Discussion Paper Series – CRC TR 224

Discussion Paper No. 795
Project B 02

Dynamic Test Design

Doruk Cetemen[1]
Francesc Dilmé[2]

November 2026

[1]LUISS and EIEF; Email: dcetemen@luiss.it
[2]University of Bonn; Email: fdilme@uni-bonn.de

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Dynamic Test Design*

Doruk Cetemen† Francesc Dilmé‡

November 3, 2026

We study dynamic test design under the threat of manipulation. At each instant, a principal who cannot commit to future actions chooses whether to hire an agent, fire him, or keep him on costly probation and design a test. The test provides information about the agent's persistent type, which is either good or bad. After observing the test choice, the bad type can exert costly effort to manipulate the evidence and mimic the good type. Deterring manipulation imposes an endogenous information budget that limits how intensively the principal can test. We show that, except at beliefs that remain unchanged in the absence of news, equilibrium tests are one-sided: they generate only passes, which trigger hiring, or only failures, which trigger firing. Equilibrium outcomes feature either correct hiring and firing decisions after costly delay or immediate decisions that risk misallocation. With commitment, an optimal policy either stops immediately or keeps the test constant throughout probation, even as the principal becomes nearly certain of the agent's type.

Keywords: Dynamic testing, Poisson monitoring, optimal stopping

JEL codes: C73, D82, D83, D86

*This work was funded by European Research Council grant 101165373 (Cetemen). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. Dilmé gratefully acknowledges support from the German Research Foundation (DFG) through CRC TR 224 (Project B02) and under Germany's Excellence Strategy – EXC 2126/1–390838866 and EXC 2047–390685813.

†LUISS and EIEF; dcetemen@luiss.it

‡University of Bonn; fdilme@uni-bonn.de

1 Introduction

In many long-term relationships, a principal chooses how to learn about an agent's persistent characteristics before taking a consequential action. The agent can, at a cost, shape the evidence generated by this learning process. An employer, for instance, chooses how intensively to evaluate a new hire and whether to search primarily for misconduct or strong performance; the worker chooses how much effort to devote to avoiding adverse signals or building a favorable record before the firm grants permanence or terminates the match. A regulator chooses the frequency and design of audits, while the firm invests in compliance or documentation that shapes the evidence it produces. A digital platform chooses how aggressively to verify contributors, who incur costs to establish reliability or provenance. In each case, the principal chooses the direction and intensity of evidence acquisition, and the agent responds in anticipation of how the resulting evidence will affect future decisions. Both the principal's incentive to acquire information and the agent's incentive to manipulate it thus depend on the continuation stakes attached to the evidence.

These two sides of the problem have largely been analyzed separately. Models of flexible learning study what a decision maker chooses to learn over time from exogenously available information sources [Che and Mierendorff, 2019, Zhong, 2022]. Models of dynamic signaling study how an agent manipulates evidence generated by a technologically fixed signal process [Dilmé, 2019, Ekmekci et al., 2022, Cetemen and Margaria, 2025].

In this paper we combine the two: the principal chooses the information process and the agent decides how much effort to devote to manipulating the evidence. Our results identify a trade-off between accuracy and delay. Equilibrium probation eliminates allocation errors on path but delays decisions, whereas immediate stopping is fast but risks misallocation. Moreover, the direction of monitoring determines who bears the cost of waiting: performance-seeking monitoring recognizes good types sooner and leaves bad types waiting, whereas misconduct-seeking monitoring does the reverse. Thus, monitoring design governs not only what is learned, but also how the delay required for learning is allocated across types.

We study a continuous-time principal–agent model in which the agent privately knows whether his persistent type is good or bad. At each instant, the principal fires the agent, hires him, or keeps him on probation; in the last case, she publicly chooses a test specifying the arrival rates of two conclusive signals, a pass and a fail. The good type always exerts high effort, whereas the bad type either shirks at no cost or imitates the good type by exerting high effort at a flow cost. A pass arrives only under high effort and leads to hiring; a fail arrives only under low effort and leads to firing. The principal incurs a flow cost while the agent is on probation, but testing itself is free, and there is no exogenous cap on test intensities. She does not value effort per se: she wants

to hire the good type and fire the bad type, while both types want to be hired. Finally, she cannot commit to future tests or stopping decisions.

The central friction is that, because test outcomes depend on effort, learning requires the bad type to shirk, yet sufficiently informative monitoring induces him to imitate the good type. A stronger pass test raises the return to imitation, whereas a stronger fail test raises the cost of shirking. Accordingly, incentive compatibility imposes an endogenous information budget: a linear constraint on the pass and fail intensities whose weights depend on the bad type's continuation value. As that value evolves, so does the principal's preferred direction of monitoring.

A key observation is that fail-only tests are attractive to the principal when the bad type's continuation value is low: failure then entails little incremental loss for him, so an intense fail test does not induce imitation. Pass-only tests are instead attractive when his continuation value is high, because passing then yields him little incremental gain. This asymmetry yields a sharp implication of strategic manipulation: a fail-only region, in which no news raises the posterior, cannot end in hiring. It must instead end at a stationary belief, at which pass and fail tests have equal intensity and the posterior remains constant until news arrives. Symmetrically, a pass-only region, in which no news lowers the posterior, cannot end in firing. Thus, conditional on entering probation, monitoring separates types ex post: only the good type is hired and only the bad type is fired, albeit with delay. In either region, the active test becomes less intense as time passes without news: the bad type's continuation value rises under fail-only monitoring, increasing his loss from failing, and falls under pass-only monitoring, increasing his gain from passing.

We show that, within a class of Markov assessments in which the posterior belief is the state and the bad type shirks on path, the equilibrium is essentially unique, and we characterize it. In every regime, the principal fires at sufficiently low beliefs and hires at sufficiently high ones. Fail-only monitoring arises only when manipulation is sufficiently costly: otherwise, deterring imitation requires a fail test too weak to justify probation. Pass-only monitoring arises only when probation is sufficiently cheap, because the principal must be willing to wait for good news. When imitation is costly and probation is cheap, monitoring is fail-only at low beliefs and pass-only at high beliefs, with the two regions joined at an interior stationary belief. For intermediate parameters, probation uses a single kind of test: fail-only when probation is relatively costly, and pass-only when imitation is relatively cheap. When imitation is cheap relative to the cost of probation, incentive-compatible learning is not worthwhile, and the principal stops immediately. Although the model features two long-lived players interacting over time, the thresholds that separate these regimes, as well as the stationary beliefs, are in closed form.

We use two benchmarks to isolate the forces shaping the equilibrium. The first removes manipulation and instead imposes a fixed total testing capacity, chosen so that

it would not bind in our equilibrium; this yields a specialization of Che and Mierendorff [2019]. Its solution features outward learning, in which no news moves the posterior toward a stopping boundary and the agent may be misclassified, and, when attention is sufficiently cheap, inward learning toward a stationary belief, which classifies him correctly. The comparison shows that the bad type's incentive constraint eliminates precisely outward learning: a pass-only region cannot end in firing, and a fail-only region cannot end in hiring. Incentive compatibility therefore replaces outward learning with immediate stopping or inward learning toward a stationary belief.

The second benchmark allows the principal to commit ex ante to the entire testing and stopping policy, subject to a cap on test intensities. There is then an optimal policy that either stops immediately or screens with constant test intensities, so that no-news beliefs move monotonically toward certainty rather than toward an interior stationary belief. Commitment matters because it allows the principal to persist with a policy she would later prefer to abandon: under maximal fail monitoring, she continues to withhold hiring even as no news makes the agent arbitrarily likely to be good; under perpetual pass-only monitoring, she continues to withhold firing even as no news makes him arbitrarily likely to be bad. Thus, the first benchmark isolates the effect of manipulation, whereas the second isolates the role of commitment to future testing and stopping.

Two extensions show that the baseline equilibria survive relaxing two modeling assumptions. First, they persist when the good type chooses effort strategically at a lower cost than the bad type: the same tests that deter imitation also strictly induce high effort from the good type. Second, they survive an enlarged menu of monotone imperfect tests, which allow failures under high effort and passes under low effort. These errors add effort-independent hiring or firing without relaxing the incentive constraint on informative testing, so conclusive testing remains optimal.

1.1 Related literature

Empirical work documents both behavioral responses to monitoring and the importance of continuation stakes. Monitoring changes opportunistic behavior in call centers, public works, and teacher attendance [Nagin et al., 2002, Olken, 2007, Duflo et al., 2012]; among newly hired manufacturing workers, enhanced monitoring raises performance but can also increase turnover [Cai et al., 2025]. Behavior also responds to the stakes attached to evaluation. Absence rises after Italian bank employees acquire dismissal protection [Ichino and Riphahn, 2005], whereas greater discretion to dismiss probationary teachers reduces absence [Jacob, 2013]. A credible dismissal threat also increases the attrition of low-performing teachers and improves the subsequent performance of those who remain [Dee and Wyckoff, 2015].

Agents also respond to the composition of the evidence used in screening. Accountability incentives induce manipulation of standardized-test outcomes [Jacob and Levitt,

2003], while increasing the weight on a performance measure in tenure decisions redirects effort toward that measure and changes who is retained [Dinerstein and Opper, 2022]. On the principal's side, information acquired after hiring affects separation decisions: employers update in response to objective performance estimates [Rockoff et al., 2012], and realized productivity gradually displaces readily observed prior information [Farber and Gibbons, 1996, Altonji and Pierret, 2001, Lange, 2007]. Decision makers also redirect attention in response to prior signals and the selection problem they face [Bartoš et al., 2016]. Yet these findings do not characterize how a principal dynamically chooses the direction and intensity of monitoring when the agent can reshape the resulting evidence.

Our paper connects flexible information acquisition with dynamic signaling through endogenous monitoring. Work on flexible learning studies how a decision maker chooses what to learn before acting. Che and Mierendorff [2019] analyze the dynamic allocation of attention across sources biased toward different actions and obtain confirmatory learning at extreme beliefs and contrarian learning at intermediate beliefs. Zhong [2022] allows a broad class of information processes and shows that optimal learning can be implemented with Poisson signals. Building on this flexible approach, Georgiadis-Harris [forthcoming] studies information acquisition when opportunities to act arrive exogenously, while Ilinov et al. [2025] apply flexible experiment choice to sequential search. These papers explain why optimal learning may be directional and, in several of these models, why Poisson breakthroughs arise, but no other player can influence the evidence. Here the learner faces a strategic agent whose hidden action changes the distribution of evidence, so experiment choice must also satisfy incentive compatibility.

Work on monitoring and test design introduces this incentive problem. With commitment, Georgiadis and Szentes [2020] and Li and Yang [2020] endogenize monitoring in principal–agent problems, while Perez-Richet and Skreta [2022] design approve–reject tests whose inputs an approval-seeking agent can falsify at a cost. In dynamic contracting, Piskorski and Westerfield [2016] study how costly monitoring interacts with performance pay and termination, and Varas et al. [2020] characterize the choice between random inspections and periodic reviews when monitoring has both informational and incentive roles. Bhaskar [2025] combines adverse selection and moral hazard in a self-enforcing relationship, where monitoring screens low types and shapes incentives after screening. Dai et al. [2024] and Wong [2023] instead allow the principal to choose flexible monitoring technologies that seek evidence tied to rewards or punishments.

As in Dai et al. [2024] and Wong [2023], our principal chooses the direction and intensity of monitoring. The key difference is that our principal cannot commit: she chooses each test optimally given the current belief and future play, rather than committing to a future monitoring path. The posterior belief is the state, and the bad type's shirking constraint creates a belief-dependent information budget. Within the simple Markov class studied below, this interaction yields one-sided monitoring away from stationary beliefs, and every one-sided region leads, in the absence of news, to a stationary belief.

A parallel literature studies dynamic signaling and reputation under a fixed monitoring technology. Dilmé [2019], Ekmekci et al. [2022], and Cetemen and Margaria [2025] study continuous-time signaling or stopping games in which hidden actions affect a technologically fixed noisy signal. Board and Meyer-ter-Vehn [2013] and Faingold and Sannikov [2011] develop related continuous-time reputation models, while Kuvalekar and Lipnowski [2020] study job insecurity with observable behavior and Dilmé [2025] studies repeated trade when future sellers observe noisy histories of a privately informed buyer's past transactions. We instead endogenize which kind of conclusive evidence the principal seeks, and at what intensity, subject to the bad type's incentive constraint. Our contribution lies in the interaction between belief-based stopping and this endogenous information budget.

Roadmap. Section 2 introduces the model, admissible Markov assessments, posterior dynamics, and continuation values. Section 3 derives the bad type's incentive constraint, the principal's payoff guarantees, and her pointwise experiment-choice problem. Section 4 defines simple Markov equilibrium and characterizes the phase diagram. Sections 5 and 6 analyze the two benchmarks. Section 7 discusses strategic effort by the good type and monotone imperfect tests; Section 8 concludes. Appendix A contains the baseline proofs and supplementary results. The Online Appendix provides additional proofs and extensions.

2 Model

Time is continuous, $t \in \mathbb{R}_+$, and the horizon is infinite. There is a principal (she) and an agent (he). The agent's type is fixed and privately known to the agent:

$$\theta \in \{B, G\}, \qquad B < 0 < G.$$

The principal's prior belief that the agent is type G is $\beta_0 \in (0, 1)$.

The agent starts on probation, during which the principal can hire or fire him at any moment. If she keeps the agent on probation, she chooses a test

$$(\lambda_G, \lambda_B) \in \mathbb{R}^2_+.$$

The component λ_G is the arrival rate of a pass under high effort, and λ_B is the arrival rate of a fail under low effort.

At each instant t during probation, play proceeds as follows.

1. The principal either stops immediately, choosing $a_t \in \{\mathsf{f}, \mathsf{h}\}$ (fire or hire), or keeps the agent on probation, choosing $a_t = \mathsf{k}$ and a test $(\lambda_G, \lambda_B) \in \mathbb{R}^2_+$.
2. The agent observes the test and privately chooses effort $e_t \in [0, 1]$, where $e_t = 1$ is

high effort and $e_t = 0$ is low effort. Effort e generates a pass at rate $e\lambda_G$ and a fail at rate $(1-e)\lambda_B$.

3. A pass ends the game with hiring, and a fail ends it with firing (see Remark 4.1 for the role of this assumption). Otherwise, probation continues.

The principal's payoff from firing is normalized to 0. Her payoff from hiring is the agent's type, θ. Keeping the agent on probation costs the principal $k > 0$ per unit of time. The agent receives terminal payoff $F < 0$ if fired and $H > 0$ if hired. While on probation, the agent receives flow payoff normalized to zero before effort costs. For the bad type, effort e costs ce per unit of time, where $c > 0$. The principal and the agent discount future payoffs at the common rate $r > 0$.

Assumption 1. While on probation, the good type always exerts high effort: $e_G \equiv 1$.

Hence, the good type passes at rate λ_G and never fails, whereas the bad type with effort e passes at rate $e\lambda_G$ and fails at rate $(1-e)\lambda_B$.[1]

Assumption 1 keeps the analysis simple by making the good type behavioral: no good-type effort incentive constraint is imposed, and no effort cost enters his continuation value. Section 7.1 shows, however, that every equilibrium characterized in Section 4 remains valid, with the same policies and thresholds, when the good type is strategic and incurs effort cost $c_H e$ per unit of time, with $0 < c_H < c$. The reason is that the bad type's incentive constraint is the binding restriction on testing: at every probation belief, the equilibrium test strictly satisfies the good type's incentive constraint for high effort.

2.1 Simple Markov assessments

We focus on simple Markov assessments that use the public posterior probability β that the agent is type G as the state variable and prescribe on-path shirking by the bad type, $e_B \equiv 0$. This restriction is natural: a test that induces imitation gives both types the same signal law and is thus uninformative, and Section 3 shows that it is never a profitable pointwise deviation for the principal.

Definition 2.1. A *simple Markov assessment* is a triple $\alpha = (a, \lambda_G, \lambda_B)$ with the following components.

1. $a : [0,1] \to \{\mathsf{f}, \mathsf{k}, \mathsf{h}\}$ is a Borel measurable action rule. The value $a(\beta)$ determines whether the principal fires the agent, keeps him on probation, or hires him at belief β.
2. $\lambda_G, \lambda_B : [0,1] \to \mathbb{R}_+$ are Borel measurable test rules. At belief β, the value $\lambda_G(\beta)$ is the pass arrival rate under effort 1, while $\lambda_B(\beta)$ is the fail arrival rate under effort 0.

[1] Equivalently, one can interpret the tests as being informative about type, but the bad type can manipulate them through costly effort, with $e = 1$ replicating the good type's signal law.

The prescribed test values $\lambda_G(\beta)$ and $\lambda_B(\beta)$ are payoff-irrelevant at beliefs with $a(\beta) \neq \mathsf{k}$; they are defined there only so that the test rules are functions on $[0,1]$. Since the bad type shirks on path, at a probation belief β the good type passes at rate $\lambda_G(\beta)$ and never fails, while the bad type fails at rate $\lambda_B(\beta)$ and never passes. The equilibrium conditions below verify that shirking is optimal and, at nonstationary probation beliefs (defined below), require the principal's test to be optimal among tests that induce shirking.

We now impose admissibility conditions on simple Markov assessments to ensure a unique no-news belief path and unique continuation values for the principal and the bad type. Similar requirements appear in Klein and Rady [2011], Board and Meyer-ter-Vehn [2013], and Che and Mierendorff [2019].

Definition 2.2. Let $\alpha = (a, \lambda_G, \lambda_B)$ be a simple Markov assessment. Let

$$\mathcal{K} := \{\beta \in [0,1] : a(\beta) = \mathsf{k}\}, \qquad \mathcal{S} := \{\beta \in \mathcal{K} : \lambda_G(\beta) = \lambda_B(\beta)\}$$

be the sets of probation beliefs and of stationary probation beliefs, respectively. The assessment is admissible if the following conditions hold.

1. $\mathcal{K} \setminus \mathcal{S}$ is a finite union of open intervals of $\mathbb{R}$.
2. $\mathcal{S}$ is finite.
3. On each connected component I of $\mathcal{K} \setminus \mathcal{S}$, which we call a nonstationary probation component, the function $\lambda_B - \lambda_G : I \to \mathbb{R}$ is locally Lipschitz (and hence, being nonzero on I, has constant sign), and $\lambda_G + \lambda_B : I \to \mathbb{R}_+$ is bounded.

These regularity requirements define the assessment class studied below. In particular, they exclude unbounded on-path test rates, infinitely many nonstationary probation components or stationary beliefs, and a reversal of the no-news drift within a nonstationary component.

Belief dynamics. Fix an admissible simple Markov assessment. While the agent is on probation and no news arrives, Bayes' rule implies that the posterior evolves according to $\dot{\beta}_t = b(\beta_t)$, where

$$b(\beta) := \beta(1-\beta)(\lambda_B(\beta) - \lambda_G(\beta)).$$

Thus no news raises the posterior if $\lambda_B(\beta) > \lambda_G(\beta)$ and lowers it if $\lambda_G(\beta) > \lambda_B(\beta)$; at each stationary belief $\beta \in \mathcal{S}$, the posterior remains constant until news arrives.[2]

Continuation values. An admissible simple Markov assessment generates a unique pair of continuation values $\Pi : [0,1] \to \mathbb{R}$ and $V : [0,1] \to \mathbb{R}$ for the principal and the

[2]Lemma A.1 in Appendix A.1 formally establishes that belief dynamics are uniquely determined for admissible assessments, up to the first time the belief exits the nonstationary probation region.

bad type, respectively.[3] At stopping beliefs,

$$\begin{aligned} a(\beta) = \mathsf{f} \quad &\Rightarrow \quad (\Pi(\beta), V(\beta)) = (0, F), \\ a(\beta) = \mathsf{h} \quad &\Rightarrow \quad (\Pi(\beta), V(\beta)) = (h(\beta), H), \end{aligned}$$

where $h(\beta)$ is the principal's expected payoff from hiring at belief β:

$$h(\beta) := \beta G + (1-\beta)B. \tag{2.1}$$

At each stationary belief $\beta \in \mathcal{S}$, news arrives at rate $\lambda := \lambda_G(\beta) = \lambda_B(\beta)$ for either type, so

$$\Pi(\beta) = \frac{\beta\lambda G - k}{r+\lambda}, \qquad V(\beta) = \frac{\lambda F}{r+\lambda}.$$

On each nonstationary probation component, the continuation values are locally absolutely continuous and satisfy the Bellman equations almost everywhere:

$$r\Pi(\beta) = -k + b(\beta)\Pi'(\beta) + \beta\lambda_G(\beta)(G - \Pi(\beta)) - (1-\beta)\lambda_B(\beta)\Pi(\beta), \tag{2.2}$$

$$rV(\beta) = b(\beta)V'(\beta) + \lambda_B(\beta)(F - V(\beta)). \tag{2.3}$$

Both values also satisfy value matching at the belief where the no-news path exits the component, whenever it does so in finite time.

The next section takes these continuation values as given and asks which tests make shirking optimal for the bad type and which of these are optimal for the principal.

3 Incentive and optimality conditions

This section derives the local incentive and optimality conditions used in the equilibrium definition in Section 4.

3.1 The bad type's incentive constraint

Fix an admissible assessment $(a, \lambda_G, \lambda_B)$ and let V denote the bad type's continuation value. At belief $\beta \in (0,1)$, suppose the principal offers the test (ℓ_G, ℓ_B), which may be different from the on-path test $(\lambda_G(\beta), \lambda_B(\beta))$. As in the assessment, the principal believes that the bad type shirks under this test; we refer to this as *assessed shirking*. Then, heuristically, the no-news posterior after a small interval of length $dt > 0$ is

$$\widehat{\beta}_{dt} := \frac{\beta e^{-\ell_G dt}}{\beta e^{-\ell_G dt} + (1-\beta)e^{-\ell_B dt}}, \tag{3.1}$$

[3] See Lemma A.1 in Appendix A.1.

which is lower than β if $\ell_G > \ell_B$, higher than β if $\ell_G < \ell_B$, and equal to β if $\ell_G = \ell_B$. If the bad type shirks over this interval, his payoff is

$$\ell_B dt\, F + (1 - \ell_B dt)(1 - r dt) V(\widehat{\beta}_{dt}) + o(dt).$$

If instead he exerts high effort over the interval, his payoff is

$$-c dt + \ell_G dt\, H + (1 - \ell_G dt)(1 - r dt) V(\widehat{\beta}_{dt}) + o(dt).$$

Shirking is then locally optimal only if the former expression is at least the latter.[4] Dividing the difference between the high-effort and shirking payoffs by dt and taking its upper limit as $dt \downarrow 0$ gives the necessary first-order condition[5]

$$\ell_G\Big[H - V(\beta^{\operatorname{sign}(\ell_B - \ell_G)})\Big] - \ell_B\Big[F - V(\beta^{\operatorname{sign}(\ell_B - \ell_G)})\Big] \leq c. \tag{3.2}$$

Where V is continuous, (3.2) reads $\ell_G[H - V(\beta)] + \ell_B[V(\beta) - F] \leq c$: the flow benefit of imitation, which yields the gain $H - V(\beta)$ from passing at rate ℓ_G and avoids the loss $V(\beta) - F$ from failing at rate ℓ_B, cannot exceed its flow cost. This linear constraint is the endogenous information budget; its weights move with the bad type's continuation value.

For a given continuation value function V, define the set of (first-order) incentive-compatible tests at belief β by

$$\mathcal{A}(\beta; V) := \{(\ell_G, \ell_B) \in \mathbb{R}^2_+ : (3.2) \text{ holds}\}. \tag{3.3}$$

At points where V is continuous and $F < V(\beta) < H$, this set is the triangle with vertices

$$(0,0), \qquad \left(\frac{c}{H - V(\beta)}, 0\right), \qquad \left(0, \frac{c}{V(\beta) - F}\right).$$

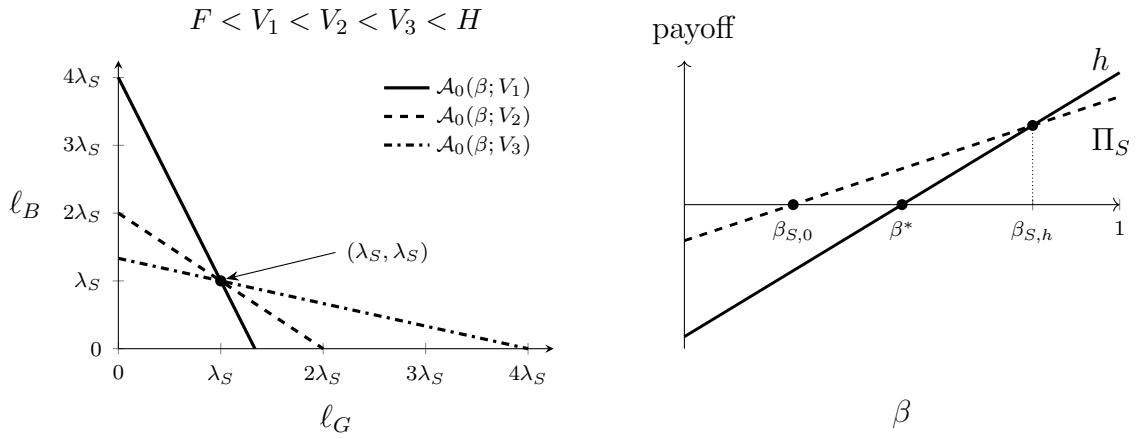
The tests for which this first-order constraint binds are

$$\mathcal{A}_0(\beta; V) := \{(\ell_G, \ell_B) \in \mathbb{R}^2_+ : (3.2) \text{ holds with equality}\}. \tag{3.4}$$

The binding incentive sets $\mathcal{A}_0(\beta; V)$ are depicted in panel (a) of Figure 1 for a fixed belief β and several continuation values V. At continuity points with $F < V < H$, each set is the downward-sloping line segment joining $(c/(H-V), 0)$ and $(0, c/(V-F))$, while the corresponding incentive-compatible set $\mathcal{A}(\beta; V)$ is the triangle below it. All these

[4] For the induced value V, Lemma A.1 supplies the required Bellman equation and value matching. If (3.2) holds for the prescribed test at every probation belief, Lemma A.2 then shows that shirking is optimal against all progressively measurable effort deviations, not only one-shot deviations.

[5] $V(\beta^-)$ and $V(\beta^+)$ denote the left and right limits of V at β, respectively. When the corresponding limit does not exist, $V(\beta^-)$ denotes the left-hand $\liminf$ and $V(\beta^+)$ denotes the right-hand $\limsup$. We set $V(\beta^0) = V(\beta)$.



(a) Binding incentive sets.

(b) Payoff crossings.

Figure 1: Payoff benchmarks and binding incentive sets. Panel (a) shows $\mathcal{A}_0(\beta; V_i)$ for fixed β and $F < V_1 < V_2 < V_3 < H$. Panel (b) shows the pairwise crossings of 0, h, and Π_S.

segments pass through the stationary test (λ_S, λ_S) defined in (3.6): with equal rates, V cancels from (3.2). As V increases, the horizontal intercept rises and the vertical intercept falls, so the boundary becomes flatter. Intuitively, a higher continuation value reduces the bad type's gain from passing, $H - V$, but increases his gain from avoiding failure, $V - F$.

Whenever the relevant one-sided limit exists, (3.2) characterizes the bad type's first-order effort incentives under assessed shirking. Since his first-order payoff is linear in effort, if $(\ell_G, \ell_B) \notin \mathcal{A}(\beta; V)$, high effort uniquely maximizes his first-order payoff. If $(\ell_G, \ell_B) \in \mathcal{A}(\beta; V) \setminus \mathcal{A}_0(\beta; V)$, zero effort uniquely maximizes it, whereas on $\mathcal{A}_0(\beta; V)$ he is indifferent to first order.

3.2 Payoff guarantees for the principal

Fix an admissible assessment, and let Π denote the principal's continuation value. Firing yields the principal 0 at every belief, and immediate hiring yields $h(\beta)$, defined in (2.1). The static cutoff is

$$h(\beta) \geq 0 \qquad \Longleftrightarrow \qquad \beta \geq \beta^* := \frac{-B}{G - B}. \tag{3.5}$$

The principal can also keep a stationary test (ℓ, ℓ) in place until a pass or fail occurs. Stationarity means that the no-news posterior is constant; pass and fail outcomes still arrive and remain informative. For every V,

$$(\ell, \ell) \in \mathcal{A}(\beta; V) \qquad \Longleftrightarrow \qquad \ell \leq \lambda_S := \frac{c}{H - F}. \tag{3.6}$$

Because, whatever his continuation value, the bad type's unique best response to a stationary test (ℓ, ℓ) with $\ell < \lambda_S$ is shirking, the principal can ensure a payoff arbitrarily

close to

$$\Pi_S(\beta) := \frac{\lambda_S \beta G - k}{r + \lambda_S}. \tag{3.7}$$

Under the on-path selection of shirking at indifference maintained above, the endpoint $\ell = \lambda_S$ attains $\Pi_S(\beta)$. We write the bad type's value under this stationary test as

$$V_S := \frac{\lambda_S F}{r + \lambda_S}. \tag{3.8}$$

The crossings of Π_S with 0 and h are

$$\beta_{S,0} := \frac{k}{\lambda_S G}, \qquad \beta_{S,h} := \frac{-k - (r + \lambda_S)B}{r(G - B) - \lambda_S B}, \tag{3.9}$$

which need not belong to $[0, 1]$. Panel (b) of Figure 1 depicts the three payoff benchmarks. Since the principal can fire, hire, or switch to the stationary test at any belief, these benchmarks are payoffs she can guarantee herself.

3.3 Principal's optimality

A test that induces high effort by the bad type gives both types the same signal law and leaves the posterior unchanged. At a continuation value $\Pi(\beta) \geq \max\{0, h(\beta)\}$, its net first-order payoff gain is

$$-k - r\Pi(\beta) + \ell_G[h(\beta) - \Pi(\beta)] < 0,$$

since $-k - r\Pi(\beta) < 0$ and $h(\beta) \leq \Pi(\beta)$. Hence such a test is never a profitable pointwise deviation, and the equilibrium class below imposes pointwise test optimality only among tests that induce shirking.

Fix an admissible assessment and a belief $\beta \in \mathcal{K} \setminus \mathcal{S}$ at which Π is differentiable. Define

$$\Psi_P(\beta, \ell_G, \ell_B; \Pi) := \beta(1 - \beta)(\ell_B - \ell_G)\Pi'(\beta) + \beta\ell_G(G - \Pi(\beta)) - (1 - \beta)\ell_B\Pi(\beta),$$

so that, under assessed shirking, continuing probation with the test (ℓ_G, ℓ_B) yields the principal the first-order payoff gain $-k - r\Pi(\beta) + \Psi_P(\beta, \ell_G, \ell_B; \Pi)$. Pointwise optimality means that the test maximizes this gain, and hence Ψ_P, among incentive-compatible tests.

Lemma 3.1. *Consider an admissible assessment with continuation values (Π, V), and let $\beta \in \mathcal{K} \setminus \mathcal{S}$. Suppose that Π is differentiable at β, V is continuous at β, and $F < V(\beta) < H$. A test is pointwise optimal among incentive-compatible tests if and only if it maximizes Ψ_P over $\mathcal{A}_0(\beta; V)$.*

Formally, at a nonstationary probation belief, pointwise optimality requires

$$(\lambda_G(\beta), \lambda_B(\beta)) \in \underset{(\ell_G,\ell_B)\in\mathcal{A}(\beta;V)}{\arg\max} \Psi_P(\beta, \ell_G, \ell_B; \Pi).$$

By Lemma 3.1, the maximization can be restricted to $\mathcal{A}_0(\beta; V)$ whenever V is continuous at β and $F < V(\beta) < H$.

Requiring the principal's continuation value to be continuous has a direct economic interpretation. Strictly incentive-compatible one-sided tests let the principal move the posterior locally in either direction at a cost that vanishes with the distance traveled, so any jump in her value could be exploited at negligible cost.

3.4 Endpoint restrictions for learning phases

The principal's payoff guarantees restrict how a nonstationary learning phase can end in finite no-news time. An open interval $I \subseteq \mathcal{K} \setminus \mathcal{S}$ is a *pass-learning phase* if $\lambda_G > \lambda_B$ on I, and a *fail-learning phase* if $\lambda_B > \lambda_G$ on I. By admissibility, every nonstationary probation component is one or the other. No news moves the posterior toward the lower endpoint of a pass-learning phase and toward the upper endpoint of a fail-learning phase.

Lemma 3.2. *Consider an admissible assessment with induced continuation values* (Π, V) *such that*

$$(\lambda_G(\beta), \lambda_B(\beta)) \in \mathcal{A}(\beta; V) \qquad \text{for all } \beta \in \mathcal{K} \setminus \mathcal{S},$$

and

$$\Pi(\beta) \geq \max\{0, h(\beta), \Pi_S(\beta)\} \qquad \text{for all } \beta \in [0, 1].$$

Let $I = (\beta', \beta'')$ be a connected component of $\mathcal{K} \setminus \mathcal{S}$. If I is pass-learning, then its lower endpoint β', if reached in finite no-news time, cannot be a firing endpoint: $a(\beta') \neq \mathsf{f}$. If I is fail-learning, then its upper endpoint β'', if reached in finite no-news time, cannot be a hiring endpoint: $a(\beta'') \neq \mathsf{h}$.

The intuition is simple. Near a firing endpoint, the bad type's continuation value is close to F, so his gain from avoiding a failure, $V(\beta) - F$, is small. The binding incentive frontier is then steep: it permits a large fail intensity ℓ_B but caps the pass intensity near λ_S. A pass-learning phase therefore learns no faster than the stationary test while drifting toward the zero payoff from firing, which violates the principal's guarantees $\Pi \geq \max\{0, \Pi_S\}$. Symmetrically, near a hiring endpoint, the bad type's continuation value is close to H, so the frontier is flat and caps the fail intensity near λ_S; a fail-learning phase ending in hiring would then violate $\Pi \geq \max\{h, \Pi_S\}$.

4 Equilibrium

Definition 4.1. An admissible simple Markov assessment is a *simple Markov equilibrium* if the following conditions hold.

1. Every stationary probation belief uses the binding stationary test:

$$\lambda_G(\beta) = \lambda_B(\beta) = \lambda_S \qquad \text{for all } \beta \in \mathcal{S}.$$

2. The principal's value is continuous and satisfies the payoff guarantees:

$$\Pi \in C([0,1]), \qquad \Pi(\beta) \geq \max\{0, h(\beta), \Pi_S(\beta)\} \qquad \text{for all } \beta \in [0,1].$$

3. At every nonstationary probation belief, Π is differentiable, the prescribed test is pointwise optimal,

$$(\lambda_G(\beta), \lambda_B(\beta)) \in \underset{(\ell_G,\ell_B)\in\mathcal{A}(\beta;V)}{\arg\max} \Psi_P(\beta, \ell_G, \ell_B; \Pi), \qquad \beta \in \mathcal{K} \setminus \mathcal{S},$$

   and the principal's Bellman equation (2.2) holds pointwise there.

Throughout, equilibrium means an assessment satisfying Definition 4.1. The definition exploits the recursive structure of the problem to express equilibrium directly in terms of tests and continuation values. Prescribed shirking, including zero effort at binding tests, is a best response rather than a behavioral restriction: given the principal's policy, the incentive constraint and the continuation-value equations make it optimal against arbitrary history-dependent effort deviations (see Lemma A.2).

Off-path tests are evaluated with the same continuation values that support prescribed play. For any test, condition (3.2), evaluated on the side toward which the test moves the posterior, determines whether shirking is locally optimal. At nonstationary probation beliefs, the definition requires the prescribed test to maximize the principal's first-order payoff gain among all tests satisfying this condition. Tests inducing imitation are uninformative and never profitable, and the payoff guarantees rule out gains from hiring, firing, or switching to stationary testing.

The remaining conditions extend this discipline to beliefs where a smooth Bellman equation alone would not suffice. Inside stopping regions, the incentive constraint caps the allocation-relevant signal intensity at λ_S, making the stationary test the relevant local benchmark. As we argued in Section 3.3, the guarantee $\Pi \geq \Pi_S$, the one-sided Bellman inequalities, and continuity of Π, rule out profitable local deviations in these regions and at stationary beliefs and region boundaries, even when the agent's value jumps. The simple formulation thus retains the state-by-state incentive discipline of Markov equilibrium: off-path beliefs and alternative tests are governed by the same continuation values that support prescribed play.

4.1 Equilibrium characterization

We now present our main characterization result. A *fail-only region* is a nonstationary probation component on which $\lambda_B > \lambda_G = 0$, and a *pass-only region* is one on which $\lambda_G > \lambda_B = 0$. The thresholds in the statement are obtained in closed form in Section 4.2.

Theorem 4.1. *For every parameter vector satisfying the maintained assumptions, there is an equilibrium, unique up to the qualifications stated at the end. There are three thresholds: k_F (independent of c), $c_N(k)$, and $c_P(k)$. They satisfy $c_N(k) < c_P(k)$ for every $k < k_F$ and partition the (k, c)-parameter space as follows:*

(i) ***(No probation)*** *If $c \leq c_N(k)$, then $\mathcal{K} \setminus \mathcal{S}$ is empty.*

(ii) ***(Fail-only)*** *If $k \geq k_F$ and $c > c_N(k)$, then $\mathcal{K} \setminus \mathcal{S}$ has exactly one connected component; this component is a fail-only region, and its right endpoint is a stationary belief.*

(iii) ***(Pass-only)*** *If $k < k_F$ and $c_N(k) < c \leq c_P(k)$, then $\mathcal{K} \setminus \mathcal{S}$ has exactly one connected component; this component is a pass-only region, and its left endpoint is a stationary belief.*

(iv) ***(Fail–pass)*** *If $k < k_F$ and $c > c_P(k)$, then $\mathcal{K} \setminus \mathcal{S}$ has exactly two connected components, one a fail-only region and the other a pass-only region; the right endpoint of the fail-only region and the left endpoint of the pass-only region coincide at a stationary belief.*

In each case, there is no probation at extreme posteriors: there exists $\varepsilon > 0$ such that $\mathcal{K} \subset [\varepsilon, 1 - \varepsilon]$. The equilibrium is unique apart from tests outside the probation set and the principal's stopping choice at β^*, where she is indifferent between firing and hiring, when $c \leq c_N(k)$. At $c = c_N(k)$, stationary probation at the isolated belief β^* is also permitted and leaves the principal's value unchanged.

Theorem 4.1 implies that equilibrium tests in $\mathcal{K} \setminus \mathcal{S}$ are one-sided: either $\lambda_G = 0$ or $\lambda_B = 0$. The intuition is as follows. The bad type's incentive constraint is a linear information budget, and the principal's objective Ψ_P is linear in the pass and fail intensities. Consequently, the principal allocates the entire budget to the test with the larger return per unit of budget. A test with both intensities positive can be optimal only when the two tests tie. Since every binding budget line passes through (λ_S, λ_S), such a tie is equivalent, by the Bellman equation, to the principal obtaining the stationary-testing value Π_S. This is incompatible with the continuation-value dynamics in a nonstationary region: moving away from a stationary belief raises the bad type's value in a pass-only region and lowers it in a fail-only region, whereas the tie condition can hold only on the opposite side of the stationary belief. Hence the tie cannot occur in $\mathcal{K} \setminus \mathcal{S}$, and the optimal test is necessarily pass-only or fail-only.

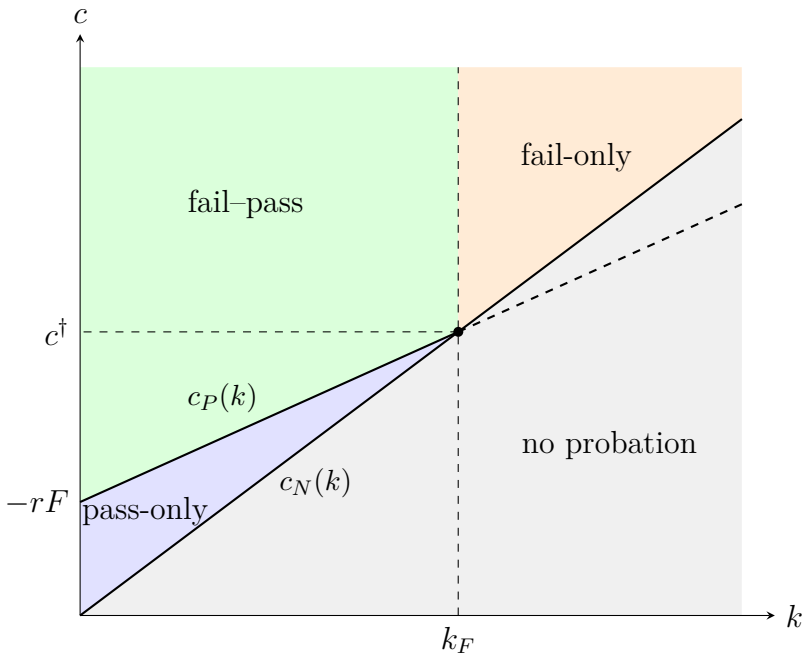


Figure 2: Phase diagram in the (k, c)-space. The line $c_N(k)$ separates the regions with and without nondegenerate probation. For $k < k_F$, the region between $c_N(k)$ and $c_P(k)$ is pass-only, while the region above $c_P(k)$ is fail–pass. For $k > k_F$, the region above $c_N(k)$ is fail-only.

4.2 Regions

This subsection derives the thresholds $c_N(k)$, $c_P(k)$, and k_F that generate the phase diagram in Theorem 4.1, depicted in Figure 2, and constructs the equilibrium in each resulting region of the (k, c)-parameter space. Figure 3 provides a roadmap for these constructions by displaying the stopping regions, probation regions, stationary beliefs, and direction of no-news learning across the four regimes. We consider in turn the no-probation, fail–pass, fail-only, and pass-only cases.

No probation. In case (i) of Theorem 4.1, there is no nondegenerate probation. The threshold $c_N(k)$ has a direct interpretation: it is the effort cost at which the stationary test breaks even at the static cutoff. Indeed, evaluating Π_S at the static cutoff gives

$$\Pi_S(\beta^*) = \frac{G(-B)}{(G-B)[r(H-F)+c]}[c - c_N(k)],$$

where

$$c_N(k) := \frac{k(H-F)(G-B)}{G(-B)}. \tag{4.1}$$

Thus $\Pi_S(\beta^*) > 0$ if and only if $c > c_N(k)$. In that case the stationary test gives the principal a strictly positive payoff at the static cutoff, so immediate stopping at the static cutoff violates the payoff guarantee $\Pi \geq \Pi_S$.

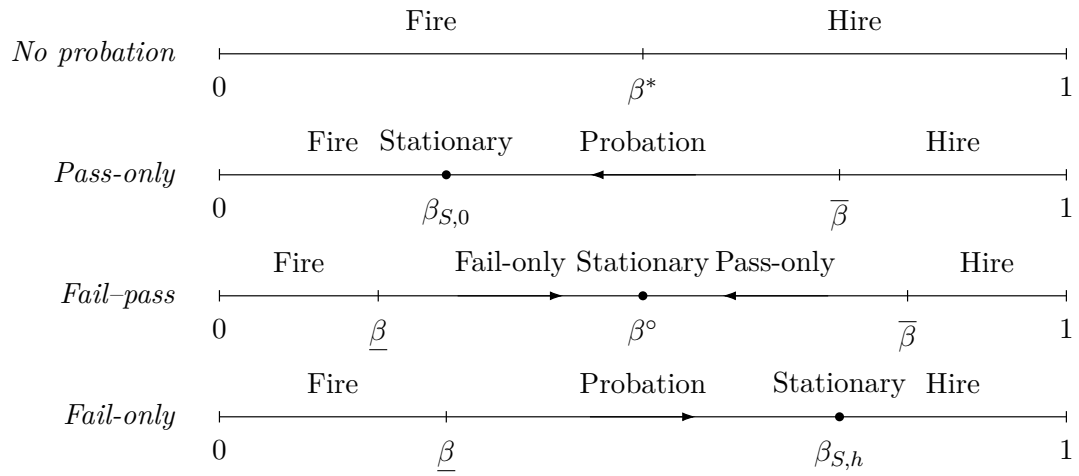


Figure 3: No-news belief dynamics. Arrows indicate the direction of belief updating in the absence of a signal; filled dots denote stationary probation beliefs. Cutoff locations are schematic and are not comparable across rows. The optional isolated stationary belief at $c = c_N(k)$ is omitted.

For $c \leq c_N(k)$, the principal fires below the static cutoff and hires above it. At $\beta = \beta^*$, firing and hiring are payoff-equivalent. At the knife edge $c = c_N(k)$, one may also insert the isolated stationary probation belief

$$\beta^* = \beta_{S,0} = \beta_{S,h}, \qquad \lambda_G(\beta^*) = \lambda_B(\beta^*) = \lambda_S,$$

without changing the principal's value.

Fail–pass regime. In case (iv), let β° denote the stationary belief. At this belief, $\lambda_G(\beta^\circ) = \lambda_B(\beta^\circ) = \lambda_S$, $\Pi(\beta^\circ) = \Pi_S(\beta^\circ)$, and $V(\beta^\circ) = V_S$, as defined in (3.6), (3.7), and (3.8). The proof of Theorem 4.1 shows that the two extreme one-sided tests yield the same local return at β°:

$$\underbrace{\beta^\circ \frac{c}{H - V_S} \frac{rG + k}{r + \lambda_S}}_{\text{pass-only local return}} = \underbrace{(1 - \beta^\circ) \frac{c}{V_S - F} \frac{k}{r + \lambda_S}}_{\text{fail-only local return}}.$$

Intuitively, at the maximal pass rate $c/(H - V_S)$, a pass when the agent is good produces the flow-equivalent hiring payoff rG and ends probation, saving the flow cost k. At the maximal fail rate $c/(V_S - F)$, a fail when the agent is bad ends probation and saves k, since the firing payoff is zero. Equating these local returns gives the stationary belief

$$\beta^\circ := \frac{k(Hr + c)}{k(Hr + c) - rF(Gr + k)} \in (0, 1). \tag{4.2}$$

For a nondegenerate fail–pass equilibrium to exist, the stationary belief β° must lie strictly inside the interval $(\beta_{S,0}, \beta_{S,h})$, on which Π_S exceeds both stopping payoffs.

Simple algebra shows that

$$\beta^\circ > \beta_{S,0} \quad \Longleftrightarrow \quad c > c_P(k) := \frac{k(H-F)}{G} - rF, \tag{4.3}$$

$$\beta^\circ < \beta_{S,h} \quad \Longleftrightarrow \quad k < k_F := \frac{rBF}{H-F}. \tag{4.4}$$

The lines $c = c_N(k)$ and $c = c_P(k)$ cross at $(k_F, c^\dagger)$, where $c^\dagger := c_P(k_F)$; see Figure 2.[6]

In case (iv), both conditions hold, so the stationary payoff at β° strictly exceeds both stopping payoffs, and the probation set is

$$\mathcal{K} = (\underline{\beta}, \overline{\beta}), \qquad 0 < \underline{\beta} < \beta_{S,0} < \beta^\circ < \beta_{S,h} < \overline{\beta} < 1.$$

The principal fires for $\beta \leq \underline{\beta}$ and hires for $\beta \geq \overline{\beta}$. On $(\underline{\beta}, \beta^\circ)$, the fail-only system (A.18) is solved to the left from the stationary values at β°. The lower threshold $\underline{\beta}$ is uniquely determined by $\Pi(\underline{\beta}) = 0$. On $(\beta^\circ, \overline{\beta})$, the pass-only system (A.20) is solved to the right from the same stationary values. The upper threshold $\overline{\beta}$ is uniquely determined by $\Pi(\overline{\beta}) = h(\overline{\beta})$.

Fail-only one-sided equilibrium. In case (ii), there must be a belief $\underline{\beta} < \beta_{S,h}$ such that the principal fires for $\beta \leq \underline{\beta}$, uses fail-only tests for $\beta \in (\underline{\beta}, \beta_{S,h})$, offers the stationary test at $\beta_{S,h}$, and hires for $\beta > \beta_{S,h}$. The lower boundary is uniquely determined by value matching with firing, $\Pi(\underline{\beta}) = 0$.

Pass-only one-sided equilibrium. Case (iii) is analogous to case (ii): there must be a belief $\overline{\beta} > \beta_{S,0}$ such that the principal hires for $\beta \geq \overline{\beta}$, uses pass-only tests for $\beta \in (\beta_{S,0}, \overline{\beta})$, offers the stationary test at $\beta_{S,0}$, and fires for $\beta < \beta_{S,0}$. The upper boundary is uniquely determined by value matching with hiring, $\Pi(\overline{\beta}) = h(\overline{\beta})$.

Remark 4.1. For tractability, we require the principal to hire after a pass and fire after a fail, so that any news arrival ends probation immediately. For fixed continuation values, the bad type's instantaneous best response is $e = 0$ when the shirking incentive constraint holds strictly (on $\mathcal{A}(\beta; V) \setminus \mathcal{A}_0(\beta; V)$) and $e = 1$ when it is violated (outside $\mathcal{A}(\beta; V)$). This dichotomy can fail if the principal instead retains discretion after observing test outcomes. Consider a pass-only test $(\lambda_G, 0) \notin \mathcal{A}(\beta; V)$, holding the candidate continuation values fixed. Such a test cannot sustain $e = 1$: a pass would then be uninformative and leave the principal's continuation decision unchanged, removing the bad type's incentive to exert costly effort. Nor can it sustain $e = 0$, since a pass would then be conclusive and lead to hiring, making imitation profitable. The bad type must therefore mix.

[6] Their inverses, used in the Online Appendix, are $k_N(c) := cG(-B)/[(H-F)(G-B)]$ and $k_P(c) := G(c + rF)/(H-F)$.

Our analysis of this alternative formulation, available upon request, shows that the equilibria characterized here remain equilibria for a range of parameter values, but also identifies an open region in which equilibria involve mixed strategies on the equilibrium path. The full analysis therefore requires a richer class of Markov assessments, allowing the principal to randomize over tests and the bad type to exert interior effort on path. Characterizing the resulting mixing probabilities and belief dynamics substantially complicates the analysis while adding limited insight into our central question of how manipulation incentives shape the direction and intensity of testing.

The prescribed consequences of test outcomes can be interpreted as an institutional constraint. For example, a probationary appointment may guarantee confirmation following a successful assessment and require termination following a definitive failure. The employer may retain discretion over the evaluation process and over early stopping while remaining bound by the consequences of a formal assessment once its outcome is observed. The assumption thus captures a limited form of commitment to the consequences of realized test outcomes, without requiring commitment to a future testing policy.

5 No-manipulation attention benchmark

To isolate the role of manipulation, suppose that the bad type cannot imitate the good type (shirking is enforceable) and that the principal allocates a fixed attention capacity between pass and fail signals:

$$\lambda_G + \lambda_B \leq \bar{\lambda}, \qquad \bar{\lambda} > 0. \tag{5.1}$$

This is the attention problem of Che and Mierendorff [2019], specialized to our payoffs. For a clean comparison, choose $\bar{\lambda}$ above the largest total testing intensity used in the strategic equilibrium:

$$\bar{\lambda} > \sup_{\beta \in \mathcal{K}} \left\{ \lambda_G^S(\beta) + \lambda_B^S(\beta) \right\}, \tag{5.2}$$

where the superscript S denotes strategic-equilibrium intensities; the supremum is taken to be zero when $\mathcal{K} = \varnothing$. Thus, imposing the same capacity constraint on the strategic model would leave its equilibrium unchanged. In Che and Mierendorff [2019], the source that produces conclusive evidence for the good state corresponds to our pass-only monitoring; the other source corresponds to fail-only monitoring.[7]

In this benchmark, the principal optimizes jointly over stopping rules and measurable attention controls adapted to the observed history, subject to (5.1). The restrictions

[7]To obtain the strictly positive correct-action payoffs required by Che and Mierendorff [2019], add $k/(2r)$ to every terminal payoff and replace the flow cost by $k/2$. Every policy's discounted value then increases by exactly $k/(2r)$, including when testing never ends, so optimal testing and stopping are unchanged.

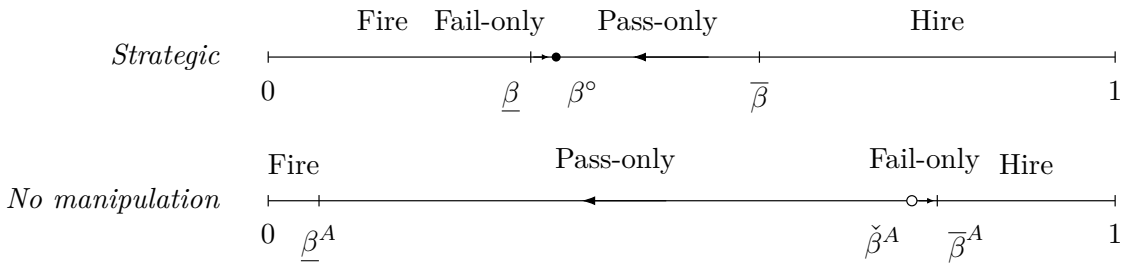


Figure 4: Strategic and no-manipulation policies on a common belief scale. Parameters: $G = H = 1$, $B = F = -1$, $r = 1$, $c = 8/5$, $k = 1/4$, and $\bar{\lambda} = 4$. Arrows indicate the direction of no-news belief drift. The filled point is the strategic stationary belief; the open point is the no-manipulation benchmark's repelling switch, not a stationary belief. Cutoff locations are rounded.

on admissible simple Markov assessments in Definition 2.2, used for the strategic equilibrium analysis, are not imposed on this benchmark. Markov policies may have isolated switching beliefs; at a repelling switch, their specification includes a choice of outgoing no-news trajectory.

Proposition 5.1. *There are thresholds $0 \leq \underline{k}^A \leq \overline{k}^A$ such that an optimal Markov policy has the following form, where pass-only and fail-only monitoring use rates $(\bar{\lambda}, 0)$ and $(0, \bar{\lambda})$, respectively:*

(i) ***(No probation)*** *If $k \geq \overline{k}^A$, the principal stops immediately.*

(ii) ***(Pass–fail)*** *If $\underline{k}^A \leq k < \overline{k}^A$, probation occurs on $(\underline{\beta}^A, \overline{\beta}^A)$. The principal uses pass-only monitoring below a switching belief $\check{\beta}^A$ and fail-only monitoring above it. In the absence of news, the posterior moves toward a stopping boundary.*

(iii) ***(Pass–fail–pass–fail)*** *If $0 < k < \underline{k}^A$, there are cutoffs*

$$\underline{\beta}^A < \beta_-^A < \beta_o^A < \beta_+^A < \overline{\beta}^A$$

such that monitoring is pass-only on $(\underline{\beta}^A, \beta_-^A) \cup (\beta_o^A, \beta_+^A)$, fail-only on $(\beta_-^A, \beta_o^A) \cup (\beta_+^A, \overline{\beta}^A)$, and stationary at $\beta_o^A = k/(rG + 2k)$, where $\lambda_G = \lambda_B = \bar{\lambda}/2$. In the absence of news, the posterior moves toward a stopping boundary in the two outer regions and toward β_o^A in the two inner regions.

Accuracy. Strategic probation classifies both types correctly: no news leads to a stationary test rather than to a stopping decision. The no-manipulation benchmark is less accurate in its outer regions. Under pass-only monitoring, a good type is mistakenly fired if no pass arrives before the lower boundary; under fail-only monitoring, a bad type is mistakenly hired if no fail arrives before the upper boundary. In the two inner regions of case (iii), the benchmark classifies both types correctly, as does strategic probation.

In Figure 4, strategic monitoring is more accurate on the common probation interval $(\underline{\beta}, \overline{\beta})$. The no-manipulation benchmark is more accurate on $(\underline{\beta}^A, \underline{\beta})$ and $(\overline{\beta}, \overline{\beta}^A)$, where it continues to learn while the strategic principal stops; the two models coincide outside $(\underline{\beta}^A, \overline{\beta}^A)$.

Delay. On the two intervals where the strategic principal stops immediately but the no-manipulation benchmark continues, the benchmark's additional accuracy comes with delay. On the common probation interval in Figure 4, the comparison is subtler. The benchmark tests more intensively, since it uses the full capacity $\bar{\lambda}$, which exceeds every strategic total testing intensity. Yet, in the absence of news, its posterior travels to the stopping cutoff $\underline{\beta}^A$, whereas strategic testing reaches β° and then continues until a conclusive signal arrives. On $(\underline{\beta}, \beta^\circ)$, the two models also seek different evidence: strategic monitoring is fail-only, whereas benchmark monitoring is pass-only. Rates and destinations alone therefore do not rank decision times across the two models in general. For these parameters, direct computation shows that the benchmark reaches a decision sooner in expectation, for each type and at every belief in $(\underline{\beta}, \overline{\beta})$; on this interval, strategic probation thus buys accuracy with delay.

Within the strategic model, however, the ranking of expected decision times across types is sharp and depends on the direction of monitoring. Under pass-only monitoring, the good type may be identified by a pass before beliefs reach the stationary belief, whereas the bad type cannot generate a pass and must wait for the no-news path to reach that belief and then for a fail. The bad type therefore waits longer. Under fail-only monitoring the ranking is reversed: the bad type may be identified by a fail, while the good type must wait for the stationary test. Hence, in the fail–pass equilibrium shown in Figure 4, the good type waits longer below β° and the bad type waits longer above it. Economically, the principal's choice of what evidence to seek determines not only the direction of learning but also which type bears the delay required for perfect classification.

6 Commitment benchmark

The commitment problem provides a useful benchmark for the equilibrium analysis: it isolates what the principal can achieve when she chooses the entire testing and stopping policy ex ante. We impose an upper bound $\bar{\lambda}$ on each test intensity, rather than on their sum as in Section 5. The bound matters under commitment because the principal may wish to use the most aggressive feasible test. By contrast, the same bound would not affect the equilibrium characterization if it exceeded all equilibrium intensities for the parameter vector under consideration. We assume $\bar{\lambda} \geq c/H$, so the pass-only policy below remains feasible.

Theorem 6.1. *Fix a prior $\beta \in (0,1)$. Suppose that the principal can commit to a possibly randomized testing and stopping policy, subject to the upper bounds $0 \leq \lambda_G \leq \bar{\lambda}$ and $0 \leq \lambda_B \leq \bar{\lambda}$ at every history, and that the bad type best responds, shirking when indifferent. Assume that $\bar{\lambda} \geq c/H$. At least one of the following policies is optimal:*

(i) *Immediate firing.*

(ii) *Immediate hiring.*

(iii) *If $c + rF > 0$, maximal fail monitoring with $\lambda_B = \bar{\lambda}$, $\lambda_G = \lambda_G^{\bar{\lambda}}$, and no stopping in the absence of news, where $\lambda_G^{\bar{\lambda}} \in (0, c/H)$ is the unique constant pass intensity that makes the bad type indifferent between shirking and exerting high effort.*

(iv) *Perpetual pass-only monitoring with $\lambda_G = c/H$ and $\lambda_B = 0$, and no stopping in the absence of a pass.*

The theorem reduces the commitment problem to a date-zero comparison among two immediate decisions and the available screening policies. The substantive new cases are (iii) and (iv). Under policy (iii), which is needed only when $c + rF > 0$, the fail rate is set at the upper bound and the pass rate is chosen so that the bad type's incentive constraint binds. Since $\lambda_G^{\bar{\lambda}} < c/H \leq \bar{\lambda}$, the posterior belief that the type is good increases toward one as long as no news arrives. The cap binds on the fail rate, which screens out the bad type most aggressively, while the pass rate keeps him indifferent between shirking and exerting high effort.

Policy (iv) has the opposite belief dynamics. With pass-only monitoring, the posterior belief that the type is good decreases toward zero as long as no news arrives. A good type is eventually hired at the first pass, whereas a bad type shirks, produces no signal, and remains on probation forever. The principal therefore weighs the cost of keeping a bad type on probation against the benefit of screening a good type without inducing imitation.

Policies (iii) and (iv) require strong commitment by the principal. Under policy (iii), the absence of news makes her increasingly confident that the agent is good, yet she withholds hiring until a pass occurs. Under policy (iv), the absence of news makes her increasingly confident that the agent is bad, yet she continues probation indefinitely unless a pass occurs. Both policies therefore require her to continue testing even when she would prefer to hire or fire immediately. In practice, such commitment may be difficult to sustain.

7 Extensions

We examine two extensions, each varying one aspect of the baseline model. The characterized equilibria survive when the good type chooses effort strategically or, separately, when the test menu includes monotone imperfect tests. The Online Appendix

analyzes the cases of productive probation and heterogeneous discounting, which preserve the phase-diagram structure under restrictions stated there and clarify how the benefits and costs of waiting shape monitoring.

7.1 Strategic effort by the good type

The baseline assumes that the good type exerts high effort. This extension asks whether inducing that effort imposes an additional restriction on monitoring when both types are strategic. We relax Assumption 1 by allowing the good type to choose effort $e_G \in [0, 1]$ at flow cost $c_H e_G$, where $0 < c_H < c$. All other primitives are unchanged. We retain the equilibrium notion in Definition 4.1, augmented by sequential optimality of the good type's effort.

Proposition 7.1. *Every equilibrium characterized in Theorem 4.1 admits an extension to the model with a strategic good type. The principal uses the same stopping and testing policy, the bad type shirks, and the good type exerts high effort. The principal's value, the bad type's continuation value, and all equilibrium thresholds are unchanged. The good type's high-effort incentive constraint is strict at every probation belief.*

The result is not immediate from $c_H < c$. Write V_G and V for the good and bad types' continuation values. Under the prescribed policy and effort choices, $V_G > V$, which reduces the good type's gain from passing even as it increases his loss from failing. The proof in the Online Appendix shows that, at every probation belief,

$$\frac{H - V_G}{c_H} > \frac{H - V}{c}, \qquad \frac{V_G - F}{c_H} > \frac{V - F}{c}.$$

Thus, relative to effort cost, both the gain from passing and the loss from failing are larger for the good type. Since each baseline test makes the bad type indifferent between shirking and high effort, it strictly induces high effort by the good type.

The principal's prescribed tests also remain pointwise optimal. With pure best responses at off-path tests, the same inequalities rule out tests that induce high effort only from the bad type. Tests that induce high effort only from the good type belong to the baseline incentive-compatible set and generate the same signal laws as in the baseline model. By contrast, tests that induce the same effort from both types are uninformative and dominated. The bad type's incentive constraint therefore remains the binding restriction on information acquisition: the baseline monitoring policy both deters imitation and provides the good type with incentives to work.

7.2 Imperfect tests

The restriction to conclusive tests raises a design question: could the principal improve screening by introducing errors that weaken the bad type's incentive to manipulate?

We enlarge the test menu to allow passes under low effort and failures under high effort, subject to monotonicity. The baseline equilibria survive because these errors add random hiring or firing without relaxing the incentive constraint on the informative part of a test.

Specifically, the principal chooses four rates

$$(\lambda_G^1, \lambda_G^0, \lambda_B^1, \lambda_B^0) \in \mathbb{R}_+^4.$$

Under effort $e \in [0,1]$, passes arrive at rate $e\lambda_G^1 + (1-e)\lambda_G^0$, and failures arrive at rate $e\lambda_B^1 + (1-e)\lambda_B^0$. We impose $\lambda_G^1 \geq \lambda_G^0$ and $\lambda_B^0 \geq \lambda_B^1$: high effort weakly increases the pass rate and weakly decreases the failure rate. All other baseline assumptions are unchanged. In particular, the good type is again assumed to exert high effort, and passes and failures still require hiring and firing, respectively.

Proposition 7.2. *Every equilibrium characterized in Theorem 4.1 remains an equilibrium when the principal's test menu is enlarged to monotone imperfect tests, within the same equilibrium class and with the same stationary and boundary conventions. The principal's policy, both continuation values, and all equilibrium thresholds are unchanged. At every nonstationary probation belief of these equilibria, tests with positive error rates are strictly dominated.*

To see the mechanism, fix the continuation values (Π, V) of a baseline equilibrium and a nonstationary probation belief β. Every monotone test consists of a conclusive test with rates

$$\widetilde{\lambda}_G := \lambda_G^1 - \lambda_G^0 \geq 0, \qquad \widetilde{\lambda}_B := \lambda_B^0 - \lambda_B^1 \geq 0,$$

together with effort-independent pass and failure hazards λ_G^0 and λ_B^1. Only the informative components enter the bad type's shirking constraint:

$$\widetilde{\lambda}_G[H - V(\beta)] + \widetilde{\lambda}_B[V(\beta) - F] \leq c.$$

Under shirking, the common hazards also cancel from the difference between the types' total signal rates. Removing them therefore preserves both local effort incentives and the no-news posterior drift, holding continuation values fixed.

The common hazards instead replace continuation by hiring or firing without providing information about type. For a shirking-compatible test, the test-dependent part of the principal's first-order payoff gain is

$$\begin{aligned} \Psi_P^{\text{imp}} = \Psi_P(\beta, \widetilde{\lambda}_G, \widetilde{\lambda}_B; \Pi) \\ + \lambda_G^0[h(\beta) - \Pi(\beta)] - \lambda_B^1 \Pi(\beta). \end{aligned} \tag{7.1}$$

Since $\Pi(\beta) > \max\{0, h(\beta)\}$ throughout each nonstationary baseline region, either positive error rate strictly reduces this payoff. Tests that induce high effort from the

bad type give both types the same signal law and remain uninformative and dominated. The enlarged menu therefore creates no profitable pointwise testing deviation.

The proof in the Online Appendix also verifies the stationary and boundary conditions. The result concerns the principal's choice to introduce errors, not a technology in which errors are unavoidable. Monotonicity is substantive: it permits the decomposition into a conclusive test and effort-independent stopping hazards on which the argument relies.

8 Conclusion

This paper studies dynamic test design when an agent can manipulate evidence and the principal cannot commit to future testing or stopping decisions. The central insight is that information is constrained by the incentives tests create. More intensive monitoring can induce the bad type to mimic the good type, generating an endogenous information budget that depends on his continuation value. The threat of manipulation therefore shapes monitoring even when no manipulation occurs in equilibrium.

Within the simple Markov equilibrium class studied here, monitoring is one-sided except at beliefs that remain unchanged in the absence of news. Fail-only tests are attractive when the bad type's continuation value is low, because he has less to lose from failure; pass-only tests are attractive when it is high, because he has less to gain from passing. In either one-sided region, testing becomes less intense in the absence of news. These results explain why an employer might concentrate on detecting misconduct rather than recognizing strong performance, or vice versa. They also show why declining testing intensity can reflect changing manipulation incentives rather than technological limitations.

Manipulation also changes the relationship between accuracy and delay. Without manipulation, the principal may end testing without a conclusive signal, accepting some risk of misclassification. In equilibrium, incentive compatibility rules out these stopping patterns: conditional on entering probation, both types are eventually classified correctly. When incentive-compatible learning is too costly, the principal instead stops immediately and risks misallocation. Accurate screening therefore entails delay, and monitoring determines who bears it: bad types wait longer than good types under pass-only monitoring, while fail-only monitoring reverses this ranking. The choice of evidence thus affects both how long suitable workers wait for confirmation and how long unsuitable workers remain under evaluation.

The commitment benchmark highlights a separate constraint: a testing policy that is initially incentive-compatible for the agent can become unattractive to the principal as information accumulates. Commitment can sustain testing even as beliefs approach certainty and the principal would otherwise prefer to hire or fire. This distinction matters for firms whose managers retain discretion over confirmation and dismissal: an announced evaluation policy need not bind a manager who prefers to stop rather than

prolong costly probation. Designing a test is therefore distinct from committing to use it until the prescribed outcome occurs.

The one-sided structure survives heterogeneous discounting and productive probation under restrictions stated in the Online Appendix. The baseline equilibria also remain valid when the good type can choose effort strategically and faces a lower effort cost than the bad type. Allowing monotone imperfect tests likewise leaves these equilibria intact: introducing failures under high effort or passes under low effort does not benefit the principal. More broadly, monitoring is not merely a technology for collecting information. Its design must account for what evidence can be obtained without inducing manipulation, how delay is distributed across types, and whether the principal can credibly maintain testing as her beliefs change.

A Proofs and supplementary results

A.1 Value characterization and effort verification

This subsection collects two technical results used throughout the analysis. The first establishes the no-news dynamics and characterizes the induced continuation values by their Bellman equations. The second turns pointwise effort optimality into optimality against arbitrary progressively measurable effort deviations. The proofs for Section 3 follow.

Lemma A.1. *Let $\alpha = (a, \lambda_G, \lambda_B)$ be an admissible simple Markov assessment prescribing shirking by the bad type. From each nonstationary probation belief β, the no-news path $\hat{\beta}(\cdot, \beta)$ is uniquely determined up to its first exit time $\sigma(\beta)$ from the component containing β. At a finite exit, write $\hat{\beta}(\sigma(\beta), \beta)$ for its limiting endpoint. The induced continuation values are the unique bounded functions $\Pi, V : [0,1] \to \mathbb{R}$ with the following properties.*

At stopping beliefs, $(\Pi(\beta), V(\beta)) = (h(\beta), H)$ if $a(\beta) = \mathsf{h}$, and $(\Pi(\beta), V(\beta)) = (0, F)$ if $a(\beta) = \mathsf{f}$. At every $\beta \in \mathcal{S}$,

$$\Pi(\beta) = \frac{\beta\lambda_G(\beta)G - k}{r + \lambda_G(\beta)}, \qquad V(\beta) = \frac{\lambda_B(\beta)F}{r + \lambda_B(\beta)}.$$

On each connected component I of $\mathcal{K} \setminus \mathcal{S}$, both functions are locally absolutely continuous and satisfy (2.2) and (2.3) for almost every $\beta \in I$. Finally, for every $\beta \in \mathcal{K} \setminus \mathcal{S}$ with $\sigma(\beta) < \infty$, writing $\beta_t := \hat{\beta}(t, \beta)$ and $\sigma := \sigma(\beta)$,

$$\lim_{t \uparrow \sigma} \Pi(\beta_t) = \Pi(\beta_\sigma), \qquad \lim_{t \uparrow \sigma} V(\beta_t) = V(\beta_\sigma). \tag{A.1}$$

Proof. Fix $\beta_0 \in \mathcal{K} \setminus \mathcal{S}$, and let I be its connected component. Admissibility makes $\lambda_B - \lambda_G$, and hence b, locally Lipschitz on I. Picard–Lindelöf gives a unique maximal

classical path $\hat{\beta}(\cdot, \beta_0) : [0, \sigma(\beta_0)) \to I$ satisfying

$$\hat{\beta}(0, \beta_0) = \beta_0, \qquad \dot{\hat{\beta}}(t, \beta_0) = b\Big(\hat{\beta}(t, \beta_0)\Big). \tag{A.2}$$

Since b has constant nonzero sign on I, the path is strictly monotone. Dividing by $\hat{\beta}(1 - \hat{\beta})$ gives

$$\frac{d}{dt} \log \frac{\hat{\beta}(t, \beta_0)}{1 - \hat{\beta}(t, \beta_0)} = \lambda_B\Big(\hat{\beta}(t, \beta_0)\Big) - \lambda_G\Big(\hat{\beta}(t, \beta_0)\Big). \tag{A.3}$$

At a finite exit, monotonicity gives a limit in $\overline{I}$. An interior limit would allow extension, contrary to maximality, so the limit lies in ∂I. There the assessment either stops or, at a stationary belief, keeps the posterior fixed; these continuations are specified by the assessment rather than by (A.2). On each compact interval $J \subseteq I$, continuity and the nonzero sign of b give $0 < m_J \leq |b(\beta)| \leq M_J < \infty$. Integrating these bounds shows that the no-news flow and its inverse are Lipschitz while the path stays in J.

Immediate stopping gives the stated terminal values. At a stationary belief, $\lambda_G = \lambda_B$ and the posterior remains fixed; conditioning on the first signal gives the stated stationary values. Terminal payoffs are bounded, and discounted probation costs are at most k/r, so both induced continuation values are bounded.

Fix $\beta \in \mathcal{K} \setminus \mathcal{S}$, let I be its connected component, and write $\beta_t := \hat{\beta}(t, \beta)$ and $\sigma := \sigma(\beta)$. For $0 \leq t < \sigma$, define the discounted survival probabilities

$$D_P(t, \beta) := \exp\left\{-rt - \int_0^t \Big[\beta_s \lambda_G(\beta_s) + (1 - \beta_s)\lambda_B(\beta_s)\Big] ds\right\},$$
$$D_B(t, \beta) := \exp\left\{-rt - \int_0^t \lambda_B(\beta_s)\, ds\right\}.$$

Under assessed shirking, public pass and fail intensities are $\beta_t \lambda_G(\beta_t)$ and $(1 - \beta_t)\lambda_B(\beta_t)$; conditional on the bad type, only failure occurs, at rate $\lambda_B(\beta_t)$. Conditioning on the first signal before any $0 \leq T < \sigma$ therefore gives

$$\Pi(\beta) = \int_0^T D_P(t, \beta)\Big[-k + \beta_t \lambda_G(\beta_t) G\Big] dt + D_P(T, \beta)\Pi(\beta_T), \tag{A.4}$$
$$V(\beta) = \int_0^T D_B(t, \beta)\lambda_B(\beta_t) F\, dt + D_B(T, \beta) V(\beta_T). \tag{A.5}$$

On compact subintervals of I, these representations and the positive, absolutely continuous survival factors imply absolute continuity of $\Pi(\beta_t)$ and $V(\beta_t)$ in time. The local Lipschitz property of the no-news flow and its inverse then makes both values locally absolutely continuous in belief, and differentiation yields (2.2)–(2.3) almost everywhere. If $\sigma < \infty$, conditioning up to the exit gives (A.1): bounded rates imply that, as the remaining travel time vanishes, the accumulated payoff terms tend to zero and the discounted survival probabilities tend to one.

Conversely, let $(\widetilde{\Pi}, \widetilde{V})$ be any bounded pair satisfying the stated conditions. Along each no-news path, multiply its Bellman equations by D_P and D_B, respectively, and integrate. The chain rule gives (A.4)–(A.5) with the candidate values in place of (Π, V). At a finite exit, value matching and the stopping or stationary formulas determine the terminal terms, since the reached endpoint lies outside $\mathcal{K} \setminus \mathcal{S}$. At an infinite exit, boundedness and $D_P(T,\beta), D_B(T,\beta) \le e^{-rT}$ make those terms vanish as $T \to \infty$. Thus the candidate pair equals the induced values at every belief, proving uniqueness. □

Lemma A.2. *Let $\alpha = (a, \lambda_G, \lambda_B)$ be an admissible simple Markov assessment, and let $\bar{c} > 0$ be an effort-cost parameter. Let W be a bounded candidate continuation value for an agent whose effort $e \in [0,1]$ generates a pass at rate $e\lambda_G$, a fail at rate $(1-e)\lambda_B$, and flow cost $\bar{c}e$. Suppose that:*

(i) *$W = H$ after hiring or a pass and $W = F$ after firing or a fail;*

(ii) *W is locally absolutely continuous on each nonstationary component and satisfies value matching at every endpoint that the no-news path reaches in finite time; and*

(iii) *at almost every nonstationary probation belief, and algebraically at every stationary probation belief,*

$$rW = bW' + \lambda_B(F - W) + \max_{u \in [0,1]} u\,\Gamma_W, \qquad \Gamma_W := \lambda_G(H - W) - \lambda_B(F - W) - \bar{c}, \tag{A.6}$$

where the term bW' is omitted at a stationary belief.

Then every measurable pointwise selector $e^*(\beta) \in \arg\max_{u \in [0,1]} u\,\Gamma_W(\beta)$ is optimal against all progressively measurable effort deviations. More precisely, for every initial probation belief β_0 and every progressively measurable process $e_t \in [0,1]$,

$$J_{\bar{c}}(e;\beta_0) - W(\beta_0) = \mathbb{E}\int_0^{\tau_{\text{end}}} e^{-rs}\left[e_s \Gamma_W(\beta_s) - \max_{u \in [0,1]} u\Gamma_W(\beta_s)\right] ds \le 0, \tag{A.7}$$

where $J_{\bar{c}}(e;\beta_0)$ is the agent's expected discounted terminal payoff net of effort costs and τ_{end} is the first signal or the assessed stopping time.

In particular, if $W = V$, $\bar{c} = c$, and

$$\lambda_G(\beta)\big[H - V(\beta)\big] - \lambda_B(\beta)\big[F - V(\beta)\big] \le c \tag{A.8}$$

throughout probation, then $e_B \equiv 0$ is globally optimal. If the inequality is strict along every state occupied with positive discounted time, shirking is the unique optimum. Symmetrically, if $\Gamma_W > 0$ throughout probation, high effort is the unique optimum.

Proof. Fix a progressively measurable effort process $e_t \in [0,1]$. Admissibility bounds both intensities on the probation set: there are only finitely many nonstationary

components and stationary beliefs, and the rates are bounded on each. Under the deviation, the pass and fail intensities are $e_t\lambda_G(\beta_t)$ and $(1-e_t)\lambda_B(\beta_t)$, while the public no-news drift remains b.

Apply the chain rule and the compensator formula to discounted W, less accumulated effort costs, up to $\tau_{\text{end}} \wedge n$, for an integer $n \geq 1$. The controlled Bellman equation makes its predictable drift $e^{-rt}\left[e_t\Gamma_W(\beta_t) - \max_{u\in[0,1]} u\Gamma_W(\beta_t)\right]$. The same identity holds at stationary beliefs by the algebraic Bellman equation. Matching at each reached endpoint cancels the intermediate boundary terms. Bounded rates and bounded W make the stopped compensated terms integrable, so taking expectations gives the stopped version of (A.7).

Let $n \to \infty$. The terminal continuation term vanishes on $\{\tau_{\text{end}} = \infty\}$, and the discounted effort cost is bounded by $\bar{c}/r$. The drift integrand is also bounded by a constant times e^{-rt}. Dominated convergence therefore gives (A.7). Its integrand is nonpositive for every effort process and is zero for a maximizing selector. This proves optimality and both sign-based conclusions; strict optimality is understood up to changes on sets of zero time–probability measure before termination. □

A.2 Proofs of results in Section 3

Proof of Lemma 3.1. For an incentive-compatible test (ℓ_G, ℓ_B), the principal's payoff over dt, followed by continuation, is

$$-kdt + \beta\ell_G G dt + \Big[1 - \beta\ell_G dt - (1-\beta)\ell_B dt\Big](1 - rdt)\Pi(\widehat{\beta}_{dt}) + o(dt).$$

Since $\widehat{\beta}_{dt} - \beta = \beta(1-\beta)(\ell_B - \ell_G)\,dt + o(dt)$, this payoff minus $\Pi(\beta)$, divided by dt, converges to $-k - r\Pi + \Psi_P(\beta, \ell_G, \ell_B; \Pi)$. We call Ψ_P the principal's Hamiltonian; pointwise optimality is thus equivalent to maximizing it over $\mathcal{A}(\beta; V)$.

Moreover, $\Pi(\beta) < \beta G$: the principal's discounted terminal payoff is at most G conditional on the good type and at most zero conditional on the bad type, and a strict continuation belief entails a strictly positive expected probation cost. For every $0 < \ell \leq c/(H-F)$, the common-rate test is feasible and has Hamiltonian $\ell(\beta G - \Pi) > 0$. The zero test therefore cannot maximize.

Because $F < V < H$, the feasible set is the triangle with vertices $(0,0)$, $(c/(H-V), 0)$, and $(0, c/(V-F))$. Any nonzero test with slack incentives and a positive Hamiltonian can be scaled up to increase that Hamiltonian. Thus every maximizer over $\mathcal{A}(\beta; V)$ lies on the edge joining the two nonzero vertices, which is exactly $\mathcal{A}_0(\beta; V)$. Since Ψ_P is linear and $\mathcal{A}(\beta; V)$ is compact, a maximizer exists, so maximizing over $\mathcal{A}(\beta; V)$ and over $\mathcal{A}_0(\beta; V)$ yields the same maximizers. □

Proof of Lemma 3.2. Use posterior log-odds and their inverse,

$$x := \log \frac{\beta}{1-\beta}, \qquad \beta(x) := \frac{e^x}{1+e^x}. \tag{A.9}$$

Bounded rates imply that a finite no-news endpoint lies in $(0,1)$. For a belief in a component with such an endpoint, let $\sigma(\beta)$ be the remaining no-news travel time. This time tends to zero as the endpoint is approached from within the component.

Suppose first that a pass-learning component ends in firing at its lower endpoint β'. Value matching (Lemma A.1) and the payoff guarantees imply $\Pi \to 0$, $V \to F$, and $\beta' \le \beta_{S,0}$. Incentive compatibility gives

$$\lambda_G \le \lambda_S + \frac{(\lambda_G - \lambda_B)(V-F)}{H-F}, \qquad \limsup_{\beta \downarrow \beta'} \lambda_G \le \lambda_S.$$

If $\beta' < \beta_{S,0}$, the principal's payoff representation has strictly negative flow $-k + \beta\lambda_G G$ near the endpoint and terminal value zero, contradicting $\Pi \ge 0$.

If $\beta' = \beta_{S,0}$, the same representation and discounted survival bounded by one give, for a constant $C > 0$ and all sufficiently nearby $\beta > \beta'$,

$$0 \le \Pi(\beta) \le C(\beta - \beta')\sigma(\beta) + C \sup_{\beta' < \xi \le \beta} |V(\xi) - F| \, [x(\beta) - x(\beta')] = o(\beta - \beta').$$

Here $\int_0^{\sigma(\beta)} (\lambda_G - \lambda_B) dt = x(\beta) - x(\beta')$, and log-odds are locally Lipschitz at β'. But $\Pi_S(\beta') = 0$ and $\Pi_S' = \lambda_S G/(r+\lambda_S) > 0$, contradicting $\Pi \ge \Pi_S$.

Now suppose that a fail-learning component ends in hiring at its upper endpoint β''. Value matching and the payoff guarantees imply $\Pi - h \to 0$, $V \to H$, and $\beta'' \ge \beta_{S,h}$, while

$$\lambda_B \le \lambda_S + \frac{(\lambda_B - \lambda_G)(H-V)}{H-F}, \qquad \limsup_{\beta \uparrow \beta''} \lambda_B \le \lambda_S.$$

Subtracting the payoff of immediate hiring from the principal's pathwise representation gives

$$(\Pi - h)(\beta) = -\int_0^{\sigma(\beta)} D_P(t,\beta)\Big[rh(\beta_t) + k + (1-\beta_t)\lambda_B(\beta_t)B\Big] dt.$$

At $\lambda_B = \lambda_S$, the bracket is affine and strictly increasing in β, with its zero at $\beta_{S,h}$. Since $B < 0$, the rate bound makes the bracket strictly positive near the endpoint if $\beta'' > \beta_{S,h}$, contradicting $\Pi \ge h$.

If $\beta'' = \beta_{S,h}$, the same bound instead yields, for a constant $C > 0$ and all sufficiently nearby $\beta < \beta''$,

$$0 \le (\Pi - h)(\beta) \le C(\beta'' - \beta)\sigma(\beta) + C \sup_{\beta \le \xi < \beta''} |H - V(\xi)| \, [x(\beta'') - x(\beta)] = o(\beta'' - \beta).$$

Yet $\Pi_S - h$ vanishes at β'' and has strictly negative posterior derivative $\lambda_S G/(r+\lambda_S) -$

$(G - B)$. Thus $\Pi - h < \Pi_S - h$ immediately to its left, again contradicting $\Pi \geq \Pi_S$. $\square$

A.3 Proof of Theorem 4.1

Proof of Theorem 4.1. The thresholds $c_N(k)$, k_F, and $c_P(k)$ are positive, and $\beta^\circ \in (0, 1)$.

Step 1: Local test choice. On a nonstationary component, $F < V < H$: under shirking, the bad type receives no flow payoff and a terminal payoff in $[F, H]$, and bounded rates imply that this payoff arrives after a positive duration; since $F < 0 < H$, discounting places V strictly between F and H. In posterior log-odds $x = \log(\beta/(1-\beta))$, write $\Delta := \lambda_B - \lambda_G$. Since V is continuous there (Lemma A.1), Lemma 3.1 implies that the chosen test is nonzero and incentive compatibility binds:

$$(H - V)\lambda_G + (V - F)\lambda_B = c. \tag{A.10}$$

Hence

$$\lambda_G = \frac{c - (V - F)\Delta}{H - F}, \qquad \lambda_B = \frac{c + (H - V)\Delta}{H - F}. \tag{A.11}$$

Since admissibility makes Δ locally Lipschitz, (A.11) makes both rates continuous. Hence every *mixed* point, at which both rates are positive, has an open mixed neighborhood. Since $\Delta \neq 0$, (2.3) gives

$$V_x = \frac{rV - \lambda_B(F - V)}{\Delta} \quad \text{almost everywhere.}$$

The right-hand side is continuous. Local absolute continuity of V and the fundamental theorem of calculus therefore imply $V \in C^1$, with this equation holding pointwise on each nonstationary component.

On a mixed neighborhood, both nonzero extremes maximize the linear Hamiltonian, so every point on their connecting edge does as well. In particular, the common-rate test (λ_S, λ_S) is optimal. Its Hamiltonian is $\lambda_S(\beta G - \Pi)$, so the Bellman equation gives

$$\Pi = \Pi_S, \tag{A.12}$$

and, since $\Pi_x = (\Pi_S)_x$ there, the equality of marginal returns becomes

$$\Gamma(\beta, V) = 0, \qquad \Gamma(\beta, V) := \beta(rG + k)(V - F) - (1 - \beta)k(H - V). \tag{A.13}$$

Equivalently,

$$V(\beta) = \frac{\beta(rG + k)F + (1 - \beta)kH}{rG\beta + k}. \tag{A.14}$$

Since $\Gamma_\beta = rG(V - F) + k(H - F) > 0$, $\Gamma_V = rG\beta + k > 0$, $\Gamma(\beta^\circ, V_S) = 0$, the mixed locus satisfies

$$V > V_S \iff \beta < \beta^\circ, \qquad V < V_S \iff \beta > \beta^\circ. \tag{A.15}$$

Differentiating $\Gamma = 0$ and using (2.3) gives $V_x = -\frac{(H-V)(V-F)}{H-F}$ and

$$\Delta = -\frac{[r(H-F)+c](V-V_S)}{2(H-V)(V-F)}. \tag{A.16}$$

Therefore mixed pass-learning points lie below β°, and mixed fail-learning points lie above it.

At nonmixed points, (A.10) selects the corresponding pure region. In a fail-only region, the active test and value equations are

$$\lambda_G(\beta) = 0, \qquad \lambda_B(\beta) = \frac{c}{V(\beta) - F}, \tag{A.17}$$

and

$$V_x = \frac{(V-F)(rV+c)}{c}, \qquad \Pi_x = \frac{(V-F)(r\Pi+k)}{c} + (1-\beta)\Pi. \tag{A.18}$$

In a pass-only region, they are

$$\lambda_G(\beta) = \frac{c}{H - V(\beta)}, \qquad \lambda_B(\beta) = 0, \tag{A.19}$$

and

$$V_x = -\frac{(H-V)rV}{c}, \qquad \Pi_x = -\frac{(H-V)(r\Pi+k)}{c} + \beta(G-\Pi). \tag{A.20}$$

Define the Hamiltonians at the pass-only and fail-only nonzero extremes by

$$\Phi_P(\beta; V, \Pi) := \Psi_P\left(\beta, \frac{c}{H - V(\beta)}, 0; \Pi\right), \quad \Phi_F(\beta; V, \Pi) := \Psi_P\left(\beta, 0, \frac{c}{V(\beta) - F}; \Pi\right).$$

Suppressing their arguments, direct substitution yields

$$\Phi_P - \Phi_F = \frac{(H-F)(r+\lambda_S)}{V-F}(\Pi - \Pi_S) \qquad \text{in a pass-only region,} \tag{A.21}$$

$$\Phi_F - \Phi_P = \frac{(H-F)(r+\lambda_S)}{H-V}(\Pi - \Pi_S) \qquad \text{in a fail-only region.} \tag{A.22}$$

Step 2: Every nonstationary component attaches to a stationary belief. A component *attaches* to the endpoint toward which its no-news path moves; we call that endpoint its attachment. For a pass-learning component, the no-news direction points to its lower endpoint. At pure points, $-\Delta = \frac{c}{H-V} > \lambda_S$. At mixed points, a failure of $-\Delta$ to be bounded away from zero would, by (A.14) and (A.16), force $(\beta, V) \to (\beta^\circ, V_S)$. This cannot occur at a lower endpoint equal to β°: mixed pass-learning points satisfy $\beta < \beta^\circ$, whereas a pass-learning component approaches its lower endpoint through beliefs strictly to its right. Thus $-\Delta$ is bounded away from zero near every admissible interior lower endpoint, which is reached in finite time. It cannot be a firing endpoint by Lemma 3.2. At a hiring endpoint, value matching gives $V \to H$; then $\Gamma(\beta, H) > 0$

excludes mixed points, while a pure region has $\lambda_G = c/(H-V) \to \infty$, violating admissibility. Finally, the endpoint cannot be zero. To see this, note first that mixed points cannot accumulate there: (A.14) gives $V \to H$, and (A.16) then gives $|\Delta| \to \infty$. Thus a sufficiently far left tail is pure pass-only. With $R := V/(H-V)$, (A.20) gives $R_x = -\frac{rH}{c}R$, $R(x) = R(x_0)e^{-rH(x-x_0)/c}$. A positive solution has $V \to H$ as $x \to -\infty$ and hence an unbounded pass rate; a negative solution reaches $V = F$ at finite x. Hence the only solution extending to $x \to -\infty$ with bounded rates is $V \equiv 0$. In that case, with $W := \Pi + k/r$,

$$W_x = -\left(\frac{Hr}{c} + \beta\right) W + \beta\left(G + \frac{k}{r}\right).$$

Boundedness and variation of constants give $W \to 0$ as $x \to -\infty$, so $\Pi \to -k/r < 0$, contrary to the payoff guarantees. Thus the lower endpoint is stationary.

For a fail-learning component, the no-news direction points to its upper endpoint and, at pure points, $\Delta = \frac{c}{V-F} > \lambda_S$. If Δ were not bounded away from zero along mixed points, the mixed locus would again force $(\beta, V) \to (\beta^\circ, V_S)$. But mixed fail-learning points satisfy $\beta > \beta^\circ$, whereas a fail-learning component approaches its upper endpoint through beliefs strictly to its left. Hence mixed fail points cannot accumulate at an upper endpoint equal to β°, and every admissible interior upper endpoint is reached in finite time. Hiring is excluded by Lemma 3.2; firing would give $V \to F$, which excludes mixed points and makes $\lambda_B = c/(V-F)$ unbounded. Finally, the endpoint cannot be one. To see this, note first that mixed points cannot accumulate there: (A.14) gives $V \to F$, and (A.16) gives $|\Delta| \to \infty$. Thus a sufficiently far right tail is pure fail-only. With $Y := V - F$, (A.18) gives $Y_x = \frac{Y(c+rF+rY)}{c}$. The scalar phase portrait shows that a solution extending to $x \to \infty$ while remaining in $F < V < H$ can only have $V \to F$ or, when $c + rF < 0$, $V \to -c/r$. The former makes $\lambda_B = c/(V-F)$ unbounded. In the latter case, with $W := \Pi + k/r$, $W_x = \left[\frac{r(V-F)}{c} + 1 - \beta\right] W - (1-\beta)\frac{k}{r}$. The coefficient converges to $-(c+rF)/c > 0$, so boundedness and variation of constants give $W \to 0$. Hence $\Pi \to -k/r < h(1)$, contrary to the payoff guarantees. Thus the upper endpoint is stationary.

Step 3: Stationary feasibility. At a stationary belief, $V = V_S$, $\Pi = \Pi_S(\beta)$. The payoff guarantees are equivalent to $\beta \in [\beta_{S,0}, \beta_{S,h}]$, and

$$\beta_{S,h} - \beta_{S,0} = \frac{[r(H-F)+c](-B)[c - c_N(k)]}{c\{rG(H-F) - B[r(H-F)+c]\}}. \tag{A.23}$$

Thus the set of feasible stationary beliefs is empty if $c < c_N(k)$, a singleton $\{\beta^*\}$ if $c = c_N(k)$, and a nondegenerate interval if $c > c_N(k)$.

If $c < c_N(k)$, Step 2 excludes nonstationary probation, so $\Pi = \max\{0, h\}$. At equality, no component can attach to β^*. Mixed points are excluded by $\Gamma(\beta^*, V_S) \neq 0$, or, when $\beta^* = \beta^\circ$, by the mixed-point side restrictions from Step 1. A pure pass region to the right would satisfy $(\Pi - h)_x(\beta^*+) = -\frac{(H-V_S)k}{c} < 0$, and a pure fail region to the

left would satisfy $\Pi_x(\beta^*-) = \frac{(V_S-F)k}{c} > 0$; both violate the payoff guarantees. Hence only the isolated stationary convention, payoff-equivalent for the principal, remains.

Suppose $c > c_N(k)$. Then $0 < \beta_{S,0} < \beta^* < \beta_{S,h} < 1$, $\Pi_S > \max\{0, h\}$ on $(\beta_{S,0}, \beta_{S,h})$. In particular, probation is required at β^*, so Step 2 implies the existence of a stationary belief.

Step 4: Global region configuration. Let s be the stationary attachment of a component, and set $Q := \Pi - \Pi_S$. Mixed points cannot approach s: continuity would force $s = \beta^\circ$, but a mixed pass component lies to the right of s and must be below β°; the fail case is symmetric. Thus each region is pure near its attachment.

Direct substitution gives, at every pure point,

$$Q_x = \frac{\Gamma(\beta, V)}{r(H-F)+c} - \left[\frac{r(H-V)}{c} + \beta\right] Q \qquad \text{in a pass-only region,} \tag{A.24}$$

$$Q_x = \frac{\Gamma(\beta, V)}{r(H-F)+c} + \left[\frac{r(V-F)}{c} + 1 - \beta\right] Q \qquad \text{in a fail-only region.} \tag{A.25}$$

At the stationary attachment, $Q = 0$, so the sign of the limiting Q_x is the sign of $\Gamma(s, V_S)$, hence the sign of $s - \beta^\circ$. Since $Q \geq 0$, a pass region requires $s \geq \beta^\circ$, and a fail region requires $s \leq \beta^\circ$. The mixed-point side restrictions in Step 1 then exclude mixed points on the entire component.

In the resulting pure pass region, $V_S < V < 0$, so $V_x > 0$. In the pure fail region, $V > \max\{F, -c/r\}$ and $V_x > 0$: the initial value satisfies $V_S > -c/r$, and uniqueness prevents crossing a stationary solution of the scalar value equation. Consequently,

$$\frac{d}{dx}\Gamma(\beta(x), V(x)) = \Gamma_\beta \beta(1-\beta) + \Gamma_V V_x > 0.$$

Apply the integrating-factor formula to (A.24)–(A.25), starting from $Q(s) = 0$. In a pass region, the forcing term is strictly positive to the right of the attachment. In a fail region it is strictly negative to the left, and the integration direction is reversed. Thus, in both cases,

$$\Pi > \Pi_S \qquad \text{away from the stationary attachment.} \tag{A.26}$$

The same integral remains strictly positive at any finite opposite endpoint: its coefficients are bounded there because $V \in [F, H]$, and its integrand has a strict sign on a nonempty interval. This argument also covers $s = \beta^\circ$, without a separate second-derivative calculation.

There is only one stationary belief. If two existed, choose two consecutive ones $s_1 < s_2$. Since both lie in the feasible stationary interval, $\Pi_S > \max\{0, h\}$ throughout (s_1, s_2). The payoff guarantees therefore force every intermediate belief to be on probation. By consecutiveness, none is stationary, so (s_1, s_2) is a nonstationary component with $Q = 0$

at both endpoints. This contradicts the previous paragraph, which shows that $Q > 0$ at the endpoint opposite the attachment.

Let s denote the unique stationary belief. Every nonstationary component attaches to s, so there is at most one component on each side of s: a pass-only region to its right and a fail-only region to its left. If only the pass-only region exists, it must attach at $s = \beta_{S,0}$. Indeed, stationary feasibility gives $s \geq \beta_{S,0}$; if the inequality were strict, then Π_S would strictly dominate stopping immediately to the left of s, forcing an additional fail-only region. Its attachment condition is therefore $\beta_{S,0} \geq \beta^\circ$. Symmetrically, if only the fail-only region exists, it must attach at $s = \beta_{S,h}$, with $\beta_{S,h} \leq \beta^\circ$. If both components exist, then $s = \beta^\circ$, $\beta_{S,0} < \beta^\circ < \beta_{S,h}$; equality at either endpoint would violate the payoff guarantees, by the same derivative argument used in Step 3.

The comparisons are

$$\beta_{S,0} - \beta^\circ = \frac{k[r(H-F)+c]}{c\{k(Hr+c) - rF(Gr+k)\}}[c_P(k) - c], \tag{A.27}$$

$$\beta^\circ - \beta_{S,h} = \frac{(H-F)(Gr+k)[r(H-F)+c]}{\{k(Hr+c) - rF(Gr+k)\}\{rG(H-F) - B[r(H-F)+c]\}}[k - k_F]. \tag{A.28}$$

Their prefactors are positive. Finally,

$$c_P(k) - c_N(k) = \frac{(H-F)(k-k_F)}{B}. \tag{A.29}$$

Since $B < 0$, $c_N(k) < c_P(k)$ exactly when $k < k_F$. Combined with Step 3, these signs give the four regions in the theorem. At $c = c_P(k)$ the fail region has zero length; at $k = k_F$ the pass region has zero length.

Step 5: Construction and uniqueness. The region-specific systems are locally Lipschitz on $F < V < H$; the stationary attachment and $V = V_S, \Pi = \Pi_S$ therefore determine the solution in each region uniquely. The strict inequality (A.26) and the gap formulas (A.21)–(A.22) verify the selected nonzero extreme relative to the other nonzero extreme.

The solution in a pass-only region reaches a unique hiring threshold. Continuing this solution, $V \uparrow 0$ and $\Pi(\beta) \longrightarrow \frac{cG-Hk}{c+Hr} < G\,(\beta \uparrow 1)$. Since $\Pi - h > 0$ at the attachment, a crossing exists. At any crossing $\Pi = h$, $h = \Pi > \Pi_S \geq 0$, and $(\Pi - h)_x = -\frac{(H-V)(rh+k)}{c} < 0$. Thus every zero is crossed from positive to negative, so the crossing is unique. Since $\Pi > \Pi_S \geq h$ for $\beta \leq \beta_{S,h}$, it lies above $\beta_{S,h}$.

The solution in a fail-only region reaches a unique firing threshold. Every regime requiring such a region has $c > -rF$. Indeed, in the fail–pass region $c > c_P(k) > -rF$, while in the fail-only region $c > c_N(k) \geq c_N(k_F) = -rF\frac{G-B}{G} > -rF$. With $Y := V - F$ and $Z := \Pi/Y$, $Y_x = \frac{Y(c+rF+rY)}{c}$, $Z_x = \frac{k}{c} - \left(\beta + \frac{rF}{c}\right) Z$.

As $x \to -\infty$,

$$Y \sim C_0 e^{(c+rF)x/c} \quad \text{for some } C_0 > 0, \qquad \beta = o(Y), \qquad Z \longrightarrow \frac{k}{rF} < 0,$$

where the last limit follows from variation of constants in the linear Z-equation. Thus $\Pi < 0$ near zero, while $\Pi > 0$ at the attachment, so a crossing exists. At every zero, $\Pi_x = \frac{(V-F)k}{c} > 0$. Thus every zero is crossed from negative to positive, so the crossing is unique. Since $\Pi > \Pi_S \geq 0$ for $\beta \geq \beta_{S,0}$, it lies below $\beta_{S,0}$.

For $c \leq c_N(k)$, stop immediately, with either stopping action at β^* and the optional isolated stationary convention at equality. For $c > c_N(k)$, attach the regions identified in Step 4, using $s = \beta_{S,h}$, $s = \beta_{S,0}$, or $s = \beta^\circ$ in the fail-only, pass-only, or fail–pass case, respectively. Terminate each outer region at its unique stopping crossing. A region of zero length at a regime boundary is simply omitted.

The probation set consists of at most two finite open intervals and one stationary point. On each interval the pure rate is smooth and bounded, and the drift has constant nonzero sign, so the assessment is admissible. Both values match (V_S, Π_S) at the stationary no-news attachment. At the other endpoint the region ends at the unique principal-indifference crossing, which makes Π continuous; no matching of V is required there because the no-news path moves away from that endpoint.

The constructed functions are bounded, satisfy the Bellman equations, and match the prescribed values at every endpoint reached in finite no-news time. Lemma A.1 therefore identifies them with the induced continuation values.

When $c \leq c_N(k)$, stationary feasibility gives $\Pi_S \leq \max\{0, h\}$, so immediate stopping satisfies the payoff guarantees. On a pass-only side, $\Pi > \Pi_S \geq 0$ and $\Pi > h$ up to the unique hiring crossing; above it, $h \geq \max\{0, \Pi_S\}$ because the crossing lies above $\beta_{S,h}$. In a sole pass equilibrium, below $s = \beta_{S,0}$ both h and Π_S are nonpositive. On a fail-only side, the firing crossing lies below $\beta_{S,0}$; below it both h and Π_S are negative, while between it and the attachment $\Pi > 0$, and whenever $h \geq 0$ one has $\Pi > \Pi_S \geq h$. Above a sole fail attachment, $h \geq \max\{0, \Pi_S\}$. The two-sided case combines the two one-sided arguments. Thus the payoff guarantees hold globally.

In every nonstationary region, the value functions are differentiable and the Bellman equations hold. The binding incentive constraint, the strict inequality (A.26), and the gap formulas select the prescribed nonzero extreme; its Hamiltonian is $r\Pi + k > 0$, whereas the zero test has Hamiltonian zero. Thus the prescribed test is the pointwise maximizer over the full incentive-compatible triangle. The stationary test is prescribed by construction. Finally, the bad type's Bellman equation, binding incentive constraint, and stationary equality permit application of Lemma A.2. Hence the constructed assessments satisfy Definition 4.1, and the selected shirking behavior is optimal against every progressively measurable effort deviation.

Steps 2–4 show that the closure of each probation component has two endpoints

strictly inside $(0,1)$. Since there are finitely many such components and stationary points, their union has positive distance from $\{0,1\}$; hence choose $\varepsilon > 0$ accordingly. Local uniqueness of the ODEs and the crossing derivatives then give uniqueness, modulo payoff-irrelevant choices at stopping indifference beliefs, off-path tests, and the isolated stationary convention at $c = c_N(k)$. □

A.4 Proofs of results in Section 5

Proof of Proposition 5.1. Let W denote the principal's value in this benchmark, and let $m(\beta) := \max\{0, h(\beta)\}$ be the stopping payoff. We apply Che and Mierendorff [2019, Theorem 1], using the payoff and flow-cost normalization in Section 5 and identifying their states with the good and bad types and their terminal actions with hiring and firing. Hiring pays G in the good state and B in the bad state, whereas firing pays zero in both states. The two sources provide conclusive pass and fail news, respectively, with attention capacity $\bar{\lambda}$.

Step 1: Local test choice. In a continuation region with smooth values, the test's contribution to the principal's Bellman equation is

$$\beta\lambda_G(G - W) - (1-\beta)\lambda_B W + \beta(1-\beta)(\lambda_B - \lambda_G)W'.$$

Its maximum equals $rW + k > 0$, since $W \geq m \geq 0$. A maximizing test is therefore nonzero and exhausts the capacity: otherwise, scaling up both rates would strictly increase this expression. Write $\lambda_G = a\bar{\lambda}$ and $\lambda_B = (1-a)\bar{\lambda}$, where $a \in [0,1]$. The variational inequality is

$$\begin{aligned} 0 = \max\Big\{ & m(\beta) - W(\beta), \\ & -k - rW(\beta) + \max_{a\in[0,1]} \bar{\lambda}\big[a\beta\{G - W(\beta)\} - (1-a)(1-\beta)W(\beta) \\ & \qquad + (1-2a)\beta(1-\beta)W'(\beta)\big]\Big\}. \end{aligned} \tag{A.30}$$

At nonsmooth beliefs, this equation is interpreted in the viscosity sense. The Hamiltonian is affine in a, so a pure test maximizes whenever its two endpoint values differ. The endpoints $a = 1$ and $a = 0$ are pass-only and fail-only monitoring, with rates $(\bar{\lambda}, 0)$ and $(0, \bar{\lambda})$, respectively. In the absence of news, $\dot{\beta} = (1-2a)\bar{\lambda}\beta(1-\beta)$. Thus pass-only monitoring lowers the posterior, whereas fail-only monitoring raises it.

Step 2: Stopping. At the stopping payoff $W = m$, the maximal continuation gain is

$$\begin{cases} \bar{\lambda}\beta G - k, & \beta < \beta^*, \\ -\bar{\lambda}(1-\beta)B - rh(\beta) - k, & \beta > \beta^*. \end{cases}$$

The first expression is increasing and the second is decreasing in β. Their common limit at β^* is $\bar{\lambda}\beta^* G - k$, which is also the maximal gain at β^*. Hence the upper probation-cost threshold is

$$\overline{k}^A := \bar{\lambda}\beta^* G = \frac{\bar{\lambda} G(-B)}{G - B}. \tag{A.31}$$

For $k \geq \overline{k}^A$, the stopping payoff satisfies (A.30), and the verification argument in the cited theorem gives immediate stopping. For $k < \overline{k}^A$, continuing briefly at β^* is strictly profitable.

For $0 < k < \overline{k}^A$, the same theorem gives a continuation interval $(\underline{\beta}^A, \overline{\beta}^A)$, with pass-only monitoring near its lower endpoint and fail-only monitoring near its upper endpoint. Value matching and smooth pasting give

$$\begin{aligned} W(\underline{\beta}^A) &= 0, & W'(\underline{\beta}^A+) &= 0, \\ W(\overline{\beta}^A) &= h(\overline{\beta}^A), & W'(\overline{\beta}^A-) &= G - B. \end{aligned}$$

Evaluating the corresponding region-specific equations at the endpoints yields $k = \bar{\lambda}\underline{\beta}^A G$, $k + rh(\overline{\beta}^A) = -\bar{\lambda}(1 - \overline{\beta}^A)B$. Consequently,

$$\underline{\beta}^A = \frac{k}{\bar{\lambda} G}, \qquad \overline{\beta}^A = \frac{-k - (r + \bar{\lambda})B}{r(G - B) - \bar{\lambda} B}. \tag{A.32}$$

For $0 < k < \overline{k}^A$, these boundaries satisfy $0 < \underline{\beta}^A < \beta^* < \overline{\beta}^A < 1$.

Step 3: Interior switches. A stationary continuation policy uses $\lambda_G = \lambda_B = \bar{\lambda}/2$. Its payoff is $W_S^A(\beta) := \frac{\bar{\lambda}\beta G - 2k}{2r + \bar{\lambda}}$. In the low-cost configuration of the cited theorem, the value functions in the two inner regions attach smoothly to this stationary payoff. At their attachment, $W = W_S^A$ and $W' = (W_S^A)'$, and the two pure-test Hamiltonians coincide. Their difference, evaluated at W_S^A, is

$$\bar{\lambda}\Big[\beta G + (1 - 2\beta)W_S^A - 2\beta(1 - \beta)(W_S^A)'\Big] = \frac{2\bar{\lambda}}{2r + \bar{\lambda}}\Big[\beta(rG + 2k) - k\Big].$$

The stationary attachment is therefore

$$\beta_o^A = \frac{k}{rG + 2k}. \tag{A.33}$$

The lower probation-cost threshold in the cited theorem specializes to

$$\begin{aligned} D_A &:= 1 + \left(\frac{2r + \bar{\lambda}}{\bar{\lambda}}\right)^{\bar{\lambda}/r}, \\ \underline{k}^A &:= \max\left\{0, \min\left\{\overline{k}^A, \frac{(r + \bar{\lambda})G}{D_A} - rG, \frac{(r + \bar{\lambda})(-B)}{D_A}\right\}\right\}. \end{aligned} \tag{A.34}$$

In particular, $0 \leq \underline{k}^A \leq \overline{k}^A$.

If $\underline{k}^A \leq k < \overline{k}^A$, the theorem's crossing argument gives a switching belief $\check{\beta}^A \in (\underline{\beta}^A, \overline{\beta}^A)$, with pass-only monitoring below it and fail-only monitoring above it. By the drift formula in Step 1, no news moves the posterior toward the lower stopping boundary in the pass-only region and toward the upper stopping boundary in the fail-only region. This proves (ii).

If $0 < k < \underline{k}^A$, the same argument gives cutoffs $\underline{\beta}^A < \beta_-^A < \beta_o^A < \beta_+^A < \overline{\beta}^A$. The two outer regions use pass-only and fail-only monitoring, respectively, whereas the two inner regions use fail-only and pass-only monitoring, respectively, and attach to the stationary policy at β_o^A. Thus no news moves outward in the outer regions and inward in the inner regions, as claimed in (iii).

At each repelling switch, choose an optimal outgoing no-news trajectory, as allowed in Section 5. The cited verification argument establishes optimality over the benchmark's history-adapted attention controls and stopping rules. Together with Step 2, this proves all three cases. □

A.5 Proofs of results in Section 6

Proof of Theorem 6.1. Define

$$\Pi^{\mathsf{P}}(\beta) := \beta \frac{cG - kH}{c + rH} - (1 - \beta)\frac{k}{r}, \tag{A.35}$$

$$\Pi^{\mathsf{M}}(\beta) := \beta \frac{G(c + rF) - k(H - F)}{c + rH}. \tag{A.36}$$

The first expression is the payoff from policy (iv). The second is an auxiliary coefficient used to bound the payoff from screening.

When $c + rF > 0$, define

$$\lambda_G^{\bar{\lambda}} := \frac{cr + \bar{\lambda}(c + rF)}{rH + \bar{\lambda}(H - F)} \tag{A.37}$$

and let

$$\Pi_{\bar{\lambda}}^{\mathsf{M}}(\beta) := \beta \frac{\lambda_G^{\bar{\lambda}} G - k}{r + \lambda_G^{\bar{\lambda}}} - (1 - \beta)\frac{k}{r + \bar{\lambda}} \tag{A.38}$$

be the payoff from policy (iii). We prove that the principal's optimal commitment payoff is

$$\Pi_{\bar{\lambda}}^{C}(\beta) = \begin{cases} \max\left\{0,\, h(\beta),\, \Pi^{\mathsf{P}}(\beta),\, \Pi_{\bar{\lambda}}^{\mathsf{M}}(\beta)\right\}, & c + rF > 0, \\ \max\left\{0,\, h(\beta),\, \Pi^{\mathsf{P}}(\beta)\right\}, & c + rF \leq 0. \end{cases} \tag{A.39}$$

Fix an arbitrary feasible commitment policy and a best response of the bad type. For each $\theta \in \{G, B\}$, let $\tau_\theta \in [0, \infty]$ be the terminal time, where $\tau_\theta = \infty$ if probation

never ends. When $\tau_\theta < \infty$, let $a_\theta \in \{\mathsf{h}, \mathsf{f}\}$ denote the terminal action. Define

$$X_\theta := \mathbb{E}_\theta \left[e^{-r\tau_\theta} \mathbf{1}_{\{a_\theta = \mathsf{h}\}} \right], \tag{A.40}$$

$$Z_\theta := \mathbb{E}_\theta \left[e^{-r\tau_\theta} \mathbf{1}_{\{a_\theta = \mathsf{f}\}} \right], \tag{A.41}$$

$$Q_\theta := \mathbb{E}_\theta \left[\int_0^{\tau_\theta} e^{-rt}\, dt \right]. \tag{A.42}$$

We adopt the convention $e^{-r\infty} = 0$. Since every finite terminal history ends in either hiring or firing,

$$X_\theta + Z_\theta + rQ_\theta = 1, \qquad \theta \in \{G, B\}. \tag{A.43}$$

Monotonicity and the implication of the rate bound. The discounted terminal allocations satisfy

$$X_G \geq X_B, \qquad Z_B \geq Z_G. \tag{A.44}$$

Indeed, couple the two types along their common no-news history. Construct pass arrivals from a common Poisson random measure and obtain the bad type's pass process by predictable thinning with retention probability equal to his effort. Thus every pass of the bad type is also a pass of the good type. Add a fail process only for the bad type. Under this coupling, whenever the bad type is hired, the good type is hired weakly earlier, and whenever the good type is fired, the bad type is fired weakly earlier. This gives (A.44) after discounting and taking expectations.

The bound on test intensities further implies

$$Z_B - Z_G \leq \bar{\lambda} Q_B. \tag{A.45}$$

To verify this inequality, condition on all public and private randomization. Along the resulting no-news path, let T be the time at which the principal stops in the absence of a signal, and let $A \in \{\mathsf{h}, \mathsf{f}\}$ be the corresponding action when $T < \infty$. Write g_t, b_t, and e_t for the pass intensity, fail intensity, and bad type's effort along this path. Define the discounted no-signal probabilities

$$D_G(t) := \exp\left\{ -rt - \int_0^t g_s\, ds \right\}, \qquad D_B(t) := \exp\left\{ -rt - \int_0^t \Big[e_s g_s + (1 - e_s) b_s \Big]\, ds \right\}.$$

Conditional on the fixed randomization, $Q_B = \int_0^T D_B(t)\, dt$ and

$$Z_B - Z_G = \int_0^T D_B(t)(1 - e_t) b_t\, dt + \mathbf{1}_{\{A = \mathsf{f}\}} \Big[D_B(T) - D_G(T) \Big]. \tag{A.46}$$

When $T = \infty$, the terminal term is zero and the desired rate bound follows directly from $(1 - e_t) b_t \leq \bar{\lambda}$. For $T < \infty$, let $\Delta := \int_0^T (1 - e_t)(g_t - b_t)\, dt$. Since $D_G(T) = D_B(T) e^{-\Delta}$,

the inequality $1 - e^{-z} \leq z_+$, where $z_+ := \max\{z, 0\}$, and the monotonicity of D_B imply

$$\begin{aligned} D_B(T) - D_G(T) &\leq D_B(T)\Delta_+ \\ &\leq \int_0^T D_B(t)(1 - e_t)(g_t - b_t)_+ \, dt. \end{aligned}$$

Substituting this bound into (A.46) gives

$$\begin{aligned} Z_B - Z_G &\leq \int_0^T D_B(t)(1 - e_t) \left[b_t + (g_t - b_t)_+\right] dt \\ &= \int_0^T D_B(t)(1 - e_t) \max\{g_t, b_t\} \, dt \\ &\leq \bar{\lambda} \int_0^T D_B(t) \, dt = \bar{\lambda} Q_B. \end{aligned}$$

Averaging over the conditioned randomization proves (A.45).

The full-imitation constraint. By exerting high effort forever, the bad type can generate the good type's public history distribution and receive $HX_G + FZ_G - cQ_G$. His actual effort costs are nonnegative. A necessary incentive constraint is therefore

$$HX_B + FZ_B \geq HX_G + FZ_G - cQ_G. \tag{A.47}$$

Set $x := X_B$, $y := rQ_B$, $u := X_G - X_B$, $w := Z_B - Z_G$. By (A.44), $x, y, u, w \geq 0$. The accounting identities imply $Z_B = 1 - x - y$, $Z_G = 1 - x - y - w$, $rQ_G = y + w - u$. Consequently,

$$x + y + w = 1 - Z_G \leq 1 \tag{A.48}$$

and, by (A.45),

$$w \leq \frac{\bar{\lambda}}{r} y. \tag{A.49}$$

Substituting the accounting identities into (A.47) yields

$$(c + rH)u \leq cy + (c + rF)w. \tag{A.50}$$

Upper bound. The principal's payoff under the arbitrary commitment is

$$\begin{aligned} \Pi &= \beta\Big(GX_G - kQ_G\Big) + (1 - \beta)\Big(BX_B - kQ_B\Big) \\ &= h(\beta)x + \beta\left(G + \frac{k}{r}\right)u - \frac{k}{r}y - \beta\frac{k}{r}w. \end{aligned} \tag{A.51}$$

Because $\beta(G + k/r) \geq 0$, inequality (A.50) gives

$$\Pi \leq h(\beta)x + \Pi^{\mathsf{P}}(\beta)y + \Pi^{\mathsf{M}}(\beta)w. \tag{A.52}$$

Suppose first that $c + rF \leq 0$. Then $\Pi^{\mathsf{M}}(\beta) = \beta\frac{G(c+rF)-k(H-F)}{c+rH} \leq 0$. Since $w \geq 0$, (A.52) and $x + y \leq 1$ imply $\Pi \leq \max\left\{0,\, h(\beta),\, \Pi^{\mathsf{P}}(\beta)\right\}$.

Now suppose that $c + rF > 0$. A direct substitution of (A.37) into (A.38) gives

$$\Pi^{\mathsf{M}}_{\bar{\lambda}}(\beta) = \frac{r}{r+\bar{\lambda}}\Pi^{\mathsf{P}}(\beta) + \frac{\bar{\lambda}}{r+\bar{\lambda}}\Pi^{\mathsf{M}}(\beta). \tag{A.53}$$

Indeed, $r + \lambda_G^{\bar{\lambda}} = \frac{(r+\bar{\lambda})(c+rH)}{rH+\bar{\lambda}(H-F)}$, from which (A.53) follows after collecting terms.

Using (A.53), the right-hand side of (A.52) can be written as

$$h(\beta)x + \Pi^{\mathsf{P}}(\beta)\left(y - \frac{r}{\bar{\lambda}}w\right) + \Pi^{\mathsf{M}}_{\bar{\lambda}}(\beta)\frac{r+\bar{\lambda}}{\bar{\lambda}}w.$$

The three displayed coefficients are nonnegative by (A.49), and their sum is $x+y+w \leq 1$. Therefore,

$$\Pi \leq \max\left\{0,\, h(\beta),\, \Pi^{\mathsf{P}}(\beta),\, \Pi^{\mathsf{M}}_{\bar{\lambda}}(\beta)\right\}.$$

This establishes the upper bound in (A.39).

Attainability. Immediate firing and immediate hiring attain 0 and $h(\beta)$, respectively.

For perpetual pass-only monitoring, set $\lambda_G = c/H$ and $\lambda_B = 0$, which is feasible because $\bar{\lambda} \geq c/H$. The constant bad-type value zero satisfies the controlled Bellman equation with $\Gamma_W = \lambda_G H - c = 0$. The verification argument of Lemma A.2 therefore makes every effort process optimal. Under the maintained shirking selection, the bad type never exits and the good type is hired at rate c/H, giving

$$\beta\frac{(c/H)G - k}{r + c/H} - (1-\beta)\frac{k}{r} = \Pi^{\mathsf{P}}(\beta).$$

When $c + rF > 0$, take $\lambda_B = \bar{\lambda}$ and the pass rate $\lambda_G^{\bar{\lambda}}$ in (A.37). Its numerator is positive and

$$\frac{c}{H} - \lambda_G^{\bar{\lambda}} = \frac{-\bar{\lambda}F(c+rH)}{H[rH + \bar{\lambda}(H-F)]} > 0,$$

so both rates satisfy the bound. With $V_B^{\bar{\lambda}} = \bar{\lambda}F/(r+\bar{\lambda})$, the rate formula gives

$$rV_B^{\bar{\lambda}} = \bar{\lambda}(F - V_B^{\bar{\lambda}}), \qquad \lambda_G^{\bar{\lambda}}(H - V_B^{\bar{\lambda}}) + \bar{\lambda}(V_B^{\bar{\lambda}} - F) = c.$$

Thus the same verification argument again makes every effort process optimal. Selecting shirking yields

$$\beta\frac{\lambda_G^{\bar{\lambda}}G - k}{r + \lambda_G^{\bar{\lambda}}} - (1-\beta)\frac{k}{r+\bar{\lambda}} = \Pi^{\mathsf{M}}_{\bar{\lambda}}(\beta).$$

All terms in (A.39) are therefore attainable, which completes the proof. □

Proof of Proposition 7.1. See the Online Appendix. □

Proof of Proposition 7.2. See the Online Appendix. □

References

Joseph G. Altonji and Charles R. Pierret. Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350, 2001. doi: 10.1162/003355301556329.

Vojtěch Bartoš, Michal Bauer, Julie Chytilová, and Filip Matějka. Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review*, 106(6):1437–1475, 2016. doi: 10.1257/aer.20140571.

Dhruva Bhaskar. Dynamic screening and the dual roles of monitoring. *American Economic Journal: Microeconomics*, 17(4):352–398, 2025. doi: 10.1257/mic.20230049.

Simon Board and Moritz Meyer-ter-Vehn. Reputation for quality. *Econometrica*, 81(6): 2381–2462, 2013. doi: 10.3982/ECTA9039.

Jing Cai, Sai Luo, and Shing-Yi Wang. Money or monitoring: Evidence on improving worker effort. NBER Working Paper 33977, National Bureau of Economic Research, June 2025.

Doruk Cetemen and Chiara Margaria. Dynamic signaling in Wald options. Working paper, 2025.

Yeon-Koo Che and Konrad Mierendorff. Optimal dynamic allocation of attention. *American Economic Review*, 109(8):2993–3029, 2019. doi: 10.1257/aer.20171000.

Liang Dai, Yenan Wang, and Ming Yang. Dynamic contracting with flexible monitoring. Working paper, December 2024.

Thomas S. Dee and James Wyckoff. Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, 34(2):267–297, 2015. doi: 10.1002/pam.21818.

Francesc Dilmé. Dynamic quality signaling with hidden actions. *Games and Economic Behavior*, 113:116–136, 2019. doi: 10.1016/j.geb.2018.07.010.

Francesc Dilmé. Repeated trade with imperfect information about previous transactions. *Theoretical Economics*, 20(1):209–254, 2025. doi: 10.3982/TE5694.

Michael Dinerstein and Isaac M. Opper. Screening with multitasking: Theory and empirical evidence from teacher tenure reform. NBER Working Paper 30310, National Bureau of Economic Research, July 2022.

Esther Duflo, Rema Hanna, and Stephen P. Ryan. Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4):1241–1278, 2012. doi: 10.1257/aer.102.4.1241.

Mehmet Ekmekci, Leandro Gorno, Lucas Maestri, Jian Sun, and Dong Wei. Learning from manipulable signals. *American Economic Review*, 112(12):3995–4040, 2022. doi: 10.1257/aer.20211158.

Eduardo Faingold and Yuliy Sannikov. Reputation in continuous-time games. *Econometrica*, 79(3):773–876, 2011. doi: 10.3982/ECTA7377.

Henry S. Farber and Robert Gibbons. Learning and wage dynamics. *The Quarterly Journal of Economics*, 111(4):1007–1047, 1996. doi: 10.2307/2946706.

George Georgiadis and Balázs Szentes. Optimal monitoring design. *Econometrica*, 88(5): 2075–2107, 2020. doi: 10.3982/ECTA16475.

Alkis Georgiadis-Harris. Preparing to act. *Econometrica*, forthcoming.

Andrea Ichino and Regina T. Riphahn. The effect of employment protection on worker effort: Absenteeism during and after probation. *Journal of the European Economic Association*, 3

(1):120–143, 2005. doi: 10.1162/1542476053295296.

Pavel Ilinov, Andrei Matveenko, Salil Sharma, Elias Tsakas, and Mark Voorneveld. Sequential search with flexible information. Working paper, 2025.

Brian A. Jacob. The effect of employment protection on teacher effort. *Journal of Labor Economics*, 31(4):727–761, 2013. doi: 10.1086/669942.

Brian A. Jacob and Steven D. Levitt. Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics*, 118(3):843–877, 2003. doi: 10.1162/00335530360698441.

Nicolas Klein and Sven Rady. Negatively correlated bandits. *The Review of Economic Studies*, 78(2):693–732, 2011. doi: 10.1093/restud/rdq025.

Aditya Kuvalekar and Elliot Lipnowski. Job insecurity. *American Economic Journal: Microeconomics*, 12(2):188–229, 2020. doi: 10.1257/mic.20190132.

Fabian Lange. The speed of employer learning. *Journal of Labor Economics*, 25(1):1–35, 2007. doi: 10.1086/508730.

Anqi Li and Ming Yang. Optimal incentive contract with endogenous monitoring technology. *Theoretical Economics*, 15(3):1135–1173, 2020. doi: 10.3982/TE3130.

Daniel S. Nagin, James B. Rebitzer, Seth Sanders, and Lowell J. Taylor. Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *American Economic Review*, 92(4):850–873, 2002. doi: 10.1257/00028280260344498.

Benjamin A. Olken. Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, 115(2):200–249, 2007. doi: 10.1086/517935.

Eduardo Perez-Richet and Vasiliki Skreta. Test design under falsification. *Econometrica*, 90 (3):1109–1142, 2022. doi: 10.3982/ECTA16346.

Tomasz Piskorski and Mark M. Westerfield. Optimal dynamic contracts with moral hazard and costly monitoring. *Journal of Economic Theory*, 166:242–281, 2016. doi: 10.1016/j.jet.2016.08.003.

Jonah E. Rockoff, Douglas O. Staiger, Thomas J. Kane, and Eric S. Taylor. Information and employee evaluation: Evidence from a randomized intervention in public schools. *American Economic Review*, 102(7):3184–3213, 2012. doi: 10.1257/aer.102.7.3184.

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. doi: 10.1093/restud/rdaa012.

Yu Fu Wong. Dynamic monitoring design. Working paper, 2023.

Weijie Zhong. Optimal dynamic information acquisition. *Econometrica*, 90(4):1537–1582, 2022. doi: 10.3982/ECTA17787.