Power Dynamics in Organizations*

Jin Li
Northwestern University

Niko Matouschek
Northwestern University

Michael Powell
Northwestern University

First version: December 2013
This version: March 2015

Abstract

We explore the evolution of power within organizations. To this end, we examine an infinitely repeated game in which a principal can empower an agent by letting him choose a project. The principal, however, does not know what projects are available to the agent. We characterize the optimal relational contract and explore its implications. Our results speak to how power is earned, lost, and retained. They show that entrenched power structures are consistent with managers who are managing power optimally. And they provide a new perspective on two long-standing issues in organizational economics.

Keywords: internal organization, relational contracting, power

JEL classifications: D23, D82, L23

*We are very grateful for comments from Alessandro Bonatti, Wouter Dessein, Matthias Fahn, Daniel Garrett, George Georgiadis, Bob Gibbons, Navin Kartik, Steve Martin, Patrick Rey, Mike Whinston, Yanhui Wu, and participants of various conferences and seminars. We thank Can Urgun for excellent research assistance. All remaining errors are our own. Li: Kellogg School of Management, Northwestern University, jin-li@kellogg.northwestern.edu; Matouschek: Kellogg School of Management, Northwestern University, n-matouschek@kellogg.northwestern.edu; Powell: Kellogg School of Management, Northwestern University, mike-powell@kellogg.northwestern.edu.

1 Introduction

The allocation of formal authority in a firm is carved in stone: it resides with those at the top of the hierarchy and cannot be delegated in a legally binding manner (see, for instance, Bolton and Dewatripont (2013)). Most firms, however, are not autocracies in which a central authority commands all decisions. Instead, those at the top routinely empower their subordinates by promising not to overrule their decisions or to at least exercise restraint in doing so (Baker, Gibbons, and Murphy 1999).

In contrast to formal authority, the allocation of this informal authority or "power" is fluid and changes over time (see, for instance, Pfeffer (1981)). Some of the observed dynamics are intuitive. Divisions, for example, often gain power during periods in which their products are particularly important for the overall profitability of their firms. Other dynamics, however, are more difficult to reconcile with an efficiency-based view of organizations. There are, for instance, many examples in which powerful divisions are able to hold on to their power, even when the peak of their importance has long passed and even when they are using it in ways that are clearly self-serving and harmful to the firm's profits.

Sears's catalog division, for example, was able to use the power it had gained during the heyday of mail ordering to delay its closure until the mid-90s, years after analysts started calling for its shut down (Schaefer 1998). More recently, observers blamed Microsoft's failure in mobile computing on its powerful Windows and Office divisions. As one news article reported (Eichenwald 2012):

"*Indeed, executives* [at Microsoft] *said, Microsoft failed repeatedly to jump on emerging technologies because of the company's fealty to Windows and Office. 'Windows was the god* [. . .] *Ideas about mobile computing with a user experience that was cleaner than with a P.C. were deemed unimportant by a few powerful people in that division, and they managed to kill the effort.'*"

In this paper we explore the evolution of power within firms and organizations more broadly. Our goal is to understand whether power can be earned, how it is lost, and why some are able to retain it even when they are using it in openly selfish ways. Existing economic theories of organization are not well-suited to explore these issues, since they are either static or focus on settings in which the optimal allocation of power is stationary (for the former see, for instance, Holmstrom (1984), Aghion and Tirole (1997), and Dessein (2002) and, for the latter, see Baker, Gibbons, and Murphy (1999) and Alonso and Matouschek (2007)). In this paper we therefore expand the existing literature by developing a dynamic model of power.

Our dynamic approach allows us to capture the notion that the prospect of more power tomorrow can motivate subordinates to make good use of whatever power they have today. Even though this notion of power as a form of payment or reward is absent in the economic literature on organizations, it has a long history in sociology and organizational theory. Cyert and March (1963), in particular, observed more than fifty years ago that payments within organizations often take the form of promises about future decisions and decision-making rather than monetary transfers.

As an illustration that will be familiar to many readers, take the dean of a college who wants to convince a department to make a spousal hire. It would be unusual for the dean to try to convince the department to make the hire by promising to pay its members more money. The dean may well, however, promise to reward the department by giving it more discretion in future hiring or to bias future school-wide decisions in its favor. In other words, the dean may pay the department members with power rather than in cash.

The problem with rewarding subordinates by promising them more power in the future is that they value such a promise precisely because it will allow them to bias future decisions in their favor. Paying with power therefore generates a dynamic trade-off between the current and future agency costs of power: it induces subordinates to make good use of whatever power they have today, but it does so only by allowing them to abuse their power in the future.

We show that if those at the top manage this trade-off optimally, then, initially, power is earned and lost in line with an agent's current performance: the agent becomes more powerful if he makes decisions that are good for the organization, and he becomes less powerful if he does not. Eventually, however, this link between power and performance is broken. Depending on the agent's initial performance, he then either loses his power permanently and is forced to accept whatever decisions those at the top now make in his stead, or his power becomes unchallengeable, and he is free to make whatever decisions he sees fit. In either case, the organization is no longer able to make efficient use of the agent's expertise, and its performance suffers.

Entrenched power structures, and the frustrations and apparent inefficiencies they entail, therefore need not reflect management failures. Instead, such structures may arise precisely because managers are managing power optimally, albeit in a second-best setting in which their ability to reward employees with money is constrained. These results shed light on dynamic aspects of life inside organizations that have so far received little formal attention. And in doing so, they provide new perspectives on why similar firms organize differently, even though those organizational differences are associated with differences in performance, and why established firms often have a harder time adapting to changes in their environments than their younger rivals, even when the need to adapt to those changes is widely understood. To explore these implications, it is useful to go beyond the broad sketch of intuitions we have given so far and describe our model and results in more detail.

To this end, consider our baseline model, which is an infinitely repeated game between a principal and her agent. In the stage game, both parties first decide whether to enter the relationship. If both do enter, the principal either chooses a status quo project herself or she empowers the agent to choose a project. If the principal does empower the agent, the agent then gets to choose between his preferred project and, if available, the principal's. The payoffs are such that each party prefers his or her preferred project to the status quo and the status quo to the other party's preferred project. The stage game is therefore essentially a trust game. In contrast to a standard trust game, however, the principal's preferred project may not be available, and its availability is only observed by the agent. If the agent does not choose the principal's preferred project, the principal therefore does not know whether he willingly abused his power or simply had no choice.

Under the optimal relational contract the principal starts out by empowering the agent, but she does so with the understanding that the agent chooses the principal's preferred project whenever it is available. To motivate the agent to use his power in this way, the principal rewards him for choosing her preferred project with an increase in his continuation payoff, and she punishes him for choosing his preferred project with a reduction in his continuation payoff. These changes in the agent's continuation payoffs capture his future power: the more power the agent is promised in the future, the more often he will be able to choose his preferred project, and the higher his continuation payoff is today.

This initial phase of the relationship continues until the continuation payoff crosses one of two thresholds. If it drops below the lower threshold—because the principal's preferred project was only available very rarely—the principal either takes away the agent's power forever, or she terminates the relationship altogether. If, instead, the continuation payoff rises above the upper threshold—because the principal's preferred project was available sufficiently often—the principal promises to empower the agent in all future periods, no matter what decisions he makes. The relationship therefore does not cycle between punishment and reward phases, as in Green and Porter (1984). Instead, the principal finds it optimal to delay punishments and rewards for as long as possible before administering them with maximum force. In the long run, an active relationship therefore ends up with one of two very different allocations of power: permanent disenfranchisement or permanent empowerment. Moreover, which allocation the organization ends up with is fully determined by random events during its early history.

These dynamics speak to the debate on why firms often find it difficult to change their organizations and struggle to imitate those of their more successful rivals (see Milgrom and Roberts

(1995), Brynjolfsson and Milgrom (2013), and Halac and Prat (2014)). The starting point of this debate is the claim that the competitive advantage of some successful firms, such as Lincoln Electric and Toyota, is based on how they are organized. This claim has recently been backed up by an emerging empirical literature which shows that similar firms organize differently and that those differences in organization translate into differences in performance (see Bloom, Eifert, McKenzie, Mahajan, and Roberts (2013) and, for a survey, Gibbons and Hendersen (2013)). In contrast to other sources of competitive advantage, however, a firm's organization is often well-known and typically not protected by patents. This raises the question of why under-performing firms don't simply imitate the organizations of their better-organized rivals.

Our model suggests that the answer to this question may lie in the firms' pasts. Specifically, the long-run dynamics we described above imply that firms that are initially identical can end up with different organizations in the long-run that generate different performance levels. These differences persist, not because firms are unable to observe each others' actions nor because those actions are protected by patents. Instead, they persist only because firms are constrained by their past promises. One firm, for instance, may be able to centralize decision-making without triggering resentment among its employees. In another, seemingly identical firm, however, employees may view their power as the reward for previous achievements and interpret centralization as the violation of past promises. As such, our model supports the intuitive view that, as long as there are relational aspects to a firm's organization, its history can serve as a hidden barrier to imitation.

Another key implication of the dynamics we described above is that the organization gradually gets worse at adapting to changes in the environment. Initially, the principal is able to induce the agent to adapt to those changes in a profit-maximizing manner by making promises about his future power. Eventually, however, the principal has to live up to those promises and either make the decisions herself, and thus forgo the agent's information, or allow the agent to abuse his power and bias decisions in his favor. In either case, the organization no longer adapts to changes in the environment in a profit-maximizing manner. These dynamics contrast with those in the many models on relationship building in which relationships improve over time, because the parties learn to cooperate and coordinate with each other (see, for instance, Chassang (2010) and Halac (2014)). In our setting, instead, the relationship deteriorates over time as it gets bogged down by the need to fulfill the very promises that ensured its success early on.

In our baseline model the changes in the environment that the organization fails to adapt to are privately observed by the agent. The organization's failure to adapt therefore cannot be detected by an outsider who is simply observing the organization's current decisions and instead has to be inferred from the equilibrium strategies and the organization's history. As such, our baseline model does not provide a satisfying explanation for why some firms fail to adapt to new opportunities even when those opportunities are publicly observable, and the failure to adapt is widely critiqued, as in our motivating examples.

To explore the failure to adapt to public opportunities, we build on our baseline model by allowing for a publicly observable and permanently available opportunity to arrive at some random time during the game. This new project generates a higher expected payoff for the principal than any expected payoff she can realize in its absence. And, once the project has arrived, the principal can simply choose it herself, without having to induce the agent to do so. Nevertheless, we show that if the project arrives late enough, the principal only adopts it with some delay and sometimes does not adopt it at all. The reason for this delay is that the principal now goes beyond promising to tolerate the agent's biased decision-making by pledging to distort her own decisions in his favor. Truly powerful agents therefore do not only get away with making selfish decisions. Their power extends to those at the top of the organization who are bound by their past promises to do the agents' bidding, even when doing so is known to hurt the organization overall.

These dynamics resonate with the observation that established firms often fail to respond to disruptive changes in their industries. Bower and Christensen (1995, p.43), in particular, observed that "*One of the most consistent patterns in business is the failure of leading companies to stay at the top of their industries when technologies or markets change*" and coined the term "disruptive innovation" to describe this phenomenon. Our model suggests that the very promises that allow firms to adapt to changes in their environments when they are young prevent them from doing so when they are old. The sluggishness of established firms, and the flexibility of their younger rivals, are then two sides of the same coin and both consistent with optimal management.

In summary, our model provides a simple theory of power dynamics that can answer the questions we started out with: yes, power can be earned, and it can be earned by making good use of whatever power one has today; it can, however, also be lost by making bad use of that power, even if that bad use is forced by an act of nature and is not a willing choice; and as for why some are able to retain their power even though they are using it in openly selfish ways, that too can be explained as the delayed reward for the selfless use of whatever power they had during the early days of the organization. These answers shed light on the evolution of power within organizations. And, in doing so, they provide new perspectives on two long-standing issues in the economics of organizations and management.

2 Literature Review

There is a large literature in sociology and organizational theory on power in organizations. The central question that this literature explores is why some members of organizations wield more power than others. A common answer is that power is held by those who are deemed to be particularly valuable to the organization. A member may be particularly valuable, for instance, because he controls important resources, as in the "resource dependence theory of power" (Emerson 1962, Pfeffer and Salancik 1978). Or he may be particularly valuable because he helps the organization deal with contingencies, as in the "strategic contingency theory of power" (Hickson, Hinings, Lee, Schneck, and Pennings, 1971).

The economics literature on power is closely related to, and overlaps with, the incomplete contracting literature on delegation. Most of this literature assumes that the principal can contractually commit to different allocations of decision rights and then explores the formal allocation of those decision rights (see, for instance, Aghion and Tirole (1997), Dessein (2002), and, for a survey, Bolton and Dewatripont (2013)). Courts, however, do not typically enforce contracts between different parties within the same organization (see, for instance, Aghion, Bloom, and Van Reenen (2013) and Bolton and Dewatripont (2013)). In line with this fact, a small number of papers explore the informal allocation of decision rights that arises if those at the top commit to different allocations through non-contractual means. In Aghion and Tirole (1997), for instance, the principal can commit to behaving as if formal authority had been delegated by becoming overloaded and thus staying uninformed. And in Baker, Gibbons, and Murphy (1999) and Alonso and Matouschek (2007) the principal can commit to behaving as if formal authority had been delegated by agreeing to a relational contract. We follow these last two papers in modeling power as a relational contract. In contrast to those papers, however, we explore a setting in which the optimal allocation of power is not stationary and explore how it changes over time.

Another aspect of our model that we share with those in the incomplete contracting literature on delegation is that we rule out monetary transfers. This assumption captures, albeit in a stark way, the view that the ability of members of an organization to exchange money is often limited by a variety of managerial and legal constraints. The literature on mechanism design without transfers takes this view as its starting point and then explores the optimal design of contracts when parties cannot exchange money but do not face any other constraints on the contracts they can write. In this literature, the papers closest to ours are Holmstrom (1984), Alonso and Matouschek (2008), Armstrong and Vickers (2010) and Amador and Bagwell (2013), who explore static environments, and Guo and Horner (2014), who explore a dynamic one.

Our paper is also related to the large literature that studies the economics of relationships; see Samuelson (2006) and Mailath and Samuelson (2006) for reviews. For relationships that survive in the long run, many papers show that their performance improves in general; see Ghosh and Ray (1996), Kranton (1996), Watson (1999, 2002), Mailath and Samuelson (2001), Chassang (2010), Yang (2013), and Halac (2014). Our focus is on the decline of performance among surviving relationships. One reason for relationships to decline is that the production environment worsens; see, for instance, McAdams (2011), Garrett and Pavan (2012), and Halac and Prat (2014). The production environment in our model is either stationary—in the baseline model—or improves over time—in the model with public opportunities.

Finally, our paper is related to the literature on dynamic games with one-sided private information; see Mailath and Samuelson (2006) for a general review and Malcomson (1999, 2013) for surveys of applications in labor and organizational economics. In these types of models, the parties use changes in the agent's continuation payoff to provide incentives. The long-run dynamics then depend on how these continuation payoffs are eventually delivered to the agent. In some models, long-run dynamics involve either termination of the relationship or convergence to an efficient steady state; see, for example, Clementi and Hopenhayn (2006), Biais, Marriotti, Plantin, and Rochet (2007), and DeMarzo and Fishman (2007). In other models, punishment is temporary, and relationships forever cycle between punishment and reward phases; see, for example, Padro i Miquel and Yared (2012), Fong and Li (2013), Li and Matouschek (2013), and Zhu (2013). In all these models, the average long-run performance of firms that do not exit is identical. In our model, instead, firms settle into either permanent disenfranchisement or permanent empowerment. As a result, firms that do not exit in the long-run can experience persistent organizational and performance differences.

3 The Model

A risk-neutral principal and a risk-neutral agent are in an infinitely repeated relationship. Time is discrete and denoted by $t = 1, 2, ...$ We first describe the stage game and then move on to the repeated game. In the description of the stage game we omit time subscripts for convenience.

The Stage Game At the beginning of the stage game, the principal and the agent simultaneously decide whether to enter the relationship. We denote their entry decisions by $d_j \in \{0, 1\}$ for $j = P, A$, where $d_j = 1$ denotes the decision to enter. If at least one party decides not to enter, both realize a zero payoff and time moves on to the next period.

If, instead, both parties do decide to enter, the principal next decides whether to let the agent choose a project. We refer to the principal's decision as her empowerment decision and denote it by $e \in \{0, 1\}$, where $e = 1$ denotes empowerment. In line with our discussion in the introduction and literature review, we use the term "empowerment" rather than "delegation" since the principal's decision is non-contractible. Moreover, we denote both projects and project choices by k and the principal's and the agent's stage-game payoffs, conditional on both parties having entered the relationship, by $\Pi(k)$ and $U(k)$.

If the principal decides not to empower the agent, she chooses a safe project $k = S$ that generates payoffs $\Pi(S) = U(S) = a > 0$. If, instead, the principal does empower the agent, the agent can choose between up to two projects. One of these projects is the agent's preferred project $k = A$, and the other is the principal's preferred project $k = P$. The agent's preferred project gives the agent a payoff $U(A) = B$ and the principal a payoff $\Pi(A) = b$, where $B > a > b > 0$. Analogously, the principal's preferred project gives the principal a payoff $\Pi(P) = B$ and the agent a payoff $U(P) = b$. Empowerment therefore takes the form of a trust game in which the principal only benefits from empowerment if she can trust the agent to choose her preferred project. The assumption that payoffs are symmetric facilitates the exposition, but it is not important for our results.

The key feature of the game is that the principal's preferred project is not always available and that only the agent can observe whether it is available. If the agent chooses his preferred project, the principal therefore cannot observe whether the agent abused his power–by choosing his preferred project even though the principal's was available–or simply had no choice. In particular, the principal's preferred project is only available with probability $p \in (0, 1)$, where the availability is independent across periods. Other than the availability of the principal's preferred project, all information is publicly observable.

Finally, after the parties have realized their payoffs, they observe the realization ω of a public randomization device, after which time moves on to the next period. Figure 1 summarizes the timing, payoffs, and information structure of the stage game.

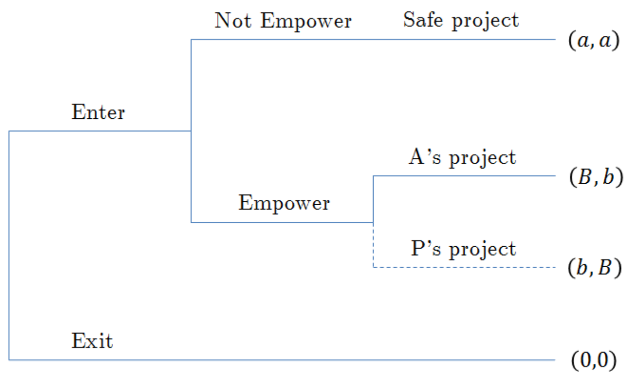

Figure 1: This figure illustrates the timing, payoffs, and information structure of the stage game. The first payoff listed is the agent's; the second is the principal's. The dashed line represents the agent's private information about the availability of the principal's preferred project.

The Repeated Game The principal and the agent have a common discount factor $\delta \in (0, 1)$. At the beginning of any period t, the principal's expected payoff is therefore given by

$$\pi_t = (1 - \delta) \, \mathrm{E}_t \left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} d_{P,t} d_{A,t} \Pi \left(k_t \right) \right],$$

and the agent's expected payoff is given by

$$u_t = (1 - \delta) \, \mathrm{E}_t \left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} d_{P,t} d_{A,t} U \left(k_t \right) \right].$$

Note that we multiply the right-hand side of each expression by $(1 - \delta)$ to express payoffs as per-period averages.

We follow the literature on repeated games with imperfect public monitoring and define a relational contract as a pure-strategy Perfect Public Equilibrium (henceforth PPE) in which the principal and the agent play public strategies and, following every history, the strategies are a Nash Equilibrium of the continuation game. Public strategies are strategies in which the players condition their actions only on publicly available information. In particular, the agent's strategy does not depend on her past private information. Our restriction to pure strategies is without loss of generality, because our game has only one-sided private information and is therefore a game with a product monitoring structure. In this case, there is no need to consider private strategies since every sequential equilibrium outcome is also a PPE outcome (see, for instance, p.330 in Mailath and Samuelson (2006)).

Formally, let $h_{t+1} = \{d_{P,\tau}, d_{A,\tau}, e_\tau, k_\tau, \omega_\tau\}_{\tau=1}^{t}$ denote the public history at the beginning of any period $t+1$, and let H_{t+1} denote the set of period $t+1$ public histories. Note that $H_1 = \emptyset$. A public strategy for the principal is a sequence of functions $\{D_{P,t}, E_t, K_{P,t}\}_{t=1}^{\infty}$, where $D_{P,t} : H_t \to \{0,1\}$, $E_t : H_t \cup \{d_{P,t}, d_{A,t}\} \to \{0,1\}$, $K_{P,t} : H_t \cup \{d_{P,t}, d_{A,t}, e_t\} \to \mathcal{K}_P$, and where $\mathcal{K}_P = \{S\}$ is the set of projects available to the principal. Similarly, a public strategy for the agent is a sequence of functions $\{D_{A,t}, K_{A,t}\}_{t=1}^{\infty}$, where $D_{A,t} : H_t \to \{0,1\}$ and $K_t : H_t \cup \{d_{P,t}, d_{A,t}, e_t\} \to \mathcal{K}_{A,t}$, and where $\mathcal{K}_{A,t} \in \{\{A\}, \{A, P\}\}$ is the set of projects available to the agent.

We define an "optimal relational contract" as a PPE that maximizes the principal's first-period equilibrium payoff. Our goal is to characterize the set of optimal relational contracts.

4 Preliminaries

In this section, we follow the approach pioneered by Abreu, Pearce, and Stacchetti (1990) to characterize the PPE payoff set. Each equilibrium payoff pair (u, π) can be supported either by randomization among several equilibrium payoff pairs or by a pure action of the stage game and a pair of continuation payoffs associated with each public outcome. Public randomization occurs at the end of the period; at the beginning of each period, the players therefore select a pure action of the stage game, receive the flow payoffs generated by that pure action, and expect to receive particular continuation payoffs. We begin our analysis in Section 4.1 by describing the constraints that have to be satisfied for an equilibrium payoff pair to be supported by a particular action and a particular set of continuation payoffs. We then characterize the PPE payoff set as the solution to a functional equation, which we describe in Section 4.2. Given our explicit characterization of the PPE payoff set, we then describe the dynamics of an optimal relational contract in Section 5.

4.1 The Constraints

We denote the PPE payoff set by $\mathcal{E}$. Any payoff pair $(u, \pi) \in \mathcal{E}$ is either generated by pure actions or by randomization among equilibrium payoff pairs that are each generated by pure actions. There are four classes of pure actions. First, both parties enter the relationship, after which the principal empowers the agent with the understanding that he chooses the principal's preferred project whenever it is available. We call this set of actions "cooperative empowerment" and denote it by E_C. Second, both parties enter the relationship, after which the principal empowers the agent with the understanding that he can always choose his preferred project. We call this action "uncooperative empowerment" and denote it by E_U. Third, both parties enter the relationship, after which the principal centralizes and chooses the safe project. We call this action "centralization" and denote it by C. Finally, neither party enters the relationship. We call this set of actions "exit" and denote it by X. In the remainder of this section, we first discuss the constraints that have to be satisfied for a payoff pair $(u, \pi) \in \mathcal{E}$ to be generated by one of these four sets of pure actions. We then conclude the section by stating the constraint that has to be satisfied if the payoff pair is generated by randomization.

Centralization A payoff pair (u, π) can be supported by centralization if the following constraints are satisfied.

(i.) Feasibility: For the continuation payoffs to be feasible, they also need to be PPE payoffs. The continuation payoffs u_C and π_C that the parties realize under centralization therefore have to satisfy the self-enforcement constraint

$$(u_C, \pi_C) \in \mathcal{E}. \tag{SE$_\mathrm{C}$}$$

(ii.) No Deviation: To ensure that neither party deviates, we need to consider both off- and on-schedule deviations. Off-schedule deviations are deviations that both parties can observe. There is no loss of generality in assuming that if an off-schedule deviation occurs, the parties terminate the relationship by opting out in all future periods, as this is the worst possible equilibrium and gives each party its minmax payoff.

The principal and the agent can deviate off schedule by opting out of the relationship. If either party does so, he or she realizes a zero payoff this period and in all future periods. Since the payoffs from the three projects are strictly positive, the parties therefore do not have an incentive to deviate off schedule by opting out of the relationship.

The principal could also deviate off schedule by empowering the agent. There is no loss of generality in assuming that the agent will then choose his preferred project. By deviating, the principal would therefore reduce her current payoff from a to $b < a$, after which she would make a zero payoff in all future periods. The principal therefore never wants to deviate off schedule by empowering the agent.

On-schedule deviations are deviations that are privately observed. Since the principal does not have any private information, and the agent does not get to choose a project, there are no on-schedule deviations in the case of centralization.

(iii.) Promise Keeping: Finally, the consistency of the PPE payoff decomposition requires that the parties' payoffs are equal to the weighted sum of current and future payoffs. The promise-keeping constraints

$$\pi = (1 - \delta)\, a + \delta \pi_C \tag{PK$_\mathrm{C}^\mathrm{P}$}$$

and

$$u = (1 - \delta)\, a + \delta u_C \tag{PK$_\mathrm{C}^\mathrm{A}$}$$

ensure that this is the case.

Cooperative Empowerment A payoff pair (u, π) can be supported by cooperative empowerment if the following constraints are satisfied.

(i.) Feasibility: For the continuation payoffs to be feasible, they also need to be PPE payoffs. Let (u_ℓ, π_ℓ) denote the parties' continuation payoffs if the agent chooses his preferred project, and let (u_h, π_h) denote their payoffs if he chooses the principal's preferred project. The self-enforcement constraint is then given by

$$(u_h, \pi_h), (u_\ell, \pi_\ell) \in \mathcal{E}. \tag{$\mathrm{SE}_{\mathrm{E_C}}$}$$

(ii.) No Deviation: As in the case of centralization, the principal and the agent never want to deviate off schedule by opting out of the relationship, since doing so gives them a zero payoff. The principal may, however, want to deviate off schedule by not empowering the agent, in which case she realizes payoff a this period and a zero payoff in all future periods. To ensure that she does not want to do so, the no-reneging constraint

$$p\left[(1-\delta)B + \delta\pi_h\right] + (1-p)\left[(1-\delta)b + \delta\pi_l\right] \geq (1-\delta)a \tag{$\mathrm{NR}_{\mathrm{E_C}}$}$$

has to be satisfied.

Since the principal does not have any private information, she cannot engage in any on-schedule deviations. The agent, however, may choose his preferred project when the principal's preferred project is available. To ensure that he does not want to do so, the incentive constraint

$$(1-\delta)b + \delta u_h \geq (1-\delta)B + \delta u_\ell \tag{$\mathrm{IC}_{\mathrm{E_C}}$}$$

has to be satisfied.

(iii.) Promise Keeping: The promise-keeping constraints are now given by

$$\pi = p\left[(1-\delta)B + \delta\pi_h\right] + (1-p)\left[(1-\delta)b + \delta\pi_\ell\right] \tag{$\mathrm{PK}^{\mathrm{P}}_{\mathrm{E_C}}$}$$

and

$$u = p\left[(1-\delta)b + \delta u_h\right] + (1-p)\left[(1-\delta)B + \delta u_\ell\right]. \tag{$\mathrm{PK}^{\mathrm{A}}_{\mathrm{E_C}}$}$$

Uncooperative Empowerment A payoff pair (u, π) can be supported by uncooperative empowerment if the following constraints are satisfied.

(i.) Feasibility: We denote continuation payoffs under uncooperative empowerment by (u_{E_U}, π_{E_U}). The self-enforcement constraint is then given by

$$(u_{E_U}, \pi_{E_U}) \in \mathcal{E}. \tag{$\mathrm{SE}_{\mathrm{E_U}}$}$$

(ii.) No Deviation: As in the case of cooperative empowerment, the principal and the agent never want to deviate off schedule by opting out of the relationship, since doing so gives them a zero payoff both this period and in all future periods. The principal may, however, want to deviate off schedule by not empowering the agent. To ensure that she does not want to do so, the no-reneging constraint

$$(1-\delta)\, b + \delta \pi_{E_U} \geq (1-\delta)\, a \qquad (\mathrm{NR}_{\mathrm{E_U}})$$

has to be satisfied.

(iii.) Promise Keeping: The promise-keeping constraints are now given by

$$\pi = (1-\delta)\, b + \delta \pi_{E_U} \qquad (\mathrm{PK}^{\mathrm{P}}_{\mathrm{E_U}})$$

for the principal and

$$u = (1-\delta)\, B + \delta u_{E_U} \qquad (\mathrm{PK}^{\mathrm{A}}_{\mathrm{E_U}})$$

for the agent.

Exit A payoff pair (u, π) can be supported by exit if the following constraints are satisfied.

(i.) Feasibility: We denote the continuation payoffs under exit by (u_X, π_X). The self-enforcement constraint is then given by

$$(u_X, \pi_X) \in \mathcal{E}. \qquad (\mathrm{SE}_{\mathrm{Exit}})$$

(ii.) No Deviation: The principal and the agent can deviate off schedule by entering the relationship. If the principal or the agent does so, he or she realizes a zero payoff this period and in all future periods. The parties therefore do not have an incentive to deviate by entering the relationship. There are no other off- or on-schedule deviations in this case.

(iii.) Promise Keeping: The promise-keeping constraints are now given by

$$\pi = \delta \pi_X \qquad (\mathrm{PK}^{\mathrm{P}}_{\mathrm{X}})$$

for the principal and

$$u = \delta u_X \qquad (\mathrm{PK}^{\mathrm{A}}_{\mathrm{X}})$$

for the agent.

Randomization Finally, a payoff pair (u, π) can be supported by randomization. In this case, there exist at most three distinct PPE payoffs $(u_i, \pi_i) \in \mathcal{E}$, $i = 1, 2, 3$ such that

$$(u, \pi) = \alpha_1 (u_1, \pi_1) + \alpha_2 (u_2, \pi_2) + \alpha_3 (u_3, \pi_3)$$

for some $\alpha_1, \alpha_2, \alpha_3 \geq 0$ and $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

Assumptions on Parameters We now make three assumptions on the parameters of the model.

ASSUMPTIONS. *(i.)* $\delta/(1-\delta) \geq (a-b)/b$, *(ii.)* $\delta/(1-\delta) \geq (B-b)/(B-a)$, *and (iii.)* $p < 1/2$.

The first assumption ensures that the principal's no-reneging constraint can be satisfied under uncooperative empowerment. The second assumption guarantees that there exist feasible pairs of continuation payoffs u_ℓ and u_h that are sufficiently spread apart so that the agent's incentive constraint is satisfied. This assumption therefore ensures that the parties can make use of cooperative empowerment in equilibrium. We make the third assumption for expositional convenience and relegate the case in which it does not hold to the appendix.

4.2 The Constrained Maximization Problem

We now use the techniques developed by Abreu, Pearce, and Stacchetti (1990) to characterize the PPE payoff set and, in particular, its frontier. For this purpose, we define the payoff frontier as

$$\pi(u) \equiv \sup\left\{\pi' : \left(u, \pi'\right) \in \mathcal{E}\right\},$$

where $\mathcal{E}$ is the PPE payoff set. We also define

$$\underline{u} \equiv \inf\{u : (u, \pi) \in \mathcal{E}\}$$

and

$$\overline{u} \equiv \sup\{u : (u, \pi) \in \mathcal{E}\}$$

as the smallest and the largest PPE payoff for the agent.

We can now state our first lemma, which establishes several properties of the PPE payoff set. The proofs of this lemma and all other results are in the appendices.

LEMMA 1. *The PPE payoff set $\mathcal{E}$ has the following properties: (i.) it is compact; (ii.) $\pi(u)$ is concave; (iii.) the payoff pair (u, π) belongs to $\mathcal{E}$ if and only if $u \in [0, B]$ and $\pi \in [bu/B, \pi(u)]$.*

The first part of the lemma shows that the PPE payoff set is compact. This result immediately follows from the assumption that there are only a finite number of actions. It implies that for any $u \in [\underline{u}, \overline{u}]$ the payoff pair $(u, \pi(u))$ is in the PPE payoff set. The second part of the lemma shows that the payoff frontier is concave, which follows directly from the availability of a public randomization device. Finally, the third part shows that the agent's smallest and largest PPE payoffs are zero and B. It also shows that, for any $u \in [0, B]$, the principal's smallest PPE payoff is bu/B and that, for any $\pi \in [bu/B, \pi(u)]$, the payoff pair (u, π) is in the PPE payoff set.

A key implication of the first lemma is that the PPE payoff set is fully characterized by its frontier. To characterize the frontier, we need to determine, for each $(u, \pi(u)) \in \mathcal{E}$, whether it is supported by a pure action $j \in \{C, E_C, E_U, X\}$ or by randomization. Moreover, if it is supported by a pure action j, we need to specify the associated continuation payoffs. The next lemma characterizes the principal's continuation payoff for any of the agent's continuation payoffs, regardless of the actions that the parties take.

LEMMA 2. *For any $(u, \pi(u))$, the continuation payoffs are also on the frontier.*

The lemma shows that payoffs on the frontier are sequentially optimal. This is the case since the principal's actions are publicly observable. It is therefore not necessary to punish her by moving below the PPE frontier. This feature of our model is similar to Spear and Srivastava (1987) and the first part of Levin (2003), in which the principal's actions are also publicly observable. In contrast, joint punishments are necessary when multiple parties have private information as, for instance, in Green and Porter (1984), Athey and Bagwell (2001), and the second part of Levin (2003).

Having characterized the principal's continuation payoff for any of the agent's continuation payoffs in the previous lemma, we now state the agent's continuation payoffs associated with each action in the next lemma.

LEMMA 3. *For any payoff pair $(u, \pi(u))$ on the frontier, the agent's continuation payoffs satisfy the following conditions:*

(*i.*) *If the payoff pair is supported by centralization, the agent's continuation payoff satisfies*

$$\delta u_C(u) = u - (1-\delta)\, a.$$

(*ii.*) *If the payoff pair is supported by cooperative empowerment, the agent's continuation payoff can be chosen to satisfy*

$$\delta u_\ell(u) = u - (1-\delta)\, B$$

and

$$\delta u_h(u) = u - (1-\delta)\, b.$$

(*iii.*) *If the payoff pair is supported by uncooperative empowerment, the agent's continuation payoff satisfies*

$$\delta u_{E_U}(u) = u - (1-\delta)\, B.$$

(*iv.*) *If the payoff pair is supported by exit, the agent's continuation payoff satisfies*

$$\delta u_X(u) = u.$$

In the cases of centralization, uncooperative empowerment, and exit, the agent's continuation payoffs follow directly from the promise-keeping constraints $\mathrm{PK}^{\mathrm{A}}_{\mathrm{C}}$, $\mathrm{PK}^{\mathrm{A}}_{\mathrm{E_U}}$, and $\mathrm{PK}^{\mathrm{A}}_{\mathrm{X}}$. In the case of cooperative empowerment, instead, the agent's continuation payoffs follow directly from combining the promise-keeping constraints with the agent's incentive constraint $\mathrm{IC}_{\mathrm{E_C}}$, where we take the incentive constraint to be binding. To see that we can do so, suppose that the incentive constraint is not binding. We can then reduce u_h and increase u_ℓ in such a way that u remains the same, and all the relevant constraints continue to be satisfied. Since the PPE frontier is concave, such a change makes the principal weakly better off.

Next we use Lemmas 2 and 3 to provide expressions for the principal's expected payoff for a given action and a given expected payoff for the agent. For this purpose, let $\pi_j(u)$ for $j \in \{C, E_C, E_U, X\}$ be the highest equilibrium payoff to the principal given action j and agent's payoff u. We then have

$$\pi_C(u) = (1-\delta)\,a + \delta\pi(u_C(u)),$$

$$\pi_{E_C}(u) = p\left[(1-\delta)\,B + \delta\pi(u_h(u))\right] + (1-p)\left[(1-\delta)\,b + \delta\pi(u_\ell(u))\right],$$

$$\pi_{E_U}(u) = (1-\delta)\,b + \delta\pi(u_{E_U}(u)),$$

and

$$\pi_X(u) = \delta\pi(u_X(u)).$$

We can now state the next lemma, which describes the constrained maximization problem that characterizes the payoff frontier.

LEMMA 4: *The PPE frontier $\pi(u)$ is the unique function that solves the following problem. For all $u \in [0, B]$*

$$\pi(u) = \max_{\alpha_j \geq 0, u_j \in [0,B]} \sum_{j \in \{C, E_C, E_U, X\}} \alpha_j \pi_j(u_j)$$

such that

$$\sum_{j \in \{C, E_C, E_U, X\}} \alpha_j = 1$$

and

$$\sum_{j \in \{C, E_C, E_U, X\}} \alpha_j u_j = u.$$

The lemma shows that any payoff pair on the frontier is generated either by a pure action j—in which case the weight α_j is equal to one—or by randomization—in which case α_j is less than one. We obtain the frontier by choosing the weights optimally. Notice that the frontier can be thought of as a fixed point to an operator that maps bounded functions to bounded functions through the constrained maximization problem. We show in the appendix that the fixed point is unique even though the operator is not a contraction mapping.

5 The Optimal Relational Contract

In this section we characterize the optimal relational contract, that is, the PPE that maximizes the principal's expected payoff. For this purpose, we first characterize the payoff frontier by solving the constrained maximization problem in Lemma 4.

LEMMA 5. *There exist two cut-off levels* $\underline{u}_{CE} \in (a, \delta a + (1-\delta) B)$ *and* $\bar{u}_{CE} = (1-\delta) b + \delta B$ *such that the PPE payoff frontier* $\pi(u)$ *is divided into four regions:*

(*i.*) *For* $u \in [0, a]$, $\pi(u) = u$ *and* $(u, \pi(u))$ *is supported by randomization between exit and centralization.*

(*ii.*) *For* $u \in [a, \underline{u}_{CE}]$, $\pi(u) = ((\underline{u}_{CE} - u) a + (u - a) \pi(\underline{u}_{CE})) / (\underline{u}_{CE} - a)$ *and* $(u, \pi(u))$ *is supported by randomization between centralization and cooperative empowerment.*

(*iii.*) *For* $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$, $\pi(u) = \pi_{E_C}(u)$ *and* $(u, \pi(u))$ *is supported by cooperative empowerment.*

(*iv.*) *For* $u \in [\bar{u}_{CE}, B]$, $\pi(u) = ((B - u) \pi(\bar{u}_{CE}) + (u - \bar{u}_{CE}) b) / (B - \bar{u}_{CE})$ *and* $(u, \pi(u))$ *is supported by randomization between cooperative and uncooperative empowerment.*

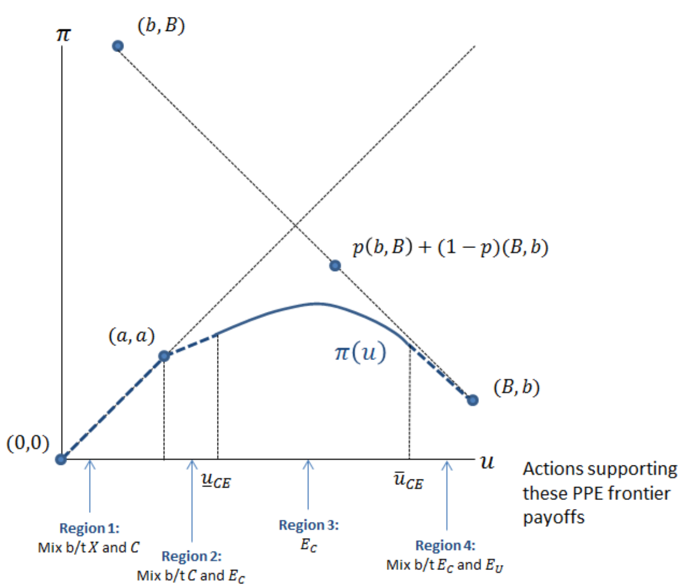

Figure 2: This figure illustrates the feasible stage-game payoffs, the PPE payoff frontier, and the actions that support each point on the frontier. The dotted linear segments are supported by public randomization between their two endpoints, and this public randomization occurs at the end of the period.

We illustrate the lemma in Figure 2. The lemma shows that the payoff frontier is divided into four regions. In three of these four regions, payoffs are supported by randomization and, as a result, the payoff frontier is linear. In any such region, payoffs can be supported by multiple types of randomization. Since for all such randomizations, payoffs end up at one of the endpoints of the region eventually, we assume that the parties randomize between the endpoints immediately. In the remaining region, payoffs are supported by pure actions, and the payoff frontier is concave.

We can now describe the optimal relational contract and how it evolves over time.

PROPOSITION 1. *First period: The agent's and the principal's payoffs are given by* $u^* \in [\underline{u}_{CE}, \bar{u}_{CE}]$ *and* $\pi(u^*) = \pi_{E_C}(u^*)$. *The parties engage in cooperative empowerment. If the agent chooses the principal's preferred project, his continuation payoff is given by*

$$u_h(u^*) = (u^* - (1-\delta)\, b)/\delta > u^*.$$

If, instead, the agent chooses his own preferred project, his continuation payoff is given by

$$u_\ell(u^*) = (u^* - (1-\delta)\, B)/\delta < u^*.$$

Subsequent periods: The agent's and the principal's payoffs are given by $u \in \{0, a\} \cup [\underline{u}_{CE}, \bar{u}_{CE}] \cup \{B\}$ *and* $\pi(u)$. *Their actions and continuation payoffs depend on which region* u *is in:*

(i.) If $u = 0$, *the parties exit. The agent's continuation payoff is given by* $u_X(0) = 0$.

(ii.) If $u = a$, *the parties engage in centralization. The agent's continuation payoff is given by* $u_C(a) = a$.

(iii.) If $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$, *the parties engage in cooperative empowerment. If the agent chooses the principal's preferred project, his continuation payoff is given by* $u_h(u) > u$. *If, instead, the agent chooses his own preferred project, his continuation payoff is given by* $u_\ell(u) < u$.

(iv.) If $u = B$, *the parties engage in uncooperative empowerment. The agent's continuation payoff is given by* $u_{E_U}(B) = B$.

The proposition shows that the principal starts out by engaging in cooperative empowerment. To motivate the agent to choose her preferred project whenever it is available, the principal increases his continuation payoff whenever he chooses her preferred project, and she decreases his continuation payoff whenever he does not.

To see how the principal optimally increases the agent's continuation payoff, suppose the agent chooses the principal's preferred project for a number of consecutive periods. The principal then continues to engage in cooperative empowerment, and the agent's continuation payoff continues to increase, until the parties reach a period in which the continuation payoff passes the threshold $\bar{u}_{CE}$.

At the end of that period, the parties engage in randomization to determine their actions in the following period. Depending on the outcome of this randomization, the principal either continues to engage in cooperative empowerment, or she moves to uncooperative empowerment. Finally, once the principal has moved to uncooperative empowerment, she stays there in all subsequent periods. We refer to this steady state as "permanent empowerment."

To see how the principal optimally decreases the agent's continuation payoff, suppose instead that the agent chooses his own preferred project for a number of consecutive periods. The principal then continues to engage in cooperative empowerment, and the agent's continuation payoff continues to decrease, until the parties reach a period in which the continuation payoff falls below the threshold $\underline{u}_{CE}$. At the end of that period, the parties engage in one of two types of randomization to determine their actions in the following period. If $u \in [a, \underline{u}_{CE}]$, the principal either continues to engage in cooperative empowerment, or she moves to centralization. And if, instead, $u \in [0, a)$, the principal either moves to centralization or she exits the relationship in the next period. Finally, once the principal has moved to either centralization or exit, she stays there in all subsequent periods. We refer to these steady states as "permanent disenfranchisement" and "termination."

The above proposition leaves open two questions about the long-run outcome of the relationship. First, does the principal always end up administering a punishment or reward? And if she ever does administer a punishment, does it take the form of termination or permanent disenfranchisement? The next proposition answers these questions.

PROPOSITION 2. *In the optimal relational contract, the principal chooses cooperative empowerment for the first τ periods, where τ is random and finite with probability one. Moreover, there exists a threshold p^* such that the relationship never terminates if $p \leq p^*$. If, instead, $p > p^*$, punishment can take the form of either termination or permanent disenfranchisement, depending on the history of the relationship.*

The proposition shows that the answer to the first question—whether the principal always ends up administering a punishment or reward—is yes. And it shows that the answer to the second question—whether the punishment takes the form of termination or permanent disenfranchisement—is that it depends on the probability p that the principal's preferred project is available. Having characterized the optimal relational contract, we now turn to its implications, which we already sketched and discussed in the introduction.

The first implication is that the allocation of power the organization converges to, and thus the long-run payoff the principal realizes, depend on random events in the organization's early history. In particular, whether the organization converges to permanent disenfranchisement—in which case the principal's payoff is given by a—or whether it converges to permanent empowerment—in which case it is given by $b < a$—depends on the randomly determined availability of projects in the periods before the organization converges to either steady state. This suggests that persistent differences in the organization of firms, and the persistent differences in performance that they generate, may be due to random differences in the early history of those firms.

Also, and related, the model suggests an explanation for why some under-performing firms do not copy the organizational practices of their more successful rivals, even though such practices are not protected by patents. In particular, it suggests that such firms may not imitate their more-successful rivals since their seemingly inefficient organizations are either a reward for past successes or a punishment for past failures. In either case, employees would view the adoption of a different organizational structure as the violation of a mutual understanding and punish the firm accordingly. A firm's history can therefore serve as a barrier to organizational imitation.

The second implication is that the principal's ability to make use of the agent's information, and thus the payoff she is able to realize, declines over time and that it does so even if the relationship does not terminate. In particular, the principal's first-period payoff $\pi(u^*)$ is strictly larger than the payoffs that the principal realizes once the relationship has converged to permanent disenfranchisement—in which case the principal makes $a < \pi(u^*)$—or permanent empowerment—in which case she makes $b < \pi(u^*)$. Notice that the result that the principal's payoff declines over time does not follow simply from the fact that we are focusing on an optimal relational contract that maximizes the principal's equilibrium payoff. An optimal relational contract could, in principle, require the parties to cycle between punishment and reward phases as it does in Padro i Miquel and Yared (2012), Fong and Li (2013), Li and Matouschek (2013), and Zhu (2013), who study games similar to ours, and as in the famous class of equilibria that Green and Porter (1984) focus on.

To see why such cycling is not optimal in our setting, notice that both rewards—letting the agent choose his preferred project even when the principal's is available—and punishments—opting out or centralizing—are costly for the principal. The threat to retract a previously promised reward, and the promise to retract a previously threatened punishment, however, do not impose any costs on the principal, yet they motivate the agent just the same. Delaying rewards and punishments therefore creates an additional and costless tool that the principal can use to motivate the agent. Because of this benefit, the principal wants to delay them as much as she can.

The fact that the relationship gets worse over time raises the question of why the principal does not eventually renegotiate the relational contract with the agent or replace him altogether. There is not yet a consensus on how to model renegotiation in repeated games (see Miller and Watson (2013) and Safronov and Strulovici (2014) for recent work). We therefore follow the literature on dynamic mechanism design and rule out renegotiation by assumption (see, for instance, Guo and Horner (2014)). We have also already ruled out replacement by assuming that there is only a single agent. This assumption captures the notion that the agent is special and generates relationship-specific surplus, which is an essential feature of economic models of relationships.

6 The Failure to Exploit Public Opportunities

A key feature of our baseline model is that the changes in the environment that the organization fails to adapt to are privately observed by the agent. The model therefore cannot explain why some firms fail to take advantage of opportunities that can be identified without any special expertise and that are apparent to those at the top of the firms' hierarchies and even to the wider public, as in our motivating examples.

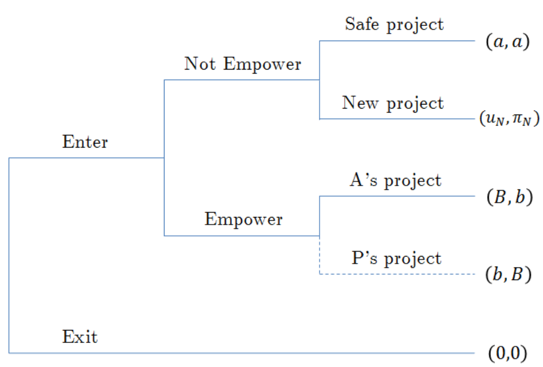

Figure 3: This figure illustrates the timing, payoffs, and information structure of the stage game in the post-opportunity phase. The first payoff listed is the agent's; the second is the principal's.

To address this issue, we now divide the periods into a pre-opportunity phase and a post-opportunity phase. The only difference between the stage game in the post-opportunity phase and the one in our baseline model is that the principal can now choose between two projects: the safe project and a new project that gives the principal a payoff π_N and the agent a payoff u_N. The availability of the new project is publicly observable. We summarize the payoffs from the stage game in the post-opportunity phase in Figure 3.

The stage game in the pre-opportunity phase is also very similar to the one in our baseline model. The only difference is that at the end of the stage game in period t, just before the realization of the public randomization device, nature determines whether the stage game in $t+1$ will again be in the pre-opportunity phase or be the first in the post-opportunity phase. The probability that the game transitions to the post-opportunity phase is given by $q \in (0,1)$, and whether this transition occurs is independent across periods.

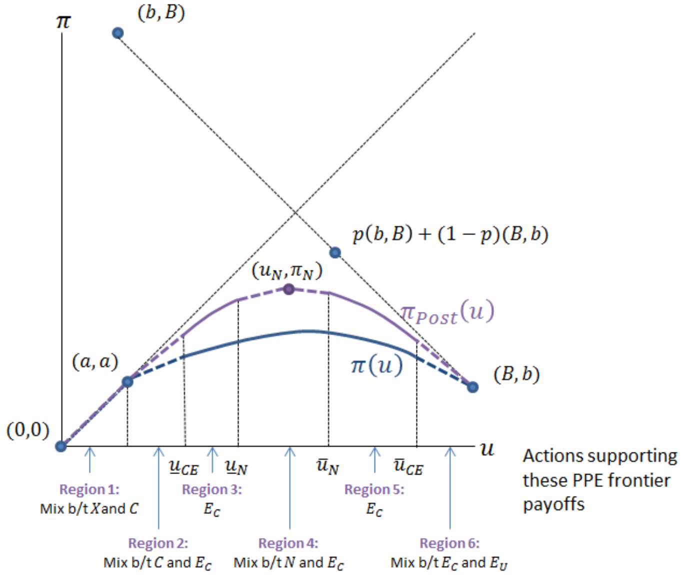

Figure 4: This figure illustrates the PPE payoff frontier for the baseline model and the PPE payoff frontier and the actions that support each point on the frontier for the post-opportunity phase. The dotted segments are supported by public randomization between their endpoints, and this public randomization occurs at the end of the period. The principal's choice of action in the first period of the post-opportunity phase is the action associated with the point on the frontier associated with the continuation payoffs determined in the pre-opportunity phase.

To make the analysis interesting, we assume that the new project is neither too attractive nor too unattractive to both parties (see conditions N_u and N_π in Appendix B). We denote by N the set of (u_N, π_N) that satisfy these conditions. Notice that the game is now a stochastic game rather than a repeated one. To characterize the optimal relational contract, we therefore have to characterize two payoff frontiers: $\pi_{\text{Pre}}(\cdot)$—the frontier in the pre-opportunity phase—and $\pi_{\text{Post}}(\cdot)$—the frontier in the post-opportunity phase. Since the game transitions from the pre- to the post-opportunity phase, but never the reverse, we first characterize $\pi_{\text{Post}}(\cdot)$ and then $\pi_{\text{Pre}}(\cdot)$.

We characterize $\pi_{\text{Post}}(\cdot)$ formally in the appendix. Figure 4 illustrates its main features and compares them to those of the payoff frontier $\pi(\cdot)$ in our baseline model. One difference is that $\pi_{\text{Post}}(\cdot)$ is everywhere above $\pi(\cdot)$. This reflects the fact that the principal's payoff from the new project π_N is higher than her highest equilibrium payoff in the baseline model. In fact, π_N is the highest payoff on the post-opportunity frontier. The principal would therefore always choose the new project if it were available in the first period. Another difference between the payoff frontiers is that the payoffs from the new project (u_N, π_N) are on $\pi_{\text{Post}}(\cdot)$ but not $\pi(\cdot)$. This is the case since choosing the new project can be sustained as a subgame perfect equilibrium of the stage game in the post-opportunity phase but obviously not in the baseline model. Finally, for values of u around u_N, the frontier payoffs in the post-opportunity phase are not supported by cooperative empowerment as they are in the baseline model. Instead, they are supported by randomization between cooperative empowerment and choosing the new project.

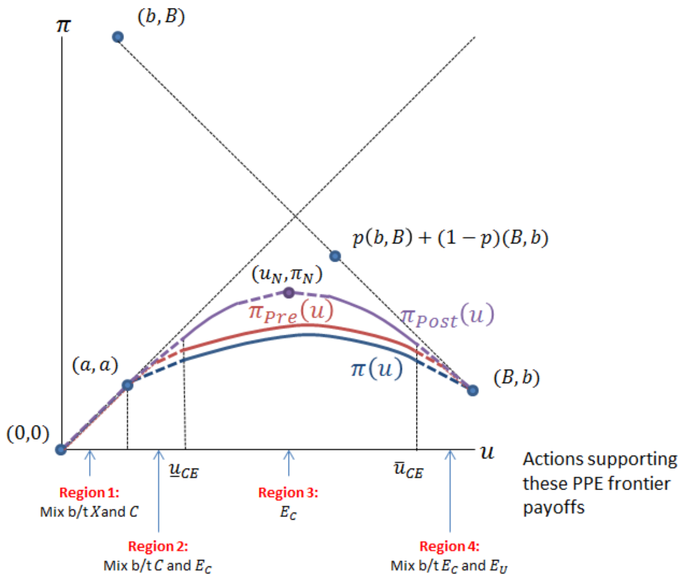

Figure 5: This figure illustrates the PPE payoff frontier for the baseline model and for the pre-opportunity and post-opportunity phases of the game with public opportunities. It describes the actions that support each point on the frontier for the pre-opportunity phase. The dotted segments are supported by public randomization between their endpoints, and this public randomization occurs at the end of the period. The agent's expected continuation payoff corresponds to a continuation payoff he will realize if the new opportunity remains unavailable and a continuation payoff he will realize if the new opportunity becomes available, both of which lie on their respective frontiers.

Now, consider the payoff frontier in the pre-opportunity phase. We illustrate its main features in Figure 5 and again relegate the formal analysis to the appendix. The figure shows that the payoff frontier in the pre-opportunity phase looks similar to the one in the baseline model. Recall, however, that in the baseline model, payoffs on the frontier are supported by continuation payoffs that are again on the same frontier. In the pre-opportunity phase, in contrast, they are supported by continuation payoffs that are either on the frontier of the pre-opportunity phase or on the frontier of the post-opportunity phase. As the game evolves during the pre-opportunity phase, it can therefore become necessary to distort the continuation payoff that the agent receives if the new project becomes available away from u_N. From Figure 5, however, it follows that there is then at least some chance that the principal will not choose the new project as soon as it becomes available and may, in fact, never do so. Proposition 3 provides conditions under which this is the case.

PROPOSITION 3. *For each* $(u_N, \pi_N) \in N$,

(i.) There exists $\overline{\pi}(u_N)$ *and* $\overline{q}(u_N)$ *such that for all* $\pi_N \leq \overline{\pi}(u_N)$ *and* $q \leq \overline{q}(u_N)$, *there exists a history* h^T *such that* $\Pr\left(u_T = u_N | h^T\right) < 1$, *where* T *is the first period in the post-opportunity phase and* h^T *is the history of public outcomes up to period* T.

(ii.) There exists a $\widehat{\pi}(u_N)$ *and* $\widehat{q}(u_N)$ *such that for all* $\pi_N \leq \widehat{\pi}(u_N)$ *and* $q \leq \widehat{q}(u_N)$, *there exists a history* h^T *such that* $\Pr\left(u_t = u_N | h^T\right) = 0$ *for all* $t \geq T$.

The first part of the proposition provides conditions under which the principal does not choose the new project as soon as it becomes available. Suppose, for instance, that during the pre-opportunity phase the agent has chosen the principal's preferred project so often that his continuation payoff if the new project becomes available exceeds u_N. The principal is then rewarding the agent for his good performance in the pre-opportunity phase by promising not to choose the new project as soon as it becomes available.

The second part of the proposition shows that the principal may in fact promise never to choose the new project. Suppose that during the pre-opportunity phase the agent has chosen the principal's preferred project even more often so that his continuation payoff if the new project becomes available does not only exceed u_N but is actually equal to B. The principal is then rewarding the agent for his performance in the pre-opportunity phase by promising him that he will always be able to choose his own preferred project, even after the new project becomes available.

The same forces that limit the organization's ability to adapt to changes in the environment that are privately observed by the agent can therefore also prevent it from taking advantage of opportunities, even when those opportunities are publicly observable and even when the principal can exploit those opportunities herself without having to induce the agent to do so.

7 Extensions

We now turn to robustness issues and explore the implications of relaxing two of the key assumptions in our baseline model.

One such assumption is that monetary transfers are not feasible. As we discussed in the literature review, this assumption is in line with the incomplete contracting literature on delegation and the broader literature on mechanism design without transfers. Even though there are often constraints on the use of monetary transfers, however, they do not typically prevent it entirely. It is therefore natural to ask what specifically is constraining the use of monetary transfers and how the constrained use of such transfers affects the predictions of our model. We address this issue in the first subsection below in which we extend our baseline model by allowing for monetary transfers from the principal to the agent. We continue to rule out transfers from the agent to the principal, however, on the basis of the agent's liquidity constraint. We then show that both the short-run and the long-run dynamics are essentially unchanged. Since liquidity constraints are pervasive among employees, the predictions of our baseline model therefore apply more broadly than a narrow interpretation of the model may at first suggest.

A particular feature of the optimal relational contracts we characterized above is that the agent either exercises his power in an entirely selfless manner—by choosing the principal's preferred project whenever it is available—or without any restraint—by always choosing his own preferred project. In the second subsection below, we show that this stark feature is an artifact of the assumption that the agent can only choose between at most two projects. In particular, there we return to the baseline model without monetary transfers and show that once the agent can choose among multiple projects, the principal can punish and reward the agent by changing his discretion more continuously. The agent's discretion then ebbs and flows during the initial phase of the relationship, before the relationship settles into an entrenched allocation of power.

7.1 Monetary Transfers

To allow for monetary transfers in our baseline model, suppose that at the end of period t, the principal can pay the agent a non-negative bonus payment $w_t \geq 0$ contingent upon the agent's project choice. A relational contract therefore specifies a bonus scheme and an action to be taken in each period. Denote $\pi_T(\cdot)$ the PPE payoff frontier of this extended game with transfers. For each payoff pair $(u, \pi_T(u))$, if it is supported by pure action $j \neq E_C$, let $w_j(u)$ denote the bonus payment paid by the principal to the agent. If $(u, \pi_T(u))$ is supported with E_C, let $w_h(u)$ and $w_\ell(u)$ be the associated bonus payments. The agent will choose the principal's preferred project whenever his incentive constraint is satisfied, that is, whenever

$$(1-\delta)(b + w_h(u)) + \delta u_h(u) \geq (1-\delta)(B + w_\ell(u)) + \delta u_\ell(u).$$

We characterize the payoff frontier in Proposition 4 in the appendix and illustrate it in Figure 6.

In the optimal relational contract, the relationship begins with cooperative empowerment. The agent is rewarded for choosing the principal's preferred project through an increase in the probability he will be able to choose his own project indefinitely in the future. When the agent chooses his preferred project, he is punished through an increase in the probability that the safe project will be chosen indefinitely in the future or that the relationship will terminate. In particular, in the first period, both rewards and punishments involve changes in continuation payoffs rather than bonus payments.

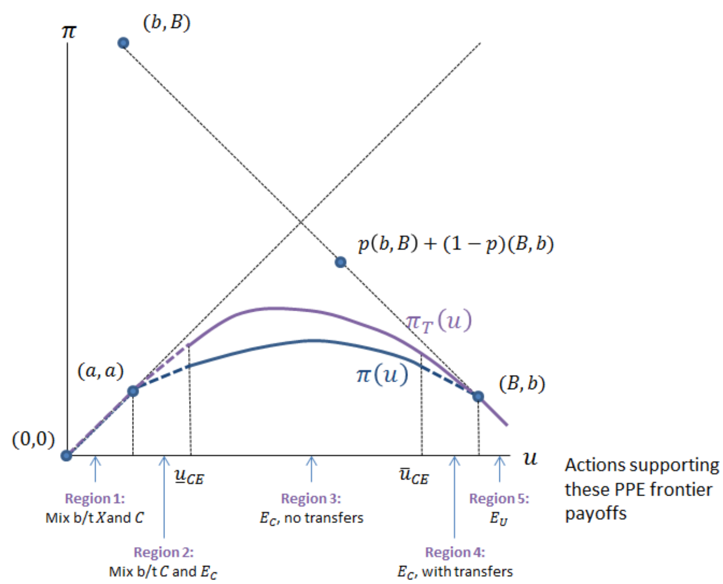

Figure 6: This figure illustrates the PPE payoff frontiers for the baseline model and for the extended model in which the principal can pay the agent. It also describes the actions that support each point on the frontier.

Since both cooperative empowerment and uncooperative empowerment yield a total surplus of $b + B$ in each period, rewards are simply a reallocation of total surplus from the principal to the agent, while punishments actually result in a decrease in total surplus. If transfers from the agent to the principal were possible, then surplus-destroying punishments could be avoided by requiring such transfers when the agent chooses his own preferred project. The principal would be indifferent between rewarding the agent through increases in his continuation payoff or through monetary transfers. If the agent is liquidity-constrained, however, punishment requires surplus destruction, and it requires more surplus destruction the lower the agent's continuation payoff is. Conversely, punishments are less costly the greater the agent's continuation payoff is. It then follows that whenever it is feasible to reward the agent with an increase in his continuation payoff, doing so is preferable to rewarding him with money. Indeed, for all $u \leq \bar{u}_{CE}$, bonus payments are never made.

In the baseline model, when $u > \bar{u}_{CE}$, the principal is unable to motivate the agent to choose her preferred project. When transfers are allowed, however, she is able to motivate him through a combination of cash and control. She therefore optimally promises him a positive bonus $w_h(u)$ and a continuation payoff $u_h(u) = B$. As a result, if the principal ever rewards the agent with money, the agent's continuation payoff is B. In the following period, it once again becomes optimal to no longer use transfers to motivate the agent, and instead to engage in permanent empowerment, which yields a payoff of b for the principal.

Notice that the point (B, b) can sometimes be implemented in other ways. For example, another way to implement (B, b) is through a bonus scheme in which the principal empowers the agent and pays him a bonus of $B - b$ whenever he chooses her preferred project. While the principal's reneging temptation under the bonus scheme is larger than her reneging temptation under permanent empowerment, for sufficiently high discount factors, the bonus scheme can indeed implement (B, b). Regardless of how the payoffs on the frontier are implemented, however, both the short-run and long-run dynamics are similar whether monetary transfers are possible.

7.2 Intermediate Allocations of Power

To explore intermediate allocations of power, we return to the baseline model without transfers but now allow the agent to choose between more than two projects. In particular, in every period, the agent can now choose between his preferred project, the principal's preferred project, if it is available, and a third project. Just like his preferred project, this third project is always available. In addition to "cooperative empowerment," "uncooperative empowerment," "centralization," and "exit," the availability of the third project gives the players the option to choose what we refer to as "partial empowerment." Under partial empowerment, the principal empowers the agent and she does so with the understanding that he will choose her preferred project whenever it is available and the safe project whenever it is not.

Formally, suppose that the agent's project-choice set in period t is $\mathcal{K}_{A,t} \in \{\{S, A\}, \{S, A, P\}\}$, where $\Pr[\mathcal{K}_{A,t} = \{S, A, P\}] = p$. Under partial empowerment, which we denote by $j = E_P$, the agent chooses $k_t = P$ whenever possible and $k_t = S$ otherwise. To motivate the agent to choose the principal's preferred project, the principal has to reward him with an increase in his continuation payoff if he does so and punish him with a decrease in his continuation payoff if he chooses the safe project. Specifically, let $u_{E_P,\ell}(u)$ denote the agent's continuation payoff under partial empowerment if he chooses the safe project and $u_{E_P,h}(u)$ if he chooses the principal's project. The agent will then choose $k_t = P$ as long as

$$(1-\delta) b + \delta u_{E_P,h}(u) \geq (1-\delta) a + \delta u_{E_P,\ell}(u)$$

is satisfied. This constraint can be taken to be binding, which in turn uniquely pins down the continuation payoffs $u_{E_P,\ell}(u)$ and $u_{E_P,h}(u)$. We denote the PPE payoff frontier of this extended game by $\pi_S(u)$. We formally characterize this frontier in Lemma 9 in the appendix and illustrate it below in Figure 7.

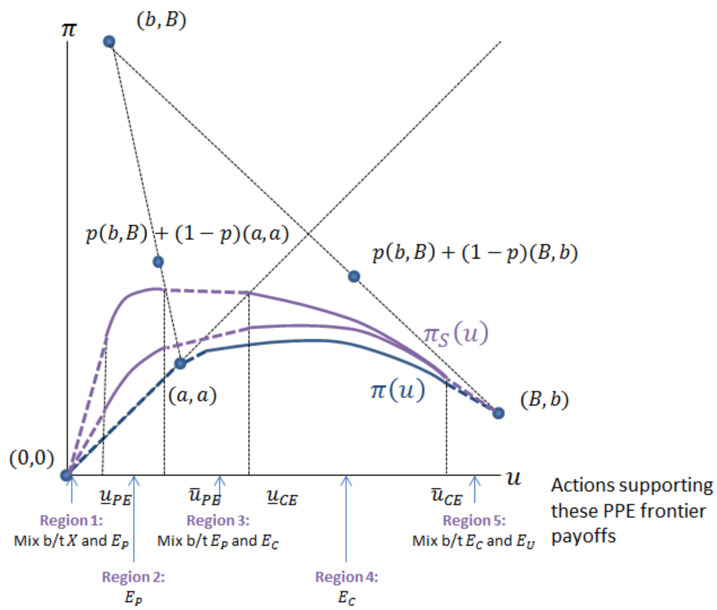

Figure 7: This figure illustrates the PPE payoff frontiers for our baseline model (bottom-most curve) and for the model in which the safe project can be chosen by the agent (top two curves). They also describe the actions that support each point on the frontier. Depending on the parameters, in the first period, either partial empowerment (corresponding to the top-most curve) or cooperative empowerment (corresponding to the middle curve) is chosen.

Partial empowerment yields higher profits for the principal than centralization does. As a result, the PPE payoff frontier lies above the point (a,a), and centralization will never be chosen in equilibrium. The frontier now consists of five regions. In the first region, the parties randomize between exit and partial empowerment in the next period. In the second region, parties engage in partial empowerment. In the third region, parties randomize between partial empowerment and cooperative empowerment in the next period. In the fourth region, parties engage in cooperative empowerment, and in the fifth region, they randomize between cooperative empowerment and uncooperative empowerment. Depending on the parameters of the model, the highest point on the frontier lies in either region 2 or region 4.

Proposition 5, in the appendix, shows that the principal starts out by engaging in either partial empowerment or cooperative empowerment. As in the baseline model, to motivate the agent to choose her preferred project whenever it is available, the principal increases his continuation value whenever he chooses her preferred project, and she decreases his continuation value when he chooses his own preferred project (under cooperative empowerment) or the safe project (under partial empowerment). The short-run dynamics are similar whether the optimal relational contract begins with partial empowerment or with cooperative empowerment.

When the agent can choose the safe project, the principal has two tools available to punish him. Under cooperative empowerment, if the agent has chosen his preferred project sufficiently often, his continuation payoff falls, and eventually the principal has to alter the agent's choice of project in order to reduce his per-period payoff. She does so by restricting the set of projects the agent can choose from, and in particular, she removes his preferred project from this set. Reduced power can therefore be a punishment for poor performance under cooperative empowerment.

Similarly, under partial empowerment, if the agent has chosen the safe project sufficiently often, his continuation payoff falls. Eventually, the principal has to punish him, and she does so by terminating the relationship with some positive probability. If the agent has chosen the principal's preferred project sufficiently often, his continuation payoff increases, and the principal eventually rewards him with increased power, allowing him to choose his own preferred project when her preferred project is not available. Finally, as in the baseline model, the possibility of the relationship moving into uncooperative empowerment serves as a potential reward for the agent.

We describe the long-run dynamics in Proposition 6 in the appendix. As in the baseline model, the relationship eventually settles into one of two steady states: termination or permanent empowerment. Since centralization is never chosen in equilibrium, however, the relationship can never settle into permanent disenfranchisement.

8 Conclusions

Power is an inherently dynamic concept. The transfer of power from the top of an organization to those further down in the hierarchy is based on informal promises and is thus necessarily relational. Moreover, the allocation of power often evolves over time, with some members experiencing increases in their power while others see theirs slip away. To understand the allocation of power within organizations, this paper therefore develops a dynamic model of power.

In this model, the central purpose of power is to serve as a reward mechanism that those at the top use to discipline their subordinates and influence their decision-making. Even though this role of power as a reward has long been noted in sociology and organizational theory, it is absent in existing economic models of organizations. We capture this role formally and show that it gives rise to rich dynamics. Our results speak to how and why power is gained, lost, and retained and thus adds to our understanding of life inside organizations. And, in doing so, it provides new perspectives on why similar firms organize differently, even though those organizational differences are associated with differences in performance, and why established firms often have a harder time adapting to changes in their environments than their younger rivals, even when those changes are publicly observable.

Our model is only a first step in developing a formal theory of power dynamics, and there are many issues that it does not address. Since we only allow for a single agent, for instance, we cannot explore horizontal differences in power across members of an organization, which are a central concern in the sociology and organizational theory literature. We also take the boundaries of the firm as given and don't allow the principal to sell the formal authority to make decisions to the agent. A richer model would allow for such a transfer of formal authority and then develop a theory of the firm in which agents may integrate precisely because it allows them to use power dynamics as a reward mechanism. We leave these and related issues for future research.

References

[1] Aghion, Philippe, Nicholas Bloom, and John Van Reenen. 2013. Incomplete Contracts and the Internal Organization of Firms. *Journal of Law, Economics, and Organization*, 30(suppl 1): 37-63.

[2] ——— and Jean Tirole. 1997. Formal and Real Authority in Organizations. *Journal of Political Economy*, 105(1): 1-29.

[3] Abreu, Dilip, David Pearce, and Ennio Stacchetti. 1990. Toward a Theory of Discounted Repeated Games with Imperfect Monitoring. *Econometrica*, 58(5): 1041-1063.

[4] Alonso, Ricardo and Niko Matouschek. 2007. Relational Delegation. *The RAND Journal of Economics*, 38(4): 1070-1089.

[5] ——— and ———. 2008. Optimal Delegation. *The Review of Economic Studies*, 75(1): 259-293.

[6] Amador, Manuel and Kyle Bagwell. 2013. The Theory of Optimal Delegation with an Application to Tariff Caps. *Econometrica*, 81(4): 1541-1599.

[7] Armstrong, Mark and John Vickers. 2010. A Model of Delegated Project Choice. *Econometrica*, 78(1): 213-244.

[8] Athey, Susan and Kyle Bagwell. 2001. Optimal Collusion with Private Information. *The RAND Journal of Economics*, 32(3): 428-465.

[9] Baker, George, Robert Gibbons and Kevin Murphy. 1999. Informal Authority in Organizations. *The Journal of Law, Economics, and Organization*, 15(1): 56-73.

[10] Biais, Bruno, Thomas Marriotti, Guillaume Plantin and Jean-Charles Rochet. 2007. Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications. *The Review of Economic Studies*, 74(2): 345-390.

[11] Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. Does Management Matter? Evidence from India. *Quarterly Journal of Economics*, 128(1): 1-51.

[12] Bolton, Patrick and Mathias Dewatripont. 2013. Authority in Organizations. In *Handbook of Organizational Economics*, eds. Robert Gibbons and John Roberts, Princeton University Press.

[13] Bower, Joseph and Clayton Christensen. 1995. Disruptive Technologies: Catching the Wave. *Harvard Business Review*, 73(1): 43–53.

[14] Brynjolfsson, Eric and Paul Milgrom. 2013. Complementarity in Organizations. In *Handbook of Organizational Economics*, eds. Robert Gibbons and John Roberts, Princeton University Press.

[15] Chassang, Sylvain. 2010. Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts. *American Economic Review*, 100(1): 448-465.

[16] Clementi, Gian Luca and Hugo Hopenhayn. 2006. A Theory of Financing Constraints and Firm Dynamics. *The Quarterly Journal of Economics*, 121(1): 229-265.

[17] Cyert, Richard and James March. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs.

[18] DeMarzo, Peter and Michael Fishman. 2007. Optimal Long-Term Financial Contracting. *The Review of Financial Studies*, 20(6): 2079-2128.

[19] Dessein, Wouter. 2002. Authority and Communication in Organizations. *The Review of Economic Studies*, 69(4): 811-838.

[20] Eichenwald, Kurt. 2012. Microsoft's Lost Decade. *Vanity Fair*, 54(8).

[21] Emerson, Richard. 1962. Power-Dependence Relations. *American Sociological Review*, 27(1): 31-41.

[22] Fong, Yuk-Fai and Jin Li. 2013. Relational Contracts, Efficiency Wages, and Employment Dynamics. Mimeo.

[23] Garrett, Daniel and Alessandro Pavan. 2012. Managerial Turnover in a Changing World. *Journal of Political Economy,* 120(5): 879-925.

[24] Ghosh, Parikshit and Debraj Ray. 1996. Cooperation in Community Interaction without Information Flows. *The Review of Economic Studies,* 63(3): 491-519.

[25] Gibbons, Robert and Rebecca Henderson. 2013. What Do Managers Do? In *Handbook of Organizational Economics*, eds. Robert Gibbons and John Roberts, Princeton University Press.

[26] Green, Edward and Robert Porter. 1984. Noncooperative Collusion under Imperfect Price Information. *Econometrica,* 52(1): 87-100.

[27] Guo, Yingni and Johannes Horner. 2014. Dynamic Mechanisms without Money. Mimeo.

[28] Halac, Marina. 2014. Relationship Building: Conflict and Project Choice Over Time. *The Journal of Law, Economics, and Organization*, 30(4): 683-708.

[29] ——— and Andrea Prat. 2014. Managerial Attention and Worker Engagement. Mimeo.

[30] Hickson, David, Robert Hinings, C.A. Lee, R.E. Schneck, and Johannes Pennings. 1971. A Strategic Contingencies' Theory of Intraorganizational Power. *Administrative Science Quarterly*, 16(2): 216-229.

[31] Holmstrom, Bengt. 1984. On the Theory of Delegation. In *Bayesian Models in Economic Theory*, eds. Marcel Boyer and Richard Kihlstrom. North-Holland.

[32] Kranton, Rachel. 1996. The Formation of Cooperative Relationships. *Journal of Law, Economics, & Organization*, 12(1): 214-233.

[33] Levin, Jonathan. 2003. Relational Incentive Contracts. *American Economic Review*, 93(3): 835-857.

[34] Li, Jin and Niko Matouschek. 2013. Managing Conflicts in Relational Contracts. *American Economic Review*, 103(6): 2328-2351.

[35] Malcomson, James. 1999. Individual Employment Contracts. In *Handbook of Labor Economics*, eds. Orley Ashenfelter and David Card, Vol. 3. Elsevier.

[36] ———. 2013. Relational Incentive Contracts. In *Handbook of Organizational Economics*, eds. Robert Gibbons and John Roberts, Princeton University Press.

[37] McAdams, David. 2011. Performance and Turnover in a Stochastic Partnership. *American Economic Journal: Microeconomics*, 3(4): 107-142.

[38] Mailath, George and Larry Samuelson. 2001. Who Wants a Good Reputation? *The Review of Economic Studies*, 68(2): 415-441.

[39] ——— AND ———. 2006. *Repeated Games and Reputations.* Oxford University Press.

[40] MILGROM, PAUL AND JOHN ROBERTS. 1995. Complementarities and Fit: Strategy, Structure, and Organizational Change in Manufacturing. *Journal of Accounting and Economics*, 19: 179-208.

[41] MILLER, DAVID AND JOEL WATSON. 2013. A Theory of Disagreement in Repeated Games with Bargaining. *Econometrica*, 81(6): 2303-2350.

[42] PADRO I MIQUEL, GERARD AND PIERRE YARED. 2012. The Political Economy of Indirect Control. *The Quarterly Journal of Economics*, 127(2): 947-1015.

[43] PFEFFER, JEFFREY. 1981. *Power in Organizations.* Pitman.

[44] ——— AND GERALD SALANCIK. 1978. *The External Control of Organizations: A Resource Dependence Perspective.* Harper & Row.

[45] SAMUELSON, LARRY. 2006. The Economics of Relationships. *Econometric Society Monographs*, 41: 136–185.

[46] SCHAEFER, SCOTT. 1998. Influence Costs, Structural Inertia and Organizational Change. *Journal of Economics and Management Strategy*, 7(2): 237–263.

[47] SPEAR, STEPHEN AND SANJAY SRIVASTAVA. 1987. On Repeated Moral Hazard with Discounting. *Review of Economic Studies*, 54(4): 599-617.

[48] SAFRONOV, MIKHAIL AND BRUNO STRULOVICI. 2014. Explicit Renegotiation in Repeated Games. Mimeo.

[49] WATSON, JOEL. 1999. Starting Small and Renegotiation. *Journal of Economic Theory*, 85(1): 52-90.

[50] ———. 2002. Starting Small and Commitment. *Games and Economic Behavior*, 38(1): 176-199.

[51] YANG, HUANXING. 2013. Non-stationary Relational Contracts with Adverse Selection. *International Economic Review*, 54(2): 525-547.

[52] ZHU, JOHN. 2013. Optimal Contracts with Shirking. *The Review of Economic Studies*, 80(2): 812-839.

Appendix (For Online Publication)

This appendix is divided into four sections. Appendix A contains proofs for the results describing the optimal relational contract in the baseline model. Appendix B contains proofs for the model with public opportunities. Appendices C and D analyze the extensions of the model to allow for monetary transfers and intermediate allocations of control.

Appendix A: Optimal Relational Contract in the Baseline Model

LEMMA 1. *The PPE payoff set $\mathcal{E}$ has the following properties: (i.) it is compact; (ii.) $\pi(u)$ is concave; (iii.) the payoff pair (u, π) belongs to $\mathcal{E}$ if and only if $u \in [0, B]$ and $\pi \in [bu/B, \pi(u)]$.*

Proof of Lemma 1: Part (i.): Note that there are finite number of actions the players can take, and standard arguments then imply that the PPE payoff set $\mathcal{E}$ is compact. Part (ii.): the concavity of π follows immediately from the availability of the public randomization device. Part (iii.): Notice that both $(0, 0)$ and (B, b) are PPE payoffs sustained by termination and permanent empowerment respectively. In addition, there exists no actions that give the agent payoffs below 0 or above B. This implies that if $(u, \pi) \in \mathcal{E}$, then $u \in [0, B]$. Moreover, one must have $\pi \leq bu/B$ because no payoff below the line segment joining $(0, 0)$ and (B, b) is feasible. Next, given the public randomization device, any payoff on the line segment between $(0, 0)$ and (B, b) can be supported as a PPE payoff. In other words, $(u, bu/B)$ is a PPE payoff for any $u \in [0, B]$. Finally, the randomization between $(u, bu/B)$ and $(u, \pi(u))$ allows us to obtain any payoff (u, π) for all $\pi \in [bu/B, \pi(u)]$. ■

LEMMA 2. *For any payoff $(u, \pi(u))$ on the frontier, the equilibrium continuation payoffs remain on the frontier.*

Proof of Lemma 2 : To show that for each payoff $(u, \pi(u))$ on the frontier, the equilibrium continuation payoffs remain on the frontier, it suffices to show that this is true if $(u, \pi(u))$ is supported by a pure action. Suppose $(u, \pi(u))$ is supported by centralization. Let (u_C, π_C) be the associated continuation payoff. Suppose to the contrary of the claim that $\pi_C < \pi(u_C)$. Now consider an alternative strategy profile that also specifies centralization but in which the continuation payoff is given by $(u_C, \widehat{\pi}_C)$, where $\widehat{\pi}_C = \pi_C + \varepsilon$ and where $\varepsilon > 0$ is small enough such that $\pi_C + \varepsilon \leq \pi(u_C)$. It follows from the promise-keeping constraints $\mathrm{PK}_{\mathrm{C}}^{\mathrm{P}}$ and $\mathrm{PK}_{\mathrm{C}}^{\mathrm{A}}$ that under this alternative strategy profile the payoffs are given by $\widehat{u} = u$ and $\widehat{\pi}_C = \pi(u) + \delta\varepsilon > \pi(u)$. It can be checked that this alternative strategy profile satisfies all the constraints and therefore constitutes a PPE. Since $\widehat{\pi}_C > \pi(u)$, this contradicts the definition of $\pi(u)$, thus it must be that $\pi_C = \pi(u_C)$. The argument is identical when $(u, \pi(u))$ is supported by cooperative or uncooperative empowerment. ■

LEMMA 3. In addition, the following hold.

(i.) If $(u, \pi(u))$ is supported with cooperative empowerment, the agent's continuation payoff can be chosen by

$$\delta u_h(u) = u - (1-\delta)\, b;$$
$$\delta u_\ell(u) = u - (1-\delta)\, B.$$

(ii.) If $(u, \pi(u))$ is supported with uncooperative empowerment, the agent's continuation payoff is given by

$$\delta u_{E_U}(u) = u - (1-\delta)\, B.$$

(iii.) If $(u, \pi(u))$ is supported with centralization, the agent's continuation payoff is given by

$$\delta u_C(u) = u - (1-\delta)\, a.$$

(iv.) If $(u, \pi(u))$ is supported by exit, the agent's continuation payoff is given by

$$\delta u_X(u) = u.$$

Proof of Lemma 3: For part $(i.)$, let $(u, \pi(u))$ be associated with the continuation payoffs $(u_\ell, u_h, \pi(u_\ell), \pi(u_h))$. Suppose that for this PPE the $\mathrm{IC}_{\mathrm{E_C}}$ is slack, that is, $(1-\delta)\, b + \delta u_h > (1-\delta)\, B + \delta u_\ell$. Now consider an alternative strategy profile with continuation payoffs given by $(\widehat{u}_\ell, \widehat{u}_h, \pi(\widehat{u}_\ell), \pi(\widehat{u}_h))$, where $\widehat{u}_\ell = u_\ell + p\varepsilon$ and $\widehat{u}_h = u_h - (1-p)\,\varepsilon$ for $\varepsilon > 0$. It follows from the promise-keeping constraints $\mathrm{PK}^{\mathrm{P}}_{\mathrm{E_C}}$ and $\mathrm{PK}^{\mathrm{A}}_{\mathrm{E_C}}$ that, under this strategy profile, the payoffs are given by $\widehat{u} = u$ and

$$\widehat{\pi} = p\left[(1-\delta)\, B + \delta\pi(\widehat{u}_h)\right] + (1-p)\left[(1-\delta)\, b + \delta\pi(\widehat{u}_l)\right].$$

From the concavity of π it then follows that

$$\widehat{\pi} \geq (1-\delta)\, b + \delta((1-p)\,\pi(u_\ell) + p\pi(u_h)) = \pi(u).$$

It can be checked that for sufficiently small ε this alternative strategy profile satisfies all the constraints and therefore constitutes a PPE. Since $\widehat{\pi} \geq \pi(u)$ this implies that for any PPE with payoffs $(\pi, u(\pi))$ for which IC is not binding there exists another PPE for which $\mathrm{IC}_{\mathrm{E_C}}$ is binding and which gives the parties weakly larger payoffs. Notice that when $\mathrm{IC}_{\mathrm{E_C}}$ is binding, we have $u_h(u) = (u - (1-\delta) b)/\delta$ and $u_\ell(u) = (u - (1-\delta) B)/\delta$. This proves part $(i.)$. Parts $(ii.) - (iv.)$ follow directly from the promise-keeping constraints $\mathrm{PK}^{\mathrm{A}}_{\mathrm{E_U}}$, $\mathrm{PK}^{\mathrm{A}}_{\mathrm{C}}$, and $\mathrm{PK}^{\mathrm{A}}_{\mathrm{X}}$. ■

LEMMA 4. *The PPE frontier $\pi(u)$ is the unique function that solves the following problem. For all $u \in [0, B]$*

$$\pi(u) = \max_{\alpha_j \geq 0, u_j \in [0,B]} \sum_{j \in \{C, E_C, E_U, X\}} \alpha_j \pi_j(u_j)$$

such that

$$\sum_{j \in \{C, E_C, E_U, X\}} \alpha_j = 1$$

and

$$\sum_{j \in \{C, E_C, E_U, X\}} \alpha_j u_j = u.$$

Proof of Lemma 4: Since the frontier is Pareto efficient, by the APS bang-bang result, for any efficient payoff pair, only using the extreme points of the payoff set is sufficient. Replacing the sup with max is valid since the payoff set is compact. To establish the uniqueness, we just observe that the problem is now a maximization problem on a compact set, even if the maximizers are not unique, the maximum is. ■

Next, instead of proving Lemma 5 directly, we establish alternative Lemmas 5A-5C, which leads to Lemma 5′. Lemma 5′ provides a complete characterization of the PPE payoff frontier for all $p \in (0, 1)$, which includes the case of $p \leq 1/2$ (Lemma 5).

LEMMA 5A. *There exists a cutoff value $\bar{u}_{CE} = (1-\delta) b + \delta B$ such that $\pi_{E_U}(u) = \pi(u)$ if and only if*

$$u \in [(1-\delta) B + \delta \bar{u}_{CE}, B].$$

Proof of Lemma 5A: First, notice that $\pi_{E_U}(B) = \pi(B)$. Next, recall that $\pi_{E_U}(u) = (1-\delta) b +$

$\delta\pi\left(u_{E_U}\left(u\right)\right)$. Taking the right derivative, we have

$$\pi'_{E_U,+}\left(u\right) = \delta\pi'_{+}\left(u_{E_U}\left(u\right)\right)u'_{E_U,+}\left(u\right) = \pi'_{+}\left(u_{E_U}\left(u\right)\right) \geq \pi'_{+}\left(u\right),$$

where we used the fact that if $u < B$, $u_{E_U}\left(u\right) < u$, and therefore $\pi'_{+}\left(u_{E_U}\left(u\right)\right) \geq \pi'_{+}\left(u\right)$ by concavity of the frontier. Since $\pi'_{E_U,+}\left(u\right) \geq \pi'_{+}\left(u\right)$ for all $u < B$, there exists u^* such that $\pi_{E_U}\left(u\right) = \pi\left(u\right)$ if and only if $u \in [u^*, B]$. Next, we show that $u^* = (1-\delta)\,B + \delta\bar{u}_{CE}$.

To do this, we first show that there exists some $u < B$ such that $\pi_{E_U}\left(u\right) = \pi\left(u\right)$, i.e., $u^* < B$. We prove this by contradiction. Suppose to the contrary that $\pi_{E_U}\left(u\right) < \pi\left(u\right)$ for all $u < B$. Choose a small enough $\varepsilon > 0$ such that $(B-\varepsilon, \pi\left(B-\varepsilon\right))$ cannot be supported by pure actions. Notice that such ε exists, because by assumption $(B-\varepsilon, \pi\left(B-\varepsilon\right))$ is not supported by E_U, and if it were supported by C or E_C, then the agent's continuation payoffs (u_C and u_h) must exceed B, leading to a contradiction. This implies that $(B-\varepsilon, \pi\left(B-\varepsilon\right))$ must be supported by randomization, and therefore the frontier is a straight line between $B-\varepsilon$ and B. Let the slope of the payoff frontier between $(B-\varepsilon, \pi\left(B-\varepsilon\right))$ and (B, b) as s. It then follows that for all $u \in [B-\delta\varepsilon, B)$ (i.e. $u_{E_U}\left(u\right) \geq B-\varepsilon$), we have

$$\pi_{E_U}\left(u\right) = \pi\left(u\right) = b + s(u-B).$$

This contradicts the assumption that $\pi_{E_U}\left(u\right) < \pi\left(u\right)$ for all $u < B$.

The above shows that $\pi_{E_U}\left(u\right) = \pi\left(u\right)$ for $u \in [u^*, B]$, where $u^* < B$. It follows that for all $u \in [u^*, B]$, $\pi'_{E_U,+}\left(u\right) = \pi'_{+}\left(u_{E_U}\left(u\right)\right) = \pi'_{+}\left(u\right)$. Since π is concave, this implies that the slope of π is constant for all agent's payoffs in $\left(u_{D_U}\left(u\right), u\right)$. It is then immediate that $\pi\left(u\right)$ is a straight line between $[u_{E_U}^{-1}(u^*), B]$. Let $(u', \pi\left(u'\right))$ be the left end point of the line segment. Notice that $(u', \pi\left(u'\right))$ is an extremal point of the payoff frontier, it must be supported by pure action. Moreover, it cannot be supported by uncooperative empowerment, because $u' \leq u_{E_U}^{-1}(u^*) < u^*$. This implies that $(u', \pi\left(u'\right))$ is supported by either cooperative empowerment or centralization. Next, we show that $(u', \pi\left(u'\right))$ cannot be supported by centralization, so we must have $\pi\left(u'\right) = \pi_{E_C}\left(u'\right)$.

Suppose to the contrary that $\pi\left(u'\right) = \pi_C\left(u'\right)$. Notice that

$$\begin{aligned} u' &= (1-\delta)\,a + \delta u_C\left(u'\right); \\ \pi\left(u'\right) &= (1-\delta)\,a + \delta\pi\left(u_C\left(u'\right)\right), \end{aligned}$$

which implies that

$$\begin{aligned} u' - a &= \delta\left(u_C\left(u'\right) - a\right); \\ \pi\left(u'\right) - a &= \delta\left(\pi\left(u_C\left(u'\right)\right) - a\right). \end{aligned}$$

Now take any $u \in (\max\{a, u'\}, B)$. We have $u_C(u) > u$, and the above implies that

$$\frac{\pi(u) - a}{u - a} = \frac{\pi(u_C(u)) - a}{u_C(u) - a} = \frac{\pi(u_C(u)) - \pi(u')}{u_C(u) - u} = \frac{\pi(u) - b}{u - B},$$

where the last inequality holds, because π is a straight line to the right of u' and $u > u'$. The equalities then imply that $(u, \pi(u))$ lies on the line segment between (a, a) and (B, b) for all $u \in [a, B]$, which contradicts Assumption 1. This implies that $\pi(u') > \pi_C(u')$, and we must have $\pi(u') = \pi_{E_C}(u')$.

Finally, we show that $u' = \bar{u}_{CE}$, i.e., $u_h(u') = B$. By $\mathrm{SE}_{\mathrm{E_C}}$, the continuation payoff $u_h(u')$ satisfies $u_h(u') \leq B$. Now suppose to the contrary that $u_h(u') < B$. Recall that s is the slope of the payoff frontier between $(u', \pi(u'))$ and (B, b). Now consider an alternative strategy profile that is supported by cooperative empowerment and whose continuation payoffs are given by $(\widehat{u}_\ell, \widehat{u}_h, \pi(\widehat{u}_\ell), \pi(\widehat{u}_h))$, where $\widehat{u}_\ell = u_\ell(u') + \varepsilon$ and $\widehat{u}_h = u_h(u') + \varepsilon$ for small $\varepsilon > 0$. It follows from the promise-keeping constraints $\mathrm{PK}^{\mathrm{P}}_{\mathrm{E_D}}$ and $\mathrm{PK}^{\mathrm{A}}_{\mathrm{E_D}}$ that, under this strategy profile, the payoffs are given by $\widehat{u} = u' + \delta\varepsilon$ and

$$\begin{aligned}
\widehat{\pi} &= p\left[(1-\delta) B + \delta\pi(\widehat{u}_h)\right] + (1-p)\left[(1-\delta) b + \delta\pi(\widehat{u}_\ell)\right] \\
&= \pi\left(u'\right) + p\delta\left[\pi\left(u_h\left(u'\right) + \varepsilon\right) - \pi\left(u_h\left(u'\right)\right)\right] \\
&\quad + (1-p)\,\delta\left[\pi\left(u_\ell\left(u'\right) + \varepsilon\right) - \pi\left(u_\ell\left(u'\right)\right)\right] \\
&> \pi\left(u'\right) + \delta s\varepsilon.
\end{aligned}$$

Notice that the strict inequality follows, because s is the smallest slope of π, and since $u_\ell(u') < u'$, we have $\pi(u_\ell(u') + \varepsilon) - \pi(u_\ell(u')) > s\varepsilon$ by the definition of u'. It can be checked that for sufficiently small ε this alternative strategy profile satisfies all the constraints and therefore constitutes a PPE. This implies that

$$\pi\left(u'\right) + \delta s\varepsilon < \widehat{\pi} \leq \pi\left(u' + \delta\varepsilon\right) = \pi(u') + \delta s\varepsilon,$$

where the weak inequality follows from the definition of π and the equality follows from that π is a straight line with slope s to the right of u'. Since the above chain of inequalities lead to a contradiction, we must have $u_h(u') = B$, or equivalently, $u' = \bar{u}_{CE}$. ■

LEMMA 5B. *There exists a cutoff probability level $p^{**} > 1/2$ such that the following hold:*

(i.) If $p \in (p^{**}, 1)$, $\pi_C(u) < \pi(u)$ for all $u \in [0, B]$.

(ii.) If $p \in (0, p^{**}]$, there exists $\underline{u}_C \leq a \leq \overline{u}_C$ such that $\pi_C(u) = \pi(u)$ if and only if $u \in [(1-\delta)\,a + \delta\underline{u}_C, (1-\delta)\,a + \delta\overline{u}_C]$. In addition, $\underline{u}_C = 0$ if and only if $p \leq 1/2$.

Proof of Lemma 5B: First, we show that if $\pi(a) < a$, then $\pi_C(u) < \pi(u)$ for all $u \in [0, B]$.

To prove this, we show that if that $\pi_C(u) = \pi(u)$ for some $u \neq a$, then $\pi(a) = a$. Now suppose there exists $u > a$ such that $\pi_C(u) = \pi(u)$. Since $u > a$, we have $u_C(u) = (u - (1-\delta)a)/\delta > u$. Recall that $\pi_C(u) = (1-\delta)a + \delta\pi(u_C(u))$. Taking left derivatives, we have

$$\pi'_{C,-}(u) = \pi'_{-}(u_C(u)) \leq \pi'_{-}(u),$$

where the inequality follows because the payoff frontier is concave and $u_C(u) > u$. Since π is concave, the above then implies that π is a straight line in $[u_C(u), u]$, and $\pi_C(u') = \pi(u')$ for all $u' \in [u_C(u), u]$. Repeating the argument at $u_C(u)$, we get that $\pi_C(u'') = \pi(u'')$ for all $u'' \in [u_C(u_C(u)), u_C(u)]$. Using the argument repeatedly and using the fact that π is continuous at a, we then have that $\pi_C(u') = \pi(u')$ for all $u' \in [a, u]$. It follows that

$$\pi(a) = \pi_C(a) = (1-\delta)a + \delta\pi(u_C(a)) = (1-\delta)a + \delta\pi(a),$$

which implies that $\pi(a) = a$, contradicting the assumption that $\pi(a) < a$. This proves that if $\pi(a) < a$, then $\pi_C(u) < \pi(u)$ for all $u > a$. The identical argument as above can be used to show that if $\pi(a) < a$, then $\pi_C(u) < \pi(u)$ for all $u < a$. This finishes showing that if $\pi(a) < a$, then $\pi_C(u) < \pi(u)$ for all $u \in [0, B]$.

Next, when $\pi(a) = a$, define $\overline{u}_C = \max\{u : \pi_C(u_C(u)) = \pi(u_C(u))\}$. Notice that the argument above implies that $\pi'(u)$ is the same for all u between a and $\overline{u}_C$, and this shows that π is a line segment in $[a, \overline{u}_C]$. Similarly, define $\underline{u}_C = \min\{u : \pi_C(u_C(u)) = \pi(u_C(u))\}$, and we then have that π is a line segment in $[\underline{u}_C, a]$.

Notice that when $p \leq 1/2$, the set of feasible payoffs lie below the 45 degree line for $u \in [0, a]$. In addition, payoffs on the 45 degree line with end points in $(0, 0)$ and (a, a) can be implemented by the randomization between the two end points (that are both PPE payoffs). This implies that $\pi(u) = u$ for all $u \in [0, a]$ and immediately implies that $\underline{u}_C = 0$.

Finally, we will show in the proof of the next lemma that there exists $p^{**} > 1/2$ such that $\pi(a) > a$ if and only $p > p^{**}$. Given this result, Lemma 5B follows immediately. We prove this result in Lemma 5C for convenience, and its proof does not depend on results established in this lemma. ■

LEMMA 5C. *There exists a cutoff probability level $p^{**} > 1/2$ such that the following holds.*

(i.) If $p \in [p^{**}, 1)$, $\pi_{E_C}(u) = \pi(u)$ if and only if $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$, where $\underline{u}_{CE} = (1-\delta)B$

(ii.) If $p \in [1/2, p^{**}]$, $\pi_{E_C}(u) = \pi(u)$ if and only if $u \in [\underline{u}^A_{CE}, \bar{u}^A_{CE}] \cup [\underline{u}_{CE}, \bar{u}_{CE}]$ for some $\bar{u}^A_{CE} < a < \underline{u}_{CE} \leq \delta a + (1-\delta)B$.

(iii.) If $p \in (0, 1/2)$, $\pi_{E_C}(u) = \pi(u)$ *if and only if* $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$ *for some* $\underline{u}_{CE} \in (a, \delta a + (1-\delta) B]$.

Proof of Lemma 5C: To prove the lemma, we take the following steps.

Step 1: We establish properties of PPE payoffs of a modified game.

Consider a modified game in which C is not feasible. Let π_T be the associated PPE payoff frontier of the modified game. We establish the following properties of π_T.

A: For $u \in [\bar{u}_{CE}, B]$, *where* $\bar{u}_{CE} = (1-\delta) b + \delta B$,

$$\pi_T(u) = \frac{u - \bar{u}_{CE}}{B - \bar{u}_{CE}} b + \frac{B - u}{B - \bar{u}_{CE}} \pi_T(\bar{u}_{CE}).$$

B: For $u \in [0, \underline{u}_{CE}]$, *where* $\underline{u}_{CE} = (1-\delta) B$,

$$\pi_T(u) = \frac{u}{\underline{u}_{CE}} \pi_T(\underline{u}_{CE}).$$

C: For $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$,

$$\pi_T(u) = p\left[(1-\delta) B + \delta \pi_T(u_h(u))\right] + (1-p)\left[(1-\delta) b + \delta \pi_T(u_\ell(u))\right].$$

The properties of π_T are established in the same way as the method used here except it is simpler. We therefore omit the proof of Step 1.

Step 2: There exists $p^{**} > 1/2$ such that $\pi_T(u, p) = \pi(u, p)$ if and only if $p \geq p^{**}$.

To see this, we first show that for all $u \in (0, B)$, $\pi_T(u, p)$ is strictly increasing in p. Consider $p_1 > p_2$. Define the operator $T_i f$, which maps bounded nonnegative functions on $[0, B]$ to bounded nonnegative functions on $[0, B]$, as

$$\begin{aligned}
T_i f(u) &= \frac{u}{\underline{u}_{CE}} f(\underline{u}_{CE}) \text{ for } u \leq \underline{u}_{CE} \\
T_i f(u) &= \frac{u - \bar{u}_{CE}}{B - \bar{u}_{CE}} b + \frac{B - u}{B - \bar{u}_{CE}} f(\bar{u}_{CE}) \text{ for } \bar{u}_{CE} \leq u \leq B \\
T_i f(u) &= p_i\left[(1-\delta) B + \delta f(u_h(u))\right] + (1-p_i)\left[(1-\delta) b + \delta f(u_\ell(u))\right] \text{ for } \underline{u}_{CE} \leq u \leq \bar{u}_{CE}
\end{aligned}$$

It is clear that T_i is bounded monotone in the sense that $T_i(f_1) \geq T_i(f_2)$ whenever $f_1 \geq f_2$ (in the sense that $f_1(u) \geq f_2(u)$ for all u). Let $Z(u) \equiv 0$ on $[0, B]$, it follows that $Z^* \equiv \lim_{n\to\infty} T_i(z)$ is a fixed point of T_i. Moreover, the fixed point is unique. Suppose to the contrary that f_1 and f_2

are two fixed point of T_i. Let $M = \sup_{u\in[0,B]}\{|f_1(u) - f_2(u)|\}$. Now notice that

$$\begin{aligned}
M &= \sup_{u\in[\underline{u}_{CE},\bar{u}_{CE}]}\{|f_1(u) - f_2(u)|\} = \sup_{u\in[\underline{u}_{CE},\bar{u}_{CE}]}\{|Tf_1(u) - Tf_2(u)|\} \\
&= \sup_{u\in[\underline{u}_{CE},\bar{u}_{CE}]}\{\delta|p_i\left(f_1\left(u_h\left(u\right)\right) - f_2\left(u_h\left(u\right)\right)\right) + (1-p_i)\left(f_1\left(u_\ell\left(u\right)\right) - f_2\left(u_\ell\left(u\right)\right)\right)|\} \\
&\leq \delta p_i \sup_{u\in[0,B]}\{|f_1(u) - f_2(u)|\} + \delta(1-p_i)\sup_{u\in[0,B]}\{|f_1(u) - f_2(u)|\}. \\
&\leq \delta M,
\end{aligned}$$

and this implies that $M = 0$. This shows that T_i has a unique fixed point.

Now let $\pi_T(u, p_1)$ be the unique fixed point of T_1 and $\pi_T(u, p_2)$ be the unique fixed point of T_2. Notice that for each $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$,

$$\begin{aligned}
T_1\pi_T(u, p_2) &= p_1\left[(1-\delta)B + \delta\pi_T\left(u_h\left(u\right), p_2\right)\right] + (1-p_1)\left[(1-\delta)b + \delta\pi_T\left(u_h\left(u\right), p_2\right)\right] \\
&= (p_1 - p_2)(1-\delta)(B-b) + T_2\pi_T(u, p_2) > \pi_T(u, p_2).
\end{aligned}$$

The monotonicity of T_1 then implies immediately that $T_1\pi_T(u, p_2) > \pi_T(u, p_2)$ for all $u \in (0, B)$. This finishes showing that for all $u \in (0, B)$, $\pi_T(u, p)$ is strictly increasing in p.

Next, we show that there exists p^{**} such that $\pi_T(u, p) = \pi(u, p)$ if and only if $p \geq p^{**}$. Choose p^{**} as the cutoff value such that $\pi_T(a, p^{**}) = a$. This implies that when $p > p^{**}$, $a < \pi_T(a, p) \leq \pi(a, p)$. By the argument in Lemma 5B, we then have that the frontier $\pi(u, p)$ is not supported by centralization for all $u \in [0, B]$, and this implies that $\pi_T(u, p) = \pi(u, p)$ for all $u \in [0, B]$. This implies that $\pi_T(u, p) = \pi(u, p)$ for all $p > p^{**}$, and since both π_T and π are continuous in p (by the maximum theorem), we then also have $\pi_T(u, p) = \pi(u, p)$.

Notice that we must have $p^{**} \geq 1/2$ because otherwise (a, a) lies strictly above the set of feasible payoffs and thus cannot be a PPE payoff. When $p = 1/2$, let u^* be the maximal equilibrium payoff of the agent such that $\pi_T(u)$ is on the 45 degree line. If $u^* > 0$, then it satisfies $u_h(u^*) = u^*$, which implies that $u^* = b < a$. Consequently, (a, a) again lies above $(a, \pi_T(a))$ when $p = 1/2$. As a result, we have $p^{**} > 1/2$. By combining Step 1 and 2, this proves part (i) of Lemma 5C.

Step 3: When $p < 1/2$, then $\pi_{E_C}(u) = \pi(u)$ if and only if $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$ for some $\underline{u}_{CE} \in (a, \delta a + (1-\delta)B]$.

To see this, notice that for $u \in [\overline{u}_C, \bar{u}_{CE}]$, where recall that $\overline{u}_C = \max\{u : \pi_C(u_C(u)) = \pi(u_C(u))\}$, $\pi(u)$ cannot be supported by centralization or uncooperative empowerment by Step 2 and Lemma 5A. Therefore, for any $u \in [\overline{u}_C, \bar{u}_{CE}]$, either $\pi(u) = \pi_{E_C}(u)$ (in which case we are done) or there exists a $\rho \in (0,1)$, a $\widehat{u}_1 \in [\overline{u}_C, u)$, and a $\widehat{u}_2 \in (u, \bar{u}_{CE}]$ such that (i.) both $(\widehat{u}_1, \pi(\widehat{u}_1))$ and

$(\widehat{u}_2, \pi(\widehat{u}_2))$ satisfy $\pi_{E_C}(\widehat{u}_i) = \pi(\widehat{u}_i)$ for $i = 1, 2$, (ii.) $(u, \pi(u)) = \rho(\widehat{u}_1, \pi(\widehat{u}_1)) + (1-\rho)(\widehat{u}_2, \pi(\widehat{u}_2))$. Let $\widehat{u}_{ih}$ and $\widehat{u}_{i\ell}$, $i = 1, 2$ be the agent's associated continuation payoffs for the two PPEs.

Now consider an alternative strategy profile in which cooperative empowerment is chosen and the continuation payoff given by $(\widehat{u}_h, \widehat{u}_\ell, \pi(\widehat{u}_h), \pi(\widehat{u}_\ell))$, where $\widehat{u}_h = \rho\widehat{u}_{1h} + (1-\rho)\widehat{u}_{2h}$ and $\widehat{u}_\ell = \rho\widehat{u}_{1\ell} + (1-\rho)\widehat{u}_{2\ell}$. It follows from the promise keeping constraints $\mathrm{PK}^{\mathrm{P}}_{\mathrm{E_C}}$ and $\mathrm{PK}^{\mathrm{A}}_{\mathrm{E_C}}$ that under this strategy profile the payoffs are given by $\widehat{u} = u$ and

$$
\begin{aligned}
\widehat{\pi} &= p\left[(1-\delta)B + \delta\pi\left(\rho\widehat{u}_{1h} + (1-\rho)\widehat{u}_{2h}\right)\right] \\
&\quad + (1-p)\left[(1-\delta)b + \delta\pi\left(\rho\widehat{u}_{1h} + (1-\rho)\widehat{u}_{2h}\right)\right] \\
&\geq \rho\pi(\widehat{u}_1) + (1-\rho)\pi(\widehat{u}_2) \\
&= \pi(u).
\end{aligned}
$$

It can be checked that this alternative strategy profile satisfies all the constraints and therefore constitutes a PPE. Therefore, $\pi_{E_C}(u) \geq \widehat{\pi} \geq \pi(u)$. This proves that $\pi_{E_C}(u) = \pi(u)$ if $u \in [\overline{u}_C, \bar{u}_{CE}]$.

Next, define $\underline{u}_{CE} = \min\{u : \pi_{E_C}(u) = \pi(u)\}$. By the above, we see that $\underline{u}_{CE} \leq \overline{u}_C$. In addition, since $p < 1/2$, it is clear that $\pi_{E_C}(u) < u = \pi(u)$ for all $u \in [0, a]$, where the equality follows from the proof of Lemma 5B. Therefore, $\underline{u}_{CE} > a$. It remains to show that $\pi_{E_C}(u) = \pi(u)$ if $u \in [\underline{u}_{CE}, \overline{u}_C]$. Suppose $u = \rho\underline{u}_{CE} + (1-\rho)\overline{u}_C$. Consider a strategy profile in which cooperative empowerment is chosen and the continuation payoff given by $(\widehat{u}_h, \widehat{u}_\ell, \pi(\widehat{u}_h), \pi(\widehat{u}_\ell))$, where $\widehat{u}_h = \rho u_h(\underline{u}_{CE}) + (1-\rho)u_h(\overline{u}_C)$ and $\widehat{u}_l = \rho u_\ell(\underline{u}_{CE}) + (1-\rho)u_\ell(\overline{u}_C)$. It follows from the promise keeping constraints $\mathrm{PK}^{\mathrm{P}}_{\mathrm{E_C}}$ and $\mathrm{PK}^{\mathrm{A}}_{\mathrm{E_C}}$ that under this strategy profile the payoffs are given by $\widehat{u} = u$ and

$$
\begin{aligned}
\widehat{\pi} &= p\left[(1-\delta)B + \delta\pi\left(\rho u_h(\underline{u}_{CE}) + (1-\rho)u_h(\overline{u}_C)\right)\right] \\
&\quad + (1-p)\left[(1-\delta)b + \delta\pi\left(\rho u_\ell(\underline{u}_{CE}) + (1-\rho)u_\ell(\overline{u}_C)\right)\right] \\
&\geq \rho\pi(\underline{u}_{CE}) + (1-\rho)\pi(\overline{u}_C) \\
&= \pi(u),
\end{aligned}
$$

where the last equality follows from the linearity of π in $[\underline{u}_{CE}, \overline{u}_C]$. Therefore, $\pi_{E_C}(u) \geq \widehat{\pi} \geq \pi(u)$, and, thus, $\pi_{E_C}(u) = \pi(u)$. This proves that $\pi_{E_C}(u) = \pi(u)$ for all $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$.

Next, we prove $\underline{u}_{CE} \leq \delta a + (1-\delta)B$, or equivalently, $u_\ell(\underline{u}_{CE}) \leq a$, and this finishes the proof of Step 3. Now suppose to the contrary that $u_\ell(\underline{u}_{CE}) > a$. This implies that $u_\ell(\overline{u}_C) > a$ since $\overline{u}_C \geq \underline{u}_{CE}$. Recall (in Lemma 5B) that π is a straight line in $[a, \overline{u}_C]$. Denote its slope as s. Now consider a strategy profile that is supported by cooperative empowerment and whose continuation payoffs are given by $(\widehat{u}_\ell, \widehat{u}_h, \pi(\widehat{u}_\ell), \pi(\widehat{u}_h))$, where $\widehat{u}_\ell = u_\ell(\overline{u}_C) - \varepsilon$ and $\widehat{u}_h = u_h(\overline{u}_C) - \varepsilon$ for $\varepsilon > 0$. We choose ε small enough so that $u_\ell(\overline{u}_C) - \varepsilon > a$ and $u_h(\overline{u}_C) - \varepsilon > \overline{u}_C$. It follows from the promise keeping constraints $\mathrm{PK}^{\mathrm{P}}_{\mathrm{E_C}}$ and $\mathrm{PK}^{\mathrm{A}}_{\mathrm{E_C}}$ that under this strategy profile the payoffs are given by $\widehat{u} = \overline{u}_C - \delta\varepsilon$ and

$$\begin{aligned}
\widehat{\pi} &= p\left[(1-\delta)B + \delta\pi\left(\widehat{u}_h\right)\right] + (1-p)\left[(1-\delta)b + \delta\pi\left(\widehat{u}_\ell\right)\right] \\
&= \pi\left(\overline{u}_C\right) + p\delta\left[\pi\left(u_h\left(\overline{u}_C\right) - \varepsilon\right) - \pi\left(u_h\left(\overline{u}_C\right)\right)\right] \\
&\quad + (1-p)\,\delta\left[\pi\left(u_\ell\left(\overline{u}_C\right) - \varepsilon\right) - \pi\left(u_\ell\left(\overline{u}_C\right)\right)\right] \\
&> \pi\left(\overline{u}_C\right) - \delta s\varepsilon.
\end{aligned}$$

Notice that the strict inequality follows because $\pi\left(u_\ell\left(\overline{u}_C\right) - \varepsilon\right) - \pi\left(u_\ell\left(\overline{u}_C\right)\right) = -s\varepsilon$ (since $u_\ell\left(\overline{u}_C\right) - \varepsilon > a$) and $\pi\left(u_h\left(\overline{u}_C\right) - \varepsilon\right) - \pi\left(u_h\left(\overline{u}_C\right)\right) > -s\varepsilon$ since $u_h(\overline{u}_C) - \varepsilon > \overline{u}_C$ and $\pi'_-(u) < s$ for all $u > \overline{u}_C$. It can be checked that this alternative strategy profile satisfies all the constraints and therefore constitutes a PPE. This implies that

$$\pi\left(\overline{u}_C\right) - \delta s\varepsilon < \widehat{\pi} \le \pi\left(\overline{u}_C - \delta\varepsilon\right) = \pi(\overline{u}_C) - \delta s\varepsilon,$$

where the weak inequality follows from the definition of π and the equality follows from the fact that π is a straight line with slope s between a and $\overline{u}_C$. Since the above chain of inequalities lead to a contradiction, we must have $u_\ell(\underline{u}_{CE}) \le u_\ell(\overline{u}_C) \le a$, or equivalently $\underline{u}_{CE} \le \delta a + (1-\delta)B$. This proves Step 3, and, thus, part (iii.) of Lemma 5C.

Step 4: If $p \in [1/2, p^*)$, $\pi_{E_C}(u) = \pi(u)$ if and only if $u \in [\underline{u}^A_{CE}, \bar{u}^A_{CE}] \cup [\underline{u}_{CE}, \bar{u}_{CE}]$ for some $\bar{u}^A_{CE} < a < \underline{u}_{CE} \le \delta a + (1-\delta)B$, where $\underline{u}^A_{CE} = (1-\delta)B$.

To prove this, notice that for $u \ge a$, we can again define $\underline{u}_{CE} = \min\{u : \pi_{E_C}(u) = \pi(u), u \ge a\}$. Using the same argument as in Step 3, we can show that when $u > a$, $\pi_{E_C}(u) = \pi(u)$ if and only if $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$. Next, notice that when $p \ge 1/2$, $\pi_{E_C}(u) \ge u$ for all u such that $u_\ell(u) \ge 0$ and $u_h(u) \le a$, i.e., $u \in [(1-\delta)B, (1-\delta)b + \delta a]$. This implies that $\pi_{E_C}(u) = \pi(u)$ for some $u \in (0, a)$ (because otherwise $\pi(u) = u$ for all $u \in (0, a)$). Now define $\bar{u}^A_{CE} = \max\{u : \pi_{E_C}(u) = \pi(u), u \le a\}$ and $\underline{u}^A_{CE} = \min\{u : \pi_D(u) = \pi(u)\}$. By the same argument as in Step 3, we can show that $\pi_{E_C}(u) = \pi(u)$ for all $u \in [\underline{u}^A_{CE}, \bar{u}^A_{CE}]$.

By the definition of $\underline{u}^A_{CE}$, it is clear that $\pi(u)$ is a straight line between 0 and $\underline{u}^A_{CE}$. Let the slope of this line segment be s. To see that $\underline{u}^A_{CE} = (1-\delta)B$, notice that $\underline{u}^A_{CE} \ge (1-\delta)B$ because otherwise cooperative empowerment is not feasible. Suppose to the contrary that $\underline{u}^A_{CE} > (1-\delta)B$. Consider a strategy profile that is supported by cooperative empowerment and whose continuation payoffs are given by $(\widehat{u}_\ell, \widehat{u}_h, \pi\left(\widehat{u}_\ell\right), \pi\left(\widehat{u}_h\right))$, where $\widehat{u}_\ell = u_\ell(\underline{u}^A_{CE}) - \varepsilon$ and $\widehat{u}_h = u_h(\underline{u}^A_{CE}) - \varepsilon$ for $\varepsilon > 0$. We choose ε small enough so that $u_\ell(\underline{u}^A_{CE}) - \varepsilon > 0$. It follows from the promise-keeping constraints

$\mathrm{PK}^{\mathrm{P}}_{\mathrm{E_C}}$ and $\mathrm{PK}^{\mathrm{A}}_{\mathrm{E_C}}$ that under this strategy profile the payoffs are given by $\widehat{u} = \underline{u}^A_{CE} - \delta\varepsilon$ and

$$\begin{aligned}
\widehat{\pi} &= p\left[(1-\delta)B + \delta\pi\left(\widehat{u}_h\right)\right] + (1-p)\left[(1-\delta)b + \delta\pi\left(\widehat{u}_\ell\right)\right] \\
&= \pi\left(\underline{u}^A_{CE}\right) + p\delta\left[\pi\left(u_h\left(\underline{u}^A_{CE}\right) - \varepsilon\right) - \pi\left(u_h\left(\underline{u}^A_{CE}\right)\right)\right] \\
&\quad + (1-p)\,\delta\left[\pi\left(u_\ell\left(\underline{u}^A_{CE}\right) - \varepsilon\right) - \pi\left(u_\ell\left(\underline{u}^A_{CE}\right)\right)\right] \\
&\geq \pi\left(\underline{u}^A_{CE}\right) - \delta s\varepsilon.
\end{aligned}$$

Notice inequality follows because π is concave so s is its maximal derivative. It can be checked that this alternative strategy profile satisfies all the constraints and therefore constitutes a PPE. This implies that

$$\pi_{E_C}\left(\underline{u}^A_{CE} - \delta\varepsilon\right) \geq \widehat{\pi} \geq \pi\left(\underline{u}^A_{CE}\right) - \delta s\varepsilon = \pi\left(\underline{u}^A_{CE} - \delta\varepsilon\right),$$

where the equality follows from that π is a straight line with slope s between 0 and $\underline{u}^A_{CE}$. By the above chain of inequalities, $\pi_{E_C}\left(\underline{u}^A_{CE} - \delta\varepsilon\right) = \pi\left(\underline{u}^A_{CE} - \delta\varepsilon\right)$, which leads to a contradiction because $\underline{u}^A_{CE}$ is defined as the smallest agent's payoff such that $\pi_{E_C}(u) = \pi(u)$.

Next, notice that we must have $\bar{u}^A_{CE} < a$. Suppose to the contrary that $\bar{u}^A_{CE} = a$, then by definition we also have $\underline{u}_{CE} = a$, and this implies that $\pi(u) = \pi_{E_C}(u)$ for all $u \in [\underline{u}^A_{CE}, \bar{u}_{CE}]$. This implies that $\pi(u) = \pi_T(u)$, where recall that $\pi_T(u)$ is defined in Step 1 (where centralization is not used). This contradicts that assumption that $p < p^{**}$. Similarly, we must also have $\underline{u}_{CE} > a$. This proves Step 4, and, thus, part (ii) of Lemma 5C. ■

Combining Lemma 5A-5C, we obtain the following lemma.

LEMMA 5'. *The PPE frontier $\pi(u)$ satisfies the following:*

(i.) If $p \in [p^{**}, 1)$, the frontier can be divided into three regions.

$$\pi(u) = \begin{cases} u\pi(\underline{u}_{CE})/\underline{u}_{CE} & u \in [0, \underline{u}_{CE}); \\ \pi_{E_C}(u) & u \in [\underline{u}_{CE}, \bar{u}_{CE}]; \\ \left((B-u)\,\pi(\bar{u}_{CE}) + (u - \bar{u}_{CE})\,b\right)/\left(B - \bar{u}_{CE}\right) & u \in (\bar{u}_{CE}, B], \end{cases}$$

where $\underline{u}_{CE} = (1-\delta)\,B$ and $\bar{u}_{CE} = (1-\delta)\,b + \delta B$.

(ii.) If $p \in (1/2, p^*)$*, the frontier can be divided into six regions.*

$$\pi(u) = \begin{cases} u\pi(\underline{u}_{CE}^A)/\underline{u}_{CE}^A & u \in [0, \underline{u}_{CE}^A); \\ \pi_{E_C}(u) & u \in [\underline{u}_{CE}^A, \bar{u}_{CE}^A]; \\ \left((a-u)\,\pi(\bar{u}_{CE}^A) + \left(u - \bar{u}_{CE}^A\right) a\right) / \left(a - \bar{u}_{CE}^A\right) & u \in (\bar{u}_{CE}^A, a); \\ ((\underline{u}_{CE} - u)\, a + (u-a)\,\pi(\underline{u}_{CE})) / (\underline{u}_{CE} - a) & u \in [a, \underline{u}_{CE}); \\ \pi_{E_C}(u) & u \in [\underline{u}_{CE}, \bar{u}_{CE}]; \\ ((B-u)\,\pi(\bar{u}_{CE}) + (u - \bar{u}_{CE})\, b) / (B - \bar{u}_{CE}) & u \in (\bar{u}_{CE}, B], \end{cases}$$

where $\underline{u}_{CE}^A = (1-\delta)\,B$, $\underline{u}_{CE} \in (a, \delta a + (1-\delta)\,B)$ *and* $\bar{u}_{CE} = (1-\delta)\,b + \delta B$.

(iii.) If $p \in (0, 1/2]$*, the frontier can be divided into four regions.*

$$\pi(u) = \begin{cases} u & u \in [0, a); \\ ((\underline{u}_{CE} - u)\, a + (u-a)\,\pi(\underline{u}_{CE})) / (\underline{u}_{CE} - a) & u \in [a, \underline{u}_{CE}); \\ \pi_{E_C}(u) & u \in [\underline{u}_{CE}, \bar{u}_{CE}]; \\ ((B-u)\,\pi(\bar{u}_{CE}) + (u - \bar{u}_{CE})\, b) / (B - \bar{u}_{CE}) & u \in (\bar{u}_{CE}, B], \end{cases}$$

where $\underline{u}_{CE} \in (a, \delta a + (1-\delta)\,B)$ *and* $\bar{u}_{CE} = (1-\delta)\,b + \delta B$.

Proposition 2', which we now turn to, contains Propositions 1 and 2 as special cases.

PROPOSITION 2'. *In the optimal equilibrium, the principal chooses cooperative empowerment if and only if* $t \leq \tau$ *for some random time period* τ. $\Pr(\tau < \infty) = 1$ *so that cooperative empowerment occurs with probability 0 in the long run. In addition, there exists* $p^* < p^{**}$ *that the following holds:*

(i.) When $p \geq p^{**}$*, the relationship either terminates or settles into permanent empowerment in the long run:*

$$\begin{aligned} \lim_{t\to\infty} \Pr(u_t &= 0) > 0, \ \lim_{t\to\infty} \Pr(u_t = B) > 0, \\ \lim_{t\to\infty} \Pr(u_t &= 0) + \Pr(u_t = B) = 1. \end{aligned}$$

(ii.) When $p \in (p^*, p^{**})$*, the relationship terminates with positive probability less than 1. When the relationship does not terminate in the long run, principal chooses either permanent disenfranchisement or permanent empowerment.*

$$\begin{aligned} \lim_{t\to\infty} \Pr(u_t &= 0) > 0, \ \lim_{t\to\infty} \Pr(u_t = a) > 0, \ \lim_{t\to\infty} \Pr(u_t = B) > 0, \\ \lim_{t\to\infty} \Pr(u_t &= 0) + \Pr(u_t = a) + \Pr(u_t = B) = 1. \end{aligned}$$

(*iii.*) *When* $p \leq p^*$, *the principal chooses either permanent disenfranchisement or permanent empowerment in the long run:*

$$\begin{aligned}\lim_{t\to\infty}\Pr(u_t &= a) > 0,\ \lim_{t\to\infty}\Pr(u_t = B) > 0,\\ \lim_{t\to\infty}\Pr(u_t &= a) + \Pr(u_t = B) = 1.\end{aligned}$$

Proof of Proposition 2': By Assumption 1, the relationship starts with cooperative empowerment. In addition, if ever centralization, uncooperative empowerment, or exit is used, Lemma 5' immediately implies that the relationship stops there forever. This establishes the existence of the random time. Moreover, since these are the only absorbing states of the relationship, and it can be easily show that there exists an $\varepsilon > 0$ and a large enough N such that the probability the relationship ends in one of the absorbing state exceeds ε every N periods. Standard arguments then imply that the relationship ends in one of the absorbing states with probability 1, and this shows that $\Pr(\tau < \infty) = 1$.

Part (i): when $p \geq p^{**}$, the PPE frontier is described as in part (i) of Lemma 5'. In this case, it is clear that the dynamics following the optimal equilibrium has only two steady states: permanent empowerment and termination. This proves part (i).

Next, suppose $p < p^{**}$. Notice that the relationship starts to the right of a, and if there exists t such that $\Pr(u_t < a) > 0$, then termination is a steady state with positive probability. Otherwise, termination is never reached and the only two steady states are permanent disenfranchisement and permanent empowerment. By Lemma 5' (part (ii) and (iii)), $u_\ell(\underline{u}_{CE}) \leq a$. It is then clear that $\Pr(u_t < a) = 0$ if and only if $u_\ell(\underline{u}_{CE}) = a$. We now show below that there exists p^* such that for all $p \leq p^*$, $u_\ell(\underline{u}_{CE}) = a$.

To do this, we show that if $u_\ell(\underline{u}_{CE}, p') = a$ then $u_\ell(\underline{u}_{CE}, p'') = a$ for $p'' < p'$. Let s_0 be the slope between (a, a) and $(\underline{u}_{CE}, \pi(\underline{u}_{CE}))$, and s_1 be the slope between $(\bar{u}_{CE}, \pi(\bar{u}_{CE}))$ and (B, b). Note that both s_0 and s_1 depend on p. Now if $u_\ell(\underline{u}_{CE}, p') = a$, this implies that

$$s_0\left(p'\right) \leq (1-p')\pi'_-(a, p') + p'\pi'_-(u_h(\underline{u}_{CE}), p').$$

To show that $u_\ell(\underline{u}_{CE}, p'') = a$ for $p'' < p'$, it then suffices to show that

$$s_0\left(p''\right) \leq (1-p'')\pi'_-(a, p'') + p''\pi'_-(u_h(\underline{u}_{CE}), p'').$$

This is because we can use the standard argument to show that $\pi(u)$ is the unique fixed point of the operator

$$Tf(u) = \max \left\{ \begin{array}{c} \max \{\pi_C(u), \pi_{E_C}(u), \pi_{E_U}(u), \pi_X(u)\}, \\ \max_{\alpha \in (0,1), u_1, u_2 \in [0,B]} \{\alpha f(u_1) + (1-\alpha) f(u_2)\} \end{array} \right\},$$

which is monotone and non-expansive (see an analogous argument in Step 2 in Lemma 5C).

A sufficient condition is then given by

$$\frac{\partial s_0(p)}{\partial p} > \partial \left((1-p)\pi'_-(a,p) + p\pi'_-(u_h(\underline{u}_{CE}),p) \right) / \partial p$$

for all $p < p'$. Notice that if $\partial s_0(p)/\partial p$ does not exist, we can replace it with the left derivative. To see that the left derivative exists, notice that $\pi(u)$ is weakly increasing in p for all u, so $s_0(p)$ is weakly increasing in all p, and therefore, the left derivative exists. Now

$$\begin{aligned} & \partial \left((1-p)\pi'_-(a,p) + p\pi'_-(u_h(\underline{u}_{CE}),p) \right) / \partial p \\ = \quad & -\pi'_-(a,p) + (1-p)\frac{\partial \pi'_-(a,p)}{\partial p} + \pi'_-(u_h(\underline{u}_{CE}),p) + p\frac{\partial \pi'_-(u_h(\underline{u}_{CE}),p)}{\partial p} \\ \leq \quad & p\frac{\partial \pi'_-(u_h(\underline{u}_{CE}),p)}{\partial p}, \end{aligned}$$

where the inequality follows because $\pi'_-(u_h(\underline{u}_{CE}),p) - \pi'_-(a,p) \leq 0$ by concavity of π and $\partial \pi'_-(a,p)/\partial p \leq 0$ because $\pi(a,p) = a$ for all $p < p^*$ and $\pi(u,p)$ is weakly increasing in p. It now follows that it suffices to show that

$$\frac{\partial s_0(p)}{\partial p} > p\frac{\partial \pi'_-(u_h(\underline{u}_{CE}),p)}{\partial p}.$$

To do this, notice that we can write

$$\pi'_-(u_h(\underline{u}_{CE}),p) = \alpha(p)s_0(p) + (1-\alpha(p))s_1(p)$$

for some $\alpha(p)$. Notice that $\partial s_1(p)/\partial p \leq 0$ since $\pi(u,p)$ increasing in p, which follows from an argument analogous to that in Step 2 in Lemma 5C. It follows that if $\partial \alpha(p)/\partial p < 0$, then

$$\begin{aligned} & p\frac{\partial \pi'_-(u_h(\underline{u}_{CE}),p)}{\partial p} \\ = \quad & p\left(\frac{\partial \alpha(p)}{\partial p}s_0(p) + \alpha(p)\frac{\partial s_0(p)}{\partial p} + (1-\alpha(p))\frac{\partial s_1(p)}{\partial p} - \frac{\partial \alpha(p)}{\partial p}s_1(p) \right) \\ \leq \quad & p\alpha(p)\frac{\partial s_0(p)}{\partial p} \\ \leq \quad & \frac{\partial s_0(p)}{\partial p}. \end{aligned}$$

To show that $\partial\alpha(p)/\partial p < 0$, notice that $\alpha(p)$ is the probability that u_t falls into $[0, a]$ before $(\bar{u}_{CE}, B]$, where $\{u_t\}$ is a sequence of the agent payoffs starting at $u_1 = \underline{u}_{CE}$ and determined by the transition rule that

$$u_{t+1} = \begin{cases} u_h(u_t) & \text{with probability } p \\ u_\ell(u_t) & \text{with probability } 1-p \end{cases}.$$

Notice that we have the open bracket at u_1 and closed bracket at a because we are looking at the left derivatives here. In general, let $F(u,p)$ be the probability that u_t falls into $[0, a]$ first before $(\bar{u}_{CE}, B]$, and the sequence starts at u. In particular, let $F(u,p) = 1$ for $u \leq 0$ and $F(u,p) = 1 = 0$ for $u \geq B$. Now consider $p_1 < p_2$. For each p_i, $F(u,p_i)$ satisfies $F(u,p_i) = T_i F(u,p_i)$, where T_i is an operator on functions with ranges in $[a, B]$ satisfying

$$T_i F(u) = \begin{cases} \frac{u-a}{\underline{u}_{CE}-a} F(\underline{u}_{CE}) + \frac{\underline{u}_{CE}-u}{\underline{u}_{CE}-a} & 0 \leq u < \underline{u}_{CE} \\ p_i F(u_h(u)) + (1-p_i) F(u_\ell(u)) & \underline{u}_{CE} \leq u \leq \bar{u}_{CE} \\ \frac{B-u}{B-\bar{u}_{CE}} F(\bar{u}_{CE}) + \frac{u-\bar{u}_{CE}}{B-\bar{u}_{CE}} b & \bar{u}_{CE} < u \leq B \end{cases}$$

Notice that T_i is a monotone operator and has a unique fixed point for each p_i. Moreover, it is clear that $F(u,p_i)$ is decreasing in u for each p_i. Let $F(u,p_1)$ be the unique solution for $F(u,p_1) = T_1 F(u,p_1)$. It is clear that $T_2 F(u,p_1) \leq F(u,p_1)$ (since $F(u,p_1)$ is decreasing in u) for all u and it follows that $F(u,p_2) \leq F(u,p_1)$ for all u. In particular, $F(\underline{u}_{CE}, p_2) \leq F(\underline{u}_{CE}, p_1)$ and this proves that $\partial\alpha(p)/\partial p < 0$. This finishes showing that if $u_\ell(\underline{u}_{CE}, p') = a$ then $u_\ell(\underline{u}_{CE}, p'') = a$ for $p'' < p'$. As a result, there exists p^* such that for all $p \leq p^*$, $u_\ell(\underline{u}_{CE}) = a$ and there are two steady states in the long run.

Finally, we show that $p^* < p^{**}$. It is clear that $p^* \leq p^{**}$, so it suffices to rule out that $p^* = p^{**}$. Recall that at $p = p^{**}$, we have $\pi^T(a, p^{**}) = a$, and $\pi^T(u, p^{**}) = \pi(u, p^{**})$. We now show that there exists some small $\varepsilon > 0$ such that for $p = p^{**} - \varepsilon$, $u_\ell(\underline{u}_{CE}, p) < a$. Then by the argument above, we have $p^* < p^{**} - \varepsilon$. Define $u_a \equiv u_\ell^{-1}(a)$ and let $s(p^{**})$ be the slope between (a, a) and $(u_a, \pi(u_a, p^{**}))$. Now suppose to the contrary that that for all $\varepsilon > 0, u_\ell(\underline{u}_{CE}, p^{**} - \varepsilon) = a$. Let $s(p)$ be the slope between (a, a) and $(u_a, \pi(u_a, p))$. Notice that for all $p < p^{**}$, $\pi(a, p) = a$. In addition, $\pi(u_a, p) \leq \pi(u_a, p^{**})$ because π is weakly increasing in p. This implies that $s(p) \leq s(p^{**})$ for all $p < p^{**}$. Now there are two cases to consider. First, $\pi(u, p^{**})$ is not a straight line between a and u_a, i.e., there exists a $u \in (a, u_a)$ such that $\pi(u, p^{**}) > a + s(p^{**})(u - a)$. In this case, we have

$$\pi(u, p^{**}) > a + s(p^{**})(u - a) \geq \lim_{\varepsilon \to 0} \pi(u, p^{**} - \varepsilon),$$

which violates the continuity of $\pi(u, p)$ in p. This is a contradiction.

In the second case, $\pi(u, p^{**})$ is a straight line between a and u_a. Notice that since $u_\ell(u_a) = a$, we have $a < u_a \leq \bar{u}_{CE}$. Lemma 5' then implies that for $u \in (a, u_a)$,

$$\pi'(u, p^{**}) = p\pi'(u_h(u), p^{**}) + (1-p)\pi'(u_\ell(u), p^{**}).$$

Since $\pi'(u, p^*)$ is a constant for $u \in (a, u_a)$ and because π is concave, we have that π' is again a constant for all $u \in (u_\ell(a), a)$. Notice that

$$a - u_\ell(a) = (u_a - a)/\delta > u_a - a.$$

Now either $u_\ell(a) \geq \underline{u}_{CE}^A \equiv (1-\delta) B$ or $u_\ell(a) < \underline{u}_{CE}^A$. Recall that π' is a constant for $u \leq \underline{u}_{CE}^A$, and denote the slope of π in this region as s_0. Notice that if $u_\ell(a) < \underline{u}_{CE}^A$, we must have $u_\ell(a) = 0$ since $\pi'_-(u) < s_0$ for all $u > \underline{u}_{CE}^A$. Now if $u_\ell(a) \geq \underline{u}_{CE}^A$, we then have that π' is a constant for $u \in (u_{\ell\ell}(a), u_\ell(a))$, and repeating the same argument, we must have that $u_{\ell^n}(a) = 0$ for some n.

Next, we claim that $u_h(u_1) \geq \bar{u}_{CD} \equiv (1-\delta) b + \delta B$. Notice that π' is a constant for $u \in (u_h(\underline{u}_{CE}^A), u_h(\bar{u}_{CE}^A))$, and

$$u_h(\bar{u}_{CE}^A) - u_h(\underline{u}_{CE}^A) = \left(\bar{u}_{CE}^A - \underline{u}_{CE}^A\right)/\delta.$$

If $u_h(u_1) \leq \bar{u}_{CE}$, we must then also have $u_h(\bar{u}_{CE}^A) \leq \bar{u}_{CE}$ (because $\pi'_-(u)$ for $u < \bar{u}_{CE}$ is strictly larger than the slope of π for $u \geq \bar{u}_{CE}$.) But if π' is a constant for $u \in (u_h(u_1), u_h(\bar{u}_{CE}^A))$, same argument as above implies that π' is a constant for $u \in (u_{\ell h}(\underline{u}_{CE}^A), u_{\ell h}(\bar{u}_{CE}^A))$, which is a line segment with length $\left(\bar{u}_{CE}^A - \underline{u}_{CE}^A\right)/\delta^2$. Applying the same argument as before, this would imply that the distance between $\bar{u}_{CE}^A$ and $\underline{u}_{CE}^A$ is bigger than $\left(\bar{u}_{CE}^A - \underline{u}_{CE}^A\right)/\delta^2$, which is a contradiction.

Finally, since $\underline{u}_{CE}^A = (1-\delta) B$, $u_h(\underline{u}_{CE}^A) \geq \bar{u}_{CE}$ then implies that

$$\frac{(1-\delta)(B-b)}{\delta} \geq B - (1-\delta)b > B - a,$$

which contradicts Assumption 1. ■

Appendix B: Optimal Relational Contract with Public Opportunities

Just as in the main section, we solve the game recursively by characterizing the PPE payoff sets. Define $\mathcal{E}_{Pre}$ as the PPE payoff set of the pre-opportunity phase and $\mathcal{E}_{Post}$ as the PPE payoff set of the post-opportunity phase. Let $\pi_i(u)$, $i = 1, 2$, be the associated payoff frontier.

Constraints in the Pre-Opportunity Phase

We first list the set of constraints for supporting the PPE payoff set $\mathcal{E}_{Post}$. Consider a PPE payoff pair $(u, \pi) \in \mathcal{E}_{Post}$. If $(u, \pi) = (0, 0)$, then the payoff pair is sustained by exit (X). Otherwise, there are five possibilities: a) support with centralization (C); b) support with cooperative empowerment (E_C); c) support with uncooperative empowerment (E_U); d) support with the new opportunity (N); e) support with randomization.

First, suppose (u, π) is supported with action $j \in \{C, E_U, N\}$. There exists a pair of continuation payoffs (u_j, π_j) that satisfies the following promise-keeping constraints:

$$u = (1-\delta)\, u(j) + \delta u_j; \tag{PK$_{\text{2Aj}}$}$$

$$\pi = (1-\delta)\, \pi(j) + \delta \pi_j, \tag{PK$_{\text{2Pj}}$}$$

where $u(j)$ and $\pi(j)$ are the agent's and principal's per period payoff associated with action j. In addition, the self-enforcing constraint requires that

$$(u_j, \pi_j) \in \mathcal{E}_{Post}. \tag{SE$_{\text{2j}}$}$$

Second, suppose (u, π) can be supported with cooperative empowerment. In this case, there exist continuation payoffs $(u_\ell, \pi_\ell, u_h, \pi_h)$ that satisfy the conditions to be described below. As in the main section, let u_ℓ and u_h be the agent's continuation payoff when he chooses his own project and the principal's project respectively, and π_ℓ and π_h are the principal's payoff defined analogously. To support (u, π), we need the following conditions to be satisfied. In order for the agent to choose the principal's action when it is available, we need

$$(1-\delta)\, b + \delta u_h \geq (1-\delta)\, B + \delta u_\ell. \tag{IC$_{\text{2D}}$}$$

In addition, the promise-keeping constraints are given by

$$u = p\left[(1-\delta)\, b + \delta u_h\right] + (1-p)\left[(1-\delta)\, B + \delta u_\ell\right] \tag{PK$_{\text{2AD}}$}$$

for the agent and

$$\pi = p\left[(1-\delta)\, B + \delta \pi_h\right] + (1-p)\left[(1-\delta)\, b + \delta \pi_\ell\right] \tag{PK$_{\text{2PD}}$}$$

for the principal. Finally, the self-enforcing conditions require that

$$(u_h, \pi_h) \in \mathcal{E}_{Post}, (u_\ell, \pi_\ell) \in \mathcal{E}_{Post.} \tag{SE$_{\text{2D}}$}$$

Third, (u, π) can be supported with randomization. In this case, there exists two distinct PPE payoffs $(u_i, \pi_i) \in \mathcal{E}_{Post}$, $i = 1, 2$ such that

$$(u, \pi) = \alpha(u_1, \pi_1) + (1-\alpha)(u_2, \pi_2) \text{ for some } \alpha \in (0, 1).$$

Constraints in the Post-Opportunity Phase

Next, we list the set of constraints for supporting the PPE payoff set $\mathcal{E}_{Pre}$. Consider a PPE payoff pair $(u, \pi) \in \mathcal{E}_{Pre}$. If $(u, \pi) = (0, 0)$, then the payoff pair is supported by exit (X). Otherwise, there are four possibilities: a) support with centralization (C); b) support with cooperative empowerment (E_C); c) support with uncooperative empowerment (E_U); d) support with randomization.

Notice that if (u, π) is supported with action $j \in \{C, E_U\}$, there exists pairs of continuation payoffs $(u_{Pre,j}, \pi_{Pre,j})$ and $(u_{Post,j}, \pi_{Post,j})$ that satisfy the following promise-keeping constraints:

$$\begin{aligned} u &= (1-\delta)\, u\,(j) + \delta\left[(1-q)\, u_{Pre,j} + q u_{Post,j}\right]; && (\mathrm{PK}_{1\mathrm{Aj}}) \\ \pi &= (1-\delta)\, \pi\,(j) + \delta\left[(1-q)\, \pi_{Pre,j} + q \pi_{Post,j}\right]. && (\mathrm{PK}_{1\mathrm{Pj}}) \end{aligned}$$

In addition, the self-enforcing constraint requires that

$$(u_{Pre,j}, \pi_{Pre,j}) \in \mathcal{E}_{Pre} \text{ and } (u_{Post,j}, \pi_{Post,j}) \in \mathcal{E}_{Post} \qquad (\mathrm{SE}_{1\mathrm{C}})$$

Next, (u, π) can be supported with cooperative delegation. The incentive compatibility condition requires the associated continuation payoffs $(u_{Pre,\ell}, \pi_{Pre,\ell}, u_{Pre,h}, \pi_{Pre,h})$ and $(u_{Post,\ell}, \pi_{Post,\ell},$ $u_{Post,h}, \pi_{Post,h})$ to satisfy

$$(1-\delta)\, b + \delta\left((1-q)\, u_{Pre,h} + q u_{Post,h}\right) \geq (1-\delta)\, B + \delta\left((1-q)\, u_{Pre,\ell} + q u_{Post,\ell}\right). \qquad (\mathrm{IC}_{1\mathrm{D}})$$

The promise-keeping constraints are given by

$$\begin{aligned} u \;=\; & p\left[(1-\delta)\, b + \delta\left((1-q)\, u_{Pre,h} + q u_{Post,h}\right)\right] && (\mathrm{PK}_{1\mathrm{AD}}) \\ & + (1-p)\left[(1-\delta)\, B + \delta\left((1-q)\, u_{Pre,\ell} + q u_{Post,\ell}\right)\right] \end{aligned}$$

for the agent and

$$\begin{aligned} \pi \;=\; & p\left[(1-\delta)\, B + \delta\left((1-q)\, \pi_{Pre,h} + q \pi_{Post,h}\right)\right] && (\mathrm{PK}_{1\mathrm{PD}}) \\ & + (1-p)\left[(1-\delta)\, b + \delta\left((1-q)\, \pi_{Pre,\ell} + q \pi_{Post,\ell}\right)\right] \end{aligned}$$

for the principal. In addition, the self-enforcing conditions require that

$$\begin{aligned} (u_{Pre,\ell}, \pi_{Pre,\ell}) &\in \mathcal{E}_{Pre}, (u_{Pre,h}, \pi_{Pre,h}) \in \mathcal{E}_{Pre}, && (\mathrm{SE}_{2\mathrm{D}}) \\ (u_{Post,\ell}, \pi_{Post,\ell}) &\in \mathcal{E}_{Post}, (u_{Post,h}, \pi_{Post,h}) \in \mathcal{E}_{Post}. \end{aligned}$$

Finally, (u, π) can be supported with randomization. In this case, there exists two distinct PPE payoffs $(u_i, \pi_i) \in \mathcal{E}_{Pre}$, $i = 1, 2$ such that

$$(u, \pi) = \alpha(u_1, \pi_1) + (1-\alpha)\,(u_2, \pi_2) \text{ for some } \alpha \in (0, 1).$$

Properties of π_{Post}

Lemma 6 below shows that $\pi_{Post}(u)$ shares similar features as the PPE payoff frontier in the main section.

LEMMA 6. *For any payoff $(u, \pi_{Post}(u))$ on the frontier, the equilibrium continuation payoffs remain on the frontier. In addition, the following holds:*

(i.) If $(u, \pi_{Post}(u))$ is supported with centralization, the agent's continuation payoff is given by

$$\delta u_C(u) = u - (1-\delta)a.$$

(ii.) If $(u, \pi_{Post}(u))$ is supported with cooperative empowerment, the agent's continuation payoff can be chosen to be

$$\begin{aligned} \delta u_h(u) &= u - (1-\delta)b; \\ \delta u_\ell(u) &= u - (1-\delta)B. \end{aligned}$$

(iii.) If $(u, \pi_{Post}(u))$ is supported with uncooperative empowerment, the agent's continuation payoff is given by

$$\delta u_{E_U}(u) = u - (1-\delta)B.$$

(iv.) If $(u, \pi_{Post}(u))$ is supported with exit, the agent's continuation payoff is given by

$$\delta u_X(u) = u.$$

(v.) If $(u, \pi_{Post}(u))$ is supported with the new opportunity, the agent's continuation payoff is given by

$$\delta u_N(u) = u - (1-\delta)u_N.$$

Proof of Lemma 6: This is proved in the same way as Lemma 3 in the main section and is omitted here.

Just as in the main section, let $\pi_{Post,j}(u)$ for $j \in \{C, E_C, E_U, X, N\}$ be the highest equilibrium payoff of the principal when the agent's payoff is u and action j is chosen. Lemma 6 implies that

$$
\begin{aligned}
\pi_{Post,C}(u) &= (1-\delta)\,a + \delta\pi_{Post}(u_C(u))\,; \\
\pi_{Post,E_C}(u) &= p\left[(1-\delta)\,B + \delta\pi_{Post}(u_h(u))\right] + (1-p)\left[(1-\delta)\,b + \delta\pi_{Post}(u_\ell(u))\right]; \\
\pi_{Post,E_U}(u) &= (1-\delta)\,b + \delta\pi_{Post}(u_{E_U}(u))\,; \\
\pi_{Post,X}(u) &= \delta\pi_{Post}(u_X(u))\,; \\
\pi_{Post,N}(u) &= (1-\delta)\,\pi_N + \delta\pi_{Post}(u_N(u))\,.
\end{aligned}
$$

Next, depending on the value of (u_N, π_N), $\pi_{Post}(u)$ can be divided up into seven regions. To focus our analysis but to allow for sufficient generality, we make the following assumptions that make sure all seven regions exist.

The first assumption restricts u_N to be in an intermediate range.

$$
u_h((1-\delta)\,B + \delta a) < u_N < u_\ell((1-\delta)\,b + \delta B)\,; \tag{N_u}
$$

Here, u_h and u_ℓ are given in Lemma 6. Notice that $u_\ell((1-\delta)\,B + \delta a) = a$ and $u_h((1-\delta)\,b + \delta B) = B$.

The second assumption restricts that π_N to be in an intermediate range.

$$
\pi(u_N) < \pi_N < \min\left\{B + b - u_N, a + \frac{pB + (1-p)\,b - a}{pb + (1-p)\,B - a}(u_N - a)\right\} \tag{N_π}
$$

The first inequality ensures that the new action will be chosen to support the payoff frontier. The second inequality ensures that cooperative empowerment will be chosen to support the payoff frontier. We denote N as the set of (u_N, π_N) such that both N_u and N_π are satisfied.

LEMMA 7. *For each* $(u_N, \pi_N) \in N$ *, there exist four cutoffs* $\underline{u}_{Post,CE}$, $\bar{u}_{Post,CE}$, $\underline{u}_N$, and $\bar{u}_N$ *such that the PPE payoff frontier* $\pi_{Post}(u)$ *is divided into seven regions:*

(i.) For $u \in [0, a]$, $\pi_{Post}(u) = u$ *and* $(u, \pi_{Post}(u))$ *is supported by randomization between exit and centralization.*

(ii.) For $u \in \left[a, \underline{u}_{Post,CE}\right]$, $\pi_{Post}(u) = \left(\left(\underline{u}_{Post,CE} - u\right) a + (u-a)\,\pi_{Post}\left(\underline{u}_{Post,CE}\right)\right) / \left(\underline{u}_{Post,CE} - a\right)$ *and* $(u, \pi_{Post}(u))$ *is supported by randomization between centralization and cooperative empowerment.*

(iii.) For $u \in \left[\underline{u}_{Post,CE}, \underline{u}_N\right]$, $\pi_{Post}(u) = \pi_{Post,E_C}(u)$ *and* $(u, \pi_{Post}(u))$ *is supported by cooperative empowerment.*

(iv.) For $u \in [\underline{u}_N, u_N]$, $\pi_{Post}(u) = ((u_N - u)\,\pi_{Post}(\underline{u}_N) + (u - \underline{u}_N)\,\pi_N) / (u_N - \underline{u}_N)$ *and* $(u, \pi_{Post}(u))$ *is supported by a randomization between cooperative empowerment and the new opportunity.*

(v.) For $u \in [u_N, \bar{u}_N]$, $\pi_{Post}(u) = ((\bar{u}_N - u)\pi_N + (u - u_N)\pi_{Post}(\bar{u}_N)) / (\bar{u}_N - u_N)$ and $(u, \pi_{Post}(u))$ is supported by a randomization between cooperative empowerment and the new opportunity.

(vi.) For $u \in [\bar{u}_N, \bar{u}_{Post,CE}]$, $\pi_{Post}(u) = \pi_{Post,E_C}(u)$ and $(u, \pi_{Post}(u))$ is supported by cooperative empowerment.

(vii.) For $u \in [\bar{u}_{Post,CE}, B]$, $\pi_{Post}(u) = ((B-u)\pi_{Post}(\bar{u}_{Post,CE}) + (u - \bar{u}_{Post,CE})b) / (B - \bar{u}_{Post,CE})$ and $(u, \pi_{Post}(u))$ is supported by a randomization between cooperative and uncooperative empowerment.

Moreover, for each u_N there exists $\widehat{\pi}_{Post}(u_N)$ such that $\pi_N = \max_u \{\pi_{Post}(u)\}$ for all $\pi_N \geq \widehat{\pi}_{Post}(u_N)$.

Proof of Lemma 7: This lemma is proved in a similar way as Lemma 5'. In particular, the same method can be used to show that for $j \in \{C, E_U, N\}$, the set of u such that $\pi_{Post}(u) = \pi_{Post,j}(u)$ must be an interval. As in the main section, we have $\pi_{Post}(u) = u$ for all $u \in [0, a]$ (part (i)). The assumption that $\pi(u_N) < \pi_N$ implies that $\pi_{Post}(u_N) = \pi_N$. Assumption N_u and N_π (the part that $u_h((1-\delta)B + \delta a) < u_N$ and $\pi_N < a + \frac{pB+(1-p)b-a}{pb+(1-p)B-a}(u_N - a)$) together imply that $\pi_{Post,E_C}((1-\delta)B + \delta a) = \pi_{Post}((1-\delta)B + \delta a)$, i.e., the payoff frontier is supported by cooperative empowerment at $(1-\delta)B+\delta a$. As a result, the payoff frontier for $u \in [a, u_N]$ is divided by three regions: randomization between centralization and cooperative empowerment (part (ii)), cooperative empowerment (part (iii)) and randomization between cooperative empowerment and the new opportunity (part (iv)).

Similarly, Assumption N_u and N_π (the part that $u_N < u_\ell((1-\delta)b + \delta B)$ and $\pi_N < B+b-u_N$) together imply that $\pi_{Post,E_C}((1-\delta)b + \delta B) = \pi_{Post}((1-\delta)b + \delta B)$, the payoff frontier is supported by cooperative empowerment at $(1-\delta)b+\delta B$. As a result, the payoff frontier for $u \in [u_N, B]$ is divided into three regions: randomization between the new opportunity and cooperative empowerment (part (v)), cooperative empowerment (part (vi)) and randomization between cooperative and uncooperative empowerment. (part (vii)).

Next, we show that for each u_N there exists $\widehat{\pi}_{Post}(u_N)$ such that $\pi_N = \max_u\{\pi_{Post}(u)\}$ for all $\pi_N \geq \widehat{\pi}_{Post}(u_N)$. To see this, suppose $\max_u\{\pi_{Post}(u)\} = \pi_N$ so that $\pi_{Post}(u)$ is maximized at u_N. Now suppose π_N increases to $\pi_N + \varepsilon$ for some $\varepsilon > 0$. It immediately follows that $\pi_{Post,N}(u_N)$ increases by ε. In addition, for each $j \in \{C, E_C, E_U\}$, the same argument as in Step 2 of Lemma 5C shows that $\pi_{Post,j}(u)$ increases by less than $\delta\varepsilon$. This implies that $\pi_{Post}(u)$ remain maximized at u_N. In other words, if $\pi_{Post}(u)$ is maximized at u_N for a given π_N, it remains so for all $\pi'_N \geq \pi_N$. This proves Lemma 7. ■

For our future analysis, it is useful to note the following.

COROLLARY 1. *$\bar{u}_{Post,CE} = (1-\delta)\,b + \delta B$ and $\pi'_{Post,+}(\bar{u}_{Post,CE}) < \pi'_{Post,-}(\bar{u}_{2,CE})$.*

Proof of Corollary 1: This is proved in the same way as that in Lemma 5A and is omitted here.

Properties of π_{Pre}

Now we characterize the payoff frontier of the pre-opportunity game. Unlike that analysis in the main section or in the post-opportunity game, there are no explicit expressions for the agent's continuation payoffs. Instead, they are pinned down by the following two conditions. First, their expected value is determined by the promise-keeping condition (with the same expressions as those in the main section). Second, we have $\pi'_{Pre}(u_{Pre,j}(u)) = \pi'_{Post}(u_{Post,j}(u))$ for $j = \{C, X, h, \ell\}$ when the payoff frontiers are differentiable. The next lemma provides the details.

LEMMA 8. *For any payoff $(u, \pi_{Pre}(u))$ on the frontier, the equilibrium continuation payoffs remain on the frontier. In addition, the following holds.*

(i.) If $(u, \pi_{Pre}(u))$ is supported with centralization, the agent's continuation payoff satisfies

$$\delta q u_{Post,C}(u) + \delta(1-q)\,u_{Pre,C}(u) = u - (1-\delta)\,a.$$

In addition,

$$\pi'_{Pre,+}(u_{Pre,C}(u)) \leq \pi'_{Post,-}(u_{Post,C}(u));\ \pi'_{Post,+}(u_{Post,C}(u)) \leq \pi'_{Pre,-}(u_{Pre,C}(u)).$$

(ii.) If $(u, \pi_{Pre}(u))$ is supported with cooperative empowerment, the agent's continuation payoff can be chosen to satisfy

$$\begin{aligned}\delta q u_{Post,l}(u) + \delta(1-q)\,u_{Pre,l}(u) &= u - (1-\delta)\,B;\\ \delta q u_{Post,h}(u) + \delta(1-q)\,u_{Pre,h}(u) &= u - (1-\delta)\,b.\end{aligned}$$

In addition, for $j \in \{h, \ell\}$,

$$\pi'_{Pre,+}(u_{Pre,j}(u)) \leq \pi'_{Post,-}(u_{Post,j}(u));\quad \pi'_{Post,+}(u_{Post,j}(u)) \leq \pi'_{Pre,-}(u_{Pre,j}(u)).$$

(iii.) If $(u, \pi_{Pre}(u))$ is supported with uncooperative empowerment, the agent's continuation payoff is given by

$$\delta q u_{Post,E_U}(u) + \delta(1-q)\,u_{Pre,E_U}(u) = u - (1-\delta)\,B.$$

In addition,

$$\pi'_{Pre,+}(u_{Pre,E_U}(u)) \leq \pi'_{Post,-}(u_{Post,E_U}(u));\ \pi'_{Post+}(u_{Post,E_U}(u)) \leq \pi'_{Pre,-}(u_{Pre,E_U}(u)).$$

(*iv.*) *If* $(u, \pi_{Pre}(u))$ *is supported with exit, the agent's continuation payoff satisfies*

$$\delta q u_{Post,X}(u) + \delta(1-q)\, u_{Pre,X}(u) = u.$$

In addition,

$$\pi'_{Pre,+}(u_{Pre,X}(u)) \leq \pi'_{Post,-}(u_{Post,X}(u));\ \pi'_{Post,+}(u_{Post,X}(u)) \leq \pi'_{Pre,-}(u_{Pre,X}(u));$$

Proof of Lemma 8: This is proved in the same way as that in Lemma 5A and is omitted here.

Now we can prove proposition 3.

PROPOSITION 3. *For each* $(u_N, \pi_N) \in N$,
(*i.*) *There exists* $\overline{\pi}(u_N)$ *and* $\overline{q}(u_N)$ *such that for all* $\pi_N \leq \overline{\pi}(u_N)$ *and* $q \leq \overline{q}(u_N)$, *there exists a history of path* h^T *such that* $\Pr\left(u_T = u_N | h^T\right) < 1$, *where* T *is the first period in the post-opportunity phase and* h^T *is the history of public outcomes up to period* T.
(*ii.*) *There exists a* $\widehat{\pi}(u_N)$ *and* $\widehat{q}(u_N)$ *such that for all* $\pi_N \leq \overline{\pi}(u_N)$ *and* $q \leq \overline{q}(u_N)$, *there exists a history of path* h^T *such that* $\Pr\left(u_t = u_N | h^T\right) = 0$ *for all* $t \geq T$.

Proof of Proposition 3: Part (*i*). We prove part (*i.*) by taking the following steps.

Step 1: To establish the upper bound $\overline{\pi}(u_N)$, define

$$\begin{aligned} s_L &= \frac{\pi\left(\underline{u}_{CE}\right) - a}{\underline{u}_{CE} - a}, \\ s_R &= \frac{b - \pi\left(\overline{u}_{CE}\right)}{B - \overline{u}_{CE}}; \end{aligned}$$

where recall π is the payoff frontier in the main section and $\underline{u}_{CE}$ ($\overline{u}_{CE}$) is the payoff level such that π is supported by randomization between centralization (uncooperative empowerment) and cooperative empowerment for $u \in [a, \underline{u}_{CE}]$ ($u \in [\overline{u}_{CE}, B]$). Recall that $\underline{u}_{CE} > a$ and $\overline{u}_{CE} < B$. Also note that $\underline{u}_{CE}$, $\overline{u}_{CE}$, and π do not depend on q and (u_N, π_N). Now define

$$\overline{\pi}(u_N) = \min\left\{a + s_L\left(u_N - a\right), b + s_R\left(u_N - B\right)\right\}.$$

For $\pi_N < \overline{\pi}(u_N)$, one then have $\pi'_{Post,+}(u_N) > s_R$ and $\pi'_{Post,-}(u_N) < s_L$.

Step 2: To establish the cutoff threshold $\overline{q}$, we introduce the following notation. Define u_C^* as the unique value satisfying

$$u_C^* = (1-\delta)\, a + \delta \left(q u_N + (1-q)\, u_C^* \right),$$

u_{UE}^* as the unique value satisfying

$$u_{UE}^* = (1-\delta)\, B + \delta \left(q u_N + (1-q)\, u_{UE}^* \right),$$

and u_{CE}^* as the unique value satisfying

$$u_{CE}^* = (1-\delta)\, b + \delta \left(q u_N + (1-q)\, u_{CE}^* \right).$$

Notice that for q close to 0, u_C^* goes to a, u_{UE}^* goes to B, and u_{CE}^* goes to b. It follows that there exists $\overline{q}$ such that for all $q < \overline{q}$, $u_C^* < \underline{u}_{CE}$, $u_{UE}^* > \overline{u}_{CE}$, and $u_{CE}^* < a$. Moreover, $u_{CE}^* < a$ implies that for all $u > a$, $u_{Pre,h}(u) > u$ if $u_{Post,h}(u) = u_N$.

Step 3: Define the set of payoffs U_N as

$$U_N = \left\{ u : \pi'_{Pre}(u) \geq \pi'_{Post,+}(u_N) \right\}.$$

Notice that we have $\pi_{Pre}(u) \leq \pi_{Post}(u)$ all u because once the opportunity has arrived, one can always replicate the actions and continuation payoffs of the game when the opportunity has not arrived. Since $\pi_{Pre}(B) = \pi_{Post}(B) = b$, we have

$$\pi'_{Post}(B) \leq \pi'(B) = s_R < \pi'_{Post,+}(u_N).$$

This implies that there exists a u' such that for all $u > u'$, $u \notin U_N$.

Step 4: If $\pi_{Pre}(u)$ is supported by cooperative empowerment and $u \notin U_N$, we must have $u_{Post,h}(u) \neq u_N$. In addition, if $\pi(u)$ is supported by an action j other than E_C and $u \notin U_N$, we must have $u_{Post,j} \neq u_N$.

To see the statement for cooperative empowerment, let $(u_{Pre,h}, u_{Pre,\ell}, u_{Post,h}, u_{Post,\ell}, \pi_{Pre}(u_{Pre,h}),$ $\pi_{Pre}(u_{Pre,\ell}), \pi_{Post}(u_{Post,h}), \pi_{Post}(u_{Post,\ell}))$ be the associated continuation payoffs. Suppose to the contrary that $u_{Post,h}(u) = u_N$. Notice that by the promise-keeping condition $\text{PK}_{\text{Post,AD}}$, we must have $u \neq B$. Now consider an alternative strategy profile with continuation payoff given by $(\widehat{u}_{Pre,h}, \widehat{u}_{Pre,\ell}, \widehat{u}_{Post,h}, \widehat{u}_{Post,\ell}, \pi_{Pre}(\widehat{u}_{Pre,h}), \pi_{Pre}(\widehat{u}_{Pre,\ell}), \pi_{Post}(\widehat{u}_{Post,h}), \pi_{Post}(\widehat{u}_{Post,\ell}))$, where

$\widehat{u}_{Pre,h} = u_{Pre,h}, \widehat{u}_{Pre,\ell} = u_{Pre,\ell}, \widehat{u}_{Post,h} = u_N + \varepsilon, \widehat{u}_{Post,\ell} = u_{Post,\ell}$. Notice that for ε small enough, $\widehat{u}_{Post,h} < \bar{u}_N$. It follows from the promise-keeping constraints $\mathrm{PK}_{\mathrm{Pre,AD}}$ and $\mathrm{PK}_{\mathrm{Pre,PD}}$ that, under the new set of continuation payoffs, the payoffs are given by

$$\begin{aligned}\widehat{u} &= u + \delta q \varepsilon; \\ \widehat{\pi} &= \pi_{Pre}(u) + \pi'_{Post,+}(u_N)\,\delta q \varepsilon.\end{aligned}$$

It can be checked that with this set of continuation payoffs $(\widehat{u}, \widehat{\pi})$ is a PPE payoff. However,

$$\begin{aligned}\pi_{Pre}(u + \delta q \varepsilon) &\leq \pi_{Pre}(u) + \pi'_{Pre,+}(u)\,\delta q \varepsilon \\ &< \pi_{Pre}(u) + \pi'_{Post,+}(u_N)\,\delta q \varepsilon,\end{aligned}$$

where the first inequality follows from the concavity of π_{Pre} and the second inequality follows because $u \notin U_N$. This is a contradiction because it violates the definition of π_{Pre}.

The statement on action j other than E_C can be proved in a similar way and is omitted here.

Step 5: Now suppose part (i) is not true in order to derive a contradiction. Suppose to the contrary that the continuation payoff following the opportunity is always (u_N, π_N). This implies that $u_t \in U_N$ for all t by Step 4. We now show that there exists a history in which the continuation payoff always increases, so and by Step 3 u_t eventually leaves U_N. Given the assumption on q, we see that if $(u_t, \pi_{Pre}(u_t))$ is supported by exit or centralization (with $u_{Post,j} = u_N$), we must have $u_{t+1} > u_t$ if the new opportunity does not arrive. In addition, if $(u_t, \pi_{Pre}(u_t))$ is supported by cooperative empowerment, the continuation payoff $u_{Pre,h}(u_t) > u_t$. In all these cases, the continuation payoffs continue to increase, and the concavity of π_{Pre} implies that the continuation payoff will eventually leave U_N. Now suppose $(u_t, \pi_{Pre}(u_t))$ is supported by E_U. If again $u_{Pre,E_U}(u_t) > u_t$, then we are done. Suppose instead $u_{Pre,E_U}(u_t) \leq u_t$. Then using the same method as in Step 4, we can show that $(u^*_{UE}, \pi_{Pre}(u^*_{UE}))$ is supported by uncooperative empowerment. Since the continuation of $(u^*_{UE}, \pi_{Pre}(u^*_{UE}))$ is either $(u^*_{UE}, \pi_{Pre}(u^*_{UE}))$ or (u_N, π_N), this implies that

$$\begin{aligned}\pi_{Pre}(u^*_{UE}) &= b + \frac{\pi_N - b}{u_N - B}(u^*_{UE} - B) \\ &< b + s_R(u^*_{UE} - B) \\ &= \pi(u^*_{UE}),\end{aligned}$$

where the inequality follows from the restriction on π_N and the last equality follows because $u^*_{UE} > \overline{u}_{CE}$ (Step 2). But this is a contradiction because the payoff frontier from the pre-opportunity game is weakly above the payoff frontier in the main section, i.e., $\pi_{Pre}(u^*_{UE}) \geq \pi(u^*_{UE})$. This proves part (*i.*).

Part (*ii.*): We now prove part (*ii.*) by taking the steps below.

Step 1: Define the maximum equilibrium supported by uncooperative empowerment as

$$\pi_{Pre,E_U}(u) = (1-\delta)\, b + \delta\left((1-q)\,\pi_{Pre}\left(u_{Pre,E_U}(u)\right) + q\pi_{Post}\left(u_{Post,E_U}(u)\right)\right).$$

We show that there exists some cutoff payoff (to be specified below) such that to its right we have $\pi_{Pre,E_U}(u) = \pi_{Pre}(u) = \pi_{Post}(u)$.

Recall from Lemma 7 that π_{Post} is a straight line to the right of $\bar{u}_{Post,CE}$ and denote its slope as s_2. Notice that for all $u \in [0, B]$, we have $\pi_{Pre}(u) \leq \pi_{Post}(u)$ because once the opportunity has arrived, one can always replicate the actions and continuation payoffs of the game when the opportunity has not arrived using the public randomization device. This implies that $\pi'_{Pre,-}(u) \geq s_2$ for all u. Now define $u^*_{Pre,UE}$ as the unique value for the agent such that

$$u^*_{Pre,UE} = (1-\delta)\, B + \delta\left(q\bar{u}_{Post,CE} + (1-q)\, u^*_{Pre,UE}\right).$$

Then for each $u \in [u^*_{Pre,UE}, B]$, let the action taken be uncooperative empowerment, and let the agent's continuation payoff be $u_{Pre,E_U}(u) = u$ when the new opportunity has not arrived and

$$u_{Post,E_U}(u) = \left(\frac{u-(1-\delta)\, B}{\delta} - (1-q)\, u\right)/q$$

when the new opportunity has arrived. Let the principal's payoff be equal to $\pi_{Pre}\left(u_{Pre,E_U}(u)\right)$ and $\pi_{Post}\left(u_{Post,E_U}(u)\right)$ respectively for these two events. With this choice of action and continuation payoffs, it is readily checked that

$$\pi_{Pre}(u) = \pi_{Pre,E_U}(u) = \pi_{Post}(u).$$

Step 2: Define

$$\bar{u}_{Pre,CE} = \max\{u : \pi_{Pre,E_C}(u) = \pi_{Pre}(u)\}.$$

We now show that $\bar{u}_{Pre,CE} = \bar{u}_{Post,CE} = (1-\delta)\, b + \delta B$ when π_N is sufficiently close to $\pi(u_N)$. Notice that we must have $\bar{u}_{Post,CE} \leq (1-\delta)\, b + \delta B$ because either $u_{Pre,h}(\bar{u}_{Post,CE}) > B$ or $u_{Post,h}(\bar{u}_{Post,CE}) > B$ otherwise. To rule out $\bar{u}_{Post,CE} < (1-\delta)\, b + \delta B$, suppose to the contrary that this is not the case. Recall that the payoff frontier in the main section π is a straight line to the right of $(1-\delta)\,b+\delta B$, and denote s as its slope. Consider a linear function $L(u)\equiv(u-B)\,s+b$, and let $\widehat{u}=u_{\ell}\left((1-\delta)\,b+\delta B\right)$. Define

$$d=L\left(\widehat{u}\right)-\pi\left(\widehat{u}\right).$$

The concavity of π implies that for all $u\le\widehat{u}$, $L(u)-\pi\left(u\right)>d$.

By Step 1, let s_2 be the slope between $(u^*_{Pre,UE},\pi_{Pre}\left(u^*_{Pre,UE}\right))$ and (B,b), and let s_1 be the slope between $(\overline{u}_{Pre,CE},\pi_{Pre}\left(\overline{u}_{1,CE}\right))$ and $(u^*_{Pre,UE},\pi_{Pre}\left(u^*_{1,UE}\right))$. For each u, $\pi_{Post}(u|\pi_N)$ is continuous in π_N (and that $\pi_{Post}(u|\pi_N)=\pi_{Pre}(u|\pi_N)=\pi(u)$ when $\pi_N=\pi(u_N)$). For each $\varepsilon_1>0$, there exists $\widehat{\pi}(u_N)$ sufficiently close to $\pi(u_N)$ such that for all $\pi_N\in(\pi(u_N),\widehat{\pi}(u_N))$, we have $s_2>s-\varepsilon_1$, $s_1\in[s-\varepsilon_1,s+\varepsilon_1]$, $\pi_{Pre}(u)<\pi(u)+\varepsilon_1$, and $\pi_{Post}(u)<\pi(u)+\varepsilon_1$.

Now by the definition of $(\overline{u}_{Pre,CE},\pi_{Pre}(\overline{u}_{1,CE}))$ can be supported by cooperative empowerment. Let $(u_{Pre,\ell},u_{Post,\ell},u_{Pre,h},u_{Post,h})$ be the associated continuation payoffs for the agent and $(\pi_{Pre}\left(u_{Pre,\ell}\right),$
$\pi_{Post}\left(u_{Post,\ell}\right),\pi_{Pre}\left(u_{Pre,h}\right),\pi_{Post}\left(u_{Post,h}\right))$ be the associated continuation payoffs for the principal. Notice that without loss of generality we may assume that $u_{Pre,h}=u_{Post,h}$. As a result, we have that $u_{Pre,h}=u_{Post,h}<B$.

Now consider an alternative strategy profile such that the principal empowers the agent. In addition, the agent's continuation payoffs are given by $(\widehat{u}_{Pre,\ell},\widehat{u}_{Post,\ell},\widehat{u}_{Pre,h},\widehat{u}_{Post,h})$ and the principal's payoffs are given by $(\pi_{Pre}\left(\widehat{u}_{Pre,\ell}\right),\pi_{Post}\left(\widehat{u}_{Post,\ell}\right),\pi_{Pre}\left(\widehat{u}_{Pre,h}\right),\pi_{Post}\left(\widehat{u}_{Post,h}\right))$, where $\widehat{u}_{Pre,\ell}=u_{Pre,\ell}+\varepsilon$, $\widehat{u}_{Post,\ell}=u_{Post,\ell}+\varepsilon$, $\widehat{u}_{Pre,h}=u_{Pre,h}+\varepsilon$, and $\widehat{u}_{Post,h}=u_{Post,h}+\varepsilon$. Sending ε to 0, we then have

$$\begin{aligned}\pi'_{Pre,+}\left(\overline{u}_{Pre,CE}\right) \quad\geq\quad & (1-p)\left((1-q)\,\pi'_{Pre,+}\left(u_{Pre,\ell}\right)+q\pi'_{Post,+}\left(u_{Post,\ell}\right)\right)\\ & +p\left((1-q)\,\pi'_{Pre,+}\left(u_{Pre,h}\right)+q\pi'_{Post,+}\left(u_{Post,h}\right)\right).\end{aligned}$$

Since $\pi'_{Pre,+}\left(\overline{u}_{1,CE}\right)\ge s_1$ (by the concavity of π_{Pre}) , and $\pi'_{Pre,+}\left(u_{1h}\right)$ and $\pi'_{Post,+}\left(u_{Post,h}\right)$ both weakly exceeds s_2, we have

$$(1-q)\,\pi'_{Pre,+}\left(u_{Pre,\ell}\right)+q\pi'_{Post,+}\left(u_{Post,\ell}\right)\le\frac{s_1-ps_2}{1-p}\le s+\varepsilon_1.$$

Next, notice that $\pi'_{Pre,+}\left(u_{Pre,\ell}\right)\ge s_2>s-\varepsilon_1$, this implies that

$$\pi'_{Post,+}\left(u_{Post,\ell}\right)\le\frac{s+\varepsilon_1-(1-q)\,\pi'_{Pre,+}\left(u_{Pre,\ell}\right)}{q}\le s+\frac{(2-q)\varepsilon_1}{q}.$$

Similarly, since $\pi'_{Post,+}\left(u_{Pre,\ell}\right)\ge s_2>s-\varepsilon_1$,

$$\pi'_{Pre,+}\left(u_{Pre,\ell}\right)\le\frac{s+\varepsilon_1-q\pi'_{Post,+}\left(u_{Pre,\ell}\right)}{1-q}\le s+\frac{(1+q)\varepsilon_1}{1-q}.$$

Now by Lemma 8 (ii),

$$(1-q)\,u_{Pre,\ell}+qu_{Post,\ell}=u_{\ell}(\overline{u}_{Pre,CE})<\widehat{u},$$

where recall that $\widehat{u}=u_{\ell}((1-\delta)\,b+\delta B)$. This implies either $u_{Pre,\ell}<\widehat{u}$ or $u_{Post,\ell}<\widehat{u}$. Suppose $u_{Pre,\ell}<\widehat{u}$. Since $\pi'_{Pre,+}\,(u_{Pre,\ell})\leq s+(1+q)\varepsilon_1/\,(1-q)$ by the above, it follows that

$$\begin{aligned}\pi_{Pre}(u_{Pre,\ell}) &= \pi_{Pre}(B)-\int_{u_{Pre,\ell}}^{B}\pi'_{Pre,+}\,(u)\,du\\ &\geq b+\left(s+\frac{(1+q)\varepsilon_1}{1-q}\right)(u_{Pre,\ell}-B)\,,\end{aligned}$$

where the inequality follows because $\pi_{Pre,}$ is concave. This implies that

$$L(u_{Pre,\ell})-\pi_{Pre,}(u_{Pre,\ell})\leq\frac{(1+q)\varepsilon_1}{1-q}\,(B-u_{Pre,\ell})\,,$$

where recall that $L(u)\equiv(u-B)\,s+b$. For small enough ε_1, this contradicts the claim that $L(u)-\pi_{Pre}(u)>L(u)-\pi(u)-\varepsilon_1>d-\varepsilon_1$ for all $u<\widehat{u}$.

Similarly, if $u_{Post,\ell}<\widehat{u}$, the same argument as above shows that

$$L(u_{Post,\ell})-\pi_{Post}(u_{Pre,\ell})\leq\frac{(2-q)\varepsilon_1}{q}\,(B-u_{Post,\ell})\,,$$

contradicts the claim that $L(u)-\pi_{Post}(u)>L(u)-\pi(u)-\varepsilon_1>d-\varepsilon_1$ for all $u<\widehat{u}$. This proves that when π_N is sufficiently close to $\pi\,(u_N)$, we must have $\overline{u}_{Pre,CE}=(1-\delta)\,b+\delta B$.

Step 3: Denote $\{u_t\}$ as the sequence of agent's payoffs along the equilibrium path. We now show that there is some probability that $u_t=B$ for some t before the new opportunity arrives (so the principal and the agent will not choose the new opportunity). Step 2 show that when $\pi_N<\widehat{\pi}\,(u_N)$, we have $\overline{u}_{Pre,CE}=(1-\delta)\,b+\delta B$, or equivalently $u_h\,(\overline{u}_{Pre,CE})=B$. To prove part (ii), it then suffices to show that a) the relationship starts with cooperative empowerment, and b) the agent's continuation payoff satisfies $u_{Pre,h}\,(u)>u$ for all $u\geq a$ (since π_{Pre} is maximized to the right of a). Notice that as q goes to 0 a) is satisfied because the relationship also starts with cooperative empowerment in the main section. In addition, b) is satisfied by the same argument as in Step 2 of the proof of part (i). As a result, there exists $\widehat{q}$ such that for all $q<\widehat{q}$ and $\pi_N<\widehat{\pi}\,(u_N)$, the relationship results in $u_T=B$ with positive probability for some T. When this occurs, $\Pr\left(u_t=B|h^T\right)=0$ *for all* $t\geq T$, and therefore, $\Pr\left(u_t=u_N|h^T\right)=0$ *for all* $t\geq T$. ■

Appendix C: Monetary Transfers

PROPOSITION 4. *When nonnegative transfers are allowed, the optimal relational contract can be implemented as follows.*

First period: The agent's and the principal's payoffs are given by $u^* \in [\underline{u}_{CE}, \bar{u}_{CE}]$ *and* $\pi_T(u^*) = \pi_{E_c}(u^*)$. *The parties engage in cooperative empowerment. If the agent chooses the principal's preferred project, his continuation payoff is given by*

$$u_h(u^*) = (u^* - (1-\delta) b)/\delta > u^*.$$

If, instead, the agent chooses his own preferred project, his continuation payoff is given by

$$u_\ell(u^*) = (u^* - (1-\delta) B)/\delta < u^*.$$

Subsequent periods: The agent's and the principal's payoffs are given by $u \in \{0, a\} \cup [\underline{u}_{CE}, B]$ *and* $\pi(u)$. *Their actions and continuation payoffs depend on what region* u *is in:*

(*i.*) *If* $u = 0$, *the parties exit. The agent's continuation payoff is given by* $u_X(0) = 0$.

(*ii.*) *If* $u = a$, *the parties engage in centralization. The agent's continuation payoff is given by* $u_C(a) = a$.

(*iii.*) *If* $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$, *the parties engage in cooperative empowerment. If the agent chooses the principal's preferred project, his continuation payoff is given by* $u_h(u) > u$ *and* $w_h(u) = 0$. *If, instead, the agent chooses his own preferred project, his continuation payoff is given by* $u_\ell(u) < u$.

(*iv.*) *If* $u \in (\bar{u}_{CE}, B)$, *the parties engage in cooperative empowerment. If the agent chooses the principal's preferred project, his continuation payoff is given by* $u_h(u) = B$ *and* $w_h(u) = (u - (1-\delta) b - \delta B)/(1-\delta)$. *If, instead, the agent chooses his own preferred project, his continuation payoff is given by* $u_\ell(u) < u$.

(*v.*) *If* $u = B$, *the parties engage in uncooperative empowerment. The agent's continuation payoff is given by* $u_{E_U}(B) = B$.

Proof of Proposition 4: For each payoff pair $(u, \pi_T(u))$ on the frontier, if it is supported with pure action $j \neq E_C$, let $w_j(u)$ be the transfer from the principal to the agent. If it is supported with E_C, let $w_h(u)$ and $w_l(u)$ be the associated transfers. A similar argument as in the proof of main text can be used to show that $\pi_T(u)$ is concave and the PPE set is compact. In addition, for any equilibrium payoff on the frontier, its continuation payoffs along the equilibrium path remain on the frontier. There is also a corresponding set of constraints as the main text. In particular, if $(u, \pi_T(u))$ is supported by cooperative empowerment, the agent's incentive compatibility constraint can be made binding. Together with the agent's promise-keeping condition, we have

$$u = (1-\delta)(b + w_h(u)) + \delta\pi_h(u) = (1-\delta)(B + w_\ell(u)) + \delta\pi_\ell(u).$$

To prove Proposition 4, we take the following steps.

Step 1: $\pi'_{T+}(u) > -1$ for all $u < B$.

Notice that for $u \geq B$, we have $\pi_T(u) = B + b - u$. This can be implemented by uncooperative empowerment and the principal pays $w_{E_U}(u) = (u - B)/(1-\delta)$. As a result, we have $\pi'_{T+}(u) \geq -1$ for all $u < B$. The concavity of π_T then implies that if $\pi'_{T+}(u) = -1$ for some u, the payoff frontier is a straight line with slope -1 in $[u, B]$. Let the left end point of the line segment be u'. Since $(u', \pi_T(u'))$ is an extremal point, it is supported by a pure action. Given $\pi_T(B) = b$, we have that $(u', \pi_T(u'))$ lies on the line segment between (B, b) and (b, B). It follows that the action associated with $(u', \pi_T(u'))$ can only be E_U or $E_{C.}$ Moreover, notice that the continuation payoffs of the players must again lie on the line segment between (B, b) and (b, B). As a result, if the agent's continuation payoff ever falls to the left of u', we get a contradiction.

Now suppose the action is E_U, the promise-keeping condition gives that

$$u' = (1-\delta)\left(B + w_{E_U}\left(u'\right)\right) + \delta u_{E_U}\left(u'\right).$$

Given that $w_{E_U}(u') \geq 0$, this implies that $u_{E_U}(u') \leq (u' - (1-\delta)B)/\delta < u'$, and, thus, violates the definition of u' be the left most point of the line segment. Next, suppose the action is E_C, the promise-keeping and the incentive-compatibility condition gives that

$$u' = (1-\delta)\left(B + w_{E_C}\left(u'\right)\right) + \delta u_\ell\left(u'\right).$$

Given that $w_{E_C}(u') \geq 0$, this implies that $u_\ell(u') \leq (u' - (1-\delta)B)/\delta < u'$, and we again derive a contradiction. Notice that concavity of π_T also implies that $\pi'_{T-}(u) > -1$ for all $u < B$.

Step 2: For all $u < B$ the following holds. If the action $j \neq E_C$, $w_j(u) = 0$. If $j = E_C$, we have $w_\ell(u) = 0$ and $w_h(u) > 0$ if and only if $u > \overline{u}_{CE} \equiv (1-\delta)b + \delta B$.

To see this, suppose to the contrary that $w_j(u) > 0$ for some $u < B$ and $j \neq E_C$. We can then show that $(u - (1-\delta)\varepsilon, \pi_T(u) + (1-\delta)\varepsilon)$ is also on the frontier for small enough ε. The reason is that we can repeat the same actions and continuation payoffs associated with $(u, \pi_T(u))$ but ask the principal to pay the agent $w_j(u) - \varepsilon$ instead. This change does not affect any of the constraints and thus results in an equilibrium whose payoff is $(u - (1-\delta)\varepsilon, \pi_T(u) + (1-\delta)\varepsilon)$. This implies that $\pi'_{T-}(u) \leq -1$, which contradicts that $\pi'_{T-}(u) > -1$ for $u < B$.

Next, suppose $(u, \pi_T(u))$ is supported with cooperative empowerment for $u < B$. Let the agent's continuation payoffs and transfers be $(u_\ell(u), u_h(u), w_\ell(u), w_h(u))$ and the principal's continuation payoffs be $(\pi_T(u_l(u)), \pi_T(u_h(u)))$. Now suppose to the contrary that $w_\ell(u) > 0$. Notice that $u_\ell(u) = (u - (1-\delta)(B + w_\ell(u)))/\delta < B$ and therefore $\pi'_{T+}(u_\ell(u)) > -1$ by Step 1. Consider an alternative strategy profile that again uses cooperative empowerment and whose continuation payoffs and transfers for the agent are given by $(u_\ell(u) + (1-\delta)\varepsilon, u_h(u), w_\ell(u) - \delta\varepsilon, w_h(u))$ and whose continuation payoffs for the principal are given by $(\pi_T(u_\ell(u) + (1-\delta)\varepsilon), \pi_T(u_h(u)))$. This alternative strategy profile satisfies all of the constraints and is again a PPE. The promise-keeping constraint implies that the agent's payoff is u. The principals' payoff is given by

$$\pi_T(u) + \delta((1-\delta)\varepsilon + \pi_T(u_\ell(u) + (1-\delta)\varepsilon) - \pi_T(u_\ell(u))) > \pi_T(u),$$

where the inequality follows because $\pi'_{T+}(u_l(u)) > -1$. This contradicts the definition of π_T.

Finally, if $w_h(u) > 0$ for some u. Let the agent's continuation payoffs be $(u_\ell(u), u_h(u))$ and the principal's continuation payoffs be $(\pi_T(u_\ell(u)), \pi_T(u_h(u)))$. Consider an alternative strategy profile that again uses cooperative empowerment and whose continuation payoffs and the transfers for the agent is given by $(u_\ell(u), u_h(u) + (1-\delta)\varepsilon)$ and $w_h(u) - \delta\varepsilon$ and whose continuation payoffs for the principal are given by $(\pi_T(u_\ell(u)), \pi_T(u_h(u) + (1-\delta)\varepsilon))$. This alternative strategy profile satisfies all of the constraints and is again a PPE. The promise-keeping constraint implies that the agent's payoff is u. The principals' payoff is given by

$$\pi_T(u) + \delta((1-\delta)\varepsilon + \pi_T(u_h(u) + (1-\delta)\varepsilon) - \pi_T(u_h(u))).$$

This value exceeds $\pi_T(u)$ if and only if $\pi'_{T+}(u_h(u)) > -1$, or equivalently, if and only if $u_h(u) < B$. Note that the agent's promise-keeping condition (together with the IC condition) gives that

$$u = (1-\delta)(b + w_h(u)) + \delta\pi_h(u).$$

It follows that $w_h(u) > 0$ if and only if $u > \overline{u}_{CE} = (1-\delta)b + \delta B$.

Step 3: We now characterize the payoff frontier $\pi_T(u)$.

The payoff frontier of $\pi_T(u)$ can be characterized in essentially the same way as the payoff frontier $\pi(u)$ in the main text. Given Step 2, we can now define for each action $j \neq E_C$, $\pi_j(u)$, the maximal equilibrium payoff sustained by action j the same way as in the main text. For $j = E_C$, we have

$$\pi_{E_C}(u) \equiv p[(1-\delta)B + \delta\pi_T(u_h(u))] + (1-p)[(1-\delta)b + \delta\pi_T(u_\ell(u))],$$

for $u \leq \overline{u}_{CE} = (1-\delta) b + \delta B$, where u_h and u_ℓ are defined in the same way as in the main text. For $u > \overline{u}_{CE}$, we define

$$\pi_{E_C}(u) \equiv p(B+b-u) + (1-p)\left[(1-\delta) b + \delta \pi_T(u_\ell(u))\right].$$

Notice that when the agent chooses the principal's preferred project, his continuation payoff weakly exceeds B by Step 2. This implies that the joint payoff of the two parties in this case is given by $B+b$, and therefore, the principal's payoff is given by $B+b-u$ for this case. Notice that one way to implement $(u, \pi_{E_C}(u))$ for $u > \overline{u}_{CE}$ is to set $u_h(u) = B$ and $w_h(u) = (u-(1-\delta) b - \delta B)/(1-\delta)$.

The same argument as in the main text shows that for each action j, the set of agent's payoff in which the payoff frontier is supported by j $(\{u : \pi_T(u) = \pi_j(u)\}$ is an interval. And just as in the main text, the payoff frontier is given by $\pi_T(u) = u$ for $u \in [0, a]$. In addition, there exists $\underline{u}_{CE} > a$ such that for $u \in [a, \underline{u}_{CE}]$, the payoff frontier is sustained by randomizing between (a, a) and $(\underline{u}_{CE}, \pi_T(\underline{u}_{CE}))$, where $(\underline{u}_{CE}, \pi_T(\underline{u}_{CE}))$ is supported by cooperative empowerment.

Next, we show that for $u \in [\underline{u}_{CE}, B]$, we have $\pi_T(u) = \pi_{E_C}(u)$. To do this, it suffices to show that $\pi_T(u) > \pi_{E_U}(u)$ for all $u < B$. Now suppose to the contrary that $\pi_T(u') = \pi_{E_U}(u')$ for some $u < B$. The same argument as in Lemma 5A implies that $\pi_T(u)$ is a straight line in $[u', B]$. Notice, however, that $\pi_T(B) = \pi_{E_C}(B)$, and $\pi'_{E_C}(B) = -1$ by construction. This implies that $\pi'_T(B) = -1$, and therefore $\pi'_T(u) = -1$ for all $u \in (u', B]$. This contradicts Step 1. It follows that $\pi_T(u) > \pi_{E_U}(u)$ for all $u \in [\underline{u}_{CE}, B)$, and therefore, $\pi_T(u) = \pi_{E_C}(u)$. This finishes the characterization of the payoff frontier, and the dynamics of the relationship can be obtained just as in Proposition 1 and 2. ■

Appendix D: Intermediate Allocations of Control

LEMMA 9. *The PPE frontier $\pi(u)$ can be divided into five regions.*

$$\pi(u) = \begin{cases} u\pi(\underline{u}_{PE})/\underline{u}_{PE} & u \in [0, \underline{u}_{PE}); \\ \pi_{E_P}(u) & u \in [\underline{u}_{PE}, \bar{u}_{PE}]; \\ ((\underline{u}_{CE} - u)\,\pi(\bar{u}_{PE}) + (u - \bar{u}_{PE})\,\pi(\underline{u}_{CE}))\,/\,(\underline{u}_{CE} - \bar{u}_{PE}) & u \in (\bar{u}_{PE}, \underline{u}_{CE}); \\ \pi_{E_C}(u) & u \in [\underline{u}_{CE}, \bar{u}_{CE}]; \\ ((B-u)\,\pi(\bar{u}_{CE}) + (u - \bar{u}_{CE})\,b)\,/\,(B - \bar{u}_{CE}) & u \in (\bar{u}_{CE}, B], \end{cases}$$

where we can choose $\underline{u}_{PE} = (1-\delta) a$, $\bar{u}_{PE} \leq a$, $\underline{u}_{CE} < (1-\delta) B + \delta a$, and $\bar{u}_{CE} = (1-\delta) b + \delta B$.

Proof of Lemma 9: The proof of this lemma follows similar steps as in characterizing the PPE payoff frontier of the model in the main section.

Step 1 Recall from the main text that the action E_P (partial empowerment) is defined as follows. If the agent finds that principal's preferred action, then it is implemented, leading to a payoff of (b, B). Otherwise, the principal takes over, resulting to a payoff of (a, a). Let the agent's continuation payoff be u_{h,E_P} and u_{ℓ,E_P} respectively.

Now if E_P is used, we have the following set of constraints.

$$u = p\left((1-\delta) b + \delta u_{h,E_P}\right) + (1-p)\left((1-\delta) a + \delta u_{\ell,E_P}\right); \tag{PK-PE}$$

$$(1-\delta) b + \delta u_{h,D_{CD}} \geq (1-\delta) a + \delta u_{\ell,D_{CD}}. \tag{IC-PE}$$

As in the main text, the IC-PE can be made to hold with equality for payoff pairs on the PPE payoff frontier, and the two conditions above imply that

$$\begin{aligned} \delta u_{h,E_P}(u) &= u - (1-\delta) b; \\ \delta u_{\ell,E_P}(u) &= u - (1-\delta) a. \end{aligned}$$

Relatedly, define

$$\pi_{E_P}(u) = p\left((1-\delta) B + \delta \pi\left(u_{h,E_P}(u)\right)\right) + (1-p)\left((1-\delta) a + \delta \pi\left(u_{\ell,E_P}(u)\right)\right).$$

This is the principal's highest equilibrium payoff given that the agent's payoff is u and the current action is partial empowerment.

Step 2: We now show that $\pi_{E_P}(a) > a$. To see this, notice that $u_{\ell,E_P}(a) = a$. It follows that

$$\pi_{E_P}(a) = p\left((1-\delta) B + \delta \pi\left(u_{h,E_P}(a)\right)\right) + (1-p)\left((1-\delta) a + \delta \pi(a)\right).$$

Since $\pi(a) \geq \pi_C(a) = a$, it suffices to show that $(1-\delta) B + \delta \pi\left(u_{h,E_P}(a)\right) > a$. Now define $x = u_{h,E_P}(a) - a$. It then follows from the promise-keeping constraint that

$$x = \frac{(1-\delta)(a-b)}{\delta}.$$

In addition,

$$\begin{aligned} & (1-\delta) B + \delta \pi\left(u_{h,E_P}(a)\right) - a \\ = \; & (1-\delta) B + \delta \pi(a + x) - a \\ \geq \; & (1-\delta) B + \delta\left(a + \frac{b-a}{B-a}\frac{(1-\delta)(a-b)}{\delta}\right) - a \\ = \; & (1-\delta)(B-a) - (1-\delta)\frac{(a-b)^2}{B-a} \\ > \; & 0, \end{aligned}$$

where the weak inequality follows because π lies above the line segment between (a, a) and (B, b), and the last inequality follows because $B+b > 2a$. This establishes that $\pi_{E_P}(a) > a$, and the same argument in the main text implies that centralization (C) will not be used on the payoff frontier.

Step 3: We now prove Lemma 9. First, we show that there exists u' such that $\pi_{E_P}(u) > \pi_{E_C}(u)$ if and only if $u < u'$. To see this, notice that $u_{h,E_P}(u) = u_{h,E_C}(u)$, so

$$\begin{aligned} & \pi_{E_P}(u) - \pi_{E_C}(u) \\ = \quad & (1-p)\left[(1-\delta)(a-b) + \delta\left(\pi\left(u_{\ell,E_P}(u)\right) - \pi\left(u_{\ell,E_C}(u)\right)\right)\right]. \end{aligned}$$

Since

$$u_{\ell,E_P}(u) - u_{\ell,E_C}(u) = \frac{(1-\delta)(B-a)}{\delta}$$

is independent of u, this implies that $\pi\left(u_{\ell,E_P}(u)\right) - \pi\left(u_{\ell,E_C}(u)\right)$ is decreasing in u given that π is concave. This proves if $\pi_{E_P}(u') - \pi_{E_C}(u') < 0$ for some u', the inequality holds for all $u > u'$. To see such u' exists, notice that a sufficient condition is that $\pi'\left(u_{\ell,E_C}(u)\right) < -(a-b)/(B-a)$ for some u. Now the largest feasible u is $\bar{u}_{E_C} = (1-\delta) b + \delta B$. Then the condition is equivalent to that $(1-\delta) b + \delta B > a$, which is implied by Assumption 1 (ii).

Next, recall from Step 2 that C cannot be used to support the payoff frontier. It follows that if a payoff pair on the supported by a pure action j, then $j \in \{E_P, E_C, E_U, X\}$. Now the same argument as in the main text shows that for $u \geq \bar{u}_{CE} = (1-\delta) b + \delta B$, the payoff frontier is a straight line sustained by the randomization between (B, b) and $(\bar{u}_{CE}, \pi(\bar{u}_{CE}))$. A similar argument shows that for $u \leq \underline{u}_{PE} = (1-\delta) a$, where $u_{\ell,E_P}(a) = 0$, the payoff frontier is sustained by the randomization between $(0,0)$ and $(\underline{u}_{PE}, \pi(\underline{u}_{PE}))$. For $u \in [\underline{u}_{PE}, \bar{u}_{CE}]$, the payoff frontier is sustained either by E_P, E_C, or the randomization between the two. Since $\pi_{E_P}(u) < \pi_{E_C}(u)$ if and only if $u < u'$ for some u', the shape of the frontier follows.

Next, we show that $\bar{u}_{PE} \leq a$. Suppose this is not the case. It then follows that

$$\bar{u}_{PE} < u_{\ell,E_C}(\bar{u}_{PE}) < u_{h,E_C}(\bar{u}_{PE}),$$

and therefore,

$$\pi'_{E_P-}(\bar{u}_{PE}) \leq p\pi'_-\left(u_{h,E_P}(\bar{u}_{PE})\right) + (1-p)\,\pi'_-\left(u_{\ell,E_P}(\bar{u}_{PE})\right) = \pi'_{E_P-}(\bar{u}_{PE}).$$

As a result, the payoff frontier must be a line segment between $\bar{u}_{PE}$ and $u_{h,E_P}(\bar{u}_{PE})$. The same type of argument as in the main text implies that the payoff frontier is a line segment between a and $u_{h,E_P}(\bar{u}_{PE})$. This implies that we can have $\bar{u}_{PE} \leq a$. Moreover, when $\bar{u}_{PE} = a$, we have

$$\pi'_{E_P-}(a) \leq p\pi'_-\left(u_{h,E_P}(a)\right) + (1-p)\,\pi'_-(a) = \pi'_{E_P-}(a),$$

so $\pi'_{-}(a) = \pi'_{+}(a) = \pi'_{-}(u_{h,E_P}(a))$.

Finally, we show that $\underline{u}_{CE} < (1-\delta) B + \delta a$. Since $\bar{u}_{PE} \leq a$, this is implied by $u_{\ell,E_P}(\underline{u}_{CE}) \leq \bar{u}_{PE}$ unless when $\bar{u}_{PE} = a$, in which case we will show that $u_{\ell,E_P}(\underline{u}_{CE}) < \bar{u}_{PE}$. By the definition of $\underline{u}_{CE}$ and $\bar{u}_{PE}$, the payoff frontier is a straight line in $[\underline{u}_{CE}, \bar{u}_{PE}]$. Let the slope of this line segment be s and $u' \geq \bar{u}_{PE}$ be the right end of the line segment. It follows that we must have

$$s = \pi'_{-}(u') \leq p\pi'_{-}(u_{h,E_C}(u')) + (1-p)\,\pi'_{-}(u_{\ell,E_C}(u')).$$

Since $\pi'_{-}(u_{h,E_C}(u')) < s$ by the definition of u', it then must be the case that $\pi'_{-}(u_{\ell,E_C}(u')) > s$. By the concavity of π, this means that $u_{\ell,E_C}(u') \leq \underline{u}_{CE}$. When $\underline{u}_{CE} = a$, recall from above that $\pi'_{-}(a) = \pi'_{+}(a) = s$. In this case, $\pi'_{-}(u_{\ell,E_C}(u')) > s$ implies that we must have $u_{\ell,E_C}(u') < \underline{u}_{CE}$. ■

PROPOSITION 5. *When partial empowerment is allowed, the optimal relational contract satisfies the following:*

First period: *The agent's and the principal's payoffs are given by $u^* \in [\underline{u}_{PE}, \bar{u}_{PE}] \cup [\underline{u}_{CE}, \bar{u}_{CE}]$ and $\pi(u^*) = \max\{\pi_{E_P}(u), \pi_{E_C}(u^*)\}$. The parties engage in either partial empowerment or cooperative empowerment. In either case, if the agent chooses the principal's preferred project, his continuation payoff increases, and it falls otherwise.*

Subsequent periods: The agent's and the principal's payoffs are given by $u \in [\underline{u}_{PE}, \bar{u}_{PE}] \cup [\underline{u}_{CE}, \bar{u}_{CE}] \cup \{B\}$ and $\pi(u)$. Their actions and continuation payoffs depend on what region u is in:

(*i.*) *If $u = 0$, the parties exit. The agent's continuation payoff is given by $u_X(0) = 0$.*

(*ii.*) *If $u \in [\underline{u}_{PE}, \bar{u}_{PE}]$, the parties engage in partial empowerment. If the agent chooses the principal's preferred project, his continuation payoff is given by $u_{E_P,h}(u) > u$. If, instead, he chooses the safe project, his continuation payoff is given by $u_{E_P,\ell}(u) < u$.*

(*iii.*) *If $u \in [\underline{u}_{CE}, \bar{u}_{CE}]$, the parties engage in cooperative empowerment. If the agent chooses the principal's preferred project, his continuation payoff is given by $u_{E_C,h}(u) > u$. If, instead, the agent chooses his own preferred project, his continuation payoff is given by $u_{E_C,\ell}(u) < u$.*

(*iv.*) *If $u = B$, the parties engage in uncooperative empowerment. The agent's continuation payoff is given by $u_{E_U}(B) = B$.*

Proof of Proposition 5: Notice that Proposition 5 follows directly from Lemma 9. For the proof of Proposition 6, we state two additional properties. First, if $u^* > a$, then $\pi(u^*) = \pi_{E_C}(u^*)$. Suppose to the contrary that $\pi(u^*) = \pi_{E_P}(u^*)$ instead. Given $u^* > a$, we have $u_{E_P,\ell}(u^*) > a$. It follows that

$$\pi'_{-}(u^*) \leq \pi'_{E_P-}(u^*) = p\pi'_{-}(u_{h,E_P}(u^*)) + (1-p)\pi'_{-}(u_{\ell,E_P}(u^*)) < 0.$$

This contradicts the optimality of u^* (which implies that $\pi(u^*) \geq 0$.)

Second, if $u^* \leq a$, then $\pi(u^*) = \pi_{E_P}(u^*)$. Now suppose to the contrary that $\pi(u^*) = \pi_{E_C}(u^*)$. Then we have $\underline{u}_{CE} \leq u^* \leq a$. This implies

$$u_{E_P,\ell}(\underline{u}_{CE}) \leq \underline{u}_{CE}.$$

Moreover, since

$$u_{E_C,\ell}(\underline{u}_{CE}) \leq u_{E_P,\ell}(\underline{u}_{CE}) \leq \underline{u}_{CE} \leq u^*$$

and π is increasing for $u \leq u^*$, this implies that

$$\pi(u_{E_C,\ell}(\underline{u}_{CE})) \leq \pi(u_{E_P,\ell}(\underline{u}_{CE})).$$

This implies, however, that $\pi_{E_C}(\underline{u}_{CE}) < \pi_{E_P}(\underline{u}_{CE})$, contradicting that $\pi_{E_C}(\underline{u}_{CE}) = \pi(\underline{u}_{CE})$.

Notice that if $u^* = a$, then we have $u_{E_P,\ell}(a) = a$, and $u_{E_P,h}(a) = (a-(1-\delta)b)/\delta > a$. A similar type of argument as in the main text implies that one must have $\pi'_{-}(u_{E_P,h}(a)) = 0$ for $u^* = a$. This implies that if π is maximized at a, we have $\pi(u) = \pi(a)$ for all $u \in [a, (a-(1-\delta)b)/\delta]$. ■

PROPOSITION 6. *In the optimal relational contract, the principal chooses either cooperative empowerment or partial empowerment for the first τ periods, where τ is random and finite with probability one. For $t > \tau$, the relationship results in either termination or permanent empowerment, depending on the history of the relationship. Both possibilities occurs with positive probability for all $p \in (0,1)$.*

Proof of Proposition 6: Notice that when either E_C or E_P is used, the continuation payoff of the agent along the frontier is always higher if the agent chooses the principal's preferred project. It follows that there is always a positive probability that $u_t = B$. It remains to show that $u_t = 0$ with positive probability. Notice that if $u^* < a$, then from the proof in Proposition 5 that $\pi(u^*) = \pi_{E_P}(u^*)$, so the relationship starts with partial empowerment. When $u < a$, $u_{E_P,\ell}(u) < u$, so a sequence of low outcomes lead to termination. When $u^* \geq a$, recall from Lemma 9 that $\bar{u}_{PE} < a$. If $\bar{u}_{PE} < a$, a sequence of low outcomes leads to $u_t = \bar{u}_{PE}$ with positive probability for some t, which then leads to the termination of the relationship with positive probability. ■