Trust with Evidence

Amirreza Ahmadzadeh* †

September 30, 2025

Please click here for the latest version.

We study a dynamic principal–agent relationship in which an agent must exert costly effort to learn a privately observed binary state before taking an action. The principal wants to match the action with the state, while the agent is biased toward one action, generating both a moral hazard (effort choice) and an adverse selection (action choice after learning the state) problem. The principal disciplines the agent through verification (at a cost), reduced workload and termination. We show reduced workload is always a valuable instrument, even when the cost of verification is small and the loss from shirking is large. By promising a reduced workload in the future, the principal can lower verification costs across multiple periods. For high biases, verification and reduced workload are insufficient instruments, and the principal must rely on firing along the equilibrium path. The threat of future firing complements verification and saves verification costs over time.

*Toulouse School of Economics, amirreza.ahmadzadeh@tse-fr.eu

†I am indebted to my advisors (in alphabetical order) Johannes Hörner, Thomas Mariotti, Anna Sanktjohanser, and Jean Tirole for their support and guidance. Part of this research was conducted while I was visiting the Department of Economics at Yale University under the supervision of Marina Halac, to whom I owe special thanks for her guidance and support. I also thank Alessandro Bonatti, Roberto Corrao, Joyee Deb, Daniel N. Hauser, Elliot Lipnowski, Paula Onuchic, Daniel Rappoport, François Salanié, Philipp Strack, Juuso Välimäki, Stephan Waizmann, Kai Hao Yang as well as audiences at the TSE Theory workshop, the Yale Micro lunch, SAET Ischia '25 for their comments and feedback. I gratefully acknowledges that this project has received funding from the European Research Council (ERC) under grant 101098319. All errors are my own.

1 Introduction

Delegation often places consequential choices in the hands of an expert who must first create information at a cost. Monitoring this process is expensive; leaving it unchecked is risky, especially when the expert has a stable tilt toward one action. I ask how a principal should govern a long-run relationship when preferences are misaligned along both effort provision and action choice.

Consider a large social platform and a human content reviewer. Each flagged item either violates policy or does not, but determining which is costly: the reviewer must read context and apply detailed rules. Operational incentives can induce a known, stable tilt toward either "remove" or "keep," independent of the realized state. The platform can verify decisions via QA second looks (at a cost), adjust the reviewer's workload over time (routing easier or harder queues, pacing), and – at the limit – terminate the employment relation. The question is how to combine these instruments intertemporally to induce effort and keep actions aligned with the underlying state.

Specifically, I study a dynamic principal–agent relationship in which the agent must exert costly effort to learn a binary state before taking an action. The agent is biased toward one action, regardless of the state – so preferences are misaligned in two dimensions: the agent dislikes exerting effort and favors one action over the other. Wages are fixed, so principal must manage the relationship using three instruments: at a cost, she can verify whether the agent has actually learnt the state (and what the state is), she can adjust the agent's workload over time, and she can terminate the relationship (thereby ceasing to pay the agent his fixed wage). My focus is on how the principal combines these tools to discipline the agent and sustain good decisions over the long run.

I fully characterize the principal-preferred equilibrium. "Trust" (the agent's promised continuation utility) initially grows whenever verification occurs and incentives are primarily provided by adjusting verification probabilities. As trust rises, the agent's less-preferred action is no longer verified; rewards and punishments are then delivered through future promises rather than audits. If a string of the agent's less-preferred actions is observed, quasi-retirement is triggered: limited shirking is tolerated to deliver promised utility. If instead a string of the agent's more-preferred actions is observed, the relationship returns to the initial, audit-led stage.

Three features stand out. First, regardless of verification costs or the magnitude of bias, quasi-retirement is eventually used as an incentive instrument. Second, when the agent's bias is large, a streak of unfavorable states (with the agent taking the corresponding action) leads to firing – even without misconduct. Third, without commitment the principal may resort to firing but never retires the agent; with commitment, the principal always retires the agent with positive probability.

Verification is a tool for disciplining the agent to both work and to match the action with the state. When the principal learns that the agent either did not work or mismatched the action with the state, she fires him. But what happens if verification repeatedly confirms that the agent has complied? Should the agent not be rewarded after several rounds in which he consistently demonstrates compliance? When verification is an efficient tool, should the principal rely entirely on it—disciplining the agent to always comply regardless of how many times he has shown his compliance?

Allowing the agent to shirk after some rounds of verification, when the agent has always complied (works, learns the state, and matches the action to the state), is always beneficial for the principal. By promising a lighter workload in the future, the principal can lower verification costs across multiple periods leading up to the present. We refer to this phenomenon as *snowballing*. In each period in which the agent is closer to being allowed to shirk, the principal can decrease the intensity of verification further. The reason is that the threat of firing is relatively stronger as the agent is closer to shirking; hence, the intensity can be decreased. The accumulation of savings on verification costs over multiple periods always dominates the loss from allowing the agent to shirk and potentially mismatching the action. This is regardless of how small the verification cost is and how big the loss from shirking is. However, the number of rounds in which verification must occur such that the agent can shirk grows as the cost of verification becomes smaller. In addition to *snowballing*, allowing the agent to shirk has another benefit. When the principal knows the agent is not working, verification is unnecessary. It entails cost but provides no additional information.

If the agent deviates by not exerting effort or by mismatching the action to the state, the principal fires him after she detects the deviation through verification. But could it ever be optimal to fire an agent who has always complied? Since termination is ex post inefficient, it eliminates all future surplus from the relationship, the question is nontrivial. When verification is an efficient, low-cost instrument, might the principal instead raise the intensity rather than rely on (on the equilibrium path) firing? More generally, conditional on no deviation by the agent, does verification always dominate termination?

The agent is eventually fired when his bias is large relative to the cost of effort. To explain the intuition of this result is important to first understand the verification strategy of the principal. In this case, the principal does not verify after the agent takes his less-preferred action. Generally verification after this action is unnecessary, except to ensure that the agent is not repeatedly shirking and choosing his less-preferred action. Once the agent's utility exceeds the utility from such a deviation (which is the case when bias is large relative to the cost of effort), the principal no longer verifies after this action. In order to discipline the agent

to match the action to the state, the principal relies primarily on differentiating verification probabilities, which is more efficient than differentiating continuation (next-period) utilities. However, when the verification probability after the less-preferred action is zero, continuation utilities must also be dispread to incentivize the agent to match the action to the state. In this case, when verification does not occur, the agent's continuation utility rises if he chooses the less-preferred action and falls if he chooses the more-preferred action. An increase in utility leads to weaker verification in the next period and moves him closer to being allowed to shirk occasionally; a decrease leads to more intensive verification and brings him closer to potential termination. After several rounds in which the agent chooses his more-preferred action and the principal does not verify, the utility decreases and ultimately the agent is fired.

When the cost of working is large relative to the bias, the relationship never terminates. In this case, the principal must verify after both actions to ensure that the agent works. However, verification after the agent's preferred action is always more intense because it increases the cost of mismatching the action. Since verification after both actions is positive, the principal can flexibly adjust their intensities to induce correct state–action matching after learning the state. Hence, she does not need to rely on differentiated continuation utilities. This increase in verification therefore substitutes for firing, cancels its snowballing effect, and, as a consequence, the agent is never fired.

Despite the fact that higher bias eventually terminates the relationship, counterintuitively, the principal may benefit from it. If the bias is large, when the agent is allowed to shirk, he can choose his preferred action. In this case, although bias makes it harder to enforce action–state matching when the agent works, it simultaneously increases the agent's motivation to work and to demonstrate that the state is his preferred one. The benefit comes from two sources: first, the agent's reward from shirking and choosing the preferred action becomes stronger, which allows the principal to sustain working over a larger interval of utilities; and second, the principal's concern about shirking diminishes, and she can lower the intensity of verification.

In the model where the principal does not have commitment power, she resorts to firing when the bias is large, but never retires (shirking forever) the agent. In contrast, if the principal could commit to her strategy, she would always retire the agent with positive probability. Snowballing operates: allowing the agent to permanently shirk reduces verification costs over multiple periods. Although both firing and retirement are ex post inefficient for the principal, firing can arise in equilibrium under no commitment whereas retirement cannot. If the principal expects that the agent does not work, she fires him. Conversely, if the agent expects that the principal fires him at every opportunity, not working is optimal.

Retirement is not part of the equilibrium since at the time of planned retirement the principal reneges.

On the technical side, the history of the game can be fully summarized by the agent's continuation utility. The upper boundary of the equilibrium payoff set is self-generating in the sense of Abreu, Pearce, and Stacchetti (1990), which allows us to formulate the principal's problem as a recursive program subject to incentive constraints. By introducing endogenous variables for verification, firing, and working, we obtain a tractable formulation that highlights the distinct role of each instrument in sustaining equilibrium.

Roadmap. Section 2 situates the paper within the existing literature. Section 3 introduces the model and the equilibrium concept. Section 4 presents the main results and builds intuition through a simple three-state automaton. Section 5 develops the recursive formulation, characterizes the upper boundary of equilibrium payoffs, and states the three main theorems. Section 5.3 provides interpretations and extensions, including the roles of commitment and transfers.

2 Literature Review

Our paper contributes to two strands of the literature. The first strand is dynamic monitoring. The work most closely related to ours is Bhaskar (2024), who study long-term principal–agent relationships with both adverse selection and moral hazard. The agent can work or shirk and has private types (high- or low-cost of effort). The principal monitors effort, using transfers and firing to provide incentives.[1] Monitoring serves to screen out low types by inducing shirking. The optimal self-enforcing contract has two phases: an initial screening phase, where the first monitored action reveals the agent's type, and a continuation phase, in which only high types remain. This second phase resembles our evidence-based structure when the agent is unbiased: effort is exerted until seniority, and monitoring intensity declines with the agent's track record. Unlike Bhaskar (2024), our agent has no persistent type but instead acquires private information about the state through costly effort (covert information acquisition).[2]

Recent work on dynamic mechanism design incorporates various forms of endogenous monitoring. Halac and Prat (2016) and Piskorski and Westerfield (2016), as well as Orlov (2022), study endogenous arrival rates of Poisson signals (good news in the former, bad news in the latter). Fahim et al. (2021) and Zeng (2022)

[1] An earlier version of their paper also considers the case without transfers.

[2] There is a large literature on dynamic moral hazard, but these papers abstract from endogenous monitoring; see, for instance, Ray (2002), Levin (2003), Board (2011), Halac (2012), Nikolowa (2017), Fong (2008), and Guo and Hörner (2021).

allow the principal to choose the precision of Brownian monitoring. Liu (2011) and Marinovic et al. (2018) endogenize the number of past observations, while Varas et al. (2020) focus on the timing of observations. Dai et al. (2024) study the joint design of monitoring and compensation in a continuous-time moral hazard model with transfers. The principal allocates limited monitoring capacity between confirmatory and contradictory evidence of effort. Wong (2023), studies general monitoring schemes under a fixed exogenous wage. He shows that the optimal scheme is non-stationary, focusing on negative evidence of effort with increasing precision but decreasing frequency over time. All of these papers focus on broader aspects of monitoring design, but they do not address the role of bias or covert information acquisition in a dynamic relationship.[3]

The second strand is dynamic delegation. The closest paper in this literature is Lipnowski and Ramos (2020), who study an infinite-horizon game where, in each period, the principal decides whether to delegate a project adoption choice to the agent. The state is binary and i.i.d. across time, while the agent's preferences are state-dependent but biased toward one action. They characterize the set of equilibrium payoffs under fixed discounting and show that, unlike in dynamic agency models with commitment, the agent's autonomy diminishes over time rather than being rewarded through backloading. Li et al. (2017) study a repeated project-selection game in which, each period, a biased agent and an uninformed principal jointly select a project and then simultaneously choose implementation effort. The effort decision effectively grants the principal commitment power, as it provides each player with a tool to unilaterally punish the other. Unlike our paper, in the both paper the principal does not have endogenous verification tool.

[3]Solan and Zhao (2021) study a repeated inspection game between a principal and two agents. In each period, the principal—who has commitment power—can inspect at most one agent, while each agent decides whether to adhere or violate. They show that postponing rewards frees up resources for verification.

3 The Setup

A principal (she) aims to match an action with an unknown state. She employs an agent (he) to learn the state and take an action. They interact repeatedly. Time is discrete, and the horizon is infinite. In each period, $t = 1, 2, ...$, the interaction unfolds as follows.

Timing:

1. The principal chooses whether to fire (F) or retain ($\neg F$) the agent. Firing ends the game.

2. The agent privately chooses to work (W) or shirk ($\neg W$). If the agent works, he privately learns the state θ and has access to a set of evidence information $S = \{\theta, \emptyset\}$. If the agent does not work, then $S = \{\emptyset\}$. The state $\theta \in \{L, R\}$ is i.i.d. across periods and the probability of $\theta = L$ is $\rho \geq 1/2$.[4]

3. The agent publicly takes an action $a \in \{L, R\}$.

4. The principal publicly verifies the evidence (V) or not ($\neg V$). If the principal verifies, the agent discloses an element $s \in S$.[5]

Equivalently, if the principal verifies, she learns whether the agent worked and if so what he learned.

Payoffs: If the agent's action matches the state, the principal gets a flow payoff of $\pi > 0$. Otherwise, she gets 0. If the agent is retained, the principal pays him a per-period fixed wage $w > 0$. Verifying the evidence costs $c_P > 0$ to the principal and working costs $c_A \geq 0$ to the agent. In addition, action R yields a flow benefit $b \geq 0$ to the agent. That is, the agent does not care about the state, and he is biased toward action R.[6] Payoffs are not observed by either player.

Once the agent is fired, if ever, the game ends and both players get 0. Let $\pi(\theta, a) = \pi \mathbb{1}_{\{\theta = a\}}$. Both players share the same discount factor $\delta \in [0, 1)$. Therefore, if the agent is fired in period $T \in \mathbb{N} \cup \{0, +\infty\}$, the principal's payoff is

$$\sum_{t=0}^{T-1} \delta^t \big(\pi(\theta^t, a^t) - w - c_P \mathbb{1}_V\big).$$

[4]The message may be interpreted as a report about the state when the agent works and learns the state. Allowing a richer message space or allowing the agent sends a message $m \in \{\hat{L}, \hat{R}\}$ to the principal does not affect our results.

[5]Note that information acquisition is modeled differently from the literature on costly state verification. Here the principal cannot learn the state directly.

[6]In Section 5.3, we discuss how the results change when the agent is biased toward $a = L$.

Let $u(L) = w$ and $u(R) = w + b$. The agent's payoff is

$$\sum_{t=0}^{T-1} \delta^t (u(a^t) - c_A \mathbb{1}_W).$$

We assume the interaction is mutually beneficial: $\pi \geq w \geq c_A$, i.e., the agent's value to the principal exceeds the wage, and the wage covers the cost of working. We also assume the wage exceeds the principal's payoff when the agent always chooses the more likely state, i.e., $w > \rho\pi$. Of course interesting dynamics arise only if the agent is sufficiently patient: $\frac{\delta}{1-\delta}(w + (1-\rho)b) \geq c_A$, meaning the present value of future benefits exceeds today's working cost.

Throughout, players condition their play on the outcome of a public randomization device (prd). Denote the outcome of the prd regarding firing by $f_\xi \in \{F, \neg F\}$, working by $w_\xi \in \{W, \neg W\}$, action recommendation by $a_\xi \in \{L, R\}$ and verification by $u_\xi \in \{V, \neg V\}$. This ensures that the equilibrium payoff set is convex and so its upper boundary is concave, a property that is used repeatedly.

Solution Concept: The solution concept we use is perfect public Bayesian equilibrium (PPBE).[7] We focus on pure strategy equilibrium. Furthermore, we focus on the principal-preferred equilibrium[8] that displays the following feature: whenever the agent works, he matches the action with the state.[9]

Histories and Strategies: Let $f \in \{F, \neg F\}$, $w \in \{W, \neg W\}$, and $v \in \{V, \neg V\}$ denote generic choices of firing, working, and verification. Let ν be the outcome of verification: if $v = V$, $\nu = s$ and if $v = \neg V$, $\nu = \emptyset$. The set $\mathcal{H}^t$ of public histories h^t of the form

$$h^t = \Big((f_\xi^1, w_\xi^1, m^1, a_\xi^1, a^1, v_\xi^1, v^1, \nu^1), ..., (f_\xi^t, w_\xi^t, m^t, a_\xi^t, a^t, v_\xi^t, v^t, \nu^t) \Big),$$

corresponds to the public history of the game at the end of period t.[10]

[7]We follow Athey and Bagwell (2008): A PPBE is a Perfect Bayesian Equilibrium (PBE) in which strategies depend on public histories and payoff-relevant variables private information, that is, information that pertains only to the current round. Allowing for more general strategies does not affect the set of equilibrium payoffs.

[8]The principal's preferred equilibrium is the equilibrium that maximizes her ex ante payoff. Without loss of generality, uniqueness should be understood in terms of the expected values of equilibrium variables, since linearity may generate indeterminacy in the variables themselves. For simplicity, whenever such indeterminacy arises, the specification is stated in terms of these expectations.

[9]In Appendix 7.2, we show that it is without loss of generality to assume that whenever the agent works, he matches the action to the state.

[10]For ease of notation we do not include the firing decision of the principal in the history of

A firing strategy of the principal is a map from the set of histories of the form (h^{t-1}, f_ξ^t) to $\{F, \neg F\}$, a verification strategy is a map from the set of histories of the form $(h^{t-1}, f_\xi^t, w_\xi^t, m^t, a_\xi^t, a^t, v_\xi^t)$ to $\{V, \neg V\}$.

A working strategy of the agent is a map from the set of histories of the form $(h^{t-1}, f_\xi^t, w_\xi^t)$ to $\{W, \neg W\}$. A messaging strategy of the agent is a map from the set of histories of the form $(h^{t-1}, f_\xi^t, w_\xi^t, S^t)$ to $\{\hat{L}, \hat{R}\}$. A strategy for the action of the agent is a map from the set of histories of the form $(h^{t-1}, f_\xi^t, w_\xi^t, S^t, m^t, a_\xi^t)$ to $\{L, R\}$. A strategy for evidence disclosure of the agent is a map from the set of histories of the form $(h^{t-1}, f_\xi^t, w_\xi^t, S^t, m^t, a_\xi^t, v_\xi^t, v^t)$ to S^t.

Solving for the principal's preferred equilibrium calls for solving for the entire upper boundary of the equilibrium payoff set. Hence, the characterization of other equilibria (for instance, the agent-preferred equilibrium) follows as a by-product.

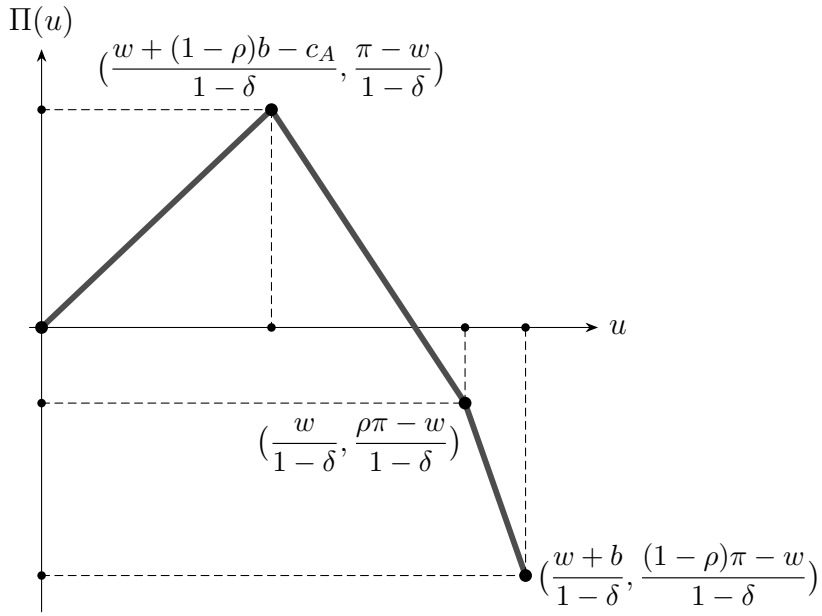

Figure 1: The upper boundary of the feasible payoff set.

Before analyzing the game, it is useful to characterize the set of feasible payoffs without incentive constraints. Figure 5 illustrates the upper boundary of this set, which exhibits several noteworthy features. On the upper boundary, the principal does not use verification and there are four extreme points: $(0, 0)$: firing, $(\frac{w+(1-\rho)b-c_A}{1-\delta}, \frac{\pi-w}{1-\delta})$: the agent works and matches the action with the state, yielding the highest feasible payoff for the principal, $(\frac{w}{1-\delta}, \frac{\rho\pi-w}{1-\delta})$: the agent does not work and always takes action L, $(\frac{w+b}{1-\delta}, \frac{\pi(1-\rho)-w}{1-\delta})$: the agent does not work and always

the game since if $f = F$ the game ends. Therefore in all histories we assume $f = \neg F$.

takes action R, yielding the highest feasible utility for the agent.[11]

Note that the principal's highest feasible payoff cannot be achieved in equilibrium, as she does not verify, and the agent has no incentive to work and learn the state.

The minmax payoff vector, is an equilibrium payoff vector: if the principal expects that the agent does not work, she fires him. Conversely, if the agent expects that the principal fires him at every opportunity, not working is optimal. Without loss of generality, in the principal's preferred equilibrium, we assume any observable deviation triggers the minmax equilibrium.[12]

There is a range of parameters for which immediate firing is the unique equilibrium. To avoid trivialities, we assume throughout that equilibria involving strictly positive payoffs exist. It will be clear from the explicit formulas what restriction this entails on the parameters.

4 Three-State Automaton: Illustrating the Main Results

To build intuition about the trade-offs involved, consider three-state automata and assume for now that the principal has commitment power. We then show how these automata can be adjusted to incorporate the principal's incentives, so that the three-state automata constitute equilibria of our game. While the principal-preferred equilibrium does not belong to this class, it shares important properties with the best equilibrium within that class.

The three states are the following. **Firing** (state $\mathcal{F}$): The principal fires the agent, both players receive zero payoff, and the game ends. **Retiring** (state $\mathcal{R}$): The agent never works, and the principal does not verify. Moreover, the agent always chooses his favorite action, $a = R$. **Matching** (state $\mathcal{M}$): The agent works, learns the state, and matches the action with the state (*complies*).[13]

Both $\mathcal{F}$ and $\mathcal{R}$ are absorbing states in which the players' actions are fixed. Hence, we focus on state $\mathcal{M}$. Determine (i) the probability of verification and (ii) the

[11]Depending on parameters, $(\frac{w}{1-\delta}, \frac{\rho\pi - w}{1-\delta})$ might not appear on the upper boundary. It appears on the upper boundary if and only if $\frac{1-\rho}{\rho}\frac{c_A}{2-\rho} > b$. Intuitively, this condition holds when the bias is sufficiently small.

[12]The following observations are also made without loss of generality: 1) If the agent works and truthfully reveals the state through message m, and the principal verifies, then the agent chooses $s = \theta$; otherwise, the principal fires the agent. 2) If the agent deviates—either by not working or by working and misreporting the state—and the principal verifies, then the agent chooses $s = \emptyset$. Even if the agent were to reveal the true state ($s = \theta$) after sending an incorrect message ($m \neq \theta$), the principal would still fire the agent. Note that, in this case, learning the state provides no benefit to the principal, since the agent's action has already been taken.

[13]The agent "complies" when he works, learns the state, and matches the action with the state.

transition probabilities from $\mathcal{M}$ to $\mathcal{R}$ and $\mathcal{F}$, as functions of verification and the agent's action, in order to maximize the principal's payoff while ensuring that the agent complies.[14]

Let v^a denote the probability of verification after the agent takes action a, and let u be the agent's utility in state $\mathcal{M}$. Rather than working with transition probabilities, it is convenient to use continuation utilities. Denote the continuation utility when action a is taken and verification occurs by u_V^a,[15] and the continuation utility when action a is taken and verification does not occur by $u_{\neg V}^a$.[16]

In state $\mathcal{M}$, the principal must account for two incentive constraints: 1) Complying must be optimal for the agent. Formally,

$$\mathbb{E}\left[u(\theta) + v^\theta \delta u_V^\theta + (1 - v^\theta)\delta u_{\neg V}^\theta\right] - c_A \geq \max_a \{u(a) + (1 - v^a)\delta u_{\neg V}^a\}. \tag{WM}$$

The left-hand side is the ex-ante utility if the agent complies. The right-hand side is the utility of the agent from shirking and choosing action a. In this case, the agent is fired after the verification and the continuation utility is zero. 2) Conditional on working, the agent must prefer to match the action with the state. Formally, for $a, a' \in \{L, R\}$,

$$u(a) + v^a \delta u_V^a + (1 - v^a)\delta u_{\neg V}^a \geq u(a') + (1 - v^{a'})\delta u_{\neg V}^{a'}. \tag{M}$$

Denote by (M–L) when $a = L,\ a' = R$, and by (M–R) when $a = R,\ a' = L$ in the constraint (M).

To understand how two distinct frictions—the temptation to shirk and mismatching the action to the state—shape the equilibrium structure, consider two special cases in turn: an unbiased agent ($b = 0$) and an agent with zero cost of effort ($c_A = 0$).

1) Unbiased agent. Assume $b = 0$. It is without loss of generality to assume 1) $v := v^L = v^R$ and 2) if the agent works, then he matches the action with the state. Intuitively, the agent does not have any incentive to mismatch the action after learning the state, and so the principal has no reason to condition verification on the chosen action. Therefore (M) is slack and the only binding incentive constraint

[14] Transitions across states are implemented with the prd. Hence, the state is common knowledge.

[15] As discussed, it is without loss of generality to assume that, after verification, if $s \neq \theta$, the agent is fired. Therefore, for simplicity, u_V^a refers to the continuation utility in case the agent provides evidence, i.e., $s = \theta$.

[16] Continuation utilities below (above) u imply a transition to $\mathcal{F}$ ($\mathcal{R}$) with positive probability. Utilities are fixed at states $\mathcal{F}$ and $\mathcal{R}$ (0 and $\frac{w+b}{1-\delta}$). Hence, transition probabilities are given by the ratio of the absolute difference between the current utility and the continuation utility to the absolute difference between the current utility and the continuation utility ($\mathcal{F}$ or $\mathcal{R}$).

is (WM).

Four observations are in order. First, the agent's utility from shirking and choosing L must be the same as shirking and choosing R. If shirking and choosing a is worse than shirking and choosing a', a mean-preserving contraction in u_V^a (verification) and $u_{\neg V}^a$ (no verification) benefits the principal.[17] Hence (WM) simplifies to

$$\mathbb{E}_\theta \left[v^\theta \delta u_V^\theta \right] \geq c_A.$$

Second, the continuation utilities after verification must be equal, i.e., $u_V := u_V^L = u_V^R$. A mean-preserving contraction of u_V^R and u_V^L does not affect (WM). Therefore, by concavity of the principal's payoff as a function of the agent's utility (the *value function*), the continuation utilities in the case of verification must be identical. Intuitively, when the agent provides evidence after verification, rewards or punishments are independent of the agent's actions.

Third, when (M) is slack, distorting $u_{\neg V}^L$ or $u_{\neg V}^R$ is not beneficial, i.e., $u_{\neg V}^L = u_{\neg V}^R = u$. The continuation utilities in the case of no verification, $u_{\neg V}^a$, do not appear in the incentive constraints. Thus, spreading them away from the initial utility u in state $\mathcal{M}$ does not help the principal. Hence, in the absence of verification, the state remains in $\mathcal{M}$.

Fourth, when (M) is slack, (WM) must bind, i.e., $v\delta u_V = c_A$. Otherwise, reduce slightly the probability of verification. Therefore, the probability of verification and the continuation utility serve as substitutes for providing incentives: reducing one requires increasing the other.

These observations extend to the principal's preferred equilibrium, without the assumption of the principal's commitment and the focus on three-state automata.

These observations, along with the agent's promise-keeping, pin down v and u_V.[18] We are left with solving for the optimal value of u in state $\mathcal{M}$. It is easy to show that $u_V \geq u$. Intuitively, if verification occurs and the agent has complied, then she does not expose the agent to firing. When the only binding constraint is (WM), the agent is not fired on the equilibrium path.

To sum up, there are only two candidates for the optimal three-state automaton. If the verification cost is small ($c_P \leq \bar{c}_P$), the state always remains at $\mathcal{M}$, i.e.,

[17] Note that if $u_V^a = u_{\neg V}^a$, then a mean-preserving contraction between continuation utilities cannot be implemented. However, she can reduce the cost of verification by decreasing the probability of verification v^a, provided that $v^a \neq 0$. If $v^a = 0$, the agent would (weakly) prefer to deviate.

[18] Formally, $u = \mathbb{E}\left[u(\theta) + v^\theta \delta u_V^\theta + (1 - v^\theta)\delta u_{\neg V}^\theta\right] - c_A$. Therefore $v = \frac{w+\delta u - u}{\delta u}$ and, $u_V = \frac{c_A}{\delta v} = \frac{c_A u}{w+\delta u - u}$.

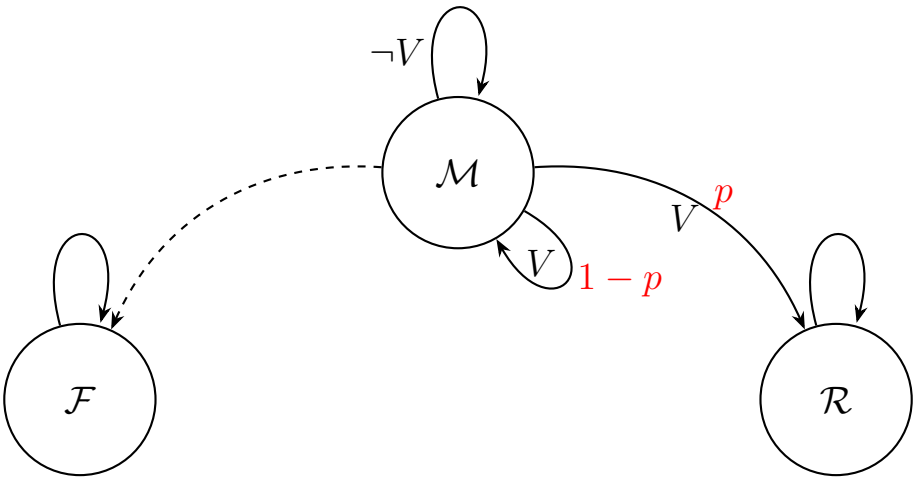

Figure 2: Three-state automaton. When verification does not occur, the state remains unchanged. If verification occurs and the agent provides evidence, the state transitions to $\mathcal{R}$ with probability $p = (u_V - u)/(\frac{w+b}{1-\delta} - u)$. If, after verification, the agent does not provide evidence ($s \neq \theta$), the state transitions to $\mathcal{F}$ (dashed arrow).

$u_V = u$.[19] In this case, the agent is neither fired nor retired (since, as noted earlier, $u^L_{\neg V} = u^R_{\neg V} = u_V = u$). If instead, $c_P > \overline{c}_P$, the state transitions from $\mathcal{M}$ to $\mathcal{R}$ with positive probability, i.e., $u_V > u$. This result is intuitive: when c_P is small, the principal prefers to verify more often and never retire the agent, rather than to reduce the probability of verification and retire the agent with positive probability.

Hence, for sufficiently low verification costs, the principal never retires the agent—verification alone suffices. Surprisingly, however, this is an artifact of the very simple automaton considered here. As more complicated automata are considered, retirement is used over a larger range of cost parameters. Indeed, for a sufficiently rich automaton, retirement is always used, no matter how small the verification cost is.

To see this, let us consider using a sequence of three-state automata. First, consider an auxiliary three-state automaton with states $\mathcal{F}$, $\mathcal{R}$, and $\mathcal{M}_1$. In state $\mathcal{M}_1$, the agent always complies; the principal retires the agent after the first verification. If verification does not occur, the state remains $\mathcal{M}_1$. The principal verifies after both actions with the same probability v, chosen so that the agent is indifferent between working and shirking.[20]

[19] Explicitly, $\overline{c}_P = \frac{\delta \frac{w}{1-\delta}\left(\frac{\rho\pi}{c_A} - 1\right)}{2 + \frac{c_A}{w - c_A}}$.

[20] Formally, $v\delta\frac{w}{1-\delta} = c_A$.

Next, replace $\mathcal{R}$ with $\mathcal{M}_1$ in the initial three-state automaton.[21] Simple algebra shows that $u_V > u$ if and only if $c_P > \overline{c}_P^1$, where $\overline{c}_P^1 < \overline{c}_P$. Intuitively, retirement now occurs after two rounds of verification: once from $\mathcal{M}$ to $\mathcal{M}_1$ and once within $\mathcal{M}_1$, from $\mathcal{M}_1$ to $\mathcal{R}$. Retirement after two rounds saves on verification costs compared to a scenario in which retirement does not occur at all. Therefore, for a larger range of verification costs, retirement with positive probability becomes optimal.

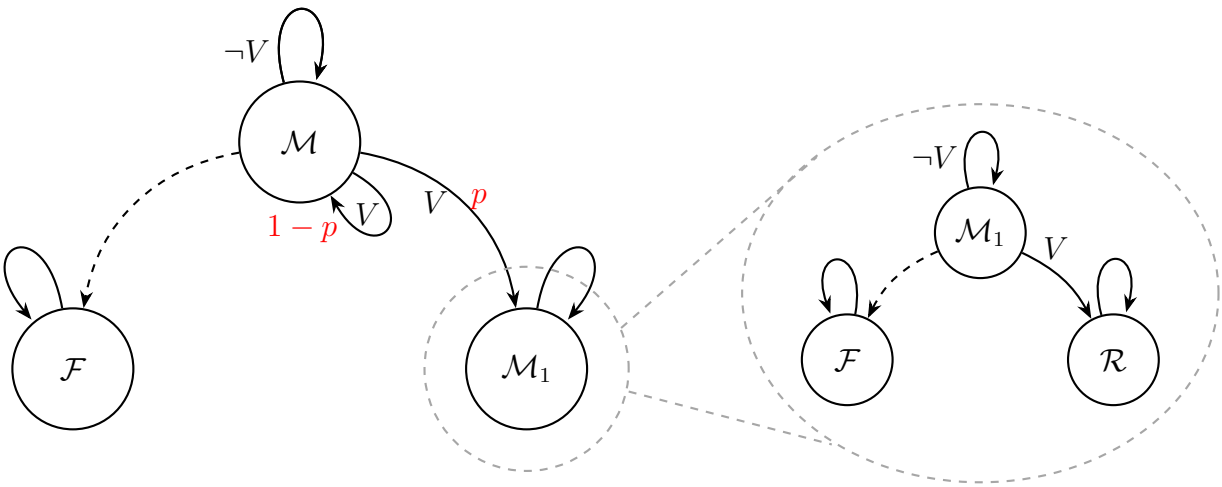

Figure 3: Replacing $\mathcal{R}$ with $\mathcal{M}_1$.

[21] Formally, replace $\mathcal{R}$ with the payoffs of $\mathcal{M}_1$ and treat it as an absorbing state.

This intuition generalizes. Introduce another auxiliary three-state automaton with states $\mathcal{F}$, $\mathcal{M}_1$, and $\mathcal{M}_2$. In state $\mathcal{M}_2$, the agent complies and after the first verification, the state transitions to $\mathcal{M}_1$, and if verification does not occur, the state remains in $\mathcal{M}_2$. The principal verifies after both actions with the same probability v, chosen so that the agent is indifferent between working and shirking. Replace $\mathcal{R}$ with $\mathcal{M}_2$ in the original three-state automaton. Simple algebra shows that $u_V > u$ if and only if $c_P > \overline{c}_P^2$, where $\overline{c}_P^2 < \overline{c}_P^1$. Similarly, one can introduce $\mathcal{M}_n$ and show that the sequence $\{\overline{c}_P^n\}$ is decreasing and converging to zero when n goes to infinity.[22]

By retiring the agent after n verifications, the principal reduces the verification probability at each of the n steps. The promise of future retirement creates incentives in all earlier periods. The savings from reduced verification costs accumulate across n periods, whereas the disutility of retirement is incurred only once. As a result, the ratio of the benefit from reduced verification costs to the disutility of retirement grows at rate n. For sufficiently large n, the benefit always outweighs the disutility. Refer to this phenomenon as "snowballing."[23]

Note that *snowballing* differs from the standard *backloading* result. While delaying rewards—i.e., *backloading*—is effective for sustaining incentives and is indeed operative in our model, the accumulating impact of such deferrals—*snowballing*—renders the deferred rewards effectively unavoidable, regardless of how small the verification cost.

2) The agent with zero cost of working. Now suppose $c_A = 0$, so that without loss of generality assume that the agent knows the state. In contrast to the unbiased case, (WM) is slack. Intuitively, since the agent knows the state, working is not a concern for the principal; the only concern is that he matches the action with the state.

Four results are worth highlighting. First, the probability of verification after R is strictly higher than after L, i.e., $v^R > v^L$. Since the agent is biased toward

[22] Explicitly $\overline{c}_P^n = \frac{\delta \frac{w}{1-\delta}\left(\frac{\rho\pi}{c_A}-1\right)}{n+2+\frac{c_A}{w-c_A}}$

[23] Note that while the *ratio* of benefits to disutility grows at rate n, the *net gain*—that is, the benefit minus the disutility—converges to zero as n becomes large. Let π_n denote the principal's payoff in state $\mathcal{M}_n$, and let π^* denote the payoff when retirement never occurs. Therefore,

$$\pi_{n-1} - \pi^* = \frac{\left(\frac{c_A}{w}\right)^n \left(n\frac{c_P}{\delta} - \frac{\rho\pi(w-c_A)}{(1-\delta)c_A}\right)}{1-\left(\frac{c_A}{w}\right)^n}.$$

Hence, the benefit from reduced verification cost is: $\frac{n\frac{c_P}{\delta}\left(\frac{c_A}{w}\right)^n}{1-\left(\frac{c_A}{w}\right)^n}$, and the disutility is: $\frac{\frac{\rho\pi(w-c_A)}{(1-\delta)c_A}\left(\frac{c_A}{w}\right)^n}{1-\left(\frac{c_A}{w}\right)^n}$.

action R, to incentivize him to match the action with the state, verification occurs more frequently after R than after L.

Second, there is no need to verify after L, i.e., $v^L = 0$. Verification after L only matters for the incentive to work, which is absent here. Shifting verification from L to R always strengthens incentives.

Third, since (WM) is not binding, the continuation utility after verification remains undistorted, i.e., $u_V^R = u$. After verification, if the agent provides evidence, the state does not transition to $\mathcal{R}$ or $\mathcal{F}$. Since (M–R) never binds and the only binding constraint is (M–L), u_V^R does not appear in any binding constraint. Thus, u_V^R is not used to provide incentives; it must therefore remain undistorted.

Fourth, in the absence of verification, the continuation utility is lower than the initial utility when the agent takes R, and higher when the agent takes L; that is, $u_{\neg V}^L \geq u \geq u_{\neg V}^R$. Intuitively, when verification does not occur, the agent is exposed to firing (punished) after taking his preferred action (R) and to retiring (rewarded) after his less preferred action (L). If $u_{\neg V}^L < u_{\neg V}^R$, then a mean-preserving contraction of $u_{\neg V}^L$ and $u_{\neg V}^R$ relaxes (M–L) and benefits the principal. Hence, it must be that $u_{\neg V}^L \geq u_{\neg V}^R$. Moreover, the envelope theorem implies that the continuation utilities, $u_{\neg V}^L$ and $u_{\neg V}^R$, must bracket u.

These four observations extend to the principal's preferred equilibrium (without the focus on three-state automata). Figure 4 displays the three-state automaton when $c_A = 0$.

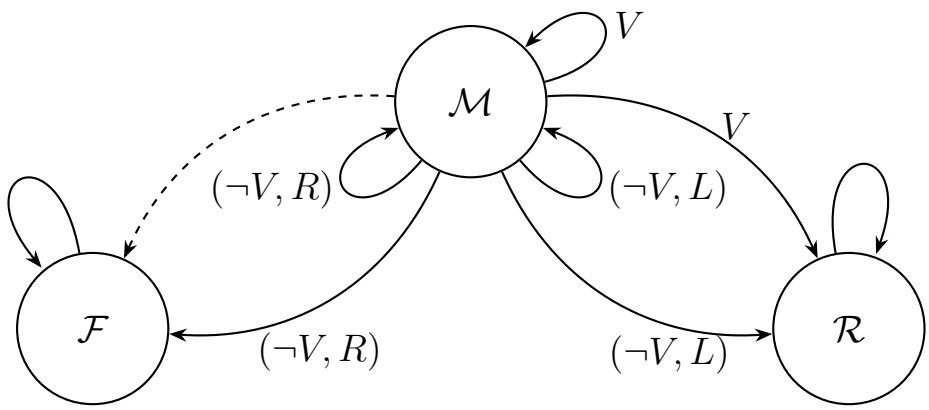

Figure 4: Three-state automaton with $c_A = 0$. The transition probability from $\mathcal{M}$ to $\mathcal{R}$ under verification is $(u_V - u)/(\frac{w+b}{1-\delta} - u)$. If verification does not occur and the action is L, the transition probability from $\mathcal{M}$ to $\mathcal{R}$ is $(u_{\neg V}^L - u)/(\frac{w+b}{1-\delta} - u)$. If verification does not occur and the action is R, the transition probability from $\mathcal{M}$ to $\mathcal{F}$ is $(u - u_{\neg V}^R)/(u - 0)$.

We can characterize the use of retirement and firing in terms of two thresholds,

$c_P^f \geq c_P^r > 0$. Retirement occurs with strictly positive probability if and only if $c_P > c_P^r$, whereas firing occurs with strictly positive probability if and only if $c_P > c_P^f$. The intuition is simple: when verification cost is low, verification is the most efficient instrument, and the principal exclusively relies on it. As c_P increases, it becomes too costly to rely solely on verification, and the principal introduces retirement as a device: promising eventual retirement allows her to reduce the frequency of verification while still giving the agent incentives to comply. Firing, by contrast, is the harshest instrument because it terminates the relationship. Hence, firing appears only when verification costs are so high that neither verification nor retirement alone can sustain incentives.

Fixing c_P, an analogous set of thresholds $b^f \geq b^r > 0$ can be defined in terms of the agent's bias b. For small biases, verification alone suffices to discipline the agent. As the bias increases b^r, the temptation to mismatch the action with the state grows, and the principal must rely on retirement as an additional incentive device. When the bias becomes large ($b > b^f$), retirement is not enough, and firing is used with positive probability.

In summary, for small biases or verification costs, the state always remain at $\mathcal{M}$ and the principal does not use firing or retiring.

Interestingly, in the principal's preferred equilibrium (without the focus on three-state automata), allowing the agent to shirk is always valuable: the principal benefits from snowballing even when c_P is small and π is large. To satisfy (M–L), the principal can either increase verification (v^R) while keeping $u^R_{\neg V}$ undistorted at u, or reward the agent by raising $u^L_{\neg V}$ above u. However, the latter is better. The cumulative savings in verification costs dominate.

Moreover, firing the agent—even without deviation—serves as an effective instrument if (M–L) binds on the equilibrium path (as in the case $c_A = 0$). To satisfy (M–L), she must either increase verification (v^R) while keeping $u^R_{\neg V}$ undistorted at u, or punish the agent by reducing $u^R_{\neg V}$ below u. The threat of future firing complements verification and reduces verification costs across multiple periods. Snowballing, by promising future punishment, renders the use of firing unavoidable, regardless of how small b is.

We now combine the two cases, allowing for both a positive bias and a cost of working. In the principal's preferred equilibrium, (1) allowing shirking is always valuable; (2) if (M–L) binds along the entire equilibrium path, then firing occurs with probability one, which happens when the expected benefit from bias exceeds the cost of working, i.e., $(1-\rho)b > c_A$; and (3) if (WM) binds only on part of the equilibrium path, the principal verifies both actions with positive probability to detect shirking, and firing does not occur. In this case, greater verification substitutes for firing and prevents its snowballing effect.

Non-commitment. So far, we assume that $\mathcal{F}$ and $\mathcal{R}$ are absorbing states

with predetermined behaviors. Also, the principal never deviates from verifying whenever she is "supposed to."[24] Since $\mathcal{F}$ corresponds to the minmax equilibrium, it is sequentially rational. However, the agent cannot be permanently retired in $\mathcal{R}$; otherwise, the principal would deviate and fire him. Sequential rationality can be restored with a slight modification of the automaton introduced above. To prevent this deviation, the agent must work at least occasionally. By replacing $\mathcal{R}$ with $\mathcal{RW}$ (reduced workload), we allow transitions from $\mathcal{RW}$ to $\mathcal{M}$. Finally, the transitions must be chosen such that the principal's payoff in $\mathcal{RW}$ exceeds $\frac{c_P}{\delta}$. This ensures that the principal has sufficient incentive to verify in $\mathcal{M}$.

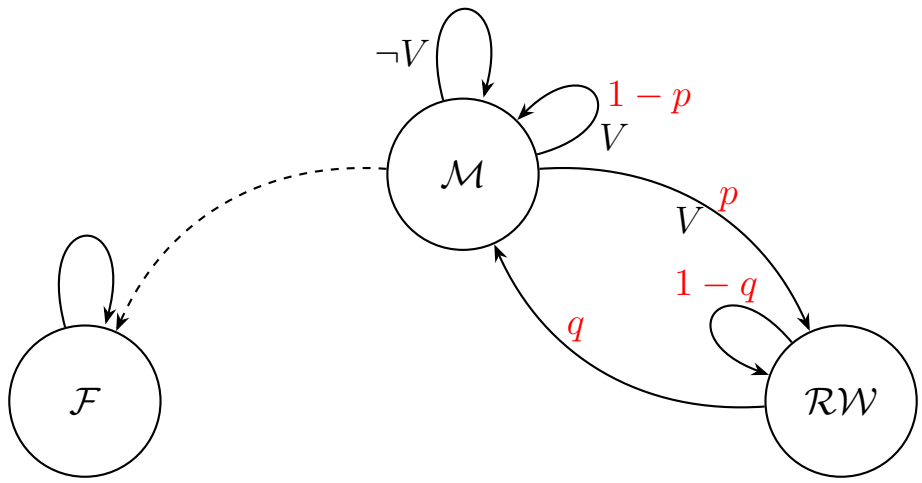

Figure 5: Three-state automaton when $b = 0$, where $\mathcal{R}$ is replaced by $\mathcal{RW}$. The state transitions from $\mathcal{RW}$ to $\mathcal{M}$ with probability q.

[24]When we say the principal is "supposed to" verify or fire, we mean that the prd recommends working, verifying or firing.

5 Results

In Section 5.1, we formalize the principal's problem. In Section 5, we (i) characterize the behavior underlying equilibria that attain payoffs on the upper boundary of the equilibrium payoff set, (ii) examine the evolution of continuation utilities along this boundary, and (iii) characterize the long-run dynamics of the principal's preferred equilibrium.

5.1 Principal's Program

Our goal is to characterize the preferred equilibrium of the principal. The history of the game can be summarized by the agent's continuation utility. The upper boundary of the equilibrium payoff set is self-generating, in the sense of Abreu, Pearce, and Stacchetti (1990). Accordingly, the problem can be formulated as a recursive program, subject to incentive constraints.

Lemma 3 in the Appendix establishes that, without loss of generality, when the agent is supposed to work, continuation utility depends only on his action and not on messages. When instead he is not supposed to work, continuation utility depends neither on action nor on messages.

If the agent is supposed to shirk, the principal does not verify. In that event, verification yields no additional information.

Therefore, along the equilibrium path, the agent's continuation utility at the end of each period takes one of three values:

1. u_V^a, when the agent is supposed to work, he sends message $a \in \{L, R\}$, and the principal verifies.

2. $u_{\neg V}^a$, when the agent is supposed to work, he sends message $a \in \{L, R\}$ and the principal does not verify.

3. $u_{\neg W}$, when the agent is recommended not to work.[25]

Let $\Pi(u)$ be the principal's highest equilibrium payoff consistent with the agent's payoff being u. Let $\Pi(u) = -\infty$ if no such equilibrium exists. Let $\overline{u}$ denote the maximum promise such that $\Pi(u) > -\infty$. The principal's program is:

$$\mathcal{P}: \quad \Pi(u) = \sup \mathbb{E}\Big[\mathbb{1}_{\neg F}\big(\pi(\theta, a) - c_P\, \mathbb{1}_V - w + \delta\, \Pi(u')\big)\Big],$$

[25]Subscript "V" stands for *working and verifying*, subscript "$\neg V$" stands for *working and not verifying* and subscript "$\neg W$" stands for *not working*.

subject to the agent's incentive constraints, namely:
1) Working and matching the action with the state

$$\mathbb{E}\left[u(\theta) + v^{\theta}\delta u_V^{\theta} + (1 - v^{\theta})\delta u_{\neg V}^{\theta}\right] - c_A \geq \max_a\{u(a) + (1 - v^a)\delta u_{\neg V}^a\}. \tag{WM}$$

2) Conditional on working, matching the action with the state,

$$u(a) + v^a\delta u_V^a + (1 - v^a)\delta u_{\neg V}^a \geq u(a') + (1 - v^{a'})\delta u_{\neg V}^{a'}. \tag{M}$$

for all $a, a' \in \{L, R\}$.
3) The promise-keeping constraint:

$$u = \mathbb{E}\big[\mathbb{1}_{\neg F}\,(u(\theta) - c_A\,\mathbb{1}_W + \delta\,u')\big], \tag{PK}$$

and subject to the principal's incentive constraints:
1) Firing

$$\mathbb{E}\Big[\pi(\theta, a) - c_P\mathbb{1}_V - w + \delta\Pi(u')\Big] \geq 0, \tag{Fi}$$

2) Verifying

$$\delta\Pi(u_V^a) - c_P \geq 0, \tag{Ve}$$

for all $a \in \{L, R\}$.

The supremum is taken over the continuation utility u', the probabilities of firing, working, verifying and of the agent taking action R when he does not work. The continuation utility u' can take one of the three aforementioned values, $u_V^a, u_{\neg V}^a, u_{\neg W} \in [0, \overline{u}]$.[26]

The working and matching constraint (WM) ensures that the agent works and matches the action with the state. The matching constraint (M) guarantees that if the agent works, he matches the action with the state. The promise-keeping constraint (PK) requires that the agent's promised utility is equal to his actual utility.

The firing incentives of the principal are as follows: (i) The principal fires the agent when she is supposed to (not firing is an observable deviation and triggers the minmax equilibrium. This is trivially satisfied, and omitted). (ii) The principal retains the agent when she is supposed to. The principal's payoff must be positive for all on-path utility levels of the agent. This incentive constraint is given by equation (Fi).

[26]Expectations are taken with respect to the random outcomes, namely the state, firing status, working status, verification status and the agent's action (when the agent is supposed to shirk).

The verification incentives of the principal are as follows: (i) The principal does not verify when she is not supposed to. This constraint is trivially satisfied, as the principal knows that the agent has worked and has matched the action to the state, making verification unnecessary. (ii) The principal verifies whenever she is supposed to. The principal's incentive to verify when she is supposed to is that the discounted expected payoff from verification, $\delta\Pi(u_V^a)$ (for $a \in \{L, R\}$), must be at least as large as the cost c_P, as represented by equation (Ve).

5.2 Analysis

To analyze the long-run dynamics of this relationship, we proceed in two steps. First, we characterize behavior as a function of the agent's utility; second, we characterize the evolution of the agent's continuation utility.

Theorem 1 characterizes the behavior as a function of the agent's utility along the upper boundary of the equilibrium payoff set. Theorem 2 examines the evolution of continuation utilities on this boundary. Theorem 3 combines these results to describe the evolution of the relationship in the principal's preferred equilibrium.

Let $\tilde{u} := \frac{w}{1-\delta}$ denote the agent's utility when he consistently shirks and chooses L, and the principal neither verifies nor fires him.

Theorem 1 (Behavior)**.** *There are unique thresholds u_f and u_e, with $0 < u_f \leq u_e < \overline{u}$, such that the behavior along the upper boundary of the equilibrium payoff set is characterized as follows:*

1. ***Firing Region,*** *$u \in [0, u_f)$**:** The firing probability declines linearly from 1 at $u = 0$ to 0 at $u = u_f$. For all $u \geq u_f$, the agent is never fired.*

2. ***Verification Region,*** *$u \in [u_f, u_e)$**:** The probability of verification after R is strictly higher than after L for all $u \in [0, \overline{u}]$, i.e., $v^R > v^L$. Both v^R and v^L are continuous in u and remain constant on the intervals $[0, u_f]$ and $[u_e, \overline{u}]$. On interval $[u_f, u_e]$, v^R is strictly decreasing, while v^L is zero for $\tilde{u} \leq u_f$; otherwise it is strictly decreasing up to $\min\{\tilde{u}, u_e\}$ and, if $\tilde{u} \leq u_e$, equal to zero in $[\tilde{u}, u_e]$.*

3. ***Reduced Workload Region,*** *$u \in [u_e, \overline{u}]$**:** The agent works with probability one whenever $u \leq u_e$. The probability of working declines strictly over $[u_e, \overline{u}]$. Moreover, the agent chooses R if the principal's marginal payoff at $\overline{u}$ is less than $\frac{\pi(1-2\rho)}{b}$, and chooses L otherwise.*

Proof in Appendix.

Figures 6 and 7 illustrate the firing, verification, and working probabilities as functions of the agent's utility. In what follows, we discuss the intuition behind Theorem 1 for each region.

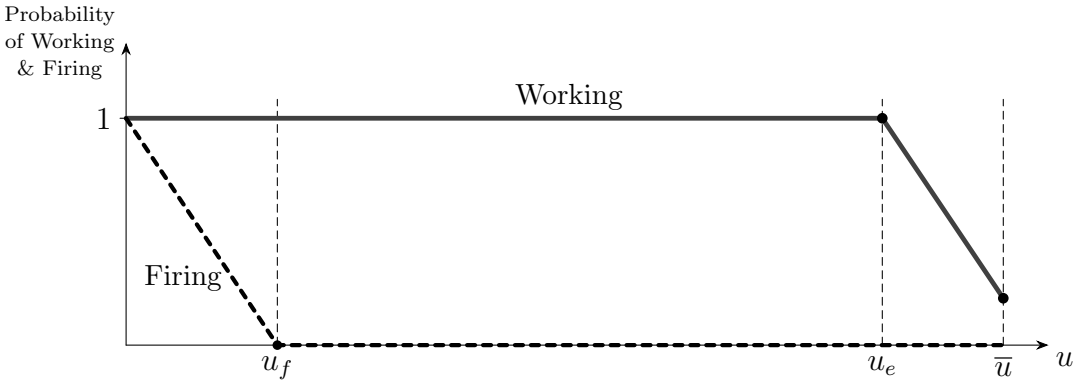

Figure 6: Working and firing probabilities as functions of the agent's utility.

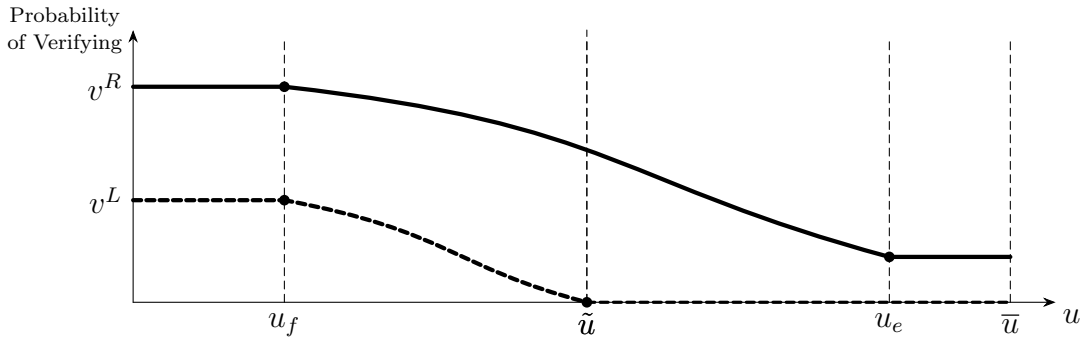

Figure 7: Verification probability as a function of the agent's utility.

Firing. When the agent's utility is low ($u < u_f$), firing with positive probability is necessary to deliver the utility to the agent. As u increases, the firing probability decreases linearly, reaching zero at u_f. Beyond this point, firing is not used.

Verification. Because the agent is biased toward R, when the state is L, verifying more frequently after R than after L (i.e., $v^R > v^L$), raises the cost of choosing R when the state is L.

Verification remains constant in regions where the principal primarily relies on other instruments to deliver the promised utility—namely, firing at low utility levels ($u \in [0, u_f)$) and reduced workload at high utility levels ($u \in [u_e, \overline{u}]$). In the low-utility region, verification is frequently used; less so for high utilities.

In the intermediate region $[u_f, u_e)$, the verification probability decreases with u. In each period in which the agent is closer to being allowed to shirk, the principal can further decrease the intensity of verification. The reason is that the threat of firing is stronger the closer the agent is to shirking; hence, the intensity can be reduced. verification after L is unnecessary, except to ensure that the agent is not

repeatedly shirking and choosing his L, thus $v^L = 0$ for all $u \geq \tilde{u}$. We call the region $[u_f, \tilde{u}]$ (if nonempty) *evidence-based* and $[\tilde{u}, u_e]$ *hybrid*.

Reduced Workload. For all $u \leq u_e$, the agent works with probability one. For $u > u_e$, some shirking is expected: the agent is no longer expected to exert effort in every period, and the probability of work declines as u increases. This shift marks a move toward implicit delegation, where the agent's past performance earns him greater discretion. Nevertheless, the agent never shirks with probability 1.[27]

In the reduced-workload region, for small biases, the agent's private gain from choosing R while shirking, is negligible relative to the principal's loss. As the bias grows large, the agent's gain becomes comparable to the principal's loss, making shirking with R part of the equilibrium. Formally, if the agent's net benefit from shirking and taking R rather than L exceeds the principal's marginal payoff, the agent takes R.

A consequence of Theorem 1 is that the principal may, in fact, benefit from the agent's bias. Although bias makes it harder to get the agent to match action with the state, it simultaneously increases the agent's motivation to work and to demonstrate that the state is R. The benefit comes from two sources. First, the agent's reward in the reduced workload region increases, since the agent chooses R, which allows the principal to sustain work over a larger interval $[u_f, u_e]$. Second, deterring shirking becomes easier and v^L can be decreased.

Before characterizing continuation utilities, we first discuss their role in shaping incentives. In the absence of transfers, continuation utilities serve as the main lever and provide two distinct instruments: the continuation utility when verification occurs and the continuation utility when verification does not occur. The former induces the agent to work, while the latter encourages matching the action to the state.

If the agent deviates by shirking whenever he is supposed to work, he obtains the same utility from taking L or R (see Lemma 3 in Appendix). That is,

$$u(L) + (1 - v^L)\delta u^L_{\neg V} = u(R) + (1 - v^R)\delta u^R_{\neg V}. \tag{I}$$

The intuition behind this lemma is as follows: if shirking and taking a is worse than taking a', then a mean-preserving contraction with respect to u^a_V (in the case of verification) and $u^a_{\neg V}$ (in the case of no verification) would benefit the principal. Narrowing the gap between continuation utilities increases the principal's payoff without affecting the agent's incentives.[28]

[27] If the agent shirks with probability 1, the principal's payoff is negative which cannot happen on the equilibrium path.

[28] Note that if $u^a_V = u^a_{\neg V}$, then the principal cannot implement a mean-preserving contraction

The above observation allows us to replace (M) with the stronger condition (I). Furthermore, (WM) can be replaced by

$$\rho v^R \delta u_V^R + (1-\rho) v^L \delta u_V^L \geq c_A. \tag{W}$$

Thus, the cost of working (c_A) must be compensated by the probability of verification and the continuation utility in the case of verification. Conditions (W) implies that the incentive to work is guaranteed through the probability of verification and the continuation utility that follows verification. Condition (I) implies that the incentive to match the action with the state is supported by the probability of not verification and the continuation utility in the absence of verification.

The upper boundary is linearly increasing in u for $u \in [0, u_f]$, concave for $u \in [u_f, u_e]$, and linearly decreasing in $u \in [u_e, \overline{u}]$ (see Claims 1 and 5 in Appendix).[29]

Due to the linearity of the upper boundary for low and high utility, a straightforward approach for implementing the equilibrium along these segments is to randomize over the segment's extreme points.

Theorem 2 characterizes continuation utilities for $u \in \{[u_f, u_e] \cup \overline{u}\}$.[30] The result highlights a crucial threshold, $\tilde{u}$, that separates regions in which the principal verifies after both actions from those where verification after R suffices.

Theorem 2 (Continuation Utilities)**.** *For all $u \in [u_f, u_e]$ in the verification region, on the equilibrium path the continuation utility after verification is independent of the agent's action:*

$$u_V := u_V^R = u_V^L.$$

In addition, the continuation utility after verification is always bigger than the continuation utility absence of verification and when the agent takes R, i.e.,

$$u_V > u_{\neg V}^R.$$

between continuation utilities. However, she can reduce the cost of verification by decreasing the probability of verification v^a.

[29]Intuitively, firing with probability f scales down the utility by $1-f$. Due to the multiplicative structure of the principal's payoff with respect to the firing probability, this also proportionally reduces the principal's value by $1-f$. The concavity of the upper boundary in the intermediate utility region arises from the prd. The principal's preferred equilibrium is achieved at an interior utility $u^* \in [u_f, u_e]$. Decreasing the probability of working (e) in $[u_e, \overline{u}]$ increases the utility linearly and decreases the principal's payoff linearly. Linearity arises for two reasons. First, in the agent's utility, the expected cost of exerting effort is given by ec_A, so the probability of shirking affects the agent's stage-game payoff linearly. Second, since the principal's payoff is scaled by the probability that the agent exerts effort, variations in shirking also induce a linear effect on the principal's continuation value.

[30]The unique equilibrium at $u = 0$ involves firing.

In the absence of verification, the continuation utilities satisfy:

- ***Evidence-based region:*** *no distortion,*

$$u^R_{\neg V} = u = u^L_{\neg V}.$$

- ***Hybrid region:*** *strict separation,*

$$u^R_{\neg V} < u < u^L_{\neg V}.$$

Moreover, at $u = \overline{u}$, the continuation utilities under working are equal to those at u_e, and $u_{\neg W} = u$.

Proof in Appendix.

Theorem 2 describes the structure of continuation utilities. The first takeaway is that when the principal verifies and the agent provides evidence, the continuation utility is independent of his action, i.e., $u_V := u^R_V = u^L_V$. The continuation utilities after verification matter only for the working constraint (W). A mean-preserving contraction of u^L_V and u^R_V does not affect the incentive constraints. Therefore, by the concavity of the value function, the continuation utilities in the case of verification must be equal.[31] Intuitively, hard information from verification dominates soft information from actions. Moreover, the agent is rewarded whenever the principal verifies and the agent provides evidence; that is, if $v^a > 0$, then $u_V > u^a_{\neg V}$.

The second takeaway is that, in the absence of verification, for continuation utilities lower than $\tilde{u}$, the agent is neither rewarded nor punished for taking L or R; that is, $u^R_{\neg V} = u = u^L_{\neg V}$. We refer to the region $[u_f, \tilde{u}]$ (if it is nonempty) as the evidence-based region, since continuation utilities evolve only based on hard information.

In the evidence-based region, the agent's continuation utility lies below what he could secure by shirking, choosing L, and repeatedly obtaining the flow payoff w. Since the principal cannot detect or punish such deviations without verification, the principal must verify with positive probability, even when the agent selects the less-preferred action. Verification probabilities are the more effective instrument: because of linearity, it yields a constant return. Hence, as long as both verification probabilities are strictly positive, the principal prefers to use differential verification rather than dispersion in the continuation utilities.

In the evidence-based region, the only binding constraint is (WM), while (M) is slack. Continuation utilities in the case of no verification are used only to

[31] Note that a mean-preserving contraction of u^R_V and u^L_V weakly increases $\min\{\Pi(u^L_V), \Pi(u^R_V)\}$, and thus does not affect (Ve).

incentivize the agent to match the action with the state. Therefore, when (M) is not binding, distorting them away from the initial utility u is unnecessary.

The third takeaway is that for utilities higher than $\tilde{u}$, the agent is punished for taking his preferred state and rewarded for taking the less-preferred one; that is, $u^R_{\neg V} < u < u^L_{\neg V}$. The relationship is now shaped by both hard (evidence) and soft (unverified) information. We refer to the region $[\tilde{u}, u_e]$ (if it is nonempty) as the hybrid region.

In the hybrid region, once $v^L = 0$ (i.e., for all $u \geq \tilde{u}$), the principal relies on spreading out continuation utilities to induce the agent to match the action to the state when the state is L. The agent is rewarded for choosing his less-preferred action, and punished otherwise. If $u^L_{\neg V} < u^R_{\neg V}$, then a mean-preserving contraction of $u^L_{\neg V}$ and $u^R_{\neg V}$ relaxes the matching constraint for state L and benefits the principal. Therefore, it must hold that $u^L_{\neg V} \geq u^R_{\neg V}$. Moreover, the envelope theorem implies that the continuation utilities $u^L_{\neg V}$ and $u^R_{\neg V}$ must bracket u. Figure 8 depicts the structure of continuation utilities in the verification region.

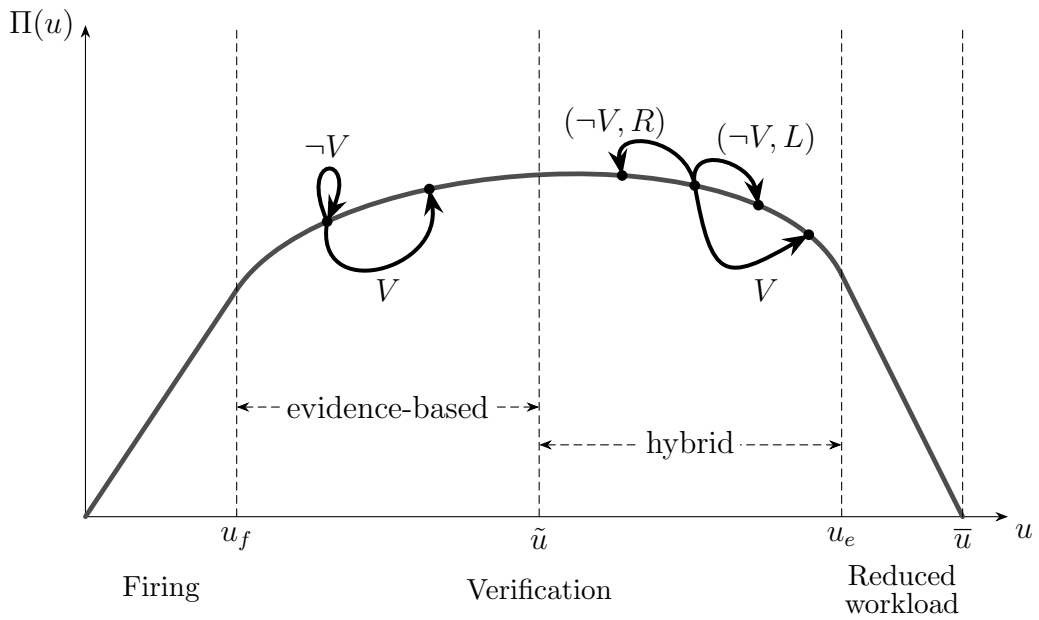

Figure 8: Continuation utilities in the verification region.

In the reduced-workload region ($[u_e, \overline{u}]$), the agent shirks with different probabilities. Accordingly, the continuation utilities when the agent is supposed to work $(u_V, u^L_{\neg V}, u^R_{\neg V})$ remain constant. Moreover, the continuation utility when the agent is not supposed to work $(u_{\neg W})$ remains undistorted at u.

Agent's preferred equilibrium. In this equilibrium, the agent's utility is

$\overline{u}$, the principal's payoff is zero ($\Pi(\overline{u}) = 0$), and the workload is minimized. If the agent is required to work, the game transitions to u_e; when u_e lies in the evidence-based region, the agent may remain indefinitely in the reduced-workload region without returning to verification, whereas if u_e lies in the hybrid region, the agent moves to the verification region with probability one. If the agent is not required to work, then depending on the extent of biases (as discussed in Theorem 1), the agent chooses either L or R. In the following period, after not working, the continuation utility stays at $u_{\neg W}$. Figures 9 displays the continuation utilities at $\overline{u}$.

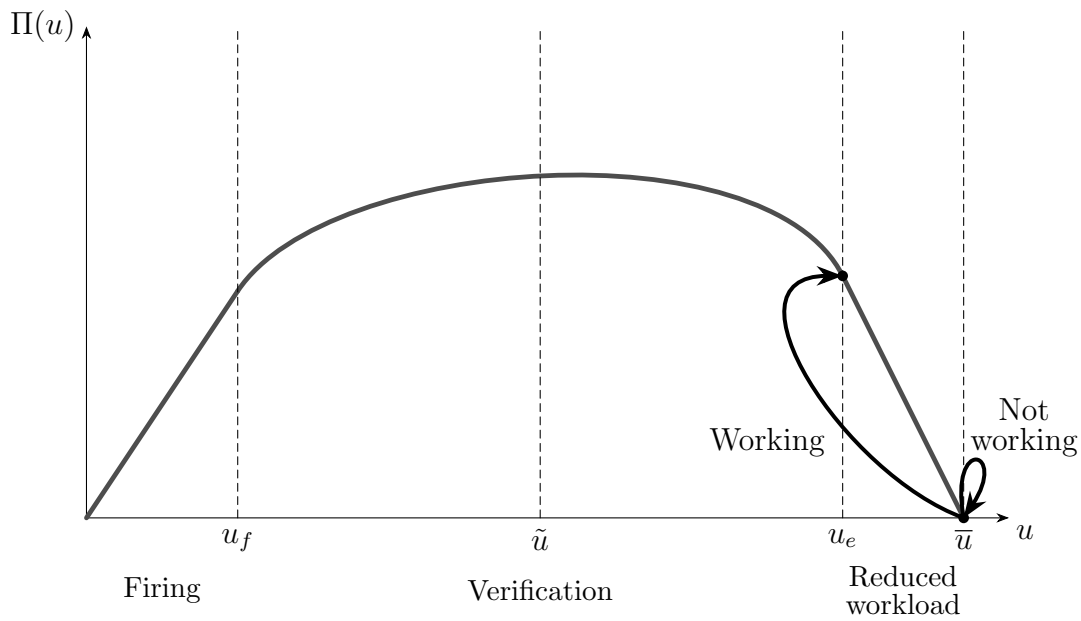

Figure 9: Continuation utilities at $\overline{u}$.

Let u^* denote the utility level of the agent that maximizes the principal's equilibrium payoff, i.e., $u^* = \arg\max_u \Pi(u)$. Note that u^* always lies in the verification region, $u^* \in [u_f, u_e]$, because the value function is strictly increasing in the firing region and strictly decreasing in the verification region. Intuitively, the relationship begins when the agent complies—works, learns the state, and matches the action with the state.

This raises the question of whether the principal ever uses firing or a reduced workload. Theorem 3 addresses this question.

Theorem 3 (Dynamics)**.** *In the principal's preferred equilibrium,*

1. *The principal uses reduced-workload with positive probability.*
2. *The principal fires the agent with probability one if and only if the evidence-based region is empty.*

The first part of Theorem 3 states that allowing the agent to shirk is always valuable, even when the cost of verification is small and the loss from not working (and thus potentially mismatching the action with the state) is large. By promising a lighter workload in the future, the cost of verification is decreasing in all proceeding periods. As explained in Section 4, this is *snowballing*.

In the evidence-based region, the agent's reward for complying takes the form of an increase in continuation utility ($u_V > u$). Such rewards move the relationship either into the hybrid region or directly into the reduced-workload region. In the hybrid region, compliance is rewarded through continuation utility when the agent chooses his less-preferred action, L, i.e., $u^L_{\neg V} > u$, which leads with positive probability to the reduced-workload region.[32]

The second part of Theorem 3 asserts that if the evidence-based region is empty, verification and reduced workload do not suffice, and the principal must rely on firing along the equilibrium path. The threat of future firing complements verification and lowers verification costs across multiple periods. Hence, snowballing—through the promise of future punishment—renders the use of firing unavoidable.

In the hybrid region, the principal punishes the agent—despite compliance—through continuation utilities when verification does not occur and the agent selects his preferred action, i.e., $u^R_{\neg V} < u$. Such punishments shift the relationship toward the firing region if the evidence-based region is empty, in which case the agent is eventually fired with probability one. By contrast, if the evidence-based region is nonempty, the principal verifies both actions with positive probability to detect

[32]Except when (Ve) binds, $\Pi'(u_V) \leq \Pi'(u)$, which implies $u_V > u$ if $\Pi(\cdot)$ is strictly concave at u. Therefore, even in the hybrid region, the agent might also be rewarded through continuation utility if verification takes place.

shirking, and firing does not occur. Greater verification substitutes for firing and prevents its snowballing effect.

When the bias is large ($(1-\rho)b > c_A$), the evidence-based region is empty. If the benefit from the agent's bias ($(1-\rho)b$) exceeds the cost of working (c_A), the agent prefers to comply rather than shirk and choose L, i.e., $w + (1-\rho)b - c_A > w$. Otherwise, (M–L) does not bind, and the evidence-based region is nonempty. In this case, the evidence-based region serves as a barrier to firing.

Principal's preferred equilibrium. The dynamics of principal's preferred equilibrium follow from theorems 1-3. As discussed earlier, the relationship begins in the verification region, i.e., $u^* \in [u_e, u_f]$. We consider three cases.

First, suppose the evidence-based region is empty ($\tilde{u} < u_f$). In this case, the entire verification region falls within the hybrid region. If the agent chooses L, the principal does not verify and the continuation utility increases, i.e., $u^L_{\neg V} > u$. The relationship may remain in the verification region, but with a lower probability of verification, or it may transition to the reduced-workload region. If it transitions to the reduced-workload region, the game may move either to u_e or to the agent's preferred equilibrium, $\overline{u}$.[33] Once the relationship enters the reduced-workload region, however, it ultimately returns to the hybrid region with probability one.

If the agent instead chooses R, the principal verifies with positive probability. If verification does not occur, the agent is punished through a reduction in continuation utility, i.e., $u^R_{\neg V} < u$. The relationship then either transitions to the firing region or remains in the verification region but with a higher probability of verification. Ultimately, the agent is fired with probability one. Figure 10 illustrates these transitions when the evidence-based region is empty.

Second, suppose the hybrid region is empty ($\tilde{u} \geq u_e$). The relationship begins in the evidence-based region, where the principal verifies after both actions, though more frequently following R. In this region, continuation utility increases only when the principal verifies and the agent provides evidence. This, in turn, implies that verification in the next period occurs with lower probability for both actions. If verification does not occur, continuation utilities remain unchanged ($u^L_{\neg V} = u^R_{\neg V} = u$), and the probability of verification does not change from one period to the next. Following verification, the relationship either remains in the evidence-based region or transitions to the reduced-workload region. Once it moves into the reduced-workload region, it remains there indefinitely. Figure 11 illustrates the equilibrium-path transitions when the hybrid region is empty.

Third, suppose both the evidence-based and hybrid regions are nonempty. If $u^* \in [u_f, \tilde{u}]$, the relationship begins in the evidence-based region. After verification,

[33] When we say the relationship moves to u, we mean it moves to the game associated with continuation utility u.

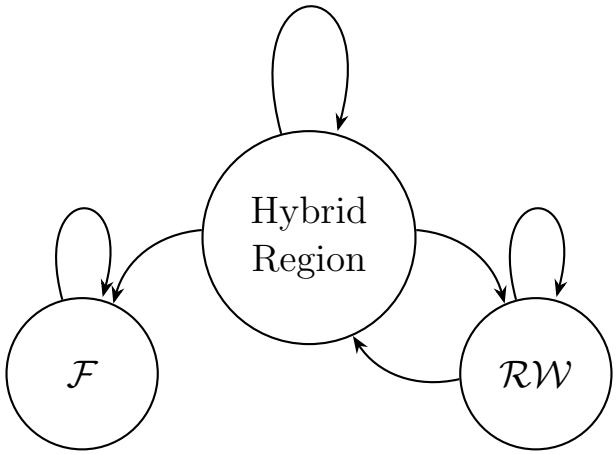

Figure 10: Equilibrium-path transitions between regions when the evidence-based region is empty.

it either remains in this region, transitions to the hybrid region, or moves to the reduced-workload region. Once the relationship leaves the evidence-based region, it cycles between the hybrid and reduced-workload regions but never returns to the evidence-based region. Figure 12 illustrates this case.

If $u^* \in [\tilde{u}, u_e]$, the relationship begins in the hybrid region, and the dynamics follow the same pattern as in the previous case after the transition to the hybrid region. In this third case, firing never occurs.

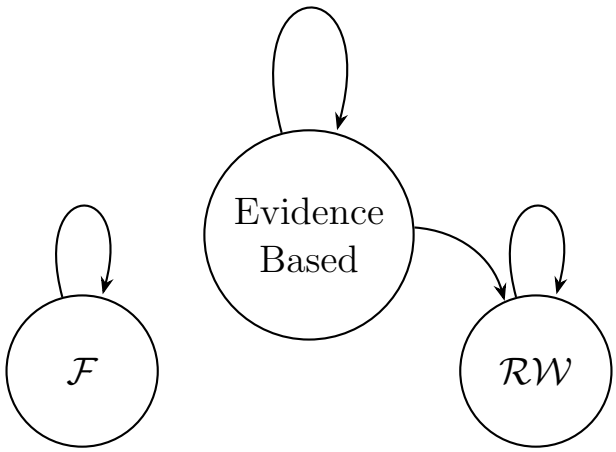

Figure 11: Equilibrium-path transitions between regions when the hybrid region is empty.

5.3 Discussion

Trust. A possible interpretation of our model is to view the absence of verification as trust: the less frequently the principal verifies, the greater the trust in the agent. More precisely, when the agent is required to work and the principal chooses not to verify, we interpret this as an act of trust.

Our results answer the following questions: how often should the principal trust and delegate decisions to the agent, and how does trust evolve over time? How is the evolution of trust related to the agent's bias?

The results show that trust builds gradually. In the evidence-based trust region, the principal never fully trusts the agent's actions, even when he chooses the less-preferred action. Nevertheless, she places relatively more trust in that case. In this region, trust increases only after the agent provides evidence—that is, the probability of verification decreases in subsequent periods.

In the hybrid trust region, the principal fully trusts the agent when he takes L, and trust increases in the following period. By contrast, if the agent takes R and verification does not occur, trust decreases. In this region, trust is shaped by both hard information (evidence obtained through verification) and soft information (the agent's actions when not verified).

At the lowest level of trust—when verification is most frequent—the relationship is fragile and may terminate. At the highest level of trust—when verification is rare—the agent benefits from a reduced workload.

Commitment. When the principal is assumed to have commitment power, the results remain largely the same. In this case, the principal's problem coincides with $\mathcal{P}$ except that the two constraints (Ve) and (Fi) are no longer required, since the

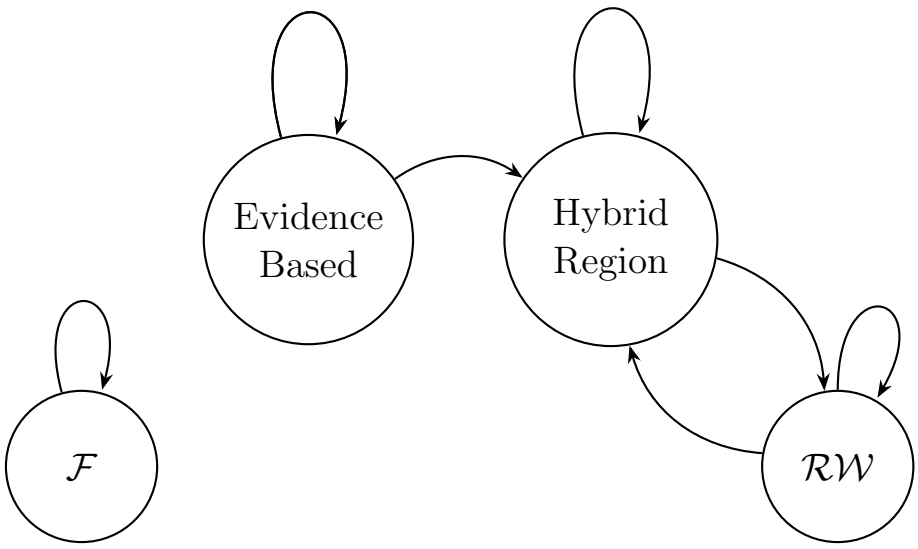

Figure 12: Equilibrium-path transitions between regions when both the hybrid and the evidence based regions are non empty.

principal's own incentives need not be considered. Theorems 1, 2, and 3 continue to hold, with the exception that the highest continuation utility is $\frac{w+b}{1-\delta}$, and the reduced-workload region may become piecewise linear (with a kink at $\frac{w}{1-\delta}$), in contrast to the linear form that arises in the no-commitment case.

Transfer. Suppose that, instead of the fixed transfer w, the principal can use a flexible and positive transfer $T \geq 0$ at the end of each period. Since transfers are a more efficient instrument, the principal never relies on reduced workload or firing on the equilibrium path. The upper boundary is linear and the principal's preferred equilibrium is stationary.

Denote T_V^a as the transfer when the agent takes action a and the principal verifies, and $T_{\neg V}^a$ as the transfer when the agent takes action a and the principal does not verify. As in the model without transfers ($u_V = u_V^L = u_V^R$), it is straightforward to show that $T_V := T_V^L = T_V^R$.

When verification occurs, transfers provide incentives for working, i.e., $(\rho v^R + (1-\rho)v^L)T_V = c_A$. When verification does not occur, transfers provide incentives to match the action with the state, i.e., $(1-v^R)T_{\neg V}^R = b + (1-v^L)T_{\neg V}^L$. The value function is linear, it follows that $T_{\neg V}^R = 0$, while the transfer when the agent takes L and the principal does not verify compensates for the agent's bias: $(1-v^L)T_{\neg V}^L = b$.

The principal's problem is therefore equivalent to minimizing her cost by selecting

transfers and verification probabilities:

$$\rho(1-v^L)T^L_{\neg V} + (\rho v^L + (1-\rho)v^R)T_V + c_P(\rho v^R + (1-\rho)v^R).$$

The first two terms are fixed by the incentive constraints. Hence the principal minimizes the last term subject to her own incentive constraints:

$$\text{verification:}\quad \delta\Pi - T_V - c_P \geq 0, \qquad \text{firing:}\quad \delta\Pi - T^L_{\neg V} \geq 0.$$

The verification constraint always binds; otherwise one could reduce v^R and increase T_V. Hence, the transfer after verification equals the principal's entire future payoff minus the current cost of verification. It follows that

$$\rho v^L + (1-\rho)v^R = \frac{c_A}{\delta\Pi - c_P}.$$

Therefore, Π satisfies

$$\Pi = \pi - b - c_A - c_P\,\frac{c_A}{\delta\Pi - c_P}.$$

Finally, two conditions must be satisfied. First, the expected probability of verification must not exceed one, $\rho v^R + (1-\rho)v^L \leq 1$, or, equivalently (using the verification incentive constraint), $\delta\Pi \geq c_A + c_P$. Second, the firing constraint requires $\delta\Pi(\delta\Pi - c_A - c_P) \geq (\delta\Pi - c_P)\,\rho b$. For sufficiently large π, both conditions are satisfied.

The agent's bias is toward the more likely state. Suppose the agent is biased toward the more likely state, i.e., $u(L) = w + b$ and $u(R) = w$. In this case, the preferences of the principal and the agent are congruent in terms of the action that should be taken when the state is unknown: both prefer L. Consequently, when the agent is expected to shirk (in the reduced-workload region), he chooses L. Aside from this difference, the structure of the principal's preferred equilibrium remains unchanged, and Theorems 1, 2, and 3 continue to apply.

6 Conclusion

We studied an infinite-horizon principal–agent relationship in which the agent exerts costly effort to learn a privately observed binary state before choosing an action. The principal wants action–state alignment, while the agent is biased toward one action. As a result, the principal faces an adverse selection problem followed by a moral hazard problem. The principal cannot observe payoffs directly but disciplines the agent through verification, workload adjustments, and termination. Continuation utilities summarize history. The upper boundary of the equilibrium payoff set is self-generating, allowing the problem to be formulated as a recursive optimization subject to the objectives of both the agent and the principal.

We showed that equilibrium behavior along the upper boundary of payoffs divides into three regions. Firing region (low continuation utilities), with firing probabilities declining as utility rises. Verification region (intermediate utilities), applied more frequently after the biased action. Reduced-workload region (high continuation utilities), with working probabilities declining as utility increases.

In the evidence-based region, continuation utilities evolve only through hard information from verification. The principal verifies after both actions with positive probability, including when the agent chooses his less-preferred action to ensure effort. Since both probabilities are strictly positive, the principal relies on differential verification rather than distorted continuation utilities. In contrast, in the hybrid region, continuation utilities evolve through both hard information (verification) and soft information (unverified actions). Here, the principal never verifies after L.

We showed that snowballing leads the principal to tolerate shirking even when the loss from shirking is high and the cost of verification is negligible. When the agent's bias is high, the threat of future firing complements verification and reduces verification costs across multiple periods. In contrast, for small biases, the evidence-based region is nonempty, and the principal must verify after both actions with positive probability. This higher level of verification substitutes for firing and prevents its snowballing effect.

References

Dhruva Bhaskar. Dynamic screening and the dual roles of monitoring. *Available at SSRN 2689503*, 2024.

Simon Board. Relational contracts and the value of loyalty. *American Economic Review*, 101(7):3349–3367, 2011.

Liang Dai, Yenan Wang, and Ming Yang. Dynamic contracting with flexible monitoring. *Available at SSRN 3496785*, 2024.

Arash Fahim, Simon Gervais, and R. Vijay Krishna. Equity prices and the dynamics of corporate governance. Working paper, 2021.

Kyna Fong. Evaluating skilled experts: Optimal scoring rules for surgeons. Working paper, 2008.

Yingni Guo and Johannes Hörner. Dynamic allocation without money. 2021.

Marina Halac. Relational contracts and the value of relationships. *American Economic Review*, 102(2):750–779, 2012.

Marina Halac and Andrea Prat. Managerial attention and worker performance. *American Economic Review*, 106(10):3104–3132, 2016.

Jonathan Levin. Relational incentive contracts. *American Economic Review*, 93 (3):835–857, 2003.

Jin Li, Niko Matouschek, and Michael Powell. Power dynamics in organizations. *American Economic Journal: Microeconomics*, 9(1):217–241, 2017.

Elliot Lipnowski and Joao Ramos. Repeated delegation. *Journal of Economic Theory*, 188:105040, 2020.

Qingmin Liu. Information acquisition and reputation dynamics. *The Review of Economic Studies*, 78(4):1400–1425, 2011.

Iván Marinovic, Andrzej Skrzypacz, and Felipe Varas. Dynamic certification and reputation for quality. *American Economic Journal: Microeconomics*, 10(2): 58–82, 2018.

Radoslawa Nikolowa. Motivate and select: Relational contracts with persistent types. *Journal of Economics & Management Strategy*, 26(3):624–635, 2017.

Dmitry Orlov. Frequent monitoring in dynamic contracts. *Journal of Economic Theory*, 206:105550, 2022.

Tomasz Piskorski and Mark M. Westerfield. Optimal dynamic contracts with moral hazard and costly monitoring. *Journal of Economic Theory*, 166:242–281, 2016.

Debraj Ray. The time structure of self-enforcing agreements. *Econometrica*, 70(2): 547–582, 2002.

Eilon Solan and Chang Zhao. Dynamic monitoring under resource constraints. *Games and Economic Behavior*, 129:476–491, 2021.

Felipe Varas, Iván Marinovic, and Andrzej Skrzypacz. Random inspections and periodic reviews: Optimal dynamic monitoring. *The Review of Economic Studies*, 87(6):2893–2937, 2020.

Yu Fu Wong. Dynamic monitoring design. *Available at SSRN 4466562*, 2023.

Yishu Zeng. *On Information Policy Design and Strategic Interactions*. PhD thesis, Stanford University, 2022.

7 Appendix

Let G denote the game presented in Section 3. Let $\tilde{\mathcal{E}}$ denote the set of all equilibria of G, and let $\mathcal{E} \subseteq \tilde{\mathcal{E}}$ be the subset of equilibria in which the agent matches the action to the state whenever the agent works. In Section 7.1 we provide proofs of the results in Section 5, which characterize the principal's preferred equilibrium in $\mathcal{E}$. Section 7.2 we show this focus is without loss of generality. Formally we show the principal's preferred equilibrium of $\tilde{\mathcal{E}}$ is also an equilibrium in $\mathcal{E}$.

7.1 Proofs of Results in Section 5

We begin by stating some claims; we then use these claims to establish the proofs of the theorems.

Define the Lagrangian associated with problem $\mathcal{P}$. Let f denote the probability of firing, e the probability of working, and $\alpha^R_{\neg W}$ the probability of choosing R when the agent is not supposed to work. Using Lemma 3 (iii), the constraints (M) and (WM) can be replaced by (I) and (W). Therefore, the Lagrangian is given by

$$\begin{aligned}
\mathcal{L} = \; (1-f)\Bigg[& e\Big(\pi - w + (\rho v^R + (1-\rho)v^L)\delta\,\Pi(u^R_V) + \rho(1-v^R)\delta\,\Pi(u^R_{\neg V}) \\
& + (1-\rho)(1-v^L)\delta\,\Pi(u^L_{\neg V}) - \big(\rho v^R + (1-\rho)v^L\big)c_p\Big) \\
& + (1-e)\Big(\pi p + \alpha^R_{\neg W}(1-2\rho)\pi + \delta\,\Pi(u_{\neg W})\Big)\Bigg] \\
& - \lambda_{PK}\Bigg[u - (1-f)\Big(e\Big(w + (1-\rho)b + (\rho v^R + (1-\rho)v^L)\delta u_V + \rho(1-v^R)\delta u^R_{\neg V} \\
& + (1-\rho)(1-v^L)\delta u^L_{\neg V}\Big) + (1-e)\big(w + \alpha^R_{\neg W} b + \delta u_{\neg W}\big)\Big)\Bigg] \\
& - \gamma\Big((\rho v^R + (1-\rho)v^L)\delta u_{\neg V} - c_A\Big) \\
& - \eta\Big(w + \delta v^L u^L_{\neg V} + \delta(1-v^L)u_V - \big(w + b + \delta(1-v^R)u^R_{\neg V}\big)\Big). \qquad \text{(L)}
\end{aligned}$$

Since the value function is concave, the firing constraint (Fi) is equivalent to assuming continuation utilities are non-negative and do not exceed $\overline{u}$. Moreover, the constraint (Ve) implies that $u_V \in [\underline{u}_V, \overline{u}_V]$ for some $\underline{u}_V, \overline{u}_V \in (0, \overline{u})$. Hence, the

additional constraints are

$$f, e, v^R, v^L, \alpha^R_{\neg W} \in [0,1], \quad u_V \in [\underline{u}_V, \overline{u}_V], \quad u^R_{\neg V}, u^L_{\neg V} \in [0, \overline{u}], \quad u_{\neg W} \in [0, \overline{u}].$$

Observe that the value function $\Pi(\cdot)$ is concave and therefore locally Lipschitz continuous on the interior of its effective domain. By Rademacher's theorem, any locally Lipschitz function is differentiable almost everywhere with respect to Lebesgue measure. This property is employed in the subsequent analysis. The following claims characterize both the solution to problem $\mathcal{P}$ and the associated value function $\Pi(\cdot)$.

Finally, note that in the linear segments of $\Pi(\cdot)$, the equilibrium can be implemented through randomization over the extreme points of the segment. Accordingly, in the proofs that follow we always assume u is not located in the interior of a linear segment of $\Pi(\cdot)$.

Claim 1 (Firing)**.** *There exists $u_f \in \left[w+b, \frac{w+b}{1-\delta}\right]$ such that $f > 0$ if and only if $u < u_f$. Moreover, $\Pi(u) = a_f u$ for some $a_f > 0$ and $f = 1 - \frac{u}{u_f}$ for $u \leq u_f$.*

Proof. If $f = 1$, then the value of $\mathcal{P}$ is zero, and by equation (PK) we obtain $u = 0$. Suppose $f < 1$. The first-order condition with respect to f implies

$$\frac{\Pi(u)}{1-f} + \lambda_{PK} \frac{u}{1-f} = 0,$$

which is equivalent to $\Pi(u) + \lambda_{PK} u = 0$ for $f \in (0,1)$. Furthermore, if $\Pi(u) + \lambda_{PK} u > 0$, then $f = 0$. Define

$$u_f \coloneqq \inf\{u \mid \Pi(u) - \Pi^+(u)u > 0\},$$

where $\Pi^+(u)$ denotes the right derivative. By concavity of $\Pi(\cdot)$, it follows that

$$\Pi(u_1) - \Pi^-(u_1)u_1 \;\geq\; \Pi(u_f) - \Pi^-(u_1)u_f \;\geq\; \Pi(u_f) - \Pi^+(u_f)u_f,$$

for all $u_1 > u_f$, where $\Pi^-(\cdot)$ denotes the left derivative. By the Subdifferential Envelope Theorem, we have $\Pi^-(u_1) \geq -\lambda_{PK}$. Thus,

$$\Pi(u_1) + \lambda_{PK} u_1 \;\geq\; \Pi(u_1) - \Pi^-(u_1)u_1 \;>\; 0,$$

which implies that $f = 0$ for all $u_1 > u_f$.

Moreover, optimality with respect to f requires

$$\Pi(u_2) - \Pi^+(u_2)u_2 \;\geq\; \Pi(u_2) + \lambda_{PK} u_2 \;\geq\; 0,$$

for all $u_2 \leq u_f$. By the definition of u_f, it follows that $\Pi(u_2) - \Pi^+(u_2)u_2 = 0$ for all

$u_2 < u_f$. This condition implies that $\Pi(u)$ is differentiable for all $u \leq u_f$. Hence, for all $u_2 < u_f$ we have $\Pi(u_2) - \Pi'(u_2)u_2 = 0$, which yields $\Pi(u) = a_f u$ for some $a_f > 0$.

Since the value function is linear on the interval $[0, u_f]$, it can be implemented as a randomization between the two points 0 (with probability $1 - \frac{u}{u_f}$) and u_f (with probability $\frac{u}{u_f}$). Therefore, $f = 1 - \frac{u}{u_f}$. □

Claim 2. *For all $u \in [0, \overline{u}]$, $\Pi(u) - u\Pi^+(u) \geq 0$.*

Proof. Suppose, for contradiction, that for some $u \in [0, \overline{u}]$ we have $\Pi(u) - u\Pi^+(u) < 0$. The first-order condition with respect to f implies that if $\Pi(u) + \lambda_{PK}u < 0$, then $f = 1$. This can occur only for $u = 0$; therefore, for all $u > 0$ we have $\Pi(u) + \lambda_{PK}u \geq 0$. The envelope theorem implies that $-\lambda_{PK} \in \partial\Pi(u)$, hence $-\lambda_{PK} \geq \Pi^+(u)$. Therefore,

$$\Pi(u) - \Pi^+(u)u \geq \Pi(u) + \lambda_{PK}u \geq 0.$$

□

Claim 3 (Continuation utility in case of not working, i.e., $u_{\neg W}$)**.**

1 *If $u_{\neg W} \leq u$, then either 1.1) $u_{\neg W} = u$ or 1.2) $\Pi^-(u_{\neg W}) = \Pi^+(u)$ and the value function is linear in $[u_{\neg W}, u]$.*

2 *If $u_{\neg W} \geq u$, then either 2.1) $u_{\neg W} = u$ or 2.2) $\Pi^+(u_{\neg W}) = \Pi^-(u)$ and the value function is linear in $[u, u_{\neg W}]$ or 2.3) $u_{\neg W} = \overline{u}$.*

Proof. Without loss of generality, we can assume $f = 0$ and $e < 1$. Suppose $u_{\neg W} < \overline{u}$. The first-order condition with respect to $u_{\neg W}$ gives us $-\lambda_{PK} \in \partial\Pi(u_{\neg W})$. By the Subdifferential Envelope Theorem, $-\lambda_{PK} \in \partial\Pi(u)$.

Suppose $u \geq u_{\neg W}$. Then, by the concavity of $\Pi(\cdot)$, either 1) $u_{\neg W} = u$ or 2) $\Pi^-(u_{\neg W}) = \Pi^+(u)$ and the value function is linear in $[u_{\neg W}, u]$.

Suppose $u \leq u_{\neg W}$. Then, by the concavity of $\Pi(\cdot)$, either 1) $u_{\neg W} = u$ or 2) $\Pi^+(u_{\neg W}) = \Pi^-(u)$ and the value function is linear in $[u, u_{\neg W}]$. Now suppose $u_{\neg W} = \overline{u}$. The first-order condition with respect to $u_{\neg W}$ implies $\Pi'(\overline{u}) \leq -\lambda_{PK}$. □

Claim 4 (Action in case of not working, i.e., $\alpha^R_{\neg W}$)**.** *There exists $\check{u} \in [u_e, \overline{u}]$ such that for all $u \leq \check{u}$, $\alpha^R_{\neg W} \in [0, 1)$, and for all $u > \check{u}$, $\alpha^R_{\neg W} = 1$. Moreover, $(1 - 2\rho)\pi + \lambda_{PK}b \geq 0$ if $\alpha^R_{\neg W} = 1$, $(1 - 2\rho)\pi + \lambda_{PK}b = 0$ if $\alpha^R_{\neg W} \in (0, 1)$, and $(1 - 2\rho)\pi + \lambda_{PK}b \leq 0$ if $\alpha^R_{\neg W} = 0$.*

Proof. Without loss of generality, we can assume $f = 0$ and $e < 1$. The first-order condition with respect to $\alpha^R_{\neg W}$ gives $(1 - 2\rho)\pi + \lambda_{PK}b \geq 0$ if $\alpha^R_{\neg W} = 1$, $(1 - 2\rho)\pi + \lambda_{PK}b = 0$ if $\alpha^R_{\neg W} \in (0, 1)$, and $(1 - 2\rho)\pi + \lambda_{PK}b \leq 0$ if $\alpha^R_{\neg W} = 0$.

By the Subdifferential Envelope Theorem, $-\lambda_{PK} \in \partial\Pi(u)$. Therefore, by concavity of $\Pi(\cdot)$, there exists $\check{u} \in \left[0, \frac{w+b}{1-\delta}\right]$ such that for all $u \leq \check{u}$, $\alpha^R_{\neg W} \in [0,1)$, and for all $u > \check{u}$, $\alpha^R_{\neg W} = 1$. □

Claim 5 (Shirking region)**.** *$\Pi(\cdot)$ is linear on the interval $[u_e, \overline{u}]$. Moreover, $\alpha^R_{\neg W} = 1$ if and only if the slope of the value function is less than $\frac{\pi(1-2\rho)}{1-\delta}$; otherwise, $\alpha^R_{\neg W} = 0$. In addition, for $u \in [u_e, \overline{u}]$ suppose $\Pi(u)$ takes the form $\Pi(u) = au + d$ for some $a, d \in \mathbb{R}$. If $a \geq \frac{\pi(1-2\rho)}{1-\delta}$, then $a\left(\frac{w}{1-\delta}\right) + d = \frac{\rho\pi - w}{1-\delta}$, and if $a \leq \frac{\pi(1-2\rho)}{1-\delta}$, then $a\left(\frac{w+b}{1-\delta}\right) + d = \frac{(1-\rho)\pi - w}{1-\delta}$.*

Proof. Without loss of generality, we can assume $f = 0$ and $e \in (0,1)$. Let

$$A = \mathbb{E}\Big[\pi(\theta,a) - c_P\,\mathbb{1}_V - w + \delta\,\Pi(u') \;\Big|\; W\Big] - \mathbb{E}\Big[\pi(\theta,a) - c_P\,\mathbb{1}_V - w + \delta\,\Pi(u') \;\Big|\; \neg W\Big],$$

$$B = \mathbb{E}\big[u(a) - c_A\,\mathbb{1}_W + \delta\,u' \;\big|\; W\big] - \mathbb{E}\big[u(a) - c_A\,\mathbb{1}_W + \delta\,u' \;\big|\; \neg W\big].$$

The first-order condition with respect to e implies $A + \lambda_{PK}B = 0$. Using the definition of the value function,

$$\Pi(u) - \mathbb{E}\Big[\pi(\theta,a) - c_P\,\mathbb{1}_V - w + \delta\,\Pi(u') \;\Big|\; \neg W\Big] = eA.$$

In addition, by (PK),

$$e = \frac{u - \mathbb{E}\big[u(a) - c_A\,\mathbb{1}_W + \delta\,u' \;\big|\; \neg W\big]}{B}.$$

Therefore, by substituting e and using the first-order condition,

$$\Pi(u) - (\rho\pi + \alpha^R_{\neg W}(1-2\rho)\pi - w + \delta\Pi(u_{\neg W})) = -\lambda_{PK}(u - w - \alpha^R_{\neg W}b - \delta u_{\neg W}). \tag{1}$$

Suppose $\Pi(\cdot)$ is differentiable at u. Then, by the envelope theorem, $\Pi'(u) = -\lambda_{PK}$. Using the above equation, we consider two cases. By Claim 4, there exists $\check{u} \in [u_e, \overline{u}]$ such that for all $u \leq \check{u}$, $\alpha^R_{\neg W} \in [0,1)$, and for all $u > \check{u}$, $\alpha^R_{\neg W} = 1$.

First, consider $u \leq \check{u}$. Using Claim 4, for all $u \leq \check{u}$ the above equation simplifies to

$$\Pi(u) - (-w + \rho\pi + \delta\Pi(u_{\neg W})) = \Pi'(u)\,(u - w - \delta u_{\neg W}).$$

By Claim 3, either $\Pi'(u) = \frac{\Pi(u_{\neg W}) - \Pi(u)}{u_{\neg W} - u}$, or $u_{\neg W} = u$, or $u_{\neg W} = \overline{u}$.

First, suppose $u_{\neg W} \neq \overline{u}$. Then

$$\Pi(u) + w - \rho\pi = \Pi'(u)\,(u - w) + \delta\Pi(u) - \delta u\Pi'(u).$$

This is a first-order ODE, and it is straightforward to show that $\Pi(\cdot)$ is affine. Suppose $\Pi(u)$ takes the form $\Pi(u) = a_1 u + d_1$ for $a_1, d_1 \in \mathbb{R}$, where $a_1\left(\frac{w}{1-\delta}\right) + d_1 = \frac{\rho\pi - w}{1-\delta}$.

Now suppose $u_{\neg W} = \overline{u}$. Then

$$\Pi(u) + w - \rho\pi = \Pi'(u)\,(u - w - \delta\overline{u}).$$

This is a first-order ODE, and again it is straightforward to show that $\Pi(\cdot)$ is affine. Suppose $\Pi(u)$ takes the form $\Pi(u) = a_2 u + d_2$ for $a_2, d_2 \in \mathbb{R}$, where

$$a_2 \tfrac{w}{1-\delta} + d_2 = \tfrac{\rho\pi - w}{1-\delta} - \tfrac{\delta(a_2\overline{u} + d_2)}{1-\delta}.$$

Therefore, in both cases $\Pi(\cdot)$ is affine. Hence, for all $u \in [u_e, \check{u}]$, $\Pi(\cdot)$ is piecewise linear. However, it is straightforward to show it cannot have more than two linear parts. Formally, there exists $u_1 \in [u_e, \check{u}]$ such that $\Pi(u) = a_1 u + d_1$ for $u \in [u_e, u_1]$, and for all $u \in [u_1, \check{u}]$, $\Pi(u) = a_2 u + d_2$. Moreover, by Claim 4, $a_i \geq \frac{(1-2\rho)\pi}{b}$, and since $a_2\overline{u} + d_2 \geq 0$, in both cases $a_i \frac{w}{1-\delta} + d_i \leq \frac{\rho\pi - w}{1-\delta}$ for $i \in \{1, 2\}$.

If $\check{u} = \overline{u}$, then $a_2\overline{u} + d_2 = 0$. Since $\Pi(\cdot)$ is concave and both affine functions pass through the point $\left(\frac{w}{1-\delta}, \frac{\rho\pi - w}{1-\delta}\right)$, these two affine functions must coincide. Hence, for all $u \in [u_e, \overline{u}]$, $\Pi(u)$ takes the form $\Pi(u) = au + d$ for $a, d \in \mathbb{R}$, where $a\left(\frac{w}{1-\delta}\right) + d = \frac{\rho\pi - w}{1-\delta}$.

Second, consider $u \geq \check{u}$**.** Thus, $\alpha^R_{\neg W} = 1$. Equation 1 simplifies to

$$\Pi(u) - ((1-\rho)\pi - w + \delta\Pi(u_{\neg W})) = \Pi'(u)\,(u - w - b - \delta u_{\neg W}).$$

By Claim 3, either $\Pi'(u) = \frac{\Pi(u_{\neg W}) - \Pi(u)}{u_{\neg W} - u}$, or $u_{\neg W} = u$, or $u_{\neg W} = \overline{u}$.

First, suppose $u_{\neg W} \neq \overline{u}$. Then

$$\Pi(u) - \tfrac{(1-\rho)\pi - w}{1-\delta} = \Pi'(u)\left(u - \tfrac{w+b}{1-\delta}\right).$$

Thus, $\Pi(u)$ takes the form $\Pi(u) = a_3 u + d_3$ for $a_3, d_3 \in \mathbb{R}$, where $a_3\left(\frac{w+b}{1-\delta}\right) + d_3 = \frac{(1-\rho)\pi - w}{1-\delta}$.

Now suppose $u_{\neg W} = \overline{u}$. Then

$$\Pi(u) + w - (1-\rho)\pi = \Pi'(u)\,(u - w - b - \delta\overline{u}).$$

Thus, $\Pi(u)$ takes the form $\Pi(u) = a_4 u + d_4$ for $a_4, d_4 \in \mathbb{R}$, where

$$a_4 \tfrac{w+b}{1-\delta} + d_4 = \tfrac{(1-\rho)\pi - w}{1-\delta} - \tfrac{\delta(a_4\overline{u} + d_4)}{1-\delta}.$$

Similar to the first case, there exists $u_3 \in [\check{u}, \overline{u}]$ such that $\Pi(u) = a_3 u + d_3$ for

$u \in [\check{u}, u_3]$, and for all $u \in [u_3, \overline{u}]$, $\Pi(u) = a_4 u + d_4$. In addition, $a_4\overline{u} + d_4 = 0$. Therefore, these two affine functions must coincide. Hence, for all $u \in [\check{u}, \overline{u}]$, $\Pi(u)$ takes the form $\Pi(u) = a_3 u + d_3$ for $a_3, d_3 \in \mathbb{R}$, where $a_3\left(\frac{w+b}{1-\delta}\right) + d_3 = \frac{(1-\rho)\pi - w}{1-\delta}$.

If $\check{u} = u_e$, then for all $u \in [u_e, \overline{u}]$, $\Pi(u)$ takes the form $\Pi(u) = a_3 u + d_3$ for $a_3, d_3 \in \mathbb{R}$, where $a_3\left(\frac{w+b}{1-\delta}\right) + d_3 = \frac{(1-\rho)\pi - w}{1-\delta}$. Moreover, by Claim 4, $a_3 \leq \frac{(1-2\rho)\pi}{b}$.

Finally, suppose $\check{u} \in (u_e, \overline{u})$. We show that this case cannot occur. Combining the first and the second cases, $\Pi(\cdot)$ would be a piecewise linear function with three linear parts. Formally, $\Pi(u) = a_1 u + b_1$ for $u \in [u_e, u_1]$, $\Pi(u) = a_2 u + b_2$ for $u \in [u_1, \check{u}]$, and $\Pi(u) = a_3 u + b_3$ for $u \in [\check{u}, \overline{u}]$. Using the facts that $a_1 \frac{w}{1-\delta} + d_1 = \frac{\rho\pi - w}{1-\delta}$, $a_2 \frac{w}{1-\delta} + d_2 \leq \frac{\rho\pi - w}{1-\delta}$, $a_1, a_2 \geq \frac{(1-2\rho)\pi}{1-\delta}$, $a_3\left(\frac{w+b}{1-\delta}\right) + d_3 = \frac{(1-\rho)\pi - w}{1-\delta}$, and $a_3 \leq \frac{(1-2\rho)\pi}{1-\delta}$, a simple algebra shows that $\Pi(\cdot)$ cannot be concave. This leads to a contradiction. □

Claim 6. *Let $u_1 := \min\{u \mid \Pi^+(u) = \Pi'(\overline{u})\}$. Without loss of generality, the probability of working at $\overline{u}$, denoted by $e \in (0,1)$, satisfies*

$$\overline{u} = e u_1 + (1-e)\left(w + \alpha^R_{\neg W} b + \delta \overline{u}\right),$$

and the continuation utility in case the agent is not supposed to work ($u_{\neg W}$) is equal to $\overline{u}$. Moreover, continuation utilities and verification probabilities when the agent is supposed to work are the same as u_1.

Proof. By Claim 5, $\Pi(\cdot)$ is linear for all $u \in [u_e, \overline{u}]$. Hence, $\Pi(\cdot)$ is also linear for all $u \in [u_1, \overline{u}]$. Suppose $\Pi(u) = au + d$ for some $a, d \in \mathbb{R}$ on this interval. Define

$$e = \frac{\overline{u} - (w + \alpha^R_{\neg W} b + \delta \overline{u})}{u_1 - (w + \alpha^R_{\neg W} b + \delta \overline{u})}.$$

Note that $\overline{u} - (w + \alpha^R_{\neg W} b + \delta \overline{u}) < 0$ since $\overline{u} < \frac{w + \alpha^R_{\neg W} b}{1-\delta}$. Therefore, $e \in (0,1)$. Using $\Pi(\overline{u}) = a\overline{u} + d$, $\Pi(u_1) = au_1 + d$, and Claim 4, it follows that

$$\Pi(\overline{u}) = e\Pi(u_1) + (1-e)\left(-w + \rho\pi + \alpha^R_{\neg W}(1-2\rho)\pi + \delta\Pi(\overline{u})\right).$$

□

Claim 7. *Let $u \in [u_f, u_e]$ and suppose constraint (M–L) is slack. Then $u^L_{\neg V} = u^R_{\neg V} = u$.*

Proof. Since constraint (M–L) is not binding, we have $\eta = 0$. The first-order conditions of the Lagrangian with respect to $u^L_{\neg V}$ and $u^R_{\neg V}$ yield

$$-\lambda_{PK} \in \partial\Pi(u^L_{\neg V}), \qquad -\lambda_{PK} \in \partial\Pi(u^R_{\neg V}),$$

and by the envelope theorem, $-\lambda_{PK} \in \partial\Pi(u)$.

If $\Pi(\cdot)$ is strictly concave at u, it follows that $u^L_{\neg V} = u^R_{\neg V} = u$.

Now suppose instead that $\Pi(\cdot)$ is linear on a right neighborhood I^+ of u, so that $\Pi(x) = ax + d$ for $x \in I^+$. The envelope theorem then implies $-\lambda_{PK} \geq a$.

Consider the case $u^L_{\neg V} \geq u$. The derivative of the Lagrangian $\mathcal{L}$ with respect to $u^L_{\neg V}$ is $a + \lambda_{PK}$. If $-\lambda_{PK} > a$, then $u^L_{\neg V} = u$. Suppose instead that $-\lambda_{PK} = a$ and, for contradiction, assume $u^L_{\neg V} > u$. The value function is

$$\begin{aligned}\Pi(u) = \pi - w + (\rho v^R + (1-\rho)v^L)\,\delta\Pi(u^R_V) + \rho(1-v^R)\,\delta\Pi(u^R_{\neg V}) \\ + (1-\rho)(1-v^L)\,\delta\Pi(u^L_{\neg V}) - \big(\rho v^R + (1-\rho)v^L\big)c_p.\end{aligned}$$

Since (M–L) does not bind, the left derivative with respect to $u^L_{\neg V}$ implies $\Pi^-(u) = \Pi^-(u^L_{\neg V})$. Moreover, fixing other variables and lowering $u^L_{\neg V}$ extends the linearity of the value function beyond u, so that u lies in the interior of a linear segment of $\Pi(\cdot)$. However, without loss of generality we assume that u is not in the interior of a linear segment of $\Pi(\cdot)$. This yields a contradiction.

If instead $u^L_{\neg V} < u$, then a symmetric argument using a left neighborhood of u leads to the same contradiction.

The same reasoning applies to $u^R_{\neg V}$. Hence $u^L_{\neg V} = u^R_{\neg V} = u$. □

Claim 8. *Suppose that, for $u \in [u_f, u_e] \cap [\underline{u}_V, \overline{u}_V]$, (W) does not bind. Then $u_V = u$.*

Proof. The first-order condition of the Lagrangian with respect to u_V implies $-\lambda_{PK} \in \partial\Pi(u_V)$. By the envelope theorem, $-\lambda_{PK} \in \partial\Pi(u)$.

For a contradiction, suppose $u_V \neq u$. Then $\Pi(\cdot)$ must be linear on the interval $[\min\{u_V, u\}, \max\{u_V, u\}]$ with slope $-\lambda_{PK}$. Without loss of generality, assume $u_V > u$. The value function is

$$\begin{aligned}\Pi(u) = \pi - w + (\rho v^R + (1-\rho)v^L)\,\delta\Pi(u^R_V) + \rho(1-v^R)\,\delta\Pi(u^R_{\neg V}) \\ + (1-\rho)(1-v^L)\,\delta\Pi(u^L_{\neg V}) - \big(\rho v^R + (1-\rho)v^L\big)c_p.\end{aligned}$$

Since (W) is slack, the left-hand derivative with respect to u_V implies $\Pi^-(u) = \Pi^-(u_V)$. Moreover, holding other variables fixed and lowering u_V extends the linearity of the value function below u, so that u lies in the interior of a linear segment of $\Pi(\cdot)$. However, without loss of generality we assume that u is not in the interior of a linear segment of $\Pi(\cdot)$. This yields a contradiction.

The argument is analogous if $u_V < u$, again leading to a contradiction. □

Claim 9. *For all $u \in [u_f, u_e]$,*

$$\Pi^+(u) = \rho\Pi^+(u^L_{\neg V}) + (1-\rho)\,\Pi^+(u^R_{\neg V}) \quad \textit{and} \quad \Pi^-(u) = \rho\Pi^-(u^L_{\neg V}) + (1-\rho)\,\Pi^-(u^R_{\neg V}).$$

Proof. By the second part of Lemma 3,

$$w + \delta(1 - v^L)u^L_{\neg V} = w + b + \delta(1 - v^R)u^R_{\neg V}.$$

Substitute

$$u = w + (1 - v^L)\delta u^L_{\neg V} + (\rho v^L + (1 - \rho)v^R)\delta u_V - c_A, \quad u^R_{\neg V} = \tfrac{-b+\delta(1-v^L)u^L_{\neg V}}{\delta(1-v^R)}$$

into the expression

$$\begin{aligned}\Pi(u) =& \pi - w + (\rho v^R + (1 - \rho)v^L)\,\delta\Pi(u_V) + \rho(1 - v^R)\,\delta\Pi(u^R_{\neg V}) \\ &+ (1 - \rho)(1 - v^L)\,\delta\Pi(u^L_{\neg V}) - \big(\rho v^R + (1 - \rho)v^L\big)c_p.\end{aligned}$$

Taking the right and left derivatives with respect to $u^L_{\neg V}$ yields the stated result. □

Claim 10. *Suppose $v^R > 0$ and $u_V \in (\underline{u}_V, \overline{u}_V)$. Then*

$$\Pi(u_V) - u_V\Pi^-(u_V) - \big[\Pi(u^R_{\neg V}) - u^R_{\neg V}\Pi^+(u^R_{\neg V})\big] \leq \tfrac{c_P}{\delta},$$

$$\Pi(u_V) - u_V\Pi^+(u_V) - \big[\Pi(u^R_{\neg V}) - u^R_{\neg V}\Pi^-(u^R_{\neg V})\big] \geq \tfrac{c_P}{\delta}.$$

Proof. First suppose (W) binds. We have $e > 0$. Using (I) and the fact that (W) binds,

$$u^L_{\neg V} = \frac{\frac{u}{1-f} - (1 - e)(w + \alpha^R_{\neg W}b + \delta u_{\neg W}) - ew}{e\delta(1 - v^L)},$$

and

$$u^R_{\neg V} = \frac{\frac{u}{1-f} - (1 - e)(w + \alpha^R_{\neg W}b + \delta u_{\neg W}) - ew}{e\delta(1 - v^R)}.$$

Now substitute $u^L_{\neg V}$, $u^R_{\neg V}$, and $u_V = \frac{c_A}{\delta((1-\rho)v^R+\rho v^L)}$ into the objective

$$\begin{aligned}(1 - f)\Bigg[e\Big(&\pi - w + ((1 - \rho)v^R + \rho v^L)\delta\,\Pi(u_V) + (1 - \rho)(1 - v^R)\delta\,\Pi(u^R_{\neg V}) \\ &+ \rho(1 - v^L)\delta\,\Pi(u^L_{\neg V}) - \big((1 - \rho)v^R + \rho v^L\big)c_p\Big) \\ &+ (1 - e)\Big(\pi p + \alpha^R_{\neg W}(1 - 2\rho)\pi + \delta\,\Pi(u_{\neg W})\Big)\Bigg].\end{aligned}$$

The left and right derivatives with respect to v^R yield the result. Note that the argument is independent of whether ((M–L)) binds, or equivalently, of whether $v^L = 0$.

Now suppose ((M–L)) binds; hence $v^L = 0$ and (W) does not bind. Using (I),

$$u^L_{\neg V} = \frac{\frac{u}{1-f} - (1-e)(w + \alpha^R_{\neg W} b + \delta u_{\neg W}) - e(w + (1-\rho)v^R \delta u_V - c_A)}{e\delta},$$

and

$$u^R_{\neg V} = \frac{\frac{u}{1-f} - (1-e)(w + \alpha^R_{\neg W} b + \delta u_{\neg W}) - e(w + (1-\rho)v^R \delta u_V - c_A)}{e\delta(1-v^R)}.$$

Now substitute $u^L_{\neg V}$ and $u^R_{\neg V}$ into the objective. Employing Claim 8 and Claim 9,

$$\Pi^+(u_V) = \rho\Pi^+(u^L_{\neg V}) + (1-\rho)\,\Pi^+(u^R_{\neg V}) \quad \text{and} \quad \Pi^-(u_V) = \rho\Pi^-(u^L_{\neg V}) + (1-\rho)\,\Pi^-(u^R_{\neg V}).$$

The left and right derivatives with respect to v^R yield the result. □

Claim 11. *Suppose $u_f \leq \frac{w}{1-\delta}$. Then for every $u \in \left[u_f, \min\{\frac{w}{1-\delta}, u_e\}\right]$ such that $\frac{c_A}{\delta - 1 + \frac{w+(1-\rho)b}{u}} \in [\underline{u}_V, \overline{u}_V]$, the following statements hold:*

1 $u^L_{\neg V} = u^R_{\neg V} = u$.

2 $v^L = 1 - \frac{1}{\delta} + \frac{w}{\delta u}$ *and* $v^R = 1 - \frac{1}{\delta} + \frac{w+b}{\delta u}$.

3 $u_V = \frac{c_A}{\delta - 1 + \frac{w+(1-\rho)b}{u}} > u + \Delta$.

Proof. Consider the relaxed problem obtained by dropping constraint (M–L) from the Lagrangian $\mathcal{L}$. A mean-preserving contraction argument in this relaxed problem implies that, without loss of generality, we may impose $u^L_{\neg V} = u^R_{\neg V}$. Let $u_{\neg V} := u^L_{\neg V} = u^R_{\neg V}$ and define $v := \rho v^R + (1-\rho)v^L$.

First, note that constraint (W) binds. Otherwise, reduce v to some $v' < v$ so that (W) binds at v'. Next set the continuation utilities as follows:

- upon verification (which now occurs with probability v'), set the continuation utility to u_V;
- upon no verification (probability $1 - v'$), assign $u_{\neg V}$ with probability $1 - v$ and assign u_V with the remaining probability $v - v'$.

This modification leaves the distribution of continuation utilities—and hence the agent's expected utility—unchanged, while strictly increasing the principal's objective because the verification probability falls. Therefore (W) must bind at the optimum.

(1) Since (M–L) is slack and by Claim 7, it follows that $u_{\neg V} = u$.

(2) Using (PK) together with (I), we obtain

$$v^L \;=\; 1 - \frac{1}{\delta} + \frac{w}{\delta u} \qquad \text{and} \qquad v^R \;=\; 1 - \frac{1}{\delta} + \frac{w+b}{\delta u}.$$

(3) Because (W) binds, and using part (2), we have

$$u_V \;=\; \frac{c_A}{\delta - 1 + \frac{w+(1-\rho)b}{u}} \in [\underline{u}_V, \overline{u}_V].$$

We first show $u_V \geq u$. Suppose, to the contrary, that $u_V < u$. Dual feasibility (KKT) then implies $\gamma \leq 0$. The first-order condition with respect to u_V yields

$$-\lambda_{PK} + \gamma \;\in\; \partial\Pi(u_V),$$

and since $u_V \in [\underline{u}_V, \overline{u}_V]$, constraint (Ve) is slack.[34] By concavity of Π on $[\underline{u}_V, \overline{u}_V]$, together with $u_V < u_{\neg V} = u$, we conclude $\gamma \leq 0$ and thus $\gamma = 0$. Hence either $u_V = u_{\neg V} = u$ (contradicting $u_V < u$), or

$$\frac{\Pi(u_V) - \Pi(u_{\neg V})}{u_V - u_{\neg V}} \;=\; -\lambda_{PK}.$$

The first-order condition with respect to v in $\mathcal{L}$ is

$$\big(\Pi(u_V) - \Pi(u_{\neg V})\big) + \lambda_{PK}(u_V - u_{\neg V}) - \gamma u_V - \frac{c_P}{\delta} \;=\; 0.$$

Combining the last two displays yields $\gamma u_V - \frac{c_P}{\delta} = 0$. Since $\gamma = 0$, this implies $\frac{c_P}{\delta} = 0$, a contradiction. Therefore $u_V \geq u$.

We now show: there exists $\Delta > 0$ such that $u_V \geq u + \Delta$. Suppose not. Then there exists $\check{u}$ such that either $u_V(\check{u}) = \check{u}$, or there is a sequence $\{\tilde{u}_i\}$ with $\tilde{u}_i \downarrow \check{u}$ and $u_V(\tilde{u}_i) \to \check{u}$. In either case, using $u_{\neg V} = u$, one obtains

$$\check{u} \;=\; \frac{w + (1-\rho)b - c_A}{1-\delta}.$$

By Claim 10 and $u_{\neg V} = u$, for all $u > \check{u}$,

$$-\big(\Pi(u) - u\Pi^-(u)\big) + \big(\Pi(u_V) - u_V\Pi^+(u_V)\big) \;\geq\; \frac{c_P}{\delta}.$$

Consider a decreasing sequence $\{u_i\}$ with $u_{i+1} = u_V(u_i)$ and $u_i \downarrow \check{u}$. Since

$$\Pi(u_i) - u_i\Pi^+(u_i) \;\leq\; \Pi(u_{i+1}) - u_{i+1}\Pi^-(u_{i+1}),$$

[34] Because $u_V \in [\underline{u}_V, \overline{u}_V]$, (Ve) does not bind.

we obtain, for all i,

$$-\big(\Pi(u_i) - u_i\Pi^+(u_i)\big) \;+\; \big(\Pi(u_{i+2}) - u_{i+2}\Pi^+(u_{i+2})\big) \;\geq\; \frac{c_P}{\delta}.$$

By Claim 2, $\Pi(u) - u\Pi^+(u) \geq 0$ for all $u \in [0, \overline{u}]$. Hence

$$0 \;\geq\; \Pi(u_{2i}) - u_{2i}\Pi^+(u_{2i}) \;\geq\; \Pi(u_1) - u_1\Pi^+(u_1) - \frac{nc_p}{\delta},$$

which yields a contradiction for n large enough. Therefore $u_V > u$, and thus $u_V \geq u + \Delta$ for some $\Delta > 0$.

Finally, since the solution to the relaxed problem satisfies (M–L) and (M–R), it is also a solution to the original problem $\mathcal{P}$. □

Claim 12. *If (M–L) is slack, then (W) binds.*

Proof. If (M–L) is slack, then $v^L > 0$. Suppose, for contradiction, that (W) does not bind. Decrease v^L while keeping the distribution of continuation utilities the same. Let the reduced value be $\tilde{v}^L$. With probability $\tilde{v}^L$, the continuation utility is u_V^L. With probability $v^L - \tilde{v}^L$, the continuation utility is u_V^L. With probability $1 - v^L$, the continuation utility is $u_{\neg V}^L$. The distribution of continuation utilities remains unchanged. Since (M–R) is slack, this change does not affect (M–R). It does not affect (M–L), (Ve), or (W), and it keeps the agent's utility constant while increasing the principal's payoff. This yields a contradiction. □

Claim 13. *The constraint $u_V \geq \underline{u}_V$ is always slack.*

Proof. Consider the relaxed problem obtained by ignoring the constraint $u_V \geq \underline{u}_V$. We show that, in this relaxed problem, $\underline{u}_V < u_f$. Since $u_V \geq u$ for all $u \in [u_f, \min\{u_e, \overline{u}_V\}]$, it follows that in the relaxed problem $u_V \geq \underline{u}_V$.

Suppose, for contradiction, that for some $u \geq u_f$ we have $u_V < \underline{u}_V$. Claim 10 implies that

$$\Pi(u_V) - u_V\Pi^+(u_V) - \big[\Pi(u_{\neg V}^R) - u_{\neg V}^R\Pi^-(u_{\neg V}^R)\big] \geq \tfrac{c_P}{\delta}.$$

By Claim 2, for all u we have $\Pi(u) - u\Pi^-(u) \geq 0$. Hence,

$$\Pi(u_V) - u_V\Pi^+(u_V) \geq \tfrac{c_P}{\delta}.$$

However, we know $\Pi(u_V) \leq \Pi(\underline{u}_V) = \frac{c_P}{\delta}$, and since $\underline{u}_V$ lies on the increasing part of $\Pi(\cdot)$, we have $\Pi^+(u_V) \geq 0$, yielding a contradiction. □

Claim 14. *If (Ve) binds, then (M–L) binds.*

Proof. Suppose, for contradiction, that (M–L) does not bind, and therefore $v^L > 0$. By Claim 12, constraint (W) binds, and by Claim 7, we have $u^R_{\neg V} = u^L_{\neg V} = u$. By (I),

$$v^L = 1 - \tfrac{1}{\delta} + \tfrac{w}{\delta u} \quad \text{and} \quad v^R = 1 - \tfrac{1}{\delta} + \tfrac{w+b}{\delta u}.$$

Since (W) binds, u_V can be expressed as a function of u. Therefore

$$u_V = \frac{c_A}{\delta - 1 + \frac{w+(1-\rho)b}{u}} = \overline{u}_V.$$

Hence, except possibly at a particular knife-edge value of u, this yields a contradiction. Therefore (M–L) must bind. □

Claim 15. *Suppose $v^R > 0$, and (W) does not bind. If $(1-\rho)u_V \geq u^R_{\neg V}$, then*

$$(u_V - \tfrac{u^R_{\neg V}}{1-\rho})\Pi^+(u) = \Pi(u_V) - \Pi(u^R_{\neg V}) - \tfrac{c_P}{\delta} - \tfrac{\rho}{1-\rho} u^R_{\neg V} \Pi^-(u^L_{\neg V}),$$

$$(u_V - \tfrac{u^R_{\neg V}}{1-\rho})\Pi^-(u) = \Pi(u_V) - \Pi(u^R_{\neg V}) - \tfrac{c_P}{\delta} - \tfrac{\rho}{1-\rho} u^R_{\neg V} \Pi^+(u^L_{\neg V}).$$

If $(1-\rho)u_V \leq u^R_{\neg V}$, then the direction reverses at u: in the preceding equations, replace $\Pi^+(u)$ with $\Pi^-(u)$ and replace $\Pi^-(u)$ with $\Pi^+(u)$.

Proof. We have $e > 0$. Using (I),

$$u^L_{\neg V} = \tfrac{b + \delta(1-v^R)u^R_{\neg V}}{\delta(1-v^L)}.$$

By the promise-keeping constraint,

$$\begin{aligned} u =& (1-f)\Big[e\Big(w + (1-\rho)b + ((1-\rho)v^R + \rho v^L)\delta u_V + (1-\rho)(1-v^R)\delta u^R_{\neg V} \\ &+ \rho(1-v^L)\delta u^L_{\neg V}\Big) - (1-e)\big(w + \alpha^R_{\neg W} b + \delta u_{\neg W}\big)\Big]. \end{aligned}$$

Now substitute u and $u^R_{\neg V}$ into the expression

$$\begin{aligned} \Pi(u) = (1-f)\Big[&e\Big(\pi - w + ((1-\rho)v^R + \rho v^L)\delta\,\Pi(u_V) + (1-\rho)(1-v^R)\delta\,\Pi(u^R_{\neg V}) \\ &+ \rho(1-v^L)\delta\,\Pi(u^L_{\neg V}) - \big((1-\rho)v^R + \rho v^L\big)c_p\Big) \\ &+ (1-e)\Big(\pi p + \alpha^R_{\neg W}(1-2\rho)\pi + \delta\,\Pi(u_{\neg W})\Big)\Big]. \end{aligned}$$

The left and right derivatives with respect to v^R yield the result. □

Claim 16. *Suppose (M–L) binds and W does not bind. Then it cannot be the case that* $u = u^L_{\neg V} = u^R_{\neg V} = u_V$.

Proof. Suppose $u = u^L_{\neg V} = u^R_{\neg V} = u_V$. Claim 15 implies

$$-\frac{\rho u}{1-\rho}\Pi^+(u) = -\frac{c_P}{\delta} - \frac{\rho u}{1-\rho}\Pi^-(u),$$

$$-\frac{\rho u}{1-\rho}\Pi^-(u) = -\frac{c_P}{\delta} - \frac{\rho u}{1-\rho}\Pi^+(u).$$

There is no $\Pi^+(u)$ and $\Pi^-(u)$ such that solves the above equations. A contradiction. □

Claim 17. *Suppose (Ve) binds and (W) does not bind. Then* $\Pi^-(u) = \Pi^-(\overline{u}_V)$. *Hence* $[\min\{u, \overline{u}_V\}, \max\{u, \overline{u}_V\}]$ *is a linear segment of* $\Pi(\cdot)$.

Proof. By the promise-keeping constraint,

$$\begin{aligned} u =& (1-f)\Big[e\Big(w + (1-\rho)b + ((1-\rho)v^R + \rho v^L)\delta u_V + (1-\rho)(1-v^R)\delta u^R_{\neg V} \\ &+ \rho(1-v^L)\delta u^L_{\neg V}\Big) - (1-e)\big(w + \alpha^R_{\neg W} b + \delta u_{\neg W}\big)\Big]. \end{aligned}$$

Now substitute u into the expression

$$\begin{aligned} \Pi(u) = (1-f)\Big[e\Big(&\pi - w + ((1-\rho)v^R + \rho v^L)\delta\,\Pi(u^R_V) + (1-\rho)(1-v^R)\delta\,\Pi(u^R_{\neg V}) \\ &+ \rho(1-v^L)\delta\,\Pi(u^L_{\neg V}) - \big((1-\rho)v^R + \rho v^L\big)c_p\Big) \\ &+ (1-e)\Big(\pi p + \alpha^R_{\neg W}(1-2\rho)\pi + \delta\,\Pi(u_{\neg W})\Big)\Big]. \end{aligned}$$

The left derivative with respect to u_V yields the result. □

Claim 18. *Suppose* $(1-\rho)b \geq c_A$. *Then* $u_f \geq \tilde{u}$.

Proof. Suppose $u \in [u_f, \tilde{u}]$. Then $u^L_{\neg V} = u^R_{\neg V} = u$ and $u_V > u$. By the promise-keeping constraint,

$$\begin{aligned} u =& w + (1-\rho)b + ((1-\rho)v^R + \rho v^L)\delta u_V + (1-\rho)(1-v^R)\delta u^R_{\neg V} + \rho(1-v^L)\delta u^L_{\neg V} - c_A \\ &> w + (1-\rho)b - c_A + \delta u. \end{aligned}$$

Hence $u > \frac{w+(1-\rho)b-c_A}{1-\delta} \geq \frac{w}{1-\delta}$, a contradiction. □

Claim 19. *Suppose $u_e \geq \frac{w}{1-\delta}$. Then for all $u \in [\max\{u_f, \frac{w}{1-\delta}\}, u_e]$,*

1) $v^R > v^L = 0$.

2) There exists $\Delta > 0$ such that $u^R_{\neg V} + \Delta < u < u^L_{\neg V} - \Delta$.

Proof. **Step 1: Show $v^L = 0$ (hence $v^R > v^L$).** Suppose, for contradiction, that $v^L > 0$. Then constraint (M–L) does not bind. By Claim 12, (W) binds. Moreover, since (M–L) does not bind, Claim 7 gives $u = u^L_{\neg V} = u^R_{\neg V}$. Using the promise-keeping constraint (PK), we obtain

$$v^L \;=\; 1 - \frac{1}{\delta} + \frac{w}{\delta u}.$$

Because $u > \frac{w}{1-\delta}$, this implies $v^L < 0$, a contradiction. Therefore $v^L = 0$ and (M–L) binds.

Step 2: Show $u^R_{\neg V} < u < u^L_{\neg V}$. We consider two cases according to the multiplier η.

Case A: $\eta = 0$. The derivatives of the Lagrangian $\mathcal{L}$ with respect to $u^L_{\neg V}$ and $u^R_{\neg V}$ imply

$$-\lambda_{PK} \in \partial\Pi(u^L_{\neg V}) \quad \text{and} \quad -\lambda_{PK} \in \partial\Pi(u^R_{\neg V}),$$

and the envelope theorem implies $-\lambda_{PK} \in \partial\Pi(u)$. Since the value function is concave, if $u \neq u^a_{\neg V}$ then $\Pi(\cdot)$ is linear on the interval $[\min\{u, u^a_{\neg V}\}, \max\{u, u^a_{\neg V}\}]$ for $a \in \{L, R\}$. Without loss of generality, we assume u is not in the interior of a linear segment of $\Pi(\cdot)$. Claim 9 then implies $u^L_{\neg V} = u = u^R_{\neg V}$.

If (W) binds, straightforward algebra shows $v^L > 0$ and $u \leq \frac{w}{1-\delta}$, a contradiction; hence (W) does not bind. If (Ve) does not bind, then by Claim 8 we conclude $u^L_{\neg V} = u = u^R_{\neg V} = u_V$, which contradicts Claim 16.

Now suppose (Ve) binds.

- If $(1-\rho)u_V \geq u$, then Claim 15 implies

  $$(\overline{u}_V - \tfrac{u}{1-\rho})\Pi^+(u) = -\Pi(u) - \tfrac{\rho u}{1-\rho}\Pi^-(u), \qquad (\overline{u}_V - \tfrac{u}{1-\rho})\Pi^-(u) = -\Pi(u) - \tfrac{\rho u}{1-\rho}\Pi^+(u).$$

  Therefore $\overline{u}_V = \frac{u(1+\rho)}{1-\rho}$ and $\frac{\rho u}{1-\rho}(\Pi^+(u) + \Pi^-(u)) = -\Pi(u)$. Using Claim 17, since $u_V > u$, we have $\Pi^+(u) = \Pi^-(u)$. Hence

  $$\Pi'(u) = \frac{-\Pi(u)}{\frac{2\rho u}{1-\rho}} = \frac{\Pi(u) - \frac{c_P}{\delta}}{u - \overline{u}_V} = \frac{\Pi(u) - \frac{c_P}{\delta}}{u - \frac{u(1+\rho)}{1-\rho}},$$

  which yields $\Pi(u) = \Pi(u) - \frac{c_P}{\delta}$, a contradiction.

- If $(1-\rho)u_V < u$, then Claim 15 implies

  $$(\overline{u}_V - \tfrac{u}{1-\rho})\Pi^-(u) = -\Pi(u) - \tfrac{\rho u}{1-\rho}\Pi^-(u), \qquad (\overline{u}_V - \tfrac{u}{1-\rho})\Pi^+(u) = -\Pi(u) - \tfrac{\rho u}{1-\rho}\Pi^+(u).$$

  Therefore $\Pi^-(u) = \Pi^+(u) = \frac{-\Pi(u)}{\overline{u}_V - u}$. By Claim 17, $[\min\{u, \overline{u}_V\}, \max\{u, \overline{u}_V\}]$ is a linear segment of $\Pi(\cdot)$. Hence $\Pi'(u) = \frac{\frac{c_P}{\delta} - \Pi(u)}{\overline{u}_V - u}$. This contradicts the earlier identity $\Pi'(u) = \frac{-\Pi(u)}{\overline{u}_V - u}$. Therefore we may conclude $\eta < 0$.

Case B: $\eta < 0$. The derivative of the Lagrangian $\mathcal{L}$ with respect to $u^L_{\neg V}$ implies $-\lambda_{PK} + \eta \in \partial\Pi(u^L_{\neg V})$, and with respect to $u^R_{\neg V}$ it implies $-\lambda_{PK} - \eta \in \partial\Pi(u^R_{\neg V})$. By concavity of Π, we have $u^R_{\neg V} \leq u \leq u^L_{\neg V}$. Therefore, if $\Pi(\cdot)$ is differentiable at u, it follows that $u^R_{\neg V} < u < u^L_{\neg V}$. Suppose instead that $\Pi(\cdot)$ is not differentiable at u. Consider three cases:

1. $u^R_{\neg V} = u = u^L_{\neg V} = u_V$. If (W) binds, straightforward algebra shows $v^L > 0$ and $u \leq \frac{w}{1-\delta}$, a contradiction; hence (W) does not bind. If (Ve) does not bind, Claim 8 implies $u^L_{\neg V} = u = u^R_{\neg V} = u_V$, which contradicts Claim 16. If (Ve) binds, an argument analogous to the case $\eta = 0$ yields a contradiction.

2. $u^L_{\neg V} = u > u^R_{\neg V}$. Claim 9 yields a contradiction.

3. $u^R_{\neg V} = u > u^L_{\neg V}$. Similarly, Claim 9 yields a contradiction.

Thus $u^R_{\neg V} < u < u^L_{\neg V}$.

Step 3: Strict separation by a margin. We now show that there exists $\Delta > 0$ such that $u^R_{\neg V} + \Delta < u < u^L_{\neg V} - \Delta$. Suppose, toward a contradiction and without loss of generality, that there exists a sequence $u_i \to \check{u}$ with $u^L_{\neg V}(u_i) \to \check{u}$. Claim 9 implies $\Pi^-(u^R_{\neg V}(u_i)) \to \Pi^-(\check{u})$ and $\Pi^+(u^R_{\neg V}(u_i)) \to \Pi^+(\check{u})$. Since $\Pi(\cdot)$ is concave, the sequence $u^R_{\neg V}(u_i)$ also converges to $\check{u}$. Consider two cases:

- For a subsequence u'_i, constraint (W) binds. Simple algebra then yields $\check{u} \leq \frac{w}{1-\delta}$, a contradiction.

- There exists a subsequence u''_i such that (W) does not bind. Using Claim 15 yields a contradiction.

This completes the proof of both (1) and (2). □

Proof of Theorem 1:

Proof.

1 It follows from Claim 1.

2 It follows from Claims 11 and 19.

3 It follows from Claims 4 and 5.

□

Proof of Theorem 2:

Proof. It follows from Claims 6, 11 and 19. □

Proof of Theorem 3:

Proof. It follows from Claims 11, 18 and 19. □

7.2 Characterization of the Principal's Preferred Equilibrium in $\tilde{\mathcal{E}}$

In this section, we show that focusing on $\mathcal{E}$ to characterize the principal's preferred equilibrium is without loss of generality.

Lemma 1. *Suppose $f(\cdot) : \mathbb{R} \to \mathbb{R}$ is a continuous and concave function. Let $x, y \in \mathbb{R}$, $z \in \mathbb{R}^+$ and $a \in (0,1)$ such that $x > y$ and $x - \frac{z}{a} \geq y + \frac{z}{1-a}$ Then, the following inequality holds:*

$$af(x) + (1-a)f(y) \leq af(x - \frac{z}{a}) + (1-a)f(y + \frac{z}{1-a})$$

Furthermore, the inequality is strict if $f(\cdot)$ is not linear on $[y, x]$.

Proof. The inequality holds if and only if

$$\frac{f(x) - f(x - \frac{z}{a})}{x - (x - \frac{z}{a})} \leq \frac{f(y + \frac{z}{1-a}) - f(y)}{(y + \frac{z}{1-a}) - y}.$$

Let $\alpha \in (0,1)$ such that $\alpha x + (1-\alpha)y = x - \frac{z}{a}$ and $\beta \in (0,1)$ such that $\beta x + (1-\beta)y = y + \frac{z}{1-a}$. Therefore the inequality in the Lemma holds if and only of

$$\frac{f(x) - f(\alpha x + (1-\alpha)y)}{(1-\alpha)(x-y)} \leq \frac{f(\beta x + (1-\beta)y) - f(y)}{\beta(x-y)}.$$

Since $x > y$, equivalently

$$\beta f(x) + (1-\alpha)f(y) \leq \beta f(\alpha x + (1-\alpha)y) + (1-\alpha)f(\beta x + (1-\beta)y).$$

By concavity of $f(\cdot)$ we know $f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y)$ and $f(\beta x + (1-\beta)y) \geq \beta f(x) + (1-\beta)f(y)$. Combining both inequalities gives us the last inequality. In addition, if $f(\cdot)$ is not linear in interval $[y, x]$, then $f(\alpha x + (1-\alpha)y) > \alpha f(x) + (1-\alpha)f(y)$ and $f(\beta x + (1-\beta)y) > \beta f(x) + (1-\beta)f(y)$ which imply the inequality in the statement is a strict inequality. □

Lemma 2. *It is without loss of generality to restrict attention to direct, truthful communications in $\tilde{\mathcal{E}}$ with message space $\hat{L}, \hat{R}$.*

Proof. Let G_M be the same game as G but with message space M. We show that for any equilibrium $e_M \in G$, there exists an outcome-equivalent equilibrium $e \in G$.

Fix a history h^{t-1} in e_M. Assume the agent is not fired and is supposed to work after h^{t-1}, i.e., $f_\xi^t = \neg F$ and $w_\xi^t = W$.[35] Assume he sends a message $m^t \in M$.

Define e' to be the same as e_M, except that after $\{h^{t-1}, \neg F, W\}$ the message space is restricted to $\{\hat{L}, \hat{R}\}$ and the agent reveals the state. We show that e' is an equilibrium. Consider two cases:

1. **Separating messages in e_M:** If the agent plays a separating strategy when sending messages m^t in e_M, then truthful direct communication in e' is clearly outcome-equivalent to e_M.

2. **Pooling messages in e_M:** If m^t is pooling in e_M, suppose instead that the agent reveals the state in e'. Continuation strategies in e' following revelation are defined to coincide with continuation strategies in e_M after sending m^t. The principal's incentive to verify or not verify in e' is the same as in e_M. After the action a is taken, the principal's payoff in e' is identical to that in e_M, since the action itself is sunk. Moreover, in the continuation game both players face the same incentives in e' and e_M. Because the agent has already taken the action, verifying and learning the state is payoff-irrelevant.

The argument above is not specific to the message space following h^{t-1}; it applies to all message spaces on the equilibrium path. □

7.2.1 Principal's Preferred Equilibrium of $\tilde{\mathcal{E}}$

Our objective in this section is to characterize the principal's preferred equilibrium within $\tilde{\mathcal{E}}$. The history of the game can be summarized by the agent's continuation utility. The upper boundary of the equilibrium payoff set is self-generating in the sense of Abreu, Pearce, and Stacchetti (1990). Consequently, the problem can be

[35] If the agent is not supposed to work ($w_\xi^t = \neg W$), messages are superfluous.

formulated as a recursive program subject to the incentive constraints of both the agent and the principal.

As noted earlier, the principal does not verify when the agent is recommended not work. Thus, on the equilibrium path, the agent's continuation utility at the end of each period can be classified into three cases:

1 u_V^{ma}, when the agent is supposed to work, he sends message $m \in \{\hat{L}, \hat{R}\}$ and he takes action $a \in \{L, R\}$ and the principal verifies.

2 $u_{\neg V}^{ma}$, when the agent is supposed to work, he sends message $m \in \{\hat{L}, \hat{R}\}$ and he takes action $a \in \{L, R\}$ and the principal does not verify.

3 $u_{\neg W}^{a}$, when the agent is recommended not work and the he takes action $a \in \{L, R\}$.[36]

Let v^{ma} be the verification probability when the agent sends message $m \in \{\hat{L}, \hat{R}\}$ and he takes action $a \in \{L, R\}$. Let α_W^{ma} be the probability that the agent is recommended to take action action $a \in \{L, R\}$ when the he sends message $m \in \{\hat{L}, \hat{R}\}$ When he is supposed to work. Let $\alpha_{\neg W}^{Ra}$ be the probability that the agent is recommended to take action action $a \in \{L, R\}$ When he is supposed to shirk. Let $\tilde{\Pi}(u)$ be the principal's highest equilibrium payoff consistent with the agent's equilibrium payoff being u. Let $\tilde{\Pi}(u) = -\infty$ if no equilibrium exists that gives the agent u. Let $\tilde{\bar{u}}$ denote the maximum promise such that $\Pi(u) > -\infty$. Note that by Lemma 2 we can restrict attention to truthful direct mechanisms.

The principal's program is:

$$\tilde{\mathcal{P}}: \quad \Pi(u) = \sup \mathbb{E}\Big[\mathbb{1}_{\neg F}\big(\pi(\theta, a) - c_P\, \mathbb{1}_V - w + \delta\, \tilde{\Pi}(u')\big)\Big],$$

subject to the agent's incentive constraints: 1) Working and reporting truthfully the state

$$\mathbb{E}\left[u(a) + v^{\theta a}\delta u_V^{\theta a} + (1 - v^{\theta a})\delta u_{\neg V}^{\theta a}\right] - c_A \geq \max_{m=\{\hat{L},\hat{R}\}} \{\mathbb{E}\left[u(a) + (1 - v^{ma})\delta u_{\neg V}^{ma} \mid m\right]\}. \tag{WT}$$

2) Conditional on working, reporting truthfully the state

$$\mathbb{E}\left[u(a) + v^{\theta a}\delta u_V^{\theta a} + (1 - v^{\theta a})\delta u_{\neg V}^{\theta a} \mid \theta\right] \geq \mathbb{E}[u(a) + (1 - v^{\theta' a})\delta u_{\neg V}^{\theta' a} \mid \theta'], \tag{T}$$

for all $\theta, \theta' \in \{L, R\}$. Specifically, let (T–L) denote the constraint corresponding to $\theta = L$ and $\theta' = R$, and let (T–R) denote the constraint corresponding to $\theta = R$

[36] It is easy to see messages m is superfluous when the agent does not work.

and $\theta' = L$. 3) The promise-keeping constraint:

$$u = \mathbb{E}\big[\mathbb{1}_{\neg F}\,(u(a) - c_A\,\mathbb{1}_W + \delta\,u')\big]. \tag{PK}$$

Principal's incentive constraints: 1) Firing

$$\mathbb{E}\Big[\pi(\theta, a) - c_P \mathbb{1}_V - w + \delta\tilde{\Pi}(u')\Big] \geq 0, \tag{Fi}$$

2) Verifying

$$\delta\tilde{\Pi}(u_V^{\theta a}) - c_P \geq 0, \tag{Ve}$$

for all $a, \theta \in \{L, R\}$. The supremum is taken over the continuation utility u', the probability of firing, the probability of working, the probability of verification and the probability of actions. The continuation utility u' can take the three aforementioned forms, $u_V^{ma}, u_{\neg V}^{ma}, u_{\neg W}^{a} \in [0, \tilde{\bar{u}}]$.[37],[38]

Lemma 3. *In the principal's preferred equilibrium, without loss of generality:*

1 *The probabilities of verification and the continuation utilities are independent of the action recommendations. Formally,* $u_V^{ma} = u_V^{ma'}$, $u_{\neg V}^{ma} = u_{\neg V}^{ma'}$, $u_{\neg W}^{a} = u_{\neg W}^{a'}$ *and* $v^{ma} = v^{ma'}$ *for* $m \in \{\hat{L}, \hat{R}\}$ *and* $a, a' \in \{L, R\}$.

2 *If the agent deviates and shirks whenever he is supposed to work, he will obtain the same utility from reporting* $\hat{L}$ *or* $\hat{R}$. *Formally,*

$$\mathbb{E}\left[u(a) + (1 - v^{\hat{R}a})\delta u_{\neg V}^{\hat{R}a} \mid \hat{R}\right] = \mathbb{E}\left[u(a) + (1 - v^{\hat{L}a})\delta u_{\neg V}^{\hat{L}a} \mid \hat{L}\right].$$

3 *The continuation utility of the agent after verification is independent of the agent message. Formally* $u_V^{m} = u_V^{m}$, *for* $m, m' \in \{\hat{L}, \hat{R}\}$.

Proof. 1 By Lemma 1, a mean-preserving contraction applied to each pair $(u_V^{ma}, u_V^{ma'})$, $(u_{\neg V}^{ma}, u_{\neg V}^{ma'})$, and $(u_{\neg W}^{a}, u_{\neg W}^{a'})$ weakly increases the principal's payoff while leaving all incentive constraints unchanged.

Furthermore, since these continuation utilities are independent of messages, a mean-preserving contraction between v^{ma} and $v^{ma'}$ does not alter the principal's payoff, the agent's utility, or any incentive constraint.

[37] Expectations are taken with respect to the random outcomes, namely the state, firing status, verification status, working status, and the agent's action.

[38] The agent has another incentive constraint. The agent must follow prd's action recommendation. For now we ignore this constraint and later we show with our assumptions this constraint never binds on the equilibrium path.

2 Using part (1) of the lemma, we adopt the convention of dropping the superscript "a" from continuation utilities and verification probabilities. We proceed the proof by contradiction. Without loss of generality, suppose

$$\mathbb{E}\Big[u(a) + (1 - v^{\hat{R}})\delta u_{\neg V}^{\hat{R}} \,\Big|\, \hat{R}\Big] < \mathbb{E}\Big[u(a) + (1 - v^{\hat{L}})\delta u_{\neg V}^{\hat{L}} \,\Big|\, \hat{L}\Big].$$

Then $v^{\hat{R}} > 0$; otherwise, this would contradict constraint (T–R). Now decrease $v^{\hat{R}}$ slightly while keeping the distribution of continuation utilities the same. Let the reduced value be $\tilde{v}^{\hat{R}}$. With probability $\tilde{v}^{\hat{R}}$, the continuation utility is $u_V^{\hat{R}}$. With probability $v^{\hat{R}} - \tilde{v}^{\hat{R}}$, the continuation utility is $u_V^{\hat{R}}$. With probability $1 - v^{\hat{R}}$, the continuation utility is $u_{\neg V}^{\hat{R}}$. Therefore, the distribution of continuation utilities remains unchanged. Since (T–L) is slack, this change does not affect that constraint. It does not affect other constraints, because the distribution of continuation utilities remains unchanged. It keeps the agent's utility constant while increasing the principal's payoff. This yields a contradiction.

3 Using parts (1) and (2) of the lemma, constraints WT and T are equivalent to

$$\mathbb{E}\Big[u(a) + (1 - v^{\hat{R}a})\delta u_{\neg V}^{\hat{R}a} \,\Big|\, \hat{R}\Big] = \mathbb{E}\Big[u(a) + (1 - v^{\hat{L}a})\delta u_{\neg V}^{\hat{L}a} \,\Big|\, \hat{L}\Big],$$

and

$$\rho v^{\hat{L}}\delta u_V^{\hat{L}} + (1 - \rho)v^{\hat{R}}\delta u_V^{\hat{R}} \;\geq\; c_A.$$

By Lemma 1, a mean-preserving contraction between u_V^R and u_V^L leaves all incentive constraints unchanged while weakly increasing the principal's payoff. Hence $u_V^R = u_V^L$.

□

Claim 20. *Suppose $\alpha_W^{\hat{R}L} \in (0,1)$. Then $\Pi'(u_{\neg V}^{\hat{R}}) = \frac{\pi}{b}$. If $\alpha_W^{\hat{R}L} = 1$, then $\Pi^-(u_{\neg V}^{\hat{R}}) = \frac{\pi}{b}$. Similarly, if $\alpha_W^{\hat{L}R} \in (0,1)$, then $\Pi'(u_{\neg V}^{\hat{L}}) = -\frac{\pi}{b}$, and if $\alpha_W^{\hat{R}L} = 1$, then $\Pi^+(u_{\neg V}^{\hat{L}}) = -\frac{\pi}{b}$.*

Proof. By Lemma 3, part (2),

$$u_{\neg V}^{\hat{L}} = \frac{b\big(1 - \alpha_W^{\hat{R}L}\big) - b\,\alpha_W^{\hat{L}R} + \delta\big(1 - v^{\hat{R}}\big)u_{\neg V}^{\hat{R}}}{\delta\big(1 - v^{\hat{L}}\big)}.$$

By the promise-keeping constraint,

$$u = (1-f)\Bigg[e\Big(w + b\big(1-\alpha_W^{\hat{R}L}\big) + \big(1-v^{\hat{R}}\big)\delta\, u_{\neg V}^{\hat{R}} + \big(\rho v^{\hat{L}} + (1-\rho)v^{\hat{R}}\big)\delta\, u_V - c_A\Big)$$
$$- (1-e)\Big(w + \alpha_{\neg W}^{\hat{R}} b + \delta\, u_{\neg W}\Big)\Bigg].$$

Now substitute u and $u_{\neg V}^{\hat{L}}$ into

$$\Pi(u) = (1-f)\Bigg[e\Big(\rho\big(1-\alpha_W^{\hat{L}R}\big)\pi + (1-\rho)\big(1-\alpha_W^{\hat{R}L}\big)\pi - w$$
$$+ \big(\rho v^R + (1-\rho)v^{\hat{L}}\big)\,\delta\,\Pi(u_V) + \rho\big(1-v^{\hat{R}}\big)\,\delta\,\Pi(u_{\neg V}^{\hat{R}}) + (1-\rho)\big(1-v^{\hat{L}}\big)\,\delta\,\Pi(u_{\neg V}^{\hat{L}})$$
$$- \big(\rho v^{\hat{R}} + (1-\rho)v^{\hat{L}}\big)c_p\Big) + (1-e)\Big(\pi p + \alpha_{\neg W}^{\hat{R}}(1-2\rho)\pi + \delta\,\Pi(u_{\neg W})\Big)\Bigg].$$

Taking the derivative with respect to $\alpha_W^{\hat{R}L}$ and evaluating the appropriate one-sided limits at the boundary yields the stated identities: at interior values $\alpha_W^{\hat{R}L} \in (0,1)$ we obtain $\Pi'(u_{\neg V}^R) = \frac{\pi}{b}$, while at $\alpha_W^{\hat{R}L} = 1$ the left derivative satisfies $\Pi^-(u_{\neg V}^R) = \frac{\pi}{b}$. The analogous argument with respect to $\alpha_W^{\hat{L}R}$ delivers $\Pi'(u_{\neg V}^L) = -\frac{\pi}{b}$ for interior points and $\Pi^+(u_{\neg V}^L) = -\frac{\pi}{b}$ at the boundary. □

Claim 21. *Suppose* $\frac{\pi - w}{w + (1-\rho)b - c_A} < \frac{\pi}{b}$. *Then* $\pi^-(u) \le \frac{\pi}{b}$. *Moreover,* $\alpha_W^{\hat{R}L} = 0$.

Proof. The slope of the upper boundary of the feasible payoff set is always less than $\frac{\pi - w}{w + (1-\rho)b - c_A}$. Hence $\pi'(0) \le \frac{\pi - w}{w + (1-\rho)b - c_A}$. Since $\Pi(\cdot)$ is concave, the left derivative $\pi^-(u)$ is nonincreasing; therefore $\pi^-(u) \le \frac{\pi}{b}$. Finally, by Claim 20, we conclude that $\alpha_W^{\hat{R}L} = 0$. □

Claim 22. $\alpha_W^{\hat{L}R} = 0$.

Proof. By an argument analogous to Claim 5, one can show that $\Pi(\cdot)$ is linear on the interval $[u_e, \overline{u}]$. Moreover, $\alpha_{\neg W}^R = 1$ if and only if the slope of the value function is less than $\frac{\pi(1-2\rho)}{1-\delta}$; otherwise, $\alpha_{\neg W}^R = 0$. In addition, for $u \in [u_e, \overline{u}]$ suppose $\Pi(u) = au + d$ for some $a, d \in \mathbb{R}$. If $a \ge \frac{\pi(1-2\rho)}{1-\delta}$, then

$$a\big(\tfrac{w}{1-\delta}\big) + d = \tfrac{\rho\pi - w}{1-\delta},$$

and if $a \le \frac{\pi(1-2\rho)}{1-\delta}$, then

$$a\big(\tfrac{w+b}{1-\delta}\big) + d = \tfrac{(1-\rho)\pi - w}{1-\delta}.$$

Moreover, the slope of the upper boundary of the feasible payoff set is weakly higher than $\frac{\rho\pi}{-\rho b - c_A} > -\frac{\pi}{b}$. Therefore $\pi^+(u) > -\frac{\pi}{b}$, and hence, by Claim 20, $\alpha_W^{\hat{L}R} = 0$. $\square$

Claims 21 and 22 imply $\alpha_W^{\hat{R}L} = 0$ and $\alpha_W^{\hat{L}R} = 0$.