Leaks, Sabotage, and Information Design*

Aaron Kolb[†] Erik Madsen[‡]

September 28, 2019

Abstract

We study how an organization dynamically screens an agent of uncertain loyalty whom it suspects of committing damaging acts of undermining, for instance leaking sensitive information or sabotaging production. The organization's screening tool is the agent's access to sensitive information, i.e. the stakes of the relationship, governing both productivity and the harm from undermining. A disloyal agent strategically chooses when and how intensively to undermine, with undermining stochastically detected at a rate proportional to its intensity. When the organization can commit, it optimally guides stakes by holding them constant for a time, then gradually escalating them, and finally jumping them to their maximal level. This stakes path is also the unique equilibrium outcome when the organization cannot commit, and the disloyal agent's unique equilibrium undermining path exhibits variable, non-monotonic intensity. In an information design microfoundation, the optimal stakes path is implementable via an inconclusive contradictory news process.

JEL Classification: C70, D82, D83, D86, M51

Keywords: Principal-agent model, undermining, gradualism, information design

1 Introduction

An organization has discovered it is being undermined by an insider with access to privileged information. Sensitive documents have been leaked to the media, corporate secrets have been

*The authors thank Anton Kolotilin, Elliot Lipnowski, Laurent Mathevet, Eran Shmaya, Alex Wolitzky, and seminar audiences at Brown, NYU, the University of Hong Kong, Indiana Kelley, Harvard, MIT, Johns Hopkins, the 2018 NSF/NBER/CEME Conference, the 2019 Cowles Economic Theory Conference, the 2019 SAET conference, the 2019 Midwest Economic Theory Conference, the 2019 International Conference on Game Theory, and the 2019 MIT Applied Theory Summer Camp for helpful conversations.

[†]Department of Business Economics and Public Policy, Kelley School of Business, Indiana University (Email: `kolba@indiana.edu`).

[‡]Department of Economics, New York University (Email: `emadsen@nyu.edu`).

sold to competitors, obscure vulnerable points in production lines have been discovered and sabotaged. Executives suspect that a specific key employee is to blame — but how can they be sure? Cutting off the employee's access to information would staunch the bleeding until his loyalty can be assessed, but also leave him paralyzed and unable to perform his job effectively. Worse, if the employee is disloyal and expects the sequester to be temporary, he may simply lie low and feign loyalty until he regains trust and comprehensive access. How should the organization regulate the employee's access to information to optimally assess his loyalty and limit damage to operations?

Abuse of organizational secrets is a common concern in a broad range of organizations. Leaks are one important instance — for instance, of governmental or military secrets to the enemy during times of war, of damaging or sensitive information about government policies to the press, and of technology or strategic plans to rival firms. Another is straightforward sabotage of sensitive assets — for instance, destruction of a key railway line transporting troops and supplies for an upcoming operation, or malicious re-writing of software code causing defects or breakdowns on a production line. (We discuss several prominent examples of such behavior in Section 1.1.) We take these occurrences as evidence that disloyal agents occasionally seek to undermine their organization through abuse of privileged information.

An important feature of the examples discussed above is that undermining is typically a dynamic process whose source can be identified only with substantial delay. A mole within an organization typically leaks many secrets over an extended period, during which time the organization may realize that information is leaking, but not know who is responsible. Similarly, a saboteur may disrupt a production line multiple times or bomb several installations before being caught. (In Section 1.1 we highlight the ongoing nature of undermining activities in several of the examples discussed there.) We therefore study environments in which agents have many chances to undermine their organization, with each decision to undermine being only occasionally traced back to the agent.

In our model, a principal hires an agent to perform a task repeatedly over time. Task efficiency increases with the agent's access to information about the nature of the task. However, the agent may secretly choose to undermine the principal with variable intensity at any point in time, and the damage inflicted by each such act is also increasing in the agent's information. Access to information thus governs the stakes of the principal-agent relationship. Each act of undermining is detected only probabilistically, with the rate of discovery proportional to the intensity of undermining, so that a disloyal agent has the opportunity to undermine the principal multiple times. The agent's loyalty is private information, and a disloyal agent wishes to inflict maximum damage to the principal over the lifetime of his employment. The principal's goal is to regulate the stakes of the relationship so as to efficiently screen out disloyal agents, discovering them as quickly as possible and limiting the information fed to them without excessively restricting the productivity of loyal agents. We microfound stakes as resulting from dynamic disclosure of information about the task, imposing a natural monotonicity constraint on stakes over time. We further assume that the agent begins the relationship with some information, imposing a lower bound on stakes.

We first assume that the principal has commitment power and characterize her optimal stakes policy, whose dynamics in general fall into three distinct regimes. First, there exists an initial *quiet period* during which stakes are frozen at their minimal level. After this period, stakes rise at a constant rate during a *gradual escalation* phase. At the end of this interval, the agent is deemed *trusted* and stakes jump discontinuously upward to their maximal level, after which they remain fixed forever. In response to this policy, the disloyal agent undermines at full intensity during the quiet period and trusted phase, while during the gradual escalation phase all undermining policies are incentive-compatible and optimal for the principal. In the information design microfoundation, the optimal stakes policy is implemented via an initial period of no disclosure, followed by periodic unreliable reports which sometimes deliberately mislead the agent, and ending with a discrete signal which perfectly reveals the state of the world.

A natural concern in the high-stakes applications motivating our model is that the principal may not be able to credibly commit to future levels of access. We show that in our setting, credibility is not an issue — the optimal commitment stakes policy is implementable in a perfect Bayesian equilibrium of the game between the principal and disloyal agent, and yields the principal the same payoffs as under commitment. In fact, it is the *only* stakes policy implementable in any Bayes Nash equilibrium. Further, among the class of pure-strategy equilibria there exists a unique equilibrium undermining policy for the disloyal agent. This policy exhibits non-monotonic undermining intensity, with undermining at full intensity during the quiet period followed by a discontinuous drop in intensity at the beginning of the gradual escalation phase, and a subsequent smooth rise back to full undermining intensity by the beginning of the trusted phase. This undermining policy has the key property that ex post at each point in time, the agent's reputation induced by his undermining policy justifies the continuation stakes the principal has promised under the optimal stakes policy.

The design of the optimal stakes policy is crucially shaped by the agent's ability to strategically time undermining. Ideally, the principal would like to subject the agent to a low-stakes trial period, during which disloyal agents are rooted out with high enough certainty that the principal can comfortably trust the survivors and jump stakes to their maximal level. However, a disloyal agent will look ahead to the end of the trial phase and decide to feign loyalty until they become trusted, so such a scheme will fail to effectively identify loyalists (unless the minimum stakes level is very high). Indeed, an ineffective trial period merely delays the time until the principal begins building trust, reducing lifetime payoffs from the project. The agent must therefore face high enough stakes prior to becoming trusted that they prefer to undermine during the trial period rather than wait to build trust. A gradual rise in stakes to a level strictly below maximal stakes emerges as the cost-minimizing way to ensure disloyal agents choose to undermine and reveal themselves during the trial period.

The remainder of the paper is organized as follows. Section 1.1 presents real world examples of leaks and sabotage, while Section 1.2 reviews related literature. Section 2 presents the model. We derive the optimal contract under commitment in Section 3 and perform an equilibrium analysis without commitment in Section 4. Section 5 provides an information-design microfoundation for stakes and characterizes the optimal information policy. Section 6 discusses comparative statics and Section 7 concludes. The appendices contain formal equilibrium definitions, a discussion of robustness to relaxation of several key model features, and proofs of all results.

1.1 Evidence of leaks and sabotage

We now provide a set of examples illustrating the prevalence of undermining in a broad range of military, business, and political settings. Our examples share the feature of undermining which is persistent and often spread over a significant time-span. As our examples illustrate, organizational leaders are often acutely aware of the risk of such activities and preoccupied with rooting out their perpetrators.

Wartime provides many classic examples of leaks, sabotage, and other acts of undermining. We highlight two representative instances from the Second World War. The United States suffered a severe informational leak when Communist sympathizers Ethel and Julius Rosenberg supplied Soviet intelligence with confidential documents about American nuclear research, allowing the Soviet Union to quickly advance its nuclear program and perform testing by 1949. The Rosenbergs passed classified documents to Soviet handlers for years before they were eventually discovered, tried and convicted of the acts (Haynes and Klehr 2006; Ziegler and Jacobson 1995). In turn, the United States aided and abetted acts of sabotage by sympathizers in belligerent nations. In the declassified 1944 brief "Simple Sabotage Field Manual," the Office of Strategic Services (the predecessor to the CIA) provided detailed suggestions to sympathizers on how to engage in destructive acts over time in both military and industrial contexts. It advised that "[s]lashing tires, draining fuel tanks, starting fires, starting arguments, acting stupidly, short-circuiting electrical systems, [and] abrading machine parts will waste materials, manpower, and time." The brief also provided advice on how to minimize suspicion and risk of detection: "The potential saboteur should discover what types of faulty decisions and non-cooperation are normally found in his kind of work and should then devise his sabotage so as to enlarge that 'margin of error.' . . . If you commit sabotage on your job, you should naturally stay at your work." This advice contemplated continued and repeated acts of sabotage, often relying on intimate knowledge of organizational operations obtained through trust and experience.

Undermining is also common during peacetime in both business and politics, and major incidents are reported regularly in the news. We highlight several recent examples in corporate and government contexts. In the lead-up to the 2016 election, the social networking company Facebook attracted negative attention when one of its employees leaked information via laptop screenshots to a journalist at the technology news website Gizmodo (Thompson and Vogelstein 2018). On one occasion, the screenshot involved an internal memo from Mark Zuckerberg admonishing employees about appropriate speech. On a second occasion, it showed results of an internal poll indicating that a top question on Facebook employees' minds was about the company's responsibility in preventing Donald Trump from becoming president. Other articles followed shortly after, citing inside sources, about Facebook's suppression of conservative news and other manipulation of headlines in its Trending Topics section by way of "injecting" some stories and "blacklisting" others (Isaac 2016).

The electric car company Tesla has fallen victim to both leaks of sensitive information and sabotage of its codebase. In June 2018, CEO Elon Musk announced in an internal e-mail that an employee had been caught "making direct code changes to the Tesla Manufacturing Operating System under false usernames and exporting large amounts of highly sensitive Tesla data to unknown third parties." This e-mail was itself made public via a leak to the press. The employee was apparently a disgruntled employee who had been passed over for a promotion, but Musk warned employees to remain vigilent against the threat of sabotage at the instigation of "a long list of organizations that want Tesla to die. . . includ[ing] Wall Street short-sellers. . . [and] the oil & gas companies." (Kolodny 2018).

Meanwhile in United States politics, the Trump administration has suffered a series of high-profile leaks of sensitive or damaging information. These leaks include reports on Russian hacking in the United States election; discussions about sanctions with Russia; memos from former FBI director James Comey; transcripts of telephone conversations between Trump and Mexican President Peña Nieto regarding financing of a border wall (Kinery 2017); a timeline for military withdrawal from Syria; and a characterization of some Third-World nations as "shithole countries" (Rothschild 2018). In response to these recurring leaks, the Trump administration has begun actively attempting to trace them to their source,[1] report-

[1] Trump has said on Twitter: "Leakers are traitors and cowards, and we will find out who they are!"

edly via circulating fabricated stories to selected staffers (Cranley 2018).[2] Past presidential administrations have also suffered serious leaks, for instance the disclosure of classified documents by Edward Snowden and Chelsea Manning under the Obama administration (Savage 2017).

1.2 Related Literature

Our paper lies at the intersection of two literatures, one studying gradualism in variable-stakes relationships, and another analyzing optimal dynamic information design. Relative to both literatures, we innovate by introducing dynamic moral hazard in the form of opportunities for repeated, imperfectly observed undermining. We also depart from the gradualism literature by microfounding stakes as the outcome of an information design problem, and use this microfoundation to motivate natural restrictions on the set of feasible stakes curves.

The literature on gradualism studies long-term bilateral relationships with variable stakes where one party can betray the other or expropriate gains. Notable entries in this literature include Watson (1999, 2002), Kreps (2018), and Atakan, Koçkesen, and Kubilay (2019), in which the motives of at least one party are private; and Thomas and Worrall (1994), Albuquerque and Hopenhayn (2004), Rayo and Garicano (2017), and Fudenberg and Rayo (2018), in which preferences are common knowledge. In each of these papers the focus is either on the timing of a single exogenously relationship-ending act, or equivalently a publicly observed act which is punished by permanent termination of the relationship in equilibrium. In contrast, our model focuses on *recurring* undermining which is inflicted repeatedly and observed only stochastically. We are unaware of any other work which studies dynamic moral hazard of this sort in the context of a variable-stakes relationship. Further, the second set of papers cited focus on the timing of the end of the relationship when the motives of each party are commonly known, for instance when one party is known to have an outside option whose value grows with the stakes of the relationship. They therefore abstract from the screening problem at the heart of our model, and focus instead on aligning incentives to extract value from the relationship. Additionally, we microfound the stakes of a relationship via disclosure of information about a payoff-relevant state, in contrast to the typical interpretation of stakes as physical or human capital. This interpretation focuses our model on a different set of applications, implies a natural set of restrictions on feasible stakes curves, and allows us

[2] The practice is known in the intelligence community as a “barium meal test” (Wright and Greengrass 1987) or a “canary trap” (Clancy 1987). It has also been employed by businesses: for instance, Elon Musk has circulated personalized versions of an e-mail containing a bogus non-disclosure agreement (Thomas 2009); Amazon has strategically planted dummy packages in its delivery trucks to track down theft by drivers (Peterson 2018); and Paramount Pictures circulated personalized drafts of the screenplay for *Star Trek III* (Engel 1994).

to investigate how an optimal stakes curve is implemented through dynamic signaling.

The work of Watson (1999, 2002) and Kreps (2018) sits closest to ours by modeling uncertainty over whether agents prefer repeated cooperation to betrayal. Watson studies a binary type space, with one cooperative and one uncooperative type. (He also allows for uncertainty about both players' type.) We describe his work and its relation to ours in more detail below. Kreps extends Watson's model to accommodate a variety of cooperative and uncooperative types, who differ in how much they value betrayal, and focuses on how the optimal rate of stakes raising under commitment varies over time as uncooperative types are screened out at different rates. Atakan, Koçkesen, and Kubilay (2019) also consider a setting in which one agent's motives are private, but in the context of a stage game in which cooperation is ex ante optimal for both parties no matter the agent's type. In their model the basic incentive problem is the ex post incentive of the agent with known motives to manipulate the other agent if his type is revealed. They therefore study the role of gradualism in maintaining cooperation rather than in screening, much as in the literature on gradualism with known preferences.

Watson studies both Pareto-optimal stakes and betrayal policies under commitment to a stakes curve (in Watson (2002)), and equilibrium outcomes under a form of renegotiation-proofness for stakes reminiscent of lack of commitment under one-sided uncertainty (in Watson (1999)).[3] In both papers, he restricts attention to continuous stakes curves.[4] When only one player's cooperativeness is uncertain, his model yields several noteworthy results. First, with renegotiation there exists a unique sustainable stakes curve. Second, efficient stakes grow at the same rate under commitment and renegotiation.[5] And third, betrayal by the uncooperative agent occurs gradually under renegotiation.

Like Watson, we find that stakes grow at the same rate with and without commitment, and that lack of commitment picks out a unique stakes curve and leads to gradual undermining during periods of rising stakes.[6] However, unlike Watson we allow for general

[3]He considers stakes curves and equilibria with the property that at no time does one player want to freeze stakes for a short period of time, nor do both players wish to jump stakes to a point slightly further along the stakes curve. When only one player's type is uncertain, that player never wishes to freeze stakes and always wishes to jump stakes no matter his type. Watson's notion of renegotiation-proofness then reduces to the requirement that the player known to be cooperative does not wish to perturb stakes slightly via a freeze or a jump.

[4]More precisely, he considers stakes curves which are continuous except possibly at time zero, at which time they might jump in a non-right-continuous fashion. Such a jump gives players a chance to betray immediately at the start of the game. However, under one-sided uncertainty such jumps are never useful.

[5]In both cases stakes grow at the discount rate r, when parameters of the model under renegotiation are set to match the version studied under commitment.

[6]By contrast, under commitment Watson (2002) finds that the uncooperative agent optimally betrays either imediately or when stakes reach their maximum. This behavior arises because he assumes that payoffs are positive-sum even when one agent is uncooperative, so that betrayal that does not occur immediately

discontinuous stakes curves, and find that jumps in stakes are crucial to maximizing the principal's profits. This result is closely linked to our novel assumption of repeated undermining. In general whenever some player has flow profits which are linear in stakes, jumps in stakes have the potential to maximize profits at times when flow profits cross zero. However, in environments with discrete betrayal, as in Watson's model, the feasibility of a jump in stakes is severely curtailed by the incentives it provides to concentrate betrayal just after a jump. When undermining is repeated and stochastically detected, on the other hand, incentives for undermining may vary smoothly even across a jump in stakes. Thus repeated undermining tends to cut against the gradualism arising in models with discrete betrayal. Also in contrast to Watson, we study the implications of monotonicity and lower bound constraints on stakes arising naturally from an information design microfoundation, which do not appear in his model. These constraints lead to novel dynamics involving periods of flat stakes and non-monotonic undermining intensity.

Important related work by Sobel (1985) sits adjacent to the gradualism literature and also analyzes a repeated, variable-stakes relationship involving an agent with private motives. That paper shares with ours the modeling device of an agent who might either share the motives of the principal, or else desire opposing outcomes, as in the relationship between a government and a spy when the spy might be a double-agent. Unlike our paper, betrayal is public, stakes are limited by a random upper bound which is drawn iid in every period, and the agent rather than the principal sets stakes. In effect, the only action available to the principal is to withdraw the agent's freedom of action after observing stakes.[7] Sobel thus focuses on the dynamics of a principal-agent relationship when the principal has no tools to screen out the disloyal agent and must live with the resulting dysfunction.

Our paper also contributes to the growing literature on information design sparked by Kamenica and Gentzkow (2011). Our model features persuasion in a dynamic environment, with a privately informed receiver taking actions repeatedly over time subject to imperfect monitoring by the sender. A small set of recent papers, notably Ely, Frankel, and Kamenica (2015), Ely (2017), Ely and Szydlowski (2018), Ball (2018), and Orlov, Skrzypacz, and Zryumov (2018), have studied models of dynamic persuasion without private receiver information. Another emerging strand of the literature has studied the effects of combining persuasion with a receiver with private information. Several involve static persuasion problems. In

(which may be necessary to retain the cooperative agent) is optimally backloaded as much as possible. By contrast, in our model payoffs are zero-sum and so backloading of undermining is not necessary for optimality. Of course, one optimal undermining policy in our commitment setting does involve backloading all undermining during the gradual escalation phase.

[7]Formally, the agent privately observes a binary state of the world corresponding to a correct action choice and makes a cheap-talk report of it to the principal, who then acts. This can be equivalently viewed as the principal either delegating the action to the agent, or else refusing to delegate and dictating an action.

Kolotilin et al. (2017), the receiver's information is about his preference misalignment with the sender, while in Guo and Shmaya (2019) the private information is a signal about the state of the world. Kolotilin (2018) allows for the receiver to possess both private motives and a signal of the state of the world. Meanwhile Au (2015) and Basu (2018) study dynamic persuasion settings. In both papers the receiver chooses the timing of a single public game-ending action, and the sender's preferences are independent of both the state and the receiver's information. To the best of our knowledge, ours is the first paper studying dynamic persuasion of a receiver with private motives who acts repeatedly over time. And none of the papers cited above feature imperfect monitoring of receiver actions.

2 The model

2.1 The environment

A (female) principal hires a (male) agent to perform a task over a potentially infinite horizon in continuous time. The agent's effectiveness at performing this task depends on his knowledge or access to resources, which we model with a nondecreasing, $[0, 1]$-valued *stakes* process x: when the stakes are x_t, the principal's flow payoff from task performance is precisely x_t.[8] The stakes begin at an exogenous level $\phi \in [0, 1)$, meant to capture any information or resources the agent possesses prior to the start of the game or which the principal cannot avoid granting upon hiring the agent, and are thereafter controlled by the principal subject to the monotonicity constraint that stakes cannot be lowered below their current level.[9] Performance of this task is perfectly monitored and not subject to moral hazard, so we model it in reduced form without explicitly introducing a task action for the agent.

In addition to performing the perfectly monitored task, the agent may secretly take actions to undermine the principal. Specifically, at each moment in time the agent chooses an undermining intensity $\beta_t \in [0, 1]$ in addition to performing his publicly observed task.[10] When the stakes are x_t, undermining with intensity β_t inflicts a flow loss of $K\beta_t x_t$ on the principal, where $K > 1$. The principal discounts payoffs at rate $r > 0$, and so given an undermining policy β and a stakes process x her expected payoffs from employing the agent until the (possibly random or infinite) time T are

$$\Pi(\beta, x, T) = \mathbb{E} \int_0^T e^{-rt}(1 - K\beta_t)x_t \, dt.$$

[8] Formally, the principal chooses a right-continuous, increasing $[\phi, 1]$-valued stochastic process.

[9] In Section 5, we study an information-based microfoundation for the stakes process and the constraints we impose on its evoution.

[10] Formally, the agent chooses a càdlàg stochastic process adapted to the history of stakes.

Meanwhile, the agent's motives depend on a preference parameter $\theta \in \{G, B\}$. When $\theta = G$, the agent is *loyal.* A loyal agent has preferences which are perfectly aligned with the principal's. That is, given an undermining policy β and a stakes process x, the loyal agent's expected payoff from being employed until time T is $\Pi(\beta, x, T)$. On the other hand, when $\theta = B$ the agent is *disloyal.* The disloyal agent has interests diametrically opposed to the principal's, and his expected payoff is $-\Pi(\beta, x, T)$. Neither agent type incurs any direct costs from task performance or undermining; they care only about the payoffs the principal receives. If the agent is terminated or does not accept employment, he and the principal both receive an outside option normalized to 0.[11]

2.2 The information structure

Prior to accepting employment with the principal, the agent is privately informed of his type θ. The principal believes that $\theta = G$ with probability $q \in (0, 1)$.

Once the agent is employed, the principal perfectly observes the agent's task performance but only imperfectly observes whether the agent has undermined. In particular, whenever the agent undermines with intensity β_t over a time interval dt the principal immediately receives definitive confirmation of this fact with probability $\gamma \beta_t \, dt$, where $\gamma > 0$. Otherwise the act of undermining goes permanently undetected.

If the agent chooses undermining policy β, the cumulative probability that his undermining has gone undetected by time t is therefore $\exp\left(-\gamma \int_0^t \beta_s \, ds\right)$. Note that the detection rate is *not* cumulative in past undermining — the principal has one chance to detect an act of undermining at the time it occurs. To ensure consistency with this information structure, we assume the principal does not observe her ex post payoffs until the end of the game.

3 The commitment outcome

In this section we derive the principal's optimal contract when she has commitment power. Section 3.1 formally specifies the contracting problem. Sections 3.2 and 3.3 reduce the contracting problem to design of a deterministic stakes curve subject to the contstraint that the disloyal agent prefer to undermine at all times. Section 3.4 states the central proposition characterizing the optimal stakes curve, and provides economic intuition for its qualitative features. Section 3.5 traces the formal derivation of the optimal stakes curve, outlining its

[11] In Appendix B.1, we relax this restriction and allow the principal to replace the replace the agent after termination, possibly at a cost. The presence of such a continuation option does not substantially affect results, whether or not the agent cares about the principal's payoffs beyond his termination.

main steps and explaining their logic. Section 3.6 characterizes when the principal pre-emptively fires the agent without offering an operating contract.

3.1 Contracts

Under commitment, the principal hires the agent by offering a menu of contracts committing to employment terms over the duration of the relationship. Each contract specifies a stakes policy as well as a termination policy and a recommended undermining policy for the agent. No transfers are permitted.[12] Without loss, if the principal hires the agent at all, she commits to fire the agent the first time she observes an act of undermining.[13] Further, under commitment it is not actually necessary to specify a recommended undermining policy. This is because for each agent type, all undermining policies maximizing the agent's payoff yield the principal the same payoff. Thus a contract can be completely described by a, possibly stochastic, stakes policy $x = (x_t)_{t \geq 0}$. Note that every contract automatically satisfies the participation constraint, as there exists an undermining policy guaranteeing the disloyal agent a nonnegative payoff.

In our model the agent possesses private information about his preferences prior to accepting a contract. As a result, as is well-recognized in the mechanism design literature, the principal might benefit by offering distinct contracts targeted at each agent type. However, it turns out that in our setting this is never necessary. As the following lemma demonstrates, no menu of contracts can outperform a single contract accepted by both agent types.

Lemma 1. *Given any menu of contracts (x^1, x^2), there exists an $i \in \{1, 2\}$ such that offering x^i alone weakly increases the principal's payoffs.*

The intuition for this result is simple and relies on the zero-sum nature of the interaction between the principal and a disloyal agent. If the principal offers a menu of contracts which successfully screens the disloyal type into a different contract from the loyal one, she must be increasing the disloyal agent's payoff versus forcing him to accept the same contract as the one taken by the loyal agent. But increasing the disloyal agent's payoff decreases the principal's payoff, so it is always better to simply offer one contract and pool the disloyal agent with the loyal one. Thus screening of disloyal agents must be accomplished by direct detection of undermining rather than self-reports by the agent.

In light of the developments in this subsection, the principal's problem is to choose a

[12]See Appendix B.2 for a discussion of how transfers affect the analysis.

[13]We assume for now that the principal hires the agent. We characterize when pre-emptive firing in section 3.6.

single stakes policy x maximizing the payoff function

$$\begin{aligned}\Pi(x) = {} & q\,\mathbb{E}\int_0^\infty e^{-rt}x_t\,dt \\ & - (1-q)\sup_\beta \mathbb{E}\int_0^\infty \exp\left(-rt - \gamma\int_0^t \beta_s\,ds\right)(K\beta_t - 1)x_t\,dt,\end{aligned} \tag{1}$$

where expectations are with respect to uncertainty in x and β (in case the disloyal agent employs a mixed strategy or conditions on the outcomes of a stochastic stakes process). The term x_t appears both as a benefit of raising stakes with a loyal agent and as a cost of raising stakes with a disloyal agent, in case the latter has not yet been revealed and would currently undermine with sufficiently high intensity.

3.2 Deterministic stakes

We allow the principal to commit to a stochastic stakes policy as part of a contract. However, in this subsection we show that the possibility of randomization can be ignored without loss of generality.

Definition 1. *A stakes policy x is* deterministic *if it satisfies $x_t = \mathbb{E}[x_t]$ at all times.*

The following lemma justifies focusing on deterministic policies, by showing that the principal can pass from an arbitrary stakes policy to a deterministic counterpart yielding the same *ex ante* expected stakes and a weakly higher payoff.[14] The basic intuition is that varying the realization of the stakes allows the disloyal agent to tailor her undermining strategy more precisely to the state of the world. Pooling these states reduces the disloyal agent's flexibility without reducing the average effectiveness of the loyal agent's actions.

Lemma 2. *Given any stakes policy x, the deterministic stakes policy x' defined by $x'_t = \mathbb{E}[x_t]$ for all t yields a weakly higher payoff to the principal than x.*

When the principal employs a deterministic stakes policy, her payoff function may be written

$$\Pi(x) = q\int_0^\infty e^{-rt}x_t\,dt - (1-q)\sup_\beta \mathbb{E}\int_0^\infty \exp\left(-rt - \gamma\int_0^t \beta_s\,ds\right)(K\beta_t - 1)x_t\,dt,$$

where the expectation is over possible randomness in the agent's undermining policy. In light of Lemma 2, for most of the analysis we will restrict attention to deterministic stakes policies. We will refer to the path of stakes under a deterministic stakes policy as a *stakes curve.*

[14] In fact, we establish later in Lemma 10 that there are no optimal contracts involving randomized stakes.

3.3 Loyalty tests

We will be particularly interested in studying stakes policies which do not provide strict incentives for disloyal agents to undermine with less than full intensity at any point during their employment.

Definition 2. *A stakes policy is a* loyalty test *if undermining with full intensity at all times is an optimal strategy for the disloyal agent.*

An important first result in our analysis is that the principal cannot benefit by using the promise of future stakes increases to induce loyalty by a disloyal agent early on. Instead, the principal optimally employs only loyalty tests to screen disloyal from loyal agents. The following lemma states the result.

Lemma 3. *Suppose a stakes policy x is not a loyalty test. Then there exists a loyalty test x' yielding a strictly higher payoff to the principal than x.*

A key ingredient of this result is the opposition of interests of the principal and disloyal agent. Fixing a stakes policy, if the disloyal agent finds it optimal to defer undermining for a time, the gains from remaining employed and exploiting higher future stakes must outweigh the losses from failing to undermine the principal early on. As gains to the disloyal agent are losses to the principal, the principal suffers from the disloyal agent's decision to defer undermining. Of course, the principal cannot force the disloyal agent to engage in undermining. The best she can do is design a loyalty test to ensure the disloyal agent cannot raise his payoff by feigning loyalty. Thus testing loyalty constrains the set of implementable stakes policies, so it is not immediate from the previous logic that loyalty tests are optimal.

The additional insight leading to the result is the observation that the payoff of any stakes policy which is not a loyalty test can be improved by raising stakes more quickly. Intuitively, if over some time interval the disloyal agent strictly prefers to feign loyalty, then the principal may bring forward some increment of stakes from the end to the beginning of this interval without disturbing the optimality of feigning loyalty during the interval. By bringing forward a sufficiently large increment, the principal can make the disloyal agent indifferent among all undermining intensities over the interval. This modification also makes the policy locally a loyalty test over the interval, without improving the disloyal agent's payoff, since deferring undermining remains (weakly) optimal and thus the acceleration of stakes can't be used to harm the principal. And the modification improves the principal's payoff when the loyal agent is present, as the loyal agent is more effective at greater stakes.

The proof of Lemma 3 builds on this insight by modifying a given stakes policy beginning at the first time deferring undermining becomes optimal. The modified policy increases stakes as much as possible in a lump at that instant without improving the disloyal agent's continuation utility. Afterward, the stakes increase smoothly at the maximum rate consistent with the constraints of a loyalty test. This process is not guaranteed to yield greater stakes at every point in time versus the original policy. However, the proof shows that stakes are brought forward sufficiently to improve the loyal agent's payoff, and therefore the principal's.

3.4 The optimal stakes curve

In this section we provide an intuitive, heuristic derivation of the optimal stakes curve, culminating in its characterization in Proposition 1. To aid exposition, we defer formal arguments to section 3.5.

The derivation begins by noting that the design of an optimal stakes curve can be stated as maximization of the principal's expected profits subject to the incentive constraint that the resulting curve be a loyalty test. To grapple with this problem, it is helpful to first consider the relaxed problem without the incentive constraint, in which the disloyal agent exogenously undermines with full intensity at all times. Given a solution to this relaxed problem, if the disloyal agent does in fact prefer to undermine fully at all times, it is also a solution to the original problem. Let

$$\overline{\pi}_t \equiv \Pr(\theta = G \mid \beta = 1, \text{no observed undermining by time } t)$$

be the agent's time-t reputation in the relaxed problem. By Bayes' rule,

$$\overline{\pi}_t = \frac{q}{q + (1-q)e^{-\gamma t}},$$

where recall that $\overline{\pi}_0 = q$ is the prior probability that the agent is loyal. Meanwhile, in the relaxed problem the principal's payoff from (1) may be written

$$\Pi(x) = \int_0^\infty e^{-rt} \left(q + (1-q)e^{-\gamma t}\right) \left(\overline{\pi}_t - (K-1)(1-\overline{\pi}_t)\right) x_t \, dt.$$

Flow profits are maximized pointwise by setting $x_t = \phi$ whenever t is small enough that $\overline{\pi}_t < (K-1)/K$, and setting $x_t = 1$ afterward. We will let

$$t^* \equiv \inf\{t \ : \ \overline{\pi}_t \geq (K-1)/K\}$$

denote the cutoff time at which the optimal stakes curve in the relaxed problem jumps to 1.

If the agent's initial reputation is at least $\frac{K-1}{K}$, then profits are maximized by setting stakes to their maximal level immediately. Under this policy, the disloyal agent trivially has no incentive to defer undermining, so the policy solves the unrelaxed problem as well. On the other hand, if the agent's initial reputation is below $\frac{K-1}{K}$, then $t^* > 0$ and it must be checked that the disloyal agent at least weakly prefers to undermine at full intensity prior to time t^*. As might be expected, the disloyal agent is most strongly tempted to defer undermining just prior to time t^*, when his continuation payoff is largest relative to his flow payoff from undermining. Provided that the jump in stakes at time t^* is not too large — that is, provided that ϕ is above some threshold $\overline{x}$ — the disloyal agent is willing to undermine with full intensity at all times, and the solution to the relaxed problem is also the solution to the original problem.

If the initial stakes are below $\overline{x}$, then the relaxed solution fails to solve the full problem because it is not a loyalty test. In this case stakes must rise prior to a jump in stakes to the maximal level in order to preserve the loyalty test property. The following definition describes policies which change stakes as quickly as possible prior to a final jump without inducing incentives to defer undermining. An optimal stakes curve will turn out to satisfy this property.

Definition 3. *A stakes curve x satisfies the* optimal-growth *property if there exist time thresholds $\underline{t} \geq 0$ and $\overline{t} \geq \underline{t}$ such that*

$$x_t = \begin{cases} \phi, & t < \underline{t} \\ x_{\underline{t}} \exp\left((r + \gamma/K)(t - \underline{t})\right), & \underline{t} \leq t < \overline{t} \\ 1, & \overline{t} \leq t \end{cases}$$

and x is continuous at every $t \neq \overline{t}$. In particular, $x_{\underline{t}} = \phi$ if $0 < \underline{t} < \overline{t}$.

One requirement imposed by the optimal-growth property is that, after an initial quiet period with no stakes growth, x must grow at rate $r + \gamma/K$ until the final jump time $\overline{t}$. The intuition for the desirability of this property is as follows. Recall that, all else equal, the principal maximizes profits by minimizing the stakes prior to time t^*. Thus, looking earlier in the contract from t^*, the stakes should deteriorate as quickly as possible consistent with incentive-compatibility. Equivalently, going forward in time stakes are increased as quickly as possible. It turns out that the maximum permissible rate of increase is $r + \gamma/K$. This modification to the ideal stakes curve can be visualized graphically — the constant stakes level prior to time t^* is shifted upward to ensure undermining is optimal near time t^*, and then is tilted counterclockwise as steeply as possible consistent with incentive compatibility.

The second requirement imposed by the optimal-growth property is that x not drop below ϕ, the minimum admissible stakes level. This property must be imposed because if ϕ is large, then the shift-and-tilt described previously may produce a stakes curve which drops below ϕ early in the contract. This violates the lower bound constraint on x required for implementability. As a remedy, the principal optimally irons the stakes curve so that it does not drop below ϕ. Figure 1b illustrates this ironing in the early stage of the contract. Ironing will turn out to be necessary in an optimal contract when ϕ lies above a threshold $\underline{x} \in (0, \overline{x})$.

The "tilt-and-shift" procedure just described for adapting the relaxed-optimal stakes curve to achieve a loyalty test holds fixed the time of the jump to maximal stakes at t^*. In fact, the principal will in general also wish to delay this time, choosing a jump time $\overline{t}$ which is larger than t^*. The tradeoff faced by the principal in choosing $\overline{t}$ is that increasing $\overline{t}$ reduces stakes and losses early in the contract (when the agent's reputation is below $(K-1)/K$), but also reduces stakes and profits late in the contract (when his reputation is above $(K-1)/K$). When the loyalty test incentive constraint binds and the jump time is exactly t^*, shifting the jump a bit to the right entails no first-order losses later in the contract, as flow profits just to the right of t^* are approximately zero. On the other hand, the increase in flow profits early on in the contract is positive to first-order. The unique optimal stakes curve chooses $\overline{t} > t^*$ to balance the flow losses and gains from a rightward shift in the jump time.

To summarize, the optimal stakes proceeds through up to three distinct phases: an quiet period of no stakes growth, a gradual escalation period, and finally a jump to maximal stakes. This stakes curve satisfies the optimal growth property throughout the gradual escalation phase. Proposition 1, stated below, formally characterizes the optimal stakes curve.

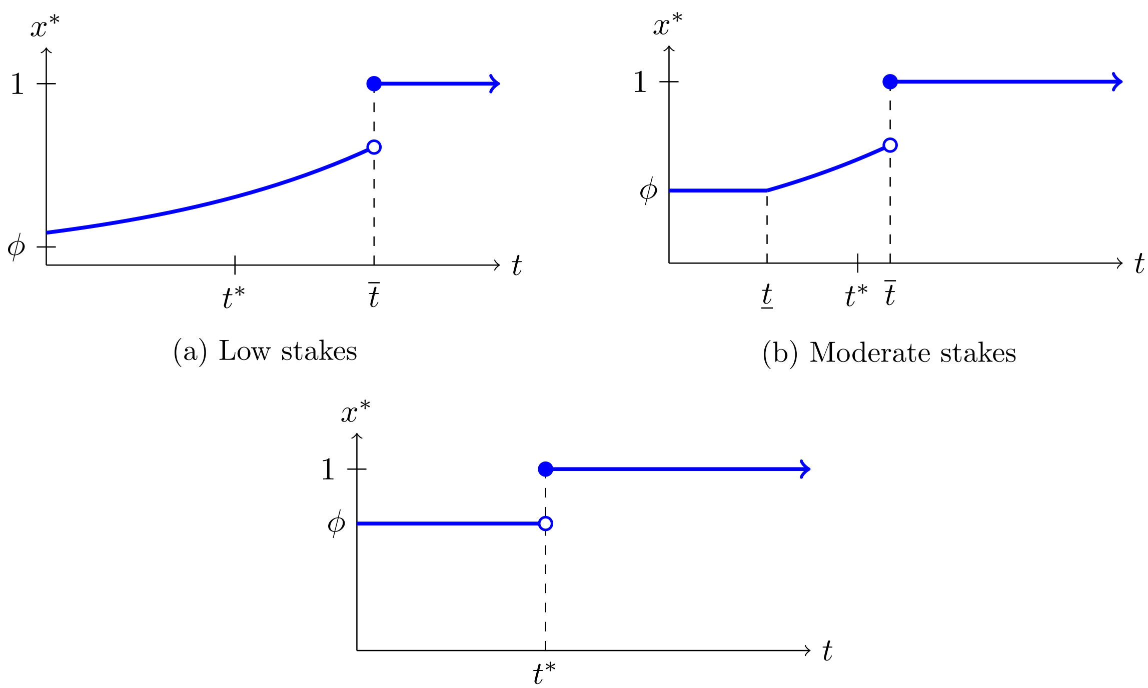
Proposition 1 (Optimal Stakes Curve)**.** *There exists a unique optimal stakes curve x^*, and time thresholds $\underline{t}$ and $\overline{t}$ such that x^* satisfies the optimal-growth property on $[\underline{t}, \overline{t}]$. If $q \geq (K-1)/K$, then $\underline{t} = \overline{t} = 0$. Otherwise, there exist stakes thresholds $\underline{x}$ and $\overline{x} > \underline{x}$ in $(0,1)$, independent of ϕ, such that:*

- *If $\phi \geq \overline{x}$, then $\underline{t} = \overline{t} = t^*$,*
- *If $\underline{x} < \phi < \overline{x}$, then $0 < \underline{t} < t^* < \overline{t}$,*
- *If $\phi \leq \underline{x}$, then $0 = \underline{t} < t^* < \overline{t}$ and $x_0^* = \underline{x}$.*

As the proposition demonstrates, whenever $q < (K-1)/K$ the optimal stakes curve exhibits distinctive behavior depending on the magnitude of ϕ. When $\phi \leq \underline{x}$, the optimal stakes curve has a gradual escalation phase but no quiet period, a case we will refer to as the *low stakes* regime. When $\phi \in (\underline{x}, \overline{x})$ the optimal stakes curve has both a quiet period and a gradual escalation phase, a case we will refer to as the *moderate stakes* regime. Finally, when $\phi \geq \overline{x}$ the optimal stakes curve has a quiet period but no gradual escalation phase, a case we will refer to as the *high stakes* regime. Figure 1 illustrates the optimal stakes curve in each of the three regimes.

Under the stakes curve x^*, the disloyal agent is indifferent over all undermining policies on the time interval $[\underline{t}, \overline{t})$, at which times the optimal-growth property ensures that the loyalty test constraint just binds locally. Meanwhile the agent strictly prefers to undermine at all other times, as the loyalty test constraint is slack locally outside $[\underline{t}, \overline{t})$. Thus the agent's set of optimal undermining policies is as described in the following lemma. Note in particular that whenever $\overline{t} > \underline{t}$, the agent has many optimal undermining policies. Due to the zero-sum property of the interaction between principal and disloyal agent, none of these policies is privileged from the point of view of maximizing principal profits.

Lemma 4. *Under stakes curve x^*, an undermining policy $\tilde{\beta}$ is optimal for the disloyal agent iff $\tilde{\beta}_t = 1$ for $t \in [0, \underline{t}) \cup [\overline{t}, \infty)$.*



(a) Low stakes

(b) Moderate stakes

(c) High stakes

Figure 1: The optimal stakes curve.

3.5 Deriving the optimal stakes curve

In this section we provide an overview of the formal proof of Proposition 1. In particular, we establish a series of lemmas which lay the groundwork for the proof. The formal proof of the proposition, stated in Appendix D.4, builds on these lemmas to complete the derivation.

In light of Lemma 3, the principal optimizes her payoff among the set of deterministic stakes policies which are loyalty tests. We must therefore characterize the constraints that testing loyalty places on stakes curves. Given a stakes curve x, define

$$U_t \equiv \int_t^\infty e^{-(r+\gamma)(s-t)}(K-1)x_s\,ds$$

to be the disloyal agent's ex ante continuation utility at time t, supposing he has not been terminated before time t and undermines with full intensity forever afterward. The following lemma characterizes a necessary and sufficient condition for a stakes curve to be a loyalty test.

Lemma 5. *A stakes curve x is a loyalty test if and only if*

$$\dot{U}_t \le \left(r + \frac{\gamma}{K}\right) U_t. \tag{2}$$

Equation (2) holds at time t if and only if the gains from undermining at that instant outweigh the accompanying risk of discovery and dismissal.[15]

In light of Lemmas 2, 3, and 5, the principal's design problem reduces to solving

$$\sup_{x\in\mathbb{X}} \int_0^\infty e^{-rt}x_t(q-(1-q)(K-1)e^{-\gamma t})\,dt \quad \text{s.t.} \quad \dot{U}_t \le (r+\gamma/K)U_t, \tag{PP}$$

where $\mathbb{X}$ is the set of monotone, càdlàg $[\phi, 1]$-valued functions.

The form of the constraint suggests transforming the problem into an optimal control problem for U. The following lemma facilitates this approach, using integration by parts to eliminate x from the objective in favor of U.

Lemma 6. *For any $x \in \mathbb{X}$,*

$$\int_0^\infty e^{-rt}x_t(q-(1-q)(K-1)e^{-\gamma t})\,dt = -\left(1 - \frac{K}{K-1}q\right)U_0 + \frac{q\gamma}{K-1}\int_0^\infty e^{-rt}U_t\,dt.$$

The final step in the transformation of the problem is deriving the admissible set of utility processes corresponding to the admissible set $\mathbb{X}$ of stakes curves. The bounds on x immediately imply that $U_t \in [\underline{U}, \overline{U}]$ for all t, where $\overline{U} \equiv (K-1)/(r+\gamma)$ and $\underline{U} \equiv \phi\overline{U}$. It is

[15]Under a stochastic stakes policy, U_t can be defined using $\mathbb{E}[x_t]$ in place of x_t, and (2) remains a necessary condition for a loyalty test. The condition is not sufficient in general, as incentive-compatibility might be achieved on average but not ex post in some states of the world if the underlying stakes policy is stochastic. However, since we are able to restrict attention to (deterministic) stakes curves by Lemma 2, ex ante and ex post incentives are identical, and (2) is sufficient.

therefore tempting to conjecture that problem (PP) can be solved by solving the auxiliary problem

$$\sup_{U \in \mathbb{U}} \left\{ -\left(1 - \frac{K}{K-1} q\right) U_0 + \frac{q\gamma}{K-1} \int_0^\infty e^{-rt} U_t \, dt \right\} \qquad \text{s.t.} \qquad \dot{U}_t \leq (r + \gamma/K) U_t, \qquad (PP')$$

where $\mathbb{U}$ is the set of absolutely continuous $[\underline{U}, \overline{U}]$-valued functions. Any solution to this problem can be mapped back into a stakes curve via the identity

$$x_t = \frac{1}{K-1} \left((r+\gamma) U_t - \dot{U}_t \right), \tag{3}$$

obtainable by differentiating the expression $U_t = \int_t^\infty e^{-(r+\gamma)(s-t)} x_s \, ds$.[16]

It turns out that this conjecture is correct if ϕ is not too large, but for large ϕ the resulting solution fails to respect the lower bound $x \geq \phi$. To correct this problem equation (3) can be combined with the lower bound $x_t \geq \phi$ to obtain an additional upper bound $\dot{U}_t \leq (r+\gamma) U_t - (K-1)\phi$ on the growth rate of U. This bound suggests solving the modified auxiliary problem

$$\begin{aligned} &\sup_{U \in \mathbb{U}} \left\{ -\left(1 - \frac{K}{K-1} q\right) U_0 + \frac{q\gamma}{K-1} \int_0^\infty e^{-rt} U_t \, dt \right\} \\ &\text{s.t.} \qquad \dot{U}_t \leq \min\{(r+\gamma/K) U_t, (r+\gamma) U_t - (K-1)\phi\} \end{aligned} \qquad (PP'')$$

The following lemma verifies that any solution to this relaxed problem for fixed $U_0 \in [\underline{U}, \overline{U}]$ yields a stakes curve respecting $x \geq \phi$. Thus a full solution to the problem allowing U_0 to vary solves the original problem.

Lemma 7. *Fix $u \in [\underline{U}, \overline{U}]$. There exists a unique solution $U^{**}(u)$ to problem (PP'') subject to the additional constraint $U_0 = u$, and the disclosure policy x^u defined by*

$$x_t^u = \frac{1}{K-1} \left((r+\gamma) U^{**}(u)_t - \dot{U}^{**}(u)_t \right)$$

satisfies $x^u \in \mathbb{X}$.

We prove Proposition 1 by solving problem (PP'') subject to $U_0 = u \in [\underline{U}, \overline{U}]$, and then in a final step optimize over u to obtain a solution to problem (PP'') and thus to problem (PP). When U_0 is held fixed, the objective is increasing in U pointwise, so optimizing the objective

[16] In general this identity need only hold a.e., given the possible non-differentiability of U on a set of measure zero. In this case many stakes curves consistent with U may exist. When U has a well-defined right-hand derivative everywhere, we will select the unique x for which the identity holds everywhere substituting the right-hand derivative of U for $\dot{U}$. As the optimal utility process lies in this class, we need not specify the selection for more general U.

amounts to maximizing the growth rate $\dot{U}$ subject to the control and the upper bound $U \leq \overline{U}$. The result is that U solves the ODE $\dot{U}_t = \min\{(r+\gamma/K)U_t, (r+\gamma)U_t - (K-1)\phi\}$ until the point at which $U_t = \overline{U}$, and then is constant afterward. Solving problem PP then reduces to a one-dimensional optimization over U_0, which can be accomplished algebraically.

The forces shaping an optimal continuation utility path can be cleanly mapped onto the forces shaping an optimal stakes curve identified in Section 3.4. An important property of the solution to problem PP'' is that only a single constraint on $\dot{U}_t$ binds at a given time, with the constraint $\dot{U}_t \leq (r+\gamma)U_t - (K-1)\phi$ binding early in the optimal contract while $\dot{U}_t \leq (r+\gamma/K)U_t$ binds later on. The first constraint binding corresponds to a constant stakes level $x_t = \phi$, while the second constraint binding corresponds to x growing at rate $r+\gamma/K$. Thus, holding fixed the jump time $\bar{t}$, the optimal stakes curve declines at rate $r+\gamma/K$ as one rewinds the contract toward time 0, until the lower bound $x = \phi$ is reached and the principal cannot set the stakes any lower. The growth rate $r+\gamma/K$ can be interpreted as the "fastest possible" growth rate for x, in a global average sense. (It is of course possible to achieve locally higher growth rates of x by designing a very concave U, with the extreme case of a concave kink corresponding to a discrete increase in stakes.)

Meanwhile the jump in stakes at time $\bar{t}$ arises due to the kink in the continuation utility curve when it reaches the upper bound $\overline{U}$. The size of the jump depends on the slope of the utility curve to the left of the kink, which is optimally made as large as possible consistent with the IC constraint and $x \geq \phi$. This exactly accords with the considerations shaping the size of the jump discussed in Section 3.4. Finally, the timing of the final jump is controlled by the choice of the initial continuation utility U_0. The smaller the choice of U_0, the longer continuation utility must grow before it reaches $\overline{U}$, and thus the later the time of the final jump. So the optimization over $\bar{t}$ discussed informally in Section 3.4 is achieved formally by optimizing U_0 in the control problem PP''.

3.6 Pre-emptive firing

We have so far ignored the possibility that the principal might fire the agent even if he is not known to be disloyal. This possibility is relevant if the agent's initial reputation q is small and minimal stakes ϕ are large. If the agent is ever fired while his type is uncertain, termination will optimally take place at time zero when his reputation is as low as possible, i.e. before the contract is signed. In this subsection we discuss when such pre-emptive firing is optimal. Equivalently, we identify when the optimal contract characterized above is profitable for the principal.

Figure 2 displays a representative division of the (q, ϕ)-parameter space into hire and no-hire regions. Regions I, II, and III represent parameter regions where $q < (K-1)/K$ and the relationship begins with high, moderate, and low stakes, respectively. Meanwhile region IV is the set of parameters for which $q \geq (K-1)/K$. The shaded region represents the parameter values for which the principal prefers to pre-emptively fire the agent immediately rather than offer an operating contract. Note that this outcome occurs only for a subset of regions I and II, when the agent begins with moderate or high stakes. By contrast, if $q \geq (K-1)/K$ or the agent begins with low stakes, the optimal operating contract always yields strictly positive payoffs, and pre-emptive firing does not occur. This is unsurprising given that in these situations, the optimal contract involves a voluntary increase of stakes by the principal at time zero.

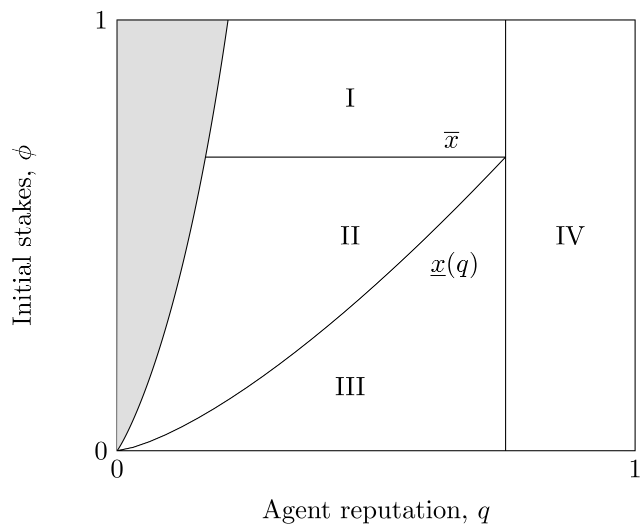


Figure 2: Optimal contract regimes in (q,ϕ)-space. Regions I, II, III represent the high, moderate, and low stakes cases when $q < (K-1)/K$, while in region IV $q \geq (K-1)/K$. Pre-emptive firing is optimal in the shaded region.

Lemma 8 formalizes the results of this subsection. For $x \in \{H, M, L\}$, let $\mathcal{S}^x$ denote the set of (q,ϕ) pairs in the high, moderate and low stakes regimes, respectively, and let $\mathcal{S}^+ = \{(q,\phi) : q \geq \frac{K-1}{K}\}$.

Lemma 8. *There exists a nonempty set $\mathcal{S}^\emptyset$ of (q,ϕ) pairs for which the principal strictly prefers not to hire the agent, and this set satisfies $\mathcal{S}^\emptyset = \{(q,\phi) : q < \underline{q}(\phi)\}$ for an increasing, continuous function $\underline{q} : [0,1] \to [0,1]$ satisfying $q(0) = 0$ and $q(1) < (K-1)/K$. The set $\mathcal{S}^\emptyset$ intersects $\mathcal{S}^H$ and $\mathcal{S}^M$ but it does not intersect $\mathcal{S}^L$ or $\mathcal{S}^+$, and neither $\mathcal{S}^H$ nor $\mathcal{S}^M$ is a subset of $\mathcal{S}^\emptyset$.*

4 The no-commitment outcome

While the optimal commitment contract sets a benchmark for efficient screening, in high-stakes applications one might reasonably doubt that a principal can credibly commit to future levels of access. In this section we study what outcomes are possible without commitment, i.e. in a setting of purely relational contracting.[17] Our main result is that the commitment stakes curve is the unique equilibrium path of any pure-strategy Bayes Nash equilibrium, and equilibrium play selects a unique incentive-compatible undermining path for the disloyal agent. Moreover, this outcome is supportable as a perfect Bayesian equilibrium. The equilibrium undermining path features rising undermining intensity throughout the gradual escalation phase of the relationship, implying a non-monotonic undermining path whenever the relationship begins in the moderate-stakes regime. The contracting outcome is therefore not only completely robust to relaxation of the commitment assumption, but in fact produces sharper predictions about disloyal agent behavior without commitment.

We establish this result in several steps. First, we develop a notion of time-consistency for contracts. Time-consistent contracts satisfy the intuitively appealing property that, were the principal able to unexpectedly renege and write a new contract at some future date, she would never want to. We show that there exists a unique time-consistent, incentive-compatible contract without randomization, which by definition implements the optimal stakes curve x^*, and characterize its undermining policy. We then prove that no contract which is not both time-consistent and incentive-compatible can be sustained in any pure-strategy Bayes Nash equilibrium. Finally, we construct a perfect Bayesian equilibrium sustaining the time-consistent, incentive-compatible contract. We conclude the section by discussing equilibria in mixed strategies.

4.1 Time-consistent contracts

We begin by recalling that under commitment, it was not necessary to explicitly specify a recommended undermining policy for the disloyal agent. This is because the zero-sum nature of the game between principal and disloyal agent implies that indifferences for the agent correspond to indifferences for the principal. Thus during any time period over which the disloyal agent is indifferent toward undermining under some stakes path, there are a multiplicity of recommended undermining policies which are optimal for the principal. In particular, in the optimal commitment contract, the undermining policy during the gradual

[17]Throughout this section, we assume that $q \geq \underline{q}(\phi)$, with $\underline{q}$ as characterized in Lemma 8. Otherwise the principal faces no commitment problem implementing the commitment outcome, which entails immediately firing the agent.

escalation phase is indeterminate.

By contrast, without commitment the disloyal agent's rate of undermining impacts principal incentives ex post and so must be specified even in regions of indifference. Throughout this section we will therefore define a *contract* as a pair (x, β), specifying both the principal's stakes curve x and the disloyal agent's undermining policy β. For most of this section, we will focus on *deterministic* contracts, in which both x and β are deterministic functions of time. We will further say that a contract (x, β) is *incentive-compatible* if β minimizes the principal's expected profits given the undermining curve x.

The rate of undermining is important without commitment because it controls an additional state variable of central interest for incentives to deviate from the equilibrium path — the principal's ex post beliefs about the agent's loyalty. The more intensively the principal expects the agent to undermine, the more quickly her beliefs about the agent's loyalty rise absent a discovery of undermining. And the level of the principal's beliefs impact her incentives to renege on a promised stakes path and either raise them more quickly, to favor an ex post trustworthy agent, or freeze them, to shield herself from an ex post suspicious agent.

It turns out that under stakes curve x^*, there is exactly one deterministic undermining policy which raises the principal's beliefs at a rate such that she prefers neither to raise stakes more quickly nor freeze stakes after a period with no discovery of undermining. To state the result precisely, we formulate a notion of time-consistency which plays a central role in our equilibrium analysis. Define an indexed set of optimal stakes curves $x^{**}(y, p)$ for $y, p \in [0, 1]$, where $x^{**}(y, p)$ is the (unique) optimal contractual stakes curve when initial stakes are y and the agent is loyal with probability p. This family of curves is characterized by Proposition 1 as ϕ and q are allowed to vary. In particular, $x^{**}(\phi, q) = x^*$.

Definition 4. *A deterministic contract (x, β) is* time-consistent *if at each time t, $(x_s)_{s \geq t} = x^{**}(x_{t-}, \pi_t)$, where π_t are the principal's time-t posterior beliefs that $\theta = G$ conditional on no observed undermining and undermining policy β.*

Note in particular that any time-consistent contract must implement the optimal stakes curve x^*.

Our first result is that there exists a unique deterministic β^* under which the contract (x^*, β^*) is incentive-compatible and time-consistent. (We defer discussion of randomized undermining policies to Section 4.3.)

Lemma 9. *There exists a unique deterministic undermining policy β^* such that (x^*, β^*) is incentive-compatible and time-consistent. β^* satisfies:*

- $\beta^*_t = 1$ *for* $t \in [0, \underline{t}) \cup [\overline{t}, \infty)$,

- *If $\underline{t} < \bar{t}$, then β^* is a strictly increasing, continuous function on $[\underline{t}, \bar{t})$ satisfying $\beta^*(\underline{t}) \geq 1/(K(1-q))$ and $\beta^*(\bar{t}-) = 1$.*

The shape of β^* is plotted in Figure 3. Note that $\beta^*(\underline{t}) < 1$, so that undermining jumps downward at the beginning of the gradual escalation phase. Further, $\beta^*(t) > 1/K$ for all $t \in [\underline{t}, \bar{t}]$, so the disloyal agent inflicts a net negative flow payoff on the principal throughout the relationship.

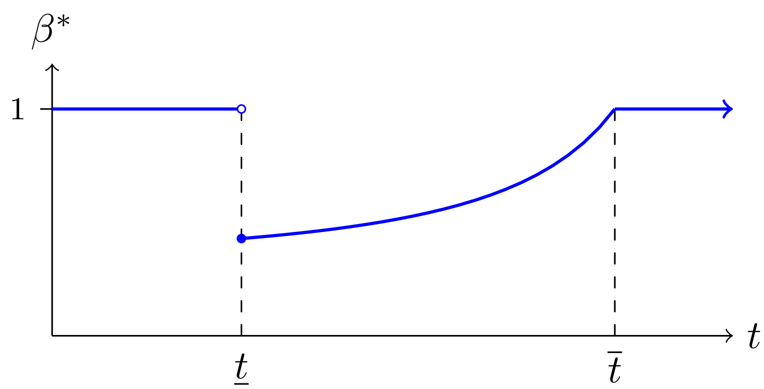


Figure 3: The time-consistent undermining path.

The basic idea behind the proof is that, for fixed minimal stakes, the optimal stakes curves for varying initial beliefs q differ only by a time shift, with higher q corresponding to a stakes curve shifted further to the left. Then as the curve $(x_s^*)_{s \geq t}$ is simply a left-shifted version of x^*, shifting q up appropriately, say to q_t', yields an optimal stakes curve $x^{**}(\phi, q_t')$ which can be made to coincide with the continuation stakes curve $(x_s^*)_{s\geq t}$. Since $x^{**}(\phi, q_t')_0 = x_t^* \geq \phi$ by construction, it follows that the lower bound constraint doesn't bind at time t even if raised from ϕ to x_{t-}^*, so that also $(x_s^*)_{s\geq t} = x^{**}(x_{t-}^*, q_t')$. Hence by guiding beliefs appropriately, the continuation stakes curve can be made to coincide at every point in time with the optimal stakes curve given current initial stakes and beliefs. The belief path q' can be shown to be uniquely determined and strictly increasing in time for $t < \bar{t}$ (after which stakes are at their maximum and the principal's set of continuation strategies is a singleton). These beliefs can then be inverted to deduce the unique undermining policy β^* under which the posterior belief process π satisfies $\pi_t = q_t'$ for $t < \bar{t}$. The most involved step in the proof establishes that $\beta^* \leq 1$, so that this policy is feasible, and that $\beta_t^* = 1$ for $t < \underline{t}$, so that the time-consistent undermining path is also incentive-compatible.[18]

By varying q and ϕ in the proof of Lemma 9, a direct corollary is that for each $y, p \in [0, 1]$, there exists a unique time-consistent, incentive-compatible undermining policy $\beta^{**}(y, p)$ cor-

[18]Recall that by Lemma 4, incentive-compatibility is automatic on the time interval $[\underline{t}, \bar{t})$.

responding to the optimal stakes curve $x^{**}(y, p)$ when initial stakes are y and beliefs are p. Another immediate consequence is that for all t, $(\beta^*)_{s \geq t} = \beta^{**}(x_t^*, \pi_t^*)$, where π^* are the posterior beliefs induced by β^*. For otherwise the undermining policy $\widetilde{\beta} = ((\beta^*)_{s \in [0,t)}, \beta^{**}(x_t^*, \pi_t^*)_{s \geq t})$ would constitute another incentive-compatible, time-consistent undermining policy under x^* and initial state (ϕ, q), contradicting the uniqueness guaranteed by Lemma 9.

4.2 Equilibrium analysis

Time-consistency is seemingly a stronger condition than necessary for equilibrium, because it requires that the equilibrium stakes curve be proof against deviations which yield the principal her optimal commitment value as a continuation payoff. In general, we might expect that the agent could respond to deviations in a way which reduces the principal's payoff compared to the commitment value, reducing their attractiveness and supporting additional stakes curves as equilibria. Our next result shows that such punishments are not possible, establishing that time-consistency and incentive-compatibility are necessary conditions for a (pure-strategy) equilibrium.[19]

Proposition 2. *If the deterministic contract (x, β) is not time-consistent and incentive-compatible, then there exists no Bayes Nash equilibrium with equilibrium path (x, β).*

The idea behind the proof is the recognition that the principal's optimal contractual profits are calculated assuming that the (disloyal) agent responds by choosing a strategy which minimizes the principal's profits under the optimal stakes curve. Thus, should the principal deviate at any point, *no* continuation strategy by the agent can impose a continuation payoff less than the principal's optimal contractual profits. Even if the agent forms expectations about future stakes from which the principal further deviates, any change in the agent's actions can only do worse at minimizing the principal's payoff. Thus after every history, the principal must collect at least her optimal contractual profits at current stakes and beliefs.

Further, after no history can the principal collect strictly more than her commitment profits, as any such outcome would be sustainable under commitment as well. The principal's profits after each history are therefore uniquely pinned down, and as these profits are uniquely achievable by the optimal stakes curve $x^{**}(x_{t-}, \pi_t)$, this curve is the unique possible continuation after every history. In particular, at time 0 the stakes curve $x^*(x_{t-}, q)$ is the only possible equilibrium stakes curve, and to eliminate profitable deviations it must be that

[19] Games with observable actions in continuous time suffer from well-known issues with the coherent definition of strategies. Because only one player's actions are observable in our setting, we can sidestep these issues through an appropriate mapping from strategies to outcomes tailored to our setting. See Appendix A for a formal definition of the equilibrium machinery used in the following results.

$(x^*_s)_{s\geq t} = x^{**}(x^*_{t-}, \pi_t)$ for all t. Thus the agent must employ the undermining policy β^* on the equilibrium path.

The previous result establishes necessity of (x^*, β^*) as the equilibrium path of any pure-strategy Bayes Nash equilibrium. However, existence of an equilibrium is not automatic in a continuous-time environment. Further, if we desire some notion of subgame perfection, the proposition does not provide guidance as to what sort of credible off-path behavior might support the equilibrium path. We therefore complete the analyis by directly constructing a perfect Bayesian equilibrium supporting (x^*, β^*).

Proposition 3. *There exists a perfect Bayesian equilibrium in pure strategies with equilibrium path* (x^*, β^*).

This equilibrium is constructed using agent and principal strategies which are Markovian in current stakes and posterior beliefs about the agent's loyalty. Given current stakes x_{t-} and beliefs π_t, the principal's continuation strategy is to play the stakes curve $(x_s)_{s\geq t} = x^{**}(x_{t-}, \pi_t)$. Meanwhile whenever the agent does not expect stakes to immediately jump upward, i.e. whenever $x_t = x^{**}(x_t, \pi_t)_0$, he undermines at the corresponding time-consistent rate $\beta_t = \beta^{**}(x_t, \pi_t)_0$. If the agent does expect stakes to immediately jump upward, then whenever $x_t < \overline{x}$ or $\pi_t < (K-1)/K$ he refrains from undermining in anticipation of the upward jump, and otherwise he undermines at full intensity.[20] The agent's undermining rule is depicted in Figure 4.

The optimality of the agent's strategy in every subgame is immediate given the incentive-compatibility of $\beta^{**}(y, p)$ whenever the principal's expected continuation stakes path is $x^{**}(y, p)$. Most of the work of the proof involves checking that the principal does not want to deviate from the desired stakes path, and in particular does not want to delay raising stakes in order to enjoy a period without undermining. It turns out that such deviations cannot be profitable, because the payoff from the equilibrium continuation when current stakes are y and the stakes lower bound does not bind is precisely y/r. Thus delaying the equilibrium continuation to collect flow payoffs at stakes y for a period of time with no undermining yields no more payoff than proceeding to raise stakes as planned.

[20]Technically, in the region where the principal immediately jumps stakes, any undermining policy is optimal for the agent in this region, as he expects to the remain in the region for a zero measure of time. And from the principal's perspective, it is optimal to increase stakes to 1 immediately no matter what the agent does if she deviates. The agent's strategy in this region is therefore not uniquely pinned down by the requirements of equilibrium. However, the undermining policy we have specified in our equilibrium is the unique one which remains optimal in a discrete-time approximation, in which the agent collects flow payoffs from undermining for a short period before stakes jump.

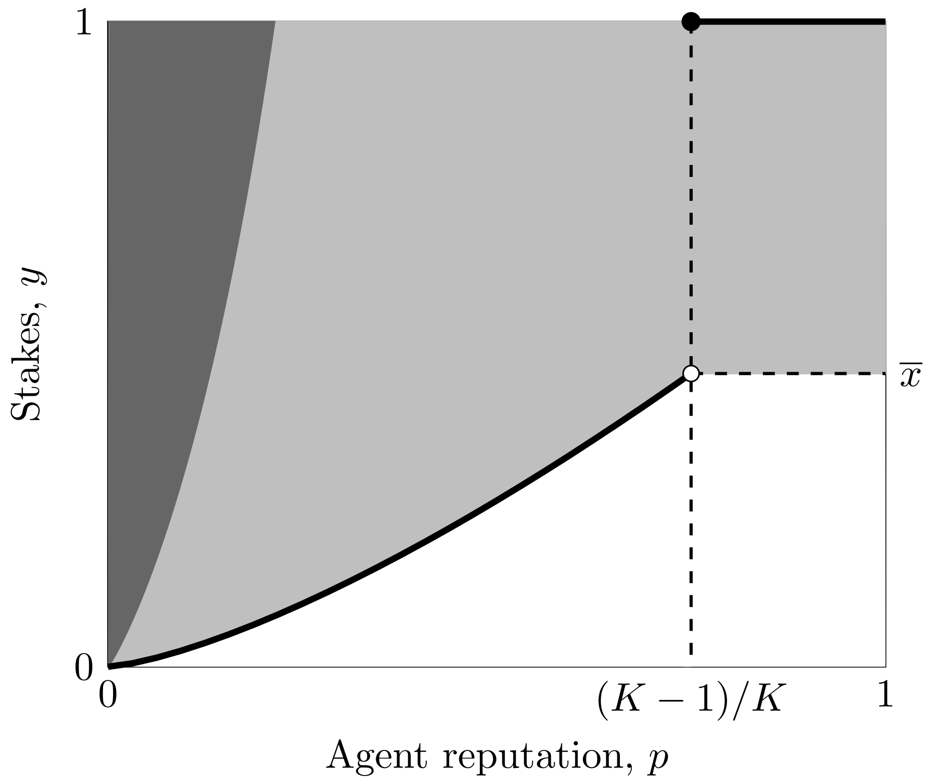


Figure 4: The agent's equilibrium undermining rule. In the darkly shaded and lightly shaded regions, the agent undermines at full intensity; in the darkly shaded region, he expects to be fired immediately. In the unshaded region, the agent refrains from undermining. Along the black curve, the agent undermines at interior intensity if $p < (K-1)/K$ and at full intensity if $p \geq (K-1)/K$.

4.3 Mixed-strategy equilibria

While analyzing the no-commitment outcome, we have so far restricted attention to equilibrium implementation of deterministic contracts. We now consider randomization of both stakes paths and undermining in equilibrium. The following lemma disposes of the possibility of randomized stakes, by showing that the stakes curve x^* is uniquely optimal in the commitment problem in the class of randomized policies.

Lemma 10. *If $\widetilde{x}$ is a (possibly random) stakes policy satisfying $\Pi(\widetilde{x}) = \Pi(x^*)$, then $\widetilde{x} = x^*$ a.s.*

This result generalizes the uniqueness claim of Proposition 1, which held only among deterministic stakes curves. As no randomized policy is optimal under commitment, the arguments in the proof of Proposition 2 imply that no random stakes policy can be sustained as the equilibrium path of any Bayes Nash equilibrium, for the principal could secure a strictly higher payoff by simply playing x^* instead.

On the other hand, randomization by the agent is possible in equilibrium. While the logic of Proposition 2 continues to guarantee that x^* is the unique equilibrium stakes path even when the agent randomizes, the path β^* is no longer the unique equilibrium path of undermining on the interval $[\underline{t}, \bar{t}]$ when randomization is allowed. Multiplicity arises because to sustain principal incentives along the equilibrium path, it is necessary and sufficient that principal posterior beliefs follow the path $(q'_t)_{t\geq 0}$ characterized in the proof of Lemma 9. With proper randomization, this can be achieved through a mixed undermining strategy. In particular, consider a mixed strategy under which the agent stops undermining at time $\underline{t}$, and then resumes with full intensity forever after at a random time $\tau^\beta \sim F(\cdot)$, where F is a distribution over $\mathbb{R}$ with support contained in the interval $[\underline{t}, \bar{t}]$. Since the agent is indifferent about undermining in this interval, any such mixed strategy is incentive-compatible. The next lemma establishes existence of a unique such F under which the sequence of the principal's posterior beliefs are precisely q'.

Lemma 11. *Suppose $\underline{t} < \bar{t}$. There exists a unique distribution function F^* with support contained in $[\underline{t}, \bar{t}]$ such that the principal's posterior beliefs induced by F^* satisfy $\pi_t = q'_t$ for every $t \in [\underline{t}, \underline{t}]$. The function F^* is strictly increasing on $[\underline{t}, \bar{t}]$, continuous except at $t = \underline{t}$, and satisfies $F^*(\underline{t}) = \beta^*_{\underline{t}}$ and $F^*(\bar{t}) = 1$.*

It follows that there exists a Bayes Nash equilibrium whose equilibrium path induces stakes path x^* and a distribution F^* over the time at which the agent begins undermining subsequent to time $\underline{t}$. (An agent strategy supporting such an equilibrium can be described by specifying that following any deviation by the principal, the agent restarts his strategy conditioning on current stakes and beliefs, and randomizes anew the next time a gradual undermining interval is reached.)[21] A multiplicity of mixed-strategy equilibria of this type can be constructed, differing in the details of the undermining path followed once the random time for undermining to begin has been reached. All such constructions have the property that the "effective" rate of undermining at each time t, i.e. the hazard rate of arrival of an observation of undermining normalized by the detection rate γ, is exactly the rate β^*_t specified in the pure-strategy equilibrium.

Mixed-strategy equilibria are notable because in a discrete-time approximation of our model, the unique equilibrium (in behavioral strategies) involves mixed strategies. To see this, note that in discrete time the Bellman equation governing the agent's optimal under-

[21] This equilibrium is also in spirit a perfect Bayesian equilibrium, although we do not attempt to formally define a notion of mixed-strategy perfect Bayesian equilibrium in our setting. Such notions are difficult to formalize in continuous time due to the general impossibility of specifying strategies featuring independent randomization of actions at all points in time, a property generally imposed in discrete-time definitions based on behavioral strategies.

mining strategy is

$$U_t = \sup_{\beta \in [0,1]} \left\{ (K\beta_t - 1)x_t^* + \exp(-(r + \gamma\beta)\Delta t) U_{t+\Delta t} \right\}.$$

Because the exponential function is strictly convex, the interior of the right-hand side is strictly convex in β whenever continuation values are positive. So any optimal undermining policy, whether deterministic or random, must have support only on $\{0, 1\}$. In particular, at times in the interval $[\underline{t}, \overline{t}]$, in equilibrium the agent must randomize over undermining at full intensity or not undermining at all to induce the proper sequence of ex post beliefs for the principal.

Despite the necessity of randomization in discrete time, our pure-strategy equilibrium is the closest in spirit to the limit of discrete-time equilibria in behavioral strategies. When the time step Δt is taken very small, paths of undermining in the time interval $[\underline{t}, \overline{t}]$ exhibit many rapid oscillations between undermining or not, closely approximating the flow payoffs of a deterministic interior undermining strategy.[22] The mixed strategy equilibria outlined above, by contrast, involve smooth undermining paths with correlation between undermining at different points in time. So our pure-strategy equilibrium is best considered as the unique limit outcome of discrete-time equilibria in behavioral strategies, while all mixed-strategy equilibria are the limit of discrete-time equilibria involving correlated randomization across periods.

5 Information design

In this section, we microfound the stakes process in our model as arising from dynamic provision of information about an underlying task-relevant state of the world. We also characterize an information policy inducing the optimal stakes curve x^*.

Suppose that the agent's task is to match a binary action $a_t \in \{L, R\}$ to a persistent binary state $\omega \in \{L, R\}$ at each time $t \in \mathbb{R}_+$.[23] The principal's ex post flow payoff at time t from task performance is

$$\pi(a_t, \omega) \equiv \mathbf{1}\{a_t = \omega\} - \mathbf{1}\{a_t \neq \omega\}.$$

[22] As is well-known, it is not possible to specify stochastic processes in continuous time which exhibit independence of values of the process at all points in time and also retain basic joint measurability properties. So the limit of behavioral strategies taken in the most literal sense is not well-defined in the continuous-time game.

[23] We explore an alternate specification with a continuous state space in Section 5.3, and extend the model to allow for an evolving binary state in Section 5.4.

Under this payoff specification, the principal-optimal task action given the agent's posterior belief $\mu_t = \Pr_t(\omega = R)$ about the state is $a_t^* = R$ if $\mu_t > 1/2$, and $a_t^* = L$ if $\mu_t < 1/2$. Since the task action is observable, optimal task performance (i.e. implementation of a^*) can be costlessly enforced by the threat of immediate termination, yielding an expected flow payoff $|2\mu_t - 1|$ to the principal. Undermining with intensity β_t inflicts expected flow damage $K\beta_t|2\mu_t - 1|$.

The principal and agent are initially uninformed and both parties assign probability $1/2$ that $\omega = R$. Upon hiring the agent, the principal becomes perfectly informed of ω while the agent receives a public signal $s \in \{L, R\}$ of ω which is correct with probability $\rho \geq 1/2$. This exogenous public signal is designed to capture several different pre-play channels of information potentially available to the agent. First, new hires might inevitably obtain *some* sense of the nature of their project through discussions with coworkers or initial training. Second, exogenous initial information arises naturally in a larger model in which an organization becomes aware it is being undermined only after employees have already been on the job for some time, and hence the screening process begins in the presence of a partially informed cohort.

The agent receives no further exogenous news about ω after observing the signal s. However, the principal has access to a disclosure technology allowing her to send noisy public signals of the state to the agent over time. This technology allows the principal to commit to signal structures inducing arbitrary posterior belief processes about the state. Formally, in line with the Bayesian persuasion literature, the principal may choose any $[0, 1]$-valued martingale process μ, where μ_t represents the agent's time-t posterior beliefs that $\omega = R$. The only restriction on μ is that $\mathbb{E}[\mu_0]$ be equal to the agent's posterior beliefs after receipt of the signal s. We will refer to any such process μ as an *information policy*.

5.1 Reduction to the abstract stakes model

As a first step toward determining an optimal information policy, we connect the principal's information design problem to the model with abstract stakes. Given an information policy μ, define the ex post precision process z by $z_t = |2\mu_t - 1|$ for all times. It is evident that z is closely related to an abstract stakes process in the baseline model, but the problems are not identical. On one hand, the fact that μ is a martingale places a new restriction on the induced stakes process; on the other hand, whereas the abstract stakes process in Section 2 was restricted to be nondecreasing, z may decrease whenever signal realizations move the agent's posterior belief closer to $1/2$. Despite these differences, we show that the information design problem may be reduced to that of designing a deterministic, monotone stakes curve, restoring equivalency with a restricted version of the abstract problem, in which stakes policies are restricted to be deterministic. (Recall that by Lemmas 2 and 10, such a restriction is without loss.)

Definition 5. *Given an information policy μ, the associated* disclosure path *x is the function defined by $x_t \equiv \mathbb{E}[z_t]$. An information policy is* deterministic *if $x_t = z_t$ at all times.*

The disclosure path traces the ex ante expected precision of the agent's beliefs over time. The fact that x must be generated by a martingale belief process places several key restrictions on its form. First, the martingality of μ implies that z is a submartingale by Jensen's inequality. The disclosure path x must therefore be an increasing function, reflecting the fact that on average the agent must become (weakly) more informed over time. Second, ex post precision can never exceed 1, so $x_t \leq 1$ at all times. Finally, upon receiving the initial signal s, the initial precision of the agent's beliefs prior to any information disclosure is $\phi \equiv |2\rho - 1|$. This places the lower bound $x \geq \phi$ on the average precision of the agent's beliefs all times. Lemma 12 establishes that the properties just outlined, plus a technical right-continuity condition, are all of the restrictions placed on a disclosure path by the requirement that it be generated by a martingale belief process.

Lemma 12. *A function x is the disclosure path for some information policy iff it is right-continuous, monotone increasing, and $[\phi, 1]$-valued. When the latter conditions hold, there is a unique deterministic information policy for which x is the disclosure path.*

In light of this result, the principal's problem reduces to design of an optimal disclosure path provided that it is without loss to focus on deterministic information policies.[24] And this result is immediate from the logic of Lemma 2, whose proof does not rely on the restriction to nondecreasing stakes processes and so may be applied to ex post precision processes.

The principal's problem thus can be solved in two steps. First, the principal optimizes over the space of disclosure paths, which coincides with the space of stakes curves considered in Section 3; the solution is precisely x^* as in Proposition 1. Second, she designs an information policy for which x^* is the disclosure path.

5.2 The optimal information policy

Lemma 12 and the logic of Lemma 2 guarantee the existence of a unique deterministic policy inducing disclosure path x^*.[25] However, a more explicit characterization is necessary

[24]Importantly, deterministic information policies still induce stochastic posterior belief processes — the precision of the agent's beliefs is non-random, but the likelihood of a particular state being the true one is a random variable.

[25]Lemma 12 leaves open the possibility of other, nondeterministic optimal information policies. Such nondeterministic policies can only be optimal for parameter values in the interior of the high stakes case,

to understand how optimal information disclosure should be implemented in practice. In this subsection we characterize an optimal information policy, and describe a signal process which induces the desired belief process.

The basic problem is to find a martingale belief process μ^* satisfying $|2\mu_t^* - 1| = x_t^*$ at all times. Equivalently, at all times $\mu_t^* \in \{(1-x_t^*)/2, (1+x_t^*)/2\}$, meaning that pathwise μ^* is composed of piecewise segments of the two disclosure envelopes $(1-x_t^*)/2$ and $(1+x_t^*)/2$. In light of this fact, the requirement that μ^* be a martingale implies that it must be a (time-inhomogeneous) Markov chain transitioning between the upper and lower disclosure envelopes, and uniquely pins down transition probabilities for this process. Thus the process μ^* implementing x^* is uniquely defined, and can be explicitly constructed by calculating the transition rates between the envelopes which make μ^* a martingale.

We will now describe a signal process to which the principal may commit in order to induce the posterior belief process μ^*. To begin, define the agent's *state conjecture* at a given time t to be the state he currently views as mostly likely; that is, his state conjecture is R if $\mu_t^* \geq 1/2$, and L otherwise. The principal influences the agent's beliefs by sending periodic recommendations about the correct state conjecture.[26] The signal structure is designed such that each time a recommendation is received, the agent changes his state conjecture. The overall arrival rate of recommendations is calibrated to induce the appropriate pace of disclosure, as follows.

First, at any time corresponding to a discontinuity of the disclosure path, the principal recommends a change in state conjecture with a probability which is higher when the agent's state conjecture is incorrect than when it is correct. In particular, when the agent reaches trusted status at time $\bar{t}$, the principal recommends a change in state conjecture if and only if the agent's conjecture is wrong, revealing the state.

Second, during the gradual escalation phase the principal recommends changes in the agent's state conjecture via a Poisson process whose intensity is higher when the agent's conjecture is currently wrong. This sort of Poisson signal process is sometimes referred to as an inconclusive contradictory news process.[27] Interestingly, while the intensity of signal arrival depends on the correctness of the agent's current conjecture, the ex ante expected intensity is constant throughout the gradual escalation phase. Thus the agent switches

where the principal can modify z during the quiet period using small mean-preserving spreads without changing the agent's behavior. Such policies do not carry new insights, so we ignore them.

[26] Equivalently, the principal sends periodic action recommendations.

[27] This news process has been featured, for instance, as one of the learning processes available to a decision-maker in the optimal learning model of Che and Mierendorff (2019). In that paper, the news process has fixed informativeness, as measured by the ratio of arrival rates in each state of the world. By contrast, our optimal disclosure process features an informativeness which rises over time. Such inhomogeneous contradictory news processes are considered, for instance, in the more general learning model of Zhong (2018).

conjectures at a constant average rate throughout this phase.[28] However, as the phase progresses the informativeness of the Poisson signals, as measured by the gap in the arrival rate when the agent's conjecture is wrong versus right, increases.

The following proposition formally establishes the claims made above. It also provides additional details of the signal process implementing the optimal belief process μ^*.

Proposition 4 (Optimal Information Policy)**.** *There exists a unique deterministic information policy μ^* implementing x^*. It is induced by the following signal process:*

- *If $\underline{t} > 0$, then on the time interval $[0, \underline{t}]$, no signals arrive,*
- *If $\underline{t} = 0$, then at time $t = 0$ a signal arrives with probability $\frac{(1+x_0^*)(x_0^*-\phi)}{2x_0^*(1-\phi)}$ if the agent's time $t = 0-$ state conjecture is incorrect, and with probability $\frac{(1-x_0^*)(x_0^*-\phi)}{2x_0^*(1+\phi)}$ otherwise,*
- *On the time inteval $(\underline{t}, \bar{t})$, a signal arrives with intensity $\overline{\lambda}_t \equiv \frac{1}{2}\left(r + \frac{\gamma}{K}\right)\frac{1+x_t^*}{1-x_t^*}$ if the agent's state conjecture is incorrect, and with intensity $\underline{\lambda}_t \equiv \frac{1}{2}\left(r + \frac{\gamma}{K}\right)\frac{1-x_t^*}{1+x_t^*}$ otherwise. Conditional on the agent's information, the expected intensity of signal arrival is $\frac{1}{2}\left(r + \frac{\gamma}{K}\right)$ at all times on this interval.*
- *At time $t = \bar{t}$, a signal arrives with probability 1 if the agent's time $t = \bar{t}-$ state conjecture is incorrect, and with probability 0 otherwise.*

Figure 5 shows a sample path of μ^*, the agent's posterior belief that $\omega = R$, in an optimal contract under the low stakes regime.

5.3 A continuous state space model

We have modeled uncertainty about the correct task action using a binary state space. This choice affords significant tractability, as the agent's posterior belief process is then one-dimensional. Richer state spaces would typically require a more complex analysis considering the optimal influence of several moments of the agent's beliefs over time. However, when preferences take a quadratic-loss form, expected flow payoffs depend only on the second moment of beliefs in general, and much larger state spaces can be accommodated within the existing analysis with minor changes. In this subsection we provide a brief overview of this extension. The main finding is that while, naturally, the details of the optimal information

[28] Our optimal information policy during the escalation phase bears some resemblance to the suspense-maximizing policy in Ely, Frankel, and Kamenica (2015), which also induces a deterministic belief precision path and features inconclusive contradictory news in the form of plot twists. However, suspense-maximization dictates a *decreasing* rate of plot twists, whereas our optimal policy features a constant arrival rate of contradictory news.

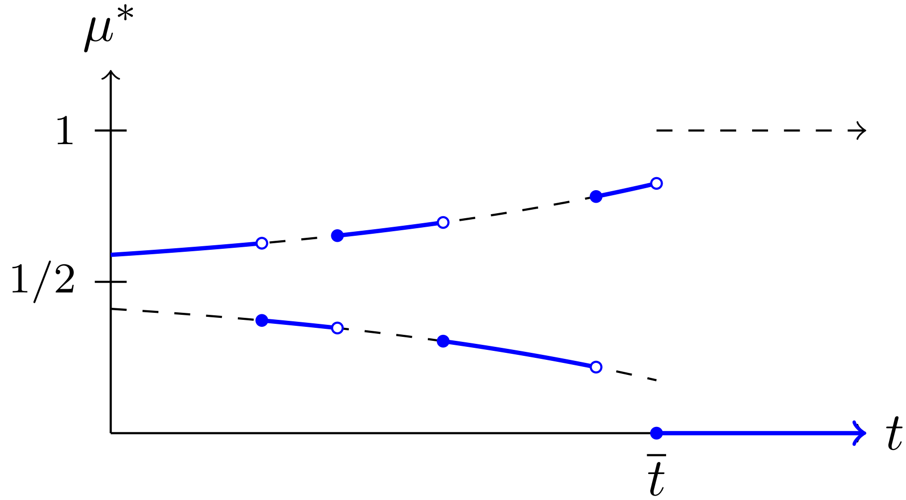


Figure 5: A sample path of the agent's posterior beliefs under the optimal information policy.

policy may change depending on the state space, the optimal disclosure path x^* is robust and continues to capture the optimal rate at which information should be disclosed.

Concretely, suppose that the state space is the real line $\Omega = \mathbb{R}$, with the agent initially possessing a diffuse prior over Ω. Upon becoming employed, the agent observes an exogenous public signal $s_i \sim N(\omega, \eta^2)$ for some $\eta > 0$, inducing time-zero posterior beliefs $\omega | s_i \sim N(s_i, \eta^2)$. (Since preferences will depend only on the second moment of beliefs, the form of the agent's prior and the exogenous signal are not very important and chosen mainly to simplify exposition.)

Principal payoffs depend on the agent's posterior beliefs as follows. At each moment in time the agent undertakes a (perfectly monitored) task action $a_t \in \mathbb{R}$, yielding a flow payoff to the principal of

$$\pi(a_t, \omega) = 1 - (a_t - \omega)^2 / C$$

for some constant $C \geq \eta^2$. (This inequality is without loss, as otherwise the principal would optimally immediately release information to avoid negative expected flow payoffs.) When the agent's posterior beliefs have mean μ_t, the optimal task action is $a_t = \mu_t$. If the variance of the agent's posterior beliefs is σ_t^2, then the induced expected flow payoffs to the principal from this choice of task action are

$$\pi(\sigma_t^2) = 1 - \sigma_t^2 / C.$$

We will take this expression to be the principal's flow payoff function given an induced agent belief process. As in the binary model, by undermining with intensity β_t the agent inflicts a flow loss of $K\beta_t\pi(\sigma_t^2)$ on the principal.

Now define the disclosure path $x_t \equiv \mathbb{E}[1-\sigma_t^2/C]$ to be the principal's ex ante time-t flow payoff under a given information policy. This process plays the same role as the x process defined in the binary case. The following lemma proves an analogue of Lemma 12 in the binary model, establishing a set of necessary conditions for disclosure paths which can be induced by some information policy.

Lemma 13. *For any information policy, x_t is right-continuous, monotone increasing, and $[\phi, 1]$-valued, where $\phi = 1 - \eta^2/C$.*

Suppose that any disclosure path satisfying the necessary conditions of Lemma 13 can be implemented by some deterministic information policy, i.e. with the posterior variance process σ^2 a deterministic function of time. Then for the same reasons as in Lemma 2 the principal might as well choose a deterministic information policy. So the problem reduces to the optimal design of x, with an objective function and constraints identical to the baseline model. The solution x^* is thus also identical. The final step is to show that x^* can in fact be implemented by some deterministic information policy.[29] The following lemma establishes this fact.

Lemma 14. *There exists a deterministic information policy whose disclosure path is x^*.*

The proof of the lemma constructs a particular signal process whose disclosure path is x^*. We will describe the process in the low-stakes case, with the other cases following along similar lines. The principal first releases a discrete normally distributed signal at time zero in the low-stakes case, with mean ω and variance calibrated to achieve the desired time-zero posterior variance. Then during the gradual escalation phase, the principal discloses information via a continuous signal process with drift ω obscured by a Brownian noise term whose variance declines as the phase progresses. Finally, at time $\bar{t}$ the principal reveals ω, eliminating all residual uncertainty.

This extension shows that the details of the optimal signal process will depend on the underlying state space, and in particular Brownian signals become a natural choice when the state space is continuous and preferences are linear in the agent's posterior variance. Nonetheless, the optimal disclosure path is robust to alternative specifications of the state space in which flow payoffs remain linear in a one-dimensional summary statistic of the agent's posterior beliefs.

[29]This implementation problem is similar to one faced in Ball (2018), in which the agent's posterior variance process is designed in a relaxed problem and afterward it is proven that there actually exists an information policy inducing the desired sequence of posterior variances. In that paper delayed reports of an exogenous Brownian state process serve as a suitable information policy. As no such process exists in our model, our proof proceeds along different lines.

5.4 An evolving task

We have so far assumed that the underlying state of the world is fixed, and therefore that the agent's information never becomes obsolete. However, in many environments a fluctuating state may be a more natural assumption, capturing environments in which the priorities of the organization change over time. These changes may be driven, for instance, by a changing competitive environment or the development of new product lines. In this subsection we explore how the optimal information policy is impacted by a fluctuating state. We find that whenever the optimal contract of the baseline model begins with a quiet period, the principal can profitably exploit a changing state by allowing the agent's information to deteriorate for a time. In all other cases, the changing state does not change the optimal disclosure path at all.

Suppose that instead of being held fixed, $\omega \in \{L, R\}$ evolves as a Markov chain which transitions between two states at rate $\lambda > 0$. This change affects the model by enlarging the set of admissible disclosure paths x. Recall that when ω was fixed, Lemma 12 required that any disclosure path be a monotone function. This is because for a fixed state, the precision of the agent's beliefs cannot drift down on average over time. However, when ω fluctuates, absent disclosures the precision of the agent's beliefs naturally deteriorates toward 0. In particular, if the agent's beliefs are μ_t at some time t, and no further information is disclosed, then the agent's beliefs evolve according to the ODE

$$\dot{\mu}_t = \lambda(1 - 2\mu_t).$$

Accordingly, the disclosure path $x_t = \mathbb{E}|2\mu_t - 1|$ decays at rate 2λ over any time interval when no information is disclosed.

The model can therefore be adapted to accommodate the changing state by requiring that x, rather than being a monotone function, satisfy the weaker condition that $e^{2\lambda t} x_t$ be monotone. In the low stakes case, when the lower bound constraint was non-binding for a fixed state, this relaxation of the problem does not change the optimal disclosure path at all. However, in the moderate and high stakes cases, the principal takes advantage of the changing state to initially degrade the precision of the agent's information over time. A consequence is that the optimal length of the gradual escalation phase lengthens and the agent attains trusted status later.[30] Figure 6 illustrates how the optimal contract in the moderate stakes case is impacted by a changing state.

[30]To establish these facts, note that extending the length of the gradual disclosure phase, and the corresponding final disclosure time, incurs the same marginal cost at the end of the contract for all λ, but provides increasing savings at the start of the contract due to lower average disclosure levels when λ is larger.

Note that the information policy implementing a given disclosure path changes somewhat under a fluctuating state. In particular, the agent must be kept abreast of changes in the state with at least some probability in order to maintain a constant disclosure level. So an agent who has reached trusted status will receive an immediate report whenever the state has changed. Further, during the escalation phase status updates serve a dual role — they both inform the agent in case he was previously off-track, and update him about changes in the state. Of course, the agent is always left uncertain as to the motivation for a particular status update!

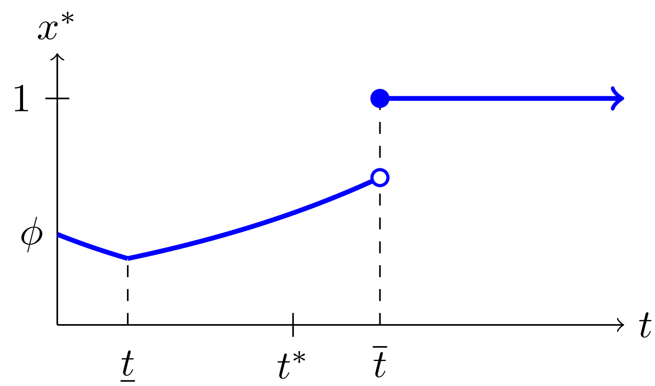


Figure 6: The optimal stakes curve in the moderate stakes case for an evolving state.

6 Comparative statics

In this section we discuss comparative statics for important features of the optimal contract with respect to changes in the key model parameters. For clarity of exposition we will focus on the moderate stakes case; the comparative statics for other cases are similar and are provided in Appendix G, along with proofs for all results reported in this section.

Table 1 below reports comparative statics for each model output (by column) with respect to each model input (by row). Here Π are the principal's expected profits under the optimal contract; $\underline{t}$ and $\overline{t}$ are the threshold times denoting the end of the quiet period and the time at which the agent reaches trusted status, respectively; $\Delta \equiv \overline{t} - \underline{t}$ is the length of the escalation phase; and $\overline{x}$ is the precision of the agent's beliefs just prior to reaching trusted status.

An increase in the damage K from undermining has a negative direct effect on the principal. This direct effect implies that the principal's payoff under the optimal contract unambiguously decreases with K: fixing a stakes curve, loyalty test or otherwise, and a disloyal agent strategy, the principal's expected flow payoff is made lower at each instant in time, and hence her ex ante payoff, after minimizing over disloyal agent strategies and

Table 1: Comparative statics

	$d\overline{x}/$	$d\Delta/$	$d\underline{t}/$	$d\overline{t}/$	$d\Pi/$
$d\gamma$	$+$	$\pm$	$-$	$-$	$+$
dq	0	0	$-$	$-$	$+$
dK	$+$	$+$	$+$	$+$	$-$
dr	$-$	$-$	$+$	$-$	$-$
$d\phi$	0	$-$	$+$	$-$	$-$

maximizing over stakes curves, must decrease. The impact of higher K on the stakes curve is a bit more subtle, due to effects on the agent's incentives. Despite that the disloyal agent enjoys inflicting larger damage on the principal, the optimal contract must provide *stronger* incentives to make the disloyal agent willing to undermine. With higher K, the agent obtains both more value from undermining today but also a higher continuation value from being employed and able to undermine in the future. Importantly, the former value is proportional to K while the latter value is proportional to $K - 1$, which grows faster in relative terms, and hence the net effect on the disloyal agent is a stronger incentive to feign loyalty. In response, fixing the final jump time, the principal must front load stakes by starting the escalation period sooner and more slowly increasing stakes thereafter, which is reflected in both a smaller final jump and a longer escalation period. In a final optimization, the principal then delays both the start and end of the escalation period, as this delay now has a larger (helpful) effect of reducing stakes in the early part of the contract, when the agent is likely to be disloyal, and a smaller (harmful) effect of reducing stakes in the later part of the contract.

An increase in the quality γ of the monitoring technology directly benefits the principal through faster detection of disloyal types. Moreover, as with K, this benefit occurs independently of the stakes curve and disloyal agent's strategy, which allows us to conclude without further calculation that the principal's payoff increases with γ. In addition to the direct benefit, the principal is also indirectly affected in two ways through the disloyal agent's incentives. While better monitoring makes the disloyal agent more likely to be detected if he undermines today and this increases the cost of undermining, it also increases the agent's effective discount rate — remaining employed in the future becomes less valuable — which decreases the cost of undermining. The first effect dominates just prior to the final jump in stakes, and to keep the policy a loyalty test, the final jump must be smaller. But the second effect dominates during the escalation period (when much of the agent's continuation value depends on future increases in stakes) and the agent's IC constraint is loosened, reflected in the growth condition $\dot{x}_t \leq (r + \gamma/K)x_t$. After optimizing the start time of this escalation period, we find that the period both starts and ends sooner with higher γ, but its length may increase or decrease.

Using similar arguments to those above about the principal's flow payoff, we see that higher proportion of loyal agents q improves the principal's payoff. It has no direct effect on the agent's incentives, and thus $\overline{x}$ is unchanged, as is the length of the escalation period. But with lower risk of facing a disloyal agent, the principal finds it optimal to start this period (and therefore reach the final jump) sooner. A higher initial stakes level ϕ hurts the principal, and it does so by simply restricting the space of feasible stakes curves. As with q, it has no direct effect on the agent's incentives, and thus $\overline{x}$ is unchanged, but since $x_0 \geq \phi$ is a binding constraint in the moderate stakes case, the escalation period is shortened and its start is delayed. In response, the principal shifts the escalation period slightly earlier in time (the cost of increasing stakes prior to t^* affects a smaller interval of times), resulting in an earlier final jump with higher ϕ.

An increase in the discount rate has two opposing effects on the principal's payoff.[31] Supposing that a stakes curve remains a loyalty test as the discount rate changes, the increase in the discount rate hurts the principal as it increases the weight she places on the flow payoffs earned in the early part of the contract, which are negative since the agent's reputation is low. Unlike the arguments about the direct effect on the principal's flow payoffs of changes in γ, K and q, which apply independent of the stakes curve and disloyal agent strategy, this claim uses the fact that the stakes curve is a loyalty test, and hence we must consider the effect of an increase in the discount rate on the set of loyalty tests over which the principal optimizes. The accompanying increase in the agent's discount rate makes the disloyal agent more willing to undermine today at the risk of being detected, forgoing continuation utility. Consequently, the disloyal agent's IC constraint is relaxed, which enlarges the set of loyalty tests. Although this effect runs opposite the first one, we show analytically that the first effect dominates — the principal is always made worse off when the discount rate increases.

7 Conclusion

This paper studies dynamic screening of disloyal agents in a variable-stakes principal-agent problem, when higher stakes both facilitate efficient performance of the agent's job and enhance the harm of leaks or sabotage. With some probability the agent opposes the interests of the principal, and can undermine her in a gradual and imperfectly detectable manner. Ideally, the principal would like to hold stakes to a very low level until she can sufficiently

[31]For comparative statics with respect to the discount rate, we study the normalized payoff $r\Pi$ to eliminate the mechanical diminuation of flow payoffs at all times due to a rise in the discount rate.

ascertain loyalty by detecting and tracing the source of any undermining. However, such a scheme is self-defeating, as it gives a disloyal agent strong incentives to defer undermining to avoid detection until stakes rise. An optimal stakes curve must therefore escalate stakes sufficiently gradually that the disloyal agent undermines and reveals himself when stakes are low.

We show that under commitment, the optimal stakes path features a quiet period of fixed stakes, a gradual escalation phase with a smooth rise in stakes, and a deterministic date at which the agent is deemed trusted and stakes jump discretely to their maximum level. This jump in stakes is a signature feature of an optimal stakes process in our setting, which does not arise in existing models of variable-stakes relationships with discrete betrayal actions. We also study the model without commitment, and find that the optimal stakes curve is the unique path of stakes supportable without commitment, with the agent's undermining strategy uniquely pinned down in the no-commitment game and exhibiting a non-monotonic intensity of undermining. We further connect our model to the information-sensitive environments motivating it by microfounding stakes as the outcome of an information design problem. We show that the optimal stakes path in the microfounded model can be implemented with an inconclusive contradictory news process: the agent is gradually fed unreliable reports of a task-relevant state until a deterministic date at which the principal reveals the true state.

Several assumptions afford our model significant tractability and present opportunities for future work. First, our model assumes that agents are either fully aligned or totally opposed to the principal's long-run interests. This restriction allows us to focus on the principal's problem of screening out disloyal agents, by abstracting from the incentive-alignment issues often studied in dynamic contracting models. However, in some situations it may be realistic to suppose that even disloyal agents can be "converted" into loyal ones through sufficient monetary compensation, team-building exercises, or other incentives. We leave open as an interesting direction for future work the broader question of when a principal might prefer to screen versus convert agents who face temptations to undermine her.

Our model also assumes that the nature of the underlying task is exogenous, and that the agent's loyalty is independent of the task. Relaxing these assumptions would allow one to study how a firm or organizations should proceed when facing severe challenges: while some alternatives might be mild and uncontroversial, other more drastic course corrections might carry a risk of alienating a subset of agents. The question of how firms should choose among alternatives with different implications for agent loyalty is an important one for future study.

Finally, we have assumed that the principal can only detect undermining as it takes place, as in the interception of telecommunication. It seems plausible that in many real-world settings, past undermining may eventually be detected with a lag, exposing the agent to larger cumulative risk the earlier she undermines. Extending our model to capture such a technology would introduce interesting additional forces shaping the disloyal agent's optimal undermining policy, and thus the resulting optimal stakes curve. Such a model would also be technically challenging to analyze, as it would involve a privately observed persistent state variable controlled by the agent.

References

Albuquerque, Rui and Hugo A. Hopenhayn (2004). Optimal Lending Contracts and Firm Dynamics. In: *The Review of Economic Studies* 71.2, pp. 285–315.

Atakan, Alp, Levent Koçkesen, and Elif Kubilay (2019). Starting Small to Communicate. Working paper.

Au, Pak Hung (2015). Dynamic information disclosure. In: *RAND Journal of Economics* 46.4, pp. 791–823.

Ball, Ian (2018). Dynamic Information Provision: Rewarding the Past and Guiding the Future. Working paper.

Basu, Pathikrit (2018). Dynamic Bayesian Persuasion with a Privately Informed Receiver. Working paper.

Capasso, Vincenzo and David Bakstein (2005). An introduction to continuous-time stochastic processes. In: *Birkhäuser Boston*, p. 1.

Che, Yeon-Koo and Konrad Mierendorff (2019). Optimal Dynamic Allocation of Attention. In: *American Economic Review* 109, pp. 2993–3029.

Clancy, Tom (1987). *Patriot Games*. G. P. Putnam's Sons.

Cranley, Ellen (2018). A senior White House aide is reportedly sharing fake stories with staffers to identify who's leaking to the press. In: *Business Insider*. URL: https://www.businessinsider.com/trump-leaks-leakers-2018-6.

Ely, Jeffrey (2017). Beeps. In: *The American Economic Review* 107.1, pp. 31–53.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica (2015). Suspense and surprise. In: *Journal of Political Economy* 123.1, pp. 215–260.

Ely, Jeffrey and Martin Szydlowski (2018). Moving the Goalposts. In: *Journal of Political Economy*, forthcoming.

Engel, Joel (1994). *Gene Roddenberry: The Myth and the Man Behind "Star Trek"*. Hyperion.

Fudenberg, Drew and Luis Rayo (2018). Training and Effort Dynamics in Apprenticeship. Working paper.

Guo, Yingni and Eran Shmaya (2019). The interval structure of optimal disclosure. In: *Econometrica* 87.2, pp. 653–675.

Haynes, John Earl and Harvey Klehr (2006). *Early Cold War Spies: The Espionage Trials That Shaped American Politics*. Cambridge University Press.

Isaac, Mike (2016). Facebook 'Trending' List Skewed by Individual Judgment, Not Institutional Bias. In: *The New York Times*. URL: https://www.nytimes.com/2016/05/21/technology/facebook-trending-list-skewed-by-individual-judgment-not-institutional-bias.

Kamenica, Emir and Matthew Gentzkow (2011). Bayesian persuasion. In: *The American Economic Review* 101.6, pp. 2590–2615.

Kinery, Emma (2017). As Jeff Sessions Announces Leak Crackdown, Here are 12 Major Leaks of the Trump Era. In: *USA Today*. URL: https://www.usatoday.com/story/news/politics/onpolitics/2017/08/04/here-12-major-leaks-donald-trump-era/539674001.

Kolodny, Lora (2018). Elon Musk Emails Employees about 'Extensive and Damaging Sabotage' by Employee. In: *CNBC*. URL: https://www.cnbc.com/2018/06/18/elon-musk-email-employee-conducted-extensive-and-damaging-sabotage.html.

Kolotilin, Anton (2018). Optimal information disclosure: A linear programming approach. In: *Theoretical Economics* 13.2, pp. 607–635.

Kolotilin, Anton et al. (2017). Persuasion of a privately informed receiver. In: *Econometrica* 85.6, pp. 1949–1964.

Kreps, David (2018). Starting Small to Screen for Mr. Good Bob. Working paper.

Liptser, Robert S and Albert N Shiryaev (2013). *Statistics of random processes: I. General theory*. Vol. 5. Springer Science & Business Media.

Orlov, Dmitry, Andrzej Skrzypacz, and Pavel Zryumov (2018). Persuading the Principal To Wait. In: *Journal of Political Economy*, forthcoming.

Peterson, Hayley (2018). Amazon plants fake packages in delivery trucks as part of an undercover ploy to 'trap' drivers who are stealing. In: *Business Insider*. URL: https://www.businessinsider.com/amazon-sets-traps-for-drivers-2018-9.

Rayo, Luis and Luis Garicano (2017). Relational Knowledge Transfers. In: *The American Economic Review* 107.9, pp. 2695–2730.

Rothschild, Neal (2018). The 16 Juiciest White House Leaks of 2018. In: *Axios*. URL: https://www.axios.com/white-house-leaks-president-trump-2018-92899c04-0dfe-4b64-bc42-18fcdb09e28f.

Savage, Charlie (2017). Chelsea Manning Leaves Prison, Closing an Extraordinary Leak Case. In: *The New York Times*. URL: https://www.nytimes.com/2017/05/16/us/politics/chelsea-manning-leaves-prison.
Sobel, Joel (1985). A Theory of Credibility. In: *The Review of Economic Studies* 52.4, pp. 557–573.
Thomas, Jonathan and Timothy Worrall (1994). Foreign Direct Investment and the Risk of Expropriation. In: *Review of Economic Studies* 61.1, pp. 81–108.
Thomas, Owen (2009). Tesla CEO in Digital Witch Hunt. In: *Gawker Media*. URL: https://gawker.com/5164035/tesla-ceo-in-digital-witch-hunt.
Thompson, Nicholas and Fred Vogelstein (2018). Inside the Two Years That Shook Facebook - and the World. In: *Wired*. URL: https://www.wired.com/story/inside-facebook-mark-zuckerberg-2-years-of-hell.
Watson, Joel (1999). Starting small and renegotiation. In: *Journal of Economic Theory* 85.1, pp. 52–90.
— (2002). Starting small and commitment. In: *Games and Economic Behavior* 38.1, pp. 176–199.
Wright, Peter and Paul Greengrass (1987). *Spycatcher: The Candid Autobiography of a Senior Intelligence Officer*. Heinemann.
Zhong, Weijie (2018). Optimal Dynamic Information Acquisition. Working paper.
Ziegler, Charles A and David Jacobson (1995). *Spying without spies: Origins of America's secret nuclear surveillance system*. Greenwood Publishing Group.

A Equilibrium definitions

In this appendix we develop a notion of strategies in the no-commitment game between the principal and disloyal agent tailored to our setting, and use them to define Bayes Nash equilbrium and perfect Bayesian equilibrium.[32] We leverage the fact that only one player (the principal) acts observably, to sidestep technical issues arising in more general settings.

The game takes place on a state space $\Omega = \mathbb{X} \times \mathbb{B} \times \mathbb{T} \times \Omega^P \times \Omega^A$, where:

- $\mathbb{X}$ is the set of increasing, $[\phi, 1]$-valued, increasing functions,
- $\mathbb{B}$ is the set of càdlàg, $[0, 1]$-valued functions,
- $\mathbb{T}$ is set of elements $\tau \in (\mathbb{R}_+ \cup \{\infty\})^{\mathbb{N}}$ such that:

[32] For simplicity, we do not explicitly model the type of the agent, and simply define payoffs appropriately to reflect the probabilistic presence of the disloyal agent.

1. $\tau_k \leq \tau_{k+1}$ for all k, with the inequality strict whenever $\tau_k < \infty$,
2. $\lim_{k\to\infty} \tau_k = \infty$,

- Ω^P and Ω^A are arbitrary state spaces.

Each element $\omega \in \Omega$ is a vector $\omega = (x, \beta, \tau, \omega^P, \omega^A)$, where x is the path of stakes; β is the path of undermining; τ is the vector of times at which undermining is observed; and ω^P and ω^A are realizations of the principal and agent's randomization devices. Note that τ may have entries of ∞, which correspond to histories with only a finite number of observations of undermining. Also, at most a finite number of observations of undermining are allowed by any finite time. $\mathbb{X}$, $\mathbb{B}$, and $\mathbb{T}$ are endowed with the Borel Σ-algebras generated by sup norm. Ω^P and Ω^A are endowed with arbitrary Σ-algebras and probability measures μ^P and μ^A.

Let X and B be the coordinate processes $X(\omega) = \omega(x)$ and $B(\omega) = \omega(\beta)$, and define the counting process N by

$$N_t(\omega) = \sum_{k=1}^{\infty} \mathbf{1}\{t \geq \omega(\tau)_k\}.$$

Define $\mathbb{F}^P$ to be the filtration induced by N and the Σ-algebra on Ω^P, and $\mathbb{F}^A$ to be the filtration induced by X, N, and the Σ-algebra on Ω^A.

Definition A.1. *A strategy profile is a triple* (χ, Λ, ζ), *where:*

- χ *is an* $\mathbb{F}^P$*-adapted, càdlàg,* $[\phi, 1]$*-valued, increasing stochastic process,*
- Λ *is an* $\mathbb{F}^P$*-stopping time,*
- ζ *is an* $\mathbb{F}^A$*-adapted, càdlàg,* $[0, 1]$*-valued stochastic process.*

(χ, Λ) are chosen by the principal, and represent her control of stakes and termination, respectively. ζ is chosen by the agent and represents his undermining policy, depending on the history of stakes, observed undermining, and his randomization device. Intuitively, (χ, ζ) jointly determine a probability measure over stakes paths and undermining. However, as both condition on the path of observed undermining, whose distribution in turn depends on ζ, the mapping must be carefully defined and shown to be coherent. We next formally a mapping from (χ, ζ) to probability measures on Ω.

Fix a stakes-undermining pair (χ, ζ). $\mathbb{F}^P$-adaptedness of χ implies existence of a unique measurable function $\widetilde{X}^\chi : \mathbb{T} \times \Omega^P \to \mathbb{X}$ such that $\widetilde{X}^\chi(\omega(\tau), \omega^P) = \chi(\omega)$ for every ω. Similarly, $\mathbb{F}^A$-adaptedness of ζ implies existence of a unique measurable function $\widetilde{B}^\zeta : \mathbb{X} \times \mathbb{T} \times \Omega^A \to \mathbb{B}$ such that $\widetilde{B}^\zeta(\omega(x), \omega(\tau), \omega^A) = \zeta(\omega)$ for every ω. Also, for each $\tau \in \mathbb{T}$ and $k = 0, 1, \ldots$, define τ^k to be the vector such that $\tau^k_j = \tau_j$ for $j \leq k$, and $\tau^k_j = \infty$ for $j > k$.

Then given a pair (χ, ζ), a regular conditional probability $\nu^\tau(\chi, \zeta)$ over $\mathbb{T} \times \Omega^P \times \Omega^A$ given (ω^P, ω^A) is characterized via the conditional distributions

$$\begin{aligned}
&\Pr(\tau_k \leq \Delta t + \tau_{k-1} \mid \tau_1, ..., \tau_{k-1}, \omega^P, \omega^A) \\
&= \begin{cases} 0, & \Delta t < 0, \\ 1 - \exp\left(-\gamma \int_0^{\Delta t} \widetilde{B}^\zeta_{s+\tau_{k-1}}(\widetilde{X}^\chi(\tau^{k-1}, \omega^P), \tau^{k-1}, \omega^A)\, ds\right), & \Delta t \geq 0. \end{cases}
\end{aligned}$$

whenever $\tau_{k-1} < \infty$, and $\Pr(\tau_k = \infty) = 1$ whenever $\tau_{k-1} = \infty$. These conditional distributions define the probability of the kth jump using the expected cumulative hazard rate under ζ, assuming that no further jumps arrive after time τ_{k-1} and stakes evolve according to χ.

For this construction to induce a well-defined conditional probability $\nu^\tau(\chi, \zeta)$ over $\mathbb{T} \times \Omega^P \times \Omega^A$, it must place probability 1 on $\mathbb{T} \times \Omega^P \times \Omega^A$ for every (ω^P, ω^A). Clearly $\nu^\tau(\chi, \zeta)$ puts probability 1 on vectors τ which are strictly increasing so long as elements are finite. Further, for each $t < \infty$,

$$\Pr(\tau_k \leq t \mid \tau_1 \leq t, ..., \tau_{k-1} \leq t, \omega^P, \omega^A) \leq 1 - \exp(-\gamma t),$$

and so

$$\nu^\tau(\chi, \zeta)(\{\lim_{k\to\infty} \tau_k \leq t\}, \omega^P, \omega^A) \leq \prod_{k=1}^{\infty} (1 - \exp(-\gamma t)) = 0.$$

In other words, $\nu^\tau(\chi, \zeta)\left(\{\lim_{k\to\infty} \tau_k = \infty\}, \omega^P, \omega^A\right) = 1$, so $\nu^\tau(\chi, \zeta)$ places probability 1 on $\mathbb{T} \times \Omega^P \times \Omega^A$, as required.

The regular conditional probability $\nu^\tau(\chi, \zeta)$ induces a probability measure $\mu^\tau(\chi, \zeta)$ over $\mathbb{T} \times \Omega^P \times \Omega^A$ in the standard way:

$$\mu^\tau(\chi, \zeta)(F) = \int \nu^\tau(\chi, \zeta)(F, \omega^P, \omega^A)\, d\mu^P(\omega^P)\, d\mu^A(\omega^A).$$

for measurable subsets $F \subset \mathbb{T} \times \Omega^P \times \Omega^A$. Finally, for every measurable subset $E \subset \Omega$, let

$$\overline{E}^{(\chi,\zeta)} = \{(\tau, \omega^P, \omega^A) \in \mathbb{T} \times \Omega^P \times \Omega^A \ : \ (\widetilde{X}^\chi(\tau, \omega^P), \widetilde{B}^\zeta(\widetilde{X}^\chi(\tau, \omega^P), \tau, \omega^A), \tau, \omega^P, \omega^A) \in E\}.$$

(Note that $\overline{E}^{(\chi,\zeta)}$ is a measurable subset of $\mathbb{T} \times \Omega^P \times \Omega^A$ whenever E is a measurable subset of Ω, given that compositions and vectors of measurable functions are measurable.) Then define $\mu(\chi, \zeta)(E) = \mu^\tau(\chi, \zeta)(\overline{E}^{(\chi,\zeta)})$. Note in particular that under $\mu(\chi, \zeta)$, $X(\omega) = \chi(\omega)$ and $B(\omega) = \zeta(\omega)$ a.s.

Using the measure $\mu(\chi, \zeta)$, payoffs to the principal and disloyal agent from a given strategy profile may be defined as

$$\Pi(\chi, \Lambda, \zeta) = q\,\mathbb{E}^{(\chi,0)}\left[\int_0^\Lambda e^{-rt} X_t(1 - KB_t)\,dt\right] + (1-q)\,\mathbb{E}^{(\chi,\zeta)}\left[\int_0^\Lambda e^{-rt} X_t(1 - KB_t)\,dt\right]$$

and

$$U^B(\chi, \Lambda, \zeta) = \mathbb{E}^{(\chi,\zeta)}\left[\int_0^\Lambda e^{-rt} X_t(KB_t - 1)\,dt\right]$$

and where $\mathbb{E}^{(\chi,\zeta)}$ represents expectations under the measure $\mu(\chi, \zeta)$. Note that the principal's payoff is a weighted average of two expected payoffs, one assuming the agent plays the strategy $\chi' = 0$, in which case $B = 0$ with probability 1, and one assuming the agent plays the strategy $\chi' = \chi$. These terms capture the payoff contributions from the loyal and disloyal agent, respectively.

Definition A.2. *A* Bayes Nash equilibrium *is a strategy profile* (χ, Λ, ζ) *such that*

$$\Pi(\chi, \Lambda, \zeta) \geq \Pi(\chi', \Lambda', \zeta) \textit{ for all } (\chi', \Lambda')$$

and

$$U^B(\chi, \Lambda, \zeta) \geq U^B(\chi, \Lambda, \zeta') \textit{ for all } \zeta'.$$

We now turn to the definition of a perfect Bayesian equilibrium. As we do not need randomization for our results involving subgame-perfect implementation, we will define PBE only in pure strategies.

Define $\widetilde{\mathbb{F}}^P$ to be the filtration induced by X (but not Ω^P) and $\widetilde{\mathbb{F}}^A$ to be the filtration induced by X and N (but not Ω^P). Given any $\widetilde{\mathbb{F}}^A$-adapted agent strategy ζ, define a belief process π^ζ by

$$\pi_t^\zeta = \begin{cases} \left(1 + \frac{1-q}{q}\exp\left(-\gamma \int_0^t \zeta_s\,ds\right)\right)^{-1}, & N_t = 0, \\ 0, & N_t > 0. \end{cases}$$

Note that π^ζ is also $\widetilde{\mathbb{F}}^A$-adapted, and accords with Bayes' rule whenever it applies. (Bayes' rule does not apply if N increments at a time when $\zeta_t = 0$. We resolve the resulting belief indeterminacy after such histories in favor of sure beliefs that the agent is disloyal.) The strategy ζ induces a natural family of continuation strategies at each time t for the continuation games beginning at that time. This family of strategies is indexed by the history $h_t = (x_s, n_s)_{s \leq t}$, and the continuation game at time t with history $(x_s, n_s)_{s \leq t}$ is isomorphic to the original game with initial stakes x_t and initial beliefs $\Pr(\theta = G) = \pi_t$. We will denote the induced strategy in each continuation game by $\zeta^{\geq t}(h_t)$.

Definition A.3. *A (pure-strategy)* perfect Bayesian equilibrium *is a strategy profile* (χ, Λ, ζ) *such that:*

- χ *is* $\widetilde{\mathbb{F}}^P$*-adapted,* Λ *is an* $\widetilde{\mathbb{F}}^P$*-stopping time, and* ζ *is* $\widetilde{\mathbb{F}}^A$*-adapted,*
- (χ, Λ, ζ) *is a Bayes Nash equilibrium,*
- *For every time* $t \geq 0$ *and history* $h_t = (x_s, n_s)_{s \leq t}$*, there exists a principal strategy* $(\chi^{\geq t}(h_t), \Lambda^{\geq t}(h_t))$ *such that* $(\chi^{\geq t}(h_t), \Lambda^{\geq t}(h_t), \zeta^{\geq t}(h_t))$ *is a pure-strategy Bayes Nash equilibrium of the game with initial stakes* X_t *and initial beliefs* π_t^ζ*.*

Our notion of perfect Bayesian equilibrium economizes on notation by not explicitly recording the principal's off-path continuation play as part of the equilibrium strategy profile. Nevertheless, the equilibrium definition requires that the agent's off-path continuation play must be part of a Nash equilibrium of the continuation game. Note that the requirement of Bayes Nash equilibrium implies this condition for all on-path histories, in which case the on-path continuation play constitutes an equilibrium of the continuation game. The extra requirement is that some other equilibrium strategy profile justify the agent's play following any off-path history.[33]

B Robustness

In this Appendix we study the robustness of our results to relaxation of several key assumptions in our baseline model.

B.1 Replacing the agent

Our model assumes that the agent is irreplaceable, so that if the agent is discovered to be disloyal and fired the principal loses all future payoffs from performance of the agent's task. This is a reasonable assumption for settings in which the agent is a specialist and the task payoff is relative to some performance benchmark for a replacement. However, in other settings it may be reasonable to believe that the principal can replace a disloyal agent with someone of similar skills, perhaps at a cost. Our analysis can be adapted to accommodate this scenario with only minor modifications. We outline the required changes, and the associated implications for the optimal stakes curve, here. We will consider two

[33]This definition also does not formally require that the *same* continuation equilibrium is played at each history along the revised equilibrium path following a deviation. However, it is easy to see that along the revised equilibrium path, continuation play of the revised equilibrium strategy continues to satisfy the equilibrium requirement for successive subgames.

alternative specifications: in one, the principal's post-termination payoff does not enter the disloyal agent's; in the other, it does. We find that in the first specification, the principal's solution is exactly the same as in the baseline model, while in the second, the optimal stakes curve is quantitatively similar and differs mainly in a tighter constraint on the rate of stakes growth during the escalation phase.

First suppose the agent cares only about firm outcomes while employed. Then the firm's continuation value following termination does not matter for his incentives, and in particular the game is no longer zero-sum. However, this turns out not to impact the optimality of offering a single contract. For it continues to be the case that any contract can be modified to create a loyalty test without changing the payoff to the disloyal agent while (weakly) improving the payoff to the principal. So without loss the principal offers only loyalty tests. Also, with a positive post-termination payoff the principal optimally recommends undermining at all times under any loyalty test, so all contracts optimally chosen by the principal must discover the disloyal agent at the same rate. Thus the contribution to the principal's profits from the post-termination option is the same for all contracts. The game is then effectively zero-sum with a disloyal agent, implying optimality of a single contract.

In light of the previous discussion, the ability to collect a continuation payoff $\Pi > 0$ following termination changes the principal's problem only by adding on a constant term $(1-q)\frac{\gamma}{r+\gamma}\Pi$ to the principal's payoff, which critically does not depend on the choice of x. So the contract design problem is completely unchanged by the possibility of replacement, and the optimal information policy will be the same. If Π is determined endogenously, say as the value of the original problem minus a search cost, it can be calculated by solving a simple fixed-point problem. Say V is the expected profit to the principal of the current agent under an optimal contract, and that new workers are drawn from the same distribution as the original, after incurring a search cost $\psi > 0$. (We will suppose ψ is sufficiently small that the principal optimally replaces the agent.) Then Π satisfies

$$\Pi = V + (1-q)\frac{\gamma}{r+\gamma}(\Pi - \psi),$$

which can be solved for Π to obtain an expression for the lifetime value of the position.

Alternatively, suppose the agent's preferences are over the lifetime value of the position, regardless of who holds it. In this case the game continues to be zero-sum, and the principal continues to optimally offer only a single contract which is a loyalty test. However, now the disloyal agent's incentive constraints for undermining tighten, reflecting the added value of holding off replacement through feigning loyalty. Specifically, the continuation value to the disloyal agent at a given time t is now

$$U_t = \int_t^\infty e^{-(r+\gamma)t}[(K-1)x_t - \gamma\Pi]\,dt,$$

which accounts for the arrival of the continuation profit Π to the principal upon discovery and termination. Letting $U_t^\dagger = U_t + \gamma\Pi/(\gamma + r)$ represent the gross continuation payoff to the agent before accounting for the continuation value to the principal, the principal's continuation payoff function is, up to a constant not depending on x,

$$-\left(1 - \frac{K}{K-1}q\right)U_0^\dagger + \frac{q\gamma}{K-1}\int_0^\infty e^{-rt}U_t^\dagger\,dt,$$

analogously to the baseline model. And the constraint $x \geq \phi$ may be written

$$\dot{U}_t^\dagger \leq (r+\gamma)U_t^\dagger - (K-1)\phi,$$

again as in the baseline model. However, appropriately modifying the proof of Lemma 5 yields the IC constraint

$$\dot{U}_t^\dagger \leq (r+\gamma/K)U_t^\dagger - \frac{r\gamma}{r+\gamma}\frac{K-1}{K}\Pi,$$

which is tighter than the analogous constraint in the baseline model. The bounds on $U^\dagger$ may also tighten — while its upper bound is $\overline{U}^\dagger = (K-1)/(r+\gamma)$ as before, it must satisfy the potentially tighter lower bound

$$\underline{U}^\dagger = \max\left\{\phi\overline{U}^\dagger, \left(\frac{\gamma}{r+\gamma} - \frac{\gamma/K}{r+\gamma/K}\right)\Pi\right\}.$$

This bound ensures that the upper bound on $\dot{U}^\dagger$ is non-negative in both constraints for all admissible levels of U. The new second term in the bound reflects the fact that when the principal's continuation payoff is high, and the current level of stakes is low, the agent may prefer to remain loyal rather than trigger turnover even absent future stakes growth.

These changes to the contracting problem yield several changes to the optimal contract. First, the growth rate of x during the escalation phase is slower than without replacement. Second, the size of the jump in stakes when the agent becomes trusted becomes smaller. Third, the optimal initial stakes level will rise, meaning the size of the quiet period shrinks and more parameter settings will fall into the low-stakes case. Otherwise, the optimal contract is qualitatively similar to the no-replacement case, and is derived similarly.

Note that all changes in the contract are driven by changes in the incentives for disloyal agents to feign loyalty. Continuation payoffs do not directly enter the principal's objective for the current agent for any given stakes policy. And replacement does not change the undermining policies which are optimally induced, as even without replacement the principal already optimally induces as much undermining as possible. So while post-termination options potentially boost the principal's lifetime profits, for a given agent they only restrict the set of loyalty tests which can be offered.

B.2 Transfers

Suppose that in addition to regulating the rate of information release, the principal could commit to a schedule of transfers to the agent. Could it use this additional instrument to improve contract performance? To study this possibility, we consider an augmentation of the model with transfers in which the agent cares about money as well as the firm's payoffs. Our main finding is that while the principal may utilize up-front transfers as part of a menu of contracts designed to screen out the disloyal agent, the contract intended for the loyal agent will be qualitatively similar to the optimal contract in the baseline model.

Specifically, if the firm's expected profits (including any spending on transfers) are Π, and the expected net present value of transfers made to the agent are T, then we model the loyal agent's total payoff as

$$U^G = (1-\delta)\Pi + \delta T$$

while the disloyal agent's payoff is

$$U^B = -(1-\delta)\Pi + \delta T$$

for some $\delta \in (0,1)$. This specification is intended to capture a world in which agents are homogeneous in their relative interest in money versus firm outcomes, but differ in their preferred firm outcomes in the same way as in the baseline model.[34] In particular, the limiting case $\delta = 0$ can be thought of as nesting the baseline model, as transfers cannot be used to usefully align incentives and the optimal contract is as in the baseline.

When $\delta > 0$, cash transfers can be used to improve outcomes by screening out the disloyal agent. In particular, suppose that under a particular operating contract involving no transfers, the disloyal agent receives a payoff of U^B. Note that a buyout contract providing an immediate transfer of $(1-\delta)U^B$ followed by termination yields the same payoff U^B to

[34] Moreover, this specification captures situations in which disloyal agents have a higher marginal value for money than the principal. This is because the disloyal agent enjoys a transfer from the principal through two channels: the direct benefit of the transfer, as well as the cost to the principal of the transfer.

the disloyal agent, since $\Pi = -(1-\delta)U^B$ for this contract. So the disloyal agent is willing to take the buyout over the operating contract. And offering this contract reduces the total cost to the principal in case the disloyal agent is present, from U^B to $(1-\delta)U^B$. So when transfers are allowed, a menu of contracts offering both an operating contract and a buyout will generally be optimal.

To better understand the optimal form of the operating contract in this environment, assume for simplicity that $\delta \leq 1/2$. This assumption ensures that the loyal agent would rather take the operating contract than a buyout.[35] In this case the optimal menu of contracts involves a buyout calibrated as in the previous paragraph, with the operating contract designed to solve

$$\sup_{x\in\mathbb{X}} \int_0^\infty e^{-rt} x_t (q - (1-q)(1-\delta)(K-1)e^{-\gamma t})\, dt \quad \text{ s.t. } \quad \dot{U}_t \leq (r + \gamma/K)U_t,$$

where now the disloyal agent's contribution to the principal's payoff is depressed by a factor $1-\delta$. Using integration by parts as in Lemma 6, the objective may be rewritten

$$-\left((1-\delta)(1-q) - \frac{1}{K-1}q\right) U_0 + \frac{q\gamma}{K-1}\int_0^\infty e^{-rt} U_t\, dt.$$

Notice that conditional on U_0, the objective is exactly as in the problem without transfers, and has the same solution. Equivalently, conditional on x_0 the path of x is unchanged by the presence of transfers. The downweighting of the disloyal agent's payoff by factor $1-\delta$ merely makes a delivery of a given utility U_0 to the bad agent less costly, increasing the optimal size of U_0. So transfers change the operating contract by boosting the optimal initial grant of information, and otherwise leave the optimal path of information release unchanged.

C Technical results

In this appendix we prove several technical lemmas aiding in the characterization of incentive-compatible undermining policies. For the lemmas in this subsection, fix a (deterministic)

[35] It can be shown that this result also holds for all δ if $K \leq 2$. When both assumptions fail, then the IC constraint of the loyal agent potentially binds. To ensure that the loyal agent chooses the operating contract, large payments in the distant future can be used to boost the value of the contract to the loyal agent without making it more valuable to the disloyal agent (who discounts future payments more heavily due to termination).

The optimal contract then either involves an operating contract with eventual payments and a buyout, or else a single operating contract with no buyout. In particular screening contracts may not be employed if both K and q are large, so that offering a buyout rarely saves money on disloyal agents but forces the principal to make large payments when the loyal agent is present.

stakes curve x.

Given any undermining policy β, define U^β to be the disloyal agent's continuation value process under β, conditional on the agent remaining employed. By definition,

$$U_t^\beta = \int_t^\infty \exp\left(-r(s-t) - \gamma \int_t^s \beta_u \, du\right)(K\beta_s - 1)x_s \, ds$$

for all t. This function is absolutely continuous with a.e. derivative

$$\frac{dU^\beta}{dt} = (r + \gamma\beta_t)U_t^\beta - (K\beta_t - 1)x_t.$$

Note that the rhs is bounded below by $f(U_t^\beta, t)$, where

$$f(u,t) \equiv \min\{(r+\gamma)u - (K-1)x_t, ru + x_t\}.$$

Lemma C.1. *Suppose $g(u,t)$ is a function which is strictly increasing in its first argument. Fix $T \in \mathbb{R}_+$, and suppose there exist two absolutely continuous functions u_1 and u_2 on $[0,T]$ such that $u_1(T) \geq u_2(T)$ while $u_1'(t) = g(u_1(t),t)$ and $u_2'(t) \geq g(u_2(t),t)$ on $[0,T]$ a.e. Then:*

1. *$u_1 \geq u_2$.*
2. *If in addition $u_1(T) = u_2(T)$ and $u_2'(t) = g(u_2(t),t)$ on $[0,T]$ a.e., then $u_1 = u_2$.*

Proof. Define $\Delta(t) \equiv u_2(t) - u_1(t)$, and suppose by way of contradiction that $\Delta(t_0) > 0$ for some $t_0 \in [0,T)$. Let $t_1 \equiv \inf\{t \geq t_0 \ : \ \Delta(t_0) \leq 0\}$. Given continuity of Δ, it must be that $t_1 > t_0$. Further, $t_1 \leq T$ given $u_1(T) \geq u_2(T)$. And by continuity $\Delta(t_1) = 0$. But also by the fundamental theorem of calculus

$$\Delta(t_1) = \Delta(t_0) + \int_{t_0}^{t_1} \Delta'(t) \, dt.$$

Now, given that $\Delta(t) > 0$ on (t_0, t_1), it must be that

$$\Delta'(t) = u_2'(t) - u_1'(t) \geq g(u_2(t),t) - g(u_1(t),t) > 0$$

a.e. on (t_0,t_1) given that g is strictly increasing in its first argument. Hence from the previous identity $\Delta(t_1) > \Delta(t_0)$, a contradiction of $\Delta(t_1) = 0$ and $\Delta(t_0) > 0$. So it must be that $\Delta \leq 0$, i.e. $u_2 \leq u_1$.

Now, suppose further that $u_1(T) = u_2(T)$ and $u_2'(t) = g(u_2(t),t)$ on $[0,T]$ a.e. Trivially $u_2(T) \geq u_1(T)$ and $u_1'(t) \geq g(u_1(t),t)$ on $[0,T]$ a.e., so reversing the roles of u_1 and u_2 in the proof of the previous part establishes that $u_1 \leq u_2$. Hence $u_1 = u_2$. $\square$

Lemma C.2. *Given any $T \in \mathbb{R}_+$ and $\overline{u} \in \mathbb{R}$, there exists a unique absolutely continuous function u such that $u(T) = \overline{u}$ and $u'(t) = f(u(t), t)$ on $[0, T]$ a.e.*

Proof. Suppose first that x is a simple function; that is, x takes one of at most a finite number of values. Given that x is monotone, this means it is constant except at a finite set of jump points $D = \{t_1, ..., t_n\}$, where $0 < t_1 < ... < t_n < T$. In this case $f(u, t)$ is uniformly Lipschitz continuous in u and is continuous in t, except on the set D. Further, f satisfies the bound $|f(u, t)| \leq (K - 1) + (r + \gamma)|u|$. Then by the Picard-Lindelhöf theorem there exists a unique solution to the ODE between each t_k and t_{k+1} for arbitrary terminal condition at t_{k+1}. Let u^0 be the solution to the ODE on $[t_n, T]$ with terminal condition $u^0(T) = \overline{u}$, and construct a sequence of functions u^k inductively on each interval $[t_{n-k}, t_{n-k+1}]$ by taking $u^k(t_{n-k+1}) = u^{k-1}(t_{n-k+1})$ to be the terminal condition for u^k. Then the function u defined by letting $u(t) = u^k(t)$ for $t \in [t_{n-k}, t_{n-k+1}]$ for $k = 0, ..., n$, with $t_0 = 0$ and $t_{n+1} = T$, yields an absolutely continuous function satisfying the ODE everywhere except on the set of jump points D.

Now consider an arbitrary x. As x is non-negative, monotone, and bounded, there exists a increasing sequence of monotone simple functions x^n such that $x^n \uparrow x$ uniformly on $[0, T]$. Further, there exists another decreasing sequence of monotone simple functions $\widetilde{x}^m$ such that $\widetilde{x}^m \downarrow x$ uniformly on $[0, T]$. For each n, let u^n be the unique solution to the ODE $u'(t) = f^n(u, t)$ with terminal condition $u(T) = \overline{u}$, where

$$f^n(u, t) \equiv \min\{(r + \gamma)u - (K - 1)\widetilde{x}^n_t, ru + x^n_t\}.$$

(The arguments used for the case of x simple ensure such a solution exists.) By construction $f^n(u, t)$ is increasing in n for fixed (u, t), hence $du^m/dt \geq f^n(u^m(t), t)$ for every $m > n$. Also observe that each f^n is strictly increasing in its first argument, hence Lemma C.1 establishes that the sequence u^n is pointwise decreasing. Then the pointwise limit of u^n as $n \to \infty$ is well-defined (though possibly infinite), a function we will denote u^∞. Further, $f^n(u, t) \leq ru + 1$, hence by Grönwall's inequality $u^n(t) \geq -1/r + (\overline{u} + 1/r) \exp(r(t - T))$. Thus the u^n are uniformly bounded below on $[0, T]$, meaning u^∞ is finite-valued.

Now define an absolutely continuous function u^* on $[0, T]$ pointwise by setting

$$u^*(t) = \overline{u} - \int_t^T f(u^\infty(s), s)\, ds.$$

Our goal is to show that u^* is the desired solution to the ODE. Once that is established, Lemma C.1 ensures uniqueness given that f is strictly increasing in its first argument.

Use the fundamental theorem of calculus to write

$$u^n(t) = \overline{u} - \int_t^T \frac{du^n}{ds}\, ds = \overline{u} - \int_t^T f^n(u^n(s), s)\, ds,$$

and take $n \to \infty$ on both sides. Note that $|f^n(u^n(s), s)| \leq (K-1) + (r+\gamma)|u^n(s)|$, and recall that the sequence u^n is pointwise decreasing and each u^n satisfies the lower bound $u^n(t) \geq -1/r + (\overline{u} + 1/r)\exp(r(t-T))$. Hence a uniform lower bound on the $u^n(t)$ is $\min\{-1/r, \overline{u}\}$. So $|u^n|$ may be bounded above by $|u^n(s)| \leq \max\{1/r, |\overline{u}|, |u^1(s)|\}$. As u^1 is continuous and therefore bounded on $[0, T]$, this bound implies that $(K-1) + (r+\gamma)\max\{1/r, |\overline{u}|, |u^1(s)|\}$ is an integrable dominating function for the sequence $|f^n(u^n(s), s)|$. The dominated convergence theorem then yields

$$u^\infty(t) = \overline{u} - \int_t^T \lim_{n\to\infty} f^n(u^n(s), s)\, ds = \overline{u} - \int_t^T f(u^\infty(s), s)\, ds = u^*(t).$$

So u^∞ and u^* are the same function. Meanwhile, differentiating the definition of u^* yields

$$\frac{du^*}{dt} = f(u^\infty(t), t) = f(u^*(t), t).$$

Then as $u^*(T) = \overline{u}$, u^* is the desired solution to the ODE. □

Lemma C.3. *An undermining policy β^* maximizes the disloyal agent's payoff under x iff $\frac{dU^{\beta^*}}{dt} = f(U_t^{\beta^*}, t)$ a.e.*

Proof. Suppose first that β^* is an undermining policy whose continuation value process U^{β^*} satisfies $\frac{dU^{\beta^*}}{dt} = f(U_t^{\beta^*}, t)$ a.e. Let $x_\infty \equiv \lim_{t\to\infty} x_t$, which must exist given that x is an increasing function. We first claim that $\limsup_{t\to\infty} U_t^{\beta^*} = (K-1)x_\infty/(r+\gamma)$. First, given that $x_t \leq x_\infty$ for all time it must also be that $U_t^{\beta^*} \leq (K-1)x_\infty/(r+\gamma)$, as the latter is the maximum achievable payoff when $x_t = x_\infty$ for all time. Suppose $\overline{U}_\infty \equiv \limsup_{t\to\infty} U_t^{\beta^*} < (K-1)x_\infty/(r+\gamma)$. Then $\limsup_{t\to\infty} f(U_t^{\beta^*}, t) \leq (r+\gamma)\overline{U}_\infty - (K-1)x_\infty < 0$. This means that for sufficiently large t, dU^{β^*}/dt is negative and bounded away from zero, meaning U^{β^*} eventually becomes negative. This is impossible, so it must be that $\overline{U}_\infty = (K-1)x_\infty/(r+\gamma)$.

Now fix an arbitrary undermining policy β with continuation value process U^β. Suppose first that at some time T, $U_T^{\beta^*} \geq U_T^\beta$. Then by Lemma C.1, $U_t^{\beta^*} \geq U_t^\beta$ for all $t < T$, and in particular the lifetime value to the disloyal agent of β^* is at least as high as that of β. So consider the remaining possibility that $U_t^{\beta^*} < U_t^\beta$ for all t. Let $\Delta U \equiv U_0^\beta - U_0^{\beta^*} > 0$. As $f(u, t)$ is strictly increasing in u and $dU^\beta/dt \geq f(U_t^\beta, t)$ for all time, it must therefore be the case that $dU^{\beta}/dt > dU^{\beta^*}/dt$ a.e. Thus

$$U_t^{\beta} = U_0^{\beta} + \int_0^t \frac{dU^{\beta}}{ds}\, ds > U_0^{\beta} + \int_0^t \frac{dU^{\beta^*}}{ds}\, ds = U_t^{\beta^*} + \Delta U$$

for all t. But then

$$\limsup_{t\to\infty} U_t^{\beta} \geq \overline{U}_{\infty} + \Delta U = (K-1)x_{\infty}/(r+\gamma) + \Delta U,$$

and as $\Delta U > 0$ it must therefore be that $U_t^{\beta} > (K-1)x_{\infty}/(r+\gamma)$ for some t sufficiently large. This contradicts the bound $U_t^{\beta} \leq (K-1)x_{\infty}/(r+\gamma)$ for all t, hence it cannot be that $U_t^{\beta^*} < U_t^{\beta}$ for all time. This proves that β^* is an optimal undermining policy for the disloyal agent.

In the other direction, fix an undermining policy β such that $\frac{dU^{\beta}}{dt} > f(U_t^{\beta}, t)$ on some positive-measure set of times. Recall that

$$\frac{dU^{\beta}}{dt} = (r - \gamma\beta_t)U_t^{\beta} - (K\beta_t - 1)x_t$$

a.e. Then as U^{β} is continuous and β, x,and $f(u,\cdot)$ are right-continuous, it must be that if $\frac{dU^{\beta}}{dt} > f(U_t^{\beta}, t)$ at some $t = t_0$, the inequality also holds a.e. on some interval $[t_0, t_1)$ with $t_1 > t_0$. Note that trivially $\frac{dU^{\beta}}{dt} = f(U_t^{\beta}, t)$ whenever $Kx_t = \gamma U_t^{\beta}$, so for a.e. $t \in [t_0, t_1)$ either $Kx_t > \gamma U_t^{\beta}$ or else $Kx_t < \gamma U_t^{\beta}$. Then by right-continuity of x and U_t^{β}, there exists a subinterval $[t_0', t_1') \subset [t_0, t_1)$ with $t_1' > t_0'$ such that either $Kx_t > \gamma U_t^{\beta}$ everywhere on this interval, or else $Kx_t < \gamma U_t^{\beta}$ everywhere on this interval.

Suppose first that $Kx_t > \gamma U_t^{\beta}$ everywhere on $[t_0', t_1')$. Then define an undermining profile β^* by setting $\beta_t^* = \beta_t$ for $t \notin [t_0', t_1')$, and $\beta_t^* = 1$ for $t \in [t_0', t_1')$. Note that this policy is càdlàg by construction. By Ito's lemma,

$$U_0^{\beta} = \int_0^T \exp\left(-rt - \gamma \int_0^t \beta_s^*\, ds\right) \left((r + \gamma\beta_t^*)U_t^{\beta} - \frac{dU^{\beta}}{dt}\right) ds + \exp\left(-rT - \gamma \int_0^T \beta_t^*\, dt\right) U_T^{\beta},$$

and taking $T \to \infty$ and noting that U^{β} is bounded yields

$$U_0^{\beta} = \int_0^{\infty} \exp\left(-rt - \gamma \int_0^t \beta_s^*\, ds\right) \left((r + \gamma\beta_t^*)U_t^{\beta} - \frac{dU^{\beta}}{dt}\right) dt.$$

Using the fact that $\beta^*_t = \beta_t$ for $t \notin [t'_0, t'_1)$, this expression may be equivalently written

$$\begin{aligned}
U_0^\beta = & \int_{[0,t'_0)\cup[t'_1,\infty)} \exp\left(-rt - \gamma \int_0^t \beta^*_s \, ds\right) \left((r+\gamma\beta_t)U_t^\beta - \frac{dU^\beta}{dt}\right) dt \\
& + \int_{t'_0}^{t'_1} \exp\left(-rt - \gamma \int_0^t \beta^*_s \, ds\right) \left((r+\gamma\beta^*_t)U_t^\beta - \frac{dU^\beta}{dt}\right) dt \\
= & \int_{[0,t'_0)\cup[t'_1,\infty)} \exp\left(-rt - \gamma \int_0^t \beta^*_s \, ds\right) (K\beta_t - 1) \, dt \\
& + \int_{t'_0}^{t'_1} \exp\left(-rt - \gamma \int_0^t \beta^*_s \, ds\right) \left((r+\gamma\beta^*_t)U_t^\beta - \frac{dU^\beta}{dt}\right) dt \\
= & \int_{[0,t'_0)\cup[t'_1,\infty)} \exp\left(-rt - \gamma \int_0^t \beta^*_s \, ds\right) (K\beta^*_t - 1) \, dt \\
& + \int_{t'_0}^{t'_1} \exp\left(-rt - \gamma \int_0^t \beta^*_s \, ds\right) \left((r+\gamma\beta^*_t)U_t^\beta - \frac{dU^\beta}{dt}\right) dt.
\end{aligned}$$

Now, by construction $dU^\beta/dt > f(U_t^\beta, t)$ for a.e. $t \in [t'_0, t'_1)$. So

$$\begin{aligned}
U_0^\beta < & \int_{[0,t'_0)\cup[t'_1,\infty)} \exp\left(-rt - \gamma \int_0^t \beta^*_s \, ds\right) (K\beta^*_t - 1) \, dt \\
& + \int_{t'_0}^{t'_1} \exp\left(-rt - \gamma \int_0^t \beta^*_s \, ds\right) \left((r+\gamma\beta^*_t)U_t^\beta - f(U_t^\beta, t)\right) dt.
\end{aligned}$$

Further,

$$f(U_t^\beta, t) = (r+\gamma)U_t^\beta - (K-1)x_t = (r+\gamma\beta^*_t)U_t^\beta - (K\beta^*_t - 1)x_t$$

for $t \in [t'_0, t'_1)$ given that $Kx_t > \gamma U_t^\beta$ and $\beta^*_t = 1$ on this interval. Thus

$$U_0^\beta < \int_0^\infty \exp\left(-rt - \gamma \int_0^t \beta^*_s \, ds\right) (K\beta^*_t - 1) \, dt = U_0^{\beta^*}.$$

Thus β^* is a strict improvement on β.

If instead $Kx_t < \gamma U_t^\beta$ on $[t'_0, t'_1)$, then defining β^* by setting $\beta^*_t = 0$ on $[t'_0, t'_1)$ and carrying through the same argument yields an identical result. Thus in all cases, there exists a càdlàg policy β^* which is a strict improvement on β, completing the proof. □

D Proofs for Section 3

D.1 Proof of Lemma 1

First note that given any menu of contracts, it is optimal for an agent to accept some contract. For the agent can always secure a value at least equal to their outside option of 0 by choosing their myopically preferred task action at each instant.

Suppose without loss that the disloyal agent prefers contract 1 to contract 2. If the loyal agent also prefers contract 1, then the principal's payoff from offering contract 1 alone is the same as from offering both contracts. On the other hand, suppose the loyal agent prefers contract 2. In this case if the principal offers contract 2 alone, her payoff when the agent is loyal is unchanged. Meanwhile by assumption the disloyal agent's payoff decreases when accepting contract 1 in lieu of contract 2. As the disloyal agent's preferences are in direct opposition to the principal's, the principal's payoff when the agent is disloyal must (weakly) increase by eliminating contract 1.

D.2 Proof of Lemma 2

Define a (deterministic) stakes policy x' by $x'_t \equiv \mathbb{E}[x_t]$. Note that the loyal agent's payoff under x is

$$\mathbb{E}\int_0^\infty e^{-rt} x_t\, dt = \int_0^\infty e^{-rt} x'_t\, dt,$$

so his payoff is the same under x' and x. As for the disloyal agent, his payoff under x and any undermining policy β which is a deterministic function of time is

$$\begin{aligned}
&\mathbb{E}\int_0^\infty \exp\left(-rt - \gamma\int_0^t \beta_s\, ds\right)(K\beta_t - 1)x_t\, dt \\
=& \int_0^\infty \exp\left(-rt - \gamma\int_0^t \beta_s\, ds\right)(K\beta_t - 1)\mathbb{E}[x_t]\, dt \\
=& \int_0^\infty \exp\left(-rt - \gamma\int_0^t \beta_s\, ds\right)(K\beta_t - 1)x'_t\, dt,
\end{aligned}$$

hence is identical to his payoff choosing β under x'. And as x'_t is a deterministic function, the disloyal agent's maximum possible payoff under x' can be achieved by an undermining policy which is deterministic in time. It follows that the disloyal agent's maximum payoff under x across *all* (not necessarily deterministic) undermining policies is at least as large as his maximum payoff under x'. Therefore the principal's payoff under x must be weakly lower than under x'.

D.3 Proof of Lemma 3

Fix x which is not a loyalty test. If x is not deterministic, pass to the deterministic policy x' defined by $x'_t = \mathbb{E}[x_t]$ for all time. First suppose that x' is a loyalty test. In this case it must be that the principal's payoff is strictly improved by passing to x'. This is because her expected profits under x and x' are the same if the disloyal agent uses undermining rule $\beta = 1$ (note that in this case expectations may be passed through integrals in the payoff function to reach the ex post stakes level); but if x is not a loyalty test then any optimal rule for the disloyal agent yields principal profits strictly lower than under $\beta = 1$, and thus $\Pi(x) < \Pi(x')$. The result of the lemma is immediate in this case. It is therefore sufficient to prove the lemma for deterministic stakes policies, as for stochastic policies the principal can first pass to the deterministic version, and then pass from there to a payoff-superior loyalty test if the deterministic version is not itself a loyalty test.

If x is a deterministic stakes policy, then the agent always has a deterministic optimal undermining policy. Fix any such policy β^*, and let U^* be its associated continuation utility path. Let $T \equiv \inf\{t \ : \ Kx_t \leq \gamma U_t^*\}$. If $T = \infty$, then $dU^*/dt = (r+\gamma)U_t^* - (K-1)x_t$ a.e. by Lemma C.3, and as also

$$\frac{dU^*}{dt} = (r + \gamma\beta_t^*)U_t^* - (K\beta_t^* - 1)x_t$$

a.e., $Kx_t > \gamma U_t^*$ for all time implies that $\beta^* = 1$ a.e. The optimality of such a policy contradicts the assumption that x is not a loyalty test. So $T < \infty$. And by right-continuity of x and continuity of U^*, it must be that $Kx_T \leq \gamma U_T^*$. Hence $U_T^* \geq \frac{Kx_T}{\gamma} \geq \frac{K\phi}{\gamma}$.

Now, define a new function $\widetilde{U}$ by

$$\widetilde{U}_t \equiv \begin{cases} U_t^*, & t < T \\ \min\{U_T^* \exp((r + \gamma/K)(t - T)), (K-1)/(r+\gamma)\}, & t \geq T \end{cases}.$$

Further define a function $\widetilde{x}$ by

$$\widetilde{x}_t \equiv \frac{1}{K-1}\left((r+\gamma)\widetilde{U}_t - \dot{\widetilde{U}}_t\right),$$

with $\dot{\widetilde{U}}$ interpreted as a right-derivative when the derivative is discontinuous.

We first claim that $\widetilde{x}$ is a valid disclosure path. Right-continuity is by construction, so it remains to show monotonicity and $[\phi, 1]$-valuedness. We proceed by explicitly computing the function. For $t < T$ it is simply

$$\widetilde{x}_t \equiv \frac{1}{K-1}\left((r+\gamma)U_t^* - \dot{U}_t^*\right),$$

and by definition $\dot{U}_t^* = (r+\gamma)U_t^* - (K-1)x_t$ for $t < T$, meaning $\widetilde{x}_t = x_t$. So $\widetilde{x}$ is monotone and $[\phi, 1]$-valued on $[0, T)$. Meanwhile for $t \geq T$ the function takes the form

$$\widetilde{x}_t = \begin{cases} \frac{\gamma U_T^*}{K}\exp\left((r+\gamma/K)(t-T)\right), & T \leq t < T' \\ 1, & t \geq T' \end{cases},$$

where T' solves $U_T^* \exp((r+\gamma/K)(t-T)) = (K-1)/(\gamma+r)$. This expression is monotone everywhere, including at the jump time T', where $\widetilde{x}_{T'-} = \frac{K-1}{K}\frac{\gamma}{\gamma+r} < 1$. And $\widetilde{x}_T = \frac{\gamma U_T^*}{K}$, which is at least ϕ given the earlier observation that $U_T^* \geq K\phi/\gamma$. Thus given that it attains a maximum of 1, the function is $[\phi, 1]$-valued on $[T, \infty)$. The last remaining check is that $\widetilde{x}$ is monotone at the jump time T. Recall that $\widetilde{x}_{T-} = x_T$ while $\widetilde{x}_T = \frac{\gamma U_T^*}{K} \geq x_T$, so monotonicity holds at T as well, proving $\widetilde{x}$ is a valid disclosure path.

Next we claim that $\widetilde{U}$ is the continuation value process of the policy $\beta = 1$ under $\widetilde{x}$, and that $\beta = 1$ is an optimal undermining policy. The first part follows from the definition of $\widetilde{x}$, which can be re-written as the total differential

$$\frac{d}{dt}\left(\exp(-(r+\gamma)t)\widetilde{U}_t\right) = -\exp(-(r+\gamma)t)(K-1)\widetilde{x}_t.$$

Integrating both sides from t to ∞ and using the fact that $\lim_{t\to\infty} \exp(-(r+\gamma)t)\widetilde{U}_t = 0$ then yields

$$\widetilde{U}_t = \int_t^\infty \exp(-(r+\gamma)(s-t))(K-1)\widetilde{x}_s\, ds,$$

as desired.

As for the optimality of $\beta = 1$, this follows from Lemma C.3 if we can show that $d\widetilde{U}/dt = f(\widetilde{U}_t, t)$ a.e., when f is defined using $\widetilde{x}$ as the underlying stakes curve. Given that $\widetilde{U}_t = U_t^*$ and $\widetilde{x}_t = x_t^*$ for $t < T$, and U_t^* is an optimal continuation utility process under x^*, Lemma C.3 ensures the ODE is satisfied for $t < T$. Meanwhile for $t > T$, the definition of $\widetilde{x}$ ensures that

$$(r+\gamma)\widetilde{U}_t - (K-1)\widetilde{x}_t = \frac{d\widetilde{U}}{dt}.$$

So we're done if we can show that $(r+\gamma)\widetilde{U}_t - (K-1)\widetilde{x}_t \leq r\widetilde{U}_t + \widetilde{x}_t$, i.e. $\widetilde{U}_t \leq \frac{K\widetilde{x}_t}{\gamma}$. Using the definition of $\widetilde{U}$ and the explicit form for $\widetilde{x}$ constructed earlier, we can read off that $\widetilde{U}_t = \frac{K\widetilde{x}_t}{\gamma}$ for $t < T$, while $\widetilde{U}_t = \frac{K-1}{\gamma+r} < \frac{K}{\gamma} = \frac{K\widetilde{x}_t}{\gamma}$ for $t \geq T$. So indeed $d\widetilde{U}/dt = f(\widetilde{U}_t, t)$ a.e., and $\beta = 1$

is an optimal undermining policy under $\widetilde{x}$.

The proposition will be proven if we can show that the principal obtains strictly higher expected profits under $\widetilde{x}$ than under x. We just saw that $\widetilde{U}_0$ is the disloyal agent's payoff under $\widetilde{x}$, and by construction $\widetilde{U}_0 = U_0^*$, where U_0^* is the disloyal agent's payoff under x. So it suffices to show that the loyal agent's payoff is higher under $\widetilde{x}$ than under x.

The loyal agent's payoff under x is just

$$V = \int_0^\infty e^{-rt} x_t \, dt.$$

Let $U_t^1 \equiv \int_t^\infty e^{-(r+\gamma)(s-t)} x_s \, ds$ be the disloyal agent's continuation utility process when he undermines at all times. Then $\frac{d}{dt}(e^{-(r+\gamma)t} U_t^1) = -e^{-(r+\gamma)t} x_t$, and using integration by parts V may be rewritten

$$V = U_0^1 + \gamma \int_0^\infty e^{\gamma t} U_t^1 \, dt.$$

Similarly, the loyal agent's payoff under $\widetilde{x}$ may be written

$$\widetilde{V} = \widetilde{U}_0 + \gamma \int_0^\infty e^{\gamma t} \widetilde{U}_t \, dt,$$

where recall $\widetilde{U}$ is the disloyal agent's continuation utility process from undermining at all times. So we're done if we can show that $\widetilde{U} \geq U^1$ and $\widetilde{U}_0 > U_0^1$.

We first establish that $\widetilde{U} \geq U^*$. In light of Lemma C.3, whenever $Kx_t \geq \gamma U_t^*$ it must be that

$$\frac{dU^*}{dt} = (r+\gamma)U_t^* - (K-1)x_t \leq (r + \gamma/K)U_t^*,$$

and whenever $Kx_t < \gamma U_t^*$, it must be that

$$\frac{dU^*}{dt} = rU_t^* + x_t \leq (r + \gamma/K)U_t^*.$$

Hence $dU^*/dt \leq (r + \gamma/K)U_t^*$ a.e. Then by Grönwall's inequality,

$$U_s^* \leq U_t^* \exp((r + \gamma/K)(s-t))$$

for any t and $s > t$.

Now, by construction $\widetilde{U}_t = U_t^*$ for $t \leq T$, and for $t \in [T, T']$ we have

$$\widetilde{U}_t = U_T^* \exp((r + \gamma/K)(t-T)) \geq U_t^*.$$

Finally, for $t \geq T$ we have $\widetilde{U}_t = (K-1)/(r+\gamma) \geq \widetilde{U}_t$ given that $(K-1)/(r+\gamma)$ is an upper bound on the value of any continuation utility. So indeed $\widetilde{U} \geq U^*$.

Finally, note that trivially $U^* \geq U^1$ given that U^* is the continuation utility process of an optimal policy, bounding above the continuation utility of an arbitrary policy at any point in time. And by assumption $U_0^* > U_0^1$, as x is not a loyalty test. So indeed $\widetilde{U} \geq U^* \geq U^1$ and $\widetilde{U}_0 \geq U_0^* > U^1$, completing the proof.

D.4 Proof of Proposition 1

We begin by providing closed form expressions for the key solution parameters and a precise statement of Proposition 1. Define

$$
\begin{aligned}
t^* &\equiv \frac{1}{\gamma} \ln \left((K-1) \frac{1-q}{q} \right) \\
\overline{x} &\equiv \frac{(K-1)\gamma}{K(\gamma + r)} \\
\underline{x} &\equiv \frac{(K-1)\gamma}{K(\gamma + r)} \left(\frac{Kq}{K-1} \right)^{1 + \frac{Kr}{\gamma}} \\
\overline{t}_L &\equiv \frac{K}{\gamma} \ln \left(\frac{K-1}{Kq} \right) \\
\underline{t}_M &\equiv t^* - \frac{1}{\gamma} \ln \left(K e^{\frac{\gamma}{K}\Delta} - (K-1) \right) \\
\Delta &\equiv \left(r + \frac{\gamma}{K} \right)^{-1} \ln \left(\frac{\overline{x}}{\phi} \right) \\
\overline{t}_M &\equiv \underline{t}_M + \Delta,
\end{aligned}
$$

and for completeness, define $\underline{t}_L \equiv 0$ and $\underline{t}_H \equiv \overline{t}_H \equiv t^*$.

We will prove that the unique optimal disclosure path x^* is as follows:

1. If $q \geq \frac{K-1}{K}$, $x_t^* = 1$ for all $t \geq 0$.

2. If $q < \frac{K-1}{K}$ and $\phi \geq \overline{x}$,

$$
x_t^* = \begin{cases} \phi & \text{if } t \in [0, t^*) \\ 1 & \text{if } t \geq t^*. \end{cases}
$$

3. If $q < \frac{K-1}{K}$ and $\phi \in (\underline{x}, \overline{x})$,

$$
x_t^* = \begin{cases} \phi & \text{if } t \in [0, \underline{t}_M) \\ \phi e^{(r+\gamma/K)(t - \underline{t}_M)} & \text{if } t \in [\underline{t}_M, \overline{t}_M) \\ 1 & \text{if } t \geq \overline{t}_M. \end{cases}
$$

4. If $q < \frac{K-1}{K}$ and $\phi \leq \underline{x}$,

$$x_t^* = \begin{cases} \underline{x} e^{(r+\gamma/K)t} & \text{if } t \in [0, \bar{t}_L) \\ 1 & \text{if } t \geq \bar{t}_L. \end{cases}$$

We solve the first two cases by solving a relaxed version of (PP), ignoring the IC constraint on $\dot{U}_t$, and checking afterward that it is satisfied. Note that the coefficient $q-(1-q)(K-1)e^{-\gamma t}$ in (1) is positive iff $q \geq (K-1)/K$ or both $q < \frac{K-1}{K}$ and $t > t^*$. Hence, the principal's relaxed problem is solved pointwise by setting $x_t = 1$ in those cases, and by setting $x_t = \phi$ otherwise. Now if $q \geq (K-1)/K$, we have $x_t^* = 1$ for all $t \geq 0$ and clearly the IC constraint is satisfied since $\dot{U}_t = 0$, solving the original problem and thus giving the first case of Proposition 1. If $q < (K-1)/K$, IC is clearly satisfied for the conjectured path at all times $t \geq t^*$. At times $t < t^*$, we have

$$\dot{U}_t = -(K-1)x_t + (r+\gamma)U_t,$$

and thus $\dot{U}_t \leq (r+\gamma/K)U_t$ if and only if $x_t \geq \frac{\gamma}{K}U_t$. Now $x_t = \phi$ and since U_t is increasing for $t < t^*$, this inequality is satisfied for all $t < t^*$ if and only if $\phi \geq \frac{\gamma}{K}U_{t^*-} = \frac{\gamma}{K}\frac{K-1}{r+\gamma} = \overline{x}$. Thus the solution to the relaxed problem solves (PP), giving the second case of Proposition 1.

For the remaining two cases of the proposition, where $q < \frac{K-1}{K}$ and $\phi < \overline{x}$, the IC constraint binds and we must solve the unrelaxed problem (PP''). We make use of notation and arguments from the proof of Lemma 7. The optimal disclosure path induces some expected utility U_0 for the disloyal agent, with U_0 lying in one of four possible regimes: $u = \underline{U}$, $u \in (\underline{U}, \widehat{U})$, $u \in [\widehat{U}, \overline{U})$ or $u = \overline{U}$, where $\widehat{U} = K\phi/\gamma$. We eliminate the first and fourth possibility as follows. If $U_0 = \underline{U}$, then $x_t^* = \phi$ for all time. But for sufficiently small $\varepsilon > 0$, an alternative contract with disclosure path $\tilde{x}_t$ defined by $\tilde{x}_t = \phi$ for $t \in [0, t^*)$ and $\tilde{x}_t = \phi + \varepsilon$ is a loyalty test in $\mathbb{X}$ and delivers a strictly higher payoff to the principal. Similarly, if $U_0 = \overline{U}$ then $x_t^* = 1$ for all time, but for sufficiently small $\varepsilon > 0$ an alternative contract with disclosure path $\tilde{x}_t$ defined by $\tilde{x}_t = 1 - \varepsilon$ for $t \in [0, t^*)$ and $\tilde{x}_t = 1$ for $t \geq t^*$ is a loyalty test in $\mathbb{X}$ and is a strict improvement over x^*. Hence, the optimal contract must satisfy either $U_0 \in (\underline{U}, \widehat{U})$ or $U_0 \in [\widehat{U}, \overline{U})$, with the associated optimal disclosure path x^* given by either equation (D.2) or equation (D.3), respectively.

Denote by $x^{*,M}$ and $x^{*,L}$ the paths defined in equations (D.2) and (D.3), respectively, with $u = U_0$. Let $\Pi^M[U_0]$ and $\Pi^L[U_0]$ be their respective payoffs to the principal. Also define the grand payoff function $\widehat{\Pi}[U_0]$ over $(\underline{U}, \overline{U})$ by

$$\widehat{\Pi}[U_0] = \begin{cases} \Pi^M[U_0], & U_0 \in (\underline{U}, \widehat{U}) \\ \Pi^L[U_0], & U_0 \in [\widehat{U}, \overline{U}). \end{cases}$$

The function $\widehat{\Pi}[U_0]$ captures the maximum profits achievable by the principal over all loyalty tests delivering an expected utility of exactly $U_0 \in (\underline{U}, \overline{U})$ to the disloyal agent. We will show that $\widehat{\Pi}$ is single-peaked with an interior maximizer, and we will derive a closed-form expression for the maximizer.

We begin with $x^{*,L}$, which has a single discontinuity at $\bar{t}_L(U_0) \equiv \frac{\ln\left(\frac{K-1}{(r+\gamma)U_0}\right)}{r+\frac{\gamma}{K}}$. The principal's profits under $x^{*,L}$ can therefore be written

$$\begin{aligned}
\Pi^L[U_0] &= -\left(1 - \frac{K}{K-1}q\right) U_0 + \frac{q\gamma}{K-1} \int_0^\infty e^{-rt} U_t dt \\
&= -\left(1 - \frac{K}{K-1}q\right) U_0 + \frac{q\gamma}{K-1} \left[\frac{K}{\gamma} U_0 \left(e^{\frac{\gamma}{K}\bar{t}_L(U_0)} - 1\right) + \frac{K-1}{r(r+\gamma)} e^{-r\bar{t}_L(U_0)}\right] \\
&= -U_0 + \frac{q\gamma}{K-1} \left(\frac{K}{\gamma} + \frac{1}{r}\right) \left(\frac{K-1}{r+\gamma}\right)^{\frac{\gamma/K}{r+\gamma/K}} U_0^{\frac{r}{r+\gamma/K}}.
\end{aligned}$$

Taking a derivative, we obtain

$$\frac{d\Pi^L}{dU_0} = -1 + \frac{Kq}{K-1} \left(\frac{K-1}{r+\gamma}\right)^{\frac{\gamma/K}{r+\gamma/K}} U_0^{-\frac{\gamma/K}{r+\gamma/K}}.$$

Now $\Pi^L[\cdot]$ is strictly concave, and for $q < \frac{K-1}{K}$, it has a unique maximizer

$$U_0^{*,L} \equiv \frac{K-1}{\gamma + r} \left(\frac{Kq}{K-1}\right)^{1+\frac{Kr}{\gamma}}.$$

Given that $q < \frac{K-1}{K}$ and $\phi < \overline{x}$, it is straightforward to verify that $U_0^{*,L} \in [\widehat{U}, \overline{U})$ iff $\phi \leq \underline{x}$, and that in this case $\bar{t}_L(U_0^{*,L}) > t^*$. On the other hand, if $\phi > \underline{x}$, then $U_0^{*,L} < \widehat{U}$ and thus $\Pi^L[U_0]$ is decreasing in U_0 for $U_0 \in [\widehat{U}, \overline{U})$.

Next, we turn to $x^{*,M}$. To simplify the algebra that follows, define the function

$$\underline{t}(U_0) \equiv \frac{1}{r+\gamma} \log\left(\frac{\widehat{U} - \underline{U}}{U_0 - \underline{U}}\right)$$

to be the time at which $x^{*,M}$ begins increasing. This is a strictly decreasing function of U_0, and so may be inverted to write $\widehat{\Pi}^L$ as a function of $\underline{t}$ over $(0, \infty)$.

With this change of variables, the single discontinuity of $x^{*,M}$ occurs at time $\bar{t}(\underline{t}) = \underline{t} + \Delta$, where Δ is defined in the beginning of this proof. The principal's payoff function may then be written

$$\begin{aligned}
\Pi^M[\underline{t}] &\equiv \int_0^\infty e^{-rt}(q-(K-1)(1-q)e^{-\gamma t})x_t dt \\
&= \int_0^{\underline{t}} e^{-rt}(q-(K-1)(1-q)e^{-\gamma t})\phi dt + + \int_{\bar{t}(\underline{t})}^\infty e^{-rt}(q-(K-1)(1-q)e^{-\gamma t})dt \\
&\quad + \int_{\underline{t}}^{\bar{t}(\underline{t})} e^{-rt}(q-(K-1)(1-q)e^{-\gamma t})\phi e^{(r+\gamma/K)(t-\underline{t})}dt.
\end{aligned}$$

Differentiating this expression wrt $\underline{t}$, and recalling that $\bar{t}'(\underline{t}) = 1$, we obtain

$$\begin{aligned}
\frac{d\Pi^M}{d\underline{t}} &= -\int_{\underline{t}}^{\bar{t}(\underline{t})} e^{-rt}(q-(K-1)(1-q)e^{-\gamma t})(r+\gamma/K)\phi e^{(r+\gamma/K)(t-\underline{t})}dt \\
&\quad - (1-\bar{x})e^{-r\bar{t}(\underline{t})}(q-(K-1)(1-q)e^{-\gamma\bar{t}(\underline{t})}) \\
&= -\left(r+\frac{\gamma}{K}\right)\phi e^{-r\underline{t}}\frac{K}{\gamma}\left[q\left(e^{\frac{\gamma}{K}\Delta}-1\right)+(1-q)e^{-\gamma\underline{t}}\left(e^{-\gamma\left(\frac{K-1}{K}\right)\Delta}-1\right)\right] \\
&\quad - (1-\bar{x})e^{-r\underline{t}}\left(qe^{-r\Delta}-(K-1)(1-q)e^{-r\Delta}e^{-\gamma(\underline{t}+\Delta)}\right) \\
&= \left(r+\frac{\gamma}{K}\right)e^{-r\underline{t}}\left[Ae^{-\gamma\underline{t}}+B\right],
\end{aligned}$$

where

$$\begin{aligned}
A &= -\phi\frac{K(1-q)}{\gamma}\left(e^{-\frac{\gamma(K-1)}{K}\Delta}-1\right)+\frac{1-q}{r+\gamma}(K-1)e^{-(r+\gamma)\Delta} \\
&= \frac{1-q}{r+\gamma}(K-1)e^{-\left(r+\frac{\gamma}{K}\right)\Delta} > 0 \\
B &= -\phi\frac{Kq}{\gamma}\left(e^{\frac{\gamma}{K}\Delta}-1\right)-\frac{1}{r+\gamma}qe^{-r\Delta} \\
&= -\frac{Kq}{\gamma(r+\gamma)}e^{-r\Delta}\left(\gamma-\phi(r+\gamma)e^{r\Delta}\right) < 0,
\end{aligned}$$

with the last inequality relying on the fact that $\phi < \bar{x}$.

This expression vanishes at the unique time $\underline{t}$ such that $Ae^{-\gamma\underline{t}} + B = 0$, in particular at time $\underline{t}_M$ as defined at the start of the proof. To the left of this time Π^M is stricty increasing, while to the right Π^M is strictly decreasing. It is straightforward to verify that $\underline{t}_M > 0$ iff $\phi > \underline{x}$, in which case $\Pi^M[U_0]$ is single-peaked with an interior maximizer

$$U_0^{*,M} \equiv \underline{t}^{-1}(\underline{t}_M) \in (\underline{U}, \widehat{U}),$$

and $0 < \underline{t}_M < t^* < \bar{t}_M$. Otherwise if $\phi \le \underline{x}$, then Π^M is monotonically decreasing in $\underline{t}$, or equivalently, monotonically increasing in U_0 for $U_0 \in (\underline{U}, \widehat{U})$.

Since $\Pi^M[\widehat{U}] = \Pi^L[\widehat{U}]$, the above arguments imply that $\widehat{\Pi}[U_0]$ is single-peaked, and its maximizer is (i) $U^{*,L} \in [\widehat{U}, \overline{U})$ if $\phi \le \underline{x}$ or (ii) $U^{*,M} \in (\underline{U}, \overline{U})$ if $\phi \in (\underline{x}, \overline{x})$. In case (i), $x^{*,L}$ with $u = U_0^{*,L}$ is the optimal stakes curve, and in case (ii), $x^{*,M}$ with $u = U_0^{*,M}$ is the optimum.

To summarize, in all parametric cases, the optimal stakes curve satisfies the optimal growth property and is uniquely detemined in closed form, characterized by the parameters stated in the beginning of this section.

D.5 Proof of Lemma 4

Recall that U^* is the disloyal agent's value function given the stakes curve x^*, and define $h(\beta, x, u, u') \equiv -\gamma\beta u + (K\beta - 1)x + u'$. We first claim that U^* satisfies the HJB equation

$$rU_t^* = \max_{\beta \in [0,1]} h(\beta, x_t^*, U_t^*, \dot{U}_t^*),$$

for all $t \in [0, \infty)$, and moreover, an undermining policy $\widetilde{\beta}$ satisfies

$$\tilde{\beta}_t \in \arg\max_{\beta \in [0,1]} h(\beta, x_t^*, U_t^*, \dot{U}_t^*)$$

for all time iff $\tilde{\beta}_t = 1$ for $t \in [0, \underline{t}) \cup [\bar{t}, \infty)$. (At times $\underline{t}$ and $\bar{t}$, set $\dot{U}_t^*$ equal to its right-hand derivative.) To see this, recall that U^* satisfies for all $t \in [0, \infty)$ the identity $x_t^* = \frac{1}{K-1}\left((r+\gamma)U_t^* - \dot{U}_t^*\right)$ and the constraint $\dot{U}_t^* \le (r + \gamma/K)U_t^*$, with equality in the latter for $t \in [\underline{t}, \bar{t})$. Combining these yields $U_t^* \le \frac{K}{\gamma} x^*$, with equality for $t \in [\underline{t}, \bar{t})$. It follows that for all $t \in [0, \infty)$, $h(\cdot, x_t^*, U_t^*, \dot{U}_t^*)$ is weakly increasing in β, and is strictly increasing on $[0, \underline{t}) \cup [\bar{t}, \infty)$. Thus for all $t \in [0, \underline{t}) \cup [\bar{t}, \infty)$, $\arg\max_{\beta \in [0,1]} h(\beta, x_t^*, U_t^*, \dot{U}_t^*) = 1$, and for all $t \in [\underline{t}, \bar{t})$, $\arg\max_{\beta \in [0,1]} h(\beta, x_t^*, U_t^*, \dot{U}_t^*) = [0, 1]$. Note that one maximizer of the HJB equation is $\widetilde{\beta}_t = 1$ for all time, so that

$$\max_{\beta \in [0,1]} h(\beta, x_t^*, U_t^*, \dot{U}_t^*) = -\gamma U_t^* + (K-1)x^* + \dot{U}_t^* = rU_t^*,$$

establishing that U^* satifies the HJB equation.

Now let $\tilde{\beta}$ be an arbitrary undermining policy, and let τ be the random time with hazard rate $\gamma\tilde{\beta}$. Then

$$e^{-rt}U_t^*\mathbf{1}\{\tau > t\} = U_0^* + \int_0^t (-rU_s^* + \dot{U}_s^*)e^{-rs}\mathbf{1}\{\tau > s\}ds - \sum_{s \le t} e^{-rs}U_s^*\mathbf{1}\{\tau = s\}.$$

Define $Q_t^{\tilde{\beta}} \equiv \exp\left(-\gamma \int_0^t \tilde{\beta}_s ds\right)$. Then, taking expectations above, we have

$$e^{-rt}U_t^*Q_t^{\tilde{\beta}} = U_0^* + \int_0^t e^{-rs}Q_s^{\tilde{\beta}}(-rU_s^* + \dot{U}_s^*)ds - \int_0^t e^{-rs}U_s^*\gamma\tilde{\beta}_s Q_s^{\tilde{\beta}}ds.$$

As U^* and $Q^{\tilde{\beta}}$ are bounded uniformly in time, taking $t \to \infty$ yields

$$U_0^* = \int_0^\infty e^{-rs}Q_s^{\tilde{\beta}}\left[(\gamma\tilde{\beta}_s + r)U_s^* - \dot{U}_s^*\right]ds. \tag{D.1}$$

Suppose first that $\widetilde{\beta}_t = 1$ for $t \in [0, \underline{t}) \cup [\overline{t}, \infty)$. Then $rU_t^* = h(\widetilde{\beta}, x_t^*, U_t^*, \dot{U}_t^*)$ for all time, in which case (D.1) becomes

$$U_0^* = \int_0^\infty e^{-rs}Q_s^{\tilde{\beta}}(K\tilde{\beta}_s - 1)x_s^* ds,$$

which is precisely the disloyal agent's expected payoff from the policy $\tilde{\beta}$. We conclude that any such $\tilde{\beta}$ achieves the payoff U_0^*. Conversely, suppose that $\widetilde{\beta}_t < 1$ for some $t \in [0, \underline{t}) \cup [\overline{t}, \infty)$. Then as $\widetilde{\beta}$ is càdlàg, it satisfies this inequality for a strictly positive measure of times on the interval, implying that $rU_t^* \le h(\widetilde{\beta}, x_t^*, U_t^*, \dot{U}_t^*)$ for all time, with the inequality holding strictly over a positive measure of times. In this case (D.1) implies

$$U_0^* > \int_0^\infty e^{-rs}Q_s^{\tilde{\beta}}(K\tilde{\beta}_s - 1)x_s^* ds.$$

Thus U_0^* is the agent's value, and is achieved precisely by the set of undermining policies described in the lemma statement.

D.6 Proof of Lemma 5

Fix a stakes curve x. Suppose there existed a time t at which $\dot{U}_t$ is defined and (2) failed. Consider any undermining policy which unconditionally sets $\beta_t = 0$ from time t to time $t+\varepsilon$ for $\varepsilon > 0$, and then unconditionally sets $\beta_t = 1$ after time $t + \varepsilon$. Let $V(\varepsilon)$ be the expected continuation payoff to the disloyal agent of such a strategy beginning at time t, conditional on not having been terminated prior to time t. Then

$$V(\varepsilon) = -\int_t^{t+\varepsilon} e^{-r(s-t)} x_s \, ds + e^{-r\varepsilon} U_{t+\varepsilon}.$$

Differentiating wrt ε and evaluating at $\varepsilon = 0$ yields

$$V'(0) = -x_t - rU_t + \dot{U}_t.$$

If $V'(0) > 0$, then for sufficiently small ε the expected payoff of unconditionally undermining forever is strictly improved on by deviating to not undermining on the interval $[t, t+\varepsilon]$ for sufficiently small $\varepsilon > 0$. (Note that undermining up to time t yields a strictly positive probability of remaining employed by time t and achieving the continuation payoff $V(\varepsilon)$.) So

$$-x_t - rU_t + \dot{U}_t \leq 0$$

is a necessary condition for unconditional undermining to be an optimal strategy for the disloyal agent.

Now, by the fundamental theorem of calculus $\dot{U}_t$ is defined whenever x_t is continuous, and is equal to

$$\dot{U}_t = (r+\gamma)U_t - (K-1)x_t.$$

As x is a monotone function it can have at most countably many discontinuities. So this identity holds a.e. Solving for x_t and inserting this identity into the inequality just derived yields (2), as desired.

In the other direction, suppose x is a stakes curve satisfying (2) a.e. To prove the converse result we invoke Lemma C.3, which requires establishing that

$$\dot{U}_t = \min\{(r+\gamma)U_t - (K-1)x_t, rU_t + x_t\}$$

a.e. By the fundamental theorem of calculus, $\dot{U}_t$ exists a.e. and is equal to

$$\dot{U}_t = (r+\gamma)U_t - (K-1)x_t$$

wherever it exists. So it is sufficient to show that $(r+\gamma)U_t - (K-1)x_t \leq rU_t + x_t$ a.e., i.e. $\gamma U_t \leq K x_t$.

Combining the hypothesized inequality $\dot{U}_t \leq \left(r + \frac{\gamma}{K}\right) U_t$ and the identity $\dot{U}_t = (r+\gamma)U_t -$

$(K-1)x_t$ yields the inequality

$$(r+\gamma)U_t - (K-1)x_t \leq \left(r + \frac{\gamma}{K}\right) U_t.$$

Re-arrangement shows that this inequality is equivalent to $x_t \geq \gamma U_t/K$, as desired.

D.7 Proof of Lemma 6

Let

$$\Pi = \int_0^\infty e^{-rt} x_t (q - (1-q)(K-1)e^{-\gamma t})\, dt.$$

Use the fact that $\dot{U}_t = -(K-1)x_t + (r+\gamma)U_t$ as an identity to eliminate x_t from the objective, yielding

$$\Pi = \int_0^\infty \left(\frac{q}{K-1}e^{\gamma t} - (1-q)\right) e^{-(r+\gamma)t}((r+\gamma)U_t - \dot{U}_t)\, dt.$$

This expression can be further rewritten as a function of U only by integrating by parts. The result is

$$\Pi = -\left(1 - \frac{K}{K-1}q\right) U_0 - \lim_{t\to\infty} \left(\frac{q}{K-1}e^{\gamma t} - (1-q)\right) e^{-(r+\gamma)t} U_t + \frac{q\gamma}{K-1}\int_0^\infty e^{-rt} U_t\, dt.$$

Since U is bounded in the interval $[0, (K-1)/(r+\gamma)]$, the surface term at $t = \infty$ must vanish, yielding the expression in the lemma statement.

D.8 Proof of Lemma 7

Fix $u \in [\underline{U}, \overline{U}]$. Ignoring the fixed term involving U_0 and all multiplicative constants, the principal's objective reduces to

$$\int_0^\infty e^{-rt} U_t\, dt,$$

which is increasing in the path of U_t pointwise. Let $g(U) \equiv \min\{(r+\gamma/K)U, (r+\gamma)U - (K-1)\phi\}$. We begin by constructing the unique pointwise maximizer of U subject to $U_0 = u$, $U \in \mathbb{U}$, and $\dot{U}_t \leq g(U_t)$ a.e.

Define a utility path $U^\dagger(u)$ as follows. Let $\widehat{U} \equiv K\phi/\gamma > \underline{U}$ be the minimal utility level at which the lower bound constraint on stakes ceases to bind, and let

$$\underline{t}^\dagger(u) \equiv \max\left\{0, \frac{1}{r+\gamma}\log\left(\frac{\widehat{U} - \underline{U}}{u - \underline{U}}\right)\right\}$$

be the corresponding first time at which the disloyal agent's continuation utility exceeds level $\widehat{U}$, ignoring the upper bound on utility. (If $u = \underline{U}$, interpret this expression as $\underline{t}^\dagger(u) = \infty$.) Then define

$$U^\dagger(u)_t = \begin{cases} e^{(r+\gamma)t}\left(u - \underline{U}\left(1 - e^{-(r+\gamma)t}\right)\right), & t < \underline{t}^\dagger(u) \\ \widehat{U}\exp\left(\left(r + \frac{\gamma}{K}\right)(t - \underline{t}^\dagger(u))\right), & t \geq \underline{t}^\dagger(u) \end{cases}.$$

It is readily checked by direct computation that $U^\dagger$ satisfies $U_0^\dagger = u$ and $\dot{U}^\dagger(u)_0 = g(U^\dagger(u)_t)$ a.e., with $g(U^\dagger(u)_t) = (r+\gamma)U^\dagger(u)_t - (K-1)\phi$ for $t < \underline{t}^\dagger(u)$ and $g(U^\dagger(u)_t) = (r+\gamma/K)U^\dagger(u)_t$ for $t \geq \underline{t}^\dagger(u)$.

Now, fix an arbitrary absolutely continuous function U satisfying $U_0 = u$ and the control constraint. Since $\dot{U}_t \leq (r+\gamma)U_t - (K-1)\phi$ a.e., the integral form of Grönwall's inequality implies that

$$U_t \leq e^{(r+\gamma)t}\left(u - \underline{U}\left(1 - e^{-(r+\gamma)t}\right)\right)$$

for all time, and in particular for $t \leq \underline{t}^\dagger(u)$, where it implies $U_t \leq U^\dagger(u)_t$. Additionally, $\dot{U}_t \leq (r + \gamma/K)U_t$ a.e. Then applying Grönwall's inequality again beginning at time $\underline{t}^\dagger(u)$ implies that

$$U_t \leq U_{\underline{t}^\dagger(u)}\exp\left(\left(r + \frac{\gamma}{K}\right)(t - \underline{t}^\dagger(u))\right) \leq U_t^\dagger$$

for all $t > \underline{t}^\dagger(u)$, with the second inequality following following from $U_{\underline{t}^\dagger(u)} \leq U^\dagger(u)_{\underline{t}^\dagger(u)} = \widehat{U}$ along with the fact that $\widehat{U} > 0$. So $U \leq U^\dagger(u)$. Thus $U^\dagger(u)$ is the unique pointwise maximizer among all absolutely continuous functions satisfying the control constraint and $U_0 = u$.

Define a function $U^{**}(u)$ by $U^{**}(u)_t = \min\{U^\dagger(u)_t, \overline{U}\}$. If $u = \underline{U}$, then the facts stated in the previous paragraph imply that the only element of $\mathbb{U}$ satisfying the control constraint and $U_0 = u$ is the constant function $U^{**}(u) = \underline{U}$, which is then trivially the pointwise maximizer of the objective. Otherwise, note that $U^\dagger(u)$ is a continuous, strictly increasing function, and there exists a unique time $\bar{t}^\dagger(u)$ at which $U^\dagger(u)_t = \overline{U}$. Then by the facts of the previous paragraph, $U^{**}(u)$ is the unique pointwise maximizer of U on the interval $[0, \bar{t}^\dagger(u)]$ subject to the control constraint and $U_0 = u$; and it is trivially the unique pointwise maximizer of U subject to $U \leq \overline{U}$ on the interval $(\bar{t}^\dagger(u), \infty)$. Thus $U^{**}(u)$ is the unique pointwise maximizer of U in $\mathbb{U}$ subject to the control constraint and $U_0 = u$.

The final step in the proof is to show that the stakes curve induced by $U^{**}(u)$ lies in $\mathbb{X}$. Let

$$x_t^u \equiv \frac{1}{K-1}\left((r+\gamma)U^{**}(u)_t - \dot{U}^{**}(u)_t\right)$$

wherever $U^{**}(u)$ is differentiable, with x_t^u selected so that the curve x^u is right-continuous at points of non-differentiability of $U^{**}(u)$. If $u = \underline{U}$, then $x^u = \phi$, which is in $\mathbb{X}$. Next, if

$u \in (\underline{U}, \widehat{U})$ then $\underline{t}^\dagger(u) < \overline{t}^\dagger(u)$ and

$$x_t^u = \begin{cases} \phi, & t \le \underline{t}^\dagger(u) \\ \phi \exp\left(\left(r + \frac{\gamma}{K}\right)\left(t - \underline{t}^\dagger(u)\right)\right), & \underline{t}^\dagger(u) < t < \overline{t}^\dagger(u) \\ 1, & t \ge \overline{t}^\dagger(u). \end{cases} \tag{D.2}$$

Note that $x^u_{\overline{t}^\dagger(u)} = 1$ while $x^u_{\overline{t}^\dagger(u)-} = \frac{\gamma}{K}\frac{K-1}{r+\gamma} < 1$. So x^u is monotone increasing, bounded below by $x_0^u = \phi$ and above by 1, meaning $x^u \in \mathbb{X}$. If $\widehat{U} < \overline{U}$ and $u \in [\widehat{U}, \overline{U})$, then

$$x_t^u = \begin{cases} \frac{u\gamma}{K} \exp\left(\left(r + \frac{\gamma}{K}\right) t\right), & t < \overline{t}^\dagger(u) \\ 1, & t \ge \overline{t}^\dagger(u). \end{cases} \tag{D.3}$$

Given $u \ge \widehat{U}$, $u\gamma/K \ge \phi$. And at $t = \overline{t}^\dagger(u)$ it satisfies $x^*_{\overline{t}^\dagger(u)} = 1$ and $x^u_{\overline{t}^\dagger(u)-} = \frac{\gamma}{K}\frac{K-1}{r+\gamma} < 1$. Thus x^u is monotone increasing everywhere, bounded below by $x_0^u = u\gamma/K \ge \phi$ and above by 1. So $x^u \in \mathbb{X}$. Finally, if $u = \overline{U}$, then $x^u = 1$ which is clearly in $\mathbb{X}$.

D.9 Proof of Lemma 8

We construct the function $\underline{q}$ piecewise over $[0, \overline{x})$ and $[\overline{x}, 1]$. It is easy to show that the principal obtains a strictly positive payoff from hiring the agent when $q \ge (K-1)/K$, i.e. $(q, \phi) \in \mathcal{S}^+$, so it suffices to consider the high, moderate, and low stakes cases, with $q < (K-1)/K$. We refer to the proof of Lemma G.1 for closed-form expressions of the principal's (normalized) payoff in each of these cases.

If $\phi \ge \overline{x}$, then the high stakes case applies. Now Π_H is strictly decreasing in ϕ and strictly increasing in q for $(q, \phi) \in \mathcal{S}^H$, and for all $\phi \ge \overline{x}$, $\Pi_H < 0$ when $q = 0$. Hence, the equation $\Pi_H = 0$ implicitly defines a continuous, increasing function $\underline{q}^H : [\overline{x}, 1] \to (0, 1]$ such that for $\phi \in [\overline{x}, 1]$, $\Pi_H < 0$ (and the principal strictly prefers not to hire the agent) iff $q < \underline{q}^H(\phi)$. Since the principal's payoff is strictly positive under the optimal contract when $q \ge (K-1)/K$, by continuity of Π_H in q, $\underline{q}^H(\phi) < (K-1)/K$ for all $\phi \in [\overline{x}, 1]$.

Now consider $\phi < \overline{x}$. Define

$$\underline{x}^\dagger(p) \equiv \frac{(K-1)\gamma}{K(\gamma + r)} \left(\frac{Kp}{K-1}\right)^{1 + \frac{Kr}{\gamma}}$$

to be the optimal initial stakes when initial beliefs are p and the lower bound constraint does not bind. (Note that $\underline{x}^\dagger(q) = \underline{x}$.) For each $\phi \in [0, \overline{x})$, the moderate stakes case applies when $q \in [0, \left(\underline{x}^\dagger\right)^{-1}(\phi))$. Now by Proposition G.1, Π_M is strictly decreasing in ϕ and strictly increasing in q, and by inspection of its formula in the proof of Lemma G.1, $\Pi_M < 0$ when $q = 0$. Thus, the equation $\Pi_M = 0$ implicitly defines a continuous, increasing function $\underline{q}^M : [0, \overline{x}) \to [0,1]$ such that for $(q,\phi) \in \mathcal{S}^M$, $\Pi_M < 0$ if and only if $q < \underline{q}^M(\phi)$. Now the low stakes case applies whenever $q \in [\left(\underline{x}^\dagger\right)^{-1}(\phi), (K-1)/K)$, and by inspection of its formula, $\Pi_L \geq 0$ with equality iff $(q,\phi) = (0,0)$. Hence, by continuity of Π_M in q, $\underline{q}^M(\phi) \leq \left(\underline{x}^\dagger\right)^{-1}(\phi) < (K-1)/K$ with the first inequality strict except when $\phi = 0$.

Now define the function $\underline{q} : [0,1] \to [0,1]$ by

$$\underline{q}(\phi) = \begin{cases} \underline{q}^M(\phi), & \phi \in [0, \underline{x}) \\ \underline{q}^H(\phi), & \phi \in [\underline{x}, 1]. \end{cases}$$

Since $\Pi_M = \Pi_H$ when both $q < (K-1)/K$ and $\phi = \overline{x}$, $\underline{q}$ is continuous at $\phi = \overline{x}$, and since $\underline{q}^H$ and $\underline{q}^M$ are continuous and increasing, so is $\underline{q}$.

Define $\mathcal{S}^\emptyset$ as in the lemma statement. By the construction above, the principal strictly prefers not to hire the agent if and only if $(q,\phi) \in \mathcal{S}^\emptyset$, and $\mathcal{S}^\emptyset$ intersects $\mathcal{S}^H$ and $\mathcal{S}^M$ but not $\mathcal{S}^L$ or $\mathcal{S}^+$. Moreover, the construction shows that $\underline{q}(\phi) < (K-1)/K$, so $\mathcal{S}^H$ is not a subset of $\mathcal{S}^\emptyset$, and $\underline{q}^M(\phi) < \left(\underline{x}^\dagger\right)^{-1}(\phi)$ for $\phi \in (0, \overline{x})$, so $\mathcal{S}^M$ is not a subset of $\mathcal{S}^\emptyset$.

E Proofs for Section 4

E.1 Proof of Lemma 9

Let U^* be the solution to problem (PP'') characterized in Lemma 7, with initial utility U_0^* as derived in the proof of Proposition 1. Fix any time $t > 0$. Then it must be the case that $(U_s^*)_{s \geq t}$ solves problem (PP'') with initial utility U_t^*. For otherwise there would exist some other utility path $\widetilde{U} \in \mathbb{U}$ satisfying $\widetilde{U}_0 = U_t^*$ and the control constraint, such that

$$\int_0^\infty e^{-rs} \widetilde{U}_s \, ds > \int_0^\infty e^{-rs} U_{s+t}^* \, ds.$$

Then the utility path $U^\dagger = ((U^*_s)_{s<t}, \widetilde{U})$ has initial utility U^*_0, lies in $\mathbb{U}$, satisfies the control constraint of problem (PP'') for all time, and by construction

$$\begin{aligned}\int_0^\infty e^{-rs} U^\dagger_s \, ds &= \int_0^t e^{-rs} U^*_s \, ds + \int_t^\infty e^{-rs} \widetilde{U}_{s-t} \, ds \\ &> \int_0^t e^{-rs} U^*_s \, ds + \int_t^\infty e^{-rs} U^*_s \, ds \\ &= \int_0^\infty e^{-rs} U^*_s \, ds,\end{aligned}$$

contradicting the assumed optimality of U^*.

Now, for $u \in [\underline{U}, \overline{U}]$ and $p \in [0, 1]$, define

$$\Pi(u, p) \equiv -\left(1 - \frac{K}{K-1} p\right) u + \frac{p\gamma}{K-1} \int_0^\infty e^{-rs} U^{**}(u)_s \, ds,$$

where recall that $U^{**}(u)$ solves problem (PP'') with $U_0 = u \in [\underline{U}, \overline{U}]$. Let

$$u^*(p) \equiv \operatorname{argmax}_{u \in [\underline{U}, \overline{U}]} \Pi(u, p).$$

The proof of Proposition 1 establishes that u^* is single-valued. We next show that u^* is a continuous, increasing function, and strictly increasing whenever $u^*(p) \in (\underline{U}, \overline{U})$. Recall from the proof of Lemma 7 that $U^{**}(u)_s$ is differentiable and increasing in u for almost every s. Thus $\Pi(u, p)$ is continuous in (u, p), so by the maximum theorem so is u^*. Further

$$\frac{\partial^2 \Pi}{\partial u \partial p} = \frac{K}{K-1} + \frac{\gamma}{K-1} \int_0^\infty e^{-rs} \frac{\partial U^{**}_s}{\partial u} \, ds > 0.$$

So Π is supermodular in (u, p), and therefore by Topkis's theorem $u^*(p)$ is (weakly) increasing in p. Further, whenever $u^*(p) \in (\underline{U}, \overline{U})$, strict positivity of $\partial^2 \Pi / (\partial u \, \partial p)$ implies that $\partial \Pi / \partial u$ is strictly positive at $u = u^*(p)$ for $p' > p$ whenever $u^*(p) > \underline{U}$. Thus $u^*(p)$ must be strictly increasing in p whenever $u^*(p) \in (\underline{U}, \overline{U})$.

Lemma E.1. *There exists a unique sequence of beliefs $(q'_t)_{t \in [0, \bar{t})}$ such that $(x^*_s)_{s \geq t} = x^{**}(x^*_{t-}, q'_t)$ for every $t \in [0, \bar{t})$. This sequence is characterized by $u^*(q'_t) = U^*_t$ for each t and is continuous, strictly increasing, and satisfies $q'_0 = q$ and $q'_{\bar{t}-} = (K-1)/K$. Further for any $t \geq \bar{t}$ and $q'_t \geq (K-1)/K$, $(x^*_s)_{s \geq t} = x^{**}(x^*_{t-}, q'_t)$.*

Proof. Recall that by definition $\bar{t} = \inf\{t \ : \ U^*_t = \overline{U}\}$. We first show that there exists a unique sequence $(q'_t)_{t \in [0, \bar{t})}$ such that $u^*(q'_t) = U^*_t$, and that this sequence is continuous and strictly increasing in t and satisfies $q'_0 = q$ and $q'_{\bar{t}-} = (K-1)/K$. Existence, uniqueness, and continuity follow directly from the facts that 1) u^* is continuous and strictly increasing in p whenever it is interior-valued; 2) $\Pi(u, 0)$ is strictly decreasing in u, so that $u^*(0) = \underline{U}$, while $\Pi(u, 1)$ is strictly increasing in u, so that $u^*(1) = \overline{U}$; and 3) $U_t^* \in (\underline{U}, \overline{U})$ for $t \in [0, \bar{t})$, with the lower bound following from the fact that x^* is non-constant, which is only possible for $U_0^* > \underline{U}$. And q_t' is strictly increasing in t because U_t^* is strictly increasing in t on $[0, \bar{t}]$.

As for the values of q' at the endpoints of the interval $[0, \bar{t})$, $u^*(q) = U_0^*$ immediately implies $q_0' = q$. Meanwhile, for $p \geq (K-1)/K$ the function $\Pi(u, p)$ is strictly increasing in u, so $u^*(p) = \overline{U}$. And for $p < (K-1)/K$, for u less than $\overline{U}$ profits satisfy

$$\Pi(u, p) = -\left(1 - \frac{K}{K-1} p\right) u + \frac{p\gamma}{K-1} \left(\int_0^{\bar{t}^\dagger(u)} e^{-rs} U^\dagger(u)_s \, ds + \int_{\bar{t}^\dagger(u)}^\infty e^{-rs} \overline{U} \, ds \right),$$

and so since $U^\dagger(u)_{\bar{t}^\dagger(u)} = \overline{U}$ and $\bar{t}^\dagger(\overline{U}) = 0$,

$$\frac{\partial \Pi}{\partial u}(\overline{U}, p) = -\left(1 - \frac{K}{K-1} p\right) < 0.$$

Thus $u^*(p) < \overline{U}$ for $p < (K-1)/K$. Suppose first that $q_{\bar{t}-}' < (K-1)/K$. Then continuity of u^* implies $\lim_{t \uparrow \bar{t}} u^*(q_t') = u^*(q_{\bar{t}-}') < \overline{U}$, contradicting the fact that $u^*(q_{\bar{t}-}') = U_{\bar{t}-}^* = \overline{U}$. Next suppose that $q_{\bar{t}-}' > (K-1)/K$. Then for t sufficiently close to $\bar{t}$, $q_t' > (K-1)/K$ and so $u^*(q_t') = \overline{U}$. But this contradicts $u^*(q_t') = U_t^* < \overline{U}$ for $t < \bar{t}$. So it must be that $q_{\bar{t}-}' = (K-1)/K$.

Meanwhile, for $t \geq \bar{t}$ we have $U_t^* = \overline{U}$. Then as $u^*(q_{\bar{t}-}') = \lim_{t \uparrow \bar{t}} U_t^* = \overline{U}$ by construction, it must be that $u^*(q_t') = U_t^*$ for all $t \geq \bar{t}$ and any $q_t' \geq q_{\bar{t}-}'$.

We next establish that if $(q_t')_{t \geq 0}$ satisfies $u^*(q_t') = U_t^*$ for every t, it also satisfies $(x_s^*)_{s \geq t} = x^{**}(x_{t-}^*, q_t')$ for every t. Let $(q_t')_{t \geq 0}$ be a sequence satisfying the hypothesis. The definition of u^* immediately implies that $(x_s^*)_{s \geq t} = x^{**}(\phi, q_t')$ for every such $t > 0$. Further, $x^{**}(\phi, q_t')_0 = x_t^* \geq x_{t-}^*$ for each t, hence the lower bound constraint does not bind when increased from ϕ to x_{t-}^*, i.e. $x^{**}(\phi, q_t') = x^{**}(x_{t-}^*, q_t')$. So for q_t' such that $(x_s^*)_{s \geq t} = x^{**}(\phi, q_t')$, it is also the case that $(x_s^*)_{s \geq t} = x^{**}(x_{t-}^*, q_t')$.

Finally, we show that for $t < \bar{t}$, $x^{**}(x_{t-}^*, p) \neq x^{**}(x_{t-}^*, q_t')$ whenever $p \neq q_t'$. Thus the unique sequence $(q_t')_{t \in [0, \bar{t})}$ satisfying $u^*(q_t') = U_t^*$ is also the unique sequence satisfying $(x_s^*)_{s \geq t} = x^{**}(x_{t-}^*, q_t')$, which completes the proof. Suppose by way of contradiction that $(x_s^*)_{s \geq t} = x^{**}(x_{t-}^*, p)$ for some $t < \bar{t}$ and $p \neq q_t'$. Then the corresponding optimal utility processes would be the same, and in particular $u^*(p; x_{t-}^*) = u^*(q_t'; x_{t-}^*)$ (where we now augment the argument of the optimal initial utility u^* to indicate the stakes bound). But $u^*(\cdot; x_{t-}^*)$ is strictly increasing whenever it does not take an extremal value, and further $u^*(\cdot; x_{t-}^*) > \underline{U}$, so it can only be that $u^*(p; x^*_{t-}) = u^*(q'_t; x^*_{t-}) = \overline{U}$. But by construction $u^*(q'_t; x^*_{t-}) = U^*_t < \overline{U}$ given $t < \bar{t}$, which is the desired contradiction. Thus $p \neq q'_t$ implies $x^{**}(x^*_{t-}, p) \neq x^{**}(x^*_{t-}, q'_t)$. □

We next establish that there exists an incentive-compatible undermining policy β^* generating posterior beliefs π such that $\pi_t = q'_t$ for all time for some sequence $(q'_t)_{t\geq 0}$ satisfying $(x^*_s)_{s\geq t} = x^{**}(x^*_{t-}, q'_t)$ for all time. The previous lemma guarantees that the sequence q' is unique for $t < \bar{t}$, and as the essentially unique incentive-compatible undermining policy for times $t \geq \bar{t}$ is $\beta^*_t = 1$, this undermining policy is essentially unique if it exists.

First, recall that any increasing sequence $q'_t \geq q'_{\bar{t}-}$ for $t \geq \bar{t}$ satisfies the desired properties. Thus in particular the posterior beliefs generated by beginning at beliefs $q'_{\bar{t}-}$ and updating assuming the undermining rule $\beta = 1$ work. So we may restrict attention to times $t < \bar{t}$, in which case $u^*(q'_t) \in (\underline{U}, \overline{U})$. For any p such that $u^*(p) < \overline{U}$, $u^*(p)$ is characterized by the first-order condition

$$\frac{\partial \Pi}{\partial u} = -\left(1 - \frac{K}{K-1}p\right) + \frac{\gamma}{K-1}p\frac{\partial}{\partial u}\int_0^\infty e^{-rs}U^{**}(u)_s\, ds = 0.$$

Lemma E.2. *For every $u \in (\underline{U}, \overline{U}]$,*

$$\frac{\partial}{\partial u}\int_0^\infty e^{-rs}U^{**}(u)_s\, ds = \frac{r\int_0^\infty e^{-rs}U^{**}(u)_s\, ds - u}{\min\{(r+\gamma/K)u, (r+\gamma)(u-\underline{U})\}}.$$

Proof. The key to proving this result is the observation that every $U^{**}(u)$ is a time-shifted version of the same curve, when $U^{**}(u)$ is appropriately extended over the entire real line. This extension may be accomplished by defining $U^{**}(u)_t = \min\{U^\dagger(u)_t, \overline{U}\}$, where, extending the logic of the proof of Lemma 7, $U^\dagger(u)$ is the unique solution for all time to the ODE

$$\dot{U}_t = \min\{(r+\gamma/K)U_t, (r+\gamma)U_t - (K-1)\phi\}$$

with initial condition $U^\dagger(u)_0 = u \in (\underline{U}, \overline{U}]$.

Recall from the proof of Lemma 7 that the binding constraint on the evolution of U_t depends on which side of the constant $\widehat{U} = K\phi/\gamma$ the current utility level lies. Suppose first that $\widehat{U} \geq \overline{U}$. Then

$$(r+\gamma)u - (K-1)\phi = (r+\gamma)(u-\underline{U}) \leq (r+\gamma/K)u$$

for all $u \in (\underline{U}, \overline{U}]$, and so whenever $U^{**}(u)_t < \overline{U}$ it must satisfy the ODE $\dot{U}^{**}(u)_t = (r+$

$\gamma)U^{**}(u)_t - (K-1)\phi$. Solving this ODE yields

$$U^{**}(u)_t = \min\left\{(u - \underline{U})e^{(r+\gamma)t} + \underline{U}, \overline{U}\right\}.$$

For every $u \in (\underline{U}, \overline{U}]$, let $\overline{\tau}(u) \equiv \inf\{t \;:\; U^{**}(u)_t = \overline{U}\}$. Some algebra reveals that

$$\overline{\tau}(u) = \frac{1}{r+\gamma} \ln \frac{\overline{U} - \underline{U}}{u - \underline{U}}.$$

In particular, note that $U^{**}(u)$ may be equivalently written

$$U^{**}(u)_t = \min\left\{(\overline{U} - \underline{U})e^{(r+\gamma)(t-\tau(u))} + \underline{U}, \overline{U}\right\}.$$

In other words, $U^{**}(u)_t = U^{**}(\overline{U})_{t-\tau(u)}$. Then

$$\begin{aligned}
&\frac{\partial}{\partial u}\int_0^\infty e^{-rt}U^{**}(u)_t\,dt \\
&= \frac{\partial}{\partial u}\int_0^\infty e^{-rt}U^{**}(\overline{U})_{t-\overline{\tau}(u)}\,dt \\
&= \frac{\partial}{\partial u}\int_{-\overline{\tau}(u)}^\infty e^{-r(t+\overline{\tau}(u))}U^{**}(\overline{U})_t\,dt \\
&= \overline{\tau}'(u)\left(-r\int_{-\overline{\tau}(u)}^\infty e^{-r(t+\overline{\tau}(u))}U^{**}(\overline{U})_t\,dt + U^{**}(\overline{U})_{-\overline{\tau}(u)}\right) \\
&= \overline{\tau}'(u)\left(-r\int_0^\infty e^{-rt}U^{**}(u)_t\,dt + U^{**}(u)_0\right).
\end{aligned}$$

Noting that $U^{**}(u)_0 = u$ and $\overline{\tau}'(u) = -((r+\gamma)(u-\underline{U}))^{-1}$ recovers the expression in the lemma statement.

Now suppose that $\widehat{U} < \overline{U}$. Extending the logic of the proof of Lemma 7, for each $u \in (\underline{U}, \overline{U}]$ there exists a unique time $\underline{\tau}(u)$, possibly negative, at which $U^{**}(u)_{\underline{\tau}(u)} = \widehat{U}$, given by

$$\underline{\tau}(u) \equiv \begin{cases} \frac{1}{r+\gamma}\ln\frac{\widehat{U}-\underline{U}}{u-\underline{U}}, & u \le \widehat{U}, \\ \frac{1}{r+\gamma/K}\ln\frac{\widehat{U}}{u}, & u > \widehat{U} \end{cases}.$$

It follows immediately by time-homogeneity of the ODE characterizing U^{**}, and global uniqueness of solutions to the ODE, that for every $u \in (\underline{U}, \overline{U}]$ and each time $t \in \mathbb{R}$, $U^{**}(u)_t = U^{**}(\widehat{U})_{t-\underline{\tau}(u)}$. Then work virtually identical to the $\widehat{U} \ge \overline{U}$ case reveals that

$$\frac{\partial}{\partial u}\int_0^\infty e^{-rt}U^{**}(u)_t\,dt = \overline{\tau}'(u)\left(-r\int_0^\infty e^{-rt}U^{**}(u)_t\,dt + u\right).$$

The result of the lemma statement follows by noting that $\underline{\tau}'(u) = -((r+\gamma)(u-\underline{U}))^{-1}$ for $u < \widehat{U}$ while $\underline{\tau}'(u) = -((r+\gamma/K)u)^{-1}$ for $u > \widehat{U}$, with the one-sided derivatives agreeing at $u = \widehat{U}$ and $(r+\gamma)(u-\underline{U}) \leq (r+\gamma/K)u$ iff $u \leq \widehat{U}$. $\square$

Lemma E.3. *For $p \in [0,1]$ such that $u^*(p) \in (\underline{U}, \overline{U}) \setminus \{\widehat{U}\}$, $u^*(p)$ is differentiable and satisfies*

$$\gamma p(1-p)\frac{du^*}{dp} = \begin{cases} (r+\gamma)(u^*(p)-\underline{U}), & u^*(p) < \widehat{U} \\ K(1-p)(r+\gamma/K)u^*(p), & u^*(p) > \widehat{U}. \end{cases}$$

Proof. Using the identity derived in Lemma E.2, the first-order condition satisfied by any $u^*(p) \in (\underline{U}, \overline{U})$ may be written

$$\frac{\gamma p\left(r\int_0^\infty e^{-rs}U^{**}(u^*(p))_s\,ds - u^*(p)\right)}{(K-1)\min\{(r+\gamma/K)u^*(p), (r+\gamma)(u^*(p)-\underline{U})\}} = 1 - \frac{K}{K-1}p.$$

This expression may be simplified by re-arranging the definition of $\Pi^*(p)$ to obtain the identity

$$\frac{\gamma p}{K-1}\int_0^\infty e^{-rs}U^{**}(u^*(p))_s\,ds = \Pi^*(p) + \left(1 - \frac{K}{K-1}p\right)u^*(p),$$

where $\Pi^*(p) \equiv \Pi(u^*(p), p)$. Inserting this identity into the FOC and rearranging yields the reduced FOC

$$r\Pi^*(p) = \begin{cases} \gamma(1-p)u^*(p) - (r+\gamma)\underline{U}\left(1 - \frac{K}{K-1}p\right), & u^*(p) \leq \widehat{U}, \\ \frac{\gamma u^*(p)}{K}, & u^*(p) > \widehat{U}. \end{cases}$$

The implicit function theorem implies that the derivative of $u^*(p)$ exists everywhere except when $u^*(p) = \widehat{U}$, at which point the FOC characterizing it is not continuously differentiable.

Meanwhile, by the envelope theorem

$$\frac{d\Pi^*}{dp} = \frac{K}{K-1}u^*(p) + \frac{\gamma}{K-1}\int_0^\infty e^{-rs}U^{**}(u^*(p))_s\,ds,$$

which may be simplified by the previous identity to read

$$\frac{d\Pi^*}{dp} = p^{-1}(\Pi^*(p) + u^*(p)).$$

Supposing first that $u^*(p) < \widehat{U}$, this derivative may be further re-written using the reduced FOC as

$$r\frac{d\Pi^*}{dp} = p^{-1}\left((\gamma(1-p)+r)u^*(p) - (r+\gamma)\underline{U}\left(1 - \frac{K}{K-1}p\right)\right).$$

Meanwhile differentiating the reduced FOC yields

$$r\frac{d\Pi^*}{dp} = \gamma(1-p)\frac{du^*}{dp} - \gamma u^*(p) + (r+\gamma)\underline{U}\frac{K}{K-1}.$$

Equating these two expressions for $d\Pi^*/dp$ and rearranging yields the identity in the lemma statement.

Now supposing that $u^*(p) > \widehat{U}$, the derivative $d\Pi^*/dp$ may be rewriten using the reduced FOC as

$$r\frac{d\Pi^*}{dp} = p^{-1}\left(r + \frac{\gamma}{K}\right)u^*(p).$$

Meanwhile differentiating the reduced FOC yields

$$r\frac{d\Pi^*}{dp} = \frac{\gamma}{K}\frac{du^*}{dp}.$$

Equating these two expressions and rearranging yields the identity in the lemma statement.

□

Now, by Bayes' rule undermining policy β yields the sequence of posterior beliefs

$$\pi_t = \frac{q}{q + (1-q)\exp\left(-\gamma\int_0^t \beta_s\,ds\right)},$$

so that

$$\dot{\pi}_t = \frac{q(1-q)\gamma\beta_t\exp\left(-\gamma\int_0^t \beta_s\,ds\right)}{\left(q + (1-q)\exp\left(-\gamma\int_0^t \beta_s\,ds\right)\right)^2} = \gamma\beta_t\pi_t(1-\pi_t).$$

Thus β_t must satisfy

$$\beta_t = \frac{\dot{q}'_t}{\gamma q'_t(1-q'_t)}$$

for almost every $t < \bar{t}$. Now, recall that q'_t is characterized by $u^*(q'_t) = U^*_t$. Differentiating both sides of this expression yields

$$\frac{du^*}{dp}(q'_t)\dot{q}'_t = \dot{U}^*_t,$$

valid everywhere except for time $\underline{t}$ (if it is nonzero and strictly less than $\bar{t}$), at which time neither u^* nor U^* is differentiable. Using Lemma E.3 to evaluate du^*/dp, this may be rewritten

$$\frac{\dot{q}'_t}{\gamma q'_t(1-q'_t)} = \begin{cases} \frac{\dot{U}^*_t}{(\gamma+r)(U^*_t-\underline{U})}, & U^*_t < \widehat{U} \\ \frac{\dot{U}^*_t}{K(1-q'_t)(r+\gamma/K)U^*_t}, & U^*_t > \widehat{U}. \end{cases}$$

Now recall that for $t < \bar{t}$,

$$\dot{U}^*_t = \begin{cases} (\gamma+r)(U^*_t - \underline{U}), & U^*_t < \widehat{U}, \\ (r+\gamma/K)U^*_t, & U^*_t > \widehat{U}. \end{cases}$$

So a time-consistent undermining policy β^* for times $t < \bar{t}$, assuming a feasible policy exists, is characterized almost everywhere by

$$\beta^*_t = \begin{cases} 1, & t < \underline{t}, \\ \frac{1}{K(1-q'_t)}, & \underline{t} \le t < \bar{t}. \end{cases}$$

Imposing right-continuity induces a unique time-consistent policy.

It remains to check that this policy is feasible and incentive-compatible. To check feasibility, we must verify that $\beta^*_t \le 1$ for all $t < \bar{t}$. This follows from the fact that $q'_t < q'_{\bar{t}-} = (K-1)/K$ for $t < \bar{t}$, so that $1/(K(1-q'_t)) < 1$, as required. Thus β^* is a feasible undermining policy. It is also incentive-compatible for times $t < \bar{t}$, as it sets $\beta^*_t = 1$ for $t < \underline{t}$, and any undermining intensity is optimal for times $t \in [\underline{t}, \bar{t})$. Extending β^* to all times by setting $\beta^*_t = 1$ for $t \ge \bar{t}$, this policy is incentive-compatible for all times, as desired. Finally, note that on the time interval $[\underline{t}, \bar{t})$, supposing it is non-empty, q'_t is continuous, strictly increasing, and satisfies $q'_{\underline{t}} \ge q'_0 = q$ and $q'_{\bar{t}-} = (K-1)/K$. Thus on this interval β^* is continuous, strictly increasing, and satisfies $\beta_{\underline{t}} \ge 1/(K(1-q))$ and $\beta_{\bar{t}-} = 1$.

E.2 Proof of Proposition 2

Fix any Bayes Nash equilibrium with equilibrium path (x, β). Payoff-maximization requires that whenever undermining is discovered on the equilibrium path (that is, β called for positive-intensity undermining the instant before discovery), the agent is immediately fired. Further, whenever undermining is off the equilibrium path, it is without loss to assume that the principal immediately fires after detecting undermining. This is because Nash equilibrium imposes no optimality requirements on the principal's play in such information sets, and increasing the probability of immediate firing does not give the agent any incentives to begin undermining at such instants. So if (x, β) is an equilibrium path of any Nash equilibrium, it is the equilibrium path of one featuring immediate firing following detected undermining, on or off the equilibrium path. We assume such firing behavior going forward.

Suppose first that (x, β) is not incentive-compatible. Then there exists another undermining path $\widetilde{\beta}$ which strictly increases the disloyal agent's ex ante payoff, supposing that the principal commits to stakes path x and immediate firing upon detecting undermining. But as the equilibrium response to detected undermining is identical to the commitment response, $\widetilde{\beta}$ must also strictly increase the disloyal agent's ex ante payoff given the principal's equilibrium strategy. This contradicts the assumption that (x, β) is an equlibrium path, so (x, β) must be incentive-compatible.

Now suppose that (x, β) is incentive-compatible but not time-consistent. Let π be the path of the principal's beliefs, conditional on no observed undermining, induced by the undermining policy β. By hypothesis, there exists a time $t \geq 0$ such that $(x_s)_{s \geq t} \neq x^{**}(x_{t-}, \pi_t)$, and this information set arises with strictly positive probability given that undermining is never detected with certainty by any given time. Let $\Pi^*(y, p)$ be the principal's optimal commitment payoff when initial stakes are $y \in [0, 1]$ and the agent is loyal with prior probability $p \in [0, 1]$. Consider a deviation by the principal to play x^{**} forever afterward. No matter the agent's continuation play, this deviation must yield the principal a payoff of at least $\Pi^*(x_{t-}, \pi_t)$. This is because payoff $\Pi^*(x_{t-}, \pi_t)$ is obtained when the principal chooses stakes curve $x^{**}(x_{t-}, \pi_t)$ and a policy of immediate firing for detected undermining, and the disloyal agent responds by minimizing the principal's expected payoff. Thus an arbitrary response by the disloyal agent yields the principal no less than this payoff level.

Now, suppose that stakes path $(x_s)_{s \geq t}$ and immediate firing after detected undermining, combined with undermining policy $(\beta_s)_{s \geq t}$, yielded a profit weakly greater than $\Pi^*(x_{t-}, \pi_t)$. By assumption $(x_s, \beta_s)_{s \geq t})$ is incentive-compatible, meaning that committing to $(x_s)_{s \geq t}$ and immediate firing must yield the principal at least $\Pi^*(x_{t-}, \pi_t)$, contradicting the unique optimality of the stakes curve $x^{**}(x_{t-}, \pi_t)$ in the commitment problem. Thus the equilibrium continuation policy must yield profits strictly less than $\Pi^*(x_{t-}, \pi_t)$, meaning the principal has a profitable deviation (from a time-zero perspective) by modifying its continuation stakes curve at time t from $(x_s)_{s \geq t}$ to $x^{**}(x_{t-}, \pi_t)$. This is a contradiction, so (x, β) must be time-consistent.

E.3 Proof of Proposition 3

Throughout this proof, we will make use of the following reference undermining and belief paths. Define a function $\beta^\dagger : [0,1] \times [0,1] \to [0,1]$ by

$$\beta^\dagger(y,p) = \begin{cases} \beta^{**}(y,p)_0, & y = x^{**}(y,p), \\ 0, & y < \min\{x^{**}(y,p), \overline{x}\}, \\ 1, & \text{otherwise.} \end{cases}.$$

$\beta^\dagger$ will be used to define a Markovian undermining policy for the agent for any current stakes y and posterior beliefs p. Note that under $\beta^\dagger$, the agent plays a best response assuming the principal plays the optimal commitment stakes curve $x^{**}(y,p)$ going forward. Note in particular that y is never strictly greater than $x^{**}(y,p)$, as this path takes into account the lower bound y; however, it is possible that $y < x^{**}(y,p)$, in which case the agent expects the stakes to jump upward at the next instant.

And given any function $x : \mathbb{R}_+ \to [0,1]$ which is càdlàg and monotone, and any initial belief $p \in [0,1]$, define a function $\pi(x,p) : \mathbb{R}_+ \to [0,1]$ as the unique solution to the ODE

$$\dot{\pi}(x,p)_t = \gamma\beta^\dagger(x_t, \pi(x,p)_t)\pi(x,p)_t(1-\pi(x,p)_t)$$

with initial condition $\pi(x,p)_0 = p$. Note in particular that $\pi(x,p)$ satisfies Bayes' rule: $\pi(x,p)_t = p/Q(x,p)_t$ for all times, where

$$Q(x,p)_t \equiv p + (1-p)\exp\left(-\gamma \int_0^t \beta^\dagger(x_s, \pi(x,p)_s)\, ds\right).$$

Define a principal strategy (χ, Λ) by $\chi_t = x_t^*$ and $\Lambda = \inf\{t \;:\; N_t > 0\}$. Define an $\widetilde{\mathbb{F}}^A$-adapted reference belief process π^* as $\pi_t^* = 0$ whenever $N_t > 0$, with $\pi_t^* = \pi(X,q)_t$ otherwise. Define an agent strategy ζ by $\zeta_t = \beta^\dagger(X_t, \pi_t^*)$ if $N_t = 0$, and $\zeta_t = 1$ otherwise. Several properties of this strategy will be important in what follows. First, π^ζ satisfies the same ODE as π^* when $N_t = 0$, and both are zero whenever $N_t > 0$, and so $\pi^\zeta = \pi^*$. Second, for any time t, if $(X_s, N_s)_{s\ge t} = (x^{**}(X_t, \pi_t^*), 0)$ then $(\zeta_s)_{s\ge t} = \beta^{**}(X_t, \pi_t^*)$.

We prove that (χ, Λ, ζ) constitute a perfect Bayesian equilibrium. Consider first the agent's strategy. Fix a time t and a subgame in which $N_t > 0$. Then the agent's strategy is trivially a best response to immediate firing. So suppose $N_t = 0$, time-t stakes are $X_t = y \in [\phi, 1]$, and $\pi_t^* = p \in [0,1]$. If $p < \underline{q}(y)$, then the agent's strategy is trivially a best response to immediate firing. So suppose $p \ge \underline{q}(y)$. The final property of the previous paragraph implies that if $(X_s, N_s)_{s\ge t} = (x^{**}(y,p), 0)$, then $(\zeta_s)_{s\ge t} = \beta^{**}(y,p)$. Thus ζ induces an undermining policy which, by definition of β^{**}, is a best response to a strategy which induces $(X_s)_{s\geq t} = x^{**}(y,p)$ and fires the first time that $N_t > 0$. In particular, ζ is a best response to (χ, Λ) in the time-zero game, as $\chi = x^* = x^{**}(\phi, q)$ and $\Lambda = \inf\{t \ : \ N_t > 0\}$.

Now consider the principal's strategy. Fix a time t and a subgame in which $N_t > 0$. In this case $\pi^*_t = 0$ and the disloyal agent undermines unconditionally at all future times under ζ, and so the principal's unique best response is immediate termination. So consider instead subgames in which $N_t = 0$, the current state is $y \in [\phi, 1]$, and current beliefs are $p \in [\phi, 1]$. We will establish that 1) if $p \geq \underline{q}(y)$, then $x^{**}(y,p)$ with firing after observing undermining is a best response to $(\zeta_s)_{s\geq t}$, and 2) if $p < \underline{q}(y)$, then immediate firing is a best response. This result will complete the proof that (χ, Λ, ζ) is a PBE, as in particular it implies that the strategy profile is a Bayes Nash equilibrium, and that the agent's continuation strategy in every subgame is part of a Bayes Nash equilibrium of that subgame.

The optimality of firing following uncovered undermining is immediate, because if this event is on-path the agent is surely disloyal and sets $(\zeta_s)_t = 1$ forever afterward, and if off-path the principal's actions afterward do not affect her ex ante payoffs. So we need only characterize optimal behavior prior to observing undermining. Further, as the agent is more likely to be loyal over time, achievable continuation payoffs rise over time. Thus an optimal firing policy either fires immediately, or else never fires prior to uncovering undermining. We therefore characterize the optimal policy assuming the latter firing policy, and show that it is $x^{**}(y,p)$. Given that this induces the undermining policy $\beta^{**}(y,p)$, the principal's payoff from firing only after uncovering undermining is therefore exactly as in the commitment game. It follows that immediate firing is an optimal policy iff $p < \underline{q}(y)$. So for the remainder of the proof we ignore the possibility of immediate firing, and characterize the optimal stakes path assuming firing only after observing undermining.

Note that the agent's strategy is Markovian in the state (X_t, π^*_t), with $\beta_t = \beta^\dagger(X_t, \pi^*_t)$, and that the evolution of π^* similarly depends only on (X_t, π^*_t). So $(x_s)_{s\geq t}$ is a best response to ζ in a subgame with $(X_t, \pi^*_t) = (y,p)$ if and only if it solves the Markovian control problem

$$\sup_{x\in\mathbb{X}(y)} \Bigg\{ p\int_0^\infty e^{-rt} x_t\,dt \\ +(1-p)\int_0^\infty \exp\left(-rt - \int_0^t \beta^\dagger(x_s, \pi(x,p)_s)\,ds\right)(1-K\beta^\dagger(x_t, \pi(x,p)_t))x_t\,dt \Bigg\},$$

where $\mathbb{X}(y) \equiv \{x \in \mathbb{X} \ : \ x \geq y\}$. So it is enough to solve the family of control problems for arbitrary initial state $(y,p) \in [0,1]\times[0,1]$, and show that $x^{**}(y,p)$ is an optimal policy. (We will drop the restriction that $y \geq \phi$ and $p \geq q$ and perform the characterization over the entire space.)

Define

$$V_-(y,p) \equiv \begin{cases} x^{**}(0,p)/r, & p < (K-1)/K, \\ p/r - (1-p)(K-1)/(r+\gamma), & p \geq (K-1)/K \end{cases}$$

for all (y,p). Note that V_- is continuous everywhere, in particular at $p = (K-1)/K$, where

$$\frac{x^{**}(0,p-)}{r} = \frac{\overline{x}}{r} = \frac{K-1}{K}\frac{\gamma}{r(r+\gamma)} = \frac{K-1}{K}\left(\frac{1}{r} - \frac{1}{r+\gamma}\right) = \frac{p}{r} - (1-p)\frac{K-1}{r+\gamma}.$$

It is also differentiable wrt p everywhere except at $p = (K-1)/K$, where it has left- and right-hand derivatives which disagree. Let $V(y,p)$ be the value function of the control problem with initial stakes y and initial beliefs p. We first show that $V_-(y,p) \geq V(y,p)$ for all (y,p), and that if $y \leq x^{**}(0,p)$ then $V(y,p) = V_-(y,p)$ and $x^{**}(y,p)$ is an optimal control under initial state (y,p).

For every path x, let $\tau(x,p)$ be the random time which equals ∞ with probability p, and otherwise has a distribution with hazard rate $\gamma\beta^\dagger(x_t, \pi(x,p)_t)$. Now fix an arbitrary càdlàg, increasing, $[y,1]$-valued path x. Note that V_- is a continuously differentiable function in both variables everywhere except on the plane $p = (K-1)/K$. But because $\pi(x,p)$ is a monotone function of time, the path of $(x_t, \pi(x,p)_t)$ crosses this plane at most once, from below, and so Ito's lemma can be applied before and after the crossing time without incident and the results concatenated. Performing this process yields

$$\begin{aligned} & e^{-rt}V_-(x_t, \pi(x,p)_t)\mathbf{1}\{\tau(x,p) > t\} \\ = {} & V_-(y,p) + \int_0^t e^{-rt}\left(-rV_-(x_s, \pi(x,p)_s) + \frac{\partial V_-}{\partial p}(x_s, \pi(x,p)_s)\dot{\pi}(x,p)_s\right)\mathbf{1}\{\tau(x,p) > s\}\, ds \\ & - \sum_{s \leq t} e^{-rs}V_-(x_s, \pi(x,p)_s)\mathbf{1}\{\tau(x,p) = s\}, \end{aligned}$$

with $\frac{\partial V_-}{\partial p}$ taken to be the left-hand derivative whenever $p = (K-1)/K$ by convention. Taking expectations of both sides yields

$$\begin{aligned} & e^{-rt}V_-(x_t, \pi(x)_t)Q(x,p)_t \\ = {} & V_-(y,p) + \int_0^t e^{-rs}Q(x,p)_s\left(-rV_-(x_s, \pi(x,p)_s) + \frac{\partial V_-}{\partial p}(x_s, \pi(x,p)_s)\dot{\pi}(x,p)_s\right) ds \\ & - \int_0^t e^{-rs}V_-(x_s, \pi(x,p)_s)(1-p)\gamma\beta^\dagger(x_s, \pi(x)_s)\exp\left(-\gamma\int_0^s \beta^\dagger(x, \pi(x,p)_u)\, du\right) ds \\ = {} & V_-(y,p) + \int_0^t e^{-rs}Q(x,p)_s\Big(-(r + (1-\pi(x,p)_s)\gamma\beta^\dagger(x_s, \pi(x,p)_s))V_-(x_s, \pi(x,p)_s) \cdots \end{aligned}$$

$$+ \frac{\partial V_-}{\partial p}(x_s, \pi(x,p)_s)\dot{\pi}(x,p)_s\Big) \, ds$$

where in the final equality we have used the Bayes' rule identity

$$1 - \pi(x,p)_s = Q(x,p)_s^{-1}(1-p)\exp\left(-\gamma \int_0^s \beta^\dagger(x_u, \pi(x,p)_u)\, du\right).$$

Inserting the identity $\dot{\pi}(x,p)_s = \gamma\beta^\dagger(x_s, \pi(x,p)_s)\pi(x,p)_s(1-\pi(x,p)_s)$ and letting $t \to \infty$ yields

$$\begin{aligned} V_-(y,p) = \int_0^\infty e^{-rs} Q(x,p)_s \Big(&(r + (1-\pi(x,p)_s)\gamma\beta^\dagger(x_s, \pi(x,p)_s))V_-(x_s, \pi(x,p)_s) \cdots \\ &- \frac{\partial V_-}{\partial p}(x_s, \pi(x,p)_s)\gamma\beta^\dagger(x_s, \pi(x,p)_s)\pi(x,p)_s(1-\pi(x,p)_s)\Big) \, ds. \end{aligned}$$

Lemma E.4. *V_- satisfies*

$$\left(r + (1-p)\gamma\beta^\dagger(y,p)\right) V_-(y,p) \geq (p + (1-p)(1-K\beta(y,p)))y + \frac{\partial V_-}{\partial p}(y,p)\gamma\beta^\dagger(y,p)p(1-p)$$

for every $y, p \in [0,1]$, with equality whenever $y = x^{**}(0,p)_0$.

Proof. Call the inequality in the lemma statement the HJB equation. Suppose first that $p \geq (K-1)/K$. In this regime $x^{**}(y,p)_0 = 1$ for all y, so the inequality $y \leq x^{**}(y,p)_0$ is automatic. Suppose first that $y < \overline{x}$. In this case $\beta^\dagger(y,p) = 0$ while

$$V_-(y,p) = \frac{p}{r} - (1-p)\frac{K-1}{r+\gamma}.$$

Thus the lhs of the HJB equation reduces to

$$rV_-(y,p) \geq rV_-(y, (K-1)/K) = \overline{x},$$

while the rhs reduces to $y < \overline{x}$. So the HJB equation holds as an inequality in this case.

Suppose instead that $y \geq \overline{x}$. In this case $\beta^\dagger(y,p) = 1$ while $V_-(y,p)$ is as stated previously. The lhs of the HJB equation is then

$$(r + (1-p)\gamma)V_-(y,p) = p + p(1-p)\frac{\gamma}{r} - r(1-p)\frac{K-1}{r+\gamma} - (1-p)^2(K-1)\frac{\gamma}{r+\gamma}.$$

Meanwhile the rhs becomes

$$(p-(1-p)(K-1))y+p(1-p)\frac{\gamma}{r}+p(1-p)(K-1)\frac{\gamma}{r+\gamma}.$$

As $p \geq (K-1)/K$, the rhs is maximized at $y=1$, at which point some simplification shows it to be equal to the lhs of the HJB equation. So the HJB equation holds as an inequality in this case, and as an equality when $y=1=x^{**}(0,p)_0$.

Next consider the regime $p<(K-1)/K$, and suppose first that $y<x^{**}(y,p)$. This means that $x^{**}(y,p)=x^{**}(0,p)$, which is bounded above by $\overline{x}$ when $p<(K-1)/K$. So automatically $y \leq \overline{x}$. Thus in this case $\beta^{\dagger}(y,p)=0$ while $V_{-}(y,p)=x^{**}(0,p)_0/r$. The lhs of the HJB equation is then just $x^{**}(0,p)_0$, while the rhs is $y<x^{**}(y,p)=x^{**}(0,p)_0$. So the HJB equation holds as inequality in this case.

Suppose instead that $y=x^{**}(y,p)_0$. Consider first the case where $x^{**}(y,p)_0>x^{**}(0,p)_0$. Then $\beta^{\dagger}(y,p)=1$, and the lhs of the HJB equation is

$$(r+(1-p)\gamma)x^{**}(0,p)_0/r$$

while the rhs is

$$(p-(1-p)(K-1))y+\frac{\partial x^{**}(0,p)_0}{\partial p}\frac{\gamma}{r}p(1-p).$$

As $p<(K-1)/K$, this expression is maximized when $y=x^{**}(0,p)_0$. Meanwhile, using the explicit expression for $\underline{x}$ derived in the proof of Proposition 1, we may calculate

$$\frac{\partial x^{**}(0,p)_0}{\partial p}=\left(1+\frac{Kr}{\gamma}\right)\frac{x^{**}(0,p)_0}{p}.$$

Hence the rhs of the HJB equation is bounded above by

$$(p-(1-p)(K-1))x^{**}(0,p)_0+\left(1+\frac{Kr}{\gamma}\right)\frac{\gamma}{r}(1-p)x^{**}(0,p),$$

which some simplification shows is identically equal to the lhs of the HJB equation. So the HJB equation holds as an inequality in this case.

Finally, suppose that $x^{**}(y,p)_0=x^{**}(0,p)_0$. In this case $\beta^{\dagger}(y,p)=\beta^{**}(0,p)_0$, and a variant of the calculations in the previous paragraph show that the lhs of the HJB equation is

$$(r+(1-p)\beta^{**}(0,p)_0\gamma)x^{**}(0,p)_0/r$$

while the rhs is

$$(p-(1-p)(K\beta^{**}(0,p)_0-1))x^{**}(0,p)_0+\left(1+\frac{Kr}{\gamma}\right)\frac{\gamma}{r}\beta^{**}(0,p)_0(1-p)x^{**}(0,p).$$

Some simplification shows that the rhs is identical to the lhs so the HJB equation holds with equality in this case. □

In light of the previous lemma,

$$V_-(y,p)\geq\int_0^\infty e^{-rs}Q(x,p)_s\left((\pi(x,p)_s+(1-\pi(x,p)_s)(1-K\beta^\dagger(x_s,\pi(x,p)_s))\right)x_s\,ds,$$

with equality if $x_t=x^{**}(0,\pi(x,p)_t)_0$ for all time. Using the Bayes' rule identities

$$\pi(x,p)_t=pQ(x,p)_t^{-1},\quad 1-\pi(x,p)_t=(1-p)Q(x,p)_s^{-1}\exp\left(-\gamma\int_0^s\beta^\dagger(x_u,\pi(x,p)_u)\,du\right),$$

the right-hand side may be rewritten in a more familiar form as

$$p\int_0^\infty e^{-rs}x_s\,ds+(1-p)\int_0^\infty\exp\left(-rt-\gamma\int_0^s\beta^\dagger(x_u,\pi(x,p)_u)\,du\right)(1-K\beta^\dagger(x_s,\pi(x,p)_s))x_s\,ds,$$

which is exactly the principal's expected profit under control x. So $V_-(y,p)$ is an upper bound on the profit of any feasible control, and the upper bound is achieved by any control satisfying $x_t=x^{**}(0,\pi(x,p)_t)_0$ for all time.

Now by definition of β^{**} and π, $x^{**}(y,p)_t=x^{**}(x^{**}(y,p)_t,\pi(x^{**}(y,p),p)_t)_0$ for all time. Further, if $y\leq x^{**}(0,p)$, then the lower bound constraint never binds at any point in time along $x^{**}(y,p)$, meaning that it must also not bind at time 0 along $x^{**}(x^{**}(y,p)_t,\pi(x^{**}(y,p),p)_t)$ for any t. So $x^{**}(y,p)_t=x^{**}(0,\pi(x^{**}(y,p),p)_t)_0$ for all time, meaning that $x^{**}(y,p)$ achieves profits $V_-(y,p)$ and is an optimal policy.

Now suppose initial states (y,p) satisfy $y>x^{**}(0,p)$. Let T be the unique time satisfying $y=x^{**}(0,\pi(y,p)_t)_0$. We claim that the policy $x^\#$ defined by $x^\#_t=y$ for $t\leq T$ and $(x^\#_t)_{t\geq T}=x^{**}(0,\pi(y,p)_T)$ is an optimal policy. Once the state has reached $(y,\pi(y,p)_T)$, the claimed optimal continuation follows immediately from the Markovian nature of the problem and the optimality result for initial states satisfying $y\leq x^{**}(0,p)$. So we need only establish that the optimal policy guides the state to $(y,\pi(y,p)_T)$ by setting $x_t=y$.

Fix an arbitrary stakes path x. As $x\geq y$, it must be that $x_t>x^{**}(0,\pi(y,p)_t)_0$ for $t<T$. Hence $\beta^\dagger(x_t,\pi(x,p)_t)=\beta^{**}(x_t,\pi(x,p)_t)_0=1$ for all $t<T$, and so $\pi(x,p)_t$ is independent of x for $t<T$. In particular, $\pi(x,p)_T=\pi(y,p)_T$. So consider the modified stakes path $\widetilde{x}$ defined by $\widetilde{x}_t=x_t$ for $t<T$, and $(\widetilde{x}_s)_{s\geq t}=x^{**}(0,\pi(y,p)_T)$. This modified path may not

be feasible, as it may require a downward jump at time T. However, it provides an upper bound on the profits of the path x, as we have modified the continuation path after time T to its optimum given beliefs $\pi(x,p)_T = \pi(y,p)_T$ ignoring the lower bound constraint. Note also that the path of $\widetilde{x}$ subsequent to time T is independent of $(\widetilde{x}_s)_{s\leq t}$.

Now, expected profits from $\widetilde{x}$, which are an upper bound on the profits of x, are just

$$\int_0^T e^{-rt} Q(y,p)_t \left(\pi(y,p)_t - \pi(y,p)_t(K-1)\right) \widetilde{x}_t \, dt + e^{-rT} Q(y,p)_T V_-(y, \pi(y,p)_T).$$

Note that the integrand is strictly negative and strictly decreasing in $\widetilde{x}_t$ at each $t < T$ given that $\widetilde{x}_t > x^{**}(0, \pi(y,p)_t)_0$ and therefore $\pi(y,p)_t < (K-1)/K$. So lowering $\widetilde{x}_t$ to y at each $t < T$ must strictly increase profits. But this results in exactly the policy $x^{\#}$, establishing its optimality.

Finally, we establish that $x^{\#} = x^{**}(y,p)$, proving that $x^{**}(y,p)$ is an optimal policy for every initial state (y,p). By construction, $(x^{**}(y,p)_s)_{s\geq t} = x^{**}(x^{**}(y,p)_t, \pi(x^{**}(y,p),p)_t)$ for all t. Let $T' \equiv \inf\{t \ : \ x^{**}(y,p)_t > y\}$. Then the path $x^{**}(x^{**}(y,p)_{T'}, \pi(x^{**}(y,p),p)_{T'}) = (x^{**}(y,p)_t)_{t\geq T}$ is unconstrained by the lower stakes bound, meaning

$$(x^{**}(y,p)_t)_{t\geq T'} = x^{**}(0, \pi(x^{**}(y,p),p)_{T'}).$$

And since $x^{**}(y,p)_t = y$ for $t \leq T'$, it must be that $\pi(x^{**}(y,p),p)_t = \pi(y,p)_t$ for $t \leq T'$. So T' satisfies $y = x^{**}(0, \pi(y,p)_{T'})_0$, an equation uniquely satisfied by T, meaning $T = T'$. Thus $x^{**}(y,p)_t = y$ for $y < T$, and $(x^{**}(y,p)_t)_{t\geq T} = x^{**}(0, \pi(x^{**}(y,p),p)_T)$, meaning $x^{**}(y,p) = x^{\#}$, as desired.

E.4 Proof of Lemma 10

Proposition 1 ensures that x^* is uniquely optimal among all deterministic stakes policies. We need only ensure that it is also optimal in the wider class of stochastic policies. Fix a stakes policy $\widetilde{x}$ satisfying $\Pi(\widetilde{x}) = \Pi(x^*)$. Lemma 2 ensures that passing to the deterministic policy $\widetilde{x}'$ defined by $\widetilde{x}'_t = \mathbb{E}[\widetilde{x}_t]$ cannot reduce the principal's profits. Then if $\Pi(\widetilde{x}) = \Pi(x^*)$, it must be that $\widetilde{x}' = x^*$. So $\mathbb{E}[\widetilde{x}_t] = x^*_t$ for all time. If $\underline{t} = \overline{t}$ then the result is immediate, as in this case $x^*_t \in \{\phi, 1\}$ for all time, and in neither case is it possible to construct a non-deterministic curve satisfying $\mathbb{E}[\widetilde{x}_t] = x^*_t$, the lower bound $\widetilde{x} \geq \phi$, and the upper bound $\widetilde{x} \leq 1$ for all time. So assume $\underline{t} < \overline{t}$. The argument just given continues to establish that $\widetilde{x}_t = x^*_t$ for all $t \in [0, \underline{t}) \cup [\overline{t}, \infty)$ a.s., so we need only consider times in the range $[\underline{t}, \overline{t})$.

We next note that $\widetilde{x}$ must be a loyalty test, as otherwise by Lemma 3 there exists a loyalty test yielding strictly higher payoffs than $\widetilde{x}$, contradicting the optimality of the stakes curve x^*. For each time t, define a stochastic process W^t by

$$W_{t'}^t = \mathbb{E}_t\left[\int_{t'}^{\infty} e^{-(r+\gamma)(s-t')}(K-1)\widetilde{x}_s\,ds\right].$$

Note that W^t is the time-t expectation of the disloyal agent's ex post continuation value process. By reasoning very similar to the proof of Lemma 5, it can be shown that a necessary condition for $\widetilde{x}$ to be a loyalty test is that

$$-\widetilde{x}_{t'} - rW_{t'}^{t'} + \dot{W}_{t'}^{t'} \le 0$$

for all t' a.s. Since $\mathbb{E}_t[W_{t'}^{t'}] = W_{t'}^t$ for every $t' > t$ by the law of iterated expectations, it follows that

$$-\mathbb{E}_t[\widetilde{x}_{t'}] - rW_{t'}^t + \dot{W}_{t'}^t \le 0$$

for all t and $t' \ge t$ a.s. Also, by the fundamental theorem of calculus

$$\dot{W}_{t'}^t = (r+\gamma)W_{t'}^t - (K-1)\mathbb{E}_t[\widetilde{x}_{t'}].$$

Combining these two equations yields the condition $W_{t'}^t \le (r+\gamma/K)\dot{W}_{t'}^t$ for all t and $t' \ge t$ a.s. Now, as $\widetilde{x}_t = 1$ for $t \ge \bar{t}$ a.s., it must be that $W_{\bar{t}}^t = (K-1)/(r+\gamma) = U_{\underline{t}}^*$ for every t a.s. Suppose that for some $t \in [\underline{t}, \bar{t})$, with positive probability $W_t^t < U_t^*$. Then by integrating the incentive constraint just derived, we find that

$$W_{\bar{t}}^t \le W_t^t \exp((r+\gamma/K)(\bar{t}-t)) < U_t^* \exp((r+\gamma/K)(\bar{t}-t)) = U_{\bar{t}}^*$$

with positive probability, a contradiction of the fact that $W_{\bar{t}}^t = U_{\underline{t}}^*$ a.s. So it must be that $W_t^t \ge U_t^*$ for all $t \in [\underline{t}, \bar{t})$ a.s. But note that also $\mathbb{E}[\widetilde{x}_s] = x_s^*$ for all s implies $\mathbb{E}[W_t^t] = U_t^*$ for all t. Hence it must be that $W_t^t = U_t^*$ for all $t \in [\underline{t}, \bar{t})$ a.s.

Now fix $t, t' \in [\underline{t}, \bar{t})$ such that $t' > t$. The fact that $W_{t'}^{t'} = U_{t'}^*$ and the latter is non-random implies by the law of iterated expectations that $W_{t'}^t = U_{t'}^*$. Since this holds for all t' sufficiently close to t, it must be that $\dot{W}_t^t = \dot{U}_t^*$ for all $t \in [\underline{t}, \bar{t})$ a.s. Then

$$\widetilde{x}_t = \frac{1}{K-1}\left((r+\gamma)W_t^t - \dot{W}_t^t\right) = \frac{1}{K-1}\left((r+\gamma)U_t^* - U_t^*\right) = x_t^*$$

for all $t \in [\underline{t}, \bar{t})$ a.s., establishing the result.

E.5 Proof of Lemma 11

Fix a distribution function F with support on $[\underline{t}, \bar{t}]$. Under F, the principal's posterior belief at time π_t for $t \in [\underline{t}, \bar{t}]$ is given by Bayes' rule as

$$\pi_t = \frac{q'_{\underline{t}}}{q'_{\underline{t}} + (1 - q'_{\underline{t}}) \left(\int_{-\infty}^{t} dF(\tau^\beta) \exp\left(-\gamma(t - \tau^\beta)\right) + (1 - F(t)) \right)}.$$

Meanwhile, q'_t is similarly characterized as the posterior beliefs under the (deterministic) undermining process β^*, which by Bayes' rule are

$$q'_t = \frac{q'_{\underline{t}}}{q'_{\underline{t}} + (1 - q'_{\underline{t}}) \exp\left(-\gamma \int_{\underline{t}}^{t} \beta^*_s \, ds\right)}.$$

Equating these two expressions reveals that $\pi_t = q'_t$ for $t \in [\underline{t}, \bar{t}]$ if

$$\int_{-\infty}^{t} dF(\tau^\beta) \exp\left(-\gamma(t - \tau^\beta)\right) + (1 - F(t)) = \exp\left(-\gamma \int_{\underline{t}}^{t} \beta^*_s \, ds\right).$$

Both sides of this expression equal 1 when $t = \underline{t}$, as $F(t) = 0$ for $t < \underline{t}$, and so the integral on the lhs evaluated at $t = \underline{t}$ is precisely $F(\underline{t})$. So the two expressions are equal everywhere on $[\underline{t}, \bar{t}]$ if their derivatives are equal everywhere on this interval. Differentiating both sides gives the required distribution function

$$F^*(t) = 1 - (1 - \beta^*_t) \exp\left(-\gamma \int_{\underline{t}}^{t} \beta^*_s \, ds\right).$$

As β^* is a strictly increasing, continuous on $[\underline{t}, \bar{t}]$, F^* is strictly increasing and continuous on $[\underline{t}, \bar{t}]$ and satisfies $F^*(\underline{t}) = \beta^*_{\underline{t}}$. Further, since $\beta^*_{\bar{t}} = 1$, $F^*(\bar{t}) = 1$. Extend F^* to the real line by setting $F^*(t) = 0$ for $t < \underline{t}$ and $F^*(t) = 1$ for $t > \bar{t}$. Then F^* is a well-defined distribution function with support on $[\underline{t}, \bar{t}]$ which is continuous everywhere except at $t = \underline{t}$.

F Proofs for Section 5

F.1 Proof of Lemma 12

We first prove the "only if" direction. Let x be the disclosure path for some information policy μ. By assumption, μ is càdlàg and thus so is x. Let $\mathcal{F}$ denote the filtration generated by μ. By the martingale property, for all $0 \leq s < t$, $2\mu_s - 1 = \mathbb{E}[2\mu_t - 1 | \mathcal{F}_s]$, and by the triangle inequality, $|2\mu_s - 1| \leq \mathbb{E}[|2\mu_t - 1||\mathcal{F}_s]$. Taking expectations at time 0 and using the law of iterated expectations, $\mathbb{E}[|2\mu_s - 1|] \leq \mathbb{E}[|2\mu_t - 1|]$, and thus by definition $x_s \leq x_t$, so x is monotone increasing. By the same logic, we have $\phi \leq x_t$ for all t, and since $\mu_t \in [0,1]$ for all t, $x_t \leq 1$. This concludes the proof for this direction.

Next, consider the "if" direction. Suppose x is càdlàg, monotone increasing and $[\phi, 1]$-valued; we show that there exists an information policy μ for which x is the disclosure path. We make use of Capasso and Bakstein (2005, Theorem 2.99).

Define $\mathcal{T} \equiv \{0-\} \cup [0, \infty)$ and $x_{0-} \equiv \phi$ and adopt the convention that $0- < t$ for all $t \geq 0$. We construct functions $q(t_1, t_2, a, Y)$ defined for $t_1, t_2 \in \mathcal{T}$ with $t_1 < t_2$, $a \in \mathbb{R}$ and $Y \in \mathcal{B}(\mathbb{R})$, which will be the desired transition probabilities for μ.

Consider any pair of times (t_1, t_2) with $0- \leq t_1 < t_2$. We begin by constructing q for cases where Y is a singleton. Define $q(t_1, t_2, a, \{y\})$ as follows. If $x_{t_1} = 1$, set $q(t_1, t_2, 0, \{0\}) = q(t_1, t_2, 1, \{1\}) = 1$ and set $q(t_1, t_2, a, \{y\}) = 0$ for all other (a, y) pairs. If $x_{t_2} = 0$, set $q(t_1, t_2, a, \{y\}) = 1$ if $a = y = \frac{1}{2}$ and set $q(t_1, t_2, a, \{y\}) = 0$ for all other (a, y) pairs. Next consider $x_{t_1} < 1$ and $x_{t_2} > 0$. For all $t \in \mathcal{T}$, define

$$m_t^{\pm} \equiv \frac{1 \pm x_t}{2}.$$

For $a \in \{m_{t_1}^+, m_{t_1}^-\}$, set $q(t_1, t_2, a, \{m_{t_2}^+\}) = \frac{a - m_{t_2}^-}{m_{t_2}^+ - m_{t_2}^-}$ and $q(t_1, t_2, a, \{m_{t_2}^-\}) = \frac{m_{t_2}^+ - a}{m_{t_2}^+ - m_{t_2}^-}$, making use of the fact that $m_{t_2}^+ - m_{t_2}^- > 0$ since $x_{t_2} > 0$. Set $q(t_1, t_2, a, \{y\}) = 0$ for all other (a, y) pairs.

Next, we extend this construction to all Borel sets Y. For any Borel set $Y \in \mathcal{B}(\mathbb{R})$, define

$$q(t_1, t_2, a, Y) \equiv \sum_{y \in Y \cap \{m_{t_2}^+, m_{t_2}^-\}} q(t_1, t_2, a, \{y\}). \tag{F.1}$$

Hence for all $a \in \mathbb{R}$, for all $0- \leq t_1 < t_2$, $q(t_1, t_2, a, \cdot)$ is a probability measure on $\mathcal{B}(\mathbb{R})$. Since $q(t_1, t_2, a, Y)$ is nonzero for at most two values of a, $q(t_1, t_2, \cdot, Y)$ is a Borel measurable function for all $\mathcal{B}(\mathbb{R})$ and all $0- \leq t_1 < t_2$. It remains to show that q satisfies the Chapman-Kolmogorov equation, which here specializes to

$$q(t_1, t_2, a, Y) = \sum_{z \in \{m_{t'}^+, m_{t'}^-\}} (q(t_1, t', a, \{z\}) q(t', t_2, z, Y)) \qquad \forall t_1 < t' < t_2. \tag{F.2}$$

By (F.1), and the fact that by construction the $q(t_1, t_2, a, \cdot)$ are probability measures for all t_1, t_2, a in the domain, it suffices to consider singletons $Y \in \{\{m_{t_2}^+\}, \{m_{t_2}^-\}\}$. Moreover, whenever $a \notin \{m_{t_1}^+, m_{t_1}^-\}$, by construction $q(t_1, t_2, a, Y) = 0$ for all t_1, t_2, a in the domain so both sides of (F.2) vanish and equality holds; we thus consider $a \in \{m^+_{t_1}, m^-_{t_1}\}$. Note that if $m^+_{t'} = m^-_{t'}$, then it must be that these both equal $a = 1/2$, and thus (F.2) reduces to $q(t_1, t_2, 1/2, \{m^+_{t_2}\}) = q(t', t_2, 1/2, \{m^+_{t_2}\})$. By construction, either (a) $x_{t_2} = 0$ so $m^+_{t_2} = 1/2$ and both sides of the equation are 1, or (b) $x_{t_2} > 0$ so $m^+_{t_2} > m^-_{t_2}$ and both sides are $\frac{1/2 - m^-_{t_2}}{m^+_{t_2} - m^-_{t_2}} = 1/2$. We conclude that (F.2) holds if $m^+_{t'} = m^-_{t'}$. Now whenever $m^+_{t'} \neq m^-_{t'}$, we have $m^+_{t_2} \neq m^-_{t_2}$ and

$$
\begin{aligned}
\sum_{z \in \{m^+_{t'}, m^-_{t'}\}} \big(q(t_1, t', a, \{z\}) q(t', t_2, z, \{m^+_{t_2}\})\big) &= \frac{a - m^-_{t'}}{m^+_{t'} - m^-_{t'}} \frac{m^+_{t'} - m^-_{t_2}}{m^+_{t_2} - m^-_{t_2}} + \frac{m^+_{t'} - a}{m^+_{t'} - m^-_{t'}} \frac{m^-_{t'} - m^-_{t_2}}{m^+_{t_2} - m^-_{t_2}} \\
&= \frac{a - m^-_{t_2}}{m^+_{t_2} - m^-_{t_2}} \\
&= q(t_1, t_2, a, \{m^+_{t_2}\}),
\end{aligned}
$$

as desired. By Capasso and Bakstein (2005, Theorem 2.99), there exists a unique Markov chain μ with q as its transition probabilities. By construction, μ is a martingale and x is its disclosure path.

F.2 Proof of Proposition 4

By Lemma 12, there exists a unique deterministic information policy μ^* for the optimal disclosure path x^* in each of the parametric cases. We now consider the three items of the proposition. Suppose that $x^*_t > x^*_{t-}$. Recall that μ^* is a martingale and has left limits, so $\mu^*_{t-} = \mathbb{E}_{t-}[\mu^*_t] \equiv \lim_{t' \uparrow t} \mathbb{E}_{t'}[\mu^*_t]$. Since μ^*_t takes values $m^\pm_t$ defined in the proof of Lemma 12, it suffices to consider binary signals $S_t \in \{L, R\}$ and without loss label these signals so that, given any μ^*_{t-}, $\mu^*_t = m^+_t$ if and only if $S_t = R$. For $\omega' \in \{L, R\}$, let $p_t(\omega')$ be the probability that $S_t = R$ conditional on information up to time t and conditional on $\omega = \omega'$. By Bayes' rule, we have two equations in two unknowns:

$$
\begin{aligned}
m^+_t &= \frac{\mu^*_{t-} p_t(R)}{\mu^*_{t-} p_t(R) + (1 - \mu^*_{t-}) p_t(L)} \\
m^-_t &= \frac{\mu^*_{t-}(1 - p_t(R))}{\mu^*_{t-}(1 - p_t(R)) + (1 - \mu^*_{t-})(1 - p_t(L))}.
\end{aligned}
$$

The solution is

$$p_t(R) = \frac{m_t^+(\mu_{t-}^* - m_t^-)}{\mu_{t-}^*(m_t^+ - m_t^-)} = \frac{(1+x_t^*)(2\mu_{t-}^* - 1 + x_t^*)}{4\mu_{t-}^* x_t^*}$$
$$p_t(L) = \frac{(1-m_t^+)(\mu_{t-}^* - m_t^-)}{(1-\mu_{t-}^*)(m_t^+ - m_t^-)} = \frac{(1-x_t^*)(2\mu_{t-}^* - 1 + x_t^*)}{4(1-\mu_{t-}^*) x_t^*}.$$

The agent's state conjecture is correct if $\mu_{t-}^* > \frac{1}{2}$ and $\omega = R$ or if $\mu_{t-}^* < \frac{1}{2}$ and $\omega = L$. By symmetry, suppose $\mu_{t-}^* < \frac{1}{2}$. If the agent's state conjecture is correct, then $\omega = L$ and he gets a contradictory signal with probability $p_t(L) = \frac{(1-x_t^*)(x_t^* - x_{t-}^*)}{2x_t^*(1+x_{t-}^*)}$. If instead the agent's state conjecture is incorrect, then $\omega = R$ and he gets a contradictory signal with probability $p_t(R) = \frac{(1+x_t^*)(x_t^* - x_{t-}^*)}{2x_t^*(1-x_{t-}^*)}$.

Next, suppose that $\dot{x}_t^* > 0$. Since beliefs move deterministically conditional on no jump occurring, we consider Poisson signals. Since beliefs move across $\frac{1}{2}$ when a jump occurs, we consider contradictory Poisson learning, that is, Poisson processes with higher arrival rate when the agent's current state conjecture is incorrect. Let $\overline{\lambda}_t, \underline{\lambda}_t$ be the arrival rates conditional on the current action being incorrect or correct, respectively. Consider $\mu_t^* > 1/2$. Then by Bayes' rule, absent an arrival,

$$\begin{aligned}\dot{x}_t^* &= 2\dot{\mu}_t^* = 2\mu_t^*(1-\mu_t^*)(\overline{\lambda}_t - \underline{\lambda}_t) \\ &= (1+x_t^*)\frac{1-x_t^*}{2}(\overline{\lambda}_t - \underline{\lambda}_t).\end{aligned} \tag{F.3}$$

In addition, conditional on an arrival, the belief updates from $\mu_t^* = \frac{1+x_t^*}{2}$ to $m_t^- = \frac{1-x_t^*}{2}$, so by Bayes' rule,

$$\frac{1-x_t^*}{2} = \frac{\frac{1+x_t^*}{2}\underline{\lambda}_t}{\frac{1+x_t^*}{2}\underline{\lambda}_t + \frac{1-x_t^*}{2}\overline{\lambda}_t}. \tag{F.4}$$

The unique solution to the linear system in (F.3) and (F.4) is $\overline{\lambda}_t$, $\underline{\lambda}_t$ as stated in the proposition. A symmetric argument shows that the same arrival rates apply when $\mu_t^* < 1/2$. To obtain the expected arrival rate, again assume wlog that $\mu_t^* > 1/2$ and recall that $\dot{x}_t^* = x_t^*(r + \gamma/K)$; by direct computation,

$$\begin{aligned}\mu_t^*\underline{\lambda}_t + (1-\mu_t^*)\overline{\lambda}_t &= \frac{1+x_t^*}{2}\frac{\dot{x}_t^*}{2x_t^*}\frac{1-x_t^*}{1+x_t^*} + \frac{1-x_t^*}{2}\frac{\dot{x}_t^*}{2x_t^*}\frac{1+x_t^*}{1-x_t^*} \\ &= \frac{1}{2}(r+\gamma/K).\end{aligned}$$

The case where $\dot{x}_t^* = 0$ and $x_{t-}^* = x_t^*$ is trivial.

F.3 Proof of Lemma 13

Fix any information policy, and let μ and σ^2 be the induced mean and variance processes for the agent's posterior beliefs. By the law of iterated expectations, μ must be a martingale. Let $\mu_t^{(2)} = \mathbb{E}_t[\omega^2]$ be the uncentered second moment of the agent's time-t beliefs about ω. Then the time-s posterior variance of the agent's beliefs may be written

$$\sigma_t^2 = \mu_t^{(2)} - \mu_t^2.$$

Taking time-s expectations of both sides for any $s < t$ and using the law of iterated expectations and Jensen's inequality yields

$$\mathbb{E}_s[\sigma_t^2] = \mu_s^{(2)} - \mathbb{E}_s\left[(\mathbb{E}_t[\omega])^2\right] \leq \mu_s^{(2)} - \mu_s^2 = \sigma_s^2.$$

So σ^2 must be a supermartingale. Taking time-zero expectations of both sides establishes that x must be monotone increasing, with in particular $x_t \geq x_{0-} = \phi$. The upper bound $x_t \leq 1$ follows from the fact that σ^2 is a non-negative process. Finally, as both μ and $\mu^{(2)}$ are martingales on a filtration satisfying the usual conditions, σ^2 is càdlàg (possibly by passing to an appropriate version). Meanwhile x is a monotone function, thus has well-defined limits from both sides everywhere. And by Fatou's lemma, $\mathbb{E}[\sigma_t^2] \leq \liminf_{s\downarrow t} \mathbb{E}[\sigma_s^2]$ for every t. Hence $x_t \geq x_{t+}$, and since x is monotone this implies $x_t = x_{t+}$. So x is càdlàg.

F.4 Proof of Lemma 14

In the high-stakes case, x^* can be implemented via disclosure of ω at time t^*.

In the moderate-stakes case, x^* is implemented by a combination of two signals. First, over the time inteval $[\underline{t}_M, \overline{t}_M]$ the agent observes a signal process ξ distributed as

$$d\xi_t = \omega\, dt + (1 - x_t^*)\sqrt{\frac{C}{(r + \gamma/K)x_t^*}}\, dB_t,$$

where B is a standard Brownian motion independent of ω and the exogenous signal. Second, at time $\overline{t}_M$ the agent observes ω. Under standard results on Brownian filtering (for example, see Liptser and Shiryaev (2013)), observing the signal process ξ induces a sequence of normal posteriors with variance process σ^2 evolving as

$$\frac{d\sigma^2}{dt} = -\sigma_t^4(1 - x_t^*)^{-2}\frac{(r + \gamma/K)x_t^*}{C},$$

with initial condition $\sigma^2_{\underline{t}_M} = \eta^2$. Note that $\sigma^2_t = C(1-x^*_t)$ satisfies this ODE with the correct initial condition, given that $\dot{x}^*_t = (r+\gamma/K)x^*_t$ on $[\underline{t}_M, \overline{t}_M)$. So the agent's posterior variance evolves as $\sigma^2_2 = C(1-x^*_t)$ on this interval. Meanwhile on $[0, \underline{t}_M]$ the agent's posterior variance is fixed at η^2, while on $[\overline{t}_M, \infty)$ the posterior variance is 0. Thus the disclosure path induced by this information policy is indeed x^*, and the policy is deterministic.

Finally, in the low-stakes case x^* is implemented by a combination of three signals. First, at time zero the agent observes a signal distributed as $\mathcal{N}\left(\omega, ((C(1-\underline{x}))^{-1} - \eta^{-2})^{-1}\right)$, independent of the exogenous signal conditional on ω. Second, over the time interval $[0, \overline{t}_L]$ the agent observes a signal process ξ distributed as in the moderate-stakes case, with the Brownian motion independent of the time-zero signal. Third, at time $\overline{t}_L$ the agent observes ω. Bayesian updating implies that after observing the time-zero signal, the agent's posterior beliefs are normally distributed with variance $\sigma^2_0 = C(1-\underline{x})$. Meanwhile on $[0, \overline{t}_L]$ the work for the moderate-stakes case shows that the agent's posterior beliefs have variance $\sigma^2_t = C(1-x^*_t)$. Finally, the agent's posterior variance on $[\overline{t}_L, \infty)$ is 0. Thus the disclosure path induced by this information policy is x^*, and the policy is deterministic.

G Proofs for Section 6

In this section we provide comparative statics results for the low stakes and high stakes cases and prove a proposition which covers all three cases.

Table 2: Comparative statics (high stakes)

	$dt^*/$	$d\Pi/$
$d\gamma$	$-$	$+$
dq	$-$	$+$
dK	$+$	$-$
dr	0	$-$
$d\phi$	0	$-$

Proposition G.1. *In the low, moderate and high stakes cases, the relationships between model inputs and outputs are as given in Tables 1, 2 and 3.*

Proof of Proposition G.1. We proceed one output at a time. The analysis of $\overline{x}$ applies to both the low and moderate stakes cases.[36] The analysis of the principal's profit — with the

[36]This value has a single definition, and the arguments used in signing its derivatives do not use any additional facts about parameter values specific to the low or moderate stakes cases.

Table 3: Comparative statics (low stakes)

	$d\overline{x}/$	$d\underline{x}/$	$d\overline{t}_L/$	$d\Pi/$
$d\gamma$	$+$	$+$	$-$	$+$
dq	0	$+$	$-$	$+$
dK	$+$	$-$	$+$	$-$
dr	$-$	$-$	0	$-$
$d\phi$	0	0	0	0

exception of comparative statics with respect to r, which we treat separately in Lemma G.1 — applies to all three cases, and is provided last.

- t^* (Table 2) and $\overline{x}$ (Tables 1 and 3): The signs of these derivatives are immediate from inspection of the respective formulas given in Section D.4.

- $\underline{x}$ (Tables 1 and 3): Recall that $\underline{x} = \frac{(K-1)\gamma}{K(\gamma+r)}\left(\frac{Kq}{K-1}\right)^{1+\frac{Kr}{\gamma}}$, which is clearly increasing in q and independent of ϕ. The first factor is decreasing in r and increasing in γ, as is the second factor since $q < \frac{K-1}{K}$, which implies $d\underline{x}/dr < 0 < d\underline{x}/d\gamma$. By direct calculation, $d\underline{x}/dK$ has the same sign as $-1 + (K-1)\ln\left(\frac{Kq}{K-1}\right) < -1$.

- Δ (Table 1): From $\Delta = \left(r + \frac{\gamma}{K}\right)^{-1} \ln\left(\frac{\overline{x}}{\phi}\right)$, it is immediate that $d\Delta/d\phi < 0 = d\Delta/dq$. Now both factors are increasing in K, so unambiguously $d\Delta/dK > 0$. By direct calculation, $d\Delta/dr = -\frac{K}{(r+\gamma)(Kr+\gamma)^2}\left[Kr + \gamma + K(r+\gamma)\ln\left(\frac{\overline{x}}{\phi}\right)\right]$. For $\phi < \overline{x}$, all terms in square brackets are positive, so $d\Delta/dr < 0$. We have $d\Delta/d\gamma = \frac{K}{(Kr+\gamma)^2}\left[\frac{r(Kr+\gamma)}{\gamma(r+\gamma)} - \ln\left(\frac{\overline{x}}{\phi}\right)\right]$; for ϕ sufficiently close to $\overline{x}$, this is strictly positive. However, evaluating at $\phi = \underline{x}$ yields $\frac{K\left[r+(r+\gamma)\ln\left(\frac{Kq}{K-1}\right)\right]}{\gamma(r+\gamma)(Kr+\gamma)}$, which is strictly negative for r sufficiently small. We conclude that $d\Delta/d\gamma$ can be positive or negative.

- $\underline{t}_M$ (Table 1): First, we have $d\underline{t}_M/dq = -\frac{1}{q(1-q)\gamma} < 0$. Next, define locally $\alpha \equiv \left(\frac{\overline{x}}{\phi}\right)^{\frac{\gamma}{Kr+\gamma}}$, where $\alpha > 1$ when the moderate stakes case applies (since $\phi > \overline{x}$). By direct calculation,

$$
\begin{aligned}
d\underline{t}_M/dK &= \frac{Kr\alpha[1+(K-1)\ln\alpha]}{(K-1)\gamma(Kr+\gamma)[1-K+K\alpha]} > 0 \\
d\underline{t}_M/dr &= \frac{K\alpha[1+K(r/\gamma+1)\ln\alpha]}{(r+\gamma)(Kr+\gamma)[1-K+K\alpha]} > 0 \\
d\underline{t}_M/d\phi &= \frac{K\alpha}{\phi(Kr+\gamma)[1-K+K\alpha]} > 0 \\
d\underline{t}_M/d\gamma &= \gamma^{-2}[\ln(X)-Y] < 0, \qquad \text{where} \\
X &\equiv \frac{q(1-K+K\alpha)}{(1-q)(K-1)} \\
Y &\equiv \frac{Kr\gamma\alpha[Kr+\gamma+K(r/\gamma+1)\ln\alpha]}{(r+\gamma)(Kr+\gamma)[1-K+K\alpha]}.
\end{aligned}
$$

Now $\ln(X) \le \ln\left[\frac{q}{(1-q)(K-1)}\left(1-K+K\left[\frac{\overline{x}}{\underline{x}}\right]^{\frac{\gamma}{Kr+\gamma}}\right)\right] = 0$, and by inspection $Y > 0$ so $d\underline{t}_M/d\gamma < 0$, as desired.

- $\overline{t}_M$ (Table 1): Since $d\Delta/dq = 0$, $d\overline{t}_M/dq = d\underline{t}_M/dq < 0$. Also, $d\overline{t}_M/dK = d\Delta/dK + d\underline{t}_M dK > 0$ unambiguously. By direct calculation, using the definition of α from above,

$$
\begin{aligned}
d\overline{t}_M/dr &= -\frac{(K-1)K(\alpha-1)[1+K(1+r/\gamma)\ln\alpha]}{(r+\gamma)(Kr+\gamma)[1-K+K\alpha]} < 0 \\
d\overline{t}_M/d\phi &= -\frac{(K-1)K(\alpha-1)}{\phi(Kr+\gamma)[1-K+K\alpha]} < 0.
\end{aligned}
$$

Finally, by tedious calculation, we have $\frac{d^2\overline{t}_M}{dqd\gamma} = \frac{1}{\gamma^2 q(1-q)} > 0$. For an upper bound on $d\overline{t}_M/d\gamma$, we increase q until $\underline{x} = \phi$. We obtain

$$
\begin{aligned}
d\overline{t}_M/d\gamma|_{\phi=\underline{x}} &= \frac{(K-1)\left[-\left(\frac{Kq}{K-1}-1\right)r\gamma + \frac{K}{K-1}(r+\gamma)[r+\gamma(1-q)]\ln\left(\frac{Kq}{K-1}\right)\right]}{(1-q)\gamma^2(r+\gamma)(Kr+\gamma)} \\
&< \frac{(K-1)\left(\frac{Kq}{K-1}-1\right)}{(1-q)\gamma^2(r+\gamma)(Kr+\gamma)}\left[-r\gamma + \frac{K}{K-1}(r+\gamma)[r+\gamma(1-q)]\right] \\
&< 0,
\end{aligned}
$$

where we have used $q < \frac{K-1}{K}$ and the inequality $\ln x < x-1$ for $x > 0$.

- $\overline{t}_L$ (Table 3): By inspection, $\overline{t}_L$ is independent of r and ϕ and is decreasing in γ and q. For K, Calculate $d\overline{t}_L/dK = \frac{1}{\gamma}\left[\frac{1}{(K-1)} - \ln\left(\frac{Kq}{K-1}\right)\right] > \frac{1}{(K-1)\gamma} > 0$.

- Π (Tables 1, 2 and 3): Recall that $\mathbb{X}$ is the set of monotone, càdlàg $[\phi, 1]$-valued functions x and let $\mathbb{B}$ denote the set of pure strategies β for the disloyal agent. The principal's payoff is the maximum of 0 and

$$\max_{x\in\mathbb{X}}\min_{b\in\mathbb{B}}\int_0^\infty x_t e^{-rt}\left[q-(1-q)(\beta_t K-1)\exp\left(-\gamma\int_0^t \beta_s ds\right)\right]dt. \tag{G.1}$$

Suppose that the expression in (G.1) is strictly positive (the other case is trivial). Now $\mathbb{X}$ and $\mathbb{B}$ are independent of q, K and γ, while the integrand is increasing in q and decreasing in K and γ, so the value in (G.1) inherits these properties. On the other hand, ϕ only enters via the principal's choice set $\mathbb{X}$. In the low stakes case, the constraint $x_0 \geq \phi$ does not bind, and the principal is unaffected by an increase in ϕ; but in the high and moderate stakes cases, this constraint does bind, so she is made worse off by an increase in ϕ. Finally, the result with respect to r is proved separately in Lemma G.1 below.

$\square$

Lemma G.1. *In the low, moderate, and high stakes cases, the principal's normalized payoff is decreasing in the discount rate.*

Proof. We begin with the high stakes case. Define $z \equiv \frac{q}{(K-1)(1-q)}$, and note that $z \in (0,1)$ as $q \in \left(0, \frac{K-1}{K}\right)$. The principal's normalized payoff is

$$\begin{aligned}\Pi_H &= \int_0^{t^*}\phi\left[q-(1-q)(K-1)e^{-\gamma t}\right]re^{-rt}dt + \int_{t^*}^\infty \left[q-(1-q)(K-1)e^{-\gamma t}\right]re^{-rt}dt\\ &= \frac{r\phi[Kq-(K-1)]+q\gamma[z^{r/\gamma}(1-\phi)+\phi]}{r+\gamma}.\end{aligned}$$

Differentiating w.r.t. r yields

$$\frac{d\Pi_H}{dr} = \frac{-\gamma[q(1-\phi)z^{r/\gamma}+(k-1)\phi(1-q)]+qz^{r/\gamma}(r+\gamma)(1-\phi)\ln(z)}{(r+\gamma)^2}.$$

The expression in square brackets is strictly positive and the last term of the numerator is strictly negative, so the full expression is unambiguously strictly negative.

For the low stakes case, the principal's normalized payoff is

$$
\begin{aligned}
\Pi_L &= \int_0^{\bar{t}_L} \underline{x} e^{(r+\gamma/k)t} \left[q-(1-q)(K-1)e^{-\gamma t}\right] re^{-rt} dt \\
&\quad + \int_{\bar{t}_L}^{\infty} \left[q-(1-q)(K-1)e^{-\gamma t}\right] re^{-rt} dt \\
&= \frac{q\gamma\left(\frac{Kq}{K-1}\right)^{Kr/\gamma}}{r+\gamma}.
\end{aligned}
$$

Now the numerator is decreasing in r as $\frac{Kq}{K-1}<1$ and the denominator is increasing in r, so we conclude that $d\Pi_L/dr<0$.

For the moderate stakes case, we have

$$
\begin{aligned}
\Pi_M &= \int_0^{\underline{t}_M} \phi \left[q-(1-q)(K-1)e^{-\gamma t}\right] re^{-rt} dt \\
&\quad + \int_{\underline{t}_M}^{\bar{t}_M} \phi e^{(r+\gamma/K)(t-\underline{t}_M)} \left[q-(1-q)(K-1)e^{-\gamma t}\right] re^{-rt} dt \\
&\quad + \int_{\bar{t}_M}^{\infty} \left[q-(1-q)(K-1)e^{-\gamma t}\right] re^{-rt} dt \\
&= \frac{-\phi r(r+\gamma)(K-1)+q(Kr+\gamma)\left[\gamma\alpha^{-Kr/\gamma}X^{r/\gamma}-(r+\gamma)\left(X^{r/\gamma}-1\right)\right]}{(r+\gamma)^2},
\end{aligned}
\tag{G.2}
$$

where $\alpha \equiv \left(\frac{\underline{x}}{\phi}\right)^{\frac{\gamma}{Kr+\gamma}}$ and $X \equiv \frac{q(1-K+K\alpha)}{(1-q)(K-1)}$. Note that in the moderate stakes case, $\alpha \in \left(1, \frac{K-1}{Kq}\right]$ and $X \in \left(\frac{q}{(1-q)(K-1)}, 1\right]$.

Differentiating (G.2), we obtain

$$
\begin{aligned}
d\Pi_M/dr &= -\frac{g(X;r,\gamma,q,K,\phi)}{(r+\gamma)^3}, \qquad \text{where} \\
g(\hat{X};r,\gamma,q,K,\phi) &\equiv \phi\gamma(r+\gamma)(K-1)\left(1-q+q\hat{X}^{r/\gamma}\right) \\
&\quad + q(r+\gamma)(Kr+\gamma)\hat{X}^{r/\gamma}\ln(\hat{X})\left(\phi[1+r/\gamma]-\alpha^{-Kr/\gamma}\right) \\
&\quad + q\hat{X}^{r/\gamma}\alpha^{-Kr/\gamma}\left(\gamma[r+\gamma(2-K)]+K[r+\gamma]^2\ln[\alpha]\right).
\end{aligned}
$$

For later purposes, we establish that $\phi[1+r/\gamma]-\alpha^{-Kr/\gamma}<0$. Using the definitions of α and

$\overline{x}$,

$$\begin{aligned}\phi[1+r/\gamma]-\alpha^{-Kr/\gamma} &= \phi\left[\frac{r+\gamma}{\gamma}-\phi^{-\frac{\gamma}{Kr+\gamma}}\overline{x}^{-\frac{Kr}{Kr+\gamma}}\right]\\ &< \phi\left[\frac{r+\gamma}{\gamma}-\overline{x}^{-\frac{\gamma}{Kr+\gamma}}\overline{x}^{-\frac{Kr}{Kr+\gamma}}\right]\\ &= \phi\frac{r+\gamma}{\gamma}\left[1-\frac{K}{K-1}\right]<0.\end{aligned}$$

We now show that for all $\hat{X}\in(0,1]$ (and in particular for $\hat{X}=X$), $g(\hat{X};r,\gamma,q,K,\phi)>0$, establishing that $d\Pi_M/dr<0$. To this end, we show that $g:\hat{X}\mapsto g(\hat{X};r,\gamma,q,K,\phi)$ is quasiconcave and that $\lim_{\hat{X}\downarrow 0} g(\hat{X};r,\gamma,q,K,\phi)>0$ and $g(1;r,\gamma,q,K,\phi)>0$:

- g is quasiconcave in $\hat{X}$: we have

$$\begin{aligned}\frac{\partial}{\partial\hat{X}}g(\hat{X};r,\gamma,q,K,\phi) = q\hat{X}^{\frac{r}{\gamma}-1}\{&\phi r(r+\gamma)(K-1)\\ &+q(r+\gamma)(Kr+\gamma)\left(\phi[1+r/\gamma]-\alpha^{-Kr/\gamma}\right)\left(\ln[\hat{X}]r/\gamma+1\right)\\ &+\alpha^{-Kr/\gamma}(r/\gamma)\left(\gamma[r+\gamma(2-K)]+K[r+\gamma]^2\ln[\alpha]\right)\}.\end{aligned}\tag{G.3}$$

  Observe that in (G.3), the factor outside the braces is positive, and let $g_2(\hat{X};r,\gamma,q,K,\phi)$ denote the expression inside the braces. As argued above, the coefficient on $\ln[\hat{X}]$ in $g_2(\hat{X};r,\gamma,q,K,\phi)$ is negative; hence, as $\hat{X}\downarrow 0$, the $g_2(\hat{X};r,\gamma,q,K,\phi)\uparrow\infty$, and thus for sufficiently small $\hat{X}$, $\frac{\partial}{\partial\hat{X}}g(\hat{X};r,\gamma,q,K,\phi)>0$. Moreover, we have that $g_2(\hat{X};r,\gamma,q,K,\phi)$ is monotonically decreasing in $\hat{X}$, giving us the quasiconcavity of $g(\hat{X};r,\gamma,q,K,\phi)$ w.r.t. $\hat{X}$.

- $\lim_{\hat{X}\downarrow 0} g(\hat{X};r,\gamma,q,K,\phi)>0$: By taking limits directly as $\hat{X}\downarrow 0$ and applying L'Hôpital's rule to obtain $\hat{X}^{r/\gamma}\ln[\hat{X}]\uparrow 0$, we have

$$\lim_{\hat{X}\downarrow 0} g(\hat{X};r,\gamma,q,K,\phi)=\phi\gamma(r+\gamma)(K-1)(1-q)>0.$$

- $g(1;r,\gamma,q,K,\phi)>0$: By plugging in $\hat{X}=1$ and simplifying, we get

$$g(1;r,\gamma,q,K,\phi)=(K-1)\gamma(r+\gamma)\phi+q\alpha^{-Kr/\gamma}\left(\gamma[r+(2-K)\gamma]+K[r+\gamma]^2\ln[\alpha]\right).\tag{G.4}$$

  Observe that α is independent of q, and thus the right hand side of (G.4) is linear in q. To show that it is positive, it suffices to show positivity for extreme values of q.

For the moderate stakes case, $\overline{x}$ is independent of q so $\phi < \overline{x}$ places no tighter lower bound on q besides $q \geq 0$. But $\underline{x} \leq \phi$ implies $q \leq \overline{q} \equiv \frac{K-1}{K}\left(\frac{\phi}{\overline{x}}\right)^{\frac{\gamma}{\gamma+Kr}}$. Evaluating at these extremes, we have

$$
\begin{aligned}
g(1; r, \gamma, 0, K, \phi) &= (K-1)\gamma(r+\gamma)\phi > 0 \\
g(1; r, \gamma, \overline{q}, K, \phi) &= (r+\gamma)^2\phi\left[1 + K(1 + r/\gamma)\ln(\alpha)\right] > 0.
\end{aligned}
$$

Hence, $g(1; r, \gamma, q, K, \phi) > 0$ for all q for which the moderate stakes case applies.

Together, these claims imply that $g(\hat{X}; r, \gamma, q, K, \phi) > 0$ for all $\hat{X} \in (0, 1]$, and thus $d\Pi_M/dr < 0$, concluding the proof. □