Embracing the Enemy*

Álvaro Delgado-Vega† Johannes Schneider‡

October 2023

— Preliminary & Incomplete—

Abstract

Two agents repeatedly compete on who can take action. A principal can partially influence this competition. Players differ in their preferences over the ideal action, and the principal is more aligned with one of the agents. We characterize the optimal contract with and without principal commitment and show that the principal provides more support to the distant than to the close agent. In doing this, a moderately biased principal outperforms an extreme and a balanced one. If interactions are frequent, all but extreme principals prefer to have two agents, even if one is very distant.

Our model applies to various situations ranging from corporate finance to political economy. It provides, e.g., a mechanism that helps explain the left turn of Italian Christian Democrats and a rationale for why managers with a mission may act as bridge-builders between divisions rather than empire-builders favoring only their home division.

1 Introduction

Groups in an organization compete about the *who* and the *how*. How should things be done? Who will be in charge of doing it? Political parties are motivated by the trappings of office but also genuinely care about governmental action. Divisions in a corporation compete to take the lead in the next project but also aim to imprint their corporate culture into the whole firm's operation. More often than not, these conflicting preferences about the *how*

*We are grateful to Antoine Loeper's early input, which very much shaped the direction of this paper. We thank Takuo Sugaya, Dana Foarta, and audiences in Bern, Konstanz, Madrid, Mannheim, and St.Gallen, at SITE 2023, and at SAET 2023 for their helpful comments and suggestions.

†ETH Zurich & University of Chicago, email: alvaro.delgado@mtec.ethz.ch

‡Universidad Carlos III de Madrid, email:jschneid@eco.uc3m.es

origin in deep worldviews making an individual *intrinsically motivated*, *mission-driven*, or *ideological* depending on context and connotation.

In the competition for the *who*, third parties (lobbies, mid-level managers, coalition partners in cabinet) can sometimes sway the outcome of the power struggle by supporting a particular agent's candidacy. These third parties have their own vested interests about the *how* and would try to use their influence to affect it. What strategy would they employ? Should they support a close ally to increase the likelihood of their preferred outcome, or should they back an adversary to incentivize his moderation? Is there a difference between their short-run and long-run incentives?

To address these questions, consider a setting where two agents compete over the right to take an action. For example, the agents could represent political parties vying for access to executive power or divisions within a firm competing to take on a prestigious project. It is known that our two agents hold opposing views on what is the right action to take. These views are deeply rooted and may represent a leftist or rightist agenda in the case of political parties or fundamentally different views on the road to success in project management. Two crucial features here are that we assume that the agents care about the action taken even when the other agent is taking it and that they attach an intrinsic value to be given the right to take the action.

Now, we add a third player to the model, the principal. The principal, too, holds preferences over the action taken. Her preference is closer to one of the agents (her "friend") than to the other (her "enemy"). The principal is incapable of taking the action herself. Moreover, while she has some power in influencing who gets the right to take the action, other exogenous forces also influence that decision.

If the principal were short-lived, her choice would be trivial. Whenever she can, she would use her influence to increase the likelihood that her friend can take action. The agent who is then given the right to act (the "leading agent" henceforth) has no incentives to take an action other than his preferred one.

However, agents in many real-world organizations regularly compete to lead the next operation, often holding diverse worldviews. Most mature democracies evolve around ideological parties who regularly compete for government; many cooperations contain divisions with distinct tribal identities that over time compete for projects and influence; and many standing committees are comprised to represent stakeholders with different missions who compete for shaping the committee's overall agenda.

If the situation of support and decision-making occurs multiple times throughout the relationship between the principal and her agents, there is a second component to consider. Suppose the principal can offer the leading agent the following contract: If you pander toward my preferred action today, I will use my power to endorse you to take action the

next time. The promise of the principal's support may incentivize the agent to take up the contract and pander toward the principal's preferred action. While in the settings we have in mind, such contracts need to be self-enforcing for the agents. We consider both settings in which the principal can credibly commit to her future behavior, and those in which the contracts are fully relational on all sides.

We show that any such contracts are particularly effective in disciplining the enemy. The result becomes particularly stark when the principal's preferences align fully with the friend's. The principal, by design, cannot guarantee the friend the right to take action. Instead, she aims to "embrace the enemy" by fully supporting him if a leading enemy panders considerably to the principal's preferred action. If the principal is more centrist, she also needs to incentivize the friend to make him pander, which she can do in two ways: either by promising to use her influence to moderate a leading enemy or by promising support to the friend. It turns out that the principal provides no support to the friend if maximum moderation of the enemy (given the principal's power) is sufficient to ensure that the friend panders fully toward the principal's preferred action whenever leading.

Two observations follow directly from our characterization. The first relates to competition. Does the principal prefer the competitive situation we have just described or a "dictatorship" in which she only is in a relationship with the friend? It turns out that, in general, the principal prefers the dictatorship only if her preferences are closely aligned with those of the friend. If, in addition, the interactions are sufficiently frequent, all principal types but those with exactly aligned preferences strictly prefer an environment with a friend and an enemy over dictatorship. Second, a principal benefits most from interacting with the two agents if her own preferences determine that there is a friend and an enemy, but only moderately so. The principal is not too close to her friend in preferences either.

The first observation implies that a principal in a relationship with an agent with whom she closely but not perfectly overlaps is willing to extend the setting to a three-player relationship, even if it means bringing in an enemy. However, given her limited control, she cannot prevent the enemy from acting at times. Introducing the enemy has two opposing effects on the principal's prospects. On the one hand, the principal gains the ability to pose a credible threat to her friend. That helps to make the friend pander to the principal. On the other hand, the enemy will take action sometimes. If interactions are sufficiently frequent, however, promising support to the enemy makes him pander enough so that the first effect is dominant.

The second observation implies that if there is self-selection into becoming a principal by paying some fixed cost, we expect, on average, to see more principals with a moderate bias than those at the extremes or in the center. The intuition behind the nonmonotonicity in the principal's expected payoff is somewhat subtle: As the principal moves in from being

extreme, she suffers less from the enemy making decisions while she remains able to provide incentives to the friend to pander toward her own preferred action fully. However, there is a maximum amount of pandering that she can incentivize. A centrist principal lacks the power to incentivize any side to play the principal's preferred action. In contrast, a moderately biased principal can convince at least the friend to choose that action. She earns a strictly higher payoff than a centrist principal. These results hint at a force in leadership selection: Leaders gain the most if they are not extreme but not fully balanced either.

Conceptually, we depart from the classical repeated moral hazard settings in five different ways. First, we consider two agents. This departure implies that we need to 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 her desire to channel her effort to push her own personal agenda. Second, we consider a principal with severely restricted power—she can neither terminate the relationship nor limit the action space. Even the selection of the leading agent which she can influence also depends on exogenous forces. Third, we assume externalities across agents. Fourth, we assume an exogenous value to become a leader. Our departures two to four have little relevance in single-agent settings but are crucial with multiple agents. Combined, they imply that each player's ability to punish in response to a deviation is limited by the fact that their payoffs inevitably depend on future decisions by the player they are aiming to punish. Fifth, our players are unable to directly transfer utility, which further strengthens our previous point. Any attempt by 2 players to reward or punish each other for past behavior has to occur through an action that—because of the externality—affects the incentives of the third player to react when it is her turn. Combined, these departures lead to an interaction between on-path and off-path behavior which both enlarges the contracting space in some dimensions yet constrains it in others.

To illustrate that point, we consider cases in which the principal has commitment power ex-ante and cases in which she only offers a relational contract. The distinction is important for two reasons. First, a committed principal can credibly promise the enemy large utilities by supporting him in future periods, even if the principal may later be tempted to renege on that promise and favor her friend instead. That allows her to ask for larger concessions and thus for overall better outcomes. Yet, this aspect loses importance if agents can collude to consistently take actions far from the principal's preferred action to punish a deviating principal.

Second, a commited principal can vouch to punish a deviating friend. Absent commitment, the friend understands that the principal has limited power to punish him because he genuinely aligns more with the friend than the enemy. Knowing the principal's bias toward him, the friend is only willing to make small concessions. Perhaps counterintuitively at first sight, the ability of the principal to punish the friend and thus extract a large concessions

depends on the friends ability to punish the principal. The greater the latter, the greater the former.

Although the lack of commitment reduces the principal's ability to ask for concessions, whether and how it affects the optimal contract is a priori unclear and requires careful analysis. Exploiting the fact that the agent's joint surplus in our model remains constant, we provide a behavioral characterization of the optimal self-enforcing contract and show that it shares the same qualitative features as the enforceable contract.

Our paper contributes to various discussions involving players with limited power and long-term horizons. How do political parties choose their coalition partners, knowing new coalition governments will be formed in the following years? How do mid-level managers allocate tasks and resources when more new tasks will come in the future? How do lobby groups allocate their electoral influence when strategizing their relationships with politicians across different elections? Each question is the center of a strand of the literature. We contribute by providing an additional, clear-cut insight into these strands: There is a rationale for seeking partners among those most distant from your positions. In discussing our analysis, we delve into two particular applications of our model. Regarding government coalition formation, we discuss the calculus of the Italian Christian Democrats for opening to coalition agreements with the Socialists in the 1960s, a crucial moment in postwar Italian politics. Regarding the internal workings of firms, we turn to the capital allocation choices of CEOs in multidivisional firms. The empirical literature has found a pattern of reverse favoritism in capital allocation, particularly pronounced when the CEO has less authority with respect to each of the divisions (Xuan, 2009). In both cases, we emphasize the crucial premise of our theoretical framework: the principal has limited power—she cannot exclude an agent entirely from taking action. This assumption applies to politics, where political parties represent constituencies whose participation rights are set by basic democratic institutions. It similarly applies to many organizations in which the power to allocate tasks or resources is separated from human resource management.

1.1 Contribution to the Literature

We study the optimal relational contract in a setting where a principal with limited power aims to govern multiple agents whose actions cause externalities on all players. The principal's key tradeoff in our paper is whether to delegate decision rights more often to the most aligned agent or to use them to discipline the less aligned agent. Although the literature has recognized all our model ingredients in isolation, we show that combining them generates new insights and exciting applications.

Although the literature on relational contracts is by now enormous and ever-growing (see, e.g., the surveys of Malcomson, 2012; MacLeod, 2014; Watson, 2021), only few papers

consider relationships with multiple 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). While these papers consider a diverse set of problems, one feature is common: deviating agents can be punished by exclusion. In our setting, such exclusion is ruled out. Because the principal's power is limited, she cannot credibly commit to excluding the agent from the game. Hence, punishments themselves are non-trivial relational contracts, and the principal's temptation to renege on herself is relevant both on and off the equilibrium path. Moreover, apart from Board (2011), Barron and Powell (2019), and Calzolari and Spagnolo (2020), all of these papers abstract from direct competition between agents.

In addition to the competition on being selected also present in Board (2011), Barron and Powell (2019), and Calzolari and Spagnolo (2020), *every agent* in our model also holds preferences about the action chosen by the *selected agent*. Unlike all three papers, we thus aim to capture relationships between long-run players who are part of the same organization and, therefore, benefit and suffer from the actions taken by their fellow members too.

Another strand of the literature has explored the allocation of power within organizations (see Bolton and Dewatripont, 2013, for a survey). While large parts of the literature consider short-run players, we consider, like, e.g., also Li, Matouschek, and Powell (2017) and Lipnowski and Ramos (2020), a setting of repeated delegation decisions. Different from us, those papers assume an informational friction. Our assumption is more basic. The principal simply cannot control the delegation process fully because external forces (voters, higher-up management, macroeconomic factors, ...) may overrule her decision with some probability.

Combined with the multi-agent setting, assuming that the principal has limited the delegation power allows us to talk about power dynamics despite having a model of perfect monitoring—a feature shared with Delgado-Vega (2023)'s rent-seeking model. The key idea in Delgado Vega (2022) is that the principal tries to divert assets from the agents' common pool into the principal's private pocket. No such common pool exists in our model. Instead, the agents' game is to divide a given pie between them, and the principal tries to influence that division. This difference provides a fundamentally different economic model. The optimal contract maximizes total welfare in our model while it minimizes it in Delgado Vega (2022); Concessions by one agent benefit the other agent in our model while they harm him in Delgado Vega (2022); whether the principal can commit has distinct effects on behaviour on and off the equilibrium path while commitment is no issue in Delgado Vega (2022).[1]

In the model with commitment, the optimal off-path behavior is asymmetric. While the

[1]Moreover, our strong non-transferable utility assumption precludes the front-loading of payments common in the contracting literature (see, e.g., Levin, 2003, for the basic argument). Frontloading is done in Delgado-Vega (2023).

enemy and the principal are best punished using a traditional grim 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 sees a carrot of returning to the on-path continuation game. There is a subtle difference, however, to Abreu's seminal stick-and-carrot idea. In Abreu (1986) the stick ensures a fixed and given period of low payoffs to *everyone*. The length of the stick in our case heavily depends on nature's choices resembling some of the ideas from Green and Porter (1984). Moreover, while R benefits from L's deviation as long as the stick is still ongoing, P (similar to L) suffers until the carrot is reached.

Long-run interaction between players in the political arena is an area of increasing interest in political economy (Padró i Miquel, 2007; Acemoglu and Robinson, 2008; Yared, 2010; Anesi and Buisseret, 2022, see, e.g.,). In particular, our paper studies situations in which a player influences the access to power of another, for example, because it is a party forming a coalition government or a lobby allocating its electoral support. The closest papers in that literature are Callander, Foarta, and Sugaya (2022) and Delgado-Vega (2023). In the language of our paper, the main conflict of interest between the principal and the agents in both these papers is how to distribute surplus between the principal and the agents. We contrast this setting in assuming that the agents' joint surplus is constant throughout and the principal aims to moderate their choices.

One result of our paper is that a principal reduces polarization between agents. Dixit, Grossman, and Gul (2000), Acemoglu, Golosov, and Tsyvinski (2011), Invernizzi and Ting (2023), and McClellan and Liao (2023) discuss forces to moderate political parties in models without a principal. Their main incentive to cooperate comes from the assumption that the parties have convex loss functions as they move away from their bliss point. Hence, moving policy slightly harms the incumbent little, while it benefits the opposition a lot providing a direct incentive for favour trading over time. Our model eliminates this channel by focusing on a system in which agents' joint surplus is constant. Instead, we concentrate on the principal's ability to promise support as a necessary and effective force of moderation

Lastly, our paper contributes to the literature on coalition government formation in parliament. At the cost of radically simplifying the stage game, we are able to study the dynamics of coalitions when parties have long-term horizons. Our main contribution is to delve into the delegation problem between coalition partners. In previous works like Baron and Diermeier (2001) office rents and policy could be exchanged freely. On the contrary, we emphasize that giving office (e.g., in the form of cabinet positions or seats in the supervisory boards of public entities) often entails transferring control rights over certain policies, allowing for what is known as ministerial drift. Hence, in our setting, both office

rent and control over policy go together, and therefore, the problem of keeping promises between coalition partners gains relevance.

2 Model

Time is discrete and infinite, indexed by $t = 1, 2, ...$ There are three players, one Principal P, and two Agents, L and R. All players discount the future exponentially through a common discount factor $\beta \in (0, 1)$. Every action is publicly observable, and the following stage game is played in each period.

Stage game. At the beginning of each 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 $k_t \in \{L, R\}$. The outcome of the coin toss depends on s_t such that the probability that $k_t = R$ is

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

and $k_t = L$ realizes with the complementary probability. If the realization is k, we say agent k was selected. The selected agent k then chooses an action $y_t \in [0, 1]$. The stage game ends and all players collect their within-period payoffs.

Within-period payoffs. Let $u_{i,t}$ denote player i's stage payoff with $i \in \{P, L, R\}$ Then, given the selected agent k and his action y,

$$u_{i,t}(k, y) = -|y - y_i^*| + b\mathbb{1}_{k=i},$$

where y_i^* denotes player i's bliss point

$$y_L^* = 0 \qquad y_P^* = \theta \in [0, 1/2) \qquad y_R^* = 1,$$

and $b > 0$ is a rent only the selected agent k enjoys. Players maximize their expected discounted stream of payoffs at the current period.

Solution Concept. We assume that the principal proposes a *contract*, $C = (\mathbf{s}, \mathbf{y}_L, \mathbf{y}_R)$, before the game starts. A contract is a complete contingent plan of action for each player. The proposed contract is *incentive compatible*, if it is a subgame-perfect Nash equilibrium of the dynamic game outlined above. The *optimal contract*, C^*, is the principal's ex-ante preferred incentive compatible contract.

We consider two settings, those in which the principal can commit ex-ante to a strategy $\mathbf{s}$ (binding contracts), and those in which she cannot (relational contracts). Agents cannot commit to an action in either scenario and optimize dynamically.

2.1 Discussion

Before turning to the analysis, we briefly discuss the key features of our model.

Externalities. The first defining feature of our model is that the selected agent's choice y_k affects the within-period utility of *all players* in every period. This is distinct from classical delegation problems, in which an agent consumes an outside option whenever she is not selected as the principal's delegate. In addition to classical moral hazard models, both the principal and the agent are motivated and hold stakes in a common institution. However, there is disagreement about the horizontal dimension of *how* actions should be taken.

Linearity. We assume (piecewise) linear payoffs over the action space for all players. This assumption is at odds with most of the literature on partisan candidate models in political economy, where parties have concave preferences around their bliss point (e.g., Dixit, Grossman, and Gul, 2000; Acemoglu, Golosov, and Tsyvinski, 2011; McClellan and Liao, 2023). As will become clear later, adding concave utility to our model would only amplify our results without affecting them qualitatively. It would, however, contaminate clarity.

Agent Competition. In addition to the horizontal disagreement, agents (but not the principal) differ ceteris paribus in their preferences about *who* is selected to take action. Our symmetry assumption implies that agents compete over a constant utility b for being selected. In addition, they compete over a constant utility -1 distributed among them via the action y. Formally, ignoring the principal, we thus have a constant sum game between agents because for every t

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

An immediate consequence of the constant-sum assumption is that the optimal contract is maximizes utilitarian welfare and is thus Pareto efficient.

Principal's Power. Contrary to standard delegation problems, we assume a principal with limited power. The principal can influence the selection process but cannot fully control it. This assumption is at the core of our model. It guarantees that all agents, at all points, expect to be selected within finite time under any principal strategy. Thus, even a principal with commitment power cannot preclude an agent from being selected.

Biased Coins. In our baseline model, we assume that the principal's decision is to bias a coin through endorsement. After endorsement, the coin is tossed and determines which agent gets selected. While this choice is the cleanest in terms of the model description, the following isomorphic alternative may fit some applications better.

Suppose nature determines at the beginning of a period whether the principal gets to select the agent (which happens with probability $2m$) or whether a fair coin determines the selected agent (which happens with probability $1 - 2m$).

3 A Simple Example

To shape the general intuition, we analyze a particularly simple special case: A setting in which a committed principal aligns with agent L: $y_L^* = y_P^* = \theta = 0$.

A Candidate Contract. A naive candidate contract in this setting would be for the principal to always endorse her "friend" L, i.e., $s = -m$. The probability of selecting L is maximized and once selected, L chooses the principal's bliss point $y = 0$. However, whenever R is selected, R has no incentives to moderate by choosing any policy other than her bliss point $y = 1$. Thus, the principal enjoys in the majority of cases her preferred action $y = 0$, and suffers from her worst action $y = 1$ in the remainder of cases. The sketched contract is trivially incentive compatible. Every agent chooses his bliss point when selected. However, we may ask: Is this contract optimal?

The Optimal Contract. It turns out that a contract that unconditionally endorses the friend is not optimal for any $m < 1/2$. Instead, (almost) the opposite is true. As the next proposition shows, the principal "embraces the enemy".

Proposition 1. *Suppose $\theta = 0$ and the principal has full commitment.*
Then, the optimal contract is such that on the equilibrium path,
(i) the principal fully endorses L until the first time R is selected,
(ii) thereafter the principal fully endorses R in all periods.
Whenever selected, L polarizes at $y_L = 0$, and R moderates at

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

Proposition 1 shows that a committed principal engages in a non-stationary strategy. She endorses her "friend" until the first time the "enemy" is selected. Then, she switches completely and remains endorsing the enemy regardless of who gets selected.

The principal's endorsement has two effects. It influences who gets selected and serves as a reward for the agent's behavior. The principal wants agent L to be selected, but endorsing L does not change L's behavior once selected: L already chooses the principal's ideal action. However, if R gets selected, the principal wants to discourage R from choosing the principal's worst action $y_R = 1$. So the principal can promise to support R in exchange for him to moderate his policies.[2]

Agents, in turn, profit in two ways from endorsement. First, they value being selected in itself, and second, selection allows them to choose that period's action. Now recall that the principal is, ceteris paribus, agnostic about who gets selected. At the optimal contract, the principal leverages her indifference about the *who*. By promising R to endorse him, the principal can ask for a large concession because R also expects to receive the selection benefit b more often. The principal is happy to distribute b to R in return for a favorable policy because she has no value for b herself—she only suffers in the policy dimension.

These considerations, however only become relevant once R is selected for the first time. The reason is simply that past endorsements are sunk when it comes to the currently selected agent's incentive. Thus, prior to R's first selection, only the distributional effect matters and the principal endorses her friend L. Once R is selected, the principal's main concern is to limit her losses from R's choices. By offering full endorsement until eternity, she can extract the largest concession from agent R.

The Role of Commitment. One may suspect that principal commitment is key in the previous exercise. After all, the principal's ability to promise endorsement the first time R is selected appears to be the driver behind the non-stationary strategy by the principal.

It turns out, that there is a cutoff value $\hat{m}$ such that whenever the principal is weak, $m < \hat{m}$, she cannot influence the agents' choices absent commitment. In equilibrium, she always endorses her "friend" L. Agents polarize by choosing their bliss points $y_L = 0, y_R = 1$. If, instead, the principal is strong, $m > \hat{m}$, she can implement the commitment contract.

Formally, the cutoff strength by the principal is given by

$$\hat{m} := \frac{1-\beta}{\beta}\left(\sqrt{\frac{b+1}{b}} - 1\right) - 1/2.$$

Proposition 2. *Suppose $\theta = 0$ and a principal without commitment power. If $m \geq \hat{m}$, Proposition 1 continues to hold. Otherwise, the optimal contract is a repetition of the unique static Nash-equilibrium $\{y_L, y_R, s\} = \{0, 1, -m\}$ in all periods.*

[2]Note that, as alluded to in Section 2, assuming concave loss functions instead of our linear ones would amplify this effect

If the principal is weak, she still wants to promise R support in return for R's concession. Yet, her limited influence when selecting the agent for the next period means that she cannot promise R reselection with high probability. In response, R is only willing to concede little. Facing a small concession, in turn, implies that the principal is tempted to break her promise and opt for endorsing her friend.

The lesson of Proposition 2 is straightforward: Whether the principal has commitment or not plays only a role when the principal is weak. The stronger the principal, the stronger her ability to moderate R. The more moderate R, the lower the temptation for the principal to renege on promising support to the enemy, R.

On a technical level, however, deriving the optimal contract and the associated behavioral strategies absent commitment is less straightforward. To pin down the principal's temptation to renege we need to determine the worst continuation game for the principal, which in turn depends on the worst continuation game for each agent.

The basic features of our model require a careful analysis here. The off-path continuation games that follow a deviation by either player need to be derived jointly because our model does not allow the replacement of a deviating agent or reversion to an exogenous outside option. Moreover, action choices always imply externalities on the other players and, thereby, on their willingness to collude in punishing a particular agent. Because direct utility transfers are not possible, deriving these punishments warrants careful analysis.

For example, take a principal with power $m = \hat{m}$. After a selected R has taken her on-path action y_R from Proposition 1, the principal is exactly indifferent between endorsing R or reneging by choosing to endorse L and entering a punishment phase.

What does her punishment look like? Once P has deviated, L and R consider P unreliable and aim to (rationally) collude on punishing the principal. Sustaining such collusion is challenging as L and R are in a constant sum competition over a pie of size $b - 1$ every period. Yet, the optimal punishment involves both of them choosing an action to the right of $\theta = 0$ to punish the principal. However, selecting $y_L > 0$ needs to be incentive compatible for agent L. Thus, L needs to fear a punishment when instead choosing $y_L = 0$. Thus, there needs to be a credible (for R and P) response to a deviation by L, for example, by reverting to the on-path contract, which entails less endorsement to L.

4 General Analysis

We now analyze the full model. We begin with some preliminary results. Then we characterize the optimal contract assuming the principal can commit before comparing it with the no-commitment solution.

4.1 Preliminaries

A first useful benchmark is to assume the principal has no power over the agents, that is, $m = 0$.

Proposition 3 (**Complete Polarization**). *If $m = 0$, the game has a unique equilibrium. Whenever agent k is selected, he chooses his bliss point $y_k = y_k^*$.*

The result follows immediately from the constant sum competition between agents, our symmetry assumption, and the linearity in payoffs. These modeling choices combined imply that any concession of the selected agent k would need to be repaid by a concession in the future which, because of discounting, is larger than the one agent k is providing today. Because agents dynamically optimize promising such a concession is only optimal if k promises an even larger one in return in the future, and so on. Thus, absent the principal, agents would fully polarize and expect to split the total payoff of $b - \frac{1}{1-\delta}$ in half.

Next, we bring back the principal by considering cases with $m > 0$ and turn to the agent's incentive problem given a certain strategy $\mathbf{s}$ by the principal. An agent's continuation value $v_k(C, h)$ is the expected sum of utility from the next period onward. At each point in time, it depends on (a) the chosen contract C and (b) the history h of past play.

Any implementable contract C must be incentive compatible. That is, each player, when given the chance, must prefer to take the action the contract prescribes over her best deviation. While such a deviation in large parts of the literature implies termination of the contract, this is not the case here. Instead, the response to a deviation is a non-trivial contract itself.

An agent, when selected, follows the action recommended by the optimal contract if and only if it is preferred to selecting the myopically optimal action y^* which provides a utility of 0 today expecting a punishment value of α_k from tomorrow onward.[3] Following the relational contracting literature, we refer to this constraint as agent k's *dynamic enforcement constraint*, which states that an action y is enforceable for a selected agent k under contract C at time t after history h_{t-1} if and only if

$$-|y - y_k^*| + \beta v_k(h_t, C) \geq \beta \alpha_k \tag{1}$$

where h_t is the history at the end of period t if the agent indeed chooses y_k. If (1) holds with equality, we say the dynamic enforcement constraint binds. We are now ready to state our next Lemma, which shows that we can restrict the set of contracts that are optimal for the principal.

[3] Despite the extensive form, regular simple penal codes á la Abreu, Pearce, and Stacchetti (1990) suffice in our setting.

Lemma 1. *It is without loss of generality to focus on contracts in which whenever an agent gets to choose y on the equilibrium path, he*

(i) either he chooses the principal's bliss point $y_{i,t} = \theta$,

(ii) or his dynamic enforcement constraint binds.

Moreover, the optimal on-path actions satisfy $y_{R,t} \geq \theta \geq y_{L,t}$ for any t.

The intuition for the first part of Lemma 1 follows from two particular aspects of the model. First, the principal's bliss point lays between the bliss points of the agents. Therefore, moving one agent's action further from his bliss point and closer to that of the principal is always beneficial for the other agent. Consider the case in which L is selected today. If the contract proposed by the principal promises that in the future, R will act less polarized, the continuation value of the contract increases for L, and hence the principal can ask him for greater concessions in the present. Therefore, there is a feedback effect by which greater moderation of R in the future allows the principal to incentivize greater moderation of L today and vice-versa. Besides, when the principal lacks commitment power, moderating R in the future has a second effect. It increases the principal's continuation value and hence, relaxes her dynamic enforcement constraint.

The second part of Lemma 1 follows because if either agent moderates beyond the principal's optimal point, then both the selected agent's and the principal's payoffs are hurt and they can—without any effect on future incentives—agree to select the principal's bliss point instead. Besides, due to the linearity of payoffs, the principal could at most compensate her loss with extra concessions from L in prior periods.

4.2 Commitment

Assume the principal has full commitment power. Hence, the principal can offer the agents a contract up front that specifies an endorsement $s(h_t)$ for any history of the game h_t.

To reduce the relevant contracting space further, we make use of the following Lemma.

Lemma 2. *Suppose there is an optimal contract such that $y_L = \theta$, and $y_R \neq \theta$. Then there is an optimal contract in which the principal either fully endorses R after the first time R is selected, or L's dynamic enforcement constraint holds with equality.*

The main reason behind Lemma 2 is the same as in the extreme case $\theta = 0$. If $y_L = \theta$, the principal has no interest in moving the action of L further. Thus, she uses her power to offer a favorable deal to R by promising her large parts of the available selection rent $b/(1-\delta)$ in return for some moderation. As a result, P has her first-best action chosen less often, yet greater moderation of R makes L more willing to uphold $y_L = \theta$. At the same time, getting selected less often tightens that willingness. Thus, if L's dynamic enforcement constraint holds with equality, the principal may not want to increase her endorsement of R.

The principal's optimal contract. Now we are ready to characterize the optimal enforceable contract. Let

$$\begin{aligned}\overline{\theta} &:= 2\frac{\beta}{1-\beta}\left(1-\beta\left(\frac{1}{2}+m\right)\right)(b+1)m,\\ \breve{\theta} &:= 1-\beta\left(\frac{1}{2}+m(1+2b)\right),\\ \underline{\theta} &:= 2\frac{\beta}{1-\beta}(b+1)m\beta\left(\frac{1}{2}+m\right).\end{aligned} \tag{2}$$

Definition 1. Let $\widehat{\mathcal{C}}$ be a class of contracts with the following on-path properties

1. the principal initially fully endorses L, $s^0 = 0m$, and continues to do so until the first time R is selected,
2. the principal fully endorses R whenever R was selected last, $s_R = m$,
3. the principal uses the same endorsement strategy s_L whenever L was selected last,
4. each agent plays a stationary strategy y_k with $y_L \leq \theta \leq y_R$.

The next proposition fully characterizes the optimal commitment contract, Figure 1 summarizes its findings.

Proposition 4 (Optimal Commitment Contract)**.** *Suppose the principal has commitment power. The optimal contract $C^* \in \widehat{\mathcal{C}}$. Moreover the following is true on the equilibrium path.*

- *Iff $\theta \geq \max\{1-\overline{\theta}, \breve{\theta}\}$, the principal attains the* first best *$y_k = \theta$.*
- *Iff $\theta \leq \overline{\theta}$ the principal the first best if L is selected, $y_L = \theta$,*
- *Iff $\theta \leq \underline{\theta}$, the principal* embraces the enemy *after R's first selection, $s_k = m$.*
- *Iff $\theta \geq \overline{\theta}$, the principal* opportunistically endorses *whoever was selected last $s_L = -m$, $s_R = m$,*
- *Iff $\theta \in (\underline{\theta}, \overline{\theta})$, then $s_L \in (-m, m)$ and decreasing in θ.*

The fact that the principal attains first-best if the vector (β, m, b) is large enough is not much of a surprise in light of traditional models. As $(\beta, m) \to (1, 1/2)$ the familiar folk-theorem logic applies for a principal who fully controls the selection process. If $b \to \infty$ the policy variable becomes irrelevant for agents compared to their benefit of being selected. Thus, they are willing to select any policy as long as it improves their selection opportunities.

The result thus becomes more interesting if interactions are not too frequent (translating to a low β), the principal's power is limited (translating to low m), and policy matters for the agents (translating to low b). In such a case L is willing to comply fully with the principal's demand as long as (i) the principal ensures that R does not polarize and (ii) the principal's bliss point is not too far away. As agents are symmetric, the same holds for R.

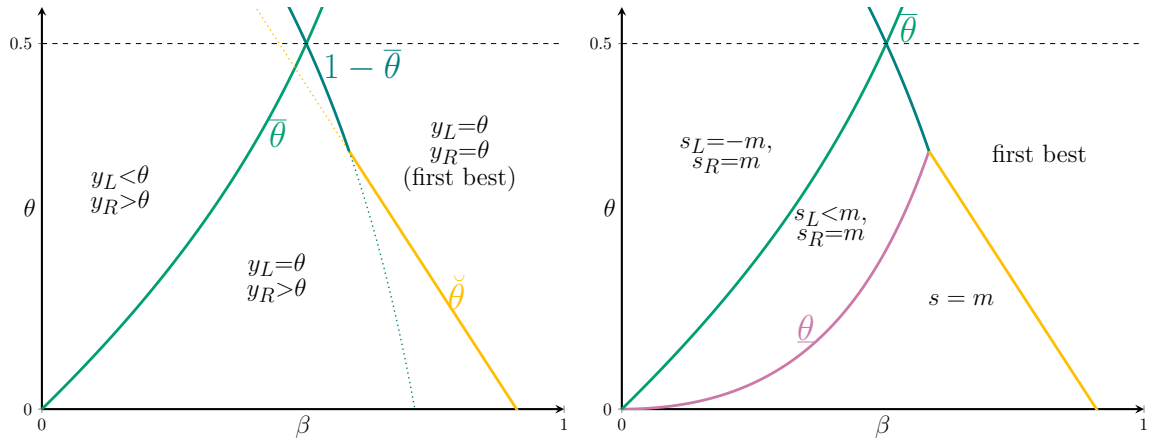

Figure 1: *Policy and Endorsement.* The left panel depicts the agents' policy choices for different values of θ and β. The right panel shows the optimal endorsement strategies for the different regions. In both cases $b = 2.5, m = 0.2$.

However, by construction, the principal's preference outside the first best region are too far away for R to fully comply. Instead, R panders partially in the direction of the principal's bliss point. If the principal is sufficiently centrist, she is unable to incentivize any agent to choose her bliss point. Instead, both agents pander partially toward the principal's first best.

To understand the principal's strategy, recall the case $\theta = 0$ again. Here, it is straightforward to see that L is willing to select the principal's optimum under any support strategy. The principal therefore uses her power to distribute b fully to make R pander toward the left. Note that the principal does not care about who gets selected per se. That means, that she is willing to provide support as long as the agent mitigates in return. Since there is no need to influence L who already selects the principal's optimum Lemma 2 applies and the principal fully endorses R.

As we increase θ slightly the same logic holds. By making R pander to the center, the principal provides enough incentives for L to select θ and no additional endorsement is needed. That changes eventually and after the principal is more centrist than $\underline{\theta}$, R's equilibrium pandering is not enough. Instead, the principal needs to provide incentives by endorsing L at least to some extent directly after L had been selected.

Once the principal becomes so centrist that even full support in these instances is not enough ($\theta > \overline{\theta}$), she acts opportunistically in the sense that she supports whoever has been selected last. One may wonder why she does not support L in periods after R had been selected. The answer here is that such support is ineffective. The joint per-period surplus of L and R is held constant at $b - 1$. In addition, if R is selected her utility is as low as possible, since P asks the highest concession it can get from R. Combining these two facts

implies that L's payoff conditional on not being selected is already maximized even under full support of R. Thus, only the support after a selection matters, which in turn is maximized at $s_L = -m$ for L.

Finally, in the initial periods, i.e., the periods until R is selected for the first time, there are no incentives to provide for R yet. Thus, the principal endorses whoever selects the best next-period action. Due to the principal's left-biased that is L.

4.3 No Commitment

We now consider the case in which the principal is unable to commit ex-ante to an endorsement strategy. Instead, she dynamically optimizes her endorsement at the beginning of every period.

Dropping commitment adds additional constraints to the problem. First, the principal must find it optimal to stay on the equilibrium-path. That on-path promise-keeping constraint becomes particularly relevant whenever the principal's on-path behavior requires her to provide some endorsement to agent R. Then, because L's preferences are more aligned with the principal's, she is tempted to renege on her promised support for R and fully endorse L instead.

To determine the principal's constraint, we need to solve for the worst relational contract for the principal. That contract will serve as the principal's punishment for reneging on her promises. There are two candidates for such a punishment. The first is the stage Nash equilibrium in which the agents polarize to actions 0 and 1 and the principal responds by fully endorsing L. However, by construction, the principal's punishment becomes relevant only if the on-path contract involves some endorsement of R. However, because agents play a constant-sum game, L prefers the principal's punishment strictly over the on-path contract. Taking these observations together, hints at a second possibility for the principal's worst contract. Agents L and R could collude in choosing actions far to the right of the principal's bliss point. The principle would still fully endorse L, which may incentivize L to take such an extreme action.

Whether this type of punishment is possible depends on the second additional contract we need to determine. Agent L's worst contract. Different from the commitment case, a grim trigger strategy of unconditionally supporting R in response to a deviation by L is not incentive compatible for the principal who would renege on her own promise to remain grim. Yet, how much the principal is willing to punish exactly, of course, also depends on how much she fears being punished herself, which thus needs to be determined jointly with that of agent L.

The optimal punishment of agent R remains at grim trigger as that is incentive compatible for all players.

As the next proposition shows, we can use that the joint surplus of L and R is constant and equal to $b-1$ in every period to narrow down the options to punish the principal to the two contracts sketched above.

Proposition 5. *The principal's worst contract is stationary with*

- *full endorsement of L in every period, $s=-m$,*
- *full polarization by R in every period $y_R=1$*
- *either full polarization by L, $y_L=0$ in every period, or $y_L=\hat{y}_L>2\theta$ in every period such that L's dynamic enforcement constraint holds with equality*

Proposition 5 reduces the problem to finding the principal's preferred contract among those that give L's the lowest continuation payoff. Due to the multidimensional constraint, however, solving for the optimum remains an involved fixed-point problem.

It is straightforward to solve for the two extreme cases $\theta=0$ and $\theta=1/2$. The former is done in Proposition 2 and relies on the fact that incentive compatibility for L is trivial. Once $\theta=1/2$ the principal favors neither side and can implement grim-trigger punishment.

For θ in a neighborhood to 0, L's temptation to deviate is small, however the principal's temptation to abandon her promises to R is large making punishments for L difficult. As we increase θ we are in a horse race between these two temptations which is non-trivial to characterize. Outside the polar cases, we have therefore so far not been able to formally characterize the optimal contract fully, however, our partial results admit the following conjecture.

Proposition 6. *For any (b,β,θ), there exists an $\hat{m}$ such that if $m>\hat{m}$ the optimal contract is qualitatively similar to the contract from Proposition 4. In particular, there are (parameter specific) cutoffs $0<\underline{\theta}<\overline{\theta}$ such that the principal does not fully endorse R when L has been selected last if $\theta>\underline{\theta}$ and $y_i\neq\theta$ when $\theta>\overline{\theta}$.*

Moreover, a deviation of L is followed by a period of intense punishment in which $s=m$ and y_R are further from θ than R's on-path action. However, eventually, continuation play in the punishment contract is identical to that on the equilibrium path.

The intuition is the following. If the optimal contract implies that L's dynamic enforcement constraint holds with equality, we know that the principal's overall preferred contract is such that it gives L the lowest possible continuation payoff *conditional on being selected*. That contract therefore solves our problem in that case. It is the best we can offer to the principal, while it is the worst we can offer to the deviating agent.

However, because the optimal contract also involved that R's dynamic enforcement constraint holds with equality *conditional on being selected*, we know by the constant-sum property, that L's continuation payoffs conditional on *not* being selected are maximized.

Punishment therefore reduces these payoffs by giving slack to R conditional on being selected right after the deviation. However, giving slack to R also reduces the principal's payoff, which implies that we can only give slack up to the point at which the principal's dynamic enforcement constraint holds with equality. To maximize the principal's continuation payoff we promise her the best possible continuation conditional on L returning to the helm which is the on-path contract. This comes at no cost, as it is also the conditionally worst contract for L. The missing step in our formal proof so far is to show that there is no contract in which we provide even larger utility to L after reelection to allow for harsher punishment in the initial stages when R is at the helm. While we are confident that this is not the case we have not yet been able to rule it out formally, hence leaving some of our results in this section at a conjecture stage.

4.4 Implications

(results in this section hold for the commitment case and for the no commitment case conditional on $m > \hat{m}$.)

Our characterization admits a set of results concerning the principal's position between the two parties. We begin with the principal's moderating effect. Absent a principal, players face a dictator game in which the dictator is chosen at random in each period. Since the agent's per-period joint payoff is fixed and the same across periods, the only equilibrium of the game is full polarization in policy choices (see Proposition 3). As we have seen, a fully biased principal is enough to moderate parties and move policy choices closer. The first implication from our characterization states that policy choices get closer to each other or agents appear less polarized, the more balanced their principal.

Proposition 7 (**Polarization**). *The distance between the agent's equilibrium actions, i.e., $y_R - y_L$ decreases as the principal's preferences become more balanced.*

The result is straightforward in light of our characterization. The more the principal moves to the center, the more she wants to move L to the center. Moving L more to the center pays a double dividend to the principal: not only does it directly affect her payoffs, but it relaxes the incentives needed for R to converge closer to the center also. The result follows.

Next, we consider the principal's ex-ante payoff as a function of her bliss point θ. A priori, it is not obvious whether an extreme principal ($\theta = 0$), who gets her first best for free when L is selected, is better or worse off than a balanced principal who needs to provide incentives to both agents toward moderating their decisions. It turns out that, indeed, both scenarios are possible depending on parameters.

Proposition 8 (**Principal's Payoff**). *If the first best is unattainable for any θ, then the principal's ex-ante payoff has a unique peak at $\theta = \overline{\theta} < 1/2$.*

For low levels of θ, the principal's ex-ante payoff increases with the principal's bliss point. Consider the principal's within-period payoff when each party is selected. Moving her bliss point marginally to the center does not change L's willingness to fully pander to the principal; hence, conditional on L being selected, the principal remains at her first best. As the principal moves closer to the center, she inevitably also moves closer to R, which mitigates her losses conditional on R being selected. In addition, since L moves to the center, R suffers less from the periods in which L is selected and in response, he is willing to move closer to the principal too. Hence, conditional on R being selected, the within-period payoff of the principal increases.

Agent R's moderation provides—if θ is low—sufficient incentives for L to choose the principal's bliss point θ, even if the principal fully endorses R. This is no longer possible when θ becomes greater. Then, shifting some of the principal's power to favor L implies a loss as R's incentives to move to the center decrease (which is partly offset by a more moderated L) and a gain, as the principal's favorite party is selected more often. Overall, her payoff still increases as long as she manages to keep L's action at her bliss point θ and hence, the principal's payoff attains its peak at $\bar{\theta}$.

When θ rises above $\bar{\theta}$, the principal is opportunistic and none of the agents selects her bliss point. The principal has moved too far away from L, and she can not even obtain her bliss point at the initial periods before R wins office for the first time. A balanced principal has neither an "enemy" nor a "friend", and that is particularly detrimental at the initial periods.

A straightforward implication of Proposition 8 is on the selection of the principal. If we suppose that becoming a principal implies an initial fixed cost, neither a player with extreme views nor one with a fully balanced view has the strongest incentives to take on the role of a principal (all else equal). Instead, those who profit the most are potential principals who lean to one side but only moderately.

Our third implication compares the analyzed game with that of a single agent L taking all decisions, a setting we refer to as the dictatorship of L. In that case, due to the lack of competition, the agent always takes his bliss point, $y_L = 0$. Intuitively, a balanced principal (weakly) prefers competing polarized agents, whereas an extreme principal (weakly) prefers a like-minded dictator. The following proposition shows that in the comparison between the two scenarios, a single crossing exists in the principal's payoffs.

Proposition 9 (**Competition**). *Fix (m, b, β). Then, 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. For $\theta < \tilde{\theta}$ she prefers a dictatorship of L.*

Moreover, if players are sufficiently patient and the principal's preferences are not fully aligned with any of the agents, i.e., $\theta > 0$, she strictly prefers to operate with two agents L and R.

The intuition follows directly from the characterization. While the principal prefers L to be selected, the presence of R allows the principal to influence L (and R) by threatening to support their rival. The cost, naturally, is that sometimes the agent choosing the policy has vastly different preferences. As the principal moves to the center, this cost declines because her preferences become more balanced.

If players are patient—which can be interpreted as frequent opportunities to act—the principal's power to moderate the agents through endorsement is strong, which means R can be tamed and hence reduces the principal's cost of adding R and creating a competitive setting. She prefers the presence of the polar opposite agent even when her own bliss point is almost to that of L.

Ally Principle. Our results provide a counterpoint to the ally principle of bureaucratic politics, by which a politician always prefers to delegate to agents closer to her preferences (Gehlbach, 2021). Proposition 4 shows that a principal prefers to delegate to her most distant agent when agents give an intrinsic value to task performance—more the more distant they are. In the real world, deviations from the ally principle are frequently observed. Companies invite activists to become board members, political leaders include their loudest critics in their government, organizations divide leadership between different factions and so on.[4] Often these type of actions are downplayed as attempts to "face-wash", to allow for "power-sharing" or simply to ensure proper representation. We provide another rational for this behavior through Proposition 9. A non-extreme principal may utilize the competition between opposing forces within the organization to make both sides move closer to her position in their respective tasks.

Preferences reports: Signalling extremism. Somewhat outside our model, we could ask whether it may payoff for an agent to make the principal (and the other agent) belief that he was more extremist than he actually is. To that extent, suppose an agent can announce his bliss point before the beginning of the game, and assume the others naively believe his claim. An equilibrium under these beliefs would involve strategies that are on-path incentive compatible given the agent's real preferences and the players' beliefs about the off-path punishments based on the reported preferences. As punishments are never realized on the

[4]For example, the German industrial manufacturer Siemens offered board positions to climate activists, Spanish Prime Minister Mariano Rajoy often selected opposing forces for similar tasks within his governments, and the German green party follows the tradition of having two leaders, one from the far-left faction of the party and one from the centrist one.

equilibrium path, punishments are fully based on the reported preferences and are never tested given the real preferences.

Would the agent report his true bliss point? He would not. The logic can be is straightforward in the case of L. As the principal believes that L's bliss point is further away from her own bliss point, she gradually begins to tilt her support in L's but never asks 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's belief about his off-path punishment worsens. In return, the principle is able to pull R's actions further in, thereby benefiting L. Thus, L gains from signaling to be extreme not only because that makes the principal more loyal toward him, and if anything reduces the concessions L is required to make, but it also benefits L because R is more afraid of polarization and willing to concede more. An interesting question is to what extent there is scope for "equilibrium signaling." However, to address this, an overall richer model is needed and we consider it beyond scope here.

5 Applications

In this section we provide two applications to illustrate how the mechanics of our model square with real-world behavior. The first concerns the historical example of the evolution of the center-right Christian Democrats in Italy to a party that successfully managed to play both sides of the political spectrum to ensure policy choices in her favor. The second concerns an observation from corporate finance that mission-driven managers often act more as bridge-builders in the sense that they allocate resources toward divisions further from their mission, than as empire builders allocating resources toward their "home" division.

5.1 The Christian Democrats and the 'opening to the left'

The Christian Democrats' decision to 'open to the left' (*apertura a sinistra*) is considered a seminal decision in postwar Italian politics. In 1962, the Christian Democrats began a series of coalition governments with the Socialists, who had been secluded from any governmental coalition until then. Though our simple model cannot capture the complexities of this historical period, we wish to emphasize how our strategic insights directly map to the observed behavior. One of our main findings, the principal's desire to embrace the enemy, may appear like a purely theoretical construct. Our example, however, provides suggestive evidence that it may be of practical relevance as well.[5]

Background. In postwar Italy, the Italian Christian Democrats (DC) were the key centrist player in Italian politics due to a strong social base and connections to the Catholic Church,

[5]A more detailed description can be found in Ginsborg (1990).

corporate interests, and the United States. However, since 1952, they lacked a majority in Parliament and had to search for coalition partners. The Christian Democrats' problem resembles that of our principal: their preferences were between those of their potential partners to the right and to the left. The DC could not fully predict which government coalition would be possible because coalition formation was plagued with uncertainties: the electoral outcomes were naturally uncertain, but so were internal developments within each political party.

The decision to 'open to the left'. In the 1950s, this search for allies was crucially curtailed to non-leftist parties by, among other reasons, a strong anti-left stance of the US government and Pope Pius XII's.

However, in line with our Proposition 9 predicts, some members of the DC understood that limiting their allies was damaging their ability to obtain their preferred policies. Amintore Fanfani was the first to propose the idea of opening to the Socialist Party. Fiercely contested internally by the DC factions closer to the Catholic Church, Fanfani was defeated in the DC Congress in 1959. However, with the start in 1961 of the Kennedy administration and Pope John XXIII's reassessment of the role of the Catholic Church, the positioning on the topic of both the United States and the Vatican changed.

Once these exogenous constraints were lifted, the DC's position changed too. In January 1962, the DC Congress backed a new direction for the party: seeking a 'center-left' coalition. Party chairman Aldo Moro envisioned that in a center-left coalition, the DC "had no intention of allowing the Socialists to push [the DC] any further than [the DC] wanted to go." (Ginsborg, 1990) Indeed, what Moro, Fanfani, and the other DC leaders wanted to do was to move the Socialists away from the communism and closer to their positions. By embracing the Socialists, they expected to moderate them.

Backed by leading Italian firms like FIAT, Pirelli, Olivetti, and ENI, the Christian Democrats envisaged a carrot-and-stick strategy aimed at transforming the Socialist Party. They believed they could attract the Socialist with the trappings of office and the chance of carrying through some reforms the DC agreed with. However, they could also discipline them with the threat of cutting the Socialist access to office and shifting to collaboration with the right. "When [Vittorio Valletta, the managing director of FIAT] went to see Kennedy in May 1962, Valletta recommended that if and when the USA began to support the Socialists financially, it should do so via the DC so that the latter could use the money 'as a lever to extract the cooperation of the Socialist party'"(Ginsborg, 1990).

The stick and the carrot. After the reformist government of Fanfani in 1962, the next center-left government—headed by Moro—slowed down reforms. Moro halted many

reforms, blaming the economic situation when the Socialists complained. In 1964, the tensions between the coalition partners mounted, and the negotiations to form Moro's second government were close to derailing.

In these circumstances, the DC benefited from the implicit threat of a turn to the right. Amid the long negotiations to form a government, the President of the Republic summoned the head of the Carabinieri, General Giovanni De Lorenzo. Though only partly understood by the public then, summoning De Lorenzo was a dubious move.[6] Socialist leader Pietro Nenni reacted to this news by immediately facilitating the formation of Moro's second government, putting an end to the Socialists' push for more ambitious reforms. Policies moved closer to those preferred by the DC.

> By now, the Socialists were very much a party of government. [...] Indeed, the gap between ideology and action, always a problem in Italy, was macroscopic in the Socialists' case. [...] after De Lorenzo, by making it clear that they would stay in the government at almost any cost, [the Socialists] allowed the pressure for reform to dwindle. (Ginsborg, 1990)

5.2 Mission-Driven CEOs

There is a long and ongoing debate at the intersection of organizational economics and corporate finance concerning managerial *empire building*(see, e.g., Gertner and Scharfstein, 2012, for an overview of the initial literature). The basic idea of managerial empire building is that managers aim to relocate resources to branches closer to them to gain influence and power, potentially, but also to establish a legacy about the direction of the company/division/unit. A standard argument following Shleifer and Vishny (1989) is that managers have an incentive to entrench themselves by giving their friends larger power within the organization. A counteracting force may be that mission-driven CEOs (Bolton and Dewatripont, 2013) may still need to *build bridges* to parts of the firm they are less aligned with to dismantle their enemies (Stein, 2003, see).

Looking at CEO behavior, Xuan (2009) finds that CEOs indeed engage in bridge-building behavior, allocating resources to the enemy rather than the friend. While Xuan (2009) interprets this as evidence for the "dark-side theories" of managerial entrenchment, our model provides an alternative explanation that does not require any risk of termination. In fact, in our model, agents have no power over the principal other than moving the company's

[6] During De Lorenzo's time as head of the Italian military intelligence, thousands of dossiers of politicians were drawn up, probably for blackmail purposes. After, as head of the Carabinieri, he developed his 'Solo' plan. Though details are unclear, the Solo plan was a supposedly counter-insurgency plan that involved the arrest of leading political figures of the left and taking control of political institutions, the radio, and the television.

policy further from the principal's vision. Embracing the enemy ensures that all divisions move closer toward the manager's vision of the firm.

6 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 more potential in embracing the enemy, thereby mitigating her action choices, than ensuring that a close ally is chosen more often. The mitigation also serves as a thread point toward the ally to pander further toward the principal's preferences.

We show that for a given conflict between agents, moderately biased principals enjoy larger utility than those at one extreme or those fully centrist, suggesting that we are likely to see these types of principals the most. Moreover, the principal in a relationship with an 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 leading to the embracing of the enemy effect, there are several aspects our model ignores that provide scope for future research. Among these are welfare concerns for the organization, the question of perfect observability of the principal's endorsement decision, and the potential role of asymmetric information.

A Proofs

We use the following Lemma in some of the proofs.

Let h describe a history ending at a node in which it is the principal's turn and, abusing notation, let $h \cup k$ describe the consecutive note in which k is selected.

Lemma 3. *Take any contract with ex-ante value v_P to the principal. Then there exists a contract with value v_P such that $y_L(\cdot) \leq y_R(\cdot)$.*

Proof. We prove this result by construction and show it for any history h including the empty set. Suppose there exists some principal-worst contract C such that at history h, $y_L(h \cup L) > y_R(h \cup R)$, and it gives a principal's value of $v_P(h)$.

Suppose first that either $y_L(h \cup L) < \theta$ or $y_R(h \cup R) > \theta$. Then, we can consider an alternative contract $\tilde{C}$ that is identical to C but with $\tilde{y}_L(h \cup L) = y_L(h \cup L) - \varepsilon$ and $\tilde{y}_R(h \cup R) = y_R(h \cup L) + \frac{(1-p(s(h)))\varepsilon}{p(s(h))}$. Note that this change relaxes L's and R's enforcement constraints at $h \cup L$ and $h \cup R$, respectively. Besides, the principal's value is unchanged, that is,

$$v_P(h) = \tilde{v}_P(h)$$

therefore, the contract is incentive compatible.

Notice that this procedure works for any $\varepsilon > 0$ as long as the original $y_L(\cdot)$ and $y_R(\cdot)$ have been on the same side of θ. If we select

$$\varepsilon = (y_L(h \cup L) - y_R(h \cup R))p(s(h))$$

we obtain a contract at which $\tilde{y}_L(h \cup L) = \tilde{y}_R(h \cup R)$ and $\tilde{v}_P(h) = v_P(h)$.

Now suppose $y_R(h \cup R) \leq \theta \leq y_L(h \cup L)$. Take the following alternative contract $\tilde{C}$ that is identical to C but with

$$\tilde{y}_L(h \cup L) = \theta - \frac{p(s(h))}{(1 - p(s(h)))}(\theta - y_R(h \cup R))$$

and

$$\tilde{y}_R(h \cup R) = \theta + \frac{1 - p(s(h))}{p(s(h))}(y_L(h \cup L) - \theta).$$

Note first that, as before, the agent's enforcement constraints are relaxed.

Also, note that the principal's value is unchanged because

$$\tilde{v}_P(h) = -\Bigg(p(s(h)) \left(\frac{1-p(s(h))}{p(s(h))} (y_L(h \cup L) - \theta) - \beta v_P(h \cup R) \right)$$
$$+ (1 - p(s(h))) \left(\frac{p(s(h))}{1-p(s(h))} (\theta - y_R(h \cup R)) - \beta v_P(h \cup L) \right) \Bigg) =$$
$$v_P(h) = -\Bigg((1 - p(s(h))) \left(\mid y_L(h \cup L) - \theta \mid - \beta w_P(h \cup L) \right)$$
$$+ p(s(h)) \left(\mid \theta - y_R(h \cup R) \mid - \beta v_P(h \cup R) \right) \Bigg).$$

Thus the contract is incentive compatible. □

A.1 Proof of Proposition 1 and 2

Because the example is a special case of the general setting, we omit direct proofs here. We offer a sketch of the proof for Proposition 2 below, given that our no-commitment result is not yet finalized.

Proof. Proposition 2 follows from recognizing that the optimal contract from Proposition 1 is identical to L's worst contract. Then, invoking proposition 5 and solving for existence yields $\hat{m}$, the principal's minimum strength to sustain the commitment contract. Otherwise, full polarization is the only option. □

A.2 Proof of Lemma 1

We prove both statements separately.

First Statment (binding DEC or θ).

Proof. The proof is constructive.

Definition 2. A history h with property Π is said to be a *first history with property* Π if there is no history leading up to history h that has property Π. In particular, a first history h of k being selected implies a history in which $-k$ was selected in all histories leading to h.

Definition 3. A history h with property Π is said to be a *last history with property* Π if there is no possible history following h with property Π.

Step 1. k's dynamic enforcement constraint has slack upon first selection. Take an arbitrary feasible contract C in the class defined by Lemma 1, that is $\mathbf{y}_L \leq \boldsymbol{\theta} \leq \mathbf{y}_R$ and assume it has a first history of k being selected at which k's dynamic enforcement constraint has slack and $y_t \neq \theta$ at that history.

Then, there is a $\tilde{y}_t$ such that $|\tilde{y}_t - \theta|$ is incentive compatible to k, and strictly preferred by P and $-k$. Thus, there exists a contract C^1 identical to C apart from k selecting $\tilde{y}_t$ which P prefers ex-ante over C. Now, either C^1 exists for $\tilde{y}_k = \theta$ which is P's most preferred action, or there is a $\tilde{y}_k \neq \theta$ such that k's dynamic enforcement constraint binds.

Step 2. k's dynamic enforcement constraint has slack upon re-selection. Consider now a contract in the class of C^1, such that no agent's dynamic enforcement constraint has slack upon first selection or that the agent chooses the principal's preferred action θ. Furthermore, assume that there is still at least one history with the property that k's dynamic enforcement constraint has slack. Take a first on-path history h with that property.

Now consider a contract C^2 which is identical to C^1, apart from that at history h, we select the associated $\tilde{y}_t$ such that either $\tilde{y}_t = \theta$ or k's dynamic enforcement constraint binds. This benefits, P leaving forward incentive constraints unchanged.

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. By assumption, k has been re-selected at an on-path history h earlier. Because we started out with a class C^1 contract, we know that at any earlier stage in which k was selected either k's dynamic enforcement constraint was binding or k selected θ. In both cases, k's dynamic enforcement constraint at that earlier stage may now be violated. To ensure that this is not the case, take the last on-path history h' at which k was selected before.[7]

Now consider a contract C^3 identical to C^2, but with the difference that at history h', agent k selects $\tilde{y}_{t'}$ such that $|\tilde{y}_{t'} - y_{t'}| = \delta(h|C^2, h')|\tilde{y}_t - y_t|$ and $d(\tilde{y}_{t'}, y_k^*) < d(y_{t'}, y_k^*)$. Thus, k decreases his pandering toward P in period t' by exactly the net present value of his own increase in pandering in period t. Thus, a selected k at time t' is indifferent between the continuation contracts from C^3 and from C^1. Thus, by construction, k's dynamic enforcement constraint holds. Because these concessions are equivalent in time t' net-present values they are equivalent in any time $t'' < t'$ net present values and thus C^3 is payoff equivalent to C^1 also for the principal, yet for all histories up to h, k's dynamic enforcement constraint holds with equality or k chooses θ whenever selected.

Using either the procedure from Step 1 or from Step 2 iteratively, we can eliminate all

[7] Analogous to the first history, the last history is defined as a history such that all other histories that follow on-path h' do not have the particular property.

on-path histories in which an agent selects any action different from θ while having slack on her dynamic enforcement constraint.

Thus, it is without loss to consider the class of contracts at which agents either have a binding dynamic enforcement constraint or select the principal's bliss point whenever selected on the equilibrium path.

Second Statement (No Overshoot). We show that for any contract violating either inequality of the Lemma 1 another contract exists that does not violate the inequalities.

Fix a contract sucht that when R is selected and $y_{R,t} < \theta$ for some on-path history h and a subsequent selection of R. 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 $y_{R,t}$ marginally. That contract improves the principal's payoff in period t by $dy_{R,t}$, relaxes R's dynamic enforcement constraint in all periods $\tau \leq t$ and leaves them unchanged in all period $\tau > t$. It is thus incentive compatible for R. Now consider the last period $\tau < t$ in the sequence leading up to h in which L has been selected if it exists. The increase of $y_{R,t}$ in period t has tightened L's dynamic enforcement constraint by the change's period-τ value $\rho dy_{R,t}$, where $\rho \in (0,1)$ describes the (discounted) conditional probability of reaching the node in which we increased $y_{R,t}$ starting from τ (see the proof of Lemma 2 on how to construct said ρ). Thus, augmenting the contract for a second time by reducing the prescribed $y_{L,\tau}$ by at most $\rho dy_{R,t}$, restores all relevant incentives for L. Now, observe that the first and second augmentation jointly are at worst ex-ante payoff neutral to the principal because in period-τ value terms her gains and losses equate. If no such period τ exists or less reduction is required, the principal is strictly better off. Hence assuming $y_{R,t} \geq \theta$ at any t is without loss. The case for $\theta \geq y_{L,t}$ is analogous.

□

A.3 Proof of Lemma 2

Proof. Let $\mathcal{H}(h)$ describe the set of all possible histories conditional on reaching history h and consider contract $C \in \mathcal{C}$. Then, $Pr(h'|C,h) \in [0,1]$ describes the conditional probability of reaching a particular history $h' \in \mathcal{H}(h)$ conditional on reaching history h under contract $C \in \mathcal{C}$. We need two pieces of notation in what follows: First—taking the interpretation of a discount factor that the game ends each period with probability $(1-\beta)$—we describe for $t' \geq t$ the probability of reaching history $h_{t'} \in \mathcal{H}(h_t)$ given C by

$$\delta(h_{t'}|C,h_t) := \beta^{t-t'} Pr(h_{t'}|C,h_t).$$

Second, the probability of reaching $h_{t'}$ and subsequently selecting $k_{t'+1} = R$ is

$$\rho(h_{t'}|C,h_t) := \delta(h_{t'}|C,h_t)\left(\frac{1}{2} + s_{t'+1}(h_{t'})\right).$$

We can then write R's on-path value from contract C when being in power after history h as

$$v_R(R;C,h) = b - (1 - y_R(C,h)) + \beta\left(\sum_{\eta\in\mathcal{H}(h)}\Big(\rho(\eta|C,h)(b-(1-y_R(C,\eta))) + (\delta(\eta|C,h)-\rho(\eta|C,h))(y_L(\eta)-1)\Big)\right)$$

Similarly, we can write P's on path value from seeing R in power after history h as

$$v_P(R;C,h) = \theta - y_R(C,h) + \beta\left(\sum_{\eta\in\mathcal{H}(h)}\Big(\rho(\eta|C,h)(\theta-y_R(C,\eta))) + (\delta(\eta|C,h)-\rho(\eta|C,h))(y_L(\eta)-\theta)\Big)\right).$$

Summing the two yields

$$v_R(R;C,h) + v_P(R;C,h) = b - (1-\theta) + \beta\left(\sum_{\eta\in\mathcal{H}(h)}\Big(\rho(\eta|C,h)(b+2\theta-2y_L)) + \delta(\eta|C,h)(2y_L(\eta)-(1+\theta))\Big)\right)$$

which for $y_L = \theta$ collapses to

$$v_R(R;C,h) + v_P(R;C,h) = b - (1-\theta) + \beta\left(\sum_{\eta\in\mathcal{H}(h)}\Big(\rho(\eta|C,h)b - \delta(\eta|C,h)(1-\theta)\Big)\right).$$

The dynamic enforcement constraint of R binds because $y_R \neq \theta$ and Lemma 1 such that $v_R(R;C,h) = b + \beta\alpha^C$ which implies

$$v_P(R;C,h) = (\theta-1)(1+\beta) + \beta\left(\sum_{\eta\in\mathcal{H}(h)}\rho(\eta|C,h)b\right) - \beta\alpha^C \tag{3}$$

and thus, the value of the contract is increasing in $s(h_{t'})$ for $h_{t'}$ following h. Thus, it is optimal to set $s(h') = m$ for all $h' \in \mathcal{H}(h_t)$ whenever $k_{t+1} = R$ which proves the claim if $(s = m, y_L = \theta)$ satisfies L's dynamic enforcement constraint. □

A.4 Proof of Proposition 3

Proof. The result follows because L and R play a constant sum game in every period and $\beta < 1$. □

A.5 Proof of Proposition 4

Proof. The proof is organized as follows.

1. We first define the worst punishment strategy and prove that it is indeed the worst (Lemma 4).
2. We then characterize contracts that achieve the first best and the parameter regions for which they are incentive compatible (Lemma 5 to 7)
3. Next, we characterize the set of contracts that ensure $y_L = \theta$ on the euqilibrium path (Lemma 8 and 9).
4. We then show that these contracts are optimal (Lemma 10 and 11)
5. Finally, we characterize the optimal contract for the remaining parameter region (Lemma 12). This last step also proves the only if part of the statement regarding the first best.

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

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

We prove both propositions jointly using a sequence of Lemmas. Throughout, let and

Definition 4 (Grim Trigger punishment)**.** The response to a deviation is said to be *grim trigger punishment* if the contract prescribes that a deviation by i leads to the principal fully endorsing $j \neq i$ in any continuation game and both agents choose their respective bliss points whenever selected, $y_L = 0$ and $y_R = 1$.

Lemma 4. *An agent's worst sustainable payoff conditional on being selected is the one obtained from grim trigger punishment if and only if* $(1 - \beta) \geq 2m\beta$*. It is given by* $b + \beta\alpha^C$*.*

Proof. The principal has commitment power and thus can promise to support the non-deviator unconditionally, which is the ceteris paribus worst outcome for the deviator. Moreover, it is trivially incentive compatible for the non-deviator to respond by playing her bliss point which results in an outcome of -1 for the deviator whenever the non-deviator is selected. The deviator's best response to any worst punishment is to select her bliss point.

Finally, the best deviation is to select said bliss point as well, resulting in a payoff of b in the deviation period and a continuation payoff of $\beta\alpha^C$.

Finally, we need to show that winning is desired on such occasions. For this step, we make use of the fact that at any period, the stage payoffs $u_L + u_R = b - 1$, meaning that L and R compete in a constant-sum manner both over the selection and the policy choices. Winning is desired if the worst continuation payoff conditional on being selected, $b + \beta\alpha^C$, is larger than the best continuation payoff conditional on not being selected, $(b-1)/(1-\beta) - (b + \beta\alpha^C)$

Thus,

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

□

The next 3 lemmas show the feasibility of the first-best allocation.

Lemma 5. *Suppose $\breve{\theta} \leq \theta \leq \hat{\theta}$. Then there is an optimal contract with $s = m$ and $y_i = \theta$ in all periods.*

Proof. Assume the grim trigger punishment and an on-path contract $s = m$ and $y_i = \theta$. Then R's dynamic enforcement constraint is

$$b - (1 - y_R) + \frac{\beta}{1-\beta}\left((1/2 + m)(b - (1 - y_R)) - (1/2 - m)(1 - \theta)\right) \geq b + \beta\alpha^C$$

which is equivalent to

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

Under this contract, $y_R = \theta$ is sustainable if

$$\begin{aligned} & \theta \geq 1 - \frac{4m(b+1) + \theta(1-2m)}{2 - \beta(1-2m)}\beta \\ \Leftrightarrow \quad & \theta \geq 1 - \left(\frac{1}{2} + m(1+2b)\right)\beta = \breve{\theta}. \end{aligned}$$

Analogously, L's dynamic enforcement constraint conditional on on-path play, $s =$

$m, y_R = \theta$, is

$$b - y_L + \frac{\beta}{1-\beta}\left((1/2 - m)(b - y_L) - (1/2 + m)\theta\right) \geq b + \beta\alpha^C$$

which implies

$$y_L \leq \frac{(1+2m)(1-\theta)\beta}{2 - \beta(1+2m)}.$$

Under this contract, $y_L = \theta$ is sustainable if

$$\begin{aligned} & \theta \leq \frac{(1+2m)(1-\theta)\beta}{2 - \beta(1+2m)} \\ \Leftrightarrow \quad & \theta \leq (1/2 + m)\beta = \hat{\theta}. \end{aligned}$$

Because both agents play the principal's first-best, this contract is optimal. □

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

Proof. We prove the claim by guessing a first-best contract. Then we show that first-best is indeed incentive compatible under that contract.

Step 1. A candidate Contract. Consider the following candidate contract with $y_i = \theta$ and on-path Markov strategies, s_L, s_R by the principal where s_i is the principal's endorsement strategy if i has been selected last 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, and there is grim trigger punishment.

Step 2. First-best is incentive compatible. Recall that in each period, the joint utility of L and R is $u_L + u_R = b - 1$ regardless of their choices and who gets selected. Further, assume, for now, we are in a setting in which both players' dynamic enforcement constraint is satisfied with equality. If the dynamic enforcement constraint of the selected agent binds in a given period, then that agent's continuation payoff is identical to deviating and entering the worst punishment, that is, $b + \beta\alpha^C$. That implies that the agent who is *not selected* receives the residual $(b-1)/(1-\beta) - (b + \beta\alpha^C)$. But then, if, under the principal's (Markov) strategy (s_L, s_R), both dynamic enforcement constraints are expected to hold with equality,

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

$$(y_R - 1) + \beta\left((1/2 + s_R)\left(b + \beta\alpha^C\right) + (1/2 - s_R)\left(\frac{b-1}{(1-\beta)} - (b + \beta\alpha^C)\right)\right) = \beta\alpha^C,$$

we can solve for y_L and y_R as

$$y_L = \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_L; \tag{5}$$

and

$$y_R = 1 - \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_R. \tag{6}$$

Note that the feature that both equations (5) and (6) are expected to hold in all future periods implies that the selected agent's choice is only a function of the next expected endorsement by the principal.

Thus, a contract with binding dynamic enforcement constraints and $y_i = \theta$ exists, if and only if there are $s_i \in [-m, m]$ such that the RHS of equations (5) and (6) are equal to θ.

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

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

which is equivalent to

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

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, first best can be attained if $\theta > \max\left\{1-\overline{\theta}, \underline{\theta}\right\}$ 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 obsolete.

Step 3. Initial period. Finally, because $y_i = \theta$ in any continuation game after the initial endorsement, the principal is indifferent between any initial endorsement which concludes the proof. □

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

Proof. As before we prove the claim by constructing a candidate contract verifying it satisfies incentive compatibility.[8]

Step 1: A candidate contract. Consider the following candidate contract with $y_i = \theta$ and on-path Markov strategies, $s_L = m$ and,

$$s_R = m - \frac{\theta - \beta(1/2+m)}{\beta\left(\beta(1/2+m)(b+1) - \theta\right)} =: \overline{s}_R.$$

The first-period endorsement strategy is arbitrary, and off-path we revert to a grim trigger punishment.

Step 2: First-best is incentive compatible. Incentive compatibility for L can be shown by describing L's value at the beginning of a period conditional on being selected from choosing $y_L = \theta$ in any period in which L is selected

$$v_L(L) = b - \theta + \beta\left((1/2+m)v_L(R) + (1/2-m)v_L(L)\right)$$

and L's value at the beginning of a period conditional on *not* being selected assuming that $y_R = \theta$ in any period in which R is selected

$$v_L(R) = -\theta + \beta\left((1/2+s_R)v_L(R) + (1/2-s_R)v_L(L)\right).$$

[8] If this region exists, there is a contract $s = m$, such that $y_L = \theta$, but it requires $y_R < \theta$ which violates Lemma 1 and does not constitute a first best. Hence, we construct a contract such that $y_R = \theta$.

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

$$v_L(L) = \frac{(1-\beta(1/2+s_R))b}{(1-\beta)(1+\beta(m-s_R))} - \frac{\theta}{(1-\beta)}$$
$$v_R(L) = \frac{(1/2-s_R)\beta b}{(1-\beta)(1+\beta(m-s_R))} - \frac{\theta}{1-\beta}.$$

The candidate contract is incentive compatible for L if and only if

$$\begin{aligned} & v_L(L) \geq b + \beta\alpha^C \\ \Leftrightarrow \qquad & s_R \leq \overline{s}_R. \end{aligned}$$

which holds by assumption. In particular, observe that $\overline{s}_R < m$ if $\underline{\theta} > \theta > \hat{\theta}$ and $\overline{s}_R = m$ if $\theta = \hat{\theta}$.

Incentive compatibility for R follows from recalling that $u_L + u_B = b - 1$ in every period which implies that

$$v_R(R) = \frac{b-1}{1-\beta} - v_L(R) = \frac{b+\theta-1}{1-\beta} - \frac{(1/2-s_R)\beta b}{(1-\beta)(1+\beta(m-s_R))}.$$

The candidate contract is incentive compatible for R if and only if

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

What remains, is to show that $\overline{s}_R \geq \underline{s}_L$ 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, the relevant 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

$$\frac{(1+b)^2(1+2m)\beta}{(1-\beta)^2}\Big(2 + \beta\left(2m + 4\beta - (1-2m)\beta^2(1+2m)^2 - 5\right)\Big) \geq 0. \tag{7}$$

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

$$8(1 + 6m)\beta^3 \geq 0$$

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 (7) is increasing on $mzin(0, 1/2)$.

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. Initial period Finally, because $y_i = \theta$ in any continuation game after the initial endorsement, the principal is indifferent between any initial endorsement which concludes the proof. □

We now characterize incentive-compatible contracts for three different regions.

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 is selected. 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 contract. Consider the following candidate contract with an on-path strategy $y_L = \theta$ whenever L is selected, 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 is selected. The principal's on-path strategy is stationary and $s_{\neg R} = -m$ until R is selected for the first time. Thereafter, her strategy is stationary and $s_R = m$ in any on-path node that follows. Any deviation triggers grim trigger punishment.

Step 2. The candidate is incentive compatible. Take the R's dynamic enforcement constraint in a contract that specifies $s = m$ and $y_L = m$ in any future on-path node, which we constructed in (4) in the proof of Lemma 5. Observe that

$$\tilde{y}_R = 1 - \frac{4m(b+1) + \theta(1-2m)}{2 - \beta(1-2m)}\beta.$$

solves (4) with equality. Because $\theta < \hat{\theta}$ the reuslting $\tilde{y}_R > \theta$.

Now consider L's dynamic enforcement constraint under the proposed contract which is

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

which 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. As we have seen, L's dynamic enforcement constraint has slack when playing the principal's bliss point, $y_L = \theta$, and receiving the worst endorsement strategy she can imagine (because $\theta < y_R$ and $b > 0$ being selected on path is preferred by L). Thus, when promised more support $s < m$, L's incentives are undistorted. Moreover, any choices prior to R's first selection are inconsequential to R's incentive when deciding on y_R for the first time. Because the principal prefers θ over $\tilde{y}_R$ she chooses $s_L = -m$ until the first time R's incentives become relevant. □

Lemma 9. *If $\theta \in [\underline{\theta}, \overline{\theta}]$, there exists a feasible contract in which both dynamic enforcement constraints hold with equality and $y_L = \theta$. In that contract $s = -m$ until the first time R is selected. Thereafter, $s_R = m$ whenever R was selected last and $s_L < m$ whenever L was selected last.*

Proof. Again, we can restrict attention to settings in which $\theta \leq (1 - \overline{\theta})$ because the complementary case is covered in Lemma 6.

Step 1. A candidate contract. Consider the following contract. The principal's on-path endorsement strategy is to endorse

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

whenever L was selected last, she endorses $s_R = m$ whenever R was selected last. When selected L chooses $y_L = \theta$ and R chooses

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

on the equilibrium path. In the first period $s_L = -m$. Any deviation is followed by grim trigger punishment.

Step 2. The candidate is incentive compatible. Suppose, that as in Lemma 6 both dynamic enforcement constraints hold with equality. Thus, recall

$$y_R = 1 - \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi m \tag{5}$$

and

$$y_L = \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_L. \tag{6}$$

The proposed s_L solves (6) 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 to the contrary that $\psi > 0$ and recall that this implies $\beta(1+2m) > 1$ and hence $1 - \beta < \beta 2m$. Moreover, recall that, for this case to have relevance, we require

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

But then because $\psi > 0$

$$2\beta m \beta(2m+1)(b+1) > (1-\beta)(b+1) > (1-\beta)$$

where the last step follows from $b > 0$. 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. Initial Period. The reasoning is analogous to that in step 3 of the proof of Lemma 8. As we have seen, L's dynamic enforcement constraint holds when playing the principal's bliss point, $y_L = \theta$. Thus, when promised more support $s = -m$, L's incentives are undistorted. Moreover, any choices before R's first selection are inconsequential to R's incentive when deciding on y_R for the first time. Because the principal prefers θ over $y_R > \theta$, she chooses $s_L = -m$ until the first time R's incentives become relevant. □

Next, we show that the contracts we constructed are optimal which also proves the only if part of the first-best statement in the proposition.

Lemma 10. *The contract constructed in the proof of Lemma 8 is optimal in its region.*

Proof. We will prove optimality as follows. First, we make use of the existing Lemmas to determine a restricted contract space that contains at least one optimal contract. Then we will show that, within that restricted contract space, no contract other than the one used in the proof of Lemma 8 can be optimal. In steps 1 and 2 we focus on histories in which R has been selected at least once.

Step 1. Other candidates. Invoking Lemma 1 we know that it is sufficient to restrict attention to contracts such that at any history the agent selected either chooses $y_i = \theta$ or that agent's dynamic enforcement constraint holds with equality. Invoking Lemma 1 we know we can, in addition restrict attention to contracts that imply $y_L \leq \theta \leq y_R$. Finally, Lemma 2 implies that within the class of contracts such that $y_L = \theta$ and $y_R \neq \theta$ we can restrict attention to those that imply $s = m$ unless either L's dynamic enforcement constraint holds with equality or R plays $y_{R,t} = \theta$ at least at some on-path node t.

Step 2. No other candidate is optimal.

Persistence of a binding dynamic enforcement constraint. First, consider a time t in which a selected L's dynamic enforcement constraint holds with equality. Then, that dynamic enforcement constraint holds in $t+1$ as well should L get selected in $t+1$. The reason is straightforward. By assumption, L receives the ceteris paribus lowest possible continuation value starting from period t. If, after being selected again, L receives a higher continuation value starting from $t+1$ we can lower L's payoff in $t+1$ without affecting any of the actions in t onward. A contradiction to having had the lowest possible continuation value in t.

If L is at her dynamic enforcement constraint, R's dynamic enforcement constraint must have slack thereafter. Second, consider a time t in which a selected L's dynamic enforcement constraint holds with equality. Then, conditional on selection, R's dynamic enforcement has to have slack in $t+1$. To see this, recall first that $u_L + u_R = b - 1$ in each period. Thus, if L's dynamic enforcement constraint holds with equality in t, and R's dynamic enforcement constraint holds in $t+1$, then L's value of the game at the beginning of t is $b + \beta\alpha^C$ while L's value of the game at the beginning of $t+1$ conditional on R being selected is $(b-1)/(1-\beta) - b - \alpha^C$. Thus, her dynamic enforcement constraint solves

$$b - y_L + \beta((1/2 - s)(b + \beta\alpha^C) + (1/2 + s)((b-1)/(1-\beta) - b - \alpha^C) = b + \beta\alpha^C$$

Which becomes equation (6):

$$y_L = \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_L, \tag{6}$$

When $\theta < \underline{\theta}$, there is no $s_L \in [-m, m]$ such that $y_L \leq \theta$ solves (6) and hence, if R is selected in $t+1$, R's dynamic enforcement constraint has to have slack. Invoking Lemma 1, this implies $y_{R,t+1} = \theta$.

No contract exists in which $y_R = \theta$. Finally, we show that there is no optimal contract in which $y_{R,t+1} = \theta$ such that R's dynamic enforcement constraint has slack. To see this, recall that under the candidate contract from the proof of Lemma 8 when R is in power she faces a contract that promises her the highest possible payoff a contract in the relevant class of contracts in all periods in which R is not selected: The principal fully supports R and L chooses the highest value, $y_L = \theta$ available in 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 $y_{R,t+1} = \theta$ if R would choose $y_{R,\tau} > \theta$ if selected at some future date $\tau > t+1$. But then, once again by Lemma 1, no such contract exists in the relevant class of contracts that is such that R's dynamic enforcement constraint holds with slack in τ. But then, from R's perspective such a promise is no different from promising to return to the candidate contract from $t+1$ onward which rules out that $y_{R,t+1} = \theta$ is incentive compatible because $\theta < 1 - \frac{\beta}{2} - (1+2b)m\beta$ and hence the best response to $s = m, y_L = \theta$ is $y_R > \theta$.

Step 3. Initial periods. The contract delivers the principal's first best until R is selected for the first time and $s = -m$ minimizes the chances of this event. That completes the proof. □

Lemma 11. *The contract constructed in the proof of Lemma 9 is optimal in its region.*

Proof. First, we determine a restricted contract space, which must contain a contract superior to the candidate contract if the candidate contract is indeed suboptimal. Second, we show no such contract exists in the restricted space.

As before we restrict attention to the continuation game after the first time R has been selected for Steps 1 and 2. We will turn to the initial periods in step 3.

Step 1. Other Candidates. We first show that all other contracts in which the dynamic enforcement constraint of both agents is satisfied with equality throughout are suboptimal. Then we show that in the class of contracts in which $y_L = \theta$, no better contract exists. Finally, we show that no better contract exists in which the dynamic enforcement constraint of either agent holds with slack at some point.

The optimal contract with binding dynamic enforcement constraints. Consider the principal's problem

$$v_P(s_R; s_L) = \max_{s_R} -(1-\beta)|\theta - y_R(s_R)| + \beta\left((1/2+s_R)v_P(s_R; s_L) + (1/2-s_R)v_P(s_L; s_R)\right)$$
$$v_P(s_L; s_R) = \max_{s_L} -(1-\beta)|\theta - y_L(s_L)| + \beta\left((1/2+s_L)v_P(s_R; s_L) + (1/2-s_L)v_P(s_L; s_R)\right),$$

subject to the agent's binding dynamic enforcement constraints, (6) and (5)

$$y_R(s_R) = 1 - \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_R \tag{5}$$
$$y_L(s_L) = \beta\left(\frac{b-1}{2(1-\beta)} - \alpha^C\right) + \psi s_L \tag{6}$$

Using the two Bellman equations above, we can solve for the principal's value of selecting the optimal strategy s_i in all periods in which i is selected, taking the choice s_j in periods in which R is selected as given. Thus, the principal'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

$$\rho_R(s_R; s_L) = \frac{1-\beta(1/2-s_L)}{1-\beta(1/2-s_L)+\beta(1/2-s_R)} \text{ and } \rho_L(s_L; s_R) = \frac{\beta(1/2+s_L)}{1-\beta(1/2+s_R)+\beta(1/2+s_L)}.$$

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 is selected.

Taking derivatives of the principal's objectives yields

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

and

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

which, after replacing via equations (5) and (6), can be written as

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

and

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

Because $\theta > \underline{\theta}$ we know that $\underline{\theta} < 1/2$ and hence $\psi < 0$ (see the proof of Lemma 9 for the formal reasoning). Thus, it is immediate that (VPL) is maximized for the largest s_L not violating any other constraint. From Lemma 1, one such constraint is that $y_L \leq \theta$ which is exactly met at the candidate contract. Because the dynamic enforcement constraint holds with equality and R's decision is unaffected conditional on the class of contracts we are in, selecting s_L such that $y_L = \theta$ is indeed optimal in the class of contracts that require the dynamic enforcement constraint of both parties to hold with equality.

Using the same argument, (VPR) is maximized for the largest s_R not violating any other constraint. Since $\theta < 1 - \overline{\theta}$ when first-best is not attainable and $\theta < \underline{\theta}$, $y_R > \theta$ for all s_R, R's dynamic enforcement constraint holds with equality by construction and L's decision is unaffected conditional on the class of contracts we are in, selectin $s_R = m$ is indeed optimal in the class of contracts that require the dynamic enforcement constraint of both parties to

hold with equality.

Thus the candidate contract is the optimal contract within the class of contracts in which both dynamic enforcement constraints hold with equality.

In the optimal contract, R's dynamic enforcement constraint holds with equality. Suppose that there exists an optimal contract in which R's dynamic enforcement constraint has slack at some node at time t. Then, because that contract is optimal and R's constraint holds with slack, there is an optimal contract in which R's dynamic enforcement constraint has slack conditional on R being selected again in $t+1$. Thus, using Lemma 1, it is without loss to assume that $y_{R,\tau} = \theta$ until L is selected the next time. Moreover, if such a contract is feasible, it is also feasible assuming that whenever L is selected the next time, L's dynamic enforcement constraint holds with equality. Because $u_L + u_R = b - 1$ in every period, such a continuation game implies the *best* possible continuation game for R conditional on losing. But then, when $\theta < 1 - \overline{\theta}$ which holds whenever $\overline{\theta}$, all contracts that satisfy R's dynamic enforcement constraint with equality for a fixed y_R until the next time L is selected imply $y_R > \theta$. By monotonicity of the dynamic enforcement constraint, no contract with $y_R \leq \theta$ exists that satisfies R's dynamic enforcement constraint. A contradiction to the premise that R's dynamic enforcement constraint has slack at some node at time t

Step 2. No other candidate is optimal. Observe that by Lemma 1, whenever L is not at her dynamic enforcement constraint we need $y_L = \theta$. Moreover, we know from the previous argument that R's dynamic enforcement constraint holds with equality. Invoking Lemma 2 and observing that because $\theta > \underline{\theta}$, $s = m, y_L = \theta$ does not satisfy L's dynamic enforcement constraint, implies that the candidate contract is optimal.

Step 3. Initial periods. Consider an initial period in which L is selected. If R was selected next period, we would enter the continuation game in which agent L selects $y_L = \theta$ if selected. Note that because $\underline{\theta} < 1/2$, $\psi < 0$, and hence $b + \beta\alpha^C > (b-1)/(1-\beta) - \beta\alpha^C$ such that being selected is preferred to not being selected even when the dynamic enforcement constraint holds with equality conditional on being selected. Thus, because in the initial period $s = -m < s_L$, L's dynamic enforcement constraint has slack, and does not affect R's incentives which makes $s = -m, y_L = \theta$ feasible. It delivers the principal's first best until R is selected for the first time and $s = -m$ minimizes the chances of this event. That completes the proof. □

Lemma 12. *If $\theta > \overline{\theta}$, then $y_i \neq \theta$ and $s_R = -s_L = m$ in the optimal contract.*

Proof. We first show that such a contract exists and is feasible before turning to optimality.

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

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

Off-path behavior is as in the optimal contract constructed in the proof of Lemma 5.

Step 2. The candidate is feasible. Because $1/2 \geq \theta > \underline{\theta}$, it follows that $\psi < 0$ and thus, the endorsement strategies imply the maximal $y_L(s_L)$ and the minimal $y_R(s_R)$ available assuming that dynamic enforcement constraints, (5) and (6), hold at all times.

Step 3. The candidate is optimal. Recall that promising the agent j is going to be brought to her dynamic enforcement constraint the next time she is selected implies the largest possible payoff for the currently selected agent i conditional on any future selection of j. By monotonicity of (5) and (6) in s_i and the fact that the agent that was selected last receives the principal's full endorsement, no other contract that satisfies that agent's dynamic enforcement constraint in the relevant class can lead to y_i closer to θ. □

□

A.6 Proof of Proposition 5

The principal's punishment is described by the principal-worst equilibrium of the game. We construct it in what follows.

Lemma 13. *There is a principal-worst contract in which $y_R = 1$ whenever R is selected.*

Proof. The existence of a worst contract is guaranteed by the compactness of the equilibrium payoff set which follows from standard arguments.

We start with a given principal's worst contract C, which gives the principal an ex-ante value of α_P.

Now we construct a contract C' with $y_R(\cdot) = 1$ with an ex-ante value of at most α_P

It is instructive to begin with the properties of the principal-worst contract if it is such that the principal's enforcement constraint never binds (except, naturally, at h_0).

Claim 1: *Suppose there is a principal's worst contract C such that for any history $h \neq h_0$, the principal's continuation value $\alpha_P(C, h) > \alpha_P$ and $y_R(h \cup R) < 1$ for some history $h \cup R$. Then there exists another contract $\widetilde{C}$ that makes the principal no better off, but has $\widetilde{y}_R(h \cup R) > y_R(h \cup R)$.*

Proof. To see this claim, assume, for a contradiction, that the principal's worst contract is such that the principal's enforcement constraint never binds, and there is a history h such that $y_R(h \cup R) < 1$. If there is such a history, then there also is a *first history* $\hat{h}$ that is, a history such that for any $h \subset \hat{h}$, either L is selected or $y_R(h \cup R) = 1$.

Now, consider a contract $\widetilde{C}$ that is identical to C but with $\tilde{y}_R(\hat{h} \cup R) = y_R(\hat{h} \cup R) + \varepsilon$, with $\varepsilon > 0$. We show that for some ε this contract is incentive compatible and distinguish between the following 3 cases:

(a) any history $h \subseteq \hat{h}$ has $k(h) = R$
(b) any history $h \subseteq \hat{h}$ such that $k(h) = L$ has slack on L's incentive constraint
(c) at least some history $h \subseteq \hat{h}$ such that $k(h) = L$ has a binding incentive constraint for L

Case (a). Incentive compatibility holds for R at $\hat{h} \cup R$ by construction. History $\hat{h}$ is a first history for $y_R(\cdot) < 1$ and $y_R(\cdot) = 1$ is R's best deviation. Thus for any $h \cup R \subset \hat{h}$, incentive compatibility for R holds by assumption. Because P's enforcement constraint does not bind and values are continuous, a $\varepsilon > 0$ exists such that incentive compatibility for P holds at all histories $h \subset \hat{h}$. Finally, the contract is unchanged for any history $h \not\subset \hat{h}$. Thus, incentive compatibility for all players holds by assumption. The last argument in particular includes any history in which L gets selected. However, the

ex-ante value of the principal is

$$v_P(h_0|\widetilde{C}) = v_P(h_0|C) - \delta(\hat{h} \cup R|h_0, C)\varepsilon, \tag{10}$$

which is strictly lower than that under contract C. Thus C is not the worst punishment.

Case (b). Incentive compatibility holds for R for the same reasons as in Case (a). It holds for P for ε small enough for the same reasons as well. Finally, it holds for L for ε small too, as L's incentive constraint had slack. Thus, as in case (a.) an $\varepsilon > 0$ small enough exists such that $\widetilde{C}$ is incentive compatible and (10) holds, implying that C is not a worst punishment.

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

Contract $\widetilde{C}'$ is incentive compatible for L at history $\check{h}$ by construction. 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 $\widetilde{C}'$ is incentive compatible for all players, provided ε is small.

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

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

so the principal is worse off under $\widetilde{C}'$. Yet, because the principal's preferences had been strict under the original contract C, there is an $\varepsilon > 0$ small enough such that $\widetilde{C}'$ is incentive compatible for the principal. If $y_L(\check{h}) > \theta$, the principal's value is:

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

making $\widetilde{C}'$ a worst punishment for the principal.

Thus, while C cannot be a worst punishment in the class discussed in Claim 1 if cases (a) and (b) apply, it may be if we are in case (c). However, even in that case, it is possible to increase y_R in the first history in which $y_R < 1$ without making the principal better off. This proves the claim. □

Claim 1 shows that dynamic enforcement of L and R 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 strictly worse that are played with positive probability. Our next claim shows that if contract C indeed was the worst punishment for the principal, then our operations do not violate the principal's enforcement constraint.

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

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

First, note that we can construct a contract $\tilde{C}$ as in the proof of Claim 1 with $\varepsilon = \hat{\varepsilon}$ such that L's and R's dynamic enforcement constraints 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 choices $\breve{h} \subset \hat{h}$.

Now, assume that $\breve{h} \subset \hat{h}$ under C and, in addition, that the constructed $\widetilde{C}$ under the required ε violates the principal's dynamic enforcement constraint at $\breve{h}$. That means the principal's continuation value $\breve{\alpha}_P(\breve{h}, \widetilde{C}) < \alpha_P$ where α_P is the principal's ex-ante value of C. If there are multiple such histories, take the 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 the principal's worst contract to begin with, which proves claim 2. □

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

Lemma 14. *The principal's worst punishment is characterized by a stationary contract $(s, y_R, y_L) = (-m, 1, \tilde{y})$ with $\tilde{y}$ either $0, 1$, or $\tilde{y} > 2\theta$ and such that L's dynamic enforcement constraint holds with equality.*

Proof. $y_R = 1$ follows from Lemma 13. Hence, no matter y_L, a selected R is always a worst outcome for P because $\theta < 1/2$. Thus, P always has an incentive to choose $s = -m$.

Finally, L chooses either $y_L = 0$, which is trivially incentive compatible, or, if this punishes the principal more and is incentive compatible, $y_L > 0$. Obviously, $y_L \notin (0, 2\theta)$ because then $y_L = 0$ is worse for P. If $y_L > 2\theta$, then the principal is worse off the higher y_L. Hence, we increase y_L to the largest value such that L's dynamic enforcement constraint is satisfied, or we hit the bound of the action space at 1.

Finally, if $y_L = 1$ was incentive compatible for some endorsement strategy s, it is incentive compatible for $s = -m$ ensuring that $s = -m$ is indeed part of a worst punishment. □

Describing the worst contract. Here we describe the principal's worst contract assuming the agent receives the same deviation payoff as under commitment, $\alpha_L = \alpha^C$.

Then the worst action $y_L \neq 0, L$ can be incentivized to select solves

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

which is

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

The principal is worse off under $y_L = \tilde{y}_L$ than under $y_L = 0$ if and only if

$$\theta < \frac{(b+1)m\beta}{1-(1/2-m)\beta} =: \tilde{\theta}$$

If, in addition, $b \geq \frac{1-(1/2+m)\beta}{2m\beta}$, the globally worst outcome for the principal, $y_i = 1$ irrespective of the selection, is implementable via an incentive compatible contract.

The ex-ante value of the principal's worst punishment is then

$$\alpha_P^D = \begin{cases} \frac{(1-2\theta)m-1/2}{1-\beta} & \text{if } \theta \geq \tilde{\theta} \\ \frac{2m(1-(2+b+m+2bm)\beta)+(2\theta-1/2)-(1-2m)\beta\theta-1/2(1-\beta)}{(1-\beta)(1-(1/2-m)\beta)} & \text{if } \theta < \tilde{\theta}. \end{cases}$$

A.7 Proof of Propsitions 7, 8, and 9

Proof. All 3 propositions directly follow from the characterization result. □

References

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

Abreu, D., D. Pearce, and E. Stacchetti (1990). "Toward a theory of discounted repeated games with imperfect monitoring". *Econometrica*, pp. 1041–1063.

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

Acemoglu, D. and J. A. Robinson (2008). "Persistence of power, elites, and institutions". *American Economic Review*, pp. 267–293.

Andrews, I. and D. Barron (2016). "The allocation of future business: Dynamic relational contracts with multiple agents". *American Economic Review*, pp. 2742–2759.

Anesi, V. and P. Buisseret (2022). "Making elections work: Accountability with selection and control". *American Economic Journal: Microeconomics*, pp. 616–44.

Baron, D. P. and D. Diermeier (2001). "Elections, governments, and parliaments in proportional representation systems". *The Quarterly Journal of Economics*, pp. 933–967.

Barron, D. and M. Powell (2019). "Policies in relational contracts". *American Economic Journal: Microeconomics*, pp. 228–49.

Board, S. (2011). "Relational contracts and the value of loyalty". *American Economic Review*, pp. 3349–3367.

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

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. (2022). "Persistence in Power of Long-Lived Parties". *Available at SSRN 4106431*.

Delgado-Vega, Á. (2023). "Which Side are You on? Interest Groups and Relational Contracts". *Working Paper*.

Dixit, A., G. M. Grossman, and F. Gul (2000). "The dynamics of political compromise". *Journal of Political Economy*, pp. 531–568.

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

Gertner, R. and D. Scharfstein (2012). "Internal Capital Markets". *The Handbook of Organizational Economics*. Princeton University Press, pp. 655–679.

Ginsborg, P. (1990). *A History of Contemporary Italy: 1943-80*. Penguin UK.

Green, E. J. and R. H. Porter (1984). "Noncooperative collusion under imperfect price information". *Econometrica: Journal of the Econometric Society*, pp. 87–100.

Invernizzi, G. M. and M. M. Ting (2023). "Institutions and Political Restraint". *American Journal of Political Science*.

Levin, J. (2002). "Multilateral contracting and the employment relationship". *The Quarterly Journal of Economics*, pp. 1075–1103.

— (2003). "Relational incentive contracts". *American Economic Review*, pp. 835–857.

Li, J., N. Matouschek, and M. Powell (2017). "Power dynamics in organizations". *American Economic Journal: Microeconomics*, pp. 217–241.

Lipnowski, E. and J. Ramos (2020). "Repeated delegation". *Journal of Economic Theory*, p. 105040.

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

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". *Working Paper*.

Nocke, V. and R. Strausz (2023). "Collective Brand Reputation". *Journal of Political Economy*, pp. 1–58.

Padró i Miquel, G. (2007). "The control of politicians in divided societies: The politics of fear". *The Review of Economic Studies*, pp. 1259–1274.

Rayo, L. (2007). "Relational incentives and moral hazard in teams". *The Review of Economic Studies*, pp. 937–963.

Shleifer, A. and R. W. Vishny (1989). "Management entrenchment: The case of manager-specific investments". *Journal of financial economics*, pp. 123–139.

Stein, J. C. (2003). "Agency, Information and Corporate Investment". *Corporate Finance*. Ed. by G. M. Constantinides, M. Harris, and R. M. Stulz. Handbook of the Economics of Finance. Elsevier, pp. 111–165.

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

Xuan, Y. (2009). "Empire-Building or Bridge-Building? Evidence from New CEOs Internal Capital Allocation Decisions". *The Review of Financial Studies*, pp. 4919–4948.

Yared, P. (2010). "Politicians, taxes and debt". *The Review of Economic Studies*, pp. 806–840.