Online Appendix for

Timing Decisions in Organizations: Communication and Authority in a Dynamic Environment

By Steven R. Grenadier, Andrey Malenko, and Nadya Malenko

American Economic Review

Section A introduces the formal definitions of equilibria in the dynamic communication game. Section B presents all supplementary proofs needed for the main results and contains additional analysis of the communication game, such as selection of equilibria and comparative statics. Section C presents the formal comparison between centralized decision-making, once-and-for-all delegation, and time-contingent delegation. Section D discusses the robustness of the results to alternative versions of the model.

A. Communication game: Equilibrium notion

This section presents the formal definition of the Perfect Bayesian equilibrium in Markov strategies of the dynamic communication game of Section III, as well as the definition of a stationary equilibrium for the case $\underline{\theta} = 0$.

This is a dynamic game with observed actions (messages and the exercise decision) and incomplete information (type θ of the agent). Heuristically, the timing of events over an infinitesimal time interval $[t, t + dt]$ prior to option exercise can be described as follows: (1) Nature determines the realization of X_t. (2) The agent sends message $m(t) \in M$ to the principal. (3) The principal decides whether to exercise the option or not. If the option is exercised, the principal obtains the payoff of $\theta X_t - I$, the agent obtains the payoff of $\theta X_t - I + b$, and the game ends. Otherwise, the game continues, and the nature draws $X_{t+dt} = X_t + dX_t$. Because the game ends when the principal exercises the option, we can only consider histories such that the option has not yet been exercised. Then, the history of the game at time t has two components: the sample path of the public state $X(t)$ and the history of messages of the agent: $\mathcal{H}_t = \{X(s), s \leq t; m(s), s < t\}$.

Definition A.1. *Strategies $m^* = \{m_t^*, t \geq 0\}$ and $e^* = \{e_t^*, t \geq 0\}$, beliefs μ^*, and a message space M constitute a **Perfect Bayesian equilibrium in Markov strategies (PBEM)** if:*

1. For every t, $\mathcal{H}_t$, $\theta \in \Theta$, and strategy m,

$$\begin{aligned} & \mathbb{E}\left[e^{-r\tau(e^*)}\left(\theta X\left(\tau\left(e^*\right)\right)-I+b\right)|\mathcal{H}_t,\theta,\mu^*\left(\cdot|\mathcal{H}_t\right),m^*,e^*\right] \\ \geq \ & \mathbb{E}\left[e^{-r\tau(e^*)}\left(\theta X\left(\tau\left(e^*\right)\right)-I+b\right)|\mathcal{H}_t,\theta,\mu^*\left(\cdot|\mathcal{H}_t\right),m,e^*\right]. \end{aligned} \tag{1}$$

2. For every t, $\mathcal{H}_t$, $m(t) \in M$, and strategy e,

$$\begin{aligned} & \mathbb{E}\left[e^{-r\tau(e^*)}\left(\theta X\left(\tau\left(e^*\right)\right)-I\right)|\mathcal{H}_t,\mu^*\left(\cdot|\mathcal{H}_t,m\left(t\right)\right),m^*,e^*\right] \\ \geq \ & \mathbb{E}\left[e^{-r\tau(e)}\left(\theta X\left(\tau\left(e\right)\right)-I\right)|\mathcal{H}_t,\mu^*\left(\cdot|\mathcal{H}_t,m\left(t\right)\right),m^*,e\right]. \end{aligned} \tag{2}$$

3. Bayes' rule is used to update beliefs $\mu^*(\theta|\mathcal{H}_t)$ to $\mu^*(\theta|\mathcal{H}_t, m(t))$ whenever possible: For every $\mathcal{H}_t$ and $m(t) \in M$, if there exists θ such that $m_t^*(\theta, \mathcal{H}_t) = m(t)$, then for all θ

$$\mu^*\left(\theta|\mathcal{H}_t,m\left(t\right)\right)=\frac{\mu^*(\theta|\mathcal{H}_t)\mathbf{1}\{m_t^*(\theta,\mathcal{H}_t)=m\left(t\right)\}}{\int_{\underline{\theta}}^1\mu^*(\tilde{\theta}|\mathcal{H}_t)\mathbf{1}\{m_t^*(\tilde{\theta},\mathcal{H}_t)=m\left(t\right)\}d\tilde{\theta}}, \tag{3}$$

where $\mu^*(\theta|\mathcal{H}_0) = \frac{1}{1-\underline{\theta}}$ for $\theta \in \Theta$ and $\mu^*(\theta|\mathcal{H}_0) = 0$ for $\theta \notin \Theta$.

4. For every t, $\mathcal{H}_t$, $\theta \in \Theta$, and $m(t) \in M$,

$$m_t^*\left(\theta,\mathcal{H}_t\right) = m^*\left(\theta,X\left(t\right),\mu^*\left(\cdot|\mathcal{H}_t\right)\right); \tag{4}$$

$$e_t^*\left(\mathcal{H}_t,m\left(t\right)\right) = e^*\left(X\left(t\right),\mu^*\left(\cdot|\mathcal{H}_t,m\left(t\right)\right)\right). \tag{5}$$

The first three conditions, given by (1)–(3), are requirements of the Perfect Bayesian equilibrium. Inequalities (1) require the equilibrium strategy m^* to be sequentially optimal for the agent for any possible history $\mathcal{H}_t$ and type realization θ. Similarly, inequalities (2) require equilibrium strategy e^* to be sequentially optimal for the principal. Equation (3) requires beliefs to be updated according to Bayes' rule. Finally, conditions (4)–(5) are requirements that the equilibrium strategies and the message space are Markov.

When we focus on the case $\underline{\theta} = 0$ in Section IV, we restrict attention to *stationary* equilibria, which are defined as follows.

Definition A.2. *Suppose $\underline{\theta} = 0$. An equilibrium (m^*, e^*, μ^*, M) is* ***stationary*** *if whenever* *posterior belief* $\mu^*(\cdot|\mathcal{H}_t)$ *is uniform over* $[0,\hat{\theta}]$ *for some* $\hat{\theta}\in(0,1)$*, then for all* $\theta\in[0,\hat{\theta}]$*:*

$$m^*(\theta, X(t), \mu^*(\cdot|\mathcal{H}_t)) = m^*\left(\frac{\theta}{\hat{\theta}}, \hat{\theta}X(t), \mu^*(\cdot|\mathcal{H}_0)\right), \tag{6}$$

$$e^*(X(t), \mu^*(\cdot|\mathcal{H}_t, m(t))) = e^*\left(\hat{\theta}X(t), \mu^*(\cdot|\mathcal{H}_0, m(t))\right). \tag{7}$$

Condition (6) means that the message of type $\theta\in[0,\hat{\theta}]$ when the public state is $X(t)$ and the posterior is uniform over $[0,\hat{\theta}]$ is the same as the message of type $\frac{\theta}{\hat{\theta}}$ when the public state is $\hat{\theta}X(t)$ and the posterior is uniform over $[0,1]$. Condition (7) means that the exercise strategy of the principal is the same when the public state is $X(t)$ and her belief is that θ is uniform over $[0,\hat{\theta}]$ as when the public state is $\hat{\theta}X(t)$ and her belief is that θ is uniform over $[0,1]$.

B. Proofs of lemmas and derivations of auxiliary results

Derivation of the optimal exercise policy for the call and put option. Let $V(X)$ be the value of the option to a risk-neutral player if the current value of $X(t)$ is X and the player perfectly knows θ. Because the player is risk-neutral, the expected return from holding the option over a small interval dt, $E\left[\frac{dV}{V}\right]$, must equal the riskless return rdt. By Itô's lemma,

$$dV(X(t)) = \left(V'(X(t))\,\alpha X(t) + \frac{1}{2}V''(X(t))\,\sigma^2 X(t)^2\right)dt + \sigma V'(X(t))\,dW(t),$$

and hence

$$\mathbb{E}\left[\frac{dV(X(t))}{V(X(t))}\right] = rdt \Leftrightarrow \frac{1}{V(X(t))}\left(V'(X(t))\,\alpha X(t) + \frac{1}{2}V''(X(t))\,\sigma^2 X(t)^2\right)dt = rdt,$$

which gives

$$rV = \alpha X\frac{\partial V}{\partial X} + \frac{1}{2}\sigma^2 X^2\frac{\partial^2 V}{\partial X^2}. \tag{8}$$

This is a second-order linear homogeneous ordinary differential equation. The general solution to this equation is $V(X) = D_1X^{\beta_1} + D_2X^{\beta_2}$, where D_1 and D_2 are the constants to be determined, and $\beta_1 < 0 < 1 < \beta_2$ are the roots of $\frac{1}{2}\sigma^2\beta(\beta-1)+\alpha\beta - r = 0$. We denote the negative root by $-\delta$ and the positive root by β. To find D_1, D_2, we use two boundary conditions. If exercise of the option occurs at trigger $\bar{X}$ and gives a payoff $p\left(\bar{X}\right)$, the first boundary condition is $V\left(\bar{X}\right) = p\left(\bar{X}\right)$.

For the call option, the second boundary condition is $\lim_{X\to 0} V(X) = 0$ because zero is an absorbing barrier for the geometric Brownian motion. Hence, $D_1 = 0$. In addition, if θ is known to the principal, then $p_{call}\left(\bar{X}\right)=\theta\bar{X}-I$. Hence,

$$V_{call}\left(X,\bar{X}\right)=\left(\frac{X}{\bar{X}}\right)^{\beta}\left(\theta\bar{X}-I\right). \tag{9}$$

Maximizing $V_{call}\left(X,\bar{X}\right)$ with respect to $\bar{X}$ to derive the optimal call option exercise policy of the principal gives $\bar{X}=\frac{\beta}{\beta-1}\frac{I}{\theta}$, i.e. (2). It also follows that the value of the option to the principal if the current value of $X\left(t\right)$ is X satisfies

$$V_{P}^{*}\left(X,\theta\right)=\begin{cases}\left(\frac{X}{X_{P}^{*}(\theta)}\right)^{\beta}\left(\theta X_{P}^{*}\left(\theta\right)-I\right), & \text{if } X\leq X_{P}^{*}\left(\theta\right)\\ \theta X-I, & \text{if } X>X_{P}^{*}\left(\theta\right).\end{cases}$$

Similarly, for the put option, the second boundary condition is $\lim_{X\to\infty}V\left(X\right)=0$, and hence $D_2=0$. Combining it with $V\left(\bar{X}\right)=p\left(\bar{X}\right)$ and using $p_{put}\left(\bar{X}\right)=\theta I-\bar{X}$, gives $V_{put}\left(X,\bar{X}\right)=\left(\frac{X}{\bar{X}}\right)^{-\delta}\left(\theta I-\bar{X}\right)$. Maximizing $V_{put}\left(X,\bar{X}\right)$ with respect to $\bar{X}$ to derive the optimal put option exercise policy of the principal gives $\bar{X}=\frac{\delta}{\delta+1}I\theta$. ■

We next prove two useful auxiliary results, which hold in any threshold-exercise PBEM. The first result shows that in any threshold-exercise PBEM, the option is exercised weakly later if the agent has less favorable information. The second auxiliary result is that it is without loss of generality to reduce the message space significantly. Specifically, for any threshold-exercise equilibrium, there exists an equilibrium with a binary message space $M=\{0,1\}$ and simple equilibrium strategies that implements the same exercise times and hence features the same payoffs of both players.

Lemma IA.1. *$\bar{X}\left(\theta_1\right)\geq\bar{X}\left(\theta_2\right)$ for any $\theta_1,\theta_2\in\Theta$ such that $\theta_2\geq\theta_1$.*

Lemma IA.2. *If there exists a threshold-exercise PBEM with thresholds $\bar{X}\left(\theta\right)$, then there exists an equivalent threshold-exercise PBEM with the binary message space $M=\{0,1\}$ and the following strategies of the agent and the principal and beliefs of the principal:*
1. The agent with type θ sends message $m\left(t\right)=1$ if and only if $X\left(t\right)\geq\bar{X}\left(\theta\right)$:

$$\bar{m}_t\left(\theta,X\left(t\right),\bar{\mu}\left(\cdot|\mathcal{H}_t\right)\right)=\begin{cases}1, & \text{if } X\left(t\right)\geq\bar{X}\left(\theta\right),\\ 0, & \text{otherwise.}\end{cases} \tag{10}$$

2. The posterior belief of the principal at any time t is that θ is distributed uniformly over $[\check{\theta}_t,\hat{\theta}_t]$ for some $\check{\theta}_t$ and $\hat{\theta}_t$ (possibly, equal).

3. The exercise strategy of the principal as a function of the state process and her beliefs is

$$\bar{e}_t(X\left(t\right), \check{\theta}_t, \hat{\theta}_t) = \begin{cases} 1, & \text{if } X\left(t\right) \geq \check{X}(\check{\theta}_t, \hat{\theta}_t), \\ 0, & \text{otherwise,} \end{cases} \tag{11}$$

for some threshold $\check{X}(\check{\theta}_t, \hat{\theta}_t)$. *Function* $\check{X}(\check{\theta}_t, \hat{\theta}_t)$ *is such that on equilibrium path the option is exercised at the first instant when the agent sends message* $m\left(t\right) = 1$, *i.e., when* $X\left(t\right)$ *hits threshold* $\bar{X}\left(\theta\right)$ *for the first time.*

Proof of Lemma IA.1. By contradiction, suppose that $\bar{X}\left(\theta_1\right) < \bar{X}\left(\theta_2\right)$ for some $\theta_2 > \theta_1$. Using the same arguments as in the derivation of (9) above but for $I - b$ instead of I, it is easy to see that if exercise occurs at a cutoff $\bar{X}$ and the current value of $X\left(t\right)$ is $X \leq \bar{X}$, then the agent's expected payoff is given by $\left(\frac{X}{\bar{X}}\right)^{\beta}\left(\theta\bar{X} - I + b\right)$. Hence, because the message strategy of type θ_1 is feasible for type θ_2, the IC condition of type θ_2 implies:

$$\left(\frac{X\left(t\right)}{\bar{X}\left(\theta_2\right)}\right)^{\beta}\left(\theta_2\bar{X}\left(\theta_2\right) - I + b\right) \geq \left(\frac{X\left(t\right)}{\bar{X}\left(\theta_1\right)}\right)^{\beta}\left(\theta_2\bar{X}\left(\theta_1\right) - I + b\right). \tag{12}$$

Similarly, because the message strategy of type θ_2 is feasible for type θ_1,

$$\left(\frac{X\left(t\right)}{\bar{X}\left(\theta_1\right)}\right)^{\beta}\left(\theta_1\bar{X}\left(\theta_1\right) - I + b\right) \geq \left(\frac{X\left(t\right)}{\bar{X}\left(\theta_2\right)}\right)^{\beta}\left(\theta_1\bar{X}\left(\theta_2\right) - I + b\right). \tag{13}$$

These inequalities imply

$$\begin{aligned} &\theta_2\bar{X}\left(\theta_1\right)\left(1 - \left(\tfrac{\bar{X}\left(\theta_1\right)}{\bar{X}\left(\theta_2\right)}\right)^{\beta-1}\right) \leq \left(I - b\right)\left(1 - \left(\tfrac{\bar{X}\left(\theta_1\right)}{\bar{X}\left(\theta_2\right)}\right)^{\beta}\right) \\ &\leq \theta_1\bar{X}\left(\theta_1\right)\left(1 - \left(\tfrac{\bar{X}\left(\theta_1\right)}{\bar{X}\left(\theta_2\right)}\right)^{\beta-1}\right), \end{aligned}$$

which is a contradiction, because $\theta_2 > \theta_1$ and $\frac{\bar{X}(\theta_1)}{\bar{X}(\theta_2)} < 1$. Thus, $\bar{X}\left(\theta_1\right) \geq \bar{X}\left(\theta_2\right)$ whenever $\theta_2 \geq \theta_1$.
■

Proof of Lemma IA.2. Consider a threshold exercise equilibrium E with an arbitrary message space M^* and equilibrium message strategy m^*, in which exercise occurs at stopping time $\tau^*\left(\theta\right) = \inf\left\{t \geq 0 | X\left(t\right) \geq \bar{X}\left(\theta\right)\right\}$ for some set of thresholds $\bar{X}\left(\theta\right)$, $\theta \in \Theta$. By Lemma IA.1, $\bar{X}\left(\theta\right)$ is weakly decreasing. Define $\theta_l\left(X\right) \equiv \inf\left\{\theta : \bar{X}\left(\theta\right) = X\right\}$ and $\theta_h\left(X\right) \equiv \sup\left\{\theta : \bar{X}\left(\theta\right) = X\right\}$ for any $X \in \mathcal{X}$. We will construct a different equilibrium, denoted by $\bar{E}$, which implements the same equilibrium exercise time $\tau^*\left(\theta\right)$ and has the structure specified in the formulation of the lemma. As we will see, it will imply that on the equilibrium path, the principal exercises the option at the first informative time $t \in \mathcal{T}$ at which she receives message $m\left(t\right) = 1$, where the set $\mathcal{T}$ of informative times is defined as

$$\mathcal{T} \equiv \left\{ t : X(t) = \bar{X} \text{ for some } \bar{X} \in \mathcal{X} \text{ and } X(s) < \bar{X} \; \forall s < t \; \right\},$$

i.e., the set of times when the process $X(t)$ reaches one of the thresholds in X for the first time.

For the collection of strategies (10) and (11) and the corresponding beliefs to be an equilibrium, we need to verify the IC conditions of the agent and the principal.

1 - IC of the agent. The IC condition of the agent requires that any type θ is better off sending a message $m(t) = 1$ when $X(t)$ first reaches $\bar{X}(\theta)$ than following any other strategy. By Assumption 1, a deviation to sending $m(t) = 1$ at any $t \notin \mathcal{T}$ does not lead the principal to change her beliefs, and hence, her behavior. Thus, it is without loss of generality to only consider deviations at $t \in \mathcal{T}$. There are two possible deviations: sending $m(t) = 1$ before $X(t)$ first reaches $\bar{X}(\theta)$ and sending $m(t) = 0$ at that moment and following some other strategy after that. Consider the first deviation: the agent of type θ can send $m(t) = 1$ when $X(t)$ hits threshold $\bar{X}(\hat{\theta})$, $\hat{\theta} > \theta_h\left(\bar{X}(\theta)\right)$ for the first time, and then the principal will exercise immediately. Consider the second deviation: if type θ deviates to sending $m(t) = 0$ when $X(t)$ hits threshold $\bar{X}(\theta)$, he can then either send $m(t) = 1$ at one of the future $t \in \mathcal{T}$ or continue sending the message $m(t) = 0$ at any future $t \in \mathcal{T}$. First, if the agent deviates to sending $m(t) = 1$ at one of the future $t \in \mathcal{T}$, the principal will exercise the option at one of the thresholds $\hat{Y} \in \mathcal{X}$, $\hat{Y} > \bar{X}(\theta)$. Note that the agent can ensure exercise at any threshold $\hat{Y} \in \mathcal{X}$ such that $\hat{Y} \geq X(t)$ by adopting the equilibrium message strategy of type $\hat{\theta}$ at which $\bar{X}(\hat{\theta}) = \hat{Y}$. Second, if the agent deviates to sending $m(t) = 0$ at all of the future $t \in \mathcal{T}$, there are two cases. If $\bar{X}(\underline{\theta}) = \infty$, the principal will never exercise the option. If $\bar{X}(\underline{\theta}) = \bar{X}_{\max} < \infty$, then the principal's belief when $X(t)$ first reaches $\bar{X}_{\max}$ is that $\theta = \underline{\theta}$, if $\bar{X}(\underline{\theta}) \neq \bar{X}(\theta) \; \forall \theta \neq \underline{\theta}$, or that $\theta \in \left[\underline{\theta}, \theta_h\left(\bar{X}_{\max}\right)\right]$, otherwise. Upon receiving $m(t) = 0$ at this moment, the principal does not change her belief by Assumption 1 and hence exercises the option at $\bar{X}_{\max} = \bar{X}(\underline{\theta})$. Finally, note that the agent cannot induce exercise at $\hat{Y} \in \mathcal{X}$ if $\hat{Y} < X(t)$: in this case, the principal's belief is that the agent's type is smaller than the type that could induce exercise at $\hat{Y}$ and this belief cannot be reversed according to Assumption 1. Combining all possible deviations, at time t, the agent can deviate to exercise at any $\hat{Y} \in \mathcal{X}$ as long as $\hat{Y} \geq X(t)$. Using the same arguments as in the derivation of (9) above but for $I - b$ instead of I, it is easy to see that the agent's expected utility given exercise at threshold $\bar{X}$ is $\left(\frac{X(t)}{\bar{X}}\right)^{\beta}\left(\theta\bar{X} - I + b\right)$. Hence, the IC condition of the agent is that

$$\left(\frac{X(t)}{\bar{X}(\theta)}\right)^{\beta}\left(\theta\bar{X}(\theta) - I + b\right) \geq \max_{\hat{Y} \in \mathcal{X}, \hat{Y} \geq X(t)} \left(\frac{X(t)}{\hat{Y}}\right)^{\beta}\left(\theta\hat{Y} - I + b\right). \tag{14}$$

Let us argue that it holds using the fact that E is an equilibrium. Suppose otherwise. Then, there exists a pair $(\theta, \hat{Y})$ with $\hat{Y} \in \mathcal{X}$ such that

$$\frac{\theta \bar{X}(\theta) - I + b}{\bar{X}(\theta)^{\beta}} < \frac{\theta \hat{Y} - I + b}{\hat{Y}^{\beta}}. \tag{15}$$

However, (15) implies that in equilibrium E type θ is better off deviating from the message strategy $m^*(\theta)$ to the message strategy $m^*\left(\tilde{\theta}\right)$ of type $\tilde{\theta}$, where $\tilde{\theta}$ is any type satisfying $\bar{X}\left(\tilde{\theta}\right) = \hat{Y}$ (since $\hat{Y} \in \mathcal{X}$, at least one such $\tilde{\theta}$ exists). This is impossible, and hence (14) holds. Hence, if the principal plays strategy (11), the agent finds it optimal to play strategy (10).

Given Lemma IA.1 and the fact that the agent plays (10), the posterior belief of the principal at any time t is that θ is distributed uniformly over $[\check{\theta}_t, \hat{\theta}_t]$ for some $\check{\theta}_t$ and $\hat{\theta}_t$ (possibly, equal). Next, consider the IC conditions of the principal. They are comprised of two parts, as evident from (11): we refer to the top line of (11) (exercising immediately when the principal "should" exercise) as the ex-post IC condition, and to the bottom line of (11) (not exercising when the principal "should" wait) as the ex-ante IC condition.

2 - "Ex-post" IC of the principal. First, consider the ex-post IC condition: we prove that the principal exercises immediately if the agent sends message $m(t) = 1$ at the first moment when $X(t)$ hits threshold $\hat{Y}$ for some $\hat{Y} \in \mathcal{X}$ (and sent message $m(t) = 0$ before). Given this message, the principal believes that $\theta \sim Uni[\theta_l(\hat{Y}), \theta_h(\hat{Y})]$. Because the principal expects the agent to play (10), the principal now expects the agent to send $m(t) = 1$ if $X(t) \geq \hat{Y}$, and $m(t) = 0$ otherwise, regardless of $\theta \in [\theta_l(\hat{Y}), \theta_h(\hat{Y})]$. Hence, the principal does not expect to learn any new information. This implies that the principal's problem is now equivalent to the standard option exercise problem with the option paying off $\frac{\theta_l(\hat{Y})+\theta_h(\hat{Y})}{2} X(t)$ upon exercise at time t. Using the same arguments as in the derivation of (9) above, the principal's expected payoff from exercise at threshold $\bar{X}$ is $\left(\frac{X(t)}{\bar{X}}\right)^{\beta} \left(\frac{\theta_l(\hat{Y})+\theta_h(\hat{Y})}{2} \bar{X} - I\right)$, which is an inverse U-shaped function with an unconditional maximum at $\frac{\beta}{\beta-1} \frac{2I}{\theta_l(\hat{Y})+\theta_h(\hat{Y})}$. Thus, the solution of the problem is to exercise the option immediately if and only if

$$X(t) \geq \frac{\beta}{\beta - 1} \frac{2I}{\theta_l(\hat{Y}) + \theta_h(\hat{Y})}. \tag{16}$$

Let us show that any threshold $\hat{Y} \in \mathcal{X}$ and the corresponding type cutoffs $\theta_l(\hat{Y})$ and $\theta_h(\hat{Y})$ in equilibrium E satisfy (16). Consider equilibrium E. For the principal to exercise at threshold $\bar{X}(\theta)$, the value that the principal gets upon exercise must be greater or equal than what she gets from delaying the exercise. The value from immediate exercise equals $\mathbb{E}[\theta | \mathcal{H}_t, m(t)] \bar{X}(\theta) - I$, where $(\mathcal{H}_t, m(t))$ is any history of the sample path of $X(t)$ and equilibrium messages that leads to exercise at time t at threshold $\bar{X}(\theta)$ in equilibrium E. Because waiting until $X(t)$ hits a threshold $\tilde{Y} > \bar{X}(\theta)$ and exercising then is a feasible strategy, the value from delaying exercise is greater or equal than the value from such a deviation, which equals $\left(\frac{\bar{X}(\theta)}{\tilde{Y}}\right)^{\beta}\left(\mathbb{E}\left[\theta|\mathcal{H}_t, m\left(t\right)\right]\tilde{Y} - I\right)$. Hence, $\bar{X}\left(\theta\right)$ must satisfy

$$\bar{X}\left(\theta\right) \in \arg\max_{\tilde{Y} \geq \bar{X}(\theta)} \left(\frac{\bar{X}\left(\theta\right)}{\tilde{Y}}\right)^{\beta}\left(\mathbb{E}\left[\theta|\mathcal{H}_t, m\left(t\right)\right]\tilde{Y} - I\right).$$

Using the fact that the unconditional maximizer of the right-hand side is $\tilde{Y} = \frac{\beta}{\beta-1}\frac{I}{\mathbb{E}[\theta|\mathcal{H}_t, m(t)]}$ and that function $\left(\frac{\bar{X}(\theta)}{\tilde{Y}}\right)^{\beta}\left(\mathbb{E}\left[\theta|\mathcal{H}_t, m_t\right]\tilde{Y} - I\right)$ is inverted U-shaped in $\tilde{Y}$, this condition can be equivalently re-written as

$$\bar{X}\left(\theta\right) \geq \frac{\beta}{\beta-1}\frac{I}{\mathbb{E}\left[\theta|\mathcal{H}_t, m\left(t\right)\right]},$$

for any history $\left(\mathcal{H}_t, m\left(t\right)\right)$ with $X\left(t\right) = \bar{X}\left(\theta\right)$ and $m\left(s\right) = m_s^*\left(\mathcal{H}_s, \theta\right)$ for some $\theta \in [\theta_l(\hat{Y}), \theta_h(\hat{Y})]$ and $s \leq t$. Let $\mathbb{H}_t^*$ denote the set of such histories. Then,

$$\bar{X}\left(\theta\right) \geq \frac{\beta}{\beta-1}\max_{\left(\mathcal{H}_t, m(t)\right) \in \mathbb{H}_t^*}\frac{I}{\mathbb{E}\left[\theta|\mathcal{H}_t, m\left(t\right)\right]},$$

or, equivalently,

$$\begin{aligned}
&\tfrac{\beta}{\beta-1}\tfrac{I}{\bar{X}(\theta)} \leq \min\nolimits_{\left(\mathcal{H}_t, m(t)\right) \in \mathbb{H}_t^*}\mathbb{E}\left[\theta|\mathcal{H}_t, m\left(t\right)\right] \\
&\leq \mathbb{E}\left[\mathbb{E}\left[\theta|\mathcal{H}_t, m\left(t\right)\right]|\theta \in [\theta_l(\hat{Y}), \theta_h(\hat{Y})], \mathcal{H}_0\right] \\
&= \mathbb{E}\left[\theta|\theta \in [\theta_l(\hat{Y}), \theta_h(\hat{Y})]\right] = \tfrac{\theta_l(\hat{Y})+\theta_h(\hat{Y})}{2},
\end{aligned}$$

where the inequality follows from the fact that the minimum of a random variable cannot exceed its mean, and the first equality follows from the law of iterated expectations. Therefore, when the principal obtains message $m = 1$ at threshold $\hat{Y} \in \mathcal{X}$, her optimal reaction is to exercise immediately. Thus, the ex-post IC condition of the principal is satisfied.

3 - "Ex-ante" IC of the principal. Finally, consider the ex-ante IC condition of the principal stating that the principal is better off waiting following a history $\mathcal{H}_t$ with $m\left(s\right) = 0$, $s \leq t$, and $\max_{s \leq t} X\left(s\right) < \bar{X}\left(\underline{\theta}\right)$. Given that the agent follows (11), for any such history $\mathcal{H}_t$, the principal's belief is that $\theta \sim Uni[\underline{\theta}, \theta_l(\hat{Y})]$ for some $\hat{Y} \in \mathcal{X}$. If the principal exercises immediately, her payoff is $\frac{\underline{\theta}+\theta_l(\hat{Y})}{2}X\left(t\right) - I$. If the principal waits, her expected payoff is

$$\int_{\underline{\theta}}^{\theta_l(\hat{Y})}\left(\frac{X\left(t\right)}{\bar{X}\left(\theta\right)}\right)^{\beta}\left(\theta\bar{X}\left(\theta\right) - I\right)\frac{1}{\theta_l(\hat{Y}) - \underline{\theta}}d\theta.$$

Suppose that there exists a pair $\hat{Y} \in \mathcal{X}$ and $\tilde{Y} < \lim_{\theta \uparrow \theta_l(\hat{Y})}\bar{X}\left(\theta\right)$ such that immediate exercise is optimal when $X(t) = \tilde{Y}$:

$$\frac{\underline{\theta} + \theta_l(\hat{Y})}{2}\tilde{Y} - I > \int_{\underline{\theta}}^{\theta_l(\hat{Y})} \left(\frac{\tilde{Y}}{\bar{X}(\theta)}\right)^{\beta} \left(\theta\bar{X}(\theta) - I\right) \frac{1}{\theta_l(\hat{Y}) - \underline{\theta}} d\theta. \tag{17}$$

We can re-write (17) as

$$\mathbb{E}_{\theta}\left[\left(\frac{1}{\tilde{Y}}\right)^{\beta}\left(\theta\tilde{Y} - I\right)|\theta < \theta_l(\hat{Y})\right] > \mathbb{E}_{\theta}\left[\left(\frac{1}{\bar{X}(\theta)}\right)^{\beta}\left(\theta\bar{X}(\theta) - I\right)|\theta < \theta_l(\hat{Y})\right]. \tag{18}$$

Let us show that if equilibrium E exists, then (18) must be violated. Consider equilibrium E, any type $\tilde{\theta} < \theta_l(\hat{Y})$, time $t < \tau^*\left(\tilde{\theta}\right)$, and any history $(\mathcal{H}_t, m(t))$ such that $X(t) = \tilde{Y}$, $\max_{s \le t, s \in \mathcal{T}} X(s) = \hat{Y}$, which is consistent with the equilibrium play of type $\tilde{\theta}$, i.e., with $m(s) = m_s^*\left(\tilde{\theta}, \mathcal{H}_s\right)$ $\forall s \le t$. Let $\mathbb{H}_t^{**}(\tilde{\theta}, \tilde{Y}, \hat{Y})$ denote the set of such histories. Because the principal prefers waiting in equilibrium E, the payoff from immediate exercise in equilibrium E cannot exceed the payoff from waiting:

$$\begin{gathered}\mathbb{E}\left[\theta\tilde{Y} - I|\mathcal{H}_t, m(t)\right] \le \mathbb{E}\left[\left(\tfrac{\tilde{Y}}{\bar{X}(\theta)}\right)^{\beta}\left(\theta\bar{X}(\theta) - I\right)|\mathcal{H}_t, m(t)\right] \Leftrightarrow \\ \mathbb{E}\left[\left(\tfrac{1}{\bar{X}(\theta)}\right)^{\beta}\left(\theta\bar{X}(\theta) - I\right) - \left(\tfrac{1}{\tilde{Y}}\right)^{\beta}\left(\theta\tilde{Y} - I\right)|\mathcal{H}_t, m(t)\right] \ge 0.\end{gathered}$$

This inequality must hold for all histories $(\mathcal{H}_t, m(t)) \in \mathbb{H}_t^{**}(\tilde{\theta}, \tilde{Y}, \hat{Y})$. In any history $(\mathcal{H}_t, m(t)) \in \mathbb{H}_t^{**}(\tilde{\theta}, \tilde{Y}, \hat{Y})$, the option is never exercised by time t if $\tilde{\theta} < \theta_l(\hat{Y})$ and is exercised before time t if $\tilde{\theta} > \theta_l(\hat{Y})$. Therefore, conditional on $\tilde{Y}$, $\hat{Y}$, and $\tilde{\theta} < \theta_l(\hat{Y})$, the distribution of $\tilde{\theta}$ is independent of the sample path of $X(s), s \le t$. Fixing $\tilde{Y}$ and $\hat{Y}$ and integrating over histories $(\mathcal{H}_t, m(t)) \in \mathbb{H}_t^{**}(\tilde{\theta}, \tilde{Y}, \hat{Y})$ and then over $\tilde{\theta} \in [\underline{\theta}, \theta_l(\hat{Y}))$, we obtain that

$$\begin{gathered}\mathbb{E}_{\tilde{\theta}}\left[\mathbb{E}_{(\mathcal{H}_t, m(t))}\left[\left(\tfrac{1}{\bar{X}(\theta)}\right)^{\beta}\left(\theta\bar{X}(\theta) - I\right) - \left(\tfrac{1}{\tilde{Y}}\right)^{\beta}\left(\theta\tilde{Y} - I\right)|(\mathcal{H}_t, m(t)) \in \mathbb{H}_t^{**}(\tilde{\theta}, \tilde{Y}, \hat{Y})\right]|\tilde{\theta} \in [\underline{\theta}, \theta_l(\hat{Y})), \hat{Y}, \tilde{Y}\right] \\ \ge 0 \Leftrightarrow \mathbb{E}_{\theta}\left[\left(\tfrac{1}{\bar{X}(\theta)}\right)^{\beta}\left(\theta\bar{X}(\theta) - I\right) - \left(\tfrac{1}{\tilde{Y}}\right)^{\beta}\left(\theta\tilde{Y} - I\right)|\theta < \theta_l(\hat{Y})\right] \ge 0,\end{gathered}$$

where we applied the law of iterated expectations and the conditional independence of the sample path of $X(t)$ and the distribution of $\tilde{\theta}$ (conditional on $\tilde{Y}$, $\hat{Y}$, and $\tilde{\theta} < \theta_l(\hat{Y})$). Therefore, (18) cannot hold. Hence, the ex-ante IC condition of the principal is also satisfied.

Thus, if there exists a threshold exercise equilibrium E where $\tau^*(\theta) = \inf\left\{t \ge 0 | X(t) \ge \bar{X}(\theta)\right\}$ for some threshold $\bar{X}(\theta)$, then there exists a threshold exercise equilibrium $\bar{E}$ of the form specified in the lemma, in which the option is exercised at the same time. Finally, let us show that on the equilibrium path, the option is indeed exercised at the first informative time t at which the principal receives message $m(t) = 1$. Because any message sent at $t \notin \mathcal{T}$ does not lead to updating of the principal's beliefs and because of the second part of (11), the principal never exercises the option prior to the first informative time $t \in \mathcal{T}$ at which she receives message $m(t) = 1$. Consider the first informative time $t \in \mathcal{T}$ at which the principal receives $m(t) = 1$. By Bayes' rule, the principal believes that θ is distributed uniformly over $(\theta_l(X(t)), \theta_h(X(t)))$. Equilibrium strategy of the agent (10) implies $X(t) = \bar{X}(\theta)\ \forall \theta \in (\theta_l(X(t)), \theta_h(X(t)))$. Therefore, in equilibrium the principal exercises the option immediately. ■

For Lemma 1, we prove the following lemma, which characterizes the structure of any incentive-compatible decision-making rule and is an analogue of Proposition 1 in Melumad and Shibano (1991) for the payoff specification in our model:

Lemma IA.3. *An incentive-compatible threshold schedule $\hat{X}(\theta)$ must satisfy the following conditions:*

1. *$\hat{X}(\theta)$ is weakly decreasing in θ.*
2. *If $\hat{X}(\theta)$ is strictly decreasing on (θ_1, θ_2), then $\hat{X}(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$.*
3. *If $\hat{X}(\theta)$ is discontinuous at $\hat{\theta}$, then the discontinuity satisfies*

$$\hat{U}_A\left(\hat{X}^-(\hat{\theta}), \hat{\theta}\right) = \hat{U}_A\left(\hat{X}^+(\hat{\theta}), \hat{\theta}\right), \tag{19}$$

$$\hat{X}(\theta) = \begin{cases} \hat{X}^-(\hat{\theta}), & \forall \theta \in \left[\frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^-(\hat{\theta})}, \hat{\theta}\right), \\ \hat{X}^+(\hat{\theta}), & \forall \theta \in \left(\hat{\theta}, \frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^+(\hat{\theta})}\right], \end{cases} \tag{20}$$

$$\hat{X}(\hat{\theta}) \in \left\{\hat{X}^-(\hat{\theta}), \hat{X}^+(\hat{\theta})\right\}, \tag{21}$$

where $\hat{X}^-(\hat{\theta}) \equiv \lim_{\theta \uparrow \hat{\theta}} \hat{X}(\theta)$ and $\hat{X}^+(\hat{\theta}) \equiv \lim_{\theta \downarrow \hat{\theta}} \hat{X}(\theta)$.

Proof of Lemma IA.3. **Proof of Part 1**. The first part of the lemma can be proven by contradiction. Suppose there exist $\theta_1, \theta_2 \in \Theta$, $\theta_2 > \theta_1$, such that $\hat{X}(\theta_2) > \hat{X}(\theta_1)$. Note that $\hat{U}_A(\hat{X}, \theta) \equiv X(0)^\beta \hat{X}^{-\beta}(\theta\hat{X} - I + b)$ and $\hat{U}_P(\hat{X}, \theta) \equiv X(0)^\beta \hat{X}^{-\beta}(\theta\hat{X} - I)$. The agent's IC constraint for $\theta = \theta_1$ and $\hat{\theta} = \theta_2$, $\hat{U}_A\left(\hat{X}(\theta_1), \theta_1\right) \geq \hat{U}_A\left(\hat{X}(\theta_2), \theta_1\right)$, can be written in the integral form:

$$\int_{\hat{X}(\theta_1)}^{\hat{X}(\theta_2)} \left(\frac{X(0)}{\hat{X}}\right)^\beta \frac{-(\beta-1)\theta_1\hat{X} + \beta(I-b)}{\hat{X}} d\hat{X} \leq 0. \tag{22}$$

Because $\theta_2 > \theta_1$ and $\beta > 1$, (22) implies

$$\int_{\hat{X}(\theta_1)}^{\hat{X}(\theta_2)} \left(\frac{X(0)}{\hat{X}}\right)^\beta \frac{-(\beta-1)\theta_2\hat{X} + \beta(I-b)}{\hat{X}} d\hat{X} < 0,$$

or, equivalently, $\hat{U}_A(\hat{X}(\theta_1), \theta_2) > \hat{U}_A(\hat{X}(\theta_2), \theta_2)$. However, this violates the agent's IC constraint $\hat{U}_A(\hat{X}(\theta_2), \theta_2) \geq \hat{U}_A(\hat{X}(\theta_1), \theta_2)$ for $\theta = \theta_2$ and $\hat{\theta} = \theta_1$. Thus, $\hat{X}(\theta)$ is weakly decreasing in θ.

Proof of Part 2. To prove the second part of the lemma, note that $\hat{U}_A\left(\hat{X}, \theta\right)$ is differentiable in θ for all $\hat{X} \in (X(0), \infty)$. Because $\hat{U}_A(\hat{X}, \theta)$ is linear in θ, it satisfies the Lipschitz condition and hence is absolutely continuous in θ for all $\hat{X} \in (X(0), \infty)$. Also, $\frac{\partial \hat{U}_A(\hat{X},\theta)}{\partial \theta} = \left(\frac{X(0)}{\hat{X}}\right)^\beta \hat{X}$, and hence $\sup_{\hat{X}\in \mathbf{X}} \left|\frac{\partial \hat{U}_A(\hat{X},\theta)}{\partial \theta}\right|$ is integrable on $\theta \in \Theta$. By the generalized envelope theorem (Milgrom and Segal, 2002), the equilibrium utility of the agent in any mechanism implementing exercise at thresholds $\hat{X}(\theta)$, $\theta \in \Theta$, denoted $V_A(\theta)$, satisfies the integral condition,

$$V_A(\theta) = V_A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \left(\frac{X(0)}{\hat{X}(s)}\right)^\beta \hat{X}(s)\, ds.$$

On the other hand, $V_A(\theta) = \hat{U}_A\left(\hat{X}(\theta), \theta\right)$. At any point θ at which $\hat{X}(\theta)$ is strictly decreasing, we have

$$\frac{dV_A(\theta)}{d\theta} = \frac{d\hat{U}_A(\hat{X}(\theta), \theta)}{d\theta} \Leftrightarrow \frac{X(0)^\beta}{\hat{X}(\theta)^\beta}\hat{X}(\theta) = \frac{X(0)^\beta}{\hat{X}(\theta)^\beta}\hat{X}(\theta) - \frac{X(0)^\beta}{\hat{X}(\theta)^\beta}\frac{(\beta-1)\theta\hat{X}(\theta) - \beta(I-b)}{\hat{X}(\theta)}\frac{d\hat{X}(\theta)}{d\theta}.$$

Because $d\hat{X}(\theta) < 0$, it must be that $(\beta-1)\theta\hat{X}(\theta) - \beta(I-b) = 0$. Thus, $\hat{X}(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$, which proves the second part of the lemma.

Proof of Part 3. Finally, consider the third part of the lemma. Eq. (19) follows from (20), continuity of $\hat{U}_A(\cdot)$, and incentive compatibility of the mechanism. Otherwise, for example, if $\hat{U}_A\left(\hat{X}^+(\hat{\theta}), \hat{\theta}\right) > \hat{U}_A\left(\hat{X}^-(\hat{\theta}), \hat{\theta}\right)$, then $\hat{U}_A\left(\hat{X}(\hat{\theta}-\varepsilon), \hat{\theta}-\varepsilon\right) = \hat{U}_A\left(\hat{X}^-(\hat{\theta}), \hat{\theta}-\varepsilon\right) < \hat{U}_A\left(\hat{X}^+(\hat{\theta}), \hat{\theta}-\varepsilon\right)$ for a sufficiently small ε, and hence types close enough to $\hat{\theta}$ from below would benefit from a deviation to $\hat{X}^+(\hat{\theta})$, i.e., from mimicking types slightly above $\hat{\theta}$.

Next, we prove (20). First, note that, (20) is satisfied at the boundaries. Indeed, denote $\theta_1^* \equiv \frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^-(\hat{\theta})}$ and suppose that $\hat{X}(\theta_1^*) \neq \hat{X}^-(\hat{\theta})$. Then, by the first part of the lemma, $\hat{X}(\theta_1^*) > \hat{X}^-(\hat{\theta})$. Because $\hat{X}^-(\hat{\theta}) \equiv \lim_{\theta\uparrow\hat{\theta}} \hat{X}(\theta)$, there exists $\varepsilon > 0$ such that $\hat{X}(\theta_1^*) > \hat{X}(\hat{\theta}-\varepsilon) \geq \hat{X}^-(\hat{\theta})$. Because the function $\hat{U}_A(x, \theta_1^*)$ has a maximum at $\hat{X}^-(\hat{\theta})$ and is strictly decreasing for $x > \hat{X}^-(\hat{\theta})$, this would imply $\hat{U}_A(\hat{X}(\theta_1^*), \theta_1^*) < \hat{U}_A(\hat{X}\left(\hat{\theta}-\varepsilon\right), \theta_1^*)$, and hence would contradict the IC condition for type θ_1^*. The proof for the boundary $\theta_2^* \equiv \frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^+(\hat{\theta})}$ is similar.

We next prove (20) for interior values of θ. First, suppose that $\hat{X}(\theta) \neq \hat{X}^-(\hat{\theta})$ for some $\theta \in \left(\frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^-(\hat{\theta})}, \hat{\theta}\right)$. By part 1 of the lemma, $\hat{X}(\theta) > \hat{X}^-(\hat{\theta})$. By IC, $\hat{U}_A(\hat{X}(\theta), \theta) \geq \hat{U}_A(\hat{X}^-(\hat{\theta}), \theta)$, which can be written in the integral form as:

$$\int_{\hat{X}^-(\hat{\theta})}^{\hat{X}(\theta)} \left(\frac{X(0)}{\xi}\right)^\beta \frac{-(\beta-1)\theta\xi + \beta(I-b)}{\xi} d\xi \geq 0.$$

The function under the integral on the left-hand side is strictly decreasing in θ and the interval $(\hat{X}^{-}(\hat{\theta}), \hat{X}(\theta))$ is non-empty. Thus, we can replace θ by $\tilde{\theta} < \theta$ under the integral and get a strict inequality: $\hat{U}_A(\hat{X}(\theta), \tilde{\theta}) > \hat{U}_A(\hat{X}^{-}(\hat{\theta}), \tilde{\theta})$ for every $\tilde{\theta} \in [\frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^{-}(\hat{\theta})}, \theta)$. However, this contradicts $\hat{X}^{-}(\hat{\theta}) = \arg\max_x \hat{U}_A(x, \frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^{-}(\hat{\theta})})$. Second, suppose that $\hat{X}(\theta) \neq \hat{X}^{+}(\hat{\theta})$ for some $\theta \in \left(\hat{\theta}, \frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^{+}(\hat{\theta})}\right)$. By part 1 of the lemma, $\hat{X}(\theta) < \hat{X}^{+}(\hat{\theta})$. By incentive compatibility, $\hat{U}_A\left(\hat{X}(\theta), \theta\right) \geq \hat{U}_A\left(\hat{X}^{+}(\theta), \theta\right)$, which can be written as

$$\int_{\hat{X}(\theta)}^{\hat{X}^{+}(\theta)} \left(\frac{X(0)}{\xi}\right)^{\beta} \frac{-(\beta-1)\theta\xi + \beta(I-b)}{\xi} d\xi \leq 0.$$

The function under the integral on the left-hand side is strictly decreasing in θ and the interval $(\hat{X}(\theta), \hat{X}^{+}(\hat{\theta}))$ is non-empty. Therefore, we can replace θ by $\tilde{\theta} > \theta$ under the integral and get a strict inequality, $\hat{U}_A\left(\hat{X}(\theta), \tilde{\theta}\right) > \hat{U}_A\left(\hat{X}^{+}(\theta), \tilde{\theta}\right)$, for every $\tilde{\theta} \in \left(\theta, \frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^{+}(\hat{\theta})}\right]$. However, this contradicts $\hat{X}^{+}(\hat{\theta}) = \arg\max_x \hat{U}_A\left(x, \frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^{+}(\hat{\theta})}\right)$.

Finally, (21) follows from the continuity of $\hat{U}_A(\cdot)$ and incentive compatibility of $\hat{X}(\theta)$. Because $\hat{\theta} \in (\frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^{-}(\hat{\theta})}, \frac{\beta}{\beta-1}\frac{I-b}{\hat{X}^{+}(\hat{\theta})})$, every policy with thresholds strictly below $\hat{X}^{-}(\hat{\theta})$ or strictly above $\hat{X}^{+}(\hat{\theta})$ is strictly dominated by $\hat{X}^{-}(\hat{\theta})$ and $\hat{X}^{+}(\hat{\theta})$, respectively, and thus cannot be incentive-compatible. Suppose that $\hat{X}(\hat{\theta}) \in \left(\hat{X}^{-}(\hat{\theta}), \hat{X}^{+}(\hat{\theta})\right)$. Incentive compatibility and (19) imply $\hat{U}_A\left(\hat{X}(\hat{\theta}), \hat{\theta}\right) \geq \hat{U}_A\left(\hat{X}^{-}(\hat{\theta}), \hat{\theta}\right) = \hat{U}_A\left(\hat{X}^{+}(\hat{\theta}), \hat{\theta}\right)$. Because $\hat{U}_A\left(x, \hat{\theta}\right)$ is strictly increasing in x for $x < \frac{\beta}{\beta-1}\frac{I-b}{\hat{\theta}}$ and strictly decreasing in x for $x > \frac{\beta}{\beta-1}\frac{I-b}{\hat{\theta}}$, the inequality must be strict: $\hat{U}_A\left(\hat{X}(\hat{\theta}), \hat{\theta}\right) > \hat{U}_A\left(\hat{X}^{-}(\hat{\theta}), \hat{\theta}\right) = \hat{U}_A\left(\hat{X}^{+}(\hat{\theta}), \hat{\theta}\right)$. However, this together with (20) and continuity of $\hat{U}_A(\cdot)$ implies that types close enough to $\hat{\theta}$ benefit from a deviation to threshold $\hat{X}(\hat{\theta})$. Hence, it must be that $\hat{X}(\hat{\theta}) \in \left\{\hat{X}^{-}(\hat{\theta}), \hat{X}^{+}(\hat{\theta})\right\}$. ■

Proof of Lemma 1. We show that for all parameter values, except the case $b = -I$ and $\underline{\theta} = 0$, there exists a unique optimal contract, and it takes the form specified in the lemma. When $b = -I$ and $\underline{\theta} = 0$, the optimal contract is not unique, but the flat contract specified in the lemma is optimal. To prove the lemma, we consider three cases: $b \geq I$, $b \in [-I, I)$, and $b < -I$. Denote the flat contract from the first part of the lemma by $\hat{X}_{flat}(\theta)$, the contract from the second part by $\hat{X}_{-}(\theta)$, and the contract from the third part by $\hat{X}_{+}(\theta)$.

Case 1: $b \geq I$. In this case, all types of agents want to exercise the option immediately. This means that any incentive-compatible contract must be flat. Among flat contracts, the one that maximizes the payoff to the principal solves

$$\arg\max_x \int_{\underline{\theta}}^{1} \frac{\theta x - I}{x^{\beta}} d\theta = \frac{2\beta}{\beta-1}\frac{I}{1+\underline{\theta}}. \tag{23}$$

Case 2: $b \in [-I, I)$. The proof for this case proceeds in two steps. First, we show that the optimal contract cannot have discontinuities, except the case $b = -I$. Second, we show that the optimal continuous contract is as specified in the lemma.

Step 1: *If $b > -I$, the optimal contract is continuous.* Indeed, by contradiction, suppose that the optimal contract $C = \left\{ \hat{X}(\theta), \theta \in \Theta \right\}$ has a discontinuity at some point $\hat{\theta} \in (\underline{\theta}, 1)$. By Lemma IA.3, the discontinuity must satisfy (19)–(21). In particular, (20) implies that there exist $\theta_1 < \hat{\theta}$ and $\theta_2 > \hat{\theta}$ such that $\hat{X}(\theta) = X_A^*(\theta_1)$ for $\theta \in \left[\theta_1, \hat{\theta}\right)$ and $\hat{X}(\theta) = X_A^*(\theta_2)$ for $\theta_2 \in \left(\hat{\theta}, \theta_2\right]$. For any $\tilde{\theta}_2 \in (\hat{\theta}, \theta_2]$, consider a contract $C_1 = \left\{ \hat{X}_1(\theta), \theta \in \Theta \right\}$, defined as

$$\hat{X}_1(\theta) = \begin{cases} \hat{X}(\theta), & \text{if } \theta \in [\underline{\theta}, \theta_1] \cup [\theta_2, 1], \\ X_A^*(\theta_1), & \text{if } \theta \in \left[\theta_1, \tilde{\theta}\right), \\ X_A^*\left(\tilde{\theta}_2\right), & \text{if } \theta \in \left(\tilde{\theta}, \tilde{\theta}_2\right], \\ X_A^*(\theta), & \text{if } \theta \in \left(\tilde{\theta}_2, \theta_2\right), \end{cases}$$

where $\tilde{\theta} = \tilde{\theta}\left(\tilde{\theta}_2\right)$ satisfies

$$\frac{\tilde{\theta} X_A^*(\theta_1) - I + b}{X_A^*(\theta_1)^\beta} = \frac{\tilde{\theta} X_A^*\left(\tilde{\theta}_2\right) - I + b}{X_A^*\left(\tilde{\theta}_2\right)^\beta}. \tag{24}$$

Because $X^{-\beta}(\theta X - I + b)$ is maximized at $X_A^*(\theta)$, the function $\pi(\theta) \equiv \frac{\theta X_A^*(\theta_1) - I + b}{X_A^*(\theta_1)^\beta} - \frac{\theta X_A^*\left(\tilde{\theta}_2\right) - I + b}{X_A^*\left(\tilde{\theta}_2\right)^\beta}$ satisfies $\pi(\theta_1) > 0 > \pi\left(\tilde{\theta}_2\right)$, and hence, by continuity of $\pi(\theta)$, there exists $\tilde{\theta} \in \left(\theta_1, \tilde{\theta}_2\right)$ such that $\pi\left(\tilde{\theta}\right) = 0$, i.e., (24) is satisfied. Intuitively, contract C_1 is the same as contract C, except that it substitutes a subset $\left[\tilde{\theta}_2, \theta_2\right]$ of the flat region with a continuous region where $\hat{X}_1(\theta) = \frac{\beta}{\beta - 1} \frac{I - b}{\theta}$. Because contract C is incentive-compatible and $\tilde{\theta}$ satisfies (24), contract C_1 is incentive-compatible too. Below we show that the payoff to the principal from contract C_1 exceeds the payoff to the principal from contract C for $\tilde{\theta}_2$ very close to θ_2. Because $\hat{X}_1(\theta) = \hat{X}(\theta)$ for $\theta \le \theta_1$ and $\theta \ge \theta_2$, it is enough to restrict attention to the payoff in the range $\theta \in (\theta_1, \theta_2)$. The payoff to the principal from contract C_1 in this range, divided by $X(0)^\beta \frac{1}{1 - \underline{\theta}}$, is

$$\int_{\theta_1}^{\tilde{\theta}(\tilde{\theta}_2)} \frac{\theta X_A^*(\theta_1) - I}{X_A^*(\theta_1)^\beta} d\theta + \int_{\tilde{\theta}(\tilde{\theta}_2)}^{\tilde{\theta}_2} \frac{\theta X_A^*\left(\tilde{\theta}_2\right) - I}{X_A^*\left(\tilde{\theta}_2\right)^\beta} d\theta + \int_{\tilde{\theta}_2}^{\theta_2} \frac{\theta X_A^*(\theta) - I}{X_A^*(\theta)^\beta} d\theta. \tag{25}$$

The derivative of (25) with respect to $\tilde{\theta}_2$, after the application of (24) and Leibniz's integral rule, is

$$(26)\qquad \int_{\tilde{\theta}}^{\tilde{\theta}_2} \frac{\beta I-(\beta-1)\,\theta X_A^*\left(\tilde{\theta}_2\right)}{X_A^*\left(\tilde{\theta}_2\right)^{\beta+1}} X_A^{*\prime}\left(\tilde{\theta}_2\right) d\theta + b\left(\frac{1}{X_A^*\left(\tilde{\theta}_2\right)^{\beta}}-\frac{1}{X_A^*\left(\theta_1\right)^{\beta}}\right)\frac{d\tilde{\theta}}{d\tilde{\theta}_2}.$$

Because $X_A^{*\prime}(\theta) = -\frac{X_A^*(\theta)}{\theta}$, the first term of (26) can be simplified to

$$(27)\quad \frac{(\beta-1)\,X_A^*\left(\tilde{\theta}_2\right)\frac{\tilde{\theta}_2^2-\tilde{\theta}^2}{2}-\beta I\left(\tilde{\theta}_2-\tilde{\theta}\right)}{X_A^*\left(\tilde{\theta}_2\right)^{\beta+1}}\frac{X_A^*\left(\tilde{\theta}_2\right)}{\tilde{\theta}_2} = \beta\frac{\tilde{\theta}_2-\tilde{\theta}}{\tilde{\theta}_2}X_A^*\left(\tilde{\theta}_2\right)^{-\beta}\left[\frac{I-b}{\tilde{\theta}_2}\frac{\tilde{\theta}_2+\tilde{\theta}}{2}-I\right].$$

From (24), $\frac{d\tilde{\theta}}{d\tilde{\theta}_2} = (\frac{\theta_1^\beta}{\theta_1}-\frac{\tilde{\theta}_2^\beta}{\tilde{\theta}_2})^{-1}\,(\beta-1)\,\tilde{\theta}_2^{\beta-2}(\tilde{\theta}-\tilde{\theta}_2)$. Using this and (24), the second term of (26) can be simplified to

$$(28)\qquad \frac{b}{X_A^*\left(\tilde{\theta}_2\right)^{\beta}}\left(1-\left(\frac{\tilde{\theta}_2}{\theta_1}\right)^{-\beta}\right)\frac{d\tilde{\theta}}{d\tilde{\theta}_2} = \beta\frac{\tilde{\theta}_2-\tilde{\theta}}{\tilde{\theta}_2}X_A^*\left(\tilde{\theta}_2\right)^{-\beta}\left(\frac{\tilde{\theta}}{\tilde{\theta}_2}\right)b.$$

Adding (27) and (28), the derivative of the principal's payoff with respect to $\tilde{\theta}_2$ is equal to $-\beta\frac{(\tilde{\theta}_2-\tilde{\theta})^2}{2\tilde{\theta}_2^2}X_A^*\left(\tilde{\theta}_2\right)^{-\beta}(I+b)$, which is strictly negative for any $b>-I$. By the mean value theorem, if $U_P\left(\tilde{\theta}_2\right)$ stands for the expected principal's utility from contract C, then $\frac{U_P(\tilde{\theta}_2)-U_P(\theta_2)}{\tilde{\theta}_2-\theta_2} = U_P'(\hat{\theta}_2) < 0$ for some $\hat{\theta}_2 \in \left(\tilde{\theta}_2, \theta_2\right)$, and hence a deviation from contract C to contract C_1 is beneficial for the principal. Hence, contract C cannot be optimal for $b>-I$.

Next, suppose $b=-I$. In this case, the derivative of (25) with respect to $\tilde{\theta}_2$ is zero for any $\tilde{\theta}_2 \in (\hat{\theta}, \theta_2]$. It can be similarly shown that if, instead, we replace a subset $\left[\theta_1, \tilde{\theta}_1\right]$ of the flat region $[\theta_1, \theta_2]$ with a continuous region where $\hat{X}_1(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$, then the derivative of the principal's utility with respect to $\tilde{\theta}_1$ is zero for any $\tilde{\theta}_1 \in [\theta_1, \hat{\theta})$. Combining the two arguments, contract C gives the principal the same expected utility as the contract where the flat region $[\theta_1, \theta_2]$ is replaced by a continuous region with $\hat{X}_1(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$, and the rest of the contract is unchanged. Thus, if a discontinuous contract is optimal, then there exists an equivalent continuous contract, which contains a strictly decreasing region and which is also optimal.

Step 2: *Optimal continuous contract.* We prove that among continuous contracts satisfying Lemma IA.3, the one specified in Lemma 1 maximizes the payoff to the principal. By Lemma IA.3 and continuity of the contract, it is sufficient to restrict attention to contracts that are combinations of, at most, one downward sloping part $\hat{X}(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$ and two flat parts: any contract that has at least two disjoint regions with $\hat{X}(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$ will exhibit discontinuity. Consider a contract such that $\hat{X}(\theta)$ is flat for $\theta \in [\underline{\theta}, \theta_1]$, is downward-sloping with $\hat{X}(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$ for $\theta \in [\theta_1, \theta_2]$, and is again flat for $\theta \in [\theta_2, 1]$, for some $\theta_1 \in [0, \theta_2]$ and $\theta_2 \in [\theta_1, 1]$. This consideration allows for all possible cases, because it can be that $\theta_1 = \underline{\theta}$ and/or $\theta_2 = 1$, or $\theta_1 = \theta_2$). The payoff to the principal, divided by $X(0)^\beta \frac{1}{1-\underline{\theta}}$, is

$$P = \int_{\underline{\theta}}^{\theta_1} \frac{\theta X_A^*(\theta_1) - I}{X_A^*(\theta_1)^\beta} d\theta + \int_{\theta_1}^{\theta_2} \frac{\theta X_A^*(\theta) - I}{X_A^*(\theta)^\beta} d\theta + \int_{\theta_2}^{1} \frac{\theta X_A^*(\theta_2) - I}{X_A^*(\theta_2)^\beta} d\theta. \tag{29}$$

Since $X_A^*(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$, the derivative with respect to θ_1 is

$$\frac{\partial P}{\partial \theta_1} = \int_{\underline{\theta}}^{\theta_1} \frac{\beta I - (\beta - 1)\theta X_A^*(\theta_1)}{X_A^*(\theta_1)^{\beta+1}} X_A^{*\prime}(\theta_1) d\theta = -\frac{\beta}{X_A^*(\theta_1)^\beta}\left[\frac{I+b}{2} - I\frac{\underline{\theta}}{\theta_1} + \left(\frac{\underline{\theta}}{\theta_1}\right)^2 \frac{I-b}{2}\right].$$

First, suppose $\underline{\theta} > 0$. Then $x = \frac{\underline{\theta}}{\theta_1}$ takes values between $\frac{\underline{\theta}}{\theta_2}$ and 1. Since $b \in [-I, I)$, the function $x^2 \frac{I-b}{2} - Ix + \frac{I+b}{2}$ is U-shaped and has two roots, 1 and $\frac{I+b}{I-b}$, which coincide for $b = 0$. If $b \in [0, I)$, this function is strictly positive for $x < 1$ because $\frac{I+b}{I-b} \geq 1$. Hence, $\frac{\partial P}{\partial \theta_1} < 0$ for $\theta_1 > \underline{\theta}$, which implies that (29) is maximized at $\theta_1 = \underline{\theta}$. If $-I < b < 0$, then $0 < \frac{I+b}{I-b} < 1$ and hence $\frac{\partial P}{\partial \theta_1} < 0$ when $\frac{\underline{\theta}}{\theta_1} < \frac{I+b}{I-b}$ or $\frac{\underline{\theta}}{\theta_1} > 1$, and $\frac{\partial P}{\partial \theta_1} > 0$ when $\frac{\underline{\theta}}{\theta_1} \in \left(\frac{I+b}{I-b}, 1\right)$. Because $\frac{\underline{\theta}}{\theta_1} \leq 1$, we conclude that (29) is increasing in θ_1 in the range $\theta_1 < \frac{I-b}{I+b}\underline{\theta}$ and decreasing in θ_1 in the range $\theta_1 > \frac{I-b}{I+b}\underline{\theta}$. Therefore, if $-I < b < 0$, (29) reaches its maximum at $\theta_1 = \min\left\{\frac{I-b}{I+b}\underline{\theta}, 1\right\}$. In particular, the maximum is achieved at $\theta_1 = \frac{I-b}{I+b}\underline{\theta}$ if $b \in [-\frac{1-\underline{\theta}}{1+\underline{\theta}}I, 0)$, and $\theta_1 = \theta_2$ if $-I < b \leq -\frac{1-\underline{\theta}}{1+\underline{\theta}}I$. Finally, if $b = -I$, then $\frac{I+b}{I-b} = 0$ and hence $\frac{\partial P}{\partial \theta_1} > 0$, i.e., (29) is increasing in θ_1. Thus, (29) is also maximized at $\theta_1 = \theta_2$.

Next, suppose $\underline{\theta} = 0$. Then $\frac{\partial P}{\partial \theta_1} < 0$ if $-I < b < I$ and $\frac{\partial P}{\partial \theta_1} = 0$, otherwise. Hence, for $-I < b < I$ and $\underline{\theta} = 0$, (29) is maximized at $\theta_1 = 0 = \underline{\theta}$. If $b = -I$ and $\underline{\theta} = 0$, the principal's utility does not depend on θ_1.

Next, the derivative of (29) with respect to θ_2 is

$$\frac{\partial P}{\partial \theta_2} = \int_{\theta_2}^{1} \frac{\beta I - (\beta - 1)\theta X_A^*(\theta_2)}{X_A^*(\theta_2)^{\beta+1}} X_A^{*\prime}(\theta_2) d\theta = \frac{\beta(1-\theta_2)}{2\theta_2^2 X_A^*(\theta_2)^\beta}(I - b - (I+b)\theta_2). \tag{30}$$

1) If $b \in [-I, 0)$, then $I - b - (I+b)\theta_2 \geq I - b - (I+b) > 0$, and hence (30) is positive for any $\theta_2 \in [\underline{\theta}, 1)$. Therefore, (29) is maximized at $\theta_2 = 1$. Combining this with the conclusions for θ_1 above, we get:

1a) For $\underline{\theta} > 0$: If $b \in [-\frac{1-\underline{\theta}}{1+\underline{\theta}}I, 0]$, then $\theta_1^* = \frac{I-b}{I+b}\underline{\theta}$ and $\theta_2^* = 1$, which together with continuity of the contract gives $\hat{X}_-(\theta)$. If $b \in [-I, -\frac{1-\underline{\theta}}{1+\underline{\theta}}I]$, then $\theta_1^* = \theta_2$ and $\theta_2^* = 1$, i.e., the optimal contract is flat. As shown above, among flat contracts, the one that maximizes the principal's payoff is $\hat{X}_{flat}(\theta)$. Note that this result implies that the optimal contract is unique among both continuous and discontinuous contracts even if $b = -I$. Indeed, Step 1 shows that the principal's utility in any discontinuous contract is the same as in a continuous contract with a strictly decreasing region. Because the optimal contract among continuous contracts is unique and is strictly flat, the principal's utility in any discontinuous contract is strictly smaller than in the flat contract, which proves uniqueness.

1b) For $\underline{\theta} = 0$: If $b \in (-I, 0)$, then $\theta_1^* = 0$ and $\theta_2^* = 1$, i.e., the optimal contract is $X_A^*(\theta)$ for all θ, consistent with $\hat{X}_-(\theta)$. If $b = -I$, then $\theta_2^* = 1$ and $\theta_1^* \in [0, 1]$, i.e., multiple optimal contracts exist (including some discontinuous contracts, as shown before). The flat contract given by $\hat{X}_{flat}(\theta)$ is one of the optimal contracts.

2) If $b \in [0, I)$, we have shown that $\theta_1^* = \underline{\theta}$ for any $\underline{\theta} \geq 0$, and hence we need to choose $\theta_2 \in [\underline{\theta}, 1]$. According to (30), $\frac{\partial P}{\partial \theta_2} > 0$ for $\theta_2 < \frac{I-b}{I+b}$ and $\frac{\partial P}{\partial \theta_2} < 0$ for $\theta_2 > \frac{I-b}{I+b}$. Since $b \geq 0$, $\frac{I-b}{I+b} < 1$. Also, $\frac{I-b}{I+b} \geq \underline{\theta} \Leftrightarrow b \leq \frac{1-\underline{\theta}}{1+\underline{\theta}} I$. Hence, if $b \geq \frac{1-\underline{\theta}}{1+\underline{\theta}} I$, then $\frac{\partial P}{\partial \theta_2} < 0$ for $\theta_2 > \underline{\theta}$, and hence (29) is maximized at $\theta_2 = \underline{\theta}$. Thus, for $b \geq \frac{1-\underline{\theta}}{1+\underline{\theta}} I$, the optimal contract is flat, which gives $\hat{X}_{flat}(\theta)$. Finally, if $b \in (0, \frac{1-\underline{\theta}}{1+\underline{\theta}} I]$, then (29) is increasing in θ_2 up to $\frac{I-b}{I+b}$ and decreasing after that. Hence, (29) is maximized at $\theta_2 = \frac{I-b}{I+b}$. Combined with $\theta_1 = \underline{\theta}$ and continuity of the contract, this gives $\hat{X}_+(\theta)$.

Case 3: $b < -I$. We show that the optimal contract is flat with $\hat{X}(\theta) = \frac{\beta}{\beta-1} \frac{2I}{\underline{\theta}+1}$. The proof proceeds in three steps. First, we show that the optimal contract cannot have any strictly decreasing regions and hence can only consist of flat regions. Second, we show that any contract with two flat regions is strictly dominated by a completely flat contract. Third, we show that any contract with at least three flat regions cannot be optimal. Combined, these steps imply that the optimal contract can only have one flat region, i.e., is completely flat. Combining this with (23) gives $\hat{X}_{flat}(\theta)$ and completes the proof of this case.

Step 1: *The optimal contract cannot have any strictly decreasing regions.*

Consider a contract with a strictly decreasing region. According to Lemma IA.3, any strictly decreasing region is characterized by $\hat{X}(\theta) = X_A^*(\theta) = \frac{\beta}{\beta-1} \frac{I-b}{\theta}$. Consider θ_1 and θ_2 such that $\hat{X}(\theta) = X_A^*(\theta)$ for $\theta \in [\theta_1, \theta_2]$. For any $\hat{\theta}_2 \in (\theta_1, \theta_2)$, consider a contract $C_2 = \left\{ \hat{X}_2(\theta), \theta \in \Theta \right\}$, defined as

$$\hat{X}_2(\theta) = \begin{cases} \hat{X}(\theta), & \text{if } \theta \in [\underline{\theta}, \theta_1] \cup [\theta_2, 1], \\ X_A^*(\theta), & \text{if } \theta \in \left[\theta_1, \hat{\theta}_2\right), \\ X_A^*(\hat{\theta}_2), & \text{if } \theta \in \left(\hat{\theta}_2, \hat{\theta}\right], \\ X_A^*(\theta_2), & \text{if } \theta \in \left(\hat{\theta}, \theta_2\right), \end{cases}$$

where $\hat{\theta} = \hat{\theta}(\hat{\theta}_2)$ satisfies

$$\text{(31)} \qquad \frac{\hat{\theta} X_A^*(\hat{\theta}_2) - I + b}{X_A^*(\hat{\theta}_2)^\beta} = \frac{\hat{\theta} X_A^*(\theta_2) - I + b}{X_A^*(\theta_2)^\beta}.$$

(such $\hat{\theta}$ always exists and lies between $\hat{\theta}_2$ and $\hat{\theta}_1$ for the same reason as in contract C_1). Intuitively, contract C_2 is the same as contract C, except that it substitutes a subset $\left[\hat{\theta}_2, \theta_2\right]$ of the decreasing region with a piecewise flat region with a discontinuity at $\hat{\theta}$. Because contract C is incentive-compatible and $\hat{\theta}$ satisfies (31), contract C_2 is incentive-compatible too. Below we show that the payoff to the principal from contract C_2 exceeds the payoff to the principal from contract C for $\hat{\theta}_2$ very close to θ_2. Because $\hat{X}_2(\theta) = \hat{X}(\theta)$ for $\theta \leq \theta_1$ and $\theta \geq \theta_2$, it is enough to restrict attention to the payoff in the range $\theta \in (\theta_1, \theta_2)$. The payoff to the principal from contract C_2 in this range, divided by $X(0)^\beta \frac{1}{1-\underline{\theta}}$, is

$$\int_{\theta_1}^{\hat{\theta}_2} \frac{\theta X_A^*(\theta) - I}{X_A^*(\theta)^\beta} d\theta + \int_{\hat{\theta}_2}^{\hat{\theta}(\hat{\theta}_2)} \frac{\theta X_A^*(\hat{\theta}_2) - I}{X_A^*(\hat{\theta}_2)^\beta} d\theta + \int_{\hat{\theta}(\hat{\theta}_2)}^{\theta_2} \frac{\theta X_A^*(\theta_2) - I}{X_A^*(\theta_2)^\beta} d\theta. \tag{32}$$

Following the same arguments as for the derivative of (25) with respect to $\tilde{\theta}_2$ above, we can check that the derivative of (32) with respect to $\hat{\theta}_2$ is given by $\beta \frac{\left(\hat{\theta}-\hat{\theta}_2\right)^2}{2\hat{\theta}_2^2} X_A^*(\hat{\theta}_2)^{-\beta}(I+b)$, which is strictly negative at any point $\hat{\theta}_2 < \theta_2$ for $b < -I$. By the mean value theorem, if $U_P(\hat{\theta}_2)$ stands for the expected principal's utility from contract C, then $\frac{U_P(\hat{\theta}_2)-U_P(\theta_2)}{\hat{\theta}_2-\theta_2} = U_P'\left(\tilde{\theta}_2\right) < 0$ for some $\tilde{\theta}_2 \in \left(\hat{\theta}_2, \theta_2\right)$, and hence a deviation from contract C to contract C_2 is beneficial for the principal. Hence, contract C cannot be optimal for $b < -I$. This result implies that any optimal contract must consist only of flat regions.

Step 2: *Any contract with two flat regions is dominated by a contract with one flat region.*

Consider a contract with two flat regions: Types $\left[\underline{\theta}, \hat{\theta}\right]$ pick exercise at $\hat{X}_L$, and types $\left[\hat{\theta}, 1\right]$ pick exercise at $\hat{X}_H < \hat{X}_L$. Type $\hat{\theta} \in (\underline{\theta}, 1)$ satisfies

$$\frac{\hat{\theta}\hat{X}_L - I + b}{\hat{X}_L^\beta} = \frac{\hat{\theta}\hat{X}_H - I + b}{\hat{X}_H^\beta}. \tag{33}$$

Consider an alternative contract with $\hat{X}(\theta) = \hat{X}_H$ for all θ. The difference between the principal's value under this pooling contract and her value under the original contract, divided by $X(0)^\beta$, is given by

$$\begin{aligned} \Delta U &= \textstyle\int_{\underline{\theta}}^1 \frac{\theta \hat{X}_H - I}{\hat{X}_H^\beta} \frac{d\theta}{1-\underline{\theta}} - \left[\int_{\underline{\theta}}^{\hat{\theta}} \frac{\theta \hat{X}_L - I}{\hat{X}_L^\beta} \frac{d\theta}{1-\underline{\theta}} + \int_{\hat{\theta}}^1 \frac{\theta \hat{X}_H - I}{\hat{X}_H^\beta} \frac{d\theta}{1-\underline{\theta}}\right] = \int_{\underline{\theta}}^{\hat{\theta}} \left(\frac{\theta \hat{X}_H - I}{\hat{X}_H^\beta} - \frac{\theta \hat{X}_L - I}{\hat{X}_L^\beta}\right) \frac{d\theta}{1-\underline{\theta}} \\ &= \frac{\hat{\theta}-\underline{\theta}}{1-\underline{\theta}} \left(\frac{\frac{\hat{\theta}+\underline{\theta}}{2}\hat{X}_H - I}{\hat{X}_H^\beta} - \frac{\frac{\hat{\theta}+\underline{\theta}}{2}\hat{X}_L - I}{\hat{X}_L^\beta}\right) = \frac{\hat{\theta}-\underline{\theta}}{1-\underline{\theta}} \left(\frac{\frac{\hat{\theta}}{2}\hat{X}_H - I}{\hat{X}_H^\beta} - \frac{\frac{\hat{\theta}}{2}\hat{X}_L - I}{\hat{X}_L^\beta} + \frac{\underline{\theta}}{2}\left(\frac{1}{\hat{X}_H^{\beta-1}} - \frac{1}{\hat{X}_L^{\beta-1}}\right)\right) \end{aligned} \tag{34}$$

Using (33) and the fact that $b \leq -I$,

$$\frac{\hat{\theta}\hat{X}_H - I}{\hat{X}_H^\beta} - \frac{\hat{\theta}\hat{X}_L - I}{\hat{X}_L^\beta} = b\left(\frac{1}{\hat{X}_L^\beta} - \frac{1}{\hat{X}_H^\beta}\right) \geq I\left(\frac{1}{\hat{X}_H^\beta} - \frac{1}{\hat{X}_L^\beta}\right) \Leftrightarrow \frac{\frac{\hat{\theta}}{2}\hat{X}_H - I}{\hat{X}_H^\beta} - \frac{\frac{\hat{\theta}}{2}\hat{X}_L - I}{\hat{X}_L^\beta} \geq 0,$$

and the inequalities are strict if $b < -I$. Combining this with $\hat{X}_H < \hat{X}_L$ and using (34), implies that $\Delta U \geq 0$ and $\Delta U > 0$ if at least one of $b < -I$ or $\underline{\theta} > 0$ holds. Thus, the contract with two flat regions is dominated by a contract with one flat region.

Step 3: *Any contract with at least three flat regions cannot be optimal.*

The proof of this step is similar to the proof of Step 3 in Proposition 4 in Melumad and Shibano (1991) for the payoff specification in our model. Suppose, on the contrary, that the optimal contract $X(\theta)$ has at least three flat regions. Consider three adjacent steps, $X_L > X_M > X_H$, of the assumed optimal contract. Types $\left(\hat{\theta}_0, \hat{\theta}_1\right)$ pick exercise at X_L, types $\left(\hat{\theta}_1, \hat{\theta}_2\right)$ pick exercise at X_M, and types $\left(\hat{\theta}_2, \hat{\theta}_3\right)$ pick exercise at X_H, where $\underline{\theta} \leq \hat{\theta}_0 < \hat{\theta}_1 < \hat{\theta}_2 < \hat{\theta}_3 \leq 1$. Incentive compatibility implies that types $\hat{\theta}_1$ and $\hat{\theta}_2$ satisfy

$$\frac{\hat{\theta}_1 X_L - I + b}{X_L^\beta} = \frac{\hat{\theta}_1 X_M - I + b}{X_M^\beta} \Leftrightarrow \hat{\theta}_1 = \frac{(I-b)\left(X_M^{-\beta} - X_L^{-\beta}\right)}{X_M^{1-\beta} - X_L^{1-\beta}}, \tag{35}$$

$$\frac{\hat{\theta}_2 X_M - I + b}{X_M^\beta} = \frac{\hat{\theta}_2 X_H - I + b}{X_H^\beta} \Leftrightarrow \hat{\theta}_2 = \frac{(I-b)\left(X_H^{-\beta} - X_M^{-\beta}\right)}{X_H^{1-\beta} - X_M^{1-\beta}}. \tag{36}$$

Consider an alternative contract with $\tilde{X}(\theta) = X_L$ for types $(\hat{\theta}_0, y)$, $\tilde{X}(\theta) = X_H$ for types $(y, \hat{\theta}_3)$, and $\tilde{X}(\theta) = X(\theta)$ otherwise, where $y \in (\hat{\theta}_1, \hat{\theta}_2)$ satisfies

$$\frac{yX_L - I + b}{X_L^\beta} = \frac{yX_H - I + b}{X_H^\beta} \Leftrightarrow y = \frac{(I-b)\left(X_H^{-\beta} - X_L^{-\beta}\right)}{X_H^{1-\beta} - X_L^{1-\beta}}. \tag{37}$$

This contract is incentive-compatible. The difference between the principal's value under this contract and the original contract, divided by $X(0)^\beta \frac{1}{1-\underline{\theta}}$, is given by

$$\begin{aligned}
&\int_{\hat{\theta}_1}^{y} \left(\tfrac{\theta X_L - I}{X_L^\beta} - \tfrac{\theta X_M - I}{X_M^\beta}\right) d\theta + \int_{y}^{\hat{\theta}_2} \left(\tfrac{\theta X_H - I}{X_H^\beta} - \tfrac{\theta X_M - I}{X_M^\beta}\right) d\theta \\
&= \left(y - \hat{\theta}_1\right)\left(\tfrac{\frac{y+\hat{\theta}_1}{2} X_L - I}{X_L^\beta} - \tfrac{\frac{y+\hat{\theta}_1}{2} X_M - I}{X_M^\beta}\right) + \left(\hat{\theta}_2 - y\right)\left(\tfrac{\frac{y+\hat{\theta}_2}{2} X_H - I}{X_H^\beta} - \tfrac{\frac{y+\hat{\theta}_2}{2} X_M - I}{X_M^\beta}\right).
\end{aligned}$$

Using the left equalities of (35) and (36), we can rewrite this as

$$\begin{aligned}
&\left(y - \hat{\theta}_1\right)\left(\tfrac{y}{2}\left(\tfrac{X_L}{X_L^\beta} - \tfrac{X_M}{X_M^\beta}\right) + \tfrac{-I+b}{2X_M^\beta} - \tfrac{-I+b}{2X_L^\beta} - I\left(\tfrac{1}{X_L^\beta} - \tfrac{1}{X_M^\beta}\right)\right) \\
&+ \left(\hat{\theta}_2 - y\right)\left(\tfrac{y}{2}\left(\tfrac{X_H}{X_H^\beta} - \tfrac{X_M}{X_M^\beta}\right) - I\left(\tfrac{1}{X_H^\beta} - \tfrac{1}{X_M^\beta}\right) + \tfrac{-I+b}{2X_M^\beta} - \tfrac{-I+b}{2X_H^\beta}\right).
\end{aligned}$$

Plugging in the values for y, $\hat{\theta}_1$, and $\hat{\theta}_2$ from the right equalities of (35), (36), and (37), we get

$$\frac{I-b}{X_M^{1-\beta}-X_L^{1-\beta}}\frac{\Sigma}{X_H^{1-\beta}-X_L^{1-\beta}}\left(\frac{\frac{(I-b)}{2}\left(X_H^{-\beta}-X_L^{-\beta}\right)\left(X_L^{1-\beta}-X_M^{1-\beta}\right)}{X_H^{1-\beta}-X_L^{1-\beta}}+\frac{b+I}{2}\left(X_M^{-\beta}-X_L^{-\beta}\right)\right)$$
$$+\frac{I-b}{X_H^{1-\beta}-X_M^{1-\beta}}\frac{\Sigma}{X_H^{1-\beta}-X_L^{1-\beta}}\left(\frac{\frac{(I-b)}{2}\left(X_H^{-\beta}-X_L^{-\beta}\right)\left(X_H^{1-\beta}-X_M^{1-\beta}\right)}{X_H^{1-\beta}-X_L^{1-\beta}}+\frac{b+I}{2}\left(X_M^{-\beta}-X_H^{-\beta}\right)\right),$$

where $\Sigma = X_H^{-\beta}X_M^{-\beta}(X_M - X_H) + X_L^{-\beta}X_H^{-\beta}(X_H - X_L) + X_M^{-\beta}X_L^{-\beta}(X_L - X_M)$. Rearranging, we obtain

$$\begin{aligned}&\frac{\Sigma\,(I-b)\,(b+I)}{2(X_H^{1-\beta}-X_L^{1-\beta})}\left(\frac{X_M^{-\beta}-X_L^{-\beta}}{X_M^{1-\beta}-X_L^{1-\beta}}+\frac{X_M^{-\beta}-X_H^{-\beta}}{X_H^{1-\beta}-X_M^{1-\beta}}\right)\\ =\;&\frac{\frac{1}{2}\Sigma\left(I^2-b^2\right)}{(X_H^{1-\beta}-X_L^{1-\beta})}\frac{-\Sigma}{(X_M^{1-\beta}-X_L^{1-\beta})(X_H^{1-\beta}-X_M^{1-\beta})},\end{aligned}$$

which is strictly positive because $I^2 - b^2 < 0$ and $X_L^{1-\beta} < X_M^{1-\beta} < X_H^{1-\beta}$. Thus, contract $X(\theta)$ is strictly dominated by contract $\tilde{X}(\theta)$ and hence cannot be optimal. ■

Supplementary analysis for the proof of Proposition 1. Proof that the principal's ex-ante IC constraint is satisfied. Let $V_P^c\left(X, \hat{\theta}; \hat{\theta}^*\right)$ denote the expected value to the principal in the equilibrium with continuous exercise (up to a cutoff) if the current value of $X(t)$ is X and the current belief is that $\theta \in \left[\underline{\theta}, \hat{\theta}\right]$ for some $\hat{\theta} > \hat{\theta}^*$. If the agent's type is $\theta > \hat{\theta}^*$, exercise occurs at threshold $\frac{\beta}{\beta-1}\frac{I-b}{\theta}$, and the principal's payoff upon exercise is $\frac{\beta}{\beta-1}(I-b) - I$. If $\theta < \hat{\theta}^*$, exercise occurs at threshold X^*. Hence,

$$\begin{aligned}\left(\hat{\theta}-\underline{\theta}\right)V_P^c\left(X,\hat{\theta};\hat{\theta}^*\right) &= \left(\frac{X}{X^*}\right)^{\beta}\int_{\underline{\theta}}^{\hat{\theta}^*}(\theta X^* - I)\,d\theta\\ &\quad +X^{\beta}\int_{\hat{\theta}^*}^{\hat{\theta}}\left(\frac{\beta}{\beta-1}\frac{I-b}{\theta}\right)^{-\beta}\left(\frac{\beta}{\beta-1}(I-b)-I\right)d\theta.\end{aligned}$$

Given belief $\theta \in \left[\underline{\theta}, \hat{\theta}\right]$, the principal can either wait and get $V_P^c\left(X, \hat{\theta}; \hat{\theta}^*\right)$ or exercise immediately and get $X\frac{\underline{\theta}+\hat{\theta}}{2} - I$. The current value of $X(t)$ satisfies $X(t) \leq X_A^*(\hat{\theta})$ because otherwise, the principal's belief would not be that $\theta \in \left[\underline{\theta}, \hat{\theta}\right]$. Hence, the ex-ante IC condition requires that for any $\hat{\theta} > \hat{\theta}^*$, $V_P^c\left(X, \hat{\theta}; \hat{\theta}^*\right) \geq X\frac{\underline{\theta}+\hat{\theta}}{2} - I$ for any $X \leq X_A^*(\hat{\theta})$. Because $X^{-\beta}V_P^c\left(X, \hat{\theta}; \hat{\theta}^*\right)$ does not depend on X, this condition is equivalent to

$$X^{-\beta}V_P^c\left(X,\hat{\theta};\hat{\theta}^*\right) \geq \max_{X\in(0,X_A^*(\hat{\theta})]}\frac{1}{X^{\beta}}\left(X\frac{\underline{\theta}+\hat{\theta}}{2}-I\right). \tag{38}$$

The function $\frac{1}{X^{\beta}}\left(X\frac{\underline{\theta}+\hat{\theta}}{2}-I\right)$ is inverse U-shaped in X and has an unconditional maximum at $\frac{\beta}{\beta-1}\frac{2I}{\underline{\theta}+\hat{\theta}}$, which is strictly greater than $X_A^*(\hat{\theta})$ for any $\hat{\theta}>\frac{I-b}{I+b}\underline{\theta}=\hat{\theta}^*$. Because $X^{-\beta}V_P^c\left(X,\hat{\theta};\hat{\theta}^*\right)$ does not depend on X, (38) is equivalent to

$$X^{-\beta}V_P^c\left(X,\hat{\theta};\hat{\theta}^*\right)\geq X_A^*(\hat{\theta})^{-\beta}\left(X_A^*(\hat{\theta})\frac{\underline{\theta}+\hat{\theta}}{2}-I\right).$$

Suppose there exists $\hat{\theta}$ for which the ex-ante IC constraint is violated, i.e.,

(39)
$$\int_{\underline{\theta}}^{\hat{\theta}^*}(X^*)^{-\beta}(\theta X^*-I)\,d\theta+\int_{\hat{\theta}^*}^{\hat{\theta}}\left(\frac{\beta}{\beta-1}\frac{I-b}{\theta}\right)^{-\beta}\left(\frac{\beta}{\beta-1}(I-b)-I\right)d\theta$$
$$<\left(\hat{\theta}-\underline{\theta}\right)X_A^*(\hat{\theta})^{-\beta}\left(X_A^*(\hat{\theta})\frac{\underline{\theta}+\hat{\theta}}{2}-I\right).$$

We show that this implies that the contract derived in Lemma 1 cannot be optimal, which is a contradiction. In particular, denote the contract from the second part of Lemma 1 by $\hat{X}_-(\theta)$. Then (39) implies that the contract $\hat{X}_-(\theta)$ is dominated by the contract with continuous exercise at $X_A^*(\theta)$ for $\theta\geq\hat{\theta}$ and exercise at $X_A^*(\hat{\theta})$ for $\theta\leq\hat{\theta}$. Indeed, the principal's expected utility under the contract $\hat{X}_-(\theta)$, divided by $X(0)^{\beta}$, is

(40)
$$\int_{\underline{\theta}}^{\hat{\theta}^*}(X^*)^{-\beta}(\theta X^*-I)\,d\theta+\int_{\hat{\theta}^*}^{1}\left(\frac{\beta}{\beta-1}\frac{I-b}{\theta}\right)^{-\beta}\left(\frac{\beta}{\beta-1}(I-b)-I\right)d\theta.$$

Similarly, the principal's expected utility under the modified contract (also divided by $X(0)^{\beta}$), where $\frac{I-b}{I+b}\underline{\theta}$ in $\hat{X}_-(\theta)$ is replaced by $\hat{\theta}$, and the cutoff $\frac{\beta}{\beta-1}\frac{I+b}{\underline{\theta}}$ in $\hat{X}_-(\theta)$ is replaced by $X_A^*(\hat{\theta})=\frac{\beta}{\beta-1}\frac{I-b}{\hat{\theta}}$, is given by

(41)
$$\int_{\underline{\theta}}^{\hat{\theta}}X_A^*(\hat{\theta})^{-\beta}\left(\theta X_A^*(\hat{\theta})-I\right)d\theta+\int_{\hat{\theta}}^{1}\left(\frac{\beta}{\beta-1}\frac{I-b}{\theta}\right)^{-\beta}\left(\frac{\beta}{\beta-1}(I-b)-I\right)d\theta.$$

Combining (40) and (41), it is easy to see that the contract with continuous exercise up to the cutoff $\hat{\theta}$ dominates the contract $\hat{X}_-(\theta)$ if and only if (39) is satisfied. Hence, the ex-ante IC constraint is indeed satisfied. ■

Supplementary analysis for the proof of Proposition 2. Part 1. Derivation of the principal's value function in the ω-equilibrium, $V_P(X(t),1;\omega)$.

The principal's value function $V_P(X(t),1;\omega)$ satisfies

(42)
$$rV_P(X,1;\omega)=\alpha XV_{P,X}(X,1;\omega)+\frac{1}{2}\sigma^2X^2V_{P,XX}(X,1;\omega).$$

The value matching condition is:

$$V_P(Y(\omega), 1; \omega) = \int_\omega^1 (\theta Y(\omega) - I)\, d\theta + \omega V_P(Y(\omega), \omega; \omega). \tag{43}$$

The intuition behind (43) is as follows. With probability $1-\omega$, θ is above ω. In this case, the agent recommends exercise, and the principal follows the recommendation. The payoff of the principal, given θ, is $\theta Y(\omega) - I$. With probability ω, θ is below ω, so the agent recommends against exercise, and the option is not exercised. The continuation payoff of the principal in this case is $V_P(Y(\omega), \omega; \omega)$. Solving (42) subject to (43), we obtain

$$V_P(X, 1; \omega) = \left(\frac{X}{Y(\omega)}\right)^\beta \left(\int_\omega^1 (\theta Y(\omega) - I)\, d\theta + \omega V_P(Y(\omega), \omega; \omega)\right). \tag{44}$$

By stationarity,

$$V_P(Y(\omega), \omega; \omega) = V_P(\omega Y(\omega), 1; \omega). \tag{45}$$

Evaluating (44) at $X = \omega Y(\omega)$ and using the stationarity condition (45), we obtain:

$$V_P(\omega Y(\omega), 1; \omega) = \omega^\beta \left[\frac{1}{2}\left(1-\omega^2\right) Y(\omega) - (1-\omega) I\right] + \omega^{\beta+1} V_P(\omega Y(\omega), 1; \omega).$$

Therefore,

$$V_P(\omega Y(\omega), 1; \omega) = \frac{\omega^\beta (1-\omega)}{1-\omega^{\beta+1}} \left[\frac{1}{2}(1+\omega) Y(\omega) - I\right]. \tag{46}$$

Plugging (46) into (44), we obtain the principal's value function (A8).

Part 2. Existence of ω-equilibria for $b < 0$.

2a. Proof that $G(\omega) > 0$, where $G(\omega) \equiv \frac{\left(1-\omega^\beta\right)(I-b)}{\omega\left(1-\omega^{\beta-1}\right)} - \frac{\beta}{\beta-1}\frac{2(I-b)}{1+\omega}$. Note that $G(\omega) = \frac{2(I-b)}{1+\omega} g(\omega)$, where $g(\omega) \equiv \frac{\left(1-\omega^\beta\right)(1+\omega)}{2\left(\omega-\omega^\beta\right)} - \frac{\beta}{\beta-1}$. We have:

$$\begin{aligned} \lim_{\omega\to 1} g(\omega) &= \lim_{\omega\to 1} \frac{1-\omega^\beta-\beta\omega^{\beta-1}(1+\omega)}{2\left(1-\beta\omega^{\beta-1}\right)} - \frac{\beta}{\beta-1} = 0, \\ g'(\omega) &= \frac{\beta\left(\omega^{\beta-1}-\omega^{\beta+1}\right)+\omega^{2\beta}-1}{2\left(\omega-\omega^\beta\right)^2}, \end{aligned}$$

where the first limit holds by l'Hopital's rule. Denote the numerator of $g'(\omega)$ by $h(\omega) \equiv \omega^{2\beta} - \beta\omega^{\beta+1} + \beta\omega^{\beta-1} - 1$. Function $h(\omega)$ is a generalized polynomial. By an extension of Descartes' Rule of Signs to generalized polynomials (Laguerre, 1883),[17] the number of positive roots of $h(\omega)$,

[17] See Theorem 3.1 in Jameson (2006).

counted with their orders, does not exceed the number of sign changes of coefficients of $h(\omega)$, i.e., three. Because $\omega = 1$ is the root of $h(\omega)$ of order three and $h(0) < 0$, then $h(\omega) < 0$ for all $\omega \in [0,1)$, and hence $g'(\omega) < 0$ for all $\omega \in [0,1)$. Combined with $\lim_{\omega\to 1} g(\omega) = 0$, this implies $g(\omega) > 0$ and hence $G(\omega) > 0$ for all $\omega \in [0,1)$.

2b. Proof of Step 1: *If $b < 0$, $V_P(X,1;\omega)$ is strictly increasing in ω for any $\omega \in (0,1)$.*

Because $V_P(X,1;\omega)$ is proportional to X^β, it is enough to prove the statement for $X = 1$. We can re-write $V_P(1,1;\omega)$ as $2^{-\beta} f_1(\omega) f_2(\omega)$, where

$$f_1(\omega) \equiv \frac{(1-\omega)(1+\omega)^\beta}{1-\omega^{\beta+1}} \text{ and } f_2(\omega) \equiv \frac{\frac{1}{2}(1+\omega)Y(\omega) - I}{\left(\frac{1}{2}(1+\omega)Y(\omega)\right)^\beta}. \tag{47}$$

Since, as shown above, $Y(\omega) > \frac{\beta}{\beta-1}\frac{2I}{1+\omega}$ for $\omega < 1$, then $\frac{1}{2}(1+\omega)Y(\omega) > \frac{\beta}{\beta-1}I > I$, and hence $f_2(\omega) > 0$ for $\omega < 1$. Because $f_1(\omega) > 0$ and $f_2(\omega) > 0$ for any $\omega \in (0,\omega^*)$, a sufficient condition for $V_P(1,1;\omega)$ to be increasing is that both $f_1(\omega)$ and $f_2(\omega)$ are increasing for $\omega \in (0,\omega^*)$.

First, consider $f_2(\omega)$. As an auxiliary result, we prove that $(1+\omega)Y(\omega)$ is strictly decreasing in ω. This follows from the fact that

$$\frac{\partial((1+\omega)Y(\omega))}{\partial\omega} = (I-b)\frac{-1+\beta\omega^{\beta-1}-\beta\omega^{\beta+1}+\omega^{2\beta}}{(\omega-\omega^\beta)^2}$$

and that as shown above, the numerator, $h(\omega)$, is strictly negative for all $\omega \in [0,1)$. Next,

$$f_2'(\omega) = \frac{(\beta-1)(1+\omega)}{4\left(\frac{1}{2}(1+\omega)Y(\omega)\right)^{\beta+1}}\left(\frac{\beta}{\beta-1}\frac{2I}{1+\omega} - Y(\omega)\right)\frac{\partial((1+\omega)Y(\omega))}{\partial\omega}.$$

Because $Y(\omega) > \frac{\beta}{\beta-1}\frac{2I}{1+\omega}$ for $\omega < 1$ as shown above, and because $(1+\omega)Y(\omega)$ is strictly decreasing in ω, $f_2'(\omega) > 0$ for any $\omega \in (0,\omega^*)$.

Second, consider $f_1(\omega)$. Note that

$$f_1'(\omega) = \frac{(1+\omega)^{\beta-1}}{1-\omega^{\beta+1}}\frac{\beta-1-(\beta+1)\omega+(\beta+1)\omega^\beta-(\beta-1)\omega^{\beta+1}}{1-\omega^{\beta+1}}.$$

Denote the numerator of the second fraction by $d(\omega) \equiv -(\beta-1)\omega^{\beta+1} + (\beta+1)\omega^\beta - (\beta+1)\omega + \beta - 1$. By an extension of Descartes' Rule of Signs to generalized polynomials, the number of positive roots of $d(\omega)$ does not exceed the number of sign changes of coefficients of $d(\omega)$, i.e., three. It is easy to show that $d(1) = d'(1) = d''(1) = 0$. Hence, $\omega = 1$ is the root of $d(\omega) = 0$ of order three, and $d(\omega)$ does not have roots other than $\omega = 1$. Since $d(0) = \beta - 1 > 0$, this implies that for any $\omega \in (0,1)$, $d(\omega) > 0$. Hence, $f_1'(\omega) > 0$, which completes the proof of this step.

2c. Proof of Step 2: $\lim_{\omega\to 1} V_P(X,1;\omega) = V_P^c(X,1)$.

According to (47), $V_P(X,1;\omega) = 2^{-\beta}X^\beta f_1(\omega) f_2(\omega)$. By l'Hopital's rule, $\lim_{\omega\to 1} f_1(\omega) =$

$\frac{2^{\beta}}{\beta+1}$, $\lim_{\omega\to 1} Y(\omega) = \frac{\beta}{\beta-1}(I-b)$, and hence $\lim_{\omega\to 1} f_1(\omega) = (\frac{\beta}{\beta-1}(I-b)-I)(\frac{\beta}{\beta-1}(I-b))^{-\beta}$.Using (A1), it is easy to see that $\lim_{\omega\to 1} V_P(X,1;\omega) = 2^{-\beta}X^{\beta}\lim_{\omega\to 1} f_1(\omega)\lim_{\omega\to 1} f_1(\omega) = V_P^c(X,1)$.

2d. Proof of Step 3. *Suppose $-I < b < I$. For ω close enough to zero, the ex-ante IC condition (A12) does not hold.*

The function $X^{-\beta}\left(\frac{1}{2}X - I\right)$ is inverse U-shaped and has a maximum at $\bar{X}_u = \frac{\beta}{\beta-1}2I$. When ω is close to zero, $Y(\omega) = \frac{\left(1-\omega^{\beta}\right)(I-b)}{\omega\left(1-\omega^{\beta-1}\right)} \to +\infty$, and hence $\max_{X\in(0,Y(\omega)]} X^{-\beta}\left(\frac{1}{2}X - I\right) = \bar{X}_u^{-\beta}\left(\frac{1}{2}\bar{X}_u - I\right)$. Hence, we can rewrite (A12) as $X^{-\beta}V_P(X,1;\omega) \geq \bar{X}_u^{-\beta}\left(\frac{1}{2}\bar{X}_u - I\right)$, and it is easy to show that it is equivalent to

$$\text{(48)}\qquad \left(\omega-\omega^{\beta}\right)^{\beta-1} H(\omega) \geq (I-b)^{\beta}\left(1-\omega^{\beta+1}\right)\left(1-\omega^{\beta}\right)^{\beta},$$

where

$$H(\omega) \equiv 2^{\beta-1}\beta^{\beta}\left(\frac{I}{\beta-1}\right)^{\beta-1}(1-\omega)\left(I(1-\omega)\left(1+\omega^{\beta}\right) - b(1+\omega)\left(1-\omega^{\beta}\right)\right).$$

Since $H(0) > 0$, then as $\omega \to 0$, the left-hand side of (48) converges to zero, while the right-hand side converges to $(I-b)^{\beta} > 0$. Hence, for ω close enough to 0, the ex-ante IC condition is violated.

2e. Proof of Step 4. *Suppose $-I < b < I$. Then (A12) is satisfied for any $\omega \geq \bar{\omega}$, where $\bar{\omega}$ is the unique solution to $Y(\omega) = \bar{X}_u$. For any $\omega < \bar{\omega}$, (A12) is satisfied if and only if $X^{-\beta}V_P(X,1;\omega) \geq \bar{X}_u^{-\beta}\left(\frac{1}{2}\bar{X}_u - I\right)$.*

Note that for any $b > -I$, $\lim_{\omega\to 1} Y(\omega) = \frac{\beta(I-b)}{\beta-1} < \frac{\beta}{\beta-1}2I = \bar{X}_u$, and hence there exists a unique $\bar{\omega}$ such that $Y(\omega) \leq \bar{X}_u \Leftrightarrow \omega \geq \bar{\omega}$. Hence, (A12) becomes

$$\text{(49)}\qquad X^{-\beta}V_P(X,1;\omega) \geq \bar{X}_u^{-\beta}\left(\frac{1}{2}\bar{X}_u - I\right) \text{ for } \omega \leq \bar{\omega},$$

$$\text{(50)}\qquad X^{-\beta}V_P(X,1;\omega) \geq Y(\omega)^{-\beta}\left(\frac{1}{2}Y(\omega) - I\right) \text{ for } \omega \geq \bar{\omega}.$$

Suppose that (50) is satisfied for some $\tilde{\omega} \geq \bar{\omega}$. Because $Y(\omega)$ is decreasing, $Y(\tilde{\omega}) \geq Y(\omega)$ for $\omega \geq \tilde{\omega}$. Because $X^{-\beta}\left(\frac{1}{2}X - I\right)$ is increasing for $X \leq \bar{X}_u$ and because $Y(\omega) \leq Y(\tilde{\omega}) \leq Y(\bar{\omega}) = \bar{X}_u$ for $\omega \geq \tilde{\omega} \geq \bar{\omega}$, we have $Y(\omega)^{-\beta}\left(\frac{1}{2}Y(\omega) - I\right) \leq Y(\tilde{\omega})^{-\beta}\left(\frac{1}{2}Y(\tilde{\omega}) - I\right)$ for any $\omega \geq \tilde{\omega}$. On the other hand, according to Step 1, $X^{-\beta}V_P(X,1;\tilde{\omega}) \leq X^{-\beta}V_P(X,1;\omega)$ for any $\omega \geq \tilde{\omega}$. Hence, if (50) is satisfied for $\tilde{\omega} \geq \bar{\omega}$, it is also satisfied for any $\omega \in [\tilde{\omega}, 1)$. Hence, to prove that (A12) is satisfied for any $\omega \geq \bar{\omega}$, it is sufficient to prove (50) for $\omega = \bar{\omega}$. Using (A8) and the fact that

$Y(\bar{\omega}) = \bar{X}_u$, (50) for $\omega = \bar{\omega}$ is equivalent to

$$\begin{aligned}
\frac{1-\bar{\omega}}{1-\bar{\omega}^{\beta+1}}\bar{X}_u^{-\beta}\left(\frac{1}{2}(1+\bar{\omega})\bar{X}_u - I\right) &\geq \bar{X}_u^{-\beta}\left(\frac{1}{2}\bar{X}_u - I\right) \\
&\Leftrightarrow \frac{1}{2}\bar{X}_u\left(\frac{1-\bar{\omega}^2}{1-\bar{\omega}^{\beta+1}} - 1\right) \geq I\left(\frac{1-\bar{\omega}}{1-\bar{\omega}^{\beta+1}} - 1\right) \\
&\Leftrightarrow \frac{1}{2I}\bar{X}_u \leq \frac{\bar{\omega}-\bar{\omega}^{\beta+1}}{\bar{\omega}^2-\bar{\omega}^{\beta+1}} \Leftrightarrow \frac{\beta}{\beta-1} \leq \frac{\bar{\omega}-\bar{\omega}^{\beta+1}}{\bar{\omega}^2-\bar{\omega}^{\beta+1}} \qquad (51)
\end{aligned}$$

Consider the function $Q(\omega) \equiv \frac{\omega-\omega^{\beta+1}}{\omega^2-\omega^{\beta+1}}$. Note that $Q'(\omega) < 0 \Leftrightarrow q(\omega) \equiv (\beta-1)\omega^\beta - \beta\omega^{\beta-1} + 1 > 0$. By an extension of Descartes' Rule of Signs to generalized polynomials (Laguerre, 1883), the number of positive roots of $q(\omega)$, counted with their orders, does not exceed the number of sign changes of coefficients of $q(\omega)$, i.e., two. Since $q(1) = q'(1) = 0$, $q(\omega)$ does not have any roots on $(0,\infty)$ other than 1. Since $q''(1) > 0$, we have $q(\omega) > 0$ for all $\omega \in (0,1)$, and hence $Q'(\omega) < 0$. By l'Hopital's rule, $\lim_{\omega\to 1} Q(\omega) = \frac{\beta}{\beta-1}$, and hence $\frac{\beta}{\beta-1} \leq Q(\omega)$ for any $\omega \in (0,1)$, which proves (51).

Part 3. Existence of ω-equilibria for $b > 0$.

3a. Proof that in the range $[0,1]$, equation $Y(\omega) = \frac{\beta}{\beta-1}\frac{2I}{\omega+1}$ has a unique solution $\omega^* \in (0,1)$, where ω^* decreases in b, $\lim_{b\to 0}\omega^* = 1$, and $\lim_{b\to I}\omega^* = 0$.

The equality $Y(\omega) = \frac{\beta}{\beta-1}\frac{2I}{\omega+1}$ is equivalent to

$$\omega = \frac{1}{\frac{\beta}{\beta-1}\frac{1-\omega^{\beta-1}}{1-\omega^\beta}\frac{2I}{I-b} - 1}. \qquad (52)$$

We next show that if $0 < b < I$, then in the range $[0,1]$, equation (52) has a unique solution $\omega^* \in (0,1)$, where ω^* decreases in b, $\lim_{b\to 0}\omega^* = 1$, and $\lim_{b\to I}\omega^* = 0$. We can rewrite (52) as

$$\begin{aligned}
\omega &= \frac{(\beta-1)\left(1-\omega^\beta\right)(I-b)}{\beta\left(1-\omega^{\beta-1}\right)2I - (\beta-1)\left(1-\omega^\beta\right)(I-b)} \\
&\Leftrightarrow \frac{2\beta I\left(\omega-\omega^\beta\right) + (\beta-1)(I-b)\left(\omega^{\beta+1}-\omega-1+\omega^\beta\right)}{\beta\left(1-\omega^{\beta-1}\right)2I - (\beta-1)\left(1-\omega^\beta\right)(I-b)} = 0.
\end{aligned}$$

Denote the left-hand side of the second equation as a function of ω by $l(\omega)$. The denominator of $l(\omega)$, $l_d(\omega)$, is nonnegative on $\omega \in [0,1]$ and equals zero only at $\omega = 1$. This follows from $l_d(0) = 2\beta I - (\beta-1)(I-b) > 0$, $l_d(1) = 0$, and $l_d'(\omega) = \beta(\beta-1)\omega^{\beta-2}(-2I + \omega(I-b)) < 0$.Therefore, $l(\omega) = 0$ if and only if the numerator of $l(\omega)$, $l_n(\omega)$, equals zero at $\omega \in (0,1)$. Since $b \in (0,I)$, then $l_n(0) = -(\beta-1)(I-b) < 0$,

$$\begin{aligned}
l_n'(\omega) &= 2\beta I\left(1-\beta\omega^{\beta-1}\right) + (\beta-1)(I-b)\left((\beta+1)\omega^\beta - 1 + \beta\omega^{\beta-1}\right), \\
l_n''(\omega) &= -2\beta^2(\beta-1)I\omega^{\beta-2} + (\beta-1)(I-b)\left(\beta(\beta+1)\omega^{\beta-1} + \beta(\beta-1)\omega^{\beta-2}\right),
\end{aligned}$$

and

$$l_n''(\omega) < 0 \Leftrightarrow (I-b)\left((\beta+1)\omega+\beta-1\right) < 2\beta I \Leftrightarrow \omega < \frac{(\beta+1)I+(\beta-1)b}{(\beta+1)(I-b)}.$$

Since $\frac{(\beta+1)I+(\beta-1)b}{(\beta+1)(I-b)} > 1$, $l_n''(\omega) < 0$ for any $\omega \in [0,1]$. Since $l_n'(0) = 2\beta I - (\beta-1)(I-b) > 0$ and $l_n'(1) = -2\beta(\beta-1)b < 0$, there exists $\hat{\omega} \in (0,1)$ such that $l_n(\omega)$ increases to the left of $\hat{\omega}$ and decreases to the right. Since $\lim_{\omega\to 1} l_n(\omega) = 0$, then $l_n(\hat{\omega}) > 0$, and hence $l_n(\omega)$ has a unique root ω^* on $(0,1)$.

Since the function $l_n(\omega)$ increases in b and is strictly increasing at the point ω^*, then ω^* decreases in b. To prove that $\lim_{b\to 0}\omega^* = 1$, it is sufficient to prove that for any small $\varepsilon > 0$, there exists $b(\varepsilon) > 0$ such that $l_n(1-\varepsilon) < 0$ for $b < b(\varepsilon)$. Since $l_n(\omega) > 0$ on $(\omega^*, 1)$, this would imply that $\omega^* \in (1-\varepsilon, 1)$, i.e., that ω^* is infinitely close to 1 when b is close to zero. Using the expression for $l_n(\omega)$, $l_n(\omega) < 0$ is equivalent to

$$\frac{2\beta}{\beta-1}\frac{\omega}{\omega+1}\frac{1-\omega^{\beta-1}}{1-\omega^{\beta}} < 1 - \frac{b}{I}. \tag{53}$$

Denote the left-hand side of (53) by $L(\omega)$. Note that $L(\omega)$ is increasing on $(0,1)$. Indeed, differentiating $L(\omega)$ and simplifying, $L'(\omega) > 0 \Leftrightarrow \Lambda(\omega) \equiv 1 - \omega^{2\beta} - \beta\omega^{\beta-1} + \beta\omega^{\beta+1} > 0$. The function $\Lambda(\omega)$ is decreasing on $(0,1)$ because $\Lambda'(\omega) < 0 \Leftrightarrow \varphi(\omega) \equiv -2\omega^{\beta+1} - (\beta-1) + (\beta+1)\omega^2 < 0$, where $\varphi'(\omega) > 0$ and $\varphi(1) = 0$. Since $\Lambda(\omega)$ is decreasing and $\Lambda(1) = 0$, then, indeed, $\Lambda(\omega) > 0$ and hence $L'(\omega) > 0$ for all $\omega \in (0,1)$. In addition, by l'Hopital's rule, $\lim_{\omega\to 1} L(\omega) = 1$. Hence, $L(1-\varepsilon) < 1$ for any $\varepsilon > 0$, and thus $l_n(1-\varepsilon) < 0$ for $b \in [0, I(1-L(1-\varepsilon)))$.

Finally, to prove that $\lim_{b\to I}\omega^* = 0$, it is sufficient to prove that for any small $\varepsilon > 0$, there exists $b(\varepsilon)$ such that $l_n(\varepsilon) > 0$ for $b > b(\varepsilon)$. Since $l_n(0) < 0$, this would imply that $\omega^* \in (0,\varepsilon)$ for $b > b(\varepsilon)$, i.e., that ω^* is infinitely close to zero when b is close to I. Based on (53), $l_n(\omega) > 0 \Leftrightarrow L(\omega) > 1 - \frac{b}{I}$. Then, for any $\varepsilon > 0$, if $b > I(1-L(\varepsilon))$, we get $1 - \frac{b}{I} < L(\varepsilon) \Leftrightarrow l_n(\varepsilon) > 0$, which completes the proof.

3b. Proof that $Y(\omega)$ is strictly decreasing in ω for $\omega \in (0,1)$. Note that

$$\frac{\partial Y(\omega)}{\partial \omega} = \frac{(I-b)}{\omega\left(\omega-\omega^{\beta}\right)^2}\left[-(\beta-1)\omega^{\beta+1} + \beta\omega^{\beta} - \omega\right],$$

where $\frac{(I-b)}{\omega\left(\omega-\omega^{\beta}\right)^2} > 0$. Thus, we need to show that $k(\omega) \equiv -(\beta-1)\omega^{\beta+1} + \beta\omega^{\beta} - \omega < 0$. According to an extension of Descartes' Rule of Signs to generalized polynomials (Laguerre, 1883), the number of positive roots of $k(\omega) = 0$, counted with their orders, does not exceed the number of change of signs of its coefficients, i.e., two. Since $k(1) = 0$, $k'(1) = 0$, and $k\text{"}(1) = -\beta(\beta-1) < 0$, $\omega = 1$ is a root of order two, and there are no other positive roots. Further, $k(0) = 0$ and $k'(0) = -1 < 0$. It follows that $k(0) = k(1) = 0$ and $k(\omega) < 0$ for all $\omega \in (0,1)$, and hence, indeed, $\frac{\partial Y(\omega)}{\partial \omega} < 0$.

3c. Proof of Step 5: *If* $b > 0$, $V_P(X, 1; \omega)$ *is strictly increasing in* ω *for any* $\omega \in (0, \omega^*)$.

The proof of this step is the same as the proof of Step 1 for the case $b < 0$ with the only difference: instead of relying on the inequality $Y(\omega) > \frac{\beta}{\beta-1}\frac{2I}{1+\omega}$ for all $\omega \in (0,1)$ as for the case $b < 0$ (which holds for $b < 0$), we rely on the inequality $Y(\omega) > \frac{\beta}{\beta-1}\frac{2I}{1+\omega}$ for all $\omega \in (0,\omega^*)$.

3d. Proof of Step 6: *If $0 < b < I$, then the ex-ante IC condition (A12) holds as a strict inequality for $\omega = \omega^*$.*

Using (A8) and $Y(\omega) = \frac{\beta}{\beta-1}\frac{2I}{1+\omega}$, we can rewrite $V_P(X,1;\omega^*)$ as $X^\beta K(\omega^*)$, where

$$K(\omega) \equiv \frac{1-\omega}{1-\omega^{\beta+1}}\left(\frac{\beta}{\beta-1}\frac{2I}{\omega+1}\right)^{-\beta}\frac{I}{\beta-1}.$$

Note that $K(0) = \bar{X}_u^{-\beta}\left(\frac{1}{2}\bar{X}_u - I\right)$ and that

$$K'(\omega) > 0 \Leftrightarrow \kappa(\omega) \equiv -(\beta-1)\,\omega^{\beta+1} + (\beta+1)\,\omega^\beta - (\beta+1)\,\omega + \beta - 1 > 0.$$

By an extension of Descartes' Rule of Signs to generalized polynomials, the number of positive roots of $\kappa(\omega)$, counted with their orders, does not exceed the number of change of signs of its coefficients, i.e., three. Note that $\omega = 1$ is the root of $\kappa(\omega)$ of order three: $\kappa(1) = \kappa'(1) = \kappa''(1) = 0$, and hence there are no other roots. Since $\kappa(0) = \beta - 1 > 0$, it follows that $\kappa(\omega) > 0$ and hence $K'(\omega) > 0$ for all $\omega \in [0,1)$. Therefore, $K(\omega)$ is strictly increasing in ω, which implies

(54) $$X^{-\beta}V_P(X,1;\omega^*) = K(\omega^*) > K(0) = \bar{X}_u^{-\beta}\left(\frac{1}{2}\bar{X}_u - I\right).$$

Because the function $X^{-\beta}\left(\frac{1}{2}X - I\right)$ achieves its global maximum at the point $\bar{X}_u$, (54) implies that (A12) holds as a strict inequality for $\omega = \omega^*$, which completes the proof of this step. ■

Figure B.1 illustrates the equilibria with partitioned exercise characterized in Proposition 2. We next compare the equilibria in Proposition 2 in terms of Pareto efficiency and derive some comparative statics results.

Proposition B.1. *If $b < 0$, the equilibrium with continuous exercise from Proposition 2 dominates all other possible equilibria in the Pareto sense: both the agent's expected payoff for each realization of θ and the principal's expected payoff are higher in this equilibrium than in any other equilibrium. If $b > 0$, the ω^*-equilibrium dominates other stationary equilibria with partitioned exercise in the following sense: both the principal's expected payoff and the ex-ante expected payoff of the agent before θ is realized are higher in the ω^*-equilibrium than in the ω-equilibrium for any $\omega < \omega^*$.*

Intuitively, when $b < 0$, the equilibrium with continuous exercise both implements the optimal mechanism for the principal and ensures that exercise occurs at the unconstrained optimal time of any type θ of the agent. When $b > 0$, the ω^*-equilibrium is the only equilibrium in which exercise is unbiased: since the principal's ex-post IC condition holds as an equality, the exercise rule maximizes the principal's payoff given that the agent's type lies in a given partition. In all other equilibria, there is both loss of information and delay in option exercise, which is detrimental for both the principal and the agent with a bias towards early exercise. Interestingly, delay in exercise in these equilibria occurs despite the fact that the agent is biased towards early exercise.

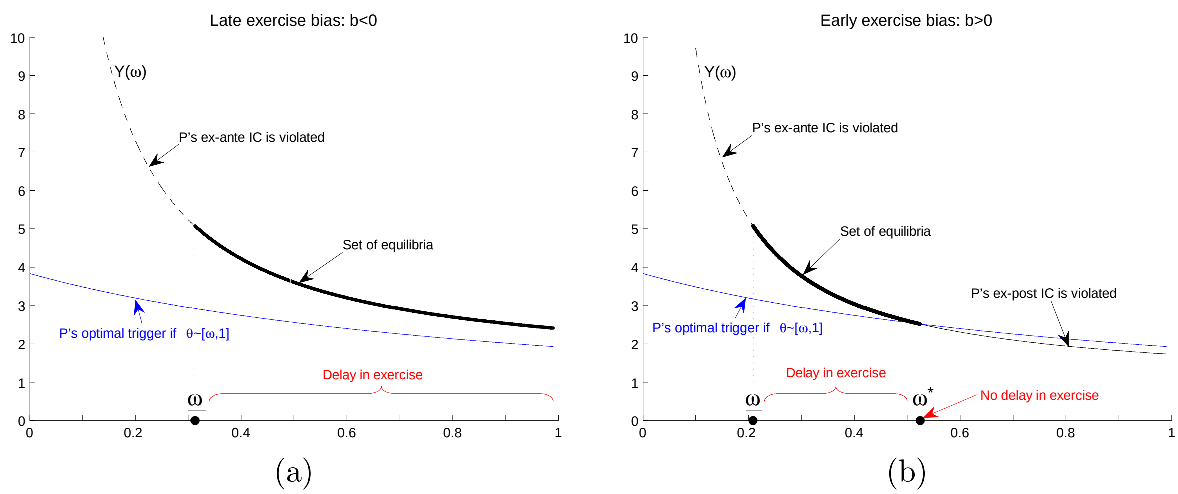

Figure B.1. Equilibria with partitioned exercise. The figures present the partition equilibria characterized in Proposition 2 for $\underline{\theta} = 0$, $r = 0.15$, $\alpha = 0.05$, $\sigma = 0.2$, and $I = 1$. The agent's bias is $b = -0.25$ in figure (a) and $b = 0.1$ in figure (b). In both figures, the black line represents the agent's IC condition, i.e., the function $Y(\omega)$, and the blue line represents the function $\frac{\beta}{\beta-1}\frac{2I}{\omega+1}$, i.e., the principal's optimal exercise trigger if θ is uniform on $[\omega, 1]$.

Proof of Proposition B.1. 1. Proof for $b < 0$. In the equilibrium with continuous exercise, exercise occurs at the unconstrained optimal time of any type θ of the agent. Therefore, the payoff of any type of the agent is higher in this equilibrium than in any other possible equilibrium. In addition, as Proposition 1 shows, the exercise times implied by the optimal mechanism if the principal could commit to any mechanism, coincide with the exercise times in the equilibrium with continuous exercise. Thus, the principal's expected payoff in this equilibrium exceeds her expected payoff under the exercise rule implied by any other equilibrium.

2. Proof for $b > 0$. The expected utility of the principal in the ω-equilibrium is $V_P(X, 1; \omega)$, given by (A8). As shown in Step 1 of the proof of Proposition 2, $V_P(X, 1; \omega)$ is strictly increasing in ω for $\omega \in (0, \omega^*)$. Hence, $V_P(X, 1; \omega^*) > V_P(X, 1; \omega)$ for any $\omega < \omega^*$. Denote the ex-ante expected utility of the agent (before the agent's type is realized) by $V_A(X, 1; \omega)$. Repeating the derivation of the principal's value function $V_P(X, 1; \omega)$ in the Online Appendix, it is easy to see that

$$V_A(X, 1; \omega) = \frac{1-\omega}{1-\omega^{\beta+1}} \left(\frac{X}{Y(\omega)}\right)^{\beta} \left(\frac{1}{2}(1+\omega) Y(\omega) - (I-b)\right).$$

The only difference of this expression from the expression for $V_P(X, 1; \omega)$ given by (A8) is that I in the second bracket of (A8) is replaced by $(I-b)$. To prove that $V_A(X, 1; \omega^*) > V_A(X, 1; \omega)$ for any $\omega < \omega^*$, we prove that $V_A(X, 1; \omega)$ is strictly increasing in ω for $\omega \in (0, \omega^*)$. The proof repeats the arguments behind Step 1 in the proof of Proposition 2. In particular, we can re-write $V_A(X, 1; \omega)$ as $2^{-\beta} X^{\beta} f_1(\omega) \tilde{f}_2(\omega)$, where

$$f_1(\omega) \equiv \frac{(1-\omega)(1+\omega)^{\beta}}{1-\omega^{\beta+1}} \text{ and } \tilde{f}_2(\omega) \equiv \frac{\frac{1}{2}(1+\omega) Y(\omega) - (I-b)}{\left(\frac{1}{2}(1+\omega) Y(\omega)\right)^{\beta}}.$$

As shown in Step 1 in the proof of Proposition 2, $f_1(\omega) > 0$ and $f_1'(\omega) > 0$. In addition, $\tilde{f}_2(\omega) > 0$ because $Y(\omega) > \frac{\beta}{\beta-1} \frac{2I}{1+\omega} > \frac{2(I-b)}{1+\omega}$ for any $\omega < \omega^*$, and $\tilde{f}_2'(\omega) > 0$ for the same reasons why $f_2'(\omega) > 0$ in Step 1 in the proof of Proposition 2. Hence, $V_A(X, 1; \omega)$ is increasing in $\omega \in (0, \omega^*)$. ■

Focusing on the most informative (ω^*) equilibrium in the early exercise bias case, the informativeness of communication exhibits interesting comparative statics.

Proposition B.2. *Consider the case of an agent biased towards early exercise, $b > 0$. Then, ω^* decreases in b and increases in β, and hence decreases in σ and α, and increases in r.*

The result that ω^* decreases in the agent's bias is in line with the result of Crawford and Sobel (1982) that less information is revealed if the misalignment of preferences is bigger. More interesting are the comparative statics results in σ, α, and r. The proposition shows that communication is less efficient when the option to wait is more valuable. For example, there is less information revelation (ω^* is lower) if the environment is more uncertain (σ is higher). Intuitively, higher uncertainty increases the value of the option to delay exercise and thus effectively increases the conflict of interest between the principal and the agent biased towards early exercise. Similarly, communication is less efficient in lower interest rate and higher growth rate environments.

Proof of Proposition B.2. The fact that ω^* decreases in b has been proved in the supplementary analysis for the proof of Proposition 2 in the Online Appendix. We next show that ω^* increases in β. From $Y(\omega) = \frac{\beta}{\beta-1}\frac{2I}{\omega+1}$, ω^* solves $F(\omega,\beta) = 0$, where $F(\omega,\beta) = \frac{\beta}{\beta-1}\frac{1-\omega^{\beta-1}}{1-\omega^{\beta}}\frac{2I}{I-b} - 1 - \frac{1}{\omega}$. Denote the unique solution by $\omega^*(\beta)$. Function $F(\omega,\beta)$ is continuously differentiable in both arguments on $\omega \in (0,1)$, $\beta > 1$. Differentiating $F(\omega^*(\beta),\beta)$ in β, $\frac{\partial\omega^*}{\partial\beta} = -\frac{F_\beta(\omega^*(\beta),\beta)}{F_\omega(\omega^*(\beta),\beta)}$. Since $F(0,\beta) < 0$, $F(1,\beta) = \frac{2b}{I-b} > 0$, and ω^* is the unique solution of $F(\omega,\beta) = 0$ in $(0,1)$, we know that $F_\omega(\omega^*(\beta),\beta) > 0$. As we prove below, $F_\beta(\omega,\beta) < 0$. Hence, ω^* is strictly increasing in $\beta > 1$. Finally, a standard calculation shows that $\frac{\partial\beta}{\partial\sigma} < 0$, $\frac{\partial\beta}{\partial\alpha} < 0$, and $\frac{\partial\beta}{\partial r} > 0$. Thus, ω^* decreases in β and α and increases in r.

Proof that $F_\beta(\omega,\beta) < 0$. Differentiating $F(\omega,\beta)$ with respect to β and reorganizing the terms, we obtain that $F_\beta(\omega,\beta) < 0$ is equivalent to

$$\frac{\left(1-\omega^{\beta-1}\right)\left(1-\omega^{\beta}\right)}{\omega^{\beta-1}(1-\omega)} + \beta(\beta-1)\ln\omega > 0.$$

Denote the left-hand side as a function of β by $N(\beta)$. Because $N(1) = 0$, a sufficient condition for $N(\beta) > 0$ for any $\beta > 1$ is that $N'(\beta) > 0$ for $\beta > 1$. Differentiating $N(\beta)$:

$$N'(\beta) = \ln\omega\left[-\frac{\omega^{1-\beta}-\omega^{\beta}}{1-\omega} + 2\beta - 1\right].$$

Because $\ln\omega < 0$ for any $\omega \in (0,1)$, condition $N'(\beta) > 0$ is equivalent to $n(\beta) \equiv \frac{\omega^{1-\beta}-\omega^{\beta}}{1-\omega} - 2\beta + 1 > 0$. Note that $\lim_{\beta\to1} n(\beta) = 0$ and $n'(\beta) = -\left(\omega^{1-\beta}+\omega^{\beta}\right)\frac{\ln\omega}{1-\omega} - 2 \equiv \eta(\beta)$. Note that

$$\eta(\beta) = \eta(1) + \int_1^\beta \eta'(x)\,dx = -\frac{(1+\omega)\ln\omega}{1-\omega} - 2 + \frac{(\ln\omega)^2}{1-\omega}\int_1^\beta\left(\left(\frac{1}{\omega}\right)^{2x-1} - 1\right)\omega^x dx. \tag{55}$$

The second term of (55) is positive, because $\left(\frac{1}{\omega}\right)^{2x-1} - 1 > 0$, since $\frac{1}{\omega} > 1$ and $2x-1 > 1$ for any $x > 1$. The first term of (55) is positive, because

$$\begin{aligned}&\lim_{\omega\to1}\left(-\frac{(1+\omega)\ln\omega}{1-\omega} - 2\right) = \lim_{\omega\to1}\left(\ln\omega + \frac{1+\omega}{\omega}\right) - 2 = 0\\ &\text{and } \frac{\partial\left(-\frac{(1+\omega)\ln\omega}{1-\omega}-2\right)}{\partial\omega} = \frac{-2\ln\omega - \frac{1}{\omega} + \omega}{(1-\omega)^2} < 0,\end{aligned}$$

where the first row is by l'Hopital's rule, and the second row is because $\left(-2\ln\omega - \frac{1}{\omega} + \omega\right)' = \frac{(1-\omega)^2}{\omega^2} > 0$ and $-2\ln\omega - \frac{1}{\omega} + \omega$ equals zero at $\omega = 1$. Thus, $\eta(\beta) > 0$ and hence $n'(\beta) > 0$ for any $\beta > 1$, which together with $n(1) = 0$ implies $n(\beta) > 0$, which in turn implies that $N(\beta) > 0$ for any $\beta > 1$. Hence, $F_\beta(\omega,\beta) < 0$. ■

Proof of Proposition 4. We start by showing that under the specified restrictions on b, the solutions θ_L and θ_H indeed exist. First, if $b \in (0, \frac{\mathbb{E}[\theta]-\underline{\theta}}{\mathbb{E}[\theta]}I)$, then $\underline{\theta}\frac{I}{I-b} < \mathbb{E}[\theta]$ and $\bar{\theta}\frac{I}{I-b} > \bar{\theta}$. Since $\mathbb{E}\left[\tilde{\theta}|\tilde{\theta} \geq \theta\right]$ is continuous in θ, the equation $\theta\frac{I}{I-b} = \mathbb{E}\left[\tilde{\theta}|\tilde{\theta} \geq \theta\right]$ has at least one solution, and all solutions are strictly below $\bar{\theta}$. Hence, $\theta_H < \bar{\theta}$. Second, if $b \in (-\frac{\bar{\theta}-\mathbb{E}[\theta]}{\mathbb{E}[\theta]}I, 0)$, then $\bar{\theta}\frac{I}{I-b} > \mathbb{E}[\theta]$ and $\underline{\theta}\frac{I}{I-b} < \underline{\theta}$. Since $\mathbb{E}\left[\tilde{\theta}|\tilde{\theta} \leq \theta\right]$ is continuous in θ, the equation $\theta\frac{I}{I-b} = \mathbb{E}\left[\tilde{\theta}|\tilde{\theta} \leq \theta\right]$ has at least one solution. By part (ii) of Assumption 2, the solution is unique and $\theta_L > \underline{\theta}$. In fact, as the proof below shows, part 1 of Proposition 4 is satisfied even if part (ii) of Assumption 2 is not satisfied and we define θ_L as the lowest solution to $\theta\frac{I}{I-b} = E\left[\tilde{\theta}|\tilde{\theta} \leq \theta\right]$.

Proof of Part 1. We show that if part (i) of Assumption 2 is satisfied and θ_L is defined as the lowest solution to $\theta\frac{I}{I-b} = E\left[\tilde{\theta}|\tilde{\theta} \leq \theta\right]$, then our problem satisfies the conditions of Proposition 1 in Amador and Bagwell (2013), once we introduce a change in process from $X(t)$ to $P(X(t)) \equiv X(t)^{-\beta+1}$. Since $\beta > 1$, $P(X)$ is strictly decreasing in X with $\lim_{X\to 0} P(X) = \infty$ and $\lim_{X\to\infty} P(X) = 0$. Thus, there is a one-to-one correspondence between X and P, so any threshold-exercise direct mechanism $\left\{\hat{X}(\theta), \theta \in \Theta\right\}$ with upper thresholds on process $X(t)$ can be equivalently written as a threshold-exercise direct mechanism $\left\{\hat{P}(\theta), \theta \in \Theta\right\}$ with lower thresholds on process $P(X(t))$, with $\hat{P}(\theta) = \hat{X}(\theta)^{-\beta+1}$. The payoffs of the principal and the agent, $\tilde{U}_P\left(\hat{P}, \theta\right)$ and $\tilde{U}_A\left(\hat{P}, \theta\right)$, divided by $X(0)^\beta$, can be written as:

$$\tilde{U}_P\left(\hat{P}, \theta\right) = \frac{\theta \hat{P}^{\frac{1}{1-\beta}} - I}{\hat{P}^{\frac{\beta}{1-\beta}}} = \theta\hat{P} - I\hat{P}^{\frac{\beta}{\beta-1}}$$

$$\tilde{U}_A\left(\hat{P}, \theta\right) = \frac{\theta \hat{P}^{\frac{1}{1-\beta}} - (I-b)}{\hat{P}^{\frac{\beta}{1-\beta}}} = \theta\hat{P} - (I-b)\hat{P}^{\frac{\beta}{\beta-1}}$$

Thus, the optimal mechanism problem (similar to (4)–(5), but for a general distribution) is a special case of the problem in Amador and Bagwell (2013) without money burning, where $\gamma = \theta$, $\pi = \hat{P}$, $\omega\left(\theta, \hat{P}\right) = \tilde{U}_P\left(\hat{P}, \theta\right)$, and $b\left(\hat{P}\right) = -(I-b)\hat{P}^{\frac{\beta}{\beta-1}}$. It is easy to check that the conditions of Assumption 1 in their paper hold for any $b < I$. By analogy with Amador and Bagwell (2013), define κ as

$$\kappa \equiv \frac{\frac{\partial^2 \tilde{U}_P\left(\hat{P}, \theta\right)}{\partial \hat{P}^2}}{b''\left(\hat{P}\right)} = \frac{I}{I-b}.$$

We next verify that the conditions of Proposition 1 in Amador and Bagwell (2013) hold for these functions. Since $X_A^*(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$, the optimal agent's threshold in terms of $P(t)$ is $P_A^*(\theta) = \left(\frac{\beta}{\beta-1}\frac{I-b}{\theta}\right)^{1-\beta}$. Then, $\frac{\partial \tilde{U}_P\left(P_A^*(\theta), \tilde{\theta}\right)}{\partial \hat{P}} = \tilde{\theta} - \frac{I}{I-b}\theta$.

Condition (c1). Since $\frac{\partial \tilde{U}_P\left(P_A^*(\theta), \theta\right)}{\partial \hat{P}} = -\frac{b}{I-b}\theta$, (c1) is satisfied if and only if $\Phi(\theta) + \frac{b}{I}\theta\phi(\theta)$ is non-decreasing for all $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$.

Condition (c2). This condition is relevant in the case $b \in (0, \frac{\mathbb{E}[\theta]-\underline{\theta}}{\mathbb{E}[\theta]}I)$, in which case $\theta_H < \bar{\theta}$. In this case, we need to verify

$$(\theta - \theta_H) \frac{I}{I-b} \geq \int_{\theta}^{\bar{\theta}} \left(\tilde{\theta} - \frac{I}{I-b}\theta_H\right) \frac{\phi\left(\tilde{\theta}\right)}{1-\Phi(\theta)} d\tilde{\theta} \;\forall\theta \in \left[\theta_H, \bar{\theta}\right]$$

with equality at θ_H. Rearranging the terms, this inequality is equivalent to $\theta\frac{I}{I-b} \geq \mathbb{E}\left[\tilde{\theta}|\tilde{\theta} \geq \theta\right]$. By definition of θ_H, this condition indeed holds as equality at θ_H. Furthermore, since θ_H is the highest $\theta \in \Theta$ at which $\theta\frac{I}{I-b} = \mathbb{E}\left[\tilde{\theta}|\tilde{\theta} \geq \theta\right]$ and because $\bar{\theta}\frac{I}{I-b} > \bar{\theta}$, this inequality holds strictly at any $\theta \in \left(\theta_H, \bar{\theta}\right]$.

Condition (c2'). This condition is relevant in the case $b < 0$. In this case, we need to verify $\frac{\partial \tilde{U}_P}{\partial \hat{P}}\left(P_A^*\left(\bar{\theta}\right), \bar{\theta}\right) \geq 0$, which is equivalent to $\bar{\theta} - \frac{I}{I-b}\bar{\theta} \geq 0$. It is satisfied if and only if $b < 0$.

Condition (c3). This condition is relevant in the case $b \in (-\frac{\bar{\theta}-\mathbb{E}[\theta]}{\mathbb{E}[\theta]}I, 0)$, in which case $\theta_L > \underline{\theta}$. In this case, we need to verify

$$(\theta - \theta_L) \frac{I}{I-b} \leq \int_{\underline{\theta}}^{\theta} \left(\tilde{\theta} - \frac{I}{I-b}\theta_L\right) \frac{\phi\left(\tilde{\theta}\right)}{\Phi(\theta)} d\tilde{\theta} \;\forall\theta \in \left[\underline{\theta}, \theta_L\right]$$

with equality at θ_L. Rearranging the terms, this inequality is equivalent to $\theta\frac{I}{I-b} \leq \mathbb{E}\left[\tilde{\theta}|\tilde{\theta} \leq \theta\right]$. By definition of θ_L, it holds as equality at θ_L. Furthermore, since θ_L is the lowest $\theta \in \Theta$ at which $\theta\frac{I}{I-b} = \mathbb{E}\left[\tilde{\theta}|\tilde{\theta} \leq \theta\right]$ and because $\underline{\theta}\frac{I}{I-b} < \underline{\theta}$, this inequality holds strictly at any $\theta \in (\underline{\theta}, \theta_L]$.

Condition (c3'). This condition is relevant in the case $b > 0$. We need to verify $\frac{\partial \tilde{U}_P}{\partial \hat{P}}\left(P_A^*\left(\underline{\theta}\right), \underline{\theta}\right) \leq 0$, which is equivalent to $\underline{\theta} - \frac{I}{I-b}\underline{\theta} \leq 0$. It is satisfied if $b \in (0, I)$ and hence is satisfied if $b \in (0, \frac{\mathbb{E}[\theta]-\underline{\theta}}{\mathbb{E}[\theta]}I)$.

Applying Proposition 1 in Amador and Bagwell (2013), we conclude that the optimal threshold-exercise decision rule is $\hat{X}(\theta) = X_A^*(\min\{\theta, \theta_H\})$ if $b \in (0, \frac{\mathbb{E}[\theta]-\underline{\theta}}{\mathbb{E}[\theta]}I)$, and $\hat{X}(\theta) = X_A^*(\max\{\theta, \theta_L\})$, if $b \in (-\frac{\bar{\theta}-\mathbb{E}[\theta]}{\mathbb{E}[\theta]}I, 0)$.

Proof of Part 2. To prove this part, we impose part (ii) of Assumption 2. Given that the principal plays the strategy stated in the proposition, the strategy of any type θ of the agent is incentive-compatible. Indeed, for any type $\theta \geq \theta_L$, exercise occurs at his most preferred time, so no type $\theta \geq \theta_L$ benefits from a deviation. Any type $\theta < \theta_L$ does not benefit from a deviation either because the agent would lose from inducing the principal to exercise earlier, and inducing exercise later than threshold $X_A^*(\theta_L)$ is not feasible given that the principal never exercises later than $X_A^*(\theta_L)$ under her strategy. Next, let us verify the optimality of the principal's strategy. We need to check that the principal has incentives to exercise the option immediately when the agent sends message $m = 1$ (the ex-post IC constraint), and not to exercise the option before getting message $m = 1$ (the ex-ante IC constraint). The ex-post IC constraint follows from the fact that the principal learns the agent's type θ if the agent sends message $m = 1$ at first-passage time of any threshold between $X_A^*\left(\bar{\theta}\right)$ and $X_A^*(\theta_L)$, and realizes that it is already too late to exercise

$(X_P^*(\theta) < X_A^*(\theta))$, and thus does not benefit from delaying exercise even further. If the agent sends a message to exercise when $X(t)$ hits $X_A^*(\theta_L) = \frac{\beta}{\beta-1}\frac{I-b}{\theta_L}$, the principal infers that $\theta \leq \theta_L$ and that she will not learn any additional information by waiting more. Given the belief that $\theta \in [\underline{\theta}, \theta_L]$, the optimal exercise threshold for the principal is given by $\frac{\beta}{\beta-1}\frac{I}{\mathbb{E}[\theta|\theta\leq\theta_L]}$, which equals $\frac{\beta}{\beta-1}\frac{I-b}{\theta_L}$ by the definition of θ_L. Hence, the ex-post IC constraint is satisfied. Finally, consider the ex-ante IC constraint.

Proof that the principal's ex-ante IC constraint is satisfied. Let $\tilde{V}_p\left(X, \hat{\theta}, \theta_L\right)$ be the expected value to the principal from following the strategy of waiting for the agent's recommendation $m = 1$ until $X(t)$ hits $X_A^*(\theta_L)$ for the first time and exercising at threshold $X_A^*(\theta_L)$ regardless of the agent's recommendation, where X is the current value of $X(t)$ and the principal believes that θ is distributed over $\left[\underline{\theta}, \hat{\theta}\right]$, $\hat{\theta} \geq \theta_L$ with p.d.f. $\phi(\theta)/\Phi(\hat{\theta})$:

$$\tilde{V}_p\left(X, \hat{\theta}, \theta_L\right) = X^\beta \left(\int_{\theta_L}^{\hat{\theta}} \frac{\theta X_A^*(\theta) - I}{X_A^*(\theta)^\beta}\frac{\phi(\theta)}{\Phi(\hat{\theta})}d\theta + \int_{\underline{\theta}}^{\theta_L} \frac{\theta X_A^*(\theta_L) - I}{X_A^*(\theta_L)^\beta}\frac{\phi(\theta)}{\Phi(\hat{\theta})}d\theta\right).$$

Because the principal's belief is that $\theta \in \left[\underline{\theta}, \hat{\theta}\right]$, the current value of $X(t)$ satisfies $X(t) \leq X_A^*(\hat{\theta})$. Hence, the ex-ante IC constraint requires $\tilde{V}_p\left(X, \hat{\theta}, \theta_L\right) \geq X\mathbb{E}\left[\theta|\theta \leq \hat{\theta}\right] - I$ for any $\hat{\theta} > \theta_L$ and $X \leq X_A^*\left(\hat{\theta}\right)$, or, equivalently,

$$\int_{\theta_L}^{\hat{\theta}} \frac{\theta X_A^*(\theta) - I}{X_A^*(\theta)^\beta}\frac{\phi(\theta)}{\Phi(\hat{\theta})}d\theta + \int_{\underline{\theta}}^{\theta_L} \frac{\theta X_A^*(\theta_L) - I}{X_A^*(\theta_L)^\beta}\frac{\phi(\theta)}{\Phi(\hat{\theta})}d\theta \geq \frac{X\mathbb{E}\left[\theta|\theta \leq \hat{\theta}\right] - I}{X^\beta}. \tag{56}$$

The right-hand side of (56) is an inverted U-shaped function that reaches its maximum at $X_{\max} \equiv \frac{\beta}{\beta-1}\frac{I}{\mathbb{E}[\theta|\theta\leq\hat{\theta}]}$. Since equation $\mathbb{E}\left[\tilde{\theta}|\tilde{\theta} \leq \theta\right] = \frac{I}{I-b}\theta$ has a unique solution $\theta_L \in \Theta$ and since $\mathbb{E}[\theta] < \frac{I}{I-b}\bar{\theta}$ for $b \in (-\frac{\bar{\theta}-\mathbb{E}[\theta]}{\mathbb{E}[\theta]}I, 0)$, then $\mathbb{E}\left[\theta|\theta \leq \hat{\theta}\right] \leq \frac{I}{I-b}\hat{\theta}$ for any $\hat{\theta} \geq \theta_L$. Therefore, $X_{\max} = \frac{\beta}{\beta-1}\frac{I}{\mathbb{E}[\theta|\theta\leq\hat{\theta}]} \geq \frac{\beta}{\beta-1}\frac{I-b}{\hat{\theta}} = X_A^*\left(\hat{\theta}\right)$, and hence the right-hand side of (56) is strictly increasing in X over $X \leq X_A^*\left(\hat{\theta}\right)$. Hence, the ex-ante IC constraint (56) is satisfied for any $X \leq X_A^*\left(\hat{\theta}\right)$ if and only if it is satisfied at $X = X_A^*\left(\hat{\theta}\right)$. Finally, suppose that (56) is violated at $X = X_A^*\left(\hat{\theta}\right)$ for some $\hat{\theta} > \theta_L$. However, this implies that threshold schedule $\hat{X}(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\max\{\theta,\theta_L\}}$ is dominated by $\hat{X}(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\max\{\theta,\hat{\theta}\}}$, which contradicts part 1 of the proposition. Indeed, violation of (56) means

$$\int_{\theta_L}^{\hat{\theta}} \frac{\theta X_A^*(\theta) - I}{X_A^*(\theta)^\beta}\frac{\phi(\theta)}{\Phi(\hat{\theta})}d\theta + \int_{\underline{\theta}}^{\theta_L} \frac{\theta X_A^*(\theta_L) - I}{X_A^*(\theta_L)^\beta}\frac{\phi(\theta)}{\Phi(\hat{\theta})}d\theta < \frac{X_A^*\left(\hat{\theta}\right)\mathbb{E}\left[\theta|\theta \leq \hat{\theta}\right] - I}{X_A^*\left(\hat{\theta}\right)^\beta}. \tag{57}$$

The principal's expected utility under the optimal contract in Part 1 of the proposition, divided by $X(0)^{\beta}$, is

$$\int_{\underline{\theta}}^{\theta_L} \frac{\theta X_A^*(\theta_L) - I}{X_A^*(\theta_L)^{\beta}} \phi(\theta)\, d\theta + \int_{\theta_L}^{\bar{\theta}} \frac{\theta X_A^*(\theta) - I}{X_A^*(\theta)^{\beta}} \phi(\theta)\, d\theta. \tag{58}$$

Consider a modified contract with θ_L replaced by $\hat{\theta}$. The principal's expected utility under this modified contract, divided by $X(0)^{\beta}$, is

$$\int_{\underline{\theta}}^{\hat{\theta}} \frac{\theta X_A^*\left(\hat{\theta}\right) - I}{X_A^*\left(\hat{\theta}\right)^{\beta}} \phi(\theta)\, d\theta + \int_{\hat{\theta}}^{\bar{\theta}} \frac{\theta X_A^*(\theta) - I}{X_A^*(\theta)^{\beta}} \phi(\theta)\, d\theta. \tag{59}$$

Rearranging the terms, it is straightforward to see that (58)<(59) is equivalent to (57). Hence, (57) implies that the contract in Part 1 of the proposition is dominated by another interval delegation contract, which is a contradiction. Thus, the ex-ante IC constraint is indeed satisfied, which completes the proof of Part 2.

Proof of Part 3. Similarly to the argument in the proof of Proposition 1, if $b > 0$, there is no equilibrium that features separation of types over some interval. Hence, the mechanism from part 1, which features separation, cannot be implemented in any equilibrium. ■

The next lemma describes the distributions that satisfy and do not satisfy Assumption 2.

Lemma IA.4. *Suppose that $-I < b < 0$ and $0 < \underline{\theta} < \bar{\theta}$.*

1. *If θ is uniformly distributed on $\left[\underline{\theta}, \bar{\theta}\right]$, Assumption 2 is satisfied.*
2. *If θ is distributed according to a truncated standard normal distribution on $\left[\underline{\theta}, \bar{\theta}\right]$, Assumption 2 is satisfied.*
3. *Suppose θ is distributed according to a power distribution with parameter α on $\left[\underline{\theta}, \bar{\theta}\right]$, i.e., $\Phi(\theta) = (\frac{\theta - \underline{\theta}}{\bar{\theta} - \underline{\theta}})^{\alpha}$. Then, Assumption 2 is satisfied if and only if $\alpha \leq 1$.*

Proof of Lemma IA.4. To prove that $\Phi(\theta) + \frac{b}{I}\theta\phi(\theta)$ is non-decreasing in θ, it is necessary and sufficient to check that

$$\left[\Phi(\theta) + \frac{b}{I}\theta\phi(\theta)\right]' = \phi(\theta) + \frac{b}{I}\phi(\theta) + \frac{b}{I}\theta\phi'(\theta) = \phi(\theta)\left(1 + \frac{b}{I}\right) + \frac{b}{I}\theta\phi'(\theta) \geq 0 \tag{60}$$

for all $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$.

Proof of Part 1. For a uniform distribution, $\phi(\theta) = \frac{1}{\bar{\theta} - \underline{\theta}}$, $\Phi(\theta) = \frac{\theta - \underline{\theta}}{\bar{\theta} - \underline{\theta}}$. First,

$$\Phi(\theta) + \frac{b}{I}\theta\phi(\theta) = \frac{\theta - \underline{\theta} + \frac{b}{I}\theta}{\bar{\theta} - \underline{\theta}},$$

which increases in θ since $b > -I$. Second,

$$\mathbb{E}[\tilde{\theta}|\tilde{\theta} \leq \theta] = \frac{I}{I-b}\theta \Leftrightarrow \frac{\theta + \underline{\theta}}{2} = \frac{I}{I-b}\theta \Leftrightarrow \theta\left(\frac{1}{2} - \frac{I}{I-b}\theta\right) + \frac{\underline{\theta}}{2} = 0,$$

which has at most one solution because $\underline{\theta} > 0$.

Proof of Part 2. Denote f and F the probability density function (pdf) and the cumulative distribution function (cdf) of the standard normal distribution $N(0,1)$. Denote the cdf of the standard normal distribution truncated on $[a,b]$ by $F_{a,b}$ and the pdf by $f_{a,b}$. Then, according to the properties of normal distribution,

$$f_{a,b}(\theta) = \frac{f(\theta)}{F(b) - F(a)}$$

and the conditional expectation is

$$\mathbb{E}[\theta|a < \theta < b] = \frac{f(b) - f(a)}{F(b) - F(a)}. \tag{61}$$

Hence, (60) is satisfied because $\phi(\theta) = \frac{f(\theta)}{F(\bar{\theta}) - F(\underline{\theta})} > 0$, $1 + \frac{b}{I} > 0$, $b < 0$, $\theta > 0$, and $\phi'(\theta) = \frac{f'(\theta)}{F(\bar{\theta}) - F(\underline{\theta})} < 0$ for $\theta > 0$. Thus, part (i) of Assumption 2 is satisfied.

Next, using (61),

$$\mathbb{E}[\tilde{\theta}|\tilde{\theta} \leq \theta] = \frac{I}{I-b}\theta \Leftrightarrow \frac{f(\theta) - f(\underline{\theta})}{F(\theta) - F(\underline{\theta})} = \frac{I}{I-b}\theta.$$

The function on the right-hand side is increasing and the function on the left-hand side is decreasing because $f(\theta)$ decreases and $F(\theta)$ increases on $\theta > 0$. Hence, this equation can have at most one solution,which proves part (ii) of Assumption 2.

Proof of Part 3. We first show that part (ii) of Assumption 2 is satisfied for any a. Since $\phi(\theta) = \frac{\alpha}{\bar{\theta} - \underline{\theta}}(\frac{\theta - \underline{\theta}}{\bar{\theta} - \underline{\theta}})^{\alpha - 1}$, we have

$$\begin{aligned} \mathbb{E}[\tilde{\theta}|\tilde{\theta} \leq \theta] &= \int_{\underline{\theta}}^{\theta} x\frac{\phi(x)}{\Phi(\theta)}dx = \frac{1}{\Phi(\theta)}\int_{\underline{\theta}}^{\theta} x\Phi'(x)\,dx = \frac{[x\Phi(x)]_{\underline{\theta}}^{\theta} - \int_{\underline{\theta}}^{\theta}\Phi(x)\,dx}{\Phi(\theta)} \\ &= \theta - \frac{\int_{\underline{\theta}}^{\theta}(\frac{x-\underline{\theta}}{\bar{\theta}-\underline{\theta}})^{\alpha}dx}{\Phi(\theta)} = \theta - \frac{\int_{\underline{\theta}}^{\theta}(x-\underline{\theta})^{\alpha}dx}{(\theta-\underline{\theta})^{\alpha}} = \theta - \frac{\frac{1}{\alpha+1}\left[(x-\underline{\theta})^{\alpha+1}\right]_{\underline{\theta}}^{\theta}}{(\theta-\underline{\theta})^{\alpha}} = \theta - \frac{\theta - \underline{\theta}}{\alpha+1} \end{aligned}$$

Hence,

$$\mathbb{E}[\tilde{\theta}|\tilde{\theta} \leq \theta] = \frac{I}{I-b}\theta \Leftrightarrow \theta - \frac{\theta - \underline{\theta}}{\alpha+1} = \frac{I}{I-b}\theta \Leftrightarrow \theta\left(\frac{\alpha}{\alpha+1} - \frac{I}{I-b}\right) + \frac{\underline{\theta}}{\alpha+1} = 0,$$

which has at most one solution because $\underline{\theta} > 0$.

Next, consider part (i) of Assumption 2:

$$\Phi(\theta) + \frac{b}{I}\theta\phi(\theta) = (\frac{\theta - \underline{\theta}}{\bar{\theta} - \underline{\theta}})^{\alpha} + \frac{b}{I}\theta\frac{\alpha}{\bar{\theta} - \underline{\theta}}(\frac{\theta - \underline{\theta}}{\bar{\theta} - \underline{\theta}})^{\alpha - 1} = \frac{(\theta - \underline{\theta})^{\alpha - 1}}{(\bar{\theta} - \underline{\theta})^{\alpha}}\left[\theta\left(1 + \frac{b}{I}\alpha\right) - \underline{\theta}\right],$$

which increases in θ if and only if

$$(\theta - \underline{\theta})^{\alpha - 1}\left(1 + \tfrac{b}{I}\alpha\right) + (\alpha - 1)(\theta - \underline{\theta})^{\alpha - 2}\left[\theta\left(1 + \tfrac{b}{I}\alpha\right) - \underline{\theta}\right] \geq 0 \Leftrightarrow$$
$$\theta\left[1 + \tfrac{b}{I}\alpha + (\alpha - 1)\left(1 + \tfrac{b}{I}\alpha\right)\right] \geq \underline{\theta}\left[1 + \tfrac{b}{I}\alpha + \alpha - 1\right] \Leftrightarrow \theta\left(1 + \tfrac{b}{I}\alpha\right) \geq \underline{\theta}\left[\tfrac{b}{I} + 1\right].$$

First, if $1 + \frac{b}{I}\alpha \leq 0 \Leftrightarrow \alpha \geq \frac{I}{-b}$, then the inequality is violated because the right-hand side is positive. Suppose $1 + \frac{b}{I}\alpha > 0 \Leftrightarrow \alpha < \frac{I}{-b}$. Then the above inequality is satisfied for any $\theta \geq \underline{\theta}$ if and only if it is satisfied for $\theta = \underline{\theta}$, or equivalently, $1 + \frac{b}{I}\alpha \geq \frac{b}{I} + 1 \Leftrightarrow \alpha \leq 1$. Note that $\frac{I}{-b} > 1$ and hence, if $\alpha \leq 1$, then $\alpha \leq 1 < \frac{I}{-b}$, and so the inequality is satisfied for all $\theta \geq \underline{\theta}$. On the other hand, if $\alpha > 1$, then either $\alpha \geq \frac{I}{-b} > 1$ and then the inequality is violated for any θ, or $\frac{I}{-b} > \alpha > 1$, and then the inequality is violated for θ close to $\underline{\theta}$. Thus, the inequality is satisfied for all $\theta \geq \underline{\theta}$ if and only if $\alpha \leq 1$. ■

Proof of Proposition 5. Consider a problem in which the agent is free to choose whether to exercise the option prior to the arrival of the news, while the principal makes the exercise decision after the arrival of the news. Let $V_A^a(X, \theta)$ be the value of the option to the agent of type θ after the arrival of the news:

$$V_A^a(X,\theta) = \begin{cases} \left(\frac{X}{X_P^*(\theta)}\right)^{\beta}(\theta X_P^*(\theta) - I + b), & \text{if } X \leq X_P^*(\theta), \\ \theta X - I + b, & \text{if } X \geq X_P^*(\theta). \end{cases}$$

Let $V_A^b(X, \theta)$ be the value of the option to the agent before the arrival of the news. Since the expected return from holding an option over a small interval $[t, t + dt]$ must be rdt, $V_A^b(X, \theta)$ satisfies

$$(r + \lambda)V_A^b(X,\theta) = \alpha X\frac{\partial V_A^b(X,\theta)}{\partial X} + \frac{1}{2}\sigma^2 X^2\frac{\partial^2 V_A^b(X,\theta)}{\partial X^2} + \lambda V_A^a(X,\theta). \tag{62}$$

First, we show that waiting is optimal in the range $X \leq X_P^*(\theta)$. Since the agent can follow the strategy of exercising the option at threshold $X_P^*(\theta)$, it must be that $V_A^b(X,\theta) \geq \left(\frac{X}{X_P^*(\theta)}\right)^{\beta}(\theta X_P^*(\theta) - I + b)$. By contradiction, suppose that there exists point $\hat{X} < X_P^*(\theta)$ at which it is optimal for the agent to exercise the option. It follows that because $V_A^b(X,\theta) \geq$

$\left(\frac{X}{X_P^*(\theta)}\right)^\beta (\theta X_P^*(\theta) - I + b)$, it must be that

$$V_A^b(X,\theta) \geq \left(\frac{X}{X_P^*(\theta)}\right)^\beta (\theta X_P^*(\theta) - I + b) \leq \left(\frac{X}{\hat{X}}\right)^\beta \left(\theta\hat{X} - I + b\right),$$

which is a contradiction since the right-hand side is strictly increasing in $\hat{X} \in (0, X_P^*(\theta))$. Thus, the exercise threshold is in the range $X \geq X_P^*(\theta)$. Eq. (62) must be solved subject to the value-matching and smooth-pasting conditions, we obtain

$$V_A^b(X,\theta) = \begin{cases} BX^{\gamma_+} + \left(\frac{X}{X_P^*(\theta)}\right)^\beta (\theta X_P^*(\theta) - I + b), & \text{if } X \leq X_P^*(\theta), \\ A_1 X^{\gamma_-} + A_2 X^{\gamma_+} + \frac{\lambda\theta}{r+\lambda-\alpha}X - \frac{\lambda(I-b)}{r+\lambda}, & \text{if } X \in \left[X_P^*(\theta), \tilde{X}_A(\theta)\right], \\ \theta X - I + b, & \text{if } X \geq \tilde{X}_A(\theta), \end{cases}$$

where $\gamma^+ > 1$ and $\gamma^- < 0$ are the roots of $\frac{1}{2}\sigma^2\gamma(\gamma-1) + \alpha\gamma - r - \lambda = 0$, constants A_1, A_2, and B are given by

$$\begin{aligned} A_1 &= \frac{X_P^*(\theta)^{-\gamma_-}}{\gamma_+ - \gamma_-}\left(\theta X_P^*(\theta)\left(\frac{(r-\alpha)(\gamma_+ - 1)}{r+\lambda-\alpha} - \beta + 1\right) - (I-b)\left(\frac{r\gamma_+}{r+\lambda} - \beta\right)\right), \\ A_2 &= \frac{\tilde{X}_A(\theta)^{-\gamma_+}}{\gamma_+ - \gamma_-}\left(\theta\tilde{X}_A(\theta)\frac{(r-\alpha)(1-\gamma_-)}{r+\lambda-\alpha} + (I-b)\frac{r\gamma_-}{r+\lambda}\right), \\ B &= A_2 - \frac{X_P^*(\theta)^{-\gamma_+}}{\gamma_+ - \gamma_-}\left(\theta X_P^*(\theta)\left(\beta - \gamma_- - \frac{\lambda(1-\gamma_-)}{r+\lambda-\alpha}\right) - (I-b)\left(\beta - \frac{r\gamma_-}{r+\lambda}\right)\right), \end{aligned}$$

and the optimal exercise threshold $\tilde{X}_A(\theta)$ satisfies

$$(63)\quad \theta\tilde{X}_A(\theta)\frac{(r-\alpha)(\gamma_+ - 1)}{r+\lambda-\alpha} - (I-b)\frac{r\gamma_+}{r+\lambda} = \left(\frac{\tilde{X}_A(\theta)}{X_P^*(\theta)}\right)^{\gamma_-}\left(\begin{array}{c} I\left(\frac{\beta}{\beta-1}\frac{(r-\alpha)(\gamma_+-1)}{r+\lambda-\alpha} - \frac{r\gamma_+}{r+\lambda}\right) \\ -b\left(\beta - \frac{r\gamma_+}{r+\lambda}\right) \end{array}\right).$$

The left-hand side is strictly increasing in $\tilde{X}_A(\theta)$. Let us see that the right-hand side is strictly decreasing in $\tilde{X}_A(\theta)$. Since $\gamma_- < 0$ and $b < 0$, it is sufficient to show that $\frac{\beta}{\beta-1}\frac{(r-\alpha)(\gamma_+-1)}{r+\lambda-\alpha} > \frac{r\gamma_+}{r+\lambda}$ and $\beta > \frac{r\gamma_+}{r+\lambda}$. Using the definition of γ_+, $\frac{\gamma_+}{\gamma_+-1}\frac{r+\lambda-\alpha}{r+\lambda} = 1 + \frac{\sigma^2\gamma_+}{2(r+\lambda)}$. Let us show that $\frac{\gamma_+}{r+\lambda}$ is strictly decreasing in λ:

$$\frac{\sigma^2\gamma_+}{r+\lambda} = \sqrt{\left(\rho\left(\alpha - \frac{\sigma^2}{2}\right)\right)^2 + 2\rho\sigma^2} - \rho\left(\alpha - \frac{\sigma^2}{2}\right),$$

where $\rho \equiv 1/(r+\lambda)$. Differentiating with respect to ρ and using $\sqrt{1+x} < 1 + \frac{x}{2}$ for $x > 0$, we obtain

$$\frac{-\rho\left(\alpha-\frac{\sigma^2}{2}\right)^2\left(\sqrt{1+\frac{2\sigma^2}{\rho\left(\alpha-\frac{\sigma^2}{2}\right)^2}}-1\right)+\sigma^2}{2\sqrt{\left(\rho\left(\alpha-\frac{\sigma^2}{2}\right)\right)^2+2\rho\sigma^2}}>0.$$

Therefore, $\frac{\gamma_+}{r+\lambda}$ is indeed strictly decreasing in λ. Hence, $\frac{\gamma_+}{\gamma_+-1}\frac{r+\lambda-\alpha}{r+\lambda}<\frac{\beta}{\beta-1}\frac{r-\lambda}{r}$ and $\frac{r\gamma_+}{r+\lambda}<\frac{r\beta}{r}$. The latter inequality proves $\beta>\frac{r\gamma_+}{r+\lambda}$. Multiplying the former inequality by $\frac{r(\gamma_+-1)}{r+\lambda-\alpha}$ proves $\frac{\beta}{\beta-1}\frac{(r-\alpha)(\gamma_+-1)}{r+\lambda-\alpha}>\frac{r\gamma_+}{r+\lambda}$. Hence, the right-hand side of (63) is strictly decreasing in $\tilde{X}_A(\theta)$. Therefore, there exists a unique $\tilde{X}_A(\theta)$ that solves (63). To show that $\tilde{X}_A(\theta)<X_A^*(\theta)$, suppose by contradiction that $\tilde{X}_A(\theta)>X_A^*(\theta)$.[18] Since waiting is optimal in $X<\tilde{X}_A(\theta)$, $V_A^b(X,\theta)>\theta X-I+b\ \forall X\in\left[X_A^*(\theta),\tilde{X}_A(\theta)\right)$. Since $V_A^*(X,\theta)=\theta X-I+b\ \forall X\geq X_A^*(\theta)$, we obtain $V_A^b(X,\theta)>V_A^*(X,\theta)\ \forall X\in\left[X_A^*(\theta),\tilde{X}_A(\theta)\right)$, which is a contradiction with $V_A^*(X,\theta)$ being the highest possible value function to the agent across all exercise policies. Thus, $\tilde{X}_A(\theta)<X_A^*(\theta)$.

Next, we show that the strategy profile stated in the proposition constitutes an equilibrium in the communication game, where $\tilde{X}_A(\theta)$ is defined by (63). The IC condition for the agent with $\theta:\tilde{X}_A(\theta)\leq\check{X}$ is satisfied, since it leads to the option being exercised at threshold $\tilde{X}_A(\theta)$, which is the optimal strategy for the agent in the constrained delegation problem, as shown above. The IC condition for the agent with $\theta:\tilde{X}_A(\theta)>\check{X}$ is satisfied, since threshold $\check{X}$ is the highest threshold at which the agent can get the option to be exercised, given the strategy of the principal, and $\check{X}$ dominates any threshold below it by monotonicity of the agent's payoff.

Finally, it remains to derive threshold $\check{X}$, at which it is optimal for the principal to exercise without waiting for the agent's recommendation. Let $V_P^b(X,\theta,Y^*)$ denote the value function to the principal prior to the arrival of the news, conditional on the type of the agent being θ and conditional on the option being exercised at threshold $Y^*\geq X$ prior to the arrival of the news. $\tilde{V}_P^b(X,\theta,Y^*)$ solves

$$(64)\qquad (r+\lambda)\tilde{V}_P^b(X,\theta,Y^*)=\alpha X\frac{\partial\tilde{V}_P^b(X,\theta,Y^*)}{\partial X}+\frac{1}{2}\sigma^2X^2\frac{\partial^2\tilde{V}_P^b(X,\theta,Y^*)}{\partial X^2}+\lambda V_P^a(X,\theta),$$

where

$$(65)\qquad V_P^a(X,\theta)=\begin{cases}\left(\frac{X}{X_P^*(\theta)}\right)^\beta\left(\theta X_P^*(\theta)-I\right), & \text{if } X\leq X_P^*(\theta)\\ \theta X-I, & \text{if } X\geq X_P^*(\theta)\end{cases}$$

is the value of the option to the principal after the arrival of the news. Eq. (64) is solved subject to the value-matching condition $\tilde{V}_P^b(Y^*,\theta,Y^*)=\theta Y^*-I$. Let $V_P^b\left(X,\hat{\theta},\check{X}\right)$ be the principal's value

[18] It is easy to see that $X_A^*(\theta)$ does not solve (63), so $\tilde{X}_A(\theta)\neq X_A^*(\theta)$.

function prior to the arrival of the news, given current state X, posterior belief that θ is uniform over $\left[\underline{\theta}, \hat{\theta}\right]$, if the principal waits for the agent's recommendation to exercise until threshold $\check{X}$:

$$V_P^b\left(X, \hat{\theta}, \check{X}\right) = \int_{\underline{\theta}}^{\hat{\theta}} \tilde{V}_P^b\left(X, \theta, \min\left\{\tilde{X}_A(\theta), \check{X}\right\}\right) \frac{1}{\hat{\theta} - \underline{\theta}} d\theta.$$

Differentiating with respect to $\check{X}$, we obtain the first-order condition that determines $\check{X}$:

$$\int_{\underline{\theta}}^{\tilde{X}_A^{-1}(\check{X})} \frac{\partial \tilde{V}_P^b\left(X, \theta, \check{X}\right)}{\partial \check{X}} d\theta = 0. \tag{66}$$

■

C. Centralized decision-making vs. delegation

In this section, we explore the implications of the results for the optimal allocation of authority. In particular, we compare centralized decision-making, where the principal keeps formal authority and plays the communication game analyzed in Sections III and IV, to delegation, where the principal delegates formal authority to exercise the option to the agent. We first analyze simple, once-and-for-all, delegation, where the principal delegates authority from the beginning and never takes it back. Next, we consider the problem where the delegation policy can be time-contingent.

C.1 Once-and-for-all delegation

First, consider the case of the late exercise bias. Because centralized decision-making implements the optimal commitment mechanism when $b < 0$, the principal is always weaker better off retaining control and getting advice from the agent rather than delegating the exercise decision. Moreover, while delegation and communication are equivalent if $\underline{\theta} = 0$, delegation is strictly inferior to communication if $\underline{\theta} > 0$: Not delegating the decision and playing the communication game implements *constrained* delegation (delegation up to a cutoff), while delegation implements *unconstrained* delegation. This result is illustrated in Figure 1 and summarized in the following corollary.

Corollary to Proposition 1. *If $b < 0$, the principal always weakly prefers retaining control and getting advice from the agent to delegating the exercise decision. The preference is strict if $\underline{\theta} > 0$. If $\underline{\theta} = 0$, retaining control and delegation are equivalent.*

This result contrasts with the implications for static decisions, such as choosing the scale of the project. Dessein (2002) shows that in the leading quadratic-uniform setting of Crawford and Sobel (1982), regardless of the direction of the agent's bias, delegation always dominates communication as long as the agent's bias is not too high so that at least some informative communication is possible. For general payoff functions, Dessein (2002) shows that delegation is optimal if the agent's bias is sufficiently small. In contrast, we show that if the agent favors late exercise, then regardless of the magnitude of his bias, the principal never wants to delegate decision-making authority once-and-for-all. Intuitively, the inability to go back in time allows the principal to commit to follow the recommendations of the agent and ensures that communication is sufficiently effective so that delegation has no further benefit.

Next, consider the case of the early exercise bias. Because the optimal commitment mechanism in Lemma 1 features constrained delegation, simple delegation does not implement the optimal mechanism. However, differently from the case of a late exercise bias, simple delegation can be preferred to centralization if the agent's bias is low enough. The following proposition summarizes our findings.

Proposition C.1. *Suppose $b > 0$, $\underline{\theta} = 0$, and consider the most informative equilibrium of the communication game, ω^*. There exist $\underline{b}$ and $\bar{b}$, such that the principal's expected value in the ω^*-equilibrium is lower than her expected value under delegation if $b < \underline{b}$, and is higher than under delegation if $b > \bar{b}$.*

The result that delegation is beneficial when the agent's bias is small enough is similar to the result of Dessein (2002) for static decisions. This similarity to the static setting is expected, given that the ω^*-equilibrium of the dynamic communication game also exists in the static communication game (see Proposition 3). Intuitively, the principal faces a trade-off: delegation leads to early exercise due to the agent's bias but uses the agent's information more efficiently. When the agent's bias is small enough, the cost from early exercise is smaller than the cost due to the loss of the agent's information, and hence delegation dominates.

C.2 Time-contingent delegation

In a dynamic setting, delegation does not need to be once-and-for-all but can instead be time-contingent: the principal may retain authority for some period of time and delegate it later, or she may take authority back from the agent. In this section, we show that there always exists a time-contingent delegation policy that implements the optimal mechanism.

We first focus on the case of an early exercise bias. Consider the following game: The principal and the agent play the communication game of Section III, but at any time, the principal may delegate decision-making authority to the agent. After authority is granted, the agent retains it until the end of the game and thus is free to choose when to exercise the option. The next result shows that for any $\underline{\theta} \geq 0$, the principal can implement the optimal mechanism by delegating the decision at the right time.

Proposition C.2. *If $b > 0$, there exists the following equilibrium. The principal delegates authority to the agent at the first moment when $X(t)$ reaches the threshold $X_d \equiv \min(\frac{\beta(I+b)}{\beta-1}, \frac{\beta}{\beta-1}\frac{2I}{\underline{\theta}+1})$ and does not exercise the option before that. For any θ, the agent sends message $m = 0$ at any point before he is given authority. If $\theta \geq \frac{I-b}{I+b}$, the agent exercises the option immediately after he is given authority, and if $\theta \leq \frac{I-b}{I+b}$, the agent exercises the option when $X(t)$ first reaches his preferred exercise threshold $X_A^*(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$. The exercise threshold in this equilibrium coincides with the optimal exercise threshold under commitment.*

Intuitively, timing delegation strategically ensures that the information of low types ($\theta \leq \frac{I-b}{I+b}$) is used efficiently, and that all types above $\frac{I-b}{I+b}$ exercise immediately at the time of delegation, exactly as in the optimal contract. The higher is the agent's bias, the later will delegation occur.

It can be similarly shown that when the agent favors late exercise, $b < 0$, the optimal mechanism from Lemma 1 is implemented by the following delegation policy. The principal delegates authority to the agent at the beginning, but then takes authority away and exercises the option at the first moment when $X(t)$ reaches the threshold $\max(\frac{\beta}{\beta-1}\frac{I+b}{\underline{\theta}}, \frac{\beta}{\beta-1}\frac{2I}{\underline{\theta}+1})$. This time-contingent delegation policy is therefore equivalent to centralized decision-making with communication.

Overall, the analysis in this section implies that in the context of timing decisions, the direction of the conflict of interest is the key driver of whether delegating authority adds value. If the agent favors late exercise, delegation, whether time-contingent or once-and-for-all, adds no additional value over centralized decision-making. In contrast, if the agent favors early exercise, once-and-for-all delegation when the agent's bias is small enough, as well as time-contingent delegation, are superior to centralized decision-making.

Proof of Proposition C.1. Let $VD(X, b)$ denote the expected value to the principal under delegation if the current value of $X(t)$ is X. If the decision is delegated to the agent, exercise occurs at threshold $X_A^*(\theta) = \frac{\beta}{\beta-1}\frac{I-b}{\theta}$, and the principal's payoff upon exercise is $\frac{\beta}{\beta-1}(I-b) - I$. Hence,

$$\begin{aligned} VD(X,b) &= \int_0^1 X^\beta \left(\frac{\beta}{\beta-1}\frac{I-b}{\theta}\right)^{-\beta}\left(\frac{\beta}{\beta-1}(I-b)-I\right) d\theta \\ &= \frac{X^\beta}{\beta+1}\left(\frac{\beta}{\beta-1}(I-b)\right)^{-\beta}\left(\frac{\beta}{\beta-1}(I-b)-I\right). \end{aligned}$$

Let $VA(X,b)$ denote the expected value to the principal in the most informative equilibrium of the communication game if the current value of $X(t)$ is X. Using (A8) and $Y(\omega) = \frac{\beta}{\beta-1}\frac{2I}{1+\omega}$,

$$VA(X,b) = X^\beta \frac{1-\omega^*(b)}{1-\omega^*(b)^{\beta+1}}\left(\frac{\beta}{\beta-1}\frac{2I}{1+\omega^*(b)}\right)^{-\beta}\frac{I}{\beta-1},$$

where $\omega^*(b)$ is the unique solution to $Y(\omega) = \frac{\beta}{\beta-1}\frac{2I}{\omega+1}$, given b. Because X^β enters as a multiplicative factor in both $VD(X,b)$ and $VA(X,b)$, it is sufficient to compare $VD(b)$ and $VA(b)$, where $VD(b) \equiv X^{-\beta}VD(X,b)$ and $VA(b) \equiv X^{-\beta}VA(X,b)$. First, consider the behavior of $VA(b)$ and $VD(b)$ around $b = I$. According to the supplementary analysis for the proof of Proposition 2 in the Online Appendix, $\lim_{b\to I}\omega^*(b) = 0$. Hence, $\lim_{b\to I} VD(b) = -\infty$ and $\lim_{b\to I} VA(b) = \left(\frac{\beta}{\beta-1}2I\right)^{-\beta}\frac{I}{\beta-1}$. By continuity of $VD(b)$ and $VA(b)$ in b, this implies that there exists $\bar{b} \in (0,I)$, such that for any $b > \bar{b}$, $VA(b) > VD(b)$. In other words, communication dominates delegation if the conflict of interest between the agent and the principal is big enough. Second, consider the behavior of $VA(b)$ and $VD(b)$ for small but positive b. Below, we prove that $\lim_{b\to 0}\frac{VA'(b)}{b} = -\infty$ and $\lim_{b\to 0}\frac{VD'(b)}{b} = -\frac{\beta^2}{\beta^2-1}\left(\frac{\beta}{\beta-1}I\right)^{-\beta-1} > -\infty$. By continuity of $VA'(b)$ and $VD'(b)$ for $b > 0$, there exists $\underline{b} > 0$ such that $VA'(b) < VD'(b)$ for any $b < \underline{b}$. Because $VA(0) = VD(0)$, $VD(b) - VA(b) = \int_0^b (VD'(y) - VA'(y))\,dy > 0$ for all $b \in (0,\underline{b}]$. Thus, delegation dominates communication if the agent favors early exercise but the bias is low enough.

Proof that $\lim_{b\to 0}\frac{VD'(b)}{b} = -\frac{\beta^2}{\beta^2-1}\left(\frac{\beta}{\beta-1}I\right)^{-\beta-1}$. By l'Hopital's rule,

$$\lim_{b\to 0+} VD(b) = \lim_{b\to 0+} VA(b) = \frac{1}{\beta+1}\left(\frac{\beta}{\beta-1}I\right)^{-\beta}\frac{I}{\beta-1}.$$

Note that $VD'(b) = -\frac{\beta b}{(\beta+1)(I-b)}\left(\frac{\beta}{\beta-1}(I-b)\right)^{-\beta}$. In particular, it follows that $\lim_{b\to 0+} VD'(b) = 0$ and $\lim_{b\to 0}\frac{VD'(b)}{b} = -\frac{\beta^2}{\beta^2-1}\left(\frac{\beta}{\beta-1}I\right)^{-\beta-1}$.

Proof that $\lim_{b\to 0}\frac{VA'(b)}{b} = -\infty$. The derivative of $VA(b)$ with respect to b can be found as

$$VA'(b) = C\frac{d\omega^*(b)}{db}\left[\frac{(1-\omega)(1+\omega)^\beta}{1-\omega^{\beta+1}}\right]'|_{\omega=\omega^*(b)}, \tag{67}$$

where $C \equiv \left(\frac{\beta}{\beta-1}2I\right)^{-\beta}\frac{I}{\beta-1}$. Recall that $\omega^*(b)$ solves (52), which is equivalent to

$$\frac{2I}{I-b}\frac{\beta}{\beta-1} = \left(\frac{1}{\omega}+1\right)\frac{1-\omega^\beta}{1-\omega^{\beta-1}}. \tag{68}$$

Differentiating this equation, we get

$$\begin{aligned} &\frac{2I}{(I-b)^2}\frac{\beta}{\beta-1}db \\ = \; &\frac{-\left(1-\omega^\beta\right)\left(1-\omega^{\beta-1}\right)+(1+\omega)\,\omega\left(-\beta\omega^{\beta-1}\left(1-\omega^{\beta-1}\right)+(\beta-1)\,\omega^{\beta-2}\left(1-\omega^\beta\right)\right)}{\omega^2\left(1-\omega^{\beta-1}\right)^2}d\omega. \end{aligned} \tag{69}$$

Because (68) is equivalent to $\frac{1}{I-b} = \frac{1}{2I}\frac{\beta-1}{\beta}\frac{1+\omega}{\omega}\frac{1-\omega^\beta}{1-\omega^{\beta-1}}$, we can rewrite the left-hand side of (69) as

$$2I\frac{\beta}{\beta-1}\left(\frac{1}{2I}\right)^2\left(\frac{\beta-1}{\beta}\right)^2\frac{(1+\omega)^2}{\omega^2}\frac{\left(1-\omega^\beta\right)^2}{\left(1-\omega^{\beta-1}\right)^2}db.$$

Substituting this into (69) and simplifying, we get

$$\frac{d\omega}{db}|_{\omega=\omega^*(b)} = \frac{1}{2I}\frac{\beta-1}{\beta}\frac{(1+\omega)^2\left(1-\omega^\beta\right)^2}{-\left(1-\omega^\beta\right)\left(1-\omega^{\beta-1}\right)+(1+\omega)\,\omega^{\beta-1}\left(-\beta\omega+\beta-1+\omega^\beta\right)}. \tag{70}$$

Plugging (70) and

$$\left[\frac{(1-\omega)(1+\omega)^\beta}{1-\omega^{\beta+1}}\right]' = \frac{(1+\omega)^{\beta-1}}{\left(1-\omega^{\beta+1}\right)^2}\left[(\beta-1)\left(1-\omega^{\beta+1}\right)-(\beta+1)\left(\omega-\omega^\beta\right)\right],$$

into (67), we get

$$VA'(b) = -D\frac{(1+\omega)^{\beta+1}\left(1-\omega^\beta\right)^2\left[(\beta-1)\left(1-\omega^{\beta+1}\right)-(\beta+1)\left(\omega-\omega^\beta\right)\right]}{\left(1-\omega^{\beta+1}\right)^2\left[\left(1-\omega^\beta\right)\left(1-\omega^{\beta-1}\right)-(1+\omega)\,\omega^{\beta-1}\left(-\beta\omega+\beta-1+\omega^\beta\right)\right]},$$

where $D \equiv \frac{C}{2I}\frac{\beta-1}{\beta}$. To find $\lim_{b\to 0}\frac{VA'(b)}{b}$, we express $\frac{1}{b}$ from (68) as

$$\frac{1}{b} = \frac{(\beta-1)(1+\omega)\left(1-\omega^\beta\right)}{I\left[(\beta-1)(1+\omega)\left(1-\omega^\beta\right)-2\beta\omega\left(1-\omega^{\beta-1}\right)\right]},$$

and hence

$$\begin{aligned} \frac{VA'(b)}{b} &= -D\frac{(1+\omega)^{\beta+1}\left(1-\omega^\beta\right)^2\left[(\beta-1)\left(1-\omega^{\beta+1}\right)-(\beta+1)\left(\omega-\omega^\beta\right)\right]}{\left(1-\omega^{\beta+1}\right)^2\left[\left(1-\omega^\beta\right)\left(1-\omega^{\beta-1}\right)-(1+\omega)\omega^{\beta-1}\left(-\beta\omega+\beta-1+\omega^\beta\right)\right]}\frac{(\beta-1)(1+\omega)\left(1-\omega^\beta\right)}{I\left[(\beta-1)(1+\omega)\left(1-\omega^\beta\right)-2\beta\omega\left(1-\omega^{\beta-1}\right)\right]} \\ &= -\frac{(\beta-1)D}{I}\frac{(1+\omega)^{\beta+2}\left(1-\omega^\beta\right)^3}{\left(1-\omega^{\beta+1}\right)^2\left[\left(1-\omega^\beta\right)\left(1-\omega^{\beta-1}\right)-(1+\omega)\omega^{\beta-1}\left(-\beta\omega+\beta-1+\omega^\beta\right)\right]}. \end{aligned}$$

Hence, $\lim_{b\to 0}\frac{VA'(b)}{b}$ equals

$$-\frac{(\beta-1)\,2^{\beta+2}D}{I}\lim_{\omega\to 1}\left[\frac{1-\omega^{\beta}}{1-\omega^{\beta+1}}\right]^2\lim_{\omega\to 1}\left[\frac{1-\omega^{\beta}}{(1-\omega^{\beta})(1-\omega^{\beta-1})-(1+\omega)\,\omega^{\beta-1}(-\beta\omega+\beta-1+\omega^{\beta})}\right].$$

By l'Hopital's rule, the first limit equals $(\frac{\beta}{\beta+1})^2$, and the second limit equals ∞, which completes the proof. ■

Proof of Proposition C.2. Note that the following three inequalities are equivalent: $b\leq\frac{1-\underline{\theta}}{1+\underline{\theta}}I\Leftrightarrow\frac{I-b}{I+b}\geq\underline{\theta}\Leftrightarrow\frac{\beta}{\beta-1}(I+b)\leq\frac{\beta}{\beta-1}\frac{2I}{\underline{\theta}+1}$. Hence, there are two cases. If $b<\frac{1-\underline{\theta}}{1+\underline{\theta}}I$, delegation occurs at threshold $\frac{\beta}{\beta-1}(I+b)=X_A^*\left(\frac{I-b}{I+b}\right)$, where $\frac{I-b}{I+b}>\underline{\theta}$. If $b\geq\frac{1-\underline{\theta}}{1+\underline{\theta}}I$, then $\frac{I-b}{I+b}\leq\underline{\theta}$ and delegation occurs at the principal's uninformed exercise threshold $\frac{\beta}{\beta-1}\frac{2I}{\underline{\theta}+1}$.

We prove that neither the agent nor the principal wants to deviate from the specified strategies. First, consider the agent. Given Assumption 1, sending a message $m=1$ is never beneficial because it does not change the principal's belief and hence her strategy. Hence, the agent cannot induce exercise before he is given authority. After the agent is given authority, his optimal strategy is to: 1) exercise immediately if $b\geq I$, or if $b<I$ and $X_d\geq X_A^*(\theta)$; 2) exercise when $X(t)$ first reaches $X_A^*(\theta)$ if $b<I$ and $X_d<X_A^*(\theta)$. Consider two cases. If $0<b<\frac{1-\underline{\theta}}{1+\underline{\theta}}I$ $(\leq I)$, then $X_d=X_A^*\left(\frac{I-b}{I+b}\right)$, and hence $X_d<X_A^*(\theta)$ if and only if $\theta<\frac{I-b}{I+b}$. Thus, types below $\frac{I-b}{I+b}$ exercise at $X_A^*(\theta)$ and types above $\frac{I-b}{I+b}$ exercise immediately at X_d, consistent with the equilibrium strategy. Second, if $b\geq\frac{1-\underline{\theta}}{1+\underline{\theta}}I$, the agent finds it optimal to exercise immediately at X_d regardless of his type: if $b\geq I$, this is always the case, and if $\frac{1-\underline{\theta}}{1+\underline{\theta}}I\leq b<I$, this is true because $X_A^*(\theta)\leq X_A^*(\underline{\theta})=\frac{\beta}{\beta-1}\frac{I-b}{\underline{\theta}}\leq\frac{\beta}{\beta-1}\frac{2I}{\underline{\theta}+1}=X_d$. Since $\frac{I-b}{I+b}\leq\underline{\theta}$, this strategy again coincides with the equilibrium strategy. Hence, the agent does not want to deviate.

Next, consider the principal. The above arguments show that the equilibrium exercise times coincide with the exercise times under the optimal mechanism in Lemma 1 for all b. Hence, the principal's expected utility in this equilibrium equals her expected utility in the optimal mechanism. Consider possible deviations of the principal, taking into account that the agent's messages are uninformative and hence the principal does not learn new information by waiting. First, the principal can exercise the option himself, before or after $X(t)$ first reaches X_d. Because a mechanism with such an exercise policy is incentive-compatible, the principal's utility from such a deviation cannot exceed her utility under the optimal mechanism and hence her equilibrium utility. Thus, such a deviation cannot be strictly profitable. Second, the principal can deviate by delegating authority to the agent before or after $X(t)$ first reaches X_d. An agent who receives authority at some point t will exercise immediately if $b\geq I$, or if $b<I$ and $X(t)\geq X_A^*(\theta)$, and will exercise when $X(t)$ first reaches $X_A^*(\theta)$ otherwise. Because a mechanism with such an exercise schedule is incentive-compatible, the principal's utility from this deviation cannot exceed

her utility under the optimal mechanism and hence her equilibrium utility. Hence, the principal does not want to deviate either. ■

D. Robustness

In this section, we show the robustness of the results to several versions of the model.

D.1 Simple compensation contracts

A reasonable question is whether simple compensation contracts, such as paying a fixed amount for exercise (if $b < 0$) or for the lack of exercise (if $b > 0$), can solve the problem and thus make the analysis less relevant. We show that this is not the case. Specifically, we allow the principal to offer the agent the following payment scheme. If the agent is biased towards late exercise ($b < 0$), the principal can promise the agent a lump-sum payment z that he will receive as soon as the option is exercised. A higher payment decreases the conflict of interest and speeds up option exercise. For example, if $z = \frac{-b}{2}$, the agent's and the principal's interests are aligned because each of them receives $\theta X - I + \frac{b}{2}$ upon exercise. However, a higher payment is also more expensive for the principal. Because of that, as the next result shows, it is always optimal for the principal to offer $z^* < \frac{-b}{2}$, and hence the conflict of interest will remain. Moreover, if the agent's bias is sufficiently small, the optimal payment is in fact zero.

Similarly, if the agent is biased towards early exercise ($b > 0$), the principal can promise the agent a flow of payments $\hat{z}dt$ up to the moment when the option is exercised. Higher $\hat{z}$ aligns the interests of the players but is expensive for the principal. The next result shows that if the initial value of the state process is sufficiently small, the optimal $\hat{z}$ is again zero. In numerical analysis, we also show that similarly to the late exercise bias case, the optimal payment is smaller than the payment that would eliminate the conflict of interest.

Proposition D.1. *Suppose $b < 0$ and the principal can promise the agent a payment $z \geq 0$ upon exercise. Then the optimal z is always strictly smaller than $\frac{-b}{2}$ and equals zero if $b > \frac{-I}{\beta - 1}$. Suppose $b > 0$ and the principal can promise the agent a flow of payments $\hat{z}dt \geq 0$ up to the moment of option exercise. Then the optimal $\hat{z}$ equals zero if $X(0)$ is sufficiently small.*

Thus, allowing simple compensation contracts often does not change the problem at all, and at most leads to an identical problem with a different bias b. We conclude that the problem and implications of our paper are robust to allowing simple compensation contracts.

Proof of Proposition D.1. First, consider $b < 0$. The payoffs of the principal and the agent upon exercise are given by $\theta X - I - z$ and $\theta X - I + b + z$, respectively. Hence, the problem is equivalent to the problem of the basic model with $I' \equiv I + z$ and $b' = 2z + b$. The interests of the principal and the agent become aligned if $b' = 0$, i.e., if $z = \frac{-b}{2}$. Note that it is never optimal to have $z > 0$ if $b' < -I'$: in this case, the equilibrium will feature uninformed exercise and hence would give the principal the same expected utility as if she did not make any payments. Similarly, it is never optimal to have $b' > 0$. Hence, we can restrict attention to $b' \in [-I, 0]$. Then, the most informative equilibrium of the communication game features continuous exercise, and according to (A1), the principal's expected utility as a function of z is

$$\begin{aligned} V(z) &= \frac{X(0)^{\beta}}{\beta+1}\left(\frac{\beta}{\beta-1}(I'-b')\right)^{-\beta}\frac{I'-\beta b'}{\beta-1} \\ &= \frac{X(0)^{\beta}}{\beta^2-1}\left(\frac{\beta}{\beta-1}\right)^{-\beta}(I-b-z)^{-\beta}\left(I-\beta b+z(1-2\beta)\right). \end{aligned}$$

Note that $V'(z) > 0 \Leftrightarrow z < z^*$, where $z^* = \frac{-(I-b)(2\beta-1)+\beta I-\beta^2 b}{(\beta-1)(2\beta-1)}$. It is easy to show that $z^* > 0 \Leftrightarrow b < \frac{-I}{\beta-1}$ and that $z^* < -\frac{b}{2} \Leftrightarrow (\beta-1)(b-2I) < 0$, which holds for any $b < 0$. This completes the proof of the first statement.

Next, consider $b > 0$. If the principal makes flow payoffs $\hat{z}dt$ before exercise, then upon exercise the agent loses $\frac{\hat{z}}{r}$, which is the present value of continuation payments at that moment. Thus, the principal's and agent's effective payoffs upon exercise are $\theta X(t) - I + \frac{\hat{z}}{r}$ and $\theta X(t) - I + b - \frac{\hat{z}}{r}$, respectively. Hence, we can consider the communication game with $I' = I - \frac{\hat{z}}{r}$ and $b' = b - 2\frac{\hat{z}}{r}$. The interests of the principal and the agent become aligned if $b = 2\frac{z}{r}$, i.e., if $z = \frac{rb}{2}$. Similarly to the case $b < 0$, it is never optimal to have $\hat{z} > 0$ if $b' \geq I'$ or $b' < 0$, and hence we can restrict attention to $b' \in [0, I')$. Denoting $\tilde{z} \equiv \frac{\hat{z}}{r}$ and using (A8), the payoff of the principal at the initial date is

$$V(\tilde{z}) = -\tilde{z} + \frac{1-\omega}{1-\omega^{\beta+1}}\left(\frac{X(0)}{Y(\omega,\tilde{z})}\right)^{\beta}\left(\frac{1}{2}(1+\omega)Y(\omega,\tilde{z}) - I + \tilde{z}\right),$$

where by (6), $Y(\omega,\tilde{z}) = \frac{\left(1-\omega^{\beta}\right)(I-b+\tilde{z})}{\omega\left(1-\omega^{\beta-1}\right)}$. By (52), the most informative equilibrium of this game is characterized by $\omega = \frac{1}{\frac{\beta}{\beta-1}\frac{1-\omega^{\beta-1}}{1-\omega^{\beta}}\frac{2(I-\tilde{z})}{I-b+\tilde{z}}-1}$. If $X(0) \to 0$, $V'(\tilde{z}) \to -1$, and hence $\tilde{z} = 0$ is optimal, which completes the proof. ■

D.2 Model with different discount rates

In our basic setup, the conflict of interest between the agent and the principal is modeled by the agent's bias b. Our results are similar in an alternative setup, where the conflict of interest arises because the agent and the principal have different discount rates. This section presents the summary of this analysis, and the full analysis is available from the authors upon request.

Suppose that the agent's discount rate is r_A, the principal's discount rate is r_P, and both players' payoff from exercise at time t is $\theta X(t) - I$. Similar to the basic model, we can define β_A and β_P, where β_i is the positive root of the quadratic equation $\frac{1}{2}\sigma^2\beta_i(\beta_i - 1) + \alpha\beta_i - r_i = 0$.

The case where the principal is more impatient than the agent ($r_P > r_A$, or equivalently, $\beta_P > \beta_A$) is similar to the case $b < 0$ in the basic model. We show that if $\underline{\theta} = 0$, then as long as $\beta_A > \frac{2\beta_P}{1+\beta_P}$, there exists an equilibrium with continuous exercise in which exercise occurs at the agent's most preferred threshold $\frac{\beta_A}{\beta_A - 1}\frac{I}{\theta}$. If $\underline{\theta} > 0$, the equilibrium features continuous exercise up to a cutoff. The case where the agent is more impatient than the principal ($r_P < r_A$) is similar to the case $b > 0$ in the basic model. We show that the equilibrium with continuous exercise does not exist and derive the analog of Proposition 2. Specifically, in the most informative stationary equilibrium, exercise is unbiased given the principal's information. This equilibrium is characterized by $\tilde{\omega}^* < 1$, which is the unique solution of

$$\frac{\left(1 - \omega^{\beta_A}\right) I}{\omega\left(1 - \omega^{\beta_A - 1}\right)} = \frac{\beta_P}{\beta_P - 1}\frac{2I}{\omega + 1}.$$

In addition, for any $\omega \in [\underline{\tilde{\omega}}, \tilde{\omega}^*)$, where $0 < \underline{\tilde{\omega}} < \tilde{\omega}^*$, there is a unique ω-equilibrium where exercise happens with delay.

D.3 Put option

So far, we have assumed that the decision problem is over the timing of exercise of a call option, such as the decision of when to invest. In this section, we show that if the decision problem is over the timing of exercise of a put option, such as the decision of when to liquidate a project, the analysis and economic insights are similar. The nature of the option, call or put, is irrelevant for the results. What matters is the asymmetric nature of time: Time moves forward and thereby creates a one-sided commitment device for the principal to follow the agent's recommendations.

Consider the model of Section I with the following change. The exercise of the option leads to the payoffs $\theta I - X(t)$ and $\theta(I + b) - X(t)$ for the principal and the agent, respectively. As before, θ is a random draw from a uniform distribution on $[\underline{\theta}, 1]$ and is privately learned by the agent at the initial date. If $\underline{\theta} = 0$, the model exhibits stationarity. For example, if the decision represents shutting down a project, $I\theta$ corresponds to the salvage value of the project, $b\theta$ represents the agent's private cost (if $b < 0$) or benefit (if $b > 0$) of liquidating the project, and $X(t)$ corresponds to the present value of the cash flows from keeping the project afloat. The solution of this model follows the same structure as the solution of the model with the call option. We summarize our findings below, and the full analysis is available from the authors upon request.

Suppose that we start with a high enough $X(0)$, so that immediate exercise does not happen. At the beginning of Online Appendix B, we show that if θ were known, the optimal exercise policy of each player would be given by a lower trigger on $X(t)$: $X_P^{**}(\theta) = \frac{\delta}{\delta+1} I\theta$, $X_A^{**}(\theta) = \frac{\delta}{\delta+1}(b+I)\theta$, where $-\delta$ is the negative root of the quadratic equation that defined β. If $b > 0$, then $X_A^{**}(\theta) > X_P^{**}(\theta)$, i.e., the agent's preferred exercise policy is to exercise earlier than the principal. Similarly, if $b < 0$, the agent is biased towards late exercise.

Suppose that $\underline{\theta} = 0$ and consider the communication game like the one in Section III. If $b \in (-\frac{I}{2}, 0)$, there is an equilibrium with full information revelation: The agent recommends waiting as long as $X(t)$ exceeds his preferred exercise threshold $X_A^{**}(\theta)$ and recommends exercising at the first moment when $X(t)$ hits $X_A^{**}(\theta)$. Upon getting the recommendation to exercise, the principal realizes it is too late and finds it optimal to exercise immediately. Prior to that, the principal prefers to wait because the value of learning θ exceeds the cost of delay. If $b > 0$, this equilibrium does not exist, and all stationary equilibria are of the form $\{(\omega, 1), (\omega^2, \omega), ...\}$, where type $\theta \in (\omega^n, \omega^{n-1})$ recommends exercise at threshold $\omega^{n-1} Y_{put}(\omega)$, where $Y_{put}(\omega) = \frac{\omega - \omega^{\gamma+1}}{1-\omega^{\gamma+1}}(I+b)$.

References

[1] Amador, Manuel, and Kyle Bagwell (2013). The theory of optimal delegation with an application to tariff caps, *Econometrica* 81, 1541-1599.

[2] Jameson, Graham, J. O. (2006). Counting zeros of generalized polynomials: Descartes' rule of signs and Laguerre's extensions, *Mathematical Gazette* 90, 518, 223-234.

[3] Laguerre, Edmond N. (1883). Sur la theorie des equations numeriques, *Journal de Mathematiques pures et appliquees* 9, 3-47. (Translation by S.A. Levin, Stanford University).

[4] Milgrom, Paul, and Ilya Segal (2002). Envelope theorems for arbitrary choice sets, *Econometrica* 70, 583-601.