$

But then because $\psi > 0$, we get that $2\beta^2 m(2m+1)(b+1) > (1-\beta)(b+1) > (1-\beta)$, which is a contradiction. Hence, $\psi < 0$ whenever $\theta \geq \underline{\theta}$. s_L increases in θ and obtains $s_L = m$ at $\overline{\theta}$ which proves $s_L < m$. By assumption, $\theta < 1 - \overline{\theta}$ and hence $y_R > \theta$ even if $s_R = m$.

Step 3. Embracing Phase. The reasoning of step 3 in the proof of Lemma 8 applies. □

Lemma 11. *The contract constructed in the proof of Lemma 10 is optimal in its region.*

Proof. In Steps 1-3 focus on $h \notin \mathcal{H}_L$. We turn to $h \in \mathcal{H}_L$ in Step 4.

Step 1. The candidate is the optimal contract with binding (DEC)**.** Consider P's problem

$$v_P(s_R; s_L) = \max_{s_R} -(1-\beta)|\theta - y_R(s_R)| + \beta\left(p(s_R)v_P(s_R; s_L) + (1-p(s_R))v_P(s_L; s_R)\right)$$
$$v_P(s_L; s_R) = \max_{s_L} -(1-\beta)|\theta - y_L(s_L)| + \beta\left(p(s_L)v_P(s_R; s_L) + (1-p(s_L))v_P(s_L; s_R)\right),$$

where agents' policies are given by their binding (DEC), which give the equations (4). Using the two Bellman equations above, we can solve for P's value of selecting the optimal strategy s_L in all periods in which L leads, taking the choice s_R in periods in which R leads as given. Thus, P's objective becomes

$$\tilde{v}_P(s_R; s_L) = (\theta - y_R(s_R))\rho_R(s_R; s_L)) + (y_L(s_L) - \theta)(1 - \rho_R(s_R; s_L)) \tag{VPR}$$

and

$$\tilde{v}_P(s_L; s_R) = (\theta - y_R(s_R))\rho_L(s_L; s_R) + (y_L(s_L) - \theta)(1 - \rho_L(s_L; s_R)). \tag{VPL}$$

with

$$\rho_R(s_R; s_L) = \frac{1-\beta(1-p(s_L))}{1-\beta(1-p(s_L))+\beta(1-p(s_R))} \text{ and } \rho_L(s_L; s_R) = \frac{\beta p(s_L)}{1-\beta p(s_R)+\beta p(s_L)}.$$

Observe that in objective $\tilde{v}_P(s_i; s_j)$, s_i is the choice whereas s_j is assumed to be chosen optimally the next time j leads. Taking derivatives of P's objectives yields

$$\frac{\partial \tilde{v}_p(s_L; s_R)}{\partial s_L} = \frac{(2-\beta(1+2s_R))}{2(1-\beta(s_R-s_L))^2}\left(2\beta\theta-\beta\left(y_L(s_L)+y_R(s_R)\right)+y_L'(s_L)\left(1+\beta(s_L-s_R)\right)\right) \quad (6)$$

and

$$\frac{\partial \tilde{v}_p(s_R; s_L)}{\partial s_R} = \frac{(2-\beta(1-2s_L))}{2(1-\beta(s_R-s_L))^2}\left(2\beta\theta-\beta(y_L(s_L)+y_R(s_R))-y_R'(s_R)\left(1+\beta(s_L-s_R)\right)\right) \quad (7)$$

which, after replacing via equation (4), can be written as

$$\begin{aligned}\frac{\partial \tilde{v}_p(s_L; s_R)}{\partial s_L} &= \frac{(2-\beta(1+2s_R))}{2(1-\beta(s_R-s_L))^2}\left(2\beta\theta-\beta\left(1+\psi(s_L+s_R)\right)+\psi+\psi\beta(s_L-s_R)\right) \\ &= \frac{(2-\beta(1+2s_R))}{(1-\beta(s_R-s_L))^2}\left(-\beta\left(\frac{1}{2}-\theta\right)+\left(\frac{1}{2}-\beta s_R\right)\psi\right)\end{aligned} \quad (8)$$

and

$$\begin{aligned}\frac{\partial \tilde{v}_p(s_R; s_L)}{\partial s_R} &= \frac{(2-\beta(1-2s_L))}{2(1-\beta(s_R-s_L))^2}\left(2\beta\theta-\beta(1+\psi(s_L+s_R))-\psi-\psi\beta(s_L-s_R)\right) \\ &= \frac{(2-\beta(1-2s_L))}{(1-\beta(s_R-s_L))^2}\left(-\beta\left(\frac{1}{2}-\theta\right)-\left(\frac{1}{2}+\beta s_L\right)\psi\right).\end{aligned} \quad (9)$$

Recall that the relevant case here is one in which $2\frac{\beta}{1-\beta}(b+1)m\beta\left(\frac{1}{2}+m\right)\underline{\theta} < 1/2$ as otherwise our case is not relevant. Thus, relevance implicitly gives a constraint on the parameter space. But then, with some algebra, we can show that (8) is always positive implying that the principal always has a desire to increase s_L (subject to not violating any constraints) no matter the choice of s_R. Recall, however, that $y_L \leq \theta$ by Lemma 2. But then, our candidate contract precisely describes the largest s_L such that $y_L \leq \theta$.

Replacing s_L with s_L^* in (9) and rearranging implies that (9) also increases in s_R. But then, no constraint is violated even for the corner solution $s_R = 1$, which is therefore optimal.

Thus the candidate contract is the optimal contract within the class of contracts in which both agents' (DEC) bind.

Step 2. R's (DEC) **binds at the optimum.** Suppose that there exists an optimal contract in which R's (DEC) has slack at some node at history h. Then, because that contract is optimal and R's (DEC) holds with slack, there is an optimal contract in which R's (DEC) has slack again in the period immediately thereafter. Using Lemma 2, it is without loss to assume that $y_R(h') = \theta$ for any $h' \supset h$ such that between h and h', L did not lead. Moreover, if such a contract is feasible, it is also feasible assuming that whenever L leads the next time, L's (DEC) holds with equality. Because $u_L + u_R = b - 1$ in every period, such a continuation game implies R's *best* possible continuation game conditional on L leading, i.e., $\overline{v}_R(L)$. But then, when $\theta < 1 - \overline{\theta}$ which holds whenever $\overline{\theta} < 1/2$, all contracts that satisfy R's (DEC) with equality for a fixed y_R until the next time L leads imply $y_R > \theta$. By monotonicity of the (DEC), no contract with $y_R \leq \theta$ exists that satisfies R's (DEC). A contradiction to the premise that R's (DEC) has slack at some history h.

Step 3. No other candidate is optimal. Observe that by Lemma 2, whenever L is not at her (DEC) we need $y_L = \theta$. Moreover, we know from the previous argument that R's (DEC) holds with equality. Invoking Lemma 4 and observing that because $\theta > \underline{\theta}$, $(s = m, y_L = \theta)$ does not satisfy L's (DEC), implies that the candidate contract is optimal.

Step 4. Embracing Phase. The reasoning of step 3 in the proof of Lemma 8 applies. □

Lemma 12. *If $\theta > \overline{\theta}$, an optimal contract is such that $y_i \neq \theta$ and $s_R = -s_L = m$.*

Proof.

Step 1. A candidate. The principal's endorsement strategy is $s_R = -s_L = m$ and the agents choose

$$\tilde{y}_L(s_L) = \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) - \psi m,$$

$$\tilde{y}_R(s_R) = 1 - \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi m,$$

which make their (DEC) bind; deviations are punished via grim trigger.

Step 2. The candidate is a contract. Because $1/2 \geq \theta > \underline{\theta}$, it follows that $\psi < 0$ and thus, $y_L(s_L = -m)$ and $y_R(s_R = m)$ satisfy (DEC)—as shown in (4).

Step 3. The candidate is optimal. Imagine agent i is currently leading. By $u_L + u_R = b - 1$, promising that agent j is brought to his (DEC) the next time he leads implies i's best payoff conditional on any future selection of j, i.e., $\overline{v}_i(j)$. By monotonicity of (4) in s_i and the fact that the agent who led last receives P's full endorsement, no other contract that satisfies agents' (DEC) and has $y_i \neq \theta$ implies y_i closer to θ. □

Lemma 13. *The characterization above implies Results 5 to 7.*

Proof. Result 5 follows from observing that y_R^* (y_L^*) decreases (increases) in θ for $\theta < \overline{\theta}$ and remains constant thereafter. Result 6 follows for $\theta < \overline{\theta}$ because $y_L^* = \theta$ and y_R^* decreases in θ. For $\theta \geq \overline{\theta}$, y_L^* is constant which implies the principal loses in the exclusion phase as θ increases without gaining in the embracing phase. The first part of Result 7 follows because the principal's payoff is continuous and she strictly prefers two agents at $\theta = 1/2$. The second part follows because, by continuity, for any $\theta > 0$, there is a β such that $0 < y_R^*(\beta) - \theta < \theta$. □

□

F Appendix: No Commitment Contracts

F.1 Proof of Proposition 3

Proof. Note first, that as $b \to \infty$, then $\breve{\theta} \to -\infty$ and $\overline{\theta} \to \infty$, thus with commitment P could implement the first best (see Result 1).

Consider now the no-commitment case. We provide a contract that can implement the first best as b becomes arbitrarily large. Consider an on-path 'opportunistic' strategy profile in which there is an initial $s_0 \in [-m, m]$ and then P endorses whoever has led in the period before. Consider also a punishment in which P, upon deviation, fully supports the agent did *not lead* in the last period, before returning to the equilibrium path. Note that, if the s.p. implements the first best, it is an equilibrium—since P is indifferent between any s. Lastly, the s.p. implements the first best because as $b \to \infty$, the loss for the agent if he deviates becomes arbitrary large, ensuring that $y_i = \theta$ satisfies (DEC). Since off- and on-path strategies are monotone, there exists some $\overline{b}$ sufficiently large to ensure the first best. □

F.2 Proof of Propositions 4 and 5

Proof. First, for $\theta = \frac{1}{2}$, P's problem is entirely symmetric, so the commitment grim-trigger punishment for a deviating agent is feasible and thus the commitment contract.

Agents prefer to lead. Consider the optimal contract when $\theta > \overline{\theta}$. Since $\overline{\theta} < \frac{1}{2}$ by assumption, recall that this implies that in the optimal commitment contract both agents are at their (DEC). Through the constant-sum property of the game between agents, that feature implies each agent receives his lowest possible continuation payoff conditional on leading ($v_L^*(L) = \underline{v}_L(L)$) and his greatest possible continuation payoff conditional on not leading ($v_L^*(R) = \overline{v}_L(R)$). Furthermore, note that under commitment $\underline{v}_k(k) \geq \overline{v}_k(-k)$, and this inequality also holds under non commitment. To see why, note that $\overline{v}_L(R)$ and $\underline{v}_R(R)$ do not change because R's punishment is unchanged while $\overline{v}_L(R)$ decreases and $\underline{v}_L(L)$ increases because L's punishment becomes milder.

As a result of the fact that $\underline{v}_L(L) \geq \overline{v}_L(R)$ and $s_L^* = -m$, no contract exists such that $y_L > \overline{\theta}$. That is, the optimal contract $y_L^* = \overline{\theta}$ is the furthest L is willing to move to the center through a stationary strategy.

P's punishment for $\theta > \overline{\theta}$. From Lemma 15 from appendix G and since $y_L^* \leq \theta$, P's worst contract is a "full polarization" contract with $s = -m$, which delivers a punishment value of

$$\alpha_P = \frac{m(1 - 2\theta) - \frac{1}{2}}{1 - \beta}. \tag{10}$$

L**'s punishment after returning: the *back-to-business* property.** We first show that if the optimal contract has L at his (DEC) whenever he leads, it is without loss to assume his punishment contract has the following feature: once L leads again after deviating, we restart the optimal contract. Formally, for any $h \notin \mathcal{H}_R$, $\sigma^{\alpha_L} = \sigma^*$. The reason is that, by construction, the optimal contract gives (i) P her greatest continuation payoff conditional on L leading ($v_P^{\alpha_L}(L) = v_P^*(L) = \overline{v}_P(L)$), *and* (ii) L the worst possible payoff conditional on leading ($v_L^{\alpha_L}(L) = v_L^*(L) = b + \beta\alpha_L$). Combined with the constant-sum property of the game between agents, implies that R receives his greatest continuation payoff conditional on L leading. Hence, no other continuation contract exists that makes L worse off, *or* provides better incentives for R and P to punish L in earlier periods.

What remains is to construct the initial punishment phase, i.e., for any $h \in \mathcal{H}_R$. We doe this separately for the two propositions.

Initial punishment of L **in Proposition 4.** Because we want to sustain the commitment solution and in that solution L's (DEC) binds, we need to construct a contract payoff equivalent to the worst contract under commitment. Naturally, that contract needs to prescribe $s_0^{\alpha_L} = m$ and $y_R^{\alpha_L} = 1$ for any $h \in \mathcal{H}_R$ and needs to satisfy P's enforcement constraint (DEC'). A necessary and sufficient condition for redundancy of P's commitment constraint is inequality

$$w_P^{\alpha_L} = \frac{1}{1-\beta p(m)}\Big(- p(m)(1-\theta) + (1-p(m))v_P^*(L)\Big) \geq \alpha_P \tag{11}$$

which we can write as:

$$\frac{1}{1-\beta p(m)}\Big(- p(m)(1-\theta) + (1-p(m))v_P^*(L)\Big) - \alpha_P \geq 0.$$

Recalling that the optimal commitment contract for $\theta > \overline{\theta}$ prescribes $y_L^* = \overline{\theta}, y_R^* = 1 - \overline{\theta}$ and $s_R^* = -s_L^* = m$ we can write

$$v_P^*(L) = \frac{\overline{\theta}}{1-\beta} - \frac{\theta}{1-2\beta m} - \frac{\beta}{(1-\beta)}\frac{1-p(m)}{(1-2\beta m)}$$

which is a linear function of θ. Since α_P given by (10) is also linear in θ, the inequality is linear in θ with slope

$$\frac{p(m)}{1-\beta p(m)} + \frac{2m}{1-\beta} - \frac{1-p(m)}{1-2\beta m}$$

which is positive for $0 < \beta < 1$ and $0 < m < \frac{1}{2}$. It is straightforward to show that if $\overline{\theta} < 1/2$, there exists a $\hat{\theta} < 1/2$ such that inequality (11) holds if and only if $\theta > \hat{\theta}$.

Lastly, observe that b enters in (11) only through $\overline{\theta}$. Since $v_P^*(L)$ increases in $\overline{\theta}$ which in turn increases in b, we can conclude that if (11) holds for some b, it holds for any $b' > b$ or, equivalently, $\hat{\theta}$ decreases in b. This proves Proposition 4.

L**'s punishment in Proposition 5.** Note that if $\theta \in (\overline{\theta}, \hat{\theta})$, the commitment punishment cannot be sustained by the previous argument. Instead, L's punishment is such that

$\alpha_L > \alpha_R = \alpha^C$.

Now, suppose that P's dynamic enforcement constraint has slack on the equilibrium path. We can invoke the back-to-business property like before, so, what remains is to construct the initial punishment phase. Consider a punishment candidate that prescribes for any $h \in \mathcal{H}_R$, $s_0^{\alpha_L} = m$ and $y_R^{\alpha_L} \geq y_R^*$ chosen such that P's dynamic enforcement constraint holds with equality. We show that this punishment candidate is indeed L's worst contract. Once more, we invoke the back-to-business property. Then, conditional on $s_0^{\alpha_L} = m$, the candidate punishment is worst for L because if $h \in \mathcal{H}_R$, $y_R^{\alpha_L}$ is the highest policy by R such that P's dynamic enforcement constraint is satisfied and if $h \notin \mathcal{H}_R$, L receives her worst payoff upon selection, $v_L^{\alpha_L}(L) = b + \beta\alpha_L$.

It remains to show that no other $s \neq s_0^{\alpha_L}$ makes L worse off. To see that this is the case, recall that since any implementable $y_R > \theta$, P's and L's payoffs are aligned regarding R's policy choice. Also notice that, for any $h \in \mathcal{H}_R$ in the optimal punishment, P is at her dynamic enforcement constraint as otherwise a harsher punishment was possible. But then, any other combination (s, y_R) that gives P the same value $w_P^{\alpha_L} = \alpha_P$ also gives L the same value *ignoring* L's payoff from leading. However, any such s increases the likelihood that L leads compared to $s_0^{\alpha_L} = m$ and thus gives L a higher overall payoff. Thus, no such strategy is worse for L which completes the proof.

P's on-path constraint. Lastly, we return to the assumption that P's dynamic enforcement constraint does not bind on the equilibrium path. Invoking again the back-to-business property of L's punishment, note that since $v_P^*(L) = v_P^{\alpha_L}(L)$ and $s_0^{\alpha_L} = s_R^* = m$, we can compare P's continuation value at the optimal contract and at the initial phase of the punishment and note that $w_P^*(R) \geq w_P^{\alpha_L}(R)$. Hence, whenever it is possible to induce $y_R^{\alpha_L} > y_R^*$ during the initial phase of punishment (or equivalently, whenever $w_P^*(R) \geq \alpha_P$), P's dynamic enforcement constraint indeed has slack on the equilibrium. This proves Proposition 5. □

F.3 Proof of Proposition 6

Proof. First, note that, since $\theta > \overline{\theta}$, Lemma 2, 10 and 12, and, in particular, equations (4) imply that there is no contract in which $y_L > 2\theta$. Hence, invoking Lemma 15 from appendix G, we know that independent of b

$$\alpha_P = \frac{m(1-2\theta) - 1/2}{1-\beta}$$

But then—because, by Lemma 2, L's (DEC) binds on the equilibrium path conditional on selection—L's punishment contract, α_L, can be constructed as in Proposition 4, assuming L's continuation value is v_L^*. Note that because the initial punishment strategies $y_R^{\alpha_L}$ and $s_0^{\alpha_L}$ are only determined by α_P, they are independent of b, too. Moreover, P's (DEC) holds with equality off-the-equilibrium path.

Now, assume $b = \hat{b}$, then, using the construction from Proposition 5, P's (DEC) binds whenever R has led last. Through the back-to-business property, that implies that a deviation by L in fact triggers a punishment phase that is identical to the on-path continuation contract when R led last. However, the principal's continuation payoff when L led last, is strictly above α_P.

So with α_P known, $\alpha_R = \alpha^C$, we can solve the following system of equations for the variables of α_L, θ_L, θ_R, s_R:

$$\begin{aligned}
v_P^*(L) &= -(\theta - \theta_L) + \beta\left((1/2 + m)v_P^*(L) + (1/2 - m)v_P^*(R)\right) \\
v_P^*(L) &= -(\theta_R - \theta) + \beta\left((1/2 + s_R)v_P^*(R) + (1/2 - s_R)v_P^*(L)\right) \\
v_L^*(L) &= b - y_L^* + \beta\left((1/2 + m)v_L^*(L) + (1/2 - m)v_L^*(R)\right) \\
v_L^*(R) &= -y_R^* + \beta\left((1/2 - s_R)v_L^*(L) + (1/2 + s_R)v_L^*(R)\right) \\
v_R^*(R) &= b - (1 - y_R^*) + \beta\left((1/2 + s_R)v_R^*(R) + (1/2 - s_R)v_R^*(L)\right) \\
v_R^*(L) &= -(1 - y_L^*) + \beta\left((1/2 + m)v_R^*(R) + (1/2 - m)v_R^*(L)\right) \\
v_L^*(L) &= b + \beta\alpha_L \\
v_R^*(R) &= b + \beta\alpha^C \\
\alpha_P &= (1/2 + s_R)v_P^*(R) + (1/2 - s_R)v_P^*(L) \\
\alpha_L &= (1/2 + s_R)v_L^*(R) + (1/2 - s_R)v_L^*(L)
\end{aligned} \tag{12}$$

Then for $b \leq \hat{b}$, we can use the first 6 Bellman equations to back out the 6 value functions $v_k^*(r)$, then we can plug into 4 remaining dynamic enforcement constraints (3 on-path, 1 off path) to determine a unique solution to our 4 variables of interest.

A solution within their domain exists if b is above a threshold, $\check{b}$.

Finally, as $b \to \check{b}$ choices converge to $s_R = -m, y_L = 0, y_R = 1$ and we arrive at a repeated static Nash equilibrium which completes the proof. □

References

Abreu, D. (1986). "Extremal equilibria of oligopolistic supergames". *JET*, pp. 191–225.

— (1988). "On the theory of infinitely repeated games with discounting". *Econometrica*, pp. 383–396.

Acemoglu, D., M. Golosov, and A. Tsyvinski (2011). "Power fluctuations and political economy". *JET*, pp. 1009–1041.

Acharya, A., E. Lipnowski, and J. Ramos (2024). "Political Accountability Under Moral Hazard". *AJPS*, forthcoming.

Akerlof, G. A. and R. E. Kranton (2005). "Identity and the Economics of Organizations". *JEP*, pp. 9–32.

Andrews, I. and D. Barron (2016). “The allocation of future business: Dynamic relational contracts with multiple agents”. *AER*, pp. 2742–2759.

Anesi, V. and P. Buisseret (2022). “Making elections work: Accountability with selection and control”. *AEJ:Micro*, pp. 616–44.

Baron, D. P. and D. Diermeier (2001). “Elections, governments, and parliaments in proportional representation systems”. *QJE*, pp. 933–967.

Barron, D. and M. Powell (2019). “Policies in relational contracts”. *AEJ:Micro*, pp. 228–49.

Bendor, J., A. Glazer, and T. Hammond (2001). “Theories of delegation”. *Annual review of political science*, pp. 235–269.

Bendor, J. and A. Meirowitz (2004). “Spatial models of delegation”. *APSR*, pp. 293–310.

Board, S. (2011). “Relational contracts and the value of loyalty”. *AER*, pp. 3349–3367.

Bolton, P. and M. Dewatripont (2013). “Authority in organizations”. *Handbook of organizational Economics*, pp. 342–372.

Callander, S. et al. (2008). “A theory of policy expertise”. *QJPS*, pp. 123–140.

Callander, S., D. Foarta, and T. Sugaya (2022). “Market Competition and Political Influence: An Integrated Approach”. *Econometrica*, pp. 2723–2753.

Calzolari, G. and G. Spagnolo (2020). “Relational contracts, procurement competition, and supplier collusion”. *mimeo*.

Delgado-Vega, Á. (2023). “Which Side are You on? Interest Groups and Relational Contracts”. *mimeo*.

Dixit, A., G. M. Grossman, and F. Gul (2000). “The dynamics of political compromise”. *JPE*, pp. 531–568.

Eliaz, K. and R. Spiegler (2020). “A model of competing narratives”. *AER*, pp. 3786–3816.

Gehlbach, S. (2021). *Formal models of domestic politics*. Cambridge University Press.

Green, E. J. and R. H. Porter (1984). “Noncooperative collusion under imperfect price information”. *Econometrica*, pp. 87–100.

Holmstörm, B. (1999). “Managerial Incentive Problems: A Dynamic Perspective”. *ReStud*, pp. 169–182.

Invernizzi, G. M. and M. M. Ting (2023). “Institutions and Political Restraint”. *AJPS*.

Levin, J. (2002). “Multilateral contracting and the employment relationship”. *QJE*, pp. 1075–1103.

— (2003). “Relational incentive contracts”. *AER*, pp. 835–857.

Li, J., N. Matouschek, and M. Powell (2017). “Power dynamics in organizations”. *AEJ:Micro*, pp. 217–241.

Lipnowski, E. and J. Ramos (2020). “Repeated delegation”. *JET*, p. 105040.

MacLeod, W. B. (2014). “The Economics of Relational Contracts”.

Mailath, G. and L. Samuelson (2006). *Repeated Games and Reputations: long-run relationships*. Oxford university press.

Malcomson, J. M. (2012). “Relational Incentive Contracts”. *The Handbook of Organizational Economics*. Princeton University Press, pp. 1014–1065.

McClellan, A. and X. Liao (2023). “Intraparty Politics and Dynamic Policy Polarization”. *mimeo*.

Nocke, V. and R. Strausz (2023). “Collective Brand Reputation”. *JPE*, pp. 1–58.

Padró i Miquel, G. and P. Yared (2012). "The political economy of indirect control". *QJE*, pp. 947–1015.

Rayo, L. (2007). "Relational incentives and moral hazard in teams". *ReStud*, pp. 937–963.

Rogoff, K. (1985). "The optimal degree of commitment to an intermediate monetary target". *QJE*, pp. 1169–1189.

Watson, J. (2021). "Theoretical foundations of relational incentive contracts". *Annual Review of Economics*, pp. 631–659.

Wright, J. (2002). *Gustav Stresemann: Weimar's greatest statesman*. Oxford University Press, USA.

Xuan, Y. (2009). "Empire-Building or Bridge-Building? Evidence from New CEOs' Internal Capital Allocation Decisions". *RFS*, pp. 4919–4948.

Yared, P. (2010). "Politicians, taxes and debt". *ReStud*, pp. 806–840.

Online Appendix

G The Principal's Punishment

Lemma 14. *There is a principal-optimal punishment in which $y_R = 1$ whenever R leads.*

Proof. A principal's optimal punishment exists because the equilibrium payoff set is compact. Now, take a principal's optimal punishment $C$and assume it give the principal an ex-ante payoff α_P. Now we construct a contract C' with $y_R(\cdot) = 1$ with an ex-ante value of at most α_P to the principal.

We prove this by showing to claims hold.

Claim 1. *Suppose there is a principal's optimal punishment C such that for any history $h \neq h_0$, the principal's dynamic enforcement constraint has slack, i.e., $w_P(h) > \alpha_P$ and for some history $h' \supset h$, $y_R(h') < 1$. Then there exists another contract $\widetilde{C}$ that makes the principal no better off, but has $\tilde{y}_R(h') > y_R(h')$.*

Proof. Let $\hat{h}$ be the first history with property $y_R(\hat{h}) < 1$ (see Definition 1 for a definition.)

Now, consider a contract $\widetilde{C}$ that is identical to C but with $\tilde{y}_R(\hat{h}) = y_R(\hat{h}) + \varepsilon$, with $\varepsilon > 0$. We show now that such a contract exists.

Note that at $\hat{h}$ R's (DEC) holds by construction. Also, the contract is unchanged for any history $h \not\subset \hat{h}$. Lastly, for any history $h \subseteq \hat{h}$, we consider all possible cases:

(a) any history $h \subseteq \hat{h}$ in which only R leads,
(b) any history $h \subseteq \hat{h}$ in which L leads his (DEC) has slack,
(c) any history $h \subseteq \hat{h}$ in which L leads and his (DEC) binds.

Case (a). At any $h \subset \hat{h}$, $y_R(h) = 1$ and thus his (DEC) holds. Because P's enforcement constraint has slack and values are continuous, a $\varepsilon > 0$ exists such that P's (DEC') holds at all histories $h \subset \hat{h}$. However, P's continuation value is

$$w_P(h_0|\widetilde{C}) = w_P(h_0|C) - \delta(\hat{h}|h_0)\varepsilon, \tag{13}$$

which is strictly lower than that under contract C. Thus C is not a optimal punishment.

Case (b). R's (DEC) holds like in Case (a). P's (DEC') holds for ε small enough too. Finally, L's (DEC) holds for ε small too, because we assume it to have slack initially. Thus a contract $\widetilde{C}$ iexists and (13) holds, implying that C is not a optimal punishment.

Case (c). In this case, $\widetilde{C}$ is not a contract because L's (DEC) is violated at least in one history. Let $\check{h} \subset \hat{h}$ be the largest such history. Now, consider another alternative contract $\widetilde{C}'$ identical to $\widetilde{C}$ but with $\tilde{y}'_L(\check{h}) = y_L(\check{h}) - \delta(\hat{h}|\check{h})\varepsilon$.

Contract $\widetilde{C}'$ satisfies L's (DEC) at $\check{h}$. Prior to $\check{h}$, $\widetilde{C}'$ yields the same continuation payoffs (for all players) as the original contract C. After $\check{h}$, the reasoning for case (b) applies, so all player's (DEC) are satisfied.

If $y_L(\check{h}) \leq \theta$, the principal's ex-ante value is:

$$w_P(h_0|\widetilde{C}') = w_P(h_0|C) - 2\delta(\hat{h}|h_0, C)\varepsilon,$$

so the principal is worse off under $\widetilde{C}'$. Yet, because the principal's preferences had been strict in favor C, there is an $\varepsilon > 0$ small enough such that $\widetilde{C}'$ is a contractm and C is not an optimal punishment for P. If $y_L(\check{h}) > \theta$, the principal's value is:

$$w_P(h_0|\widetilde{C}') = w_P(h_0|C),$$

making $\widetilde{C}'$ an optimal punishment for P.

Thus, it is possible to increase y_R in the first history in which $y_R < 1$ leaving P weakly worse off. Iterating over this argument proves the claim. □

Claim 1 shows that the agents (DEC) can be ensured by constructing an appropriate contract C'. It ignores, however, that these constructions may violate P's dynamic enforcement constraint. The reason is that in our construction of $\widetilde{C}$ and $\widetilde{C}'$ we have made some branches of the game tree that are played with positive probability strictly worse for P. Our next claim shows that if contract C was indeed the optimal punishment for P, then our operations do not violate P's enforcement constraint.

Claim 2. *If C is a principal's optimal punishment, then there exists a contract $\widetilde{C}$ as constructed in the proof of Claim 1 such that $\tilde{y}_R(\hat{h}) = y_R(\hat{h}) + \varepsilon = 1$.*

Proof. Take a contract C and a first on-path history $\hat{h}$ under that contract, such that $y_R(\hat{h}) < 1$. Assume P's continuation value under C is $w_P(\breve{h}) = \alpha_P$ for some history $\breve{h} \neq h_0$. Finally, fix $\hat{\varepsilon} = 1 - y_R(\hat{h})$.

First, note that we can construct a contract $\tilde{C}$ as in the proof of Claim 1 with $\varepsilon = \hat{\varepsilon}$ such that the agent's (DEC) hold under $\widetilde{C}$ by combining the steps in cases (b) and (c) from that proof. Next, observe that if any potential choice of $\breve{h}$ implies $\breve{h} \not\subset \hat{h}$, contract $\widetilde{C}$ is incentive compatible by construction. Thus, what remains is to consider histories $\breve{h} \subset \hat{h}$.

Now, assume that $\breve{h} \subset \hat{h}$ under C and, in addition, that the constructed $\widetilde{C}$ under the required ε violates P's (DEC') at $\breve{h}$. That means P's continuation value $w_P(\breve{h}, \widetilde{C}) < \alpha_P$ where α_P is P's ex-ante value of C. If there are multiple such histories, take a largest one where this is the case. By construction $\widetilde{C}$ is incentive compatible for any history $h \supset \breve{h}$. But then we could replace C by a contract $\breve{C}$ in which players play starting from h_0 as if they had started at history $\breve{h}$. This contract is incentive compatible and by assumption worse than C. Thus, C cannot be P's optimal punishment to begin with proving claim 2. □

Iterating forward over the constructions in the proofs of Claim 1 and 2, either the candidate contract C is not a optimal punishment or an equivalent contract exists with $\tilde{y}(\hat{h} \cup R) = 1$ for all histories. Combined with existence, that proofs the Lemma. □

G.1 Lemma 15

Lemma 15. *The principal's optimal punishment is stationary and characterized as follows:*

- *the principal endorses L, $s = -m$,*
- *R fully polarizes $y_R = 1$*
- *L either fully polarizes, $y_L = 0$, or chooses $y_L = \hat{y}_L$, where $\hat{y}_L$ is such that his dynamic enforcement constraint binds.*

Moreover, for any (β, m, b), there is a cutoff θ^P, such that both agents polarize if and only if $\theta > \theta^P$.

Proof. From Lemma 14, we have $y_R = 1$. Hence, irrespective of $y_L(\cdot)$, P is worse off if R leads because $\theta < 1/2$. Thus, P chooses $s = -m$.

To conclude, we characterize L's actions. Fixing $(s = -m, y_R = 1)$, let $\hat{y}_L$ be L's policy such that (DEC) binds. Note further, that if certain $\hat{y}_L$ is incentive compatible for some endorsement strategy s, it is incentive compatible for $s = -m$. It is straightforward to see that L's action is either $y_L = 0$, which is trivially incentive compatible, or, $y_L = \hat{y}_L$, as P prefers any intermediate $y_L \in (0, \hat{y}_L)$ to at least one these two extremes. If $\hat{y}_L < 2\theta$, $y_L = 0$ is worse for P, and otherwise, $y_L = \hat{y}_L$ is worse. $\square$