Conditional and Lenient Drug Approvals

Chenxi Xu, Giuseppe Lopomo, David Ridley, Peng Sun*

March 24, 2026

Abstract

Firms are often allowed to sell new drugs based on preliminary efficacy evidence, with final approval contingent on confirmatory testing. We characterize optimal approval policies for a government agency that maximizes consumer welfare while considering the firm's incentives. We assume that firms have private information about testing costs and the government's payoff depends on the drug's expected efficacy. The optimal policy may include partial conditional approval for drugs with low expected efficacy, and leniency in granting final approval below conventional standards. Our calibration suggests these tools can generate hundreds of millions of dollars in annual social value.

*The authors are affiliated with Duke University's Fuqua School of Business, Durham, NC 27708-0120. The authors appreciate feedback from Angelique Acquatella, Eric Budish, Pierre Dubois, Jennifer Kao, Margaret Kyle, Erzo Luttmer, Kathleen Miller, Paul Novosad, Marco Ottaviani, Jonathan Skinner, Christopher Snyder, Ariel Stern, Joe Stevens, and seminar participants at the American Society of Health Economists, the City University of Hong Kong, Congressional Budget Office, Dartmouth College, Duke University, the Shanghai University of Finance and Economics, the Toulouse School of Economics, Tsinghua University, the University of Chicago, and conference participants at the CEPR-TSE Health Economics Conference and the INFORMS annual meeting. The authors received no external funding for this research.

1 Introduction

In 1988, HIV–AIDS activists protesting outside the Food and Drug Administration (FDA) chanted, "Hey, hey, FDA, how many people have you killed today?" With tens of thousands dying annually from HIV–AIDS, activists demanded access to promising antiviral drugs still under development (United Press International, 1988). In response, the FDA introduced a conditional approval mechanism called "accelerated approval." In 2024, the FDA granted conditional approval to seven drugs, six of which targeted rare diseases (Food and Drug Administration, 2025). Similar mechanisms exist in Europe and Japan.

Conditional approval allows a new drug to reach the market after Phase II testing, with full approval contingent on confirmatory Phase III tests. Phase II tests are relatively short and inexpensive but provide inconclusive evidence. Phase III tests are costly and time-consuming but provide statistically and clinically meaningful evidence of efficacy. Without conditional approval, firms may forgo confirmatory testing for slow-progressing diseases, which require long and large trials (Budish et al., 2015).[1]

Conditional approval benefits patients through earlier access and encourages investment by increasing revenue. Firms can generate sales while funding costly Phase III testing, making marginally profitable drugs viable. Revenue rises immediately through early sales and later through diffusion.

We use the term "leniency" to describe the government agency's ability to commit to a final approval threshold below what it would apply if choosing myopically. Leniency may involve approving drugs without fully persuasive confirmatory trials or relying on surrogate endpoints (e.g., tumor shrinkage) rather than clinical outcomes (e.g., survival). Indeed, 9 of 46 cancer indications granted final approval after conditional approval lacked statistically significant evidence of clinical benefit in the published literature (Liu et al., 2024). The government can also commit to leniency before Phase III; for example, the FDA's "breakthrough" designation provides firms with enhanced interaction and early guidance on the evidence required for approval.[2]

Leniency is always socially costly, as it raises the risk of approving drugs that are ef-

[1] More generally, regulatory restrictions and testing costs can stifle innovation and reduce social welfare (Peltzman, 1973; Philipson Sun, 2008; Grennan Town, 2020).

[2] The government can be interpreted as the drug regulator, such as the FDA, if the regulator is understood to consider social benefits broadly. Under a pilot program launched in 2025, for example, the FDA awards faster review based on criteria that include novelty, unmet need, and increasing affordability—that is, lower prices (U.S. Food and Drug Administration, 2025). Alternatively, the government may be interpreted as the Department of Health and Human Services, which oversees both the FDA (drug approval) and the Centers for Medicare & Medicaid Services (drug pricing).

fectively costly placebos. Conditional approval, by contrast, may or may not be costly, depending on expected efficacy and the extent to which earlier access boosts adoption.

Government agencies have limited information about testing costs for five reasons. First, disclosing true costs may allow the government to use price controls to extract a significant share of the firm's profit. Second, cost benchmarks are limited because conditionally approved drugs are often first-in-class. Third, Phase III programs require firm-specific manufacturing scale-up that is difficult for regulators to observe. Fourth, most conditional approvals involve rare diseases, where patient recruitment is a major in-house cost borne through firms' long-standing relationships with patient groups and clinicians. Fifth, clinical programs share resources across multiple drugs, making it difficult for outside auditors to infer the costs of individual drugs. As Light Warburton (2005) observe, "this science-based industry refuses to allow independent parties to check the validity of their cost data."[3]

Private information about Phase III testing costs motivates our mechanism-design approach, yielding optimal conditional approval and leniency thresholds based on the government's beliefs about testing costs and the drug's potential efficacy. Conceptually, the government's problem is a procurement problem in which the firm supplying the drug can be "paid" in two ways: conditional approval and leniency.

Three features distinguish the government's problem from standard procurement. First, the final approval decision depends on a testing result, which does not arise in procurement settings. Second, the two forms of payment are bounded: leniency and conditional approval cannot exceed feasible limits on evidence requirements and market size. Third, the government's payoff is nonlinear in leniency, precluding standard mechanism-design tools.

If the government believes testing costs are sufficiently small, the optimal policy uses no costly incentives. Conditional approval is granted only when the drug is ex ante promising, and there is no leniency. If, instead, testing costs are high enough to deter investment, the government can implement a take-it-or-leave-it offer by setting a scale of conditional approval and a threshold for final approval. The firm accepts only if its testing cost falls below a threshold. Optimal conditional approval and leniency balance the goal of inducing confirmatory testing against the risk of paying for expensive placebos.

[3]We focus on private information about testing costs, though firms may also have private information about efficacy. Both high costs and low expected efficacy can deter development, motivate government incentives, and create information rents. The distinction between the two matters most when the government provides funding up front and ex post profit is not positive, as with philanthropic support for tropical disease research. In such settings, private information about expected efficacy creates a moral hazard problem: the firm has little incentive to exert effort because it already has the funding and has limited gain from success (Ridley et al., 2025).

We begin with a two-period model in which the government's problem can be formulated as an (infinite-dimensional) linear program. The two-period model is appropriate for rare-disease drugs, which face limited generic competition and are exempt from price regulation. Duality theory implies that a take-it-or-leave-it offer is indeed optimal, and yields its full characterization for all parameter values in our framework.

We then analyze a three-period model with price declines due to generic competition or price regulation. We derive a linear relaxation that gives an upper bound on the government's payoff and construct a feasible (and easy to implement) policy that gives a lower bound. In our calibrated numerical analysis, these bounds are remarkably close.

Our analysis yields several recommendations for lawmakers. First, governments should be permitted to conditionally approve drugs with low expected efficacy. Second, they should have the flexibility to grant partial conditional approval, such as restricting access to specific patient subgroups; insurance coverage can likewise be limited. Third, governments should be allowed to use leniency in setting the final approval threshold below the myopic standard.

1.1 Literature Review

Myerson (1981) laid the foundation for optimal mechanism design. Baron Myerson (1982) extended the framework to regulatory settings with adverse selection. In their model, a firm privately knows its production cost, and the regulator maximizes a weighted average of consumer welfare and firm profit. The basic Myersonian framework relies on two instruments: quantity produced and money transferred. In our setting, the government uses testing requirements, conditional approval, and final approval as instruments. This replaces monetary transfers with regulatory tools and introduces additional "budget" constraints, i.e., limits on the scale of conditional approval and the degree of leniency.

Some diseases lack commercial prospects because they are concentrated in lower-income countries, where profit margins are near zero (Ridley et al., 2025). In such cases, conditional approval may help patients but offers little incentive to firms, because speeding an unprofitable drug to market does not improve returns. For tropical diseases, push and pull mechanisms are more effective than conditional approval. Push mechanisms subsidize early-stage testing, often through agencies such as the National Institutes of Health (NIH) or the Gates Foundation. Pull mechanisms raise revenues, for example through advance market commitments that guarantee purchases (Berndt et al., 2007; Kremer et al., 2022), or transferable vouchers such as U.S. priority review vouchers (Gans Ridley, 2013) and exclusivity-extension vouchers (Dubois et al., 2022).

For diseases that are profitable ex post but not ex ante due to long testing timelines (Budish et al., 2015), conditional approval offers two advantages over push and pull funding. First, it grants early access and accelerates diffusion, generating benefits both immediately and over time. Second, it is often more politically feasible than direct subsidies to drug makers, which remain unpopular. Indeed, NIH funding is directed largely to universities for early-stage testing rather than to firms for late-stage trials (Zhou et al., 2023). Push, pull, and conditional approval are not mutually exclusive, however. When calibrating the model, push funding could be incorporated as reducing testing costs, and pull funding could increase revenue, complementing conditional approval.

This paper also contributes to the literature on optimal regulatory approval. Carpenter Ting (2007) analyze a government-firm interaction where approval decisions rely on imperfect information. They examine the trade-offs between false positives and negatives in regulatory decisions. Conditional approval often depends on surrogate endpoints, which must be carefully validated to ensure reliability for governments (Bognar et al., 2017). Beyond conditional approval, governments employ pathways like the FDA's "breakthrough" designation to facilitate market entry (Chandra et al., 2024).

Permissive conditional approval policies may lead to withdrawals of drug indications if further testing reveals insufficient efficacy. Orlov et al. (2020) show that firms might delay reporting negative information to postpone withdrawals. Xu et al. (2021) demonstrate that firms balance regulatory costs by choosing between extended and shorter testing periods. To address delays, government agencies like the FDA now require firms to initiate confirmatory testing before granting conditional approval. The dynamics differ for most medical devices, especially those similar to previous devices. Testing is less rigorous, markets see frequent entry, and competition is dynamic (Grennan Town, 2020; Collard-Wexler et al., 2024).

Some papers study continuous-time models in which approval decisions evolve with accumulating evidence. Henry Ottaviani (2019) extends Wald's sequential hypothesis testing model to a game between a firm and a government. The firm conducts costly testing, generating a publicly observed Brownian motion with drift based on unobservable efficacy. The firm benefits from approval; the government's payoff depends on true efficacy. The firm chooses when to stop and seek approval. Commitment by the government leads to Pareto improvements over outcomes without commitment. Henry et al. (2022) allow for approval revocation when negative evidence emerges. The results highlight the importance of flexible regulatory policies.

The continuous-time models are ideal for settings with adaptable protocols, such as early-

stage drug testing in animals or review of drugs after final approval. However, discrete time is sufficient in settings such as late-stage human testing which involves fixed endpoints, with outcomes revealed simultaneously.[4] Furthermore, our model captures the diffusion effect triggered by conditional approval, which is not present in Henry et al. (2022).

McClellan (2022) studies a continuous-time model in which the government chooses when to stop experimentation and approve a drug, accounting for the firm's incentive to quit early. Even with symmetric information, the approval threshold falls as the firm nears exit. Under asymmetric information, more promising drugs receive a "fast track" with a lower threshold and an efficacy floor. Like their model, ours allows leniency to keep makers of marginal drugs from quitting. Our key difference is that we model conditional approval, which lets firms earn revenue and begin diffusion while confirmatory testing is ongoing.

2 Model

We begin by providing an intuitive overview of the model. We connect several key elements of the model to stylized facts listed in Appendix 7. The formal model is presented in Section 2.2.

2.1 Overview and Stylized Facts Fitting the Model

We study how a government agency should decide whether to grant conditional approval, that is, allow early sales of a new drug to a fraction of the market while the drug's efficacy is still undergoing confirmatory testing. Early access can benefit patients and help offset testing costs, but it also risks wasting resources if the drug ultimately proves ineffective.

We model the interaction as a three-period game between a government agency and a pharmaceutical firm. Before the game begins, the firm has completed preliminary clinical testing.

In the first period, the government (i) decides whether to request final confirmatory testing (Fact **??**), (ii) decides whether to allow early sales through conditional approval, possibly limited to a fraction of the market (Fact **??**), and (iii) commits to the statistical threshold that will determine final approval based on the testing outcome. The government must balance incentives for testing and early access against the risk of wasting money on ineffective drugs. The firm then decides whether to undertake confirmatory testing. If testing

[4]For a discussion of the discrete time assumption, see Fact **??** in the appendix.

is not induced or the firm declines, the game ends with zero payoffs for both parties.

In the second period, the firm and the government observe the outcome of confirmatory testing. Also, the government implements the final approval decision in accordance with the threshold announced in the first period.

In the third period, there is no strategic interaction. If the drug was approved in the second period, its price falls due to generic competition or price regulation, eliminating further profit for the firm while generating additional surplus for the government (Figure 1). Appendix 9 discusses how the length of the monopoly period affects optimal policy.

If the government approves the drug for a given period, the firm earns monopoly profits during that period (Fact **??**). For drugs with high testing costs, profits may be positive only if conditional approval is granted at a sufficiently large scale (Fact **??**). Without conditional approval, cumulative profit may remain negative throughout the relevant horizon (Figure 2), leading the firm to forgo testing. Hence, conditional approval is a means of motivating testing. The government does not directly finance testing (Fact **??**).

Conditional approval benefits the firm not only by allowing earlier sales, but also by increasing later sales through faster adoption. This is because sales do not arrive all at once. Instead, they grow gradually as doctors and patients become familiar with the drug (Fact **??**), as illustrated in the top panel of Figure 2.

Preliminary testing is informative but typically insufficient to establish statistically significant efficacy, leaving residual uncertainty about whether the drug is sufficiently effective (Fact **??**). Throughout the game, the government and the firm share the same belief about efficacy because the firm is legally required to disclose all efficacy and safety information to the government (Fact **??**). FDA regulations require firms seeking conditional approval to submit all relevant human evidence on efficacy and safety, including controlled and uncontrolled studies, marketing experience, and published or unpublished research, regardless of source (U.S. Food and Drug Administration, 2024). In the model, confirmatory testing is the only source of new information about efficacy; sales and real-world use do not generate additional reliable evidence (Fact **??**).

By contrast, the cost of confirmatory testing is privately known to the firm (Fact **??**). This assumption reflects limited cost benchmarks for first-in-class drugs, firm-specific patient recruitment and manufacturing scale-up, and the difficulty of allocating shared development costs across products.

The primary risk of conditional approval is wasting public resources on an ineffective drug, rather than safety concerns. Toxicity is generally well understood after early trials and

is especially well characterized when already approved drugs are tested for new indications. Evidence on the trade-off between speed and safety suggests that the benefits of faster access tend to outweigh the safety risks (Jena et al., 2017).

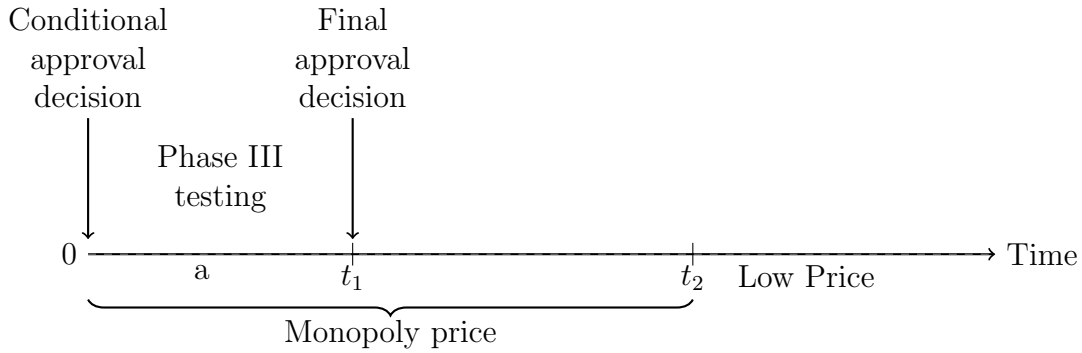


Figure 1: Timeline: at time 0 (after preliminary testing, not shown), the government decides whether to grant conditional approval and the firm decides whether to undertake Phase III testing; at time t_1, the government implements the final approval decision based on the test outcome. Following approval, the firm charges the monopoly price until time t_2, when generic entry or price regulation reduces prices.

Figure 2 illustrates key features of the optimal policy. First, annual doses typically grow gradually over time due to increasing doctor and patient familiarity (Figure 2, upper panel). Therefore, conditional approval provides additional benefits to both the firm and society by enabling the drug to reach patients earlier. Final approval is also more valuable when it follows conditional approval, because the drug already has broader use by then. Due to these dynamics, early access can be especially valuable. Second, for some drugs, cumulative profit will only be positive if the government grants conditional approval (Figure 2, lower panel). Without sufficient conditional approval and a high chance of final approval, expected profit will be negative, and the firm will not test.

2.2 Formal Model

A principal (government agency) and an agent (pharmaceutical firm) interact in a three-period game. At the beginning, the firm has already conducted preliminary clinical tests for a new drug. Based on the outcome of these tests, both parties believe that the drug's efficacy η is a random variable with support $[0, 1]$, cumulative distribution function G, and strictly positive density g. To ascertain the true efficacy of the drug $\eta \in [0, 1]$, the firm must conduct a final round of confirmatory tests which cost c. The firm knows the value of c privately, while

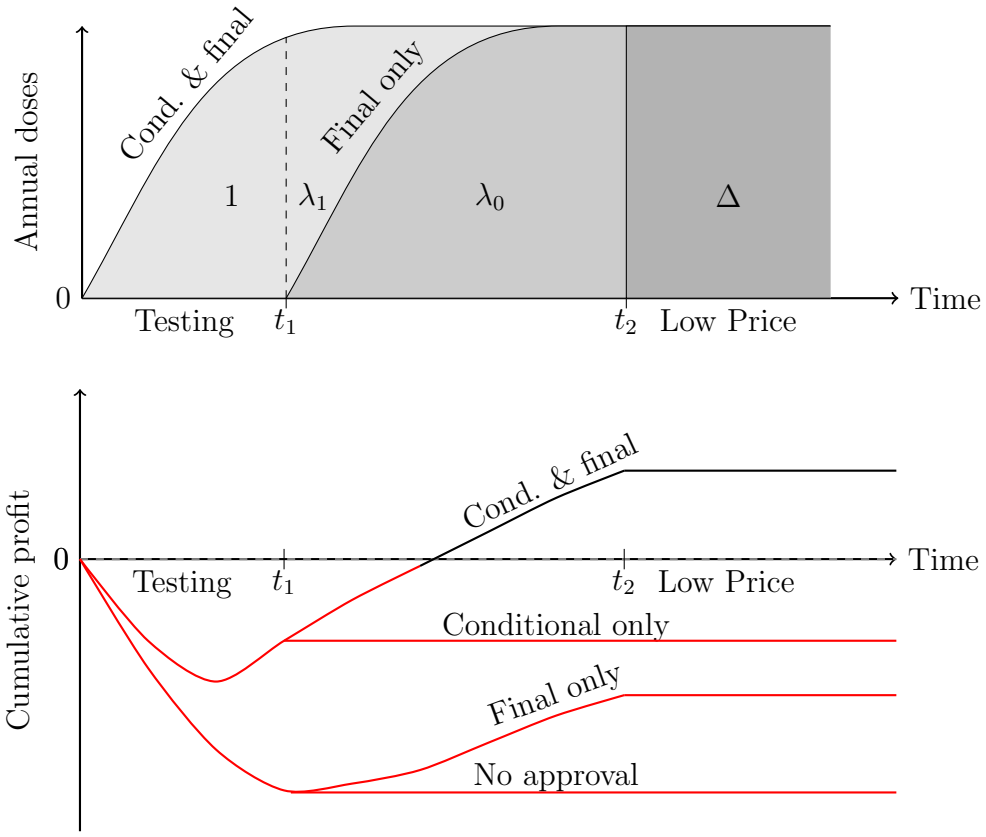


Figure 2: The upper panel shows annual doses rising earlier if conditional approval is granted. The lower panel illustrates a case in which the cumulative expected profit is positive only with conditional approval and a high probability of final approval.

the government believes that c is drawn from a distribution with a cumulative distribution function F and positive density f on its support $[c_L, c_H]$, with $0 \leq c_L \leq c_H$. We maintain the following regularity assumption throughout the paper.

Assumption 1. *The inverse hazard ratio $F(c)/f(c)$ is non-decreasing.*

Assumption 1 is routinely invoked in the mechanism design literature, following Myerson's pioneering work (Myerson, 1981). It implies that the "virtual cost" function $c + F(c)/f(c)$ is increasing, and holds for several commonly used distributions, including uniform, exponential, and Pareto distributions. In our setting, this assumption rules out stochastic regulatory policies.

Following standard results from mechanism design, we study the problem without loss of generality by focusing on direct mechanisms in which the government asks the firm to report its private information, here the testing cost c. The policy is designed so that the firm prefers to report truthfully. In practical implementation of our results, however, the government does not need to require reporting of the testing cost. Instead, we focus on a take-it-or-leave-it contract. Under this approach, the government specifies the scale of conditional approval and the efficacy threshold for final approval. The firm then decides whether to accept the offer by conducting testing or to decline it and exit. The firm accepts if and only if its testing cost is sufficiently low. We show that a take-it-or-leave-it contract is optimal in certain settings. Furthermore, we show that even when it is not exactly optimal, the best take-it-or-leave-it contract performs close to the theoretical optimum.

Formally, the game begins with the government committing publicly to a regulatory policy. Any feasible regulatory policy specifies, for each cost realization $c \in [c_L, c_H]$, a probability $\tau(c) \in [0, 1]$ of the firm conducting confirmatory testing, and a scale of conditional approval $a_1(c) \in [0, 1]$ contingent on the firm testing; and then for each pair $(c, \eta) \in [c_L, c_H] \times [0, 1]$, a probability of granting final approval $a_2(c, \eta) \in [0, 1]$.[5] If the firm chooses not to test, the game ends immediately with both parties earning zero payoffs. Otherwise, the firm enters the market in period one at scale a_1 while testing is taking place.

At time 0, the government decides whether to induce testing ($\tau(c)$) and grants conditional approval ($a_1(c)$) according to the agent's reported cost c. From time 0 to t_1, the firm sells a_1 fraction of the market at a given price. At time t_1, the test result (η) is revealed, and the government decides whether to grant final approval ($a_2(c, \eta)$). If the drug receives final approval, the firm continues selling at the monopoly profit margin until time t_2, after which

[5] Any regulatory policy can be interpreted as a direct revelation mechanism, as implied by the revelation principle (Myerson, 1981).

generic drugs are allowed to enter the market and the price falls to marginal cost. In the rest of the paper, we refer to time intervals $[0, t_1)$, $[t_1, t_2)$, and $[t_2, \infty)$ as periods one, two, and three, respectively.

Without loss of generality, we normalize both the firm's monopolistic profit per unit and the demand in period one, when full conditional approval is granted, to 1. If the firm is granted conditional approval at scale a_1, its payoff in period one is a_1.

The doses sold in period two, after the drug's final approval, depend on the doses sold in period one. As doctors and patients learn about the drug through experience and word of mouth, drug doses sold tend to rise over time. Hence, entering the market in period one at a larger scale increases later sales and profit in period two (Figure 2). We assume that the total discounted demand in period two, if final approval is granted, is $\lambda_0 + \lambda_1 \cdot a_1$, where λ_0 and λ_1 are fixed parameters. When the monopolistic period ends, generic drugs are allowed to enter the market, and the firm's net revenue becomes 0.

The firm's total expected profit is given by the testing probability multiplied by its total revenue across both periods minus the testing cost. Therefore, if the firm has a true cost c and chooses the option designed for cost c', its expected total profit is

$$\hat{\Pi}(c, c') \equiv \tau(c') \left[a_1(c') - c + [\lambda_0 + \lambda_1 \, a_1(c')] \int_0^1 a_2(c', \eta) \, \mathrm{d}G(\eta) \right]. \tag{1}$$

To be feasible, any regulatory policy must satisfy incentive compatibility (IC), meaning that the firm should not benefit from choosing any option other than the one designed for its true cost c, i.e.,

$$\forall c, c' \in [c_L, c_H]: \qquad \hat{\Pi}(c, c) \geq \hat{\Pi}(c, c'). \tag{IC}$$

Also, because the firm can earn zero profit by declining the request to conduct the final round of tests, the following individual rationality (IR) constraint must hold:

$$\forall c \in [c_L, c_H]: \qquad \hat{\Pi}(c, c) \geq 0. \tag{IR}$$

Finally, the probability of testing τ, the scale of conditional approval a_1, and the probability of final approval a_2, must be between 0 and 1:

$$\forall c \in [c_L, c_H]: \qquad 0 \leq \tau(c) \leq 1, \tag{2}$$

$$\forall c \in [c_L, c_H]: \qquad 0 \leq a_1(c) \leq 1, \text{ and} \tag{3}$$

$$\forall (c, \eta) \in [c_L, c_H] \times [0, 1]: \qquad 0 \leq a_2(c, \eta) \leq 1. \tag{4}$$

The government's payoff per unit of the drug $v(\eta)$ depends on the efficacy level. The function $v : [0,1] \to \mathbb{R}$ reflects the drug's value net of its monopoly price, which the firm charges in periods one and two. We assume that v is strictly increasing and satisfies

$$v(0) \leq -1 \text{ and } 0 < v(1), \tag{5}$$

where the price difference between the monopolistic and competitive periods equals the firm's net profit per unit, which is normalized to 1. That is, $v(0)$ equals zeros minus the drug's production cost, minus the firm's monopoly profit (which is normalized to 1), and minus the cost of toxicity.[6] Therefore, if the test outcome is sufficiently poor, i.e., η is close to 0, then the government's net payoff remains negative even after the firm loses its monopoly power.

If market access is granted in period one at scale a_1, the government's expected payoff is

$$v_E \int_{c_L}^{c_H} \tau(c)\ a_1(c)\ \mathrm{d}F(c), \tag{6}$$

where v_E is the expected value of $v(\eta)$ given by

$$v_E \equiv \int_0^1 v(\eta)\, \mathrm{d}G(\eta).$$

Given the firm's reported cost c and the revealed efficacy level η, the drug is finally approved with probability $a_2(c,\eta)$. Therefore, the government's expected payoff in period two is

$$\int_{c_L}^{c_H} \tau(c)\ [\lambda_0 + \lambda_1\, a_1(c)] \int_0^1 v(\eta)\, a_2(c,\eta)\, \mathrm{d}G(\eta)\, \mathrm{d}F(c). \tag{7}$$

At the end of period two, the firm loses its monopoly power due to generic entry or the initiation of government price controls. We assume that the price drops to the marginal cost. Therefore, the firm's profit vanishes. The government's unit payoff increases to $v(\eta) + 1$.[7] Letting Δ denote the total demand after period two, the government's payoff in period three is given by

$$\Delta \int_{c_L}^{c_H} \int_0^1 \tau(c)\ [v(\eta) + 1]\ a_2(c,\eta)\ \mathrm{d}G(\eta)\, \mathrm{d}F(c). \tag{8}$$

[6] Many conditionally approved drugs show toxicity in Phase I and II trials. For example, many conditionally approved drugs target cancer, such as chemotherapy, which is inherently toxic but potentially effective. The key questions are whether the drug is sufficiently efficacious relative to its toxicity, and whether this justifies spending on the drug.

[7] Recall that the profit margin before period three is normalized to 1.

The government maximizes its expected payoff, given by the sum of (6), (7), and (8):

$$U_R(\tau, a_1, a_2; \Delta) \equiv \int_{c_L}^{c_H} \tau(c) \left[v_E \, a_1(c) + \int_0^1 \Phi(a_1(c), \eta; \Delta) \; a_2(c, \eta) \, \mathrm{d}G(\eta) \right] \mathrm{d}F(c), \tag{9}$$

where the coefficient of a_2 in the inner integral is defined as

$$\Phi(a_1, \eta; \Delta) \equiv (\lambda_0 + \lambda_1 \, a_1) \; v(\eta) + \Delta \; [v(\eta) + 1] \, . \tag{10}$$

The monotonicity of $v(\eta)$ and assumption (5) imply the following "single-crossing" property of the function Φ.

Lemma 1. *For any $a_1 \in [0, 1]$ and $\Delta \geq 0$, the function $\Phi(a_1, \cdot; \Delta)$ is strictly increasing, and satisfies $\Phi(a_1, 0; \Delta) < 0 < \Phi(a_1, 1; \Delta)$.*

In light of Lemma 1, we can define, for each level of conditional approval $a_1 \in [0, 1]$, the "no-leniency" threshold for final approval as the unique solution to the equation $\Phi\big(a_1, \cdot\big) = 0$, i.e.,

$$\eta_{\mathrm{NL}}(a_1) \equiv v^{-1} \left(- \frac{\Delta}{\lambda_0 + \lambda_1 \, a_1 + \Delta} \right). \tag{11}$$

This threshold maximizes the government's continuation payoff after any amount of conditional approval granted in period one. A government with commitment power may choose a different threshold than $\eta_{\mathrm{NL}}(a_1)$ in order to maximize overall welfare while respecting the firm's incentive constraints.

If $\Delta = 0$, the no-leniency threshold $\eta_{\mathrm{NL}}(a_1)$ becomes

$$\bar{\eta} \equiv v^{-1}(0). \tag{12}$$

Define the government's problem as

$$\begin{aligned} U_R^*(\Delta) \equiv \max_{\tau, a_1, a_2} \; & U_R(\tau, a_1, a_2; \Delta), \\ & \text{subject to (IC), (IR), (2), (3), and (4).} \end{aligned} \tag{13}$$

The optimization problem in (13) is infinite-dimensional and non-convex, hence beyond the reach of standard solvers. Our approach consists in partitioning the parameter space into three regions, and using a different approach for each region.

In Section 3, we show that, if the drug is "low-cost," i.e., the government's belief about the firm's testing cost is concentrated on relatively low values, all (IC) and (IR) constraints

in (13) can be ignored. In this case, the highest cost is not too high to be excluded. The problem is solved by policies that provide no *costly* incentive to the firm, i.e., no leniency, and conditional approval only if the belief in the drug's efficacy is sufficiently optimistic.

In Section 4, we focus on the version of the model where the testing cost could be high, i.e., the government's belief is not as optimistic as the previous case, and there is no third period, i.e., $\Delta = 0$. We reformulate the problem in (13) as a linear optimization problem, and provide analytical expressions of optimal policies for all parameter values in this second region. The optimal solution can be implemented by a "take-it-or-leave-it" policy which specifies a conditional approval scale and a final approval threshold, in exchange for testing.

In Section 5, we let $\Delta > 0$ and provide a linear relaxation of the government's problem, which yields an upper bound for the government's expected payoff. Based on the solution to the upper-bound problem, we construct a take-it-or-leave-it policy that is feasible for the original problem (13). This policy provides a lower bound on the government's expected payoff. Under the calibrated model parameters, the lower bound is within 97.5% of the upper bound. Moreover, when $\Delta = 0$, the upper and lower bounds coincide with the optimal objective value.

Optimal conditional approval and leniency may depend on the length of exclusivity. Longer exclusivity raises firm profits but delays savings for the government. As a result, longer exclusivity can strengthen development incentives and reduce the need for conditional approval and leniency, all else equal. We report how optimal strategies vary with exclusivity length in Appendix 9.

In the next three sections, we characterize optimal regulatory policies for various configurations of parameter values.

3 Low Testing Cost

In this section, we focus on cases in which the government's belief about the firm's testing cost is relatively optimistic. For these "low testing cost" parameter configurations, the firm finds it profitable to participate without any leniency, even at the highest possible testing cost c_H.

Definition 1. *Let $\check{U}_R(a_1, \eta_m)$ denote the government's expected payoff conditional on testing,*

generated by conditional approval level a_1 and final approval threshold η_m,[8] *i.e.,*

$$\check{U}_R(a_1, \eta_m) \;\equiv\; v_E\, a_1 + \int_{\eta_m}^{1} \Big[(\lambda_0 + \lambda_1\, a_1)\, v(\eta) + \Delta\, [v(\eta) + 1]\Big]\, \mathrm{d}G(\eta), \quad \forall (a_1, \eta_m) \in [0,1]^2. \tag{14}$$

We say that a drug is

- *"promising", if*

$$\check{U}_R(0, \eta_{\mathrm{NL}}(0)) \;\leq\; \check{U}_R(1, \eta_{\mathrm{NL}}(1)); \tag{15}$$

- *"not promising", if*

$$\check{U}_R(0, \eta_{\mathrm{NL}}(0)) \;\geq\; \check{U}_R(1, \eta_{\mathrm{NL}}(1)); \tag{16}$$

- *"low-cost", if either* (15) *holds and*

$$c_H \leq 1 + (\lambda_0 + \lambda_1)\, [1 - G(\eta_{\mathrm{NL}}(1))]\,, \tag{17}$$

  or (16) *holds and*

$$c_H \leq \lambda_0\, [1 - G(\eta_{\mathrm{NL}}(0))]\,. \tag{18}$$

The expression in (14) is obtained from the expression within the square brackets in (9), after substituting $a_1(c) = a_1$ and $a_2(c, \eta) = \mathbb{1}_{\eta \geq \eta_m}$. It represents the government's expected payoff when the firm runs confirmatory testing, is granted conditional approval at scale a_1, and is promised final approval only if the revealed efficacy η exceeds the threshold η_m. The definition of "promising," as specified in (15), implies that the government achieves a higher expected payoff by granting full conditional approval ($a_1 = 1$) than not ($a_1 = 0$).

The next proposition establishes that it is optimal to provide no leniency if the drug is low-cost.

Proposition 1. *(1) If the drug is promising and low-cost, i.e.,* (15) *and* (17) *hold, it is optimal for the government to grant full conditional approval and no leniency, which*

[8] If we replace any function a_2 with the step function $a_2(c, \eta) = \mathbb{1}_{\eta \geq \eta_m(c)}$, where the threshold η_m is given by

$$1 - G(\eta_m(c)) = \int_0^1 a_2(c, \eta)\, \mathrm{d}G(\eta), \quad \forall c \in [c_L, c_H],$$

the firm's expected profit remains unchanged, and the government's expected payoff does not decrease, because the unit payoff function v is increasing. Therefore we can restrict attention without loss of generality to threshold final approval policies.

induces all cost types to test, i.e.,

$$\forall (c,\eta) \in [c_L, c_H] \times [0,1]: \qquad \tau(c) = 1, \quad a_1(c) = 1, \quad a_2(c,\eta) = \mathbb{1}_{\eta \geq \eta_{\text{NL}}(1)}. \tag{19}$$

(2) If the drug is not promising and low-cost, i.e., (16) *and* (18) *hold, on the other hand, it is optimal for the government to grant no conditional approval and no leniency, which induces all cost types to test, i.e.,*

$$\forall (c,\eta) \in [c_L, c_H] \times [0,1]: \qquad \tau(c) = 1, \quad a_1(c) = 0, \quad a_2(c,\eta) = \mathbb{1}_{\eta \geq \eta_{\text{NL}}(0)}. \tag{20}$$

When the drug is low-cost, it is optimal to provide no *costly* incentives to the firm: always no leniency, full conditional approval if the drug is promising, and no conditional approval if the drug is not promising. Therefore the optimal policies in (19) and (20) also solve both "first-best" problem, where the incentive constraints (IC) are ignored, and the "dictatorial" problem where both (IC) and (IR) are ignored. We record this result in the next corollary.

Corollary 1. *The "dictatorial" problem, defined as* (13) *without* (IC) *and* (IR)*, and the "first-best" problem, defined as* (13) *without* (IC)*, are solved by the policy in* (19) *under conditions* (15) *and* (17)*, and by the policy in* (20) *under conditions* (16) *and* (18)*.*

If the inequalities in (17) or (18) hold strictly, the optimal policies described in Proposition 1 leave the firm with a strictly positive expected profit, even when its production cost is c_H. This never happens in standard procurement problems, where the (IR) constraint for the worst type always binds.

4 High Testing Cost in the Two-Period Game ($\Delta = 0$)

Starting from this section, we allow the testing cost to violate the low-cost conditions of Definition 1. In other words, the government believes the firm's testing cost could be high relative to the drug's potential. Furthermore, we also assume prices never fall from generic entry or price controls. This implies $\Delta = 0$. In the next section, we will relax this assumption about prices falling.

Why assume prices do not fall? Generic competition may never emerge if the drug is displaced by newer technology before patent expiry. Moreover, rare-disease drugs, the majority of conditionally approved treatments (see fact **??**), often maintain high prices for

two reasons. First, markets are typically too small to support multiple firms. Second, Medicare price controls do not currently apply to rare-disease drugs.

4.1 Optimal Policies

When $\Delta = 0$, we can reformulate the government's problem (13) as a linear optimization problem, by introducing the following variables:

$$\forall c \in [c_L, c_H]: \quad \alpha_1(c) \equiv \tau(c)\, a_1(c), \text{ and} \tag{21}$$

$$\forall (c,\eta) \in [c_L, c_H] \times [0,1]: \quad \alpha_2(c,\eta) \equiv \tau(c)\ [\lambda_0 + \lambda_1\, a_1(c)]\ a_2(c,\eta). \tag{22}$$

The feasibility conditions (2)-(4) can be written as

$$\forall c \in [c_L, c_H]: \quad 0 \leq \alpha_1(c) \leq \tau(c), \tag{23}$$

$$\forall (c,\eta) \in [c_L, c_H] \times [0,1]: \quad 0 \leq \alpha_2(c,\eta) \leq \lambda_0\, \tau(c) + \lambda_1\, \alpha_1(c). \tag{24}$$

The firm's total expected profit function in (1) becomes

$$\hat{\Pi}(c, c') = \alpha_1(c') - c\, \tau(c') + \int_0^1 \alpha_2(c', \eta)\, \mathrm{d}G(\eta), \tag{25}$$

and the government's expected payoff in (9) becomes

$$U_R(\tau, \alpha_1, \alpha_2; \Delta) = \int_{c_L}^{c_H} \left[v_E\, \alpha_1(c) + \int_0^1 v(\eta)\, \alpha_2(c,\eta)\, \mathrm{d}G(\eta) \right] \mathrm{d}F(c). \tag{26}$$

The next proposition establishes that, when $\Delta = 0$, we can transform the original government's problem defined in (13) into an equivalent linear optimization formulation.

Proposition 2. *Recall $U_R^*(\Delta)$ defined in* (13)*. We have*

$$U_R^*(0) = \max_{\tau, \alpha_1, \alpha_2} \ U_R(\tau, \alpha_1, \alpha_2; \Delta), \tag{27}$$
$$\text{subject to} \ \ (\mathrm{IC}), (\mathrm{IR}), (2), (23), \ \text{and} \ (24).$$

Furthermore, given an optimal solution $\{\tau^*(c), \alpha_1^*(c), \alpha_2^*(c,\eta)\}$ to (27)*, define*

$$a_1^*(c) \equiv \frac{\alpha_1^*(c)}{\tau^*(c)} \cdot \mathbb{1}_{\tau^*(c)>0} \quad \text{and} \quad a_2^*(c,\eta) \equiv \frac{\alpha_2^*(c,\eta)}{\lambda_0\, \tau^*(c) + \lambda_1\, \alpha_1^*(c)} \cdot \mathbb{1}_{\tau^*(c)>0}. \tag{28}$$

Then, $\{\tau^*(c), a_1^*(c), a_2^*(c,\eta)\}$ *is an optimal solution to* (13).

The expressions for a_1^* and a_2^* in (28) follow directly from the change of variables defined in (21) and (22), given the optimal solution to (27). Proposition 2 provides a translation between the optimal incentive mechanism that solves (13) and an optimal solution to the linear optimization problem (27).

In the rest of this section, we focus on presenting the optimal solution to (27). A rigorous proof, based on linear programming and duality, is presented in Appendix 11.3. Here, we present an intuitive (albeit non-rigorous) line of argument to motivate the result.

We show that the optimal incentive mechanism when $\Delta = 0$ can be implemented as a take-it-or-leave-it offer. That is, the firm is granted a single conditional approval scale a_1 and is promised a single final approval threshold η, in exchange for testing, if and only if its cost is below a cutoff type c_0.

In the previous section, we examined the low-cost cases. Here, we turn to the cases where the testing cost could be high, defined as the complement of the "low-cost" condition in Definition 1, where the government may need to screen out a high cost firm. The incentive constraints and optimality conditions imply that a firm of cutoff type c_0 is indifferent between participating and not. That is, its expected revenue across both periods is equal to c_0, i.e.,

$$c_0 = a_1 + (\lambda_0 + \lambda_1\, a_1)\, [1 - G(\eta)]\,. \tag{29}$$

Therefore, a firm with testing cost below c_0 has no incentive to decline the offer. The government's problem can be rewritten as:

$$\begin{aligned} \max_{a_1,\eta,c_0}\ & \hat{U}(a_1,\eta,c_0) \equiv F(c_0)\ \left[v_E\, a_1\ +\ (\lambda_0 + \lambda_1\, a_1)\ \int_\eta^1 v(y)\,\mathrm{d}G(y)\right] \\ \text{subject to}\ \ & (29) \text{ and } (a_1,\eta,c_0) \in [0,1]\times[0,\bar{\eta}]\times[c_L,c_H], \end{aligned} \tag{30}$$

where $F(c_0)$ denotes the probability that a firm's testing cost is below the cut-off type c_0. Constraint (29) guarantees that such a firm is willing to participate. The term that multiplies $F(c_0)$ represents the government's expected payoff from committing to a conditional approval scale a_1 and final approval threshold η, conditional on the firm's participation. It is without loss of generality to restrict the search for the optimal final approval threshold from within the interval $[0,\bar{\eta}]$, following the definition (12) and the fact that v is increasing.

The optimization problem in (30) can be interpreted as a procurement problem in which the government wants to buy a new drug with the goal of maximizing total consumer sur-

plus, and has two instruments to incentivize the supplier: i) conditional approval a_1, which generates revenue in period one and stimulates demand after period one, and ii) "leniency" $\eta < \bar{\eta}$, which increases the firm's expected revenue after period one, by lowering the final approval threshold η below its no-leniency level $\bar{\eta}$, thus increasing the probability of final approval.

Now we consider first-order conditions (FOCs) to maximize $\hat{U}$. We first eliminate a_1 using (29), and substitute

$$\tilde{a}_1(\eta, c_0) \equiv \frac{c_0 - \lambda_0 \, [1 - G(\eta)]}{1 + \lambda_1 \, [1 - G(\eta)]}. \tag{31}$$

in place of a_1 in the government's expected payoff $\hat{U}$. The partial derivative with respect to η is

$$\frac{\partial \hat{U}(\tilde{a}_1(\eta, c_0), \eta, c_0)}{\partial \eta} = \frac{(\lambda_0 + \lambda_1 \, c_0) \; F(c_0) \, g(\eta)}{\left[1 + \lambda_1 \, [1 - G(\eta)]\right]^2} \cdot h(\eta), \tag{32}$$

in which

$$h(\eta) \;\equiv\; v_E - v(\eta) + \lambda_1 \int_{\eta}^{1} [v(y) - v(\eta)] \, \mathrm{d}G(y), \quad \forall \eta \in [0, \bar{\eta}]. \tag{33}$$

It is clear that the sign of h determines the sign of $\hat{U}$'s partial derivative with respect to η. The function h is decreasing and satisfies $h(0) > 0$. Therefore, for any given c_0, the objective function $\hat{U}$ admits a unique maximizer $\eta^*(c_0) \in (0, \bar{\eta}]$. If the drug is promising, i.e., $h(\bar{\eta}) \geq 0$, then the derivative is nonnegative over the entire interval, and the maximizer is attained at the upper bound, i.e., $\eta^*(c_0) = \bar{\eta}$. Conversely, if the drug is not promising, i.e., $h(\bar{\eta}) < 0$, then the maximizer $\eta^*(c_0)$ corresponds to the unique root of h in the interval $(0, \bar{\eta})$.

The function h plays a central role in determining the testing and final approval thresholds under the optimal regulatory policy. Before presenting the details, it is convenient to define the following two functions, from setting a_1 in the right-hand side of (29) to be 0 and 1, respectively:

$$\underline{c}(\eta) \equiv \lambda_0 \, [1 - G(\eta)] \,, \quad \forall \eta \in [0, 1], \tag{34}$$

and

$$\bar{c}(\eta) \equiv 1 + (\lambda_0 + \lambda_1) \, [1 - G(\eta)] \,, \quad \forall \eta \in [0, 1]. \tag{35}$$

The following lemma defines the critical testing threshold on cost c and the final approval threshold on η, to be used in the optimal mechanism.

Lemma 2. *The function h defined in* (33) *is strictly decreasing with $h(0) > 0$, leading to*

the following results:[9]

(1) If $h(\bar{\eta}) \leq 0$, there exists a unique $\check{\eta} \in [0, \bar{\eta}]$ such that $h(\check{\eta}) = 0$. Define

$$\check{c} \equiv \sup \left\{ c \in [c_L, c_H] : \frac{\mathrm{d}\hat{U}(\tilde{a}_1(\check{\eta}, c), \check{\eta}, c)}{\mathrm{d}c} \geq 0 \right\}. \tag{36}$$

(2) If $h(\bar{\eta}) \leq 0$ and $\check{c} < \underline{c}(\check{\eta})$, there exists a unique $\tilde{\eta} \in [\check{\eta}, \bar{\eta}]$ such that

$$\left. \frac{\mathrm{d}\hat{U}\big(0, \eta, \underline{c}(\eta)\big)}{\mathrm{d}\eta} \right|_{\eta = \tilde{\eta}} = 0, \tag{37}$$

and we have $h(\tilde{\eta}) \leq 0$.

(3) If $h(\bar{\eta}) > 0$ or $h(\bar{\eta}) \leq 0$ and $\check{c} > \bar{c}(\check{\eta})$, define

$$\underline{\eta} \equiv \inf \left\{ \eta \in [0, \check{\eta}] : \frac{\mathrm{d}\hat{U}(1, \eta, \bar{c}(\eta))}{\mathrm{d}\eta} \leq 0 \right\}. \tag{38}$$

We have $h(\underline{\eta}) \geq 0$.

Intuitively, the pair of testing threshold $\check{c}$ and final approval threshold $\check{\eta}$ maximizes the government's expected payoff $\hat{U}$ if we ignore feasibility constraints. The final approval thresholds $\tilde{\eta}$ and $\underline{\eta}$ correspond to the maximizers of $\hat{U}$ when the conditional approval scales are fixed at 0 and 1, respectively, and the cutoff types c_0 are set to $\underline{c}(\tilde{\eta})$ and $\bar{c}(\underline{\eta})$, respectively.

The next proposition characterizes the optimal incentive mechanism for all parameter values.

Proposition 3. *Define three values, a conditional approval scale α^*, a testing threshold c^*, and a final approval threshold η^*, that depend on model parameters in the following cases.*

(1) If $h(\bar{\eta}) \leq 0$ and $\underline{c}(\check{\eta}) \leq \check{c} \leq \bar{c}(\check{\eta})$, we have $c^* = \check{c}$ and $\eta^* = \check{\eta}$, and the optimal conditional approval scale below the threshold is $\alpha^* = \tilde{a}_1(\eta^*, c^*)$, where $\tilde{a}_1$ is defined by (31).

(2) If $h(\bar{\eta}) \leq 0$ and $\check{c} < \underline{c}(\check{\eta})$, the government grants no conditional approval, i.e., $\alpha^* = 0$, and the thresholds c^* and η^* are determined as follows:

(a) if $\underline{c}(\tilde{\eta}) < c_L$, we have $c^* = c_L$ and $\eta^* = 1$;

[9] We extend the definition of $F(c)/f(c)$ for $c \notin [c_L, c_H]$ as follows: $F(c)/f(c) = 0$ for $c < c_L$ and $F(c)/f(c) = 1/f(c_H)$ for $c > c_H$.

(b) if $c_L \leq \underline{c}(\tilde{\eta}) < c_H$*, we have* $c^* = \underline{c}(\tilde{\eta})$ *and* $\eta^* = \tilde{\eta}$*; and*

(c) if $\underline{c}(\bar{\eta}) < c_H \leq \underline{c}(\tilde{\eta})$*, we have* $c^* = c_H$ *and* $\eta^* = \sup\{\eta \in [0,1] : \underline{c}(\eta) \geq c_H\}$.

(3) If $h(\bar{\eta}) > 0$ *or* $h(\bar{\eta}) \leq 0$ *and* $\check{c} > \bar{c}(\check{\eta})$*, the government grants full conditional approval to firms whose types are below the threshold, i.e.,* $\alpha^* = 1$*. The thresholds* c^* *and* η^* *are determined as follows:*

(a) if $\bar{c}(\underline{\eta}) < c_L$*, we have* $c^* = c_L$ *and* $\eta^* = 1$*;*

(b) if $c_L \leq \bar{c}(\underline{\eta}) < c_H$*, we have* $c^* = \bar{c}(\underline{\eta})$ *and* $\eta^* = \underline{\eta}$*; and*

(c) if $\bar{c}(\bar{\eta}) < c_H \leq \bar{c}(\underline{\eta})$*, we have* $c^* = c_H$ *and* $\eta^* = \sup\{\eta \in [0,1] : \bar{c}(\eta) \geq c_H\}$.

The following solution is optimal to (27)*:*

$$\tau^*(c) = \mathbb{1}_{c\in[c_L,c^*)}, \quad \alpha_1^*(c) = \alpha^* \cdot \mathbb{1}_{c\in[c_L,c^*)}, \quad \alpha_2^*(c,\eta) = (\lambda_0 + \lambda_1\,\alpha^*) \cdot \mathbb{1}_{(c,\eta)\in[c_L,c^*)\times[\eta^*,1]}. \quad (39)$$

Cases (2) and (3) in Proposition 3 do not consider conditions $c_H \leq \underline{c}(\bar{\eta})$ or $c_H \leq \bar{c}(\bar{\eta})$, respectively, because these are special cases of Proposition 1 for $\Delta = 0$. It can be verified that when $\Delta = 0$, conditions (15)-(18) become $h(\bar{\eta}) \geq 0$, $h(\bar{\eta}) \leq 0$, $c_H \leq \bar{c}(\bar{\eta})$, and $c_H \leq \underline{c}(\bar{\eta})$, respectively.

Consider case (2) of Proposition 3. When $h(\bar{\eta}) \leq 0$ and $c_H \leq \underline{c}(\bar{\eta})$, we have

$$\check{c} \leq c_H \leq \underline{c}(\bar{\eta}) \leq \underline{c}(\check{\eta}).$$

Therefore, the condition $\check{c} < \underline{c}(\check{\eta})$ is already implied, and the optimal mechanism corresponds to case (2) in Proposition 1.

Next, consider case (3) of Proposition 3. When $c_H \leq \bar{c}(\bar{\eta})$, we have

$$\check{c} \leq c_H \leq \bar{c}(\bar{\eta}) \leq \bar{c}(\check{\eta}),$$

implying that the condition $\check{c} > \bar{c}(\check{\eta})$ is irrelevant. The optimal mechanism corresponds to case (1) in Proposition 1.

4.2 Calibration in the Two-Period Game

Next, we calibrate the model for the two-period game ($\Delta = 0$). We use the calibrated values of the model parameters based on the drug avelumab. We use these values in most of the subsequent figures in the paper. See Table 2 in Appendix 8. We use a uniform distribution of

private information c, the cdf $G(\eta) = \eta^k$ for various values of k to obtain the corresponding ex ante expected values v_E, and a linear government's utility function v. We plot c^*, α^*, and η^* against v_E.

Figure 3 shows how the optimal policy changes as the drug's expected effectiveness increases.

The optimal policy can be described according to a partition of the v_E space into five intervals. In the first interval ($v_E < v_1$), the drug's expected effectiveness is so low that it is not optimal to induce testing, even for the lowest possible cost ($c^* = c_L$), and thus the government grants no conditional approval ($\alpha^* = 0$). In the second interval, $v_E \in [v_1, v_2)$, the optimal policy grants no conditional approval but induces the firm to test if the testing cost is relatively low by providing some leniency, i.e., $\eta^* < \bar{\eta}$. In the third interval $[v_2, v_3)$, as the expected value v_E increases, the government progressively grants greater scales of conditional approval while becoming more stringent on the final approval threshold. Because the scale of conditional approval increases, the cutoff type c^* also increases. In the fourth interval $v_E \in [v_3, v_4)$, the government reduces the scale of conditional approval and continues raising the final approval threshold. The firm is now willing to test regardless of the cost ($c^* = c_H$). Here, the government lowers the conditional approval scale α^* because the higher likelihood of final approval already provides sufficient incentive for the firm to test. Finally, in the fifth interval $v_E \in [v_4, v(1)]$, the assumptions of Proposition 1 hold, hence, the optimal policy coincides with the "first-best" policy. The government grants full conditional approval and no leniency, and the firm tests regardless of the cost.

Note that the scale of conditional approval, α^*, exhibits a discontinuous jump at $v_E = v_4$. The intuition behind this behavior is as follows. When v_E approaches v_4 from the left, the final approval threshold converges to $\bar{\eta}$. Correspondingly, the coefficient of a_1 in the government's payoff function (9) is given by

$$v_E + \lambda_1 \int_{\bar{\eta}}^{1} v(\eta)\, \mathrm{d}G(\eta),$$

which crosses 0 at v_4 from below. Consequently, it becomes optimal to grant full conditional approval, i.e., $\alpha^* = 1$, when v_E increases beyond v_4.

The government's final approval threshold (η^*) is lower for drugs with lower expected value (v_E) under the parameters in our calibration exercise (Figure 3(c)). Under a higher expected value, the government can set a higher final approval threshold without discouraging the firm from testing. The particularly interesting observation in Figure 3(c) is that the

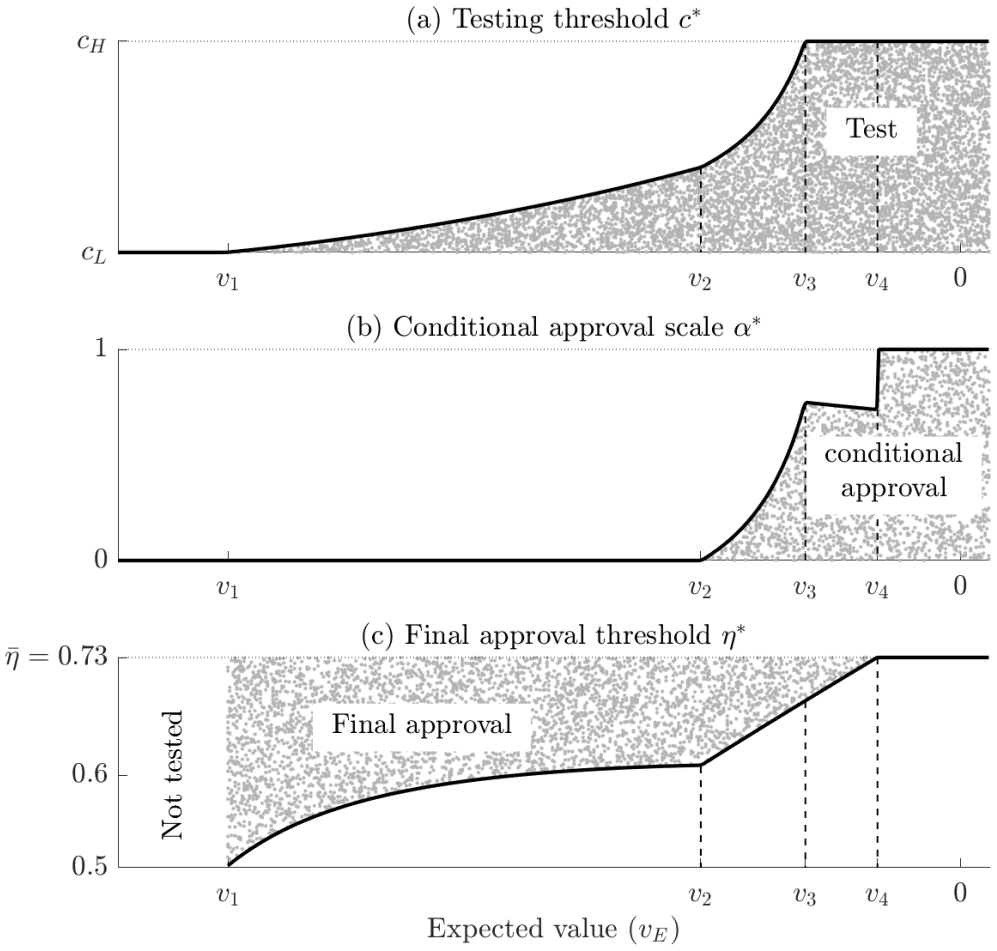


Figure 3: The points v_1–v_4 mark thresholds for the government's expected payoff. As the drug's expected value v_E increases, the testing-cost cutoff (top panel), the scale of conditional approval (middle panel), and the final-approval cutoff (bottom panel) generally rise. The exception is an intermediate range in which the drug is not attractive enough for immediate approval but sufficiently promising that the firm will test with little incentive, so the scale of conditional approval does not increase.

threshold η^* may be strictly less than the myopic threshold $\bar{\eta}$. Recall that if the efficacy is at $\bar{\eta}$, the drug's societal value $v(\bar{\eta}) = 0$. It may be counterintuitive that the government should grant final approval to a drug whose societal value is negative. The reason is that if the ex ante effectiveness v_E is fairly negative and the government still holds the final approval threshold at $\bar{\eta}$, the chance of success is too low for the firm to be willing to test. This may forfeit the possibility of developing a drug that could turn out to be efficacious. In order to ensure sufficient profit for the firm to invest in testing, the government can either use conditional approval or commit to a final approval threshold that is lower than the "no-leniency" threshold $\bar{\eta}$. That is, committing to the final approval of a drug with societal value $v(\eta) < 0$ can be interpreted as the society sharing part of the testing cost with the firm, in the hope that the drug may ultimately prove more efficacious.

In practice, missing the threshold might mean that the trials were not conclusive in demonstrating a benefit for the average patient, but the government asks for no more. The government can commit to a lower threshold during discussions with the firm before confirmatory testing. Indeed, the FDA grants what it calls "breakthrough status" to certain drugs, providing the developer with timely advice on clinical trial design (Fact **??**).

4.3 Benchmark: First-Best Optimal Policies

How do the optimal policies under private information compare to the first-best policies, where the government observes the firm's testing cost? The answer depends on whether the government believes testing costs are low or high. The previous section showed that the first-best and second-best solutions are the same under low-cost beliefs (see Corollary 1). We now examine the case where testing costs may be high. More precisely, these first-best policies solve the government's problem (13) with $\Delta = 0$, without the incentive compatibility constraints (IC). That is, we define the first-best optimization problem as

$$U_{FB} \equiv \max_{\tau, a_1, a_2} U_R(\tau, a_1, a_2; 0), \quad \text{subject to (IR), (2), (3), and (4).} \tag{40}$$

Because the firm's testing cost is potentially high, any first-best policy generates zero expected profit for the firm, thus solving the optimization problem (30) with $c_0 = c$. Among the multiple combinations of leniency and conditional approval that yield zero expected profit, it is optimal to use the combination that maximizes the government's expected payoff.

The next proposition provides a formal characterization of the first-best policy when

$\Delta = 0$.

Proposition 4. *Suppose that either* (15) *and* $c > 1+(\lambda_0+\lambda_1)\,[1-G(\eta_{\mathrm{NL}}(1))]$ *hold together, or* (16) *and* $c > \lambda_0\,[1-G(\eta_{\mathrm{NL}}(0))]$ *hold together. The combinations of conditional approval* a_1^* *and final approval threshold* η^**, as defined below, solve the optimization problem on the right-hand side of* (40)*, when restricting to policies that induce testing.*

(1) If the drug is promising, i.e., $h(\bar{\eta}) \geq 0$*, then:*

(a) if $\bar{c}(\bar{\eta}) < c \leq \bar{c}(0)$*, we have* $a_1^* = 1$ *and* $\eta^* = G^{-1}\left(1 - \dfrac{c-1}{\lambda_0+\lambda_1}\right)$*; and*

(b) if $c > \bar{c}(0)$*, the government cannot induce the firm to conduct testing.*

(2) If the drug is not promising, i.e., $h(\bar{\eta}) \leq 0$*, then:*

(a) if $\underline{c}(\bar{\eta}) < c < \underline{c}(\check{\eta})$*, we have* $a_1^* = 0$ *and* $\eta^* = G^{-1}\left(1 - \dfrac{c}{\lambda_0}\right)$*;*

(b) if $\underline{c}(\check{\eta}) \leq c \leq \bar{c}(\check{\eta})$*, we have* $a_1^* = \tilde{a}_1(\check{\eta}, c)$ *and* $\eta^* = \check{\eta}$*;*

(c) if $\bar{c}(\check{\eta}) < c \leq \bar{c}(0)$*, we have* $a_1^* = 1$ *and* $\eta^* = G^{-1}\left(1 - \dfrac{c-1}{\lambda_0+\lambda_1}\right)$*; and*

(d) if $c > \bar{c}(0)$*, the government cannot induce the firm to conduct testing.*

It is optimal to induce testing (i.e., $\tau = 1$*) if and only if the government's resulting expected payoff is nonnegative.*

When testing costs are publicly known, the government can tailor both the scale of conditional approval and the final approval threshold to each cost type. When testing costs are privately known, this type-contingent policy is infeasible. Information asymmetry, therefore, forces the government to pay information rents to firms with low testing costs. The promised revenue from conditional approval and final approval just covers the testing cost of the cutoff type, while firms with lower costs earn positive expected profits.

Figure 4 illustrates first-best optimal policies for four representative parameter configurations. At point (1), the first-best policy grants full conditional approval and no leniency. At point (2), the drug is less ex-ante efficacious (while the testing cost remains unchanged), so the first-best policy grants neither conditional approval nor leniency. These two points correspond to the two cases considered in Proposition 1 and Corollary 1. At point (3), the testing cost is higher (while efficacy remains the same as at point 2), so it is optimal to grant no conditional approval, and some leniency is required to induce testing. This corresponds to case (2)(a) in Proposition 4. At point (4), the even higher testing cost c requires a

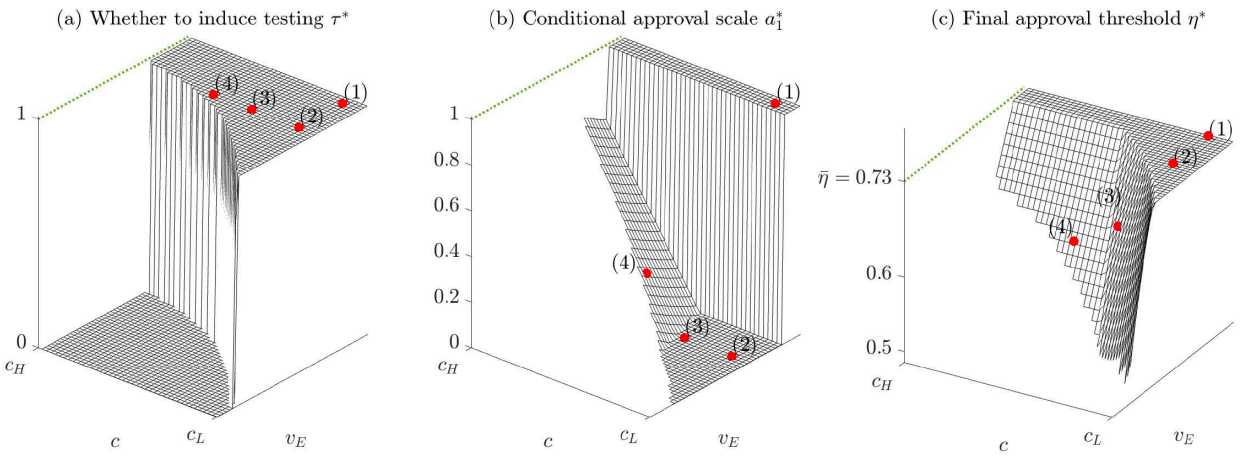


Figure 4: First-best policies for different values of the government's expected payoff v_E and the firm's testing cost c. The numbered points indicate four configurations in which the firm tests (left panel). Conditional approval is granted either because the drug is promising (point 1) or to induce testing despite higher costs (point 4, middle panel). The government applies leniency when testing costs are higher (points 3 and 4, right panel).

combination of leniency and a positive scale of conditional approval to induce testing. This corresponds to case (2)(b) in Proposition 4.

Next, we illustrate the value of information, i.e., the gap between the government's expected payoffs in the first-best and second-best scenarios (Figure 5). When the belief about efficacy is pessimistic (e.g., $v_E < -40$), the value of information is small. This indicates that learning the firm's true testing cost provides little benefit to the government, because only limited information rent is paid in this case. When the belief is optimistic (e.g., $v_E > -8$), the value of information is zero because the first-best and second-best policies coincide, as explained in Corollary 1. When the belief is intermediate, the value of information first increases and then decreases as the drug becomes more efficacious (e.g., v_E increases from -40 to -8), reflecting the government's willingness to pay to learn the true testing cost.

5 High Testing Cost in the Three-Period Game ($\Delta > 0$)

We now consider the case of potentially high testing costs and a third period in which the drug price falls due to generic competition or price controls. A lower price in the third period implies an even higher net benefit to the government. For $\Delta > 0$, the government's problem (13) cannot be formulated as an equivalent linear optimization as in Section 4.1. Therefore,

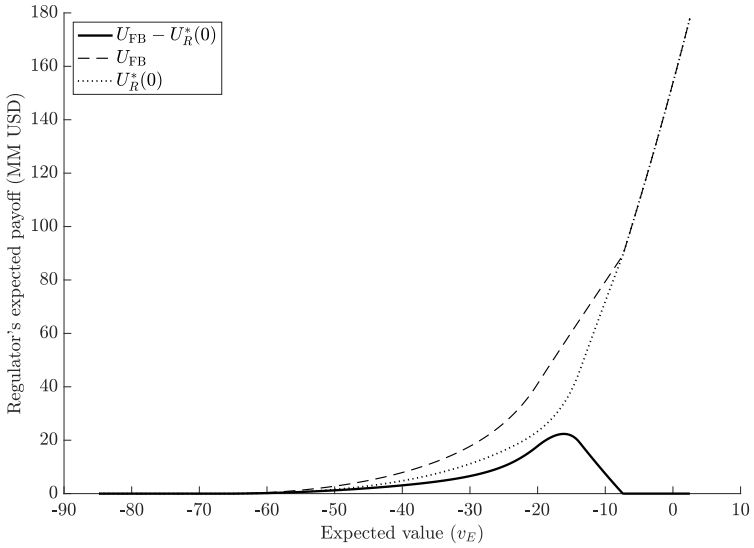


Figure 5: The government's expected payoff in the first-best (U_{FB}) and second-best ($U_R^*(0)$), and the value of information ($U_{FB} - U_R^*(0)$) as beliefs about efficacy η change.

we present a linear optimization relaxation, which not only provides an upper bound for $U_R^*(\Delta)$ but also allows us to construct a feasible mechanism.

5.1 Upper-Bound Linear Optimization

In addition to α_1 defined in (21), we define the following new decision variables for any $(c, \eta) \in [c_L, c_H] \times [0, 1]$:

$$_2(c, \eta) \equiv \tau(c)\, a_2(c, \eta), \text{ and} \tag{41}$$

$$_3(c, \eta) \equiv a_1(c)\, _2(c, \eta) = \alpha_1(c)\, a_2(c, \eta). \tag{42}$$

Using these notations, the government's payoff function (9) becomes

$$\hat{U}_R(\tau, \alpha_1, \hat{\alpha}_2, \hat{\alpha}_3; \Delta) \equiv \int_{c_L}^{c_H} \Bigg[v_E\, \alpha_1(c) + \int_0^1 \left[(\lambda_0 + \Delta)\, v(\eta) + \Delta \right] \hat{\alpha}_2(c, \eta)\, \mathrm{d}G(\eta) + \lambda_1 \int_0^1 v(\eta)\, \hat{\alpha}_3(c, \eta)\, \mathrm{d}G(\eta) \Bigg] \mathrm{d}F(c), \tag{43}$$

and the firm's payoff function (1) becomes

$$\hat{\Pi}(c, c') = \alpha_1(c') - c\,\tau(c') + \int_0^1 [\lambda_0\,{}_2(c', \eta) + \lambda_1\,{}_3(c', \eta)]\,\mathrm{d}G(\eta). \tag{44}$$

We further impose the following constraints on the decision variables ${}_2$ and ${}_3$: for any $(c, \eta) \in [c_L, c_H] \times [0, 1]$, we require

$$0 \leq {}_2\,(c, \eta) \leq \tau(c), \tag{45}$$

$$0 \leq {}_3\,(c, \eta) \leq {}_2\,(c, \eta), \text{ and} \tag{46}$$

$${}_3(c, \eta) \leq \alpha_1(c). \tag{47}$$

The government's upper-bound problem can be formulated as the following linear optimization problem:

$$\bar{U}_R^*(\Delta) \equiv \max_{\tau, \alpha_1, {}_2, {}_3} \hat{U}_R(\tau, \alpha_1, {}_2, {}_3; \Delta), \tag{48}$$
$$\text{subject to (IC), (IR), (2), (23), (45), (46), and (47),}$$

where $\hat{U}_R$ and $\hat{\Pi}$ follow (43) and (44), respectively.

The next proposition formalizes that the objective function value of the optimization problem (48) is an upper bound for the original problem (13).

Proposition 5. *$\bar{U}_R^*(\Delta) \geq U_R^*(\Delta)$ for any $\Delta \geq 0$.*

Under Assumption 1, the optimal solution to the government's upper-bound problem (48) is fully characterized by four constants: $\hat{c}$, $\hat{\alpha}$, $\hat{\eta}_1$, and $\hat{\eta}_2$. The following proposition summarizes this result.

Proposition 6. *Consider the optimization problem (48) under Assumption 1. There exists an optimal solution $\{\hat{\tau}(c), \hat{\alpha}_1(c), \hat{\alpha}_2(c, \eta), \hat{\alpha}_3(c, \eta)\}$ that can be summarized by four values: a cost threshold $\hat{c} \in [c_L, c_H]$, a scale of conditional approval $\hat{\alpha} \in [0, 1]$, and two efficacy thresholds $\hat{\eta}_1$ and $\hat{\eta}_2$ with $0 \leq \hat{\eta}_1 \leq \hat{\eta}_2 \leq \bar{\eta}$, as follows: for any $(c, \eta) \in [c_L, c_H] \times [0, 1]$,*

$$\hat{\tau}(c) = \mathbb{1}_{c < \hat{c}}, \quad \hat{\alpha}_1(c) = \hat{\alpha} \cdot \mathbb{1}_{c < \hat{c}}.$$

Furthermore, $\hat{\alpha}_2$ and $\hat{\alpha}_3$ demonstrate one of the following two cases:

Case 1:

$$\hat{\alpha}_2(c, \eta) = \mathbb{1}_{(c,\eta) \in [c_L, \hat{c}) \times [\hat{\eta}_1, 1]}, \quad \hat{\alpha}_3(c, \eta) = \hat{\alpha} \cdot \mathbb{1}_{(c,\eta) \in [c_L, \hat{c}) \times [\hat{\eta}_2, 1]}; \tag{49}$$

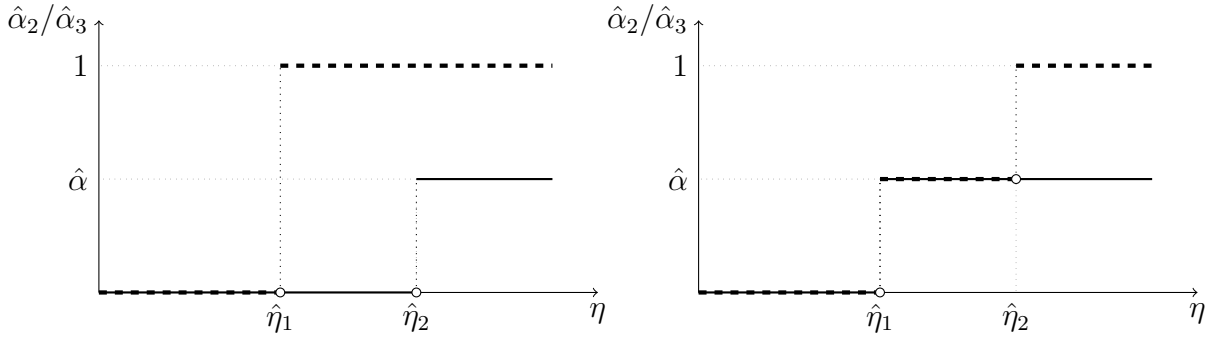


Figure 6: Illustration of $\hat{\alpha}_2$ and $\hat{\alpha}_3$ from Proposition 6. The left and right panels correspond to cases 1 and 2, respectively, with the dashed and solid lines representing $\hat{\alpha}_2$ and $\hat{\alpha}_3$, respectively, for a given $c \in [c_L, \hat{c})$.

Case 2:

$$\hat{\alpha}_2(c,\eta) = \begin{cases} 1, & \forall (c,\eta) \in [c_L, \hat{c}) \times [\hat{\eta}_2, 1], \\ \hat{\alpha}, & \forall (c,\eta) \in [c_L, \hat{c}) \times [\hat{\eta}_1, \hat{\eta}_2), \\ 0, & \textit{otherwise}, \end{cases} \qquad \hat{\alpha}_3(c,\eta) = \hat{\alpha} \cdot \mathbb{1}_{(c,\eta)\in[c_L,\hat{c})\times[\hat{\eta}_1,1]}. \quad (50)$$

The proof of the proposition, presented in Appendix 11.4, provides closed-form expressions of $\hat{c}$, $\hat{\alpha}$, $\hat{\eta}_1$, and $\hat{\eta}_2$, which vary depending on the model parameters. For simplicity of exposition, we leave the detailed expressions to Proposition 8 in the appendix.

It is worth mentioning that the solution described in Proposition 6 may not be feasible to the original problem (13). To explain this, we observe Figure 6, which illustrates the two cases of the optimal $\hat{\alpha}_2$ and $\hat{\alpha}_3$ for a given $c \in [c_L, \hat{c})$ as described in Proposition 6. As shown in the figure, the thresholds $\hat{\eta}_1$ and $\hat{\eta}_2$ may differ. As a result, we cannot identify a final approval policy based on $\hat{\alpha}_2$ and $\hat{\alpha}_3$ that achieves the optimal value of the objective function in (48).

As a special case, if the optimal $\hat{\alpha} = 0$, then $\hat{\eta}_2$ and $\hat{\eta}_1$ are irrelevant in Cases 1 and 2, respectively. This implies that the upper-bound optimal solution is feasible to the original problem (13). In particular, the government offers no conditional approval while including a firm with cost less than $\hat{c}$ to test. If Case 1 (resp. Case 2) of Proposition 6 holds, then the final approval decision follows the threshold $\hat{\eta}_1$ (resp. $\hat{\eta}_2$). This result is summarized in the following corollary.

Corollary 2. *If the optimal solution described in Proposition 6 satisfies $\hat{\alpha} = 0$, we have $\bar{U}_R^*(\Delta) = U_R^*(\Delta)$.*

If $\Delta = 0$, the two thresholds $\hat{\eta}_1$ and $\hat{\eta}_2$ always coincide,[10] ensuring that the optimal solution to the upper-bound problem remains feasible for the original one. Hence, the upper bound $\bar{U}_R^*$ is tight when $\Delta = 0$. This property is documented in the following corollary.

Corollary 3. *We have* $\bar{U}_R^*(0) = U_R^*(0)$.

5.2 Feasible Mechanism and Lower Bound

When the drug is low-cost (Definition 1), the policies described in Proposition 1 are optimal. For drugs with potentially high testing costs, the following proposition states that, given the optimal solution to the upper-bound problem (48), we can always construct a feasible solution to the original problem (13), with its objective function value serving as a lower bound.

Proposition 7. *Let* $\{\hat{\tau}(c), _1(c), _2(c,\eta), _3(c,\eta)\}$ *denote the optimal solution to the upper-bound problem* (48) *described in Proposition 6 (i.e.,* $\hat{\tau}(c) = \mathbb{1}_{c<\hat{c}}$ *and* $\hat{\alpha}_1(c) = \hat{\alpha} \cdot \mathbb{1}_{c<\hat{c}}$*). For any* $c \in [c_L, \hat{c})$*, define a threshold* $\eta_1 \in [0,1]$ *as the unique solution to*

$$\lambda_0 \int_0^1 \hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta) + \lambda_1 \int_0^1 \hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta) = (\lambda_0 + \lambda_1\,\hat{\alpha})\,[1 - G(\eta_1)]\,. \tag{51}$$

Based on these, define

$$a_1'(c) \equiv \hat{\alpha} \cdot \mathbb{1}_{c<\hat{c}} \quad \textit{and} \quad a_2'(c,\eta) \equiv \mathbb{1}_{(c,\eta)\in[c_L,\hat{c})\times[\eta_1,1]}; \tag{52}$$

and

$$\begin{aligned} &c_2 \equiv \hat{\alpha} + (\lambda_0 + \lambda_1\,\hat{\alpha})\,[1 - G(\eta_{\mathrm{NL}}(\hat{\alpha}))]\,, \quad \tau''(c) \equiv \mathbb{1}_{c<c_2}, \\ &a_1''(c) \equiv \hat{\alpha} \cdot \mathbb{1}_{c<c_2}, \quad \textit{and} \quad a_2''(c,\eta) \equiv \mathbb{1}_{(c,\eta)\in[c_L,c_2)\times[\eta_{\mathrm{NL}}(\hat{\alpha}),1]}. \end{aligned} \tag{53}$$

Both $\{\hat{\tau}(c), a_1'(c), a_2'(c,\eta)\}$ *and* $\{\tau''(c), a_1''(c), a_2''(c,\eta)\}$ *are feasible solutions to* (13)*. Therefore, we have*

$$\tilde{U}_R^*(\Delta) \equiv \max\{U_R(\hat{\tau}, a_1', a_2'; \Delta),\ U_R(\tau'', a_1'', a_2''; \Delta)\} \le U_R^*(\Delta). \tag{54}$$

Finally, if $\eta_1 > \eta_{\mathrm{NL}}(\hat{\alpha})$*, we have* $\tilde{U}_R^*(\Delta) = U_R(\tau'', a_1'', a_2''; \Delta)$.

As in Section 4, the feasible solutions to the optimization problem (13) constructed in Proposition 7 can be implemented as follows. The government chooses between two take-it-

[10] See Lemmas 5, 6, and 8 for details.

or-leave-it offers. Expression (52) (along with the function $\hat{\tau}$) indicates that the drug is tested if its cost c is below the threshold $\hat{c}$, and receives a conditional approval at scale $\hat{\alpha}$ and final approval if the efficacy result η is above η_1. Similarly, (53) specifies the testing threshold, conditional approval scale, and final approval threshold as c_2, $\hat{\alpha}$, and $\eta_{\text{NL}}(\hat{\alpha})$, respectively. Expression (54) then implies that the government chooses the better one between the two aforementioned mechanisms, which provides a lower bound on the optimal value $U_R^*(\Delta)$.

In the first construction (52), we take both $\hat{c}$ and $\hat{\alpha}$ directly from the upper-bound optimal solution to construct a feasible mechanism. In the second construction (53), by contrast, we use only $\hat{\alpha}$ as the conditional approval scale, and set the final approval threshold to $\eta_{\text{NL}}(\hat{\alpha})$.

It is instructive to explain how we obtain the other thresholds accordingly. The final approval threshold η_1 is determined by (51), which guarantees that the firm receives the same expected payoff as in the upper bound optimal solution (and hence all incentive constraints (IC) and (IR) are automatically satisfied). In particular, the left- and right-hand sides of (51) represent the firm's expected payoff in period two under the optimal solution to the upper-bound problem (48) and our candidate solution, respectively.

In the second case, we set the final approval threshold to $\eta_{\text{NL}}(\hat{\alpha})$ because it is never optimal to choose a higher value. Given this threshold, we then determine the testing threshold c_2 from (53) such that the (IR) constraint for the threshold type binds. This choice of c_2 also ensures that the incentive compatibility constraint (IC) is satisfied.

Finally, when $\eta_{\text{NL}}(\hat{\alpha}) < \eta_1$, the coefficient of a_2 remains nonnegative for all $\eta \in [\eta_{\text{NL}}(\hat{\alpha}), \eta_1]$, implying that lowering the final approval threshold from η_1 to $\eta_{\text{NL}}(\hat{\alpha})$ increases the government's expected payoff from any participating firm. Moreover, since $c_2 \geq \hat{c}$, the second policy induces greater participation. Therefore, it strictly outperforms the first policy. This explains the last sentence of Proposition 7.

5.3 Calibration in the Three-Period Game

We again use the calibrated model parameters following Appendix 8, and obtain the optimal upper-bound solution according to Proposition 8 for various G, similar to Section 4.2. Based on these, we construct feasible solutions to the original problem following Proposition 7.

Figure 13(a) illustrates the performance ratio (PR) of the candidate solution compared with the upper bound, i.e.,

$$\text{PR} = \frac{\tilde{U}_R^*}{\bar{U}_R^*}.$$

Under the given model parameters, the candidate solution consistently achieves a perfor-

mance ratio exceeding 0.975, implying that it is very close to optimal. In particular, the performance ratio equals 1 when the beliefs about efficacy are either very pessimistic ($v_E < v_2$) or sufficiently optimistic ($v_E > v_5$). In the pessimistic case ($v_E < v_2$), it is optimal to grant no conditional approval, and the upper and lower bound objective values coincide. This corresponds to Corollary 2. In the optimistic case ($v_E > v_5$), the "no-leniency" policy described in Proposition 1 is optimal, yielding the exact optimal solution to the original problem.

Our candidate mechanism shares the same characteristics as the optimal one when $\Delta = 0$: the government induces testing if the cost c is low (below the threshold $\hat{c}$) and grants final approval if the efficacy η is high (above the threshold $\hat{\eta}$). Against the corresponding expected values v_E, we plot the testing threshold $\hat{c}$ (Figure 13(b)), the scale of conditional approval $\hat{\alpha}$ (Figure 13(c)), and the threshold of final approval $\hat{\eta}$ (Figure 13(d)). Compared to Figure 3, the testing threshold $\hat{c}$ is weakly higher due to the government's positive payoff after period two, as shown in panel (b).

The discontinuous jump of $\hat{\alpha}$ in panel (c) at $v_E = v_4$ arises for the same reason as discussed in Section 4.2. When $v_E > v_4$, as the drug appears more promising, the coefficient of a_1 in the government's payoff function (9) remains nonnegative, implying that it is optimal to grant full conditional approval, i.e., $\hat{\alpha} = 1$.

In panel (d), we use the notation $\bar{\eta}_2$ to represent the highest final approval threshold. Its value coincides with η_{NL} (defined in (11)) when $a_1 = 1$, i.e., $\bar{\eta}_2 \equiv \eta_{\text{NL}}(1)$. Given $\hat{\alpha} = 1$, the coefficient of a_2 in the integrand of the objective function is nonnegative for all $\eta \in [\bar{\eta}_2, 1]$. Therefore, it is never optimal to set the final approval threshold above this value.

We compare policy performance with and without the tools of conditional approval and leniency (Figure 7). First, consider the scenario where the government uses conditional approval but not leniency. Here, it fixes a naive final approval threshold, $\bar{\eta}'$, defined as the no-leniency threshold when a_1=0:

$$\bar{\eta}' \equiv \eta_{\text{NL}}(0) = \inf\left\{\eta \in [0,1] : (\lambda_0 + \Delta)\, v(\eta) + \Delta \geq 0\right\}. \tag{55}$$

When the government uses only a leniency threshold and not conditional approval, we have $a_1(c) = 0$ for all $c \in [c_L, c_H]$. With neither tool, the final approval threshold is $\bar{\eta}'$, and only firms with low testing costs proceed with trials. For completeness, we present the optimal policies in closed form in Appendix 11.5.

In Figure 7(a), the solid line matches the curve in Figure 13(a). We observe that when either tool is unavailable, performance drops significantly under certain model parameters, highlighting the importance of both tools. It is worth noting that when the expected value

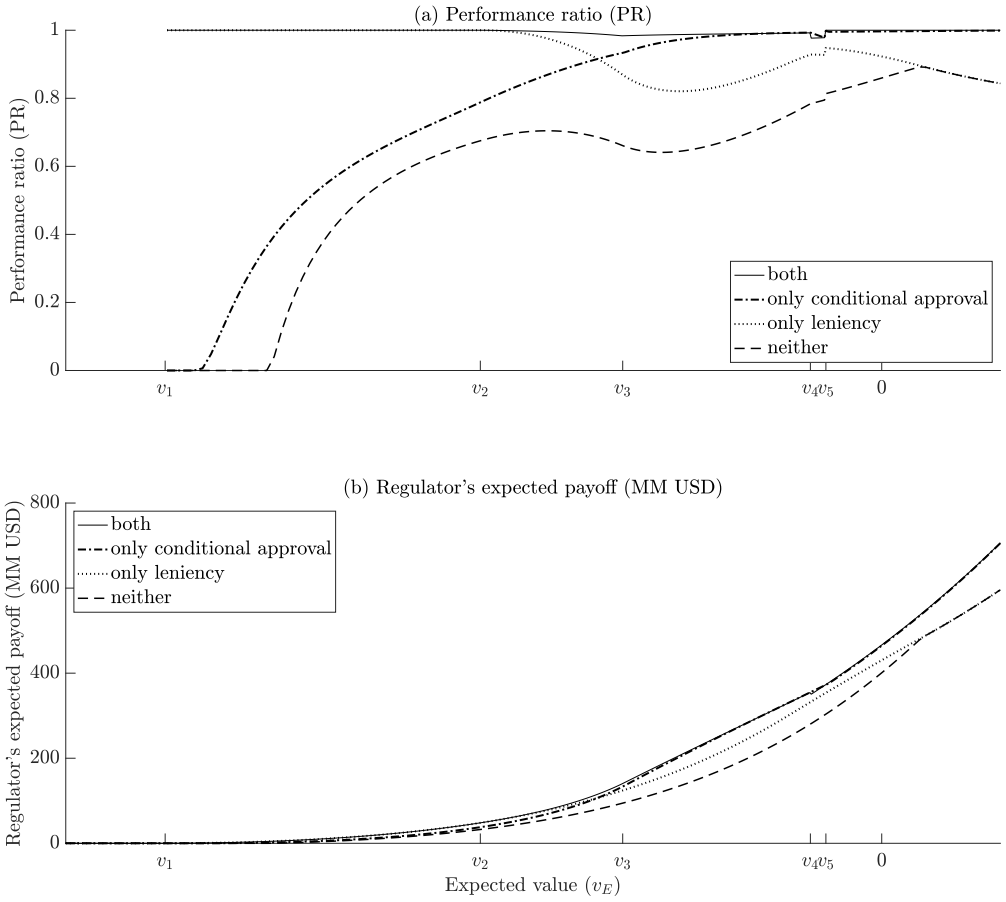


Figure 7: Performance ratios and the government's expected payoff (in million USD) of our candidate policy, the conditional approval-only policy, the leniency-only policy, and the policy without both tools.

v_E is close to v_4, the policy without leniency performs best among the four. This is because our candidate policy, which utilizes both tools, is constructed from an infeasible solution and is not necessarily optimal, making it possible for another policy to outperform it.

We plot the government's expected payoffs (in millions of USD) for the four scenarios in Figure 7(b). For instance, when $G(\eta) = \eta^{3/2}$, we have $v_E = -15.2$, which falls between v_3 and v_4. At this point, the government's payoffs are \$187, \$231, \$270, and \$270 million under the policies with neither tool, only leniency, only conditional approval, and both tools available, respectively. That is, considering both conditional approval and leniency yields a value as high as \$83 million.

Hence, conditional approval and leniency can generate more than \$6 billion in social value during a decade, compared to using neither, assuming a value of about \$83 million per drug and 79 drugs per decade. Additional value goes to shareholders, but we do not estimate it here. Nearly all of the social value comes from conditional approval rather than leniency, although leniency is relatively important for drugs with lower expected value (Figure 7(b)).

6 Conclusion

In 1992, the FDA granted its first conditional approval to an HIV-AIDS drug. The original aim was to accelerate patient access. We show that conditional approval offers a second benefit: encouraging investment in drugs that might otherwise be undeveloped.

We estimate that optimal use of conditional approval and leniency, compared to using neither, can generate hundreds of millions of dollars in social value each year, in addition to commercial value. Nearly all of the social value comes from conditional approval, rather than leniency, although leniency is relatively important for drugs with lower expected value.

Our analysis demonstrates the high value of conditional approval, even under our conservative framework that omits two additional benefits. First, conditional approval can accelerate price reductions if the end of pricing power depends on the approval date.[11] Second, conditional approval creates opportunities to learn from real-world use. Observational data may reduce the required sample size in clinical trials and lower testing costs. In this way, conditional approval can raise profits not only through earlier sales (as modeled), but also through reduced testing costs. This benefit, however, depends on the government's willingness to use non-randomized evidence, which is subject to selection bias.[12]

[11] Earlier approval can lead to earlier price declines if the patent expires before exclusivity ends, or if the drug is subject to Medicare price regulation. See Appendix 9 for details on the duration of market power.

[12] For example, if only high-income patients use the drug, their outcomes may reflect socioeconomic ad-

Our results yield three policy recommendations for lawmakers regarding regulatory authority. First, government agencies should be allowed to conditionally approve drugs with low expected efficacy, accepting that some will later be withdrawn. The FDA's actual withdrawal rate was 22% for drugs conditionally approved between 2013 and 2017 (Figure 8). Our back-of-the-envelope calibration implies that this is consistent with the optimal withdrawal rate for drugs with zero expected value (Appendix 8.5). This indicates that, in the past, the agency approved a mix of drugs with positive and negative expected values. However, parts of the agency have become more conservative, as the head of gene therapy appointed in 2025 has called for stricter conditional approval standards and less leniency (The Economist, 2025).

Second, government agencies should have the flexibility to grant partial conditional approval, such as restricting access to specific patient subgroups. Limiting insurance coverage could achieve a similar effect. Partial access should not be motivated solely by evidentiary uncertainty, though, but by whether limited access suffices to motivate continued development.

Third, government agencies should be able to commit to a final approval threshold below the myopic threshold to encourage testing. The FDA appears to follow this practice, though the policy is not formally codified.

References

Armstrong A (2024) Preventing "3 billion visits" to FDA a year with CBER's Peter Marks. *Fierce Biotech* URL `https://www.fiercebiotech.com/biotech/preventing-3-billion-visits-fda-year-cbers-peter-marks`, accessed 2024-07-02.

Baron DP, Myerson RB (1982) Regulating a monopolist with unknown costs. *Econometrica* 911–930.

Berndt ER, Glennerster R, Kremer MR, Lee J, Levine R, Weizsäcker G, Williams H (2007) Advance market commitments for vaccines against neglected diseases: Estimating costs and effectiveness. *Health Economics* 16(5):491–511.

Bognar K, Romley JA, Bae JP, Murray J, Chou JW, Lakdawalla DN (2017) The role of imperfect surrogate endpoint information in drug approval and reimbursement decisions. *Journal of Health Economics* 51:1–12.

vantages rather than drug efficacy.

Budish E, Roin BN, Williams H (2015) Do firms underinvest in long-term research? Evidence from cancer clinical trials. *American Economic Review* 105(7):2044–2085.

Carpenter D, Ting MM (2007) Regulatory errors with endogenous agendas. *American Journal of Political Science* 51(4):835–852.

Centers for Medicare & Medicaid Services (2022) CMS finalizes Medicare coverage policy for monoclonal antibodies directed against amyloid for the treatment of Alzheimer's disease. URL `https://www.cms.gov/newsroom/press-releases/cms-finalizes-medicare-coverage-policy-monoclonal-antibodies-directed-against-amyloid-treatment`.

Centers for Medicare & Medicaid Services (2024) Medicare & Medicaid spending by drug. URL `https://data.cms.gov/summary-statistics-on-use-and-payments/medicare-medicaid-spending-by-drug`.

Chandra A, Kao J, Miller KL, Stern AD (2024) Regulatory incentives for innovation: The FDA's breakthrough therapy designation. *The Review of Economics and Statistics* 1–46.

Collard-Wexler A, Grennan M, Steck A (2024) Regulating new product testing: The FDA vs. the invisible hand. *Working Paper* 1–32.

DiMasi JA, Grabowski HG, Hansen RW (2016) Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics* 47:20–33.

Dubois P, Moisson PH, Tirole J (2022) The economics of transferable patent extensions. *Toulouse School of Economics Working Paper* 1377:1–55.

Food and Drug Administration (2025) Advancing health through innovation: New drug therapy approvals 2024. URL `https://www.fda.gov/drugs/novel-drug-approvals-fda/novel-drug-approvals-2024`.

Frank RG, Shahzad M, Emanuel EJ (2022) Accelerated approval of cancer drugs: No economic reward for drug makers that conduct confirmatory trials. *Health Affairs* 41(9):1273–1280.

Gans JS, Ridley DB (2013) Innovation incentives under transferable fast-track regulatory review. *Journal of Industrial Economics* 61(3):789–816.

Grennan M, Town RJ (2020) Regulating innovation with uncertain quality: Information, risk, and access in medical devices. *American Economic Review* 110(1):120–161.

Henry E, Loseto M, Ottaviani M (2022) Regulation with experimentation: Ex ante approval, ex post withdrawal, and liability. *Management Science* 68(7):5330–5347.

Henry E, Ottaviani M (2019) Research and the approval process: The organization of persuasion. *American Economic Review* 109(3):911–955.

Jena AB, Zhang J, Lakdawalla DN (2017) The trade-off between speed and safety in drug approvals. *JAMA Oncology* 3(11):1465–1466.

Kaplan S (2016) 'Tough as nails': Storm swirls around FDA drug cop who approved controversial drug. `https://www.statnews.com/2016/09/20/janet-woodcock-sarepta-fda/`, accessed: 2025-12-16.

Kong E, Zhao O (2024) Market exclusivity and innovation: Evidence from antibiotics. `https://papers.ssrn.com/abstract=4676384`, SSRN Working Paper.

Kremer M, Levin J, Snyder CM (2022) Designing advance market commitments for new vaccines. *Management Science* 68(7):4786–4814.

Light DW, Warburton RN (2005) Extraordinary claims require extraordinary evidence. *Journal of Health Economics* 24(5):1030–1033.

Liu ITT, Kesselheim AS, Cliff ERS (2024) Clinical benefit and regulatory outcomes of cancer drugs receiving accelerated approval. *Journal of the American Medical Association* 331(17):1471–1479.

McClellan A (2022) Experimentation and approval mechanisms. *Econometrica* 90(5):2215–2247.

Myerson RB (1981) Optimal auction design. *Mathematics of Operations Research* 6(1):58–73.

Orlov D, Skrzypacz A, Zryumov P (2020) Persuading the principal to wait. *Journal of Political Economy* 128(7):2542–2578.

Peltzman S (1973) An evaluation of consumer protection legislation: The 1962 drug amendments. *Journal of Political Economy* 81(5):1049–1091.

Philipson TJ, Sun E (2008) Is the Food and Drug Administration safe and effective? *Journal of Economic Perspectives* 22(1):85–102.

Ridley DB, Lee CY (2020) Does Medicare reimbursement drive up drug launch prices? *The Review of Economics and Statistics* 102(5):980–993.

Ridley DB, Sun P, Xu C (2025) Optimal push, pull, and failure funding for global health, available at SSRN: https://ssrn.com/abstract=5266214.

Robey S, David FS (2016) Drug launch curves in the modern era. *Nature Reviews Drug Discovery* 16(1):13–14.

Snyder CM, Hoyt K, Gouglas D (2023) An optimal mechanism to fund the development of vaccines against emerging epidemics. *Journal of Health Economics* 91:102795.

The Economist (2025) America's drug regulator is in turmoil. https://www.economist.com/united-states/2025/08/10/americas-drug-regulator-is-in-turmoil, accessed August 11, 2025.

United Press International (1988) Police arrest aids protesters blocking access to fda offices. *Los Angeles Times* URL https://www.latimes.com/archives/la-xpm-1988-10-11-mn-3909-story.html.

US Congress (2023) Consolidated appropriations act, 2023. URL https://www.congress.gov/117/bills/hr2617/BILLS-117hr2617enr.pdf, pub. L. No. 117-328, div. FF, title III, §3210(d)(D), 136 Stat. 4459.

US Food and Drug Administration (2024) 21 cfr § 314.50 — content and format of an nda. https://www.ecfr.gov/current/title-21/chapter-I/subchapter-D/part-314/subpart-B/section-314.50, electronic Code of Federal Regulations; accessed: 2025-02-16.

US Food and Drug Administration (2025) FAQs: Commissioner's national priority voucher pilot program. https://www.fda.gov/news-events/press-announcements/faqs-commissioners-national-priority-voucher-pilot-program, accessed: 2025-12-14.

Wouters OJ, Kesselheim AS, Kuha J, Luyten J (2024) Sales revenues for new therapeutic agents approved by the United States Food and Drug Administration from 1995 to 2014. *Value in Health* 27(10):1373–1381.

Wouters OJ, McKee M, Luyten J (2020) Estimated research and development investment needed to bring a new medicine to market, 2009-2018. *Journal of the American Medical Association* 323(9):844–853.

Xu L, Zhao H, Petruzzi NC (2021) Inducing compliance with postmarket studies for drugs under FDA's accelerated approval pathway. *Manufacturing & Service Operations Management* 23(1):170–190.

Zhou EW, Jackson MJ, Ledley FD (2023) Spending on phased clinical development of approved drugs by the US National Institutes of Health compared with industry. *JAMA Health Forum* 4(7):e231921–e231921.

Online Appendix

This online appendix includes additional figures, facts, model calibration, and proofs.

7 Institutional background

Details about drug regulation and competition provide a foundation for the model.

fact:discreteA discrete-time model is appropriate in our setting because human clinical trials typically involve pre-specified endpoints and results that are announced all at once. Typically, Phase III trials have pre-specified endpoints, followed by a comprehensive analysis by the firm before regulatory submission. However, there are exceptions, such as early-stage animal testing and the rolling review of the COVID-19 vaccine, which have a more flexible and exploratory set of protocols.

fact:requireThe government can require initiation of confirmatory testing as a condition for granting conditional approval. Beginning in fiscal year 2023, the U.S. Congress granted the FDA the authority to mandate that trials begin before a drug receives conditional approval (US Congress, 2023). Previously, some firms selling conditionally-approved drugs cited difficulty enrolling enough patients in trials for rare diseases. This rationale was often a convenient excuse, as confirmatory testing could reveal that the drug should be withdrawn (Xu et al., 2021; Frank et al., 2022). Among 46 cancer drugs granted conditional approval between 2013 and 2017, 15% had not completed testing after a median of six years (Liu et al., 2024) (Figure 8). However, we expect that under the new policy, the share of drugs failing to do timely testing will drop.

fact:shareThe government can grant conditional approval for only a share of patients. For example, approval may be limited to those with the most severe conditions, or coverage may be restricted so that only some patients are reimbursed while others pay out-of-pocket. In the United States, the FDA determines approval scope, while Medicare sets coverage rules, both under the Secretary of Health and Human Services. After conditional approval of an Alzheimer's drug, for instance, Medicare covered it only for patients who agreed to

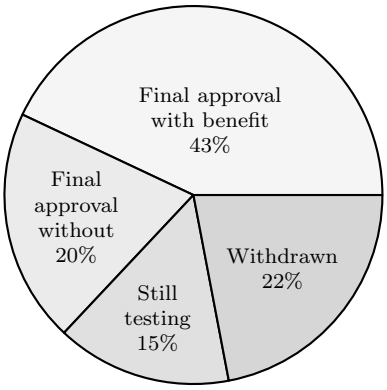


Figure 8: For 46 cancer indications conditionally approved in the U.S. between 2013 and 2017, 43% received final approval after demonstrating statistically significant benefits in confirmatory trials, 20% received final approval despite Liu et al. (2024) finding no evidence of statistically significant benefits, 22% were withdrawn, and 15% remained in ongoing testing. Source: Authors' figure based on data from Liu et al. (2024).

.

participate in a study of its efficacy (Centers for Medicare & Medicaid Services, 2022). For tractability, we assume conditional approval applies to a share of uniform patients.

fact:rareDrugs given conditional approval are typically for rare diseases and do not have competition. From 2001 to 2025, 75% of FDA conditional approvals were for rare diseases (see Figure 9). At launch, drugs for rare diseases typically have no competition, due to orphan drug exclusivity of seven years in the U.S. and ten years in Europe. Even years later, markets for rare diseases are often too small to attract multiple firms.

fact:liquidityConditional approval provides earlier revenue for some firms that would otherwise not continue testing. According to the head of the FDA center responsible for gene therapies and vaccines, "The wherewithal to do a three-year study or a four-year study without having a revenue stream, it's just beyond many companies that are startups. So having the accelerated approval process is a way to get there" (Armstrong, 2024).

fact:financingThe government does not typically directly finance late-stage testing. About 3% of NIH's total research budget was devoted to financing the clinical testing of drugs ultimately approved between 2010 and 2019 (Zhou et al., 2023). Also, NIH funding represented only about 3% of Phase III testing costs for a given approved drug (Zhou et al., 2023). Rather than supporting late-stage trials for firms, NIH funding is typically directed toward lower-cost, pre-clinical research conducted at universities. A notable exception was the financing of late-stage testing for COVID-19 vaccines through Operation Warp Speed (Snyder et al., 2023). Another exception is funding for drugs for diseases with no commercial, such

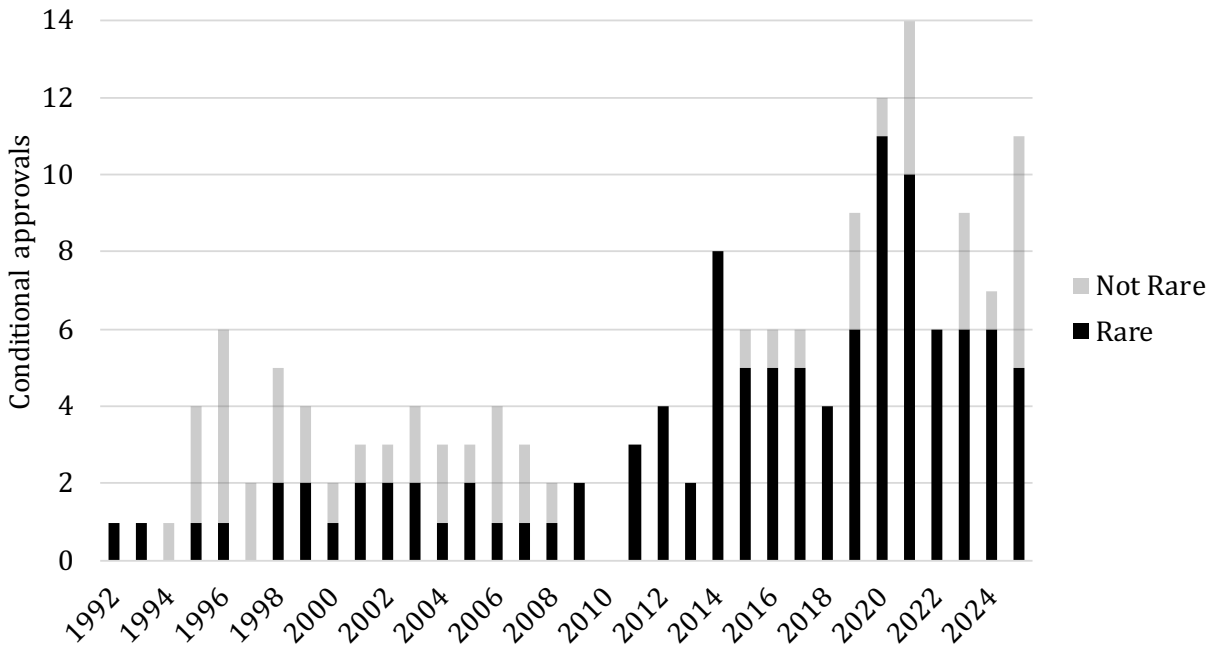


Figure 9: The FDA granted conditional approval to 84 drugs in the past decade. Most of the conditionally approved drugs treat rare diseases. The official FDA term for conditional approval is "accelerated approval" and for rare disease is "orphan disease." Source: Authors' analysis of data on approvals of new molecules by the Center for Drug Evaluation and Research.

as Gates Foundation funding for tropical disease drug development.

fact:revenueConditional approval increases the firm's revenue during conditional approval and several years afterward. Drug revenue typically rises during the first five years after approval (Robey David, 2016), as physicians and patients become more familiar with the drug. Demand is higher for drugs with which physicians and patients have more experience (Ridley Lee, 2020).

fact:phaseIIPhase II clinical testing provides limited information about a drug's efficacy. Sample sizes are typically smaller in Phase II than in Phase III and are insufficient for definitive inference. Phase I and II trials tend to be short and small to minimize patient risk and financial cost, given that most drugs entering Phase II do not progress to Phase III (DiMasi et al., 2016). We treat Phase I and II testing costs as sunk, but the model could be reframed to account for earlier decision-making by the firm.

fact:efficacyThe firm is legally required to share all efficacy and safety information about the drug with the government. Under FDA regulations, a developer seeking conditional approval must submit a comprehensive clinical data section covering both effectiveness and safety. The submission must present all relevant human evidence, not only studies supporting the proposed indication. In addition to detailed analyses of controlled trials, the sponsor is legally required to disclose any other information bearing on efficacy or safety, regardless of source. This includes evidence from controlled and uncontrolled studies of other uses, commercial marketing experience, published scientific reports, and unpublished studies, whether generated domestically or abroad (U.S. Food and Drug Administration, 2024).

fact:informationInformation from use of the drug outside randomized trials is not very informative, as it is noisy and subject to selection bias. For example, if healthier, high-income patients are more likely to purchase a drug, observed outcomes may overstate efficacy. Accordingly, our model assumes that usage after conditional approval does not affect Phase III testing. If "real-world evidence" were informative, it could reduce firms' testing costs by allowing smaller or shorter trials. Hence, we might underestimate the value of conditional approval.

fact:privateTesting costs are privately known to the firm. As Light Warburton (2005) observe, "this science-based industry refuses to allow independent parties to check the validity of their cost data and analyze it so that policy can be based on solid, objective, reproducible evidence." Phase III testing costs are plausibly private for several reasons. First, firms fear the hold-up problem: disclosing true costs would lead governments to reimburse only direct testing expenses, without compensating for failed projects or the opportunity cost

of capital. Second, benchmarks are limited because conditionally approved drugs are often first-in-class. Third, Phase III programs require firm-specific manufacturing scale-up that is difficult for regulators to observe. Fourth, most conditional approvals involve rare diseases, where patient recruitment is a major, typically in-house expense. Firms bear these costs through long-standing relationships with patient foundations, advocacy groups, and specialized clinicians, as well as through travel support and efforts to maintain enrollment. Fifth, clinical programs share resources across multiple drugs, making it hard for a government to audit costs.

fact:leniencyThe government can signal to the firm that it will grant final approval to a drug that narrowly misses the efficacy threshold. We provide evidence for both claims: (1) that the FDA sometimes grants final approval without persuasive evidence, and (2) that the FDA can signal its intended evidentiary standards. First, the FDA has granted final approval despite limited evidence of efficacy. Liu et al. (2024) report that 20% of cancer indications in their sample received full approval despite the absence of statistically significant evidence of benefit (Figure 8). In another instance, a senior FDA official acknowledged that a drug had been approved over scientific objections, describing it as "essentially a scientifically elegant placebo" (Kaplan, 2016). These cases demonstrate that, in practice, decision makers sometimes commit to and implement lenient evidentiary standards. Second, the FDA can signal or commit to a lenient threshold before confirmatory testing occurs. Under the 2012 breakthrough therapy designation, the FDA provides firms with "timely advice" and "interactive communication" to guide clinical trial design and clarify what evidence will be required for approval. Such interaction effectively communicates the future approval standard.

8 Calibration

We calibrate the model to illustrate the key economic drivers and to quantify the policy implications. The goal is not to make precise predictions for a specific drug, but rather to assess magnitudes using realistic parameter values.

We select a representative drug that received conditional approval: avelumab (brand name Bavencio). In 2017, the FDA granted conditional approval for avelumab to treat Merkel cell carcinoma, a rare and aggressive skin cancer, based on evidence that the drug produced substantial tumor shrinkage in patients who had few effective treatment options. In 2023, avelumab received final approval after confirmatory studies demonstrated concrete

clinical benefits, including improvements in overall survival.[13]

We bound the cost of Phase III testing based on other drugs in the therapeutic class. We assume that Phase III testing costs are incurred over the first three years, followed by regulatory review, with the final approval decision occurring at the start of year five. We then estimate doses and revenue over the drug's life cycle. We assume that the firm sells at monopoly prices for thirteen years, after which net revenue is negligible due to generic entry or price controls. Finally, we combine these inputs to compute the cumulative net present value and the implied optimal policy.

8.1 Estimated testing costs

Like the government, we do not observe avelumab's testing cost. Avelumab is marketed by a large firm that does not disclose drug-level development costs. To bound the testing cost, we therefore rely on data from firms that were only developing a few drugs at a time, for which reported firm-level development costs are equivalent to drug-level development costs. Wouters et al. (2020) report Phase III cost data for 13 antineoplastic and immunomodulating drugs approved between 2009 and 2018 and marketed by small firms in the same therapeutic class as avelumab. Among these drugs, we use brigatinib's Phase III cost of $98.1 million (2018 dollars) as a lower bound and sarilumab's cost of $630.7 million as an upper bound. These estimates are imprecise; in particular, Wouters et al. (2020) report low confidence in the sarilumab estimate. We therefore interpret this range as capturing plausible variation in testing costs rather than precise measurement.

8.2 Estimated doses and revenue

We estimate doses and revenue for the representative drug avelumab using Medicare and Medicaid data and extrapolate to total U.S. revenue.[14] We report values from the first full year of sales (2018) and from the fifth year (2022), the latter being more recent and more representative of long-run outcomes. Medicare and Medicaid data imply approximately 0.4 million doses in the first year and 1.4 million doses in year five (Table 1). We extrapolate to the national market by dividing these totals by 0.42, reflecting that Medicare and Medicaid account for about 42% of U.S. prescription drug spending according to National Health

[13]Following its initial approval for Merkel cell carcinoma, avelumab was subsequently approved for bladder and kidney cancers.

[14]We use sales aggregated across all indications, which overstates revenue for the initial skin cancer indication.

Expenditure data. This yields roughly one million doses in the first year and three million doses in year five. We project sales beyond year five using average drug life-cycle patterns (Robey David, 2016).

	2018	2022	Source
Medicaid+Medicare doses (MM)	0.4	1.4	Centers for Medicare & Medicaid Services (2024)
Medicaid+Medicare share	0.42	0.42	National Health Expenditures
Total doses (MM)	1.0	3.3	Medicaid+Medicare divided by .42
Medicaid+Medicare spending ($MM)	35	121	Centers for Medicare & Medicaid Services (2024)
Total sales ($MM)	83	288	Medicaid+Medicare divided by .42
Price per dose ($)	79	86	Medicare Part B average price Centers for Medicare & Medicaid Services (2024)
Willingness to pay per dose ($)	109	118	If buyer's surplus is half of seller's, willingness to pay is 1.375 times unit price
Production cost share of price	0.25	0.25	Wouters et al. (2024)

Table 1: Estimated annual doses and prices for avelumab in its first full year (2018) and its fifth year (2022).

8.3 Parameter values

For a drug granted full conditional approval, discounted doses sold (q) during that four-year period total 4.36 million (Table 2), using a 10.5% discount rate (DiMasi et al., 2016). This is about one million discounted doses per year, consistent with avelumab's first full year of sales (Table 1). Although sales increase over time, doses sold later are discounted more heavily, so higher sales are largely offset.

For a drug that receives both conditional and final approval, discounted doses sold ($\lambda_0 + \lambda_1$) total about 15 million, averaging roughly 1 million per year over the 13-year period between final approval and the loss of market power.

By contrast, a drug that receives final approval but not conditional approval sells fewer discounted doses—about 11 million over the same thirteen years (λ_0). Without conditional approval, sales begin only in year five and take several years to grow, when discounting is already substantial.

We normalize quantities by dividing all values by 4.36, so that $q = 1$.

To estimate revenue net of production costs, we assume a cost-of-goods-sold share of 25% (Wouters et al., 2024). In year five, the firm's net revenue per dose (π) is approximately \$65. If the drug is ineffective, the government's loss ($v(0)$) equals −\$86, the price paid for a drug that provides no benefit. If the drug is effective, the government's benefit ($v(1)$) equals \$32, the difference between willingness to pay and price.

Notation	**Description**	**Value**	**Source**
δ	discount rate	10.5%	DiMasi et al. (2016)
c_L	lower bound of the testing cost	98.1(\$MM)	Estimated Phase III testing cost of a similar drug (brigatinib) (Wouters et al., 2020).
c_H	upper bound of the testing cost	630.7(\$MM)	Estimated Phase III testing cost of a similar drug (sarilumab) (Wouters et al., 2020).
$v(0)$	government's loss from paying for an ineffective dose	\$86	Price (Table 1)
$v(1)$	government's payoff from paying for an effective dose	\$32	Willingness to pay minus price (Table 1)
Γ	price drop per dose in the competitive period	\$65	Price minus cost (Table 1)
q	total discounted doses with full conditional approval before final approval ($a_1 = 1$)	4.36(MM)	Equation (56).
λ_0	baseline total discounted doses if the drug is fully approved after testing is complete ($a_2 = 1$)	11.34(MM)	Equation (56).
λ_1	extra total discounted doses for a drug given the entire market in both periods ($a_1 = a_2 = 1$)	4.09(MM)	Equation (56).
Δ	total discounted doses in the competitive period	5.88(MM)	Equation (56).

Table 2: Estimated model parameters for avelumab.

$$\begin{cases} q = 4.36 \\ \lambda_0 = 11.34 \\ \lambda_0 + \lambda_1 = 15.43 \end{cases} \Rightarrow \begin{cases} q = 4.36 \\ \lambda_0 = 11.34 \\ \lambda_1 = 4.09 \end{cases} . \tag{56}$$

We also compute the discounted quantity sold after the monopoly period,

$$\Delta = \sum_{i=18}^{\infty} \frac{3.37}{(1+10.5\%)^i} = 5.88, \tag{57}$$

which captures the social value of earlier price reductions due to generic entry or regulation.

Figure 10 shows annual sales with full conditional approval and the parallel curve beginning at t_1 represents sales without conditional approval. Figure 10 illustrates doses in the first period (q), doses in the second period (λ_0), the bonus doses in the second period if the firm is granted both conditional and full approvals (λ_1), and doses in the competitive period (Δ). Hence, λ_0 denotes the number of doses sold before the competitive period under no conditional approval, and $\lambda_0 + \lambda_1$ denotes the number of doses sold under full conditional approval following conditional approval. In the model, we normalized the initial period doses (q) to 1.

8.4 Estimated cumulative net present value

We now compute the firm's cumulative net present value when testing costs are high (c_H = \$630.7 million), evenly distributed over the first three years, and full conditional approval is granted. We set the annual discount rate to $\delta = 10.5\%$, following DiMasi et al. (2016), and assume for simplicity that firms and the government discount future payoffs at the same rate. Table 3 reports discounted revenue, testing costs, and cumulative profits.

Despite large early losses, the firm eventually recovers its investment as sales grow and testing costs cease. Without conditional approval, firms facing high testing costs would forgo development entirely, generating a potential social loss.

Figure 11 plots cumulative profit net of testing costs under different assumptions about conditional approval (rows), final approval (columns), and testing costs (curves). Lower curves correspond to higher testing costs. In all cases, firms incur losses early on, but some reach positive cumulative profits depending on the approval regime. For drugs that do not receive conditional approval, cumulative profits continue to decline until one year before t_1, when the firm submits its testing data for regulatory review. Starting at t_1, cumulative profits either rise following final approval or are flat if the government denies final approval (and hence no new profits are cumulated).

Without conditional approval, only drugs with sufficiently low testing costs are profitable, and only if final approval is guaranteed (top left panel). With partial conditional approval,

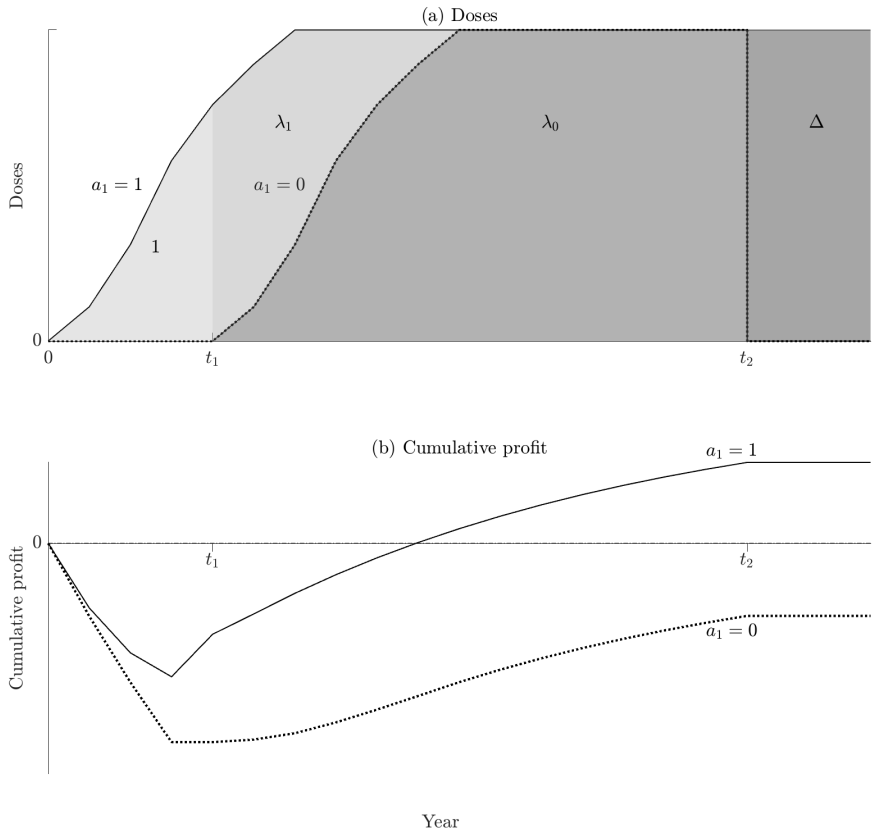


Figure 10: Annual revenue (top) and cumulative profit (bottom) depend on whether the government conditionally approves ($a_1 > 0$) or not ($a_1 = 0$) (assuming final approval). This figure matches Figure 2, but uses estimates for an actual drug avelumab. Source: Authors' analysis.

Year	Discounted net revenue (if in market) ($MM)	Discounted expected net revenue ($MM)	Annual testing cost ($MM)	Discounted annual testing cost ($MM)	Discounted annual net profit ($MM)	Cumulative net present value ($MM)
1	22	22	210	190	-169	-169
2	55	55	210	172	-117	-286
3	93	93	210	156	-63	-348
4	111	111	0	0	111	-238
5	117	53	0	0	53	-185
6	119	54	0	0	54	-131
7	108	49	0	0	49	-83
8	98	44	0	0	44	-39
9	88	40	0	0	40	1
10	80	36	0	0	36	37
11	72	33	0	0	33	70
12	65	29	0	0	29	99
13	59	27	0	0	27	126
14	54	24	0	0	24	150
15	49	22	0	0	22	172
16	44	20	0	0	20	191
17	40	18	0	0	18	209

Table 3: The firm's cumulative net present value when its testing cost is at the maximum value. "Discounted net revenue" is based on forecasted sales using average drug life cycles from Robey David (2016) and the price per dose from Table 1. Discounted expected net revenue equals discounted net revenue if the drug is in market multiplied by Pr(drug is in market), where this probability equals 1 in years 1 to 4 and 0.45 in years 5 to 17. Discounted annual net profit equals discounted expected net revenue minus discounted annual testing cost. Cumulative net present value in year i equals the sum of discounted annual net profit from years 1 to i. The annual testing cost is $630.7 million, distributed evenly over the first three years.

even high-cost drugs can eventually earn positive cumulative profits, if they are also granted final approval (middle left panel). Thus, conditional approval expands the set of drugs for which testing is privately worthwhile.

The left column of Figure 11 assumes final approval. When the possibility of failing confirmatory testing is introduced (right column), expected returns fall, which further strengthens the role of conditional approval in motivating firms to test.

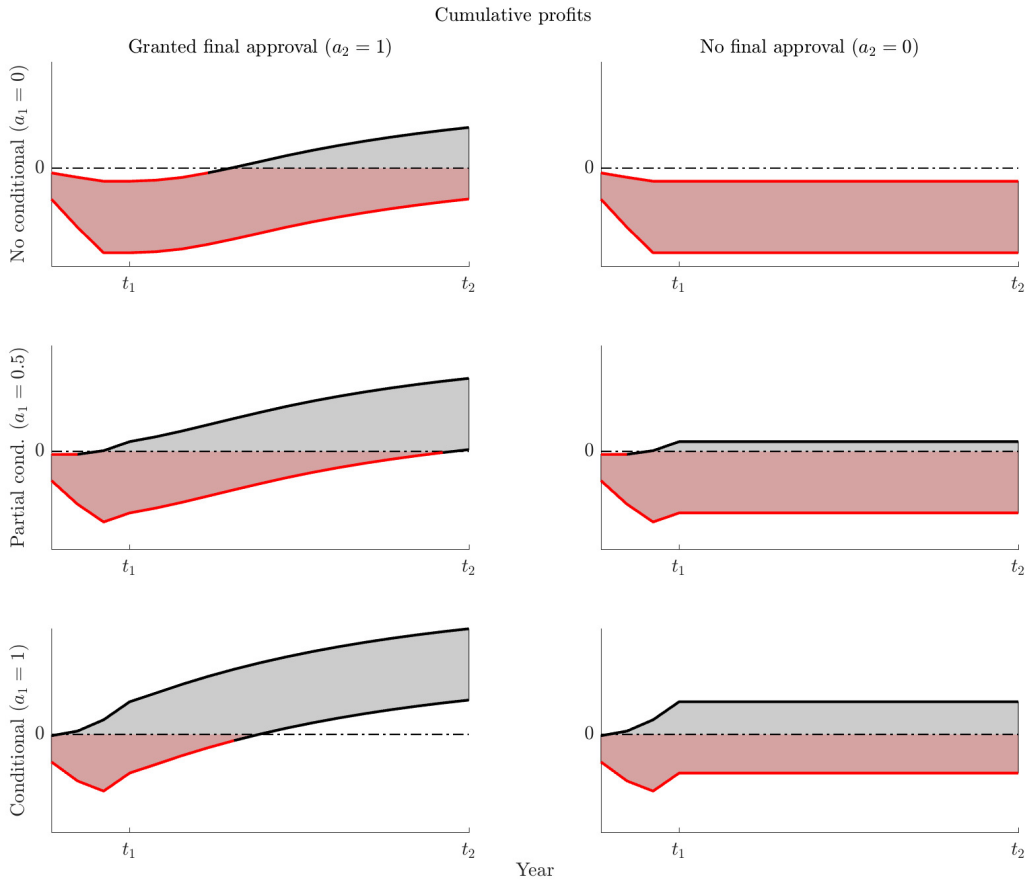


Figure 11: Cumulative net profit under different assumptions about conditional approval (rows), final approval (columns), and testing costs (curves). Lower curves correspond to higher testing costs. Source: Authors' analysis using a representative drug, avelumab.

8.5 Calibrated optimal policy

Finally, we estimate the optimal withdrawal rate: the fraction of conditionally approved drugs that are later denied final approval. Given a final approval threshold $\hat{\eta}$, the withdrawal probability is $G(\hat{\eta})$.

Figure 12 plots the withdrawal rate as a function of the drug's expected value. Two discontinuities arise. First, drugs with very low expected value ($v_E < v_2$) are not granted conditional approval. Second, at v_4, the final approval threshold jumps, increasing withdrawals.

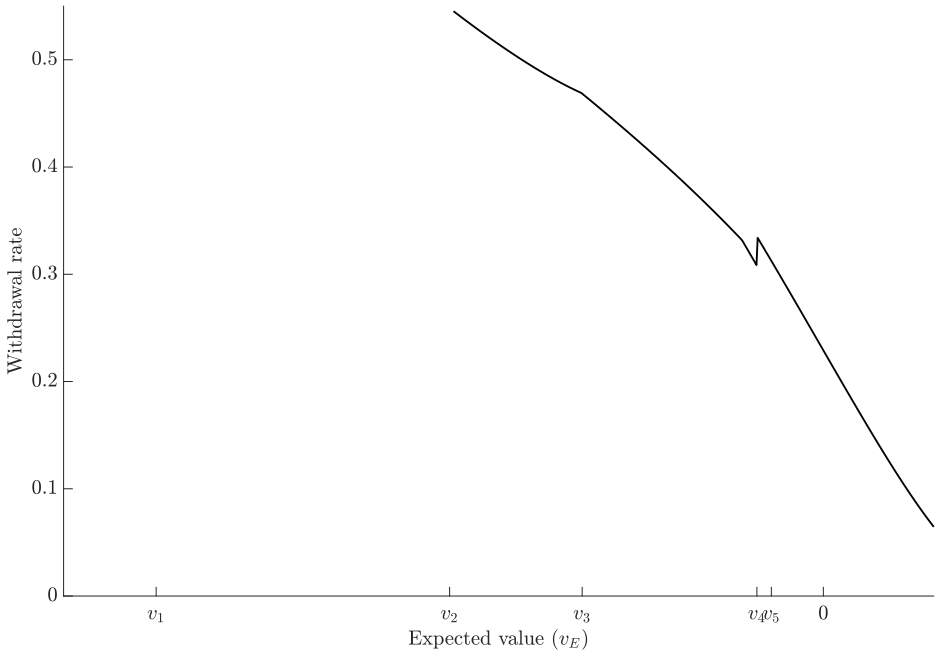


Figure 12: Optimal proportion of drugs withdrawn after receiving conditional approval. Source: Authors' analysis based on a representative drug.

For drugs with $v_E = 0$, the optimal withdrawal rate is 23%. Empirically, among cancer drugs granted conditional approval between 2013 and 2017, 22% were later withdrawn (Figure 8). This comparison suggests that historical FDA practice was close to optimal and, if anything, slightly conservative.

By contrast, the leadership of the FDA gene therapy unit appointed in 2025 has adopted a more restrictive approach to conditional approval (The Economist, 2025). Our calibration implies that such a shift will reduce social welfare.

8.6 Calibration in the Three-Period Game continued

Using calibrated parameters, we compare policies with and without conditional approval and leniency. Combining both tools performs near the upper bound. It generates over $6 billion in social value per decade, with most gains coming from conditional approval (Figure 13). See Section 5.3 for a more detailed explanation.

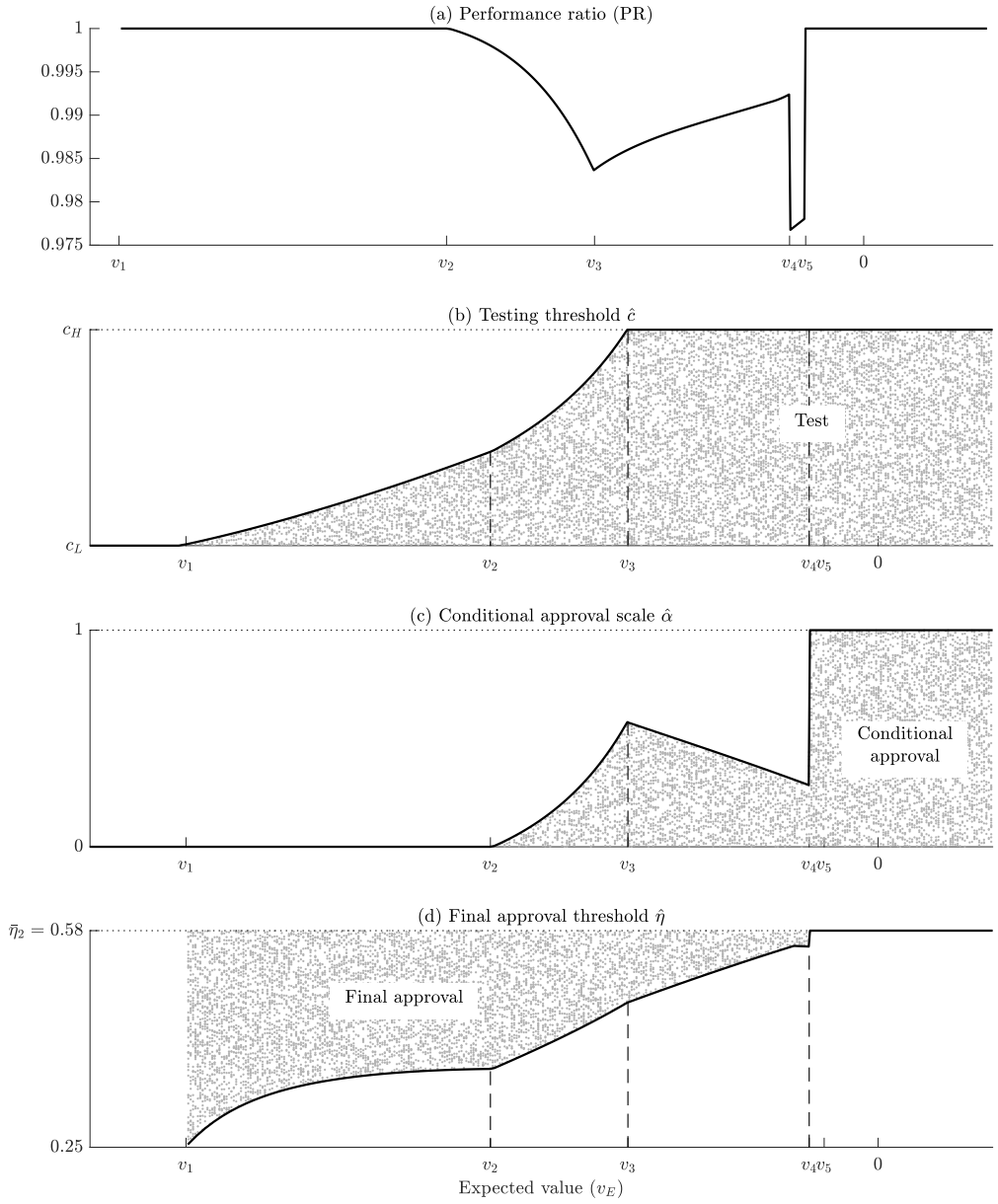


Figure 13: The lower-bound mechanism derived from the upper-bound problem performs well (upper panel). Drugs with high expected value and low testing cost are tested (second panel), receive conditional approval (third panel), and have a high final approval threshold, meaning less leniency (lower panel).

9 Flexible exclusivity

Optimal conditional approval and leniency may depend on the length of exclusivity. Some drug classes receive extensions. For example, antibiotics gain five additional years under the GAIN Act (Kong Zhao, 2024). Longer exclusivity increases monopoly sales, boosting firm profits but delaying savings for payers. Exclusivity extensions can therefore strengthen development incentives and reduce the need for conditional approval and leniency, all else equal.

Recall that in Figure 2, the curve starting at time 0 shows annual sales with full conditional approval; the parallel curve beginning at t_1 represents sales without conditional approval. Because the competitive period begins at time t_2, λ_0 and λ_1 correspond to the gray areas in the middle and on the left, respectively. The number of doses sold during the competitive period, denoted by Δ, is represented by the rectangle on the right.

Because annual sales peak in six years (Robey David, 2016), extending the firm's monopolistic period beyond 13 years does not affect λ_1 and only increases λ_0. Given that $\lambda_0+\Delta$ represents the total number of doses sold after final approval, corresponding to the area under the curve starting at t_1, any increase in λ_0 leads to an equivalent decrease in Δ. This adjustment corresponds to a rightward shift of time t_2 in Figure 2.

Let $\hat{\lambda}_0$ and Δ denote the calibrated values of λ_0 and Δ, respectively.[15] In Figure 14, we plot the testing threshold, the scale of conditional approval, the final approval threshold, and the expected payoffs of both the government and the firm (with testing costs \$200 and \$500 million) as functions of λ_0, varying from $\hat{\lambda}_0$ to $\hat{\lambda}_0 + \Delta$. The value $\hat{\lambda}_0$ corresponds to a scenario in which the firm's monopoly power ends at the end of year 13, while the upper bound $\hat{\lambda}_0 + \Delta$ reflects a case where the government extends monopoly power indefinitely. Fixing λ_0, the number of doses sold in the competitive period is $\hat{\lambda}_0 + \Delta - \lambda_0$.

As shown in panels (b) and (c), under the calibrated model parameters, when λ_0 increases, the scale of conditional approval decreases, and the final approval threshold increases. The testing threshold, illustrated in panel (a), also declines, indicating that the government induces less firm participation as the length of exclusivity increases. Finally, both the government (panel (d)) and the firm's expected payoffs (panel (e)) decrease as λ_0 increases, suggesting that it is optimal to maintain the current exclusivity timeline rather than granting additional exclusivity. Given the firm's testing cost, the monotonicity of its expected payoff in λ_0 follows from the monotonicity of the testing threshold $\hat{c}$. Specifically,

[15] See Appendix 8 for more details.

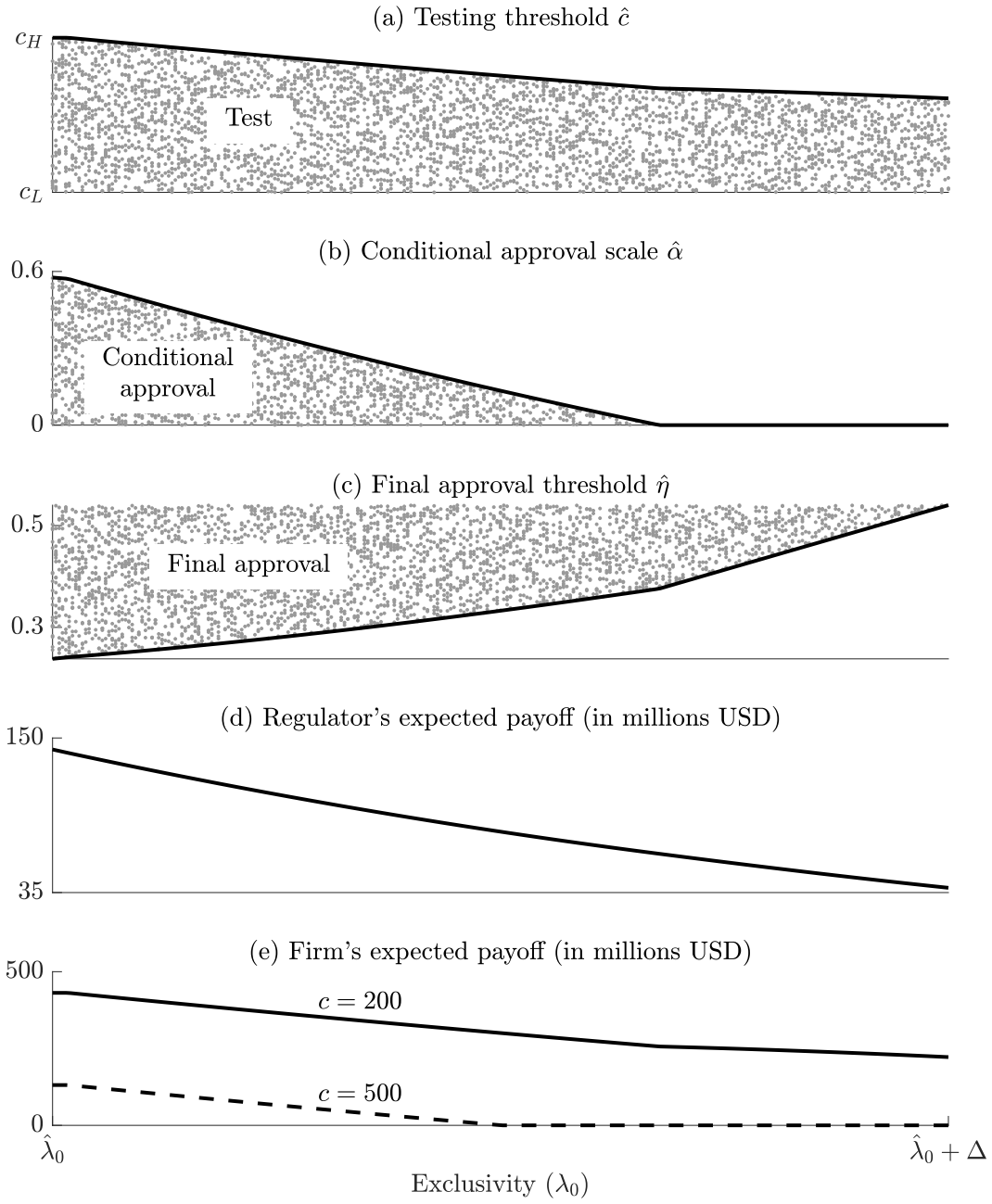


Figure 14: The testing threshold, scale of conditional approval, threshold for final approval, and the expected payoffs of the government and the firm (with different testing costs) under calibrated model parameters. The x-axis ranges from calibrated $\hat{\lambda}_0$ to $\hat{\lambda}_0 + \Delta$.

since the expected payoff of a firm with production cost $\hat{c}$ is zero, the expected payoff for a firm with cost c is given by $(c - \hat{c})^+$. This observation also highlights that, although firms generally favor extended exclusivity, they could be worse off if the government reoptimizes its policy in response to the change.

10 Private information about the efficacy distribution

In this section, we explore the case in which the distribution of efficacy η is the firm's private information. We use k to denote this private information, and the corresponding distribution of η is denoted by G_k. We assume that the family G_k satisfies the monotone likelihood ratio property (MLRP): for any $k_1 > k_2$, the likelihood ratio $g_{k_1}(\eta)/g_{k_2}(\eta)$ is non-decreasing in η. The government believes that k is a random variable with support $[k_L, k_H]$, cumulative distribution function F, and density f. A direct mechanism again specifies the probability of inducing the firm's testing $\tau(k)$, the scale of conditional approval $a_1(k)$, and the probability of granting final approval $a_2(k, \eta)$.

Unlike the base case with the cost c being the private information, even with $\Delta = 0$, the linear optimization problem is highly intractable. To be precise, the firm's expected payoff when its private information is k but it reports k' is

$$\hat{\Pi}(k, k') = \tau(k') \left[a_1(k') - c + [\lambda_0 + \lambda_1 a_1(k')] \int_0^1 a_2(k', \eta) \, \mathrm{d}G_k(\eta) \right]. \tag{58}$$

In equilibrium, the firm reports its type truthfully. The government's expected payoff (when $\Delta = 0$) is

$$U_R(\tau, a_1, a_2) \equiv \int_{k_L}^{k_H} \tau(k) \left\{ v_E(k) \, a_1(k) + [\lambda_0 + \lambda_1 \, a_1(k)] \int_0^1 v(\eta) \, a_2(k, \eta) \, \mathrm{d}G_k(\eta) \right\} \mathrm{d}F(k). \tag{59}$$

The government aims to maximize its expected payoff subject to the incentive constraints and feasibility constraints. The optimization problem remains (13), where the payoff functions are now given by (58) and (59).

The optimal solution to this problem is highly intractable and depends critically on the functional form of G_k. Figure 15 illustrates the optimal policy when the firm's private information k is uniformly distributed and the CDF is given by $G_k(\eta) = \eta^k$. The optimal mechanism features an infinite menu of contracts, with randomization over testing (i.e., τ strictly between 0 and 1) and a continuously varying scale of conditional approval and threshold for final approval. Such a mechanism is technically challenging to characterize analytically and difficult to implement in practice.

Therefore we again focus on a take-it-or-leave-it (TIOLI) offer. In this setup, the government offers a scale of conditional approval a_1 and commits to granting final approval if and only if $\eta \geq \eta^*$. The firm then decides whether to participate. The cutoff type, denoted

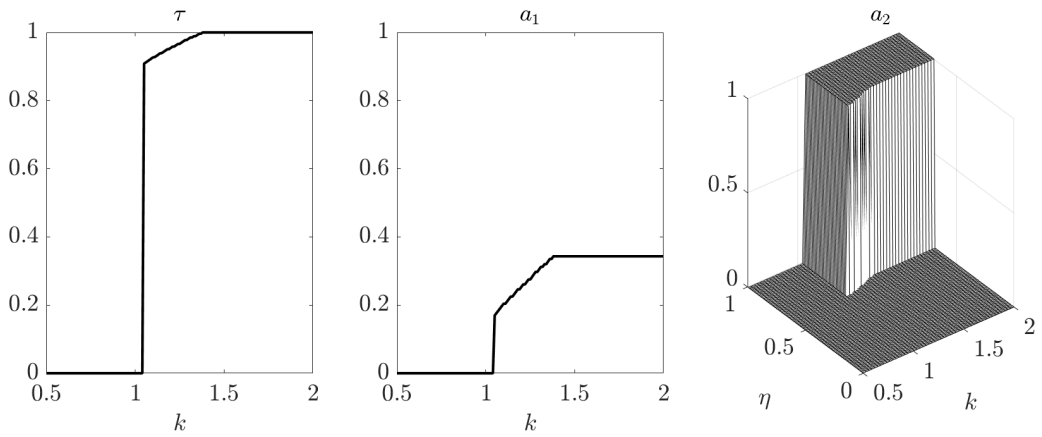


Figure 15: An optimal mechanism when the firm privately knows the distribution of efficacy η. We set $c = (c_L + c_H)/2$, $k_L = 0.5$, $k_H = 2$, $f(k) = 1/1.5$, and $G_k(\eta) = \eta^k$.

by k_0, is the type that is indifferent between testing and not testing, i.e.,

$$k_0 := \inf \{k_0 \in [k_L, k_H] : a_1 - c + (\lambda_0 + \lambda_1 a_1) [1 - G_{k_0}(\eta^*)] \geq 0\}. \tag{60}$$

The government's optimization problem under a TIOLI contract is therefore

$$\max_{a_1 \in [0,1], \eta^* \in [0,1], k_0 \in [k_L, k_H]} \int_{k_0}^{k_H} \left[v_E(k)\, a_1 + (\lambda_0 + \lambda_1 a_1) \int_{\eta^*}^{1} v(\eta)\, \mathrm{d}G_k(\eta) \right] \mathrm{d}F(k) \quad \text{s.t. (60).} \tag{61}$$

The optimization problem (61) is straightforward to solve numerically. Specifically, we discretize the values of a_1 and η^* over the unit square $[0,1]^2$. For each pair (a_1, η^*), the cutoff type k_0 is uniquely determined by (60), which then allows us to compute the corresponding objective value. Figure 16 displays the optimal TIOLI policy obtained from this numerical procedure.

Two features of the optimal policy coincide with our recommendations in the setting where the testing cost c is private information. First, when the testing cost is high, it is optimal for the government to offer (partial) conditional approval to incentivize testing. Second, for the same reason, it is optimal for the government to commit to a final approval threshold lower than $\bar{\eta}$. As shown in panel (c), the optimal threshold is almost always below the myopic threshold $\bar{\eta}$. Therefore, the structural properties in our model appear robust.

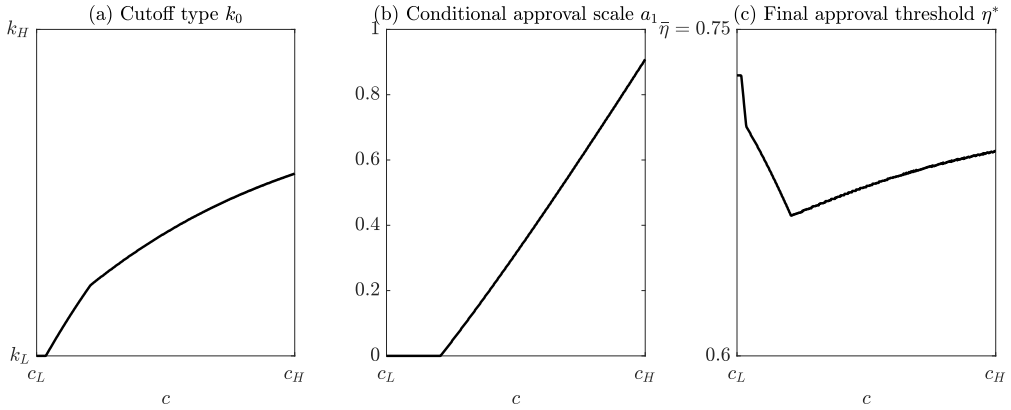


Figure 16: An optimal TIOLI policy when the firm privately knows the distribution of efficacy η. We set $k_L = 0.5$, $k_H = 2$, $f(k) = 1/1.5$, and $G_k(\eta) = \eta^k$.

11 Proofs

11.1 Proofs for Section 2

Proof. Proof of Lemma 1. By definition of Φ, we have

$$\frac{\partial \Phi(a_1, \eta; \Delta)}{\partial \eta} = (\lambda_0 + \lambda_1 a_1 + \Delta)\, v'(\eta) > 0.$$

Furthermore, by assumption (5), we have

$$\Phi(a_1, 0; \Delta) = (\lambda_0 + \lambda_1 a_1) \underbrace{v(0)}_{<0} + \Delta \underbrace{[v(0) + 1]}_{\leq 0} < 0$$

and

$$\Phi(a_1, 1; \Delta) = (\lambda_0 + \lambda_1 a_1)\, v(1) + \Delta\, [v(1) + 1] > 0.$$

This completes the proof. □

11.2 Proofs for Section 3

Proof. Proof of Proposition 1. For any c, $a_1(c)$, and $a_2(c, \eta)$, we have

$$\begin{aligned} & v_E\, a_1(c) + \int_0^1 \Big[[\lambda_0 + \lambda_1\, a_1(c)]\; v(\eta) + \Delta\; [v(\eta) + 1] \Big]\; a_2(c, \eta)\, \mathrm{d}G(\eta) \\ \leq\; & v_E\, a_1(c) + \int_{\eta_{\mathrm{NL}}(a_1(c))}^1 \Big[[\lambda_0 + \lambda_1\, a_1(c)]\; v(\eta) + \Delta\; [v(\eta) + 1] \Big]\, \mathrm{d}G(\eta) \end{aligned}$$

$$= \check{U}_R(a_1(c), \eta_{\text{NL}}(a_1(c))),$$

because the no-leniency threshold η_{NL} maximizes the government's continuation payoff after period one.

The function $\check{U}_R$ in the previous expression is the restriction of the function defined in (14), and it is convex in a_1 because

$$\frac{\mathrm{d}}{\mathrm{d}a_1}\check{U}_R(a_1, \eta_{\text{NL}}(a_1)) = v_E + \lambda_1 \int_{\eta_{\text{NL}}(a_1)}^{1} v(\eta)\,\mathrm{d}G(\eta)$$

and

$$\frac{\mathrm{d}^2}{\mathrm{d}a_1^2}\check{U}_R(a_1, \eta_{\text{NL}}(a_1)) = -\lambda_1\, v\left(\eta_{\text{NL}}(a_1)\right)\, g\left(\eta_{\text{NL}}(a_1)\right) \geq 0.$$

Therefore, the maximum is attained at either $a_1(c) = 0$ or $a_1(c) = 1$.

If the drug is promising and low-cost, i.e., (15) and (17) hold, we have

$$U_R^*(\Delta) \leq \max_{(a_1,\eta)\in[0,1]^2} \check{U}_R(a_1, \eta) = \check{U}_R(1, \eta_{\text{NL}}(1)),$$

because $\tau(c) \in [0, 1]$ and the upper bound is attained by the candidate solution defined in (19).

If the drug is not promising and low-cost, i.e., (16) and (18) hold, we have

$$U_R^*(\Delta) \leq \max_{(a_1,\eta)\in[0,1]^2} \check{U}_R(a_1, \eta) = \check{U}_R(0, \eta_{\text{NL}}(0)),$$

because $\tau(c) \in [0, 1]$ and the upper bound is attained by the candidate solution defined in (20).

This completes the proof. □

Proof. Proof of Corollary 1. The proof of Proposition 1 ignores both the (IC) and (IR) constraints. Hence, the same argument applies to the first-best problem and the dictatorial problem.

This completes the proof. □

11.3 Proofs for Section 4

Proof. Proof of Proposition 2. When $\Delta = 0$, any feasible solution $\{\tau(c), a_1(c), a_2(c, \eta)\}$ to the optimization problem (13) yields a corresponding solution $\{\tau(c), \alpha_1(c), \alpha_2(c, \eta)\}$, with α_1

and α_2 defined in (21) and (22), respectively—that is feasible for the optimization problem (27) and attains the same objective value.

Conversely, any feasible solution $\{\tau(c), \alpha_1(c), \alpha_2(c,\eta)\}$ to the optimization problem (27) yields a corresponding solution $\{\tau(c), a_1(c), a_2(c,\eta)\}$, with a_1 and a_2 defined in (28), that is feasible for the optimization problem (13) and likewise attains the same objective value.

This completes the proof. □

The following lemma is commonly used to simplify mechanism design problems involving direct payments between buyers and sellers. For completeness, we provide a proof here.

Lemma 3. *The incentive constraints* (IC) *and* (IR) *can be replaced by*

$$\tau \text{ non-increasing}, \tag{62}$$

$$\alpha_1(c) = \Pi(c_H) + \int_c^{c_H} \tau(y)\,\mathrm{d}y + c\,\tau(c) - \int_0^1 \alpha_2(c,\eta)\,\mathrm{d}G(\eta), \tag{63}$$

and

$$\Pi(c_H) \geq 0, \tag{64}$$

where $\Pi(c_H) \equiv \hat{\Pi}(c_H, c_H)$.

Proof. Proof of Lemma 3. The incentive-compatibility constraint can be rewritten as

$$\Pi(c) \geq \Pi(c') + \tau(c') \cdot (c' - c),$$

where $\Pi(c) \equiv \hat{\Pi}(c,c)$. Flipping the role of c and c' yields

$$\Pi(c') \geq \Pi(c) + \tau(c) \cdot (c - c').$$

Combining the two inequalities above yields the monotonicity of τ.

By the envelope theorem, we have

$$\Pi'(c) = -\tau(c) \leq 0,$$

which implies that

$$\Pi(c) = \Pi(c_H) + \int_c^{c_H} \tau(y)\,\mathrm{d}y.$$

Therefore, it is sufficient to impose the individual rationality constraint on type c_H and

$$\alpha_1(c) = \Pi(c_H) + \int_c^{c_H} \tau(y)\,\mathrm{d}y + c\,\tau(c) - \int_0^1 \alpha_2(c,\eta)\,\mathrm{d}G(\eta).$$

To prove the other direction, for any $c < c'$, we have

$$\begin{aligned}
\Pi(c) &= \Pi(c_H) + \int_c^{c_H} \tau(y)\,\mathrm{d}y \\
&= \Pi(c_H) + \int_{c'}^{c_H} \tau(y)\,\mathrm{d}y + \int_c^{c'} \tau(y)\,\mathrm{d}y \\
&= \Pi(c') + \int_c^{c'} \tau(y)\,\mathrm{d}y \\
&\geq \Pi(c') + \tau(c')\cdot(c'-c),
\end{aligned}$$

where the inequality follows from the monotonicity of τ. Constraint (IR) is implied by

$$\Pi(c) \geq \Pi(c_H) \geq 0.$$

This completes the proof. □

Lemma 4. *The optimization problem* (27) *is equivalent to*

$$\begin{aligned}
\max_{\{\tau,\alpha_2,\Pi(c_H)\}\in\Omega'_C} \quad & v_E\,\Pi(c_H) + \int_{c_L}^{c_H}\int_0^1 \left[v(\eta) - v_E\right]\,\alpha_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) \\
& + v_E \int_{c_L}^{c_H} \left[c + \frac{F(c)}{f(c)}\right]\,\tau(c)\,f(c)\,\,\mathrm{d}c,
\end{aligned} \tag{65}$$

where the feasible set Ω'_C *is determined by* (2), (62), (64), *and for any* $(c,\eta) \in [c_L, c_H]\times[0,1]$,

$$0 \leq \Pi(c_H) + \int_c^{c_H} \tau(y)\,\mathrm{d}y + c\,\tau(c) - \int_0^1 \alpha_2(c,\eta)\,\mathrm{d}G(\eta) \leq \tau(c), \tag{66}$$

and

$$0 \leq \alpha_2(c,\eta) \leq \lambda_0\,\tau(c) + \lambda_1\,\left[\Pi(c_H) + \int_c^{c_H} \tau(y)\,\mathrm{d}y + c\,\tau(c) - \int_0^1 \alpha_2(c,z)\,\mathrm{d}G(z)\right]. \tag{67}$$

Proof. Proof of Lemma 4. Substituting the equation (63) in Lemma 3 into the objective function yields the desired result. □

Proof. Proof of Lemma 2. We first reformulate the definitions of $\check{c}$, $\tilde{\eta}$, and $\underline{\eta}$ given in (36)–(38)

in a more explicit form:

$$\begin{aligned}
&\check{c} = \sup\left\{c \in [c_L, c_H] : v(\check{\eta})\,\left[c + \frac{F(c)}{f(c)}\right] + \lambda_0 \int_{\check{\eta}}^{1} [v(y) - v(\check{\eta})]\,\mathrm{d}G(y) \geq 0\right\},\\
&v(\tilde{\eta})\,\frac{F(\underline{c}(\tilde{\eta}))}{f(\underline{c}(\tilde{\eta}))} + \lambda_0 \int_{\tilde{\eta}}^{1} v(y)\,\mathrm{d}G(y) = 0, \text{ and}\\
&\underline{\eta} = \inf\left\{\eta \in [0, \check{\eta}] : v(\eta)\,\frac{F(\bar{c}(\eta))}{f(\bar{c}(\eta))} + v_E + (\lambda_0 + \lambda_1) \int_{\eta}^{1} v(y)\,\mathrm{d}G(y) \geq 0\right\}.
\end{aligned}$$

We start with the first part of the lemma. Because $\int_{\eta}^{1} [v(y) - v(\eta)]\,\mathrm{d}G(y)$ and $v(\eta)$ are decreasing and increasing in η, respectively, we know that function h is decreasing. Therefore, the unique existence of $\check{\eta} \in [0, \bar{\eta}]$ such that $h(\check{\eta}) = 0$ is implied by $h(\bar{\eta}) \leq 0 \leq h(0)$, or equivalently,

$$\lambda_1 \int_{\bar{\eta}}^{1} v(y)\,\mathrm{d}G(y) + v_E \leq 0 \leq \lambda_1\,(v_E - v(0)) - v(0) + v_E.$$

The second part of the inequality holds by $v_E - v(0) > 0$.

To prove the second part, we first verify that $v_E \leq 0$, which follows from

$$\check{\eta} \text{ uniquely exists} \quad \Rightarrow \quad \lambda_1 \underbrace{\int_{\bar{\eta}}^{1} v(y)\,\mathrm{d}G(y)}_{>0} + v_E \leq 0 \quad \Rightarrow \quad v_E < 0.$$

Define

$$t_1(\eta) \equiv v(\eta)\,\frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))} + \lambda_0 \int_{\eta}^{1} v(y)\,\mathrm{d}G(y), \quad \forall \eta \in [0, \bar{\eta}].$$

Taking derivative yields

$$\begin{aligned}
t_1'(\eta) &= v'(\eta)\,\frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))} + v(\eta)\,\frac{\mathrm{d}}{\mathrm{d}\eta}\left[\frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))}\right] - \lambda_0\,v(\eta)\,g(\eta)\\
&= v'(\eta)\,\frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))} - \lambda_0\,v(\eta)\,g(\eta)\;\frac{\mathrm{d}}{\mathrm{d}c}\left[c + \frac{F(c)}{f(c)}\right]\Bigg|_{c=\underline{c}(\eta)},
\end{aligned}$$

which is positive for any $\eta \in [0, \bar{\eta}]$. Therefore, to prove the unique existence of $\tilde{\eta}$, it suffices to verify that $t_1(\check{\eta}) \leq 0 \leq t_1(\bar{\eta})$, where the second inequality follows from

$$t_1(\bar{\eta}) = \lambda_0 \int_{\bar{\eta}}^{1} v(y)\,\mathrm{d}G(y) > 0.$$

Because $\check{c} < \underline{c}(\check{\eta})$, by definition of $\check{c}$, we have

$$v(\check{\eta}) \left[\underline{c}(\check{\eta}) + \frac{F(\underline{c}(\check{\eta}))}{f(\underline{c}(\check{\eta}))}\right] + \lambda_0 \int_{\check{\eta}}^{1} [v(y) - v(\check{\eta})] \, \mathrm{d}G(y) < 0,$$

which implies that

$$\begin{aligned} t_1(\check{\eta}) &= v(\check{\eta}) \, \frac{F(\underline{c}(\check{\eta}))}{f(\underline{c}(\check{\eta}))} + \lambda_0 \int_{\check{\eta}}^{1} v(y) \, \mathrm{d}G(y) \\ &< -v(\check{\eta}) \, \underline{c}(\check{\eta}) - \lambda_0 \int_{\check{\eta}}^{1} [v(y) - v(\check{\eta})] \, \mathrm{d}G(y) + \lambda_0 \int_{\check{\eta}}^{1} v(y) \, \mathrm{d}G(y) \\ &= -v(\check{\eta}) \, \underline{c}(\check{\eta}) + \lambda_0 \, v(\check{\eta}) \, [1 - G(\check{\eta})] \\ &= 0. \end{aligned}$$

Therefore, we know the unique existence of $\tilde{\eta} \in [\check{\eta}, \bar{\eta}]$, and the monotonicity property implies the desired result, i.e.,

$$h(\tilde{\eta}) \leq h(\check{\eta}) = 0.$$

It remains to prove the last part of the lemma. Define

$$t_2(\eta) \equiv v(\eta) \, \frac{F(\bar{c}(\eta))}{f(\bar{c}(\eta))} + \left[v_E + (\lambda_0 + \lambda_1) \int_{\eta}^{1} v(y) \, \mathrm{d}G(y)\right], \quad \forall \eta \in [0, \bar{\eta}].$$

Taking derivative yields

$$\begin{aligned} t_2'(\eta) &= v'(\eta) \, \frac{F(\bar{c}(\eta))}{f(\bar{c}(\eta))} + v(\eta) \, \frac{\mathrm{d}}{\mathrm{d}\eta} \left[\frac{F(\bar{c}(\eta))}{f(\bar{c}(\eta))}\right] - (\lambda_0 + \lambda_1) \, v(\eta) \, g(\eta) \\ &= v'(\eta) \, \frac{F(\bar{c}(\eta))}{f(\bar{c}(\eta))} - (\lambda_0 + \lambda_1) \, v(\eta) \, g(\eta) \, \left. \frac{\mathrm{d}}{\mathrm{d}c} \left[c + \frac{F(c)}{f(c)}\right] \right|_{c = \bar{c}(\eta)}, \end{aligned}$$

which is positive for any $\eta \in [0, \bar{\eta}]$. When $h(\bar{\eta}) > 0$, we have

$$t_2(\bar{\eta}) = v_E + (\lambda_0 + \lambda_1) \int_{\bar{\eta}}^{1} v(y) \, \mathrm{d}G(y) > \lambda_0 \int_{\bar{\eta}}^{1} v(y) \, \mathrm{d}G(y) > 0.$$

When $h(\bar{\eta}) \leq 0$ and $\check{c} > \bar{c}(\check{\eta})$, by definition of $\check{c}$, we have

$$v(\check{\eta}) \left[\bar{c}(\check{\eta}) + \frac{F(\bar{c}(\check{\eta}))}{f(\bar{c}(\check{\eta}))}\right] + \lambda_0 \int_{\check{\eta}}^{1} [v(y) - v(\check{\eta})] \, \mathrm{d}G(y) > 0,$$

which can be rewritten as

$$v(\check{\eta})\,\frac{F(\bar{c}(\check{\eta}))}{f(\bar{c}(\check{\eta}))}+\lambda_0\int_{\check{\eta}}^{1}v(y)\,\mathrm{d}G(y)>-v(\check{\eta})\,\left[1+\lambda_1\,\left[1-G(\check{\eta})\right]\right].$$

Because of the monotonicity of t_2, to prove the desired result, it suffices to verify that $t_2(\check{\eta})\geq 0$, which follows from

$$\begin{aligned}
t_2(\check{\eta}) &= v(\check{\eta})\,\frac{F(\bar{c}(\check{\eta}))}{f(\bar{c}(\check{\eta}))}+\left[v_E+(\lambda_0+\lambda_1)\int_{\check{\eta}}^{1}v(y)\,\mathrm{d}G(y)\right]\\
&> -v(\check{\eta})\,\left[1+\lambda_1\,\left[1-G(\check{\eta})\right]\right]+v_E+\lambda_1\int_{\check{\eta}}^{1}v(y)\,\mathrm{d}G(y)\\
&= \underbrace{v_E-v(\check{\eta})}_{=-\lambda_1\int_{\check{\eta}}^{1}[v(y)-v(\check{\eta})]\,\mathrm{d}G(y)}+\lambda_1\left[\int_{\check{\eta}}^{1}v(y)\,\mathrm{d}G(y)-v(\check{\eta})\,\left[1-G(\check{\eta})\right]\right]\\
&= 0.
\end{aligned}$$

Therefore, the threshold $\underline{\eta}$ is defined as

$$\underline{\eta}=\begin{cases}\mathring{\eta}, & \text{if } t_2(0)\leq 0\\ 0, & \text{if } t_2(0)>0,\end{cases}$$

where $\mathring{\eta}\in[0,\check{\eta}]$ is the unique zero of t_2.

To prove $h(\underline{\eta})\geq 0$, note that when $h(\bar{\eta})>0$, we have

$$h(\underline{\eta})\geq h(\bar{\eta})>0.$$

When $\check{c}>\bar{c}(\check{\eta})$, the monotonicity of h implies that

$$h(\underline{\eta})\geq h(\check{\eta})=0.$$

This completes the proof. □

Proof. Proof of Proposition 3. It is straightforward to verify that our candidate solutions are feasible in all cases.

To prove the optimality, we start with the first case. Define

$\kappa=-v(\check{\eta})$,

$$\nu(c,\eta) = [v(\check{\eta}) - v(\eta)]\ f(c)\, g(\eta), \quad \forall (c,\eta) \in [c_L, c_H] \times [0, \check{\eta}],$$

$$\mu(c,\eta) = [v(\eta) - v(\check{\eta})]\ f(c)\, g(\eta), \quad \forall (c,\eta) \in [c_L, c_H] \times [\check{\eta}, 1],$$

$$\sigma(c) = v_E\,[c\, f(c) + F(c)] + (\lambda_0 + \lambda_1\, c) \int_{\check{\eta}}^{1} \mu(c,\eta)\, \mathrm{d}\eta + \lambda_1 \int_{c_L}^{c} \int_{\check{\eta}}^{1} \mu(y,\eta)\, \mathrm{d}\eta\, \mathrm{d}y, \qquad \forall c \in \mathcal{C}_1, \text{ and}$$

$$\beta(c) = -v_E\,[c\, f(c) + F(c)] - (\lambda_0 + \lambda_1\, c) \int_{\check{\eta}}^{1} \mu(c,\eta)\, \mathrm{d}\eta - \lambda_1 \int_{c_L}^{c} \int_{\check{\eta}}^{1} \mu(y,\eta)\, \mathrm{d}\eta\, \mathrm{d}y, \quad \forall c \in \mathcal{C}_2,$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are subsets of $[c_L, c_H]$ satisfying: (i) $\mathcal{C}_1 \cup \mathcal{C}_2 = [c_L, c_H]$; and (ii) the measure of $\mathcal{C}_1 \cap \mathcal{C}_2$ is 0. These subsets will be specified later.

The nonnegativity of κ, ν, and μ follows from the definition of $\check{\eta}$ and the monotonicity of v directly.

The expressions of σ and β can be rewritten as

$$\sigma(c) = \underbrace{\left[v(\check{\eta})\ \left[c + \frac{F(c)}{f(c)}\right] + \lambda_0 \int_{\check{\eta}}^{1} [v(\eta) - v(\check{\eta})]\ g(\eta)\, \mathrm{d}\eta\right]}_{\text{decreasing in } c} f(c), \quad \forall c \in \mathcal{C}_1,$$

and

$$\beta(c) = -\left[v(\check{\eta})\ \left[c + \frac{F(c)}{f(c)}\right] + \lambda_0 \int_{\check{\eta}}^{1} [v(\eta) - v(\check{\eta})]\ g(\eta)\, \mathrm{d}\eta\right] f(c), \quad \forall c \in \mathcal{C}_2.$$

In the two possible cases, we have:

1. if $c_L \leq \check{c} < c_H$, we know that σ and β are nonnegative on $[c_L, \check{c}]$ and $[\check{c}, c_H]$, respectively. In this case, we have $\mathcal{C}_1 = [c_L, \check{c}]$ and $\mathcal{C}_2 = [\check{c}, c_H]$; and

2. if $\check{c} = c_H$, we know that σ is nonnegative on $[c_L, c_H]$. In this case, we have $\mathcal{C}_1 = [c_L, c_H]$ and $\mathcal{C}_2 = \emptyset$.

Because

$$-\kappa - \lambda_1 \int_{c_L}^{c_H} \int_{\check{\eta}}^{1} \mu(c,\eta)\, \mathrm{d}\eta\, \mathrm{d}c = v_E,$$

$\forall (c,\eta) \in [c_L, c_H] \times [0, \check{\eta}]$:

$$-\nu(c,\eta) + \lambda_1\, g(\eta) \int_{\check{\eta}}^{1} \mu(c,z)\, \mathrm{d}z = [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\check{\eta}, 1]$:

$$\mu(c,\eta) + \lambda_1\, g(\eta) \int_{\check{\eta}}^{1} \mu(c,z)\,\mathrm{d}z = [v(\eta) - v_E]\; g(\eta)\, f(c),$$

$\forall c \in \mathcal{C}_1$:

$$\sigma(c) - (\lambda_0 + \lambda_1\, c) \int_{\check{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1 \int_{c_L}^{c} \int_{\check{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y = v_E\,[c\, f(c) + F(c)],$$

$\forall c \in \mathcal{C}_2$:

$$-\beta(c) - (\lambda_0 + \lambda_1\, c) \int_{\check{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1 \int_{c_L}^{c} \int_{\check{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y = v_E\,[c\, f(c) + F(c)],$$

we have

$$\begin{aligned}
&\kappa\,[-\Pi(c_H)] + \int_{\mathcal{C}_1} \sigma(c)\,\tau(c)\,\mathrm{d}c + \int_{\mathcal{C}_2} \beta(c)\;[-\tau(c)]\,\mathrm{d}c + \int_{c_L}^{c_H} \int_{0}^{\check{\eta}} \nu(c,\eta)\;[-\alpha_2(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \int_{\check{\eta}}^{1} \mu(c,\eta) \left[\alpha_2(c,\eta) - \lambda_0\,\tau(c) - \lambda_1 \left[\Pi(c_H) + \int_{c}^{c_H} \tau(y)\,\mathrm{d}y + c\,\tau(c) - \int_{0}^{1} \alpha_2(c,z)\,\mathrm{d}G(z)\right]\right] \mathrm{d}\eta\,\mathrm{d}c \\
= &\left[-\kappa - \lambda_1 \int_{c_L}^{c_H} \int_{\check{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\right]\,\Pi(c_H) \\
&\quad + \int_{\mathcal{C}_1} \left[\sigma(c) - (\lambda_0 + \lambda_1\, c) \int_{\check{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1 \int_{c_L}^{c} \int_{\check{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\,\tau(c)\,\mathrm{d}c \\
&\quad + \int_{\mathcal{C}_2} \left[-\beta(c) - (\lambda_0 + \lambda_1\, c) \int_{\check{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1 \int_{c_L}^{c} \int_{\check{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\,\tau(c)\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \int_{0}^{\check{\eta}} \left[-\nu(c,\eta) + \lambda_1\, g(\eta) \int_{\check{\eta}}^{1} \mu(c,z)\,\mathrm{d}z\right]\,\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \int_{\check{\eta}}^{1} \left[\mu(c,\eta) + \lambda_1\, g(\eta) \int_{\check{\eta}}^{1} \mu(c,z)\,\mathrm{d}z\right]\,\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
= &\, v_E\,\Pi(c_H) + \int_{c_L}^{c_H} \int_{0}^{1} [v(\eta) - v_E]\;\alpha_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) + v_E \int_{c_L}^{c_H} \left[c + \frac{F(c)}{f(c)}\right]\,\tau(c)\,f(c)\;\mathrm{d}c.
\end{aligned}$$

Therefore, the objective function is bounded by

$$\int_{\mathcal{C}_1} \sigma(c)\,\mathrm{d}c.$$

Substituting our candidate solution into the objective function achieves the upper bound in

all the subcases, which follows from:

$$\begin{aligned}
&\int_{c_L}^{\check{c}} \sigma(c)\,\mathrm{d}c - \left[v_E\,\alpha^* + (\lambda_0 + \lambda_1\,\alpha^*)\int_{\check{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] F(\check{c}) \\
&= \int_{c_L}^{\check{c}} \left[v(\check{\eta})\,\left[c + \frac{F(c)}{f(c)}\right] + \lambda_0 \int_{\check{\eta}}^{1} [v(\eta) - v(\check{\eta})]\,g(\eta)\,\mathrm{d}\eta\right] f(c)\,\mathrm{d}c \\
&\quad - \left[v_E\,\alpha^* + (\lambda_0 + \lambda_1\,\alpha^*)\int_{\check{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] F(\check{c}) \\
&= v(\check{\eta})\int_{c_L}^{\check{c}} [(c - \check{c})\,f(c) + F(c)]\,\mathrm{d}c \\
&= 0,
\end{aligned}$$

where $\alpha^* = \frac{\check{c} - \lambda_0\,[1 - G(\check{\eta})]}{1 + \lambda_1\,[1 - G(\check{\eta})]}$.

This establishes the optimality in the first case.

To prove the optimality in the second case, we first define

$$\begin{aligned}
&\kappa = -v(\tilde{\eta}), \\
&\nu(c,\eta) = [v(\tilde{\eta}) - v(\eta)]\,f(c)\,g(\eta), && \forall (c,\eta) \in [c_L, c_H] \times [0, \tilde{\eta}], \\
&\mu(c,\eta) = [v(\eta) - v(\tilde{\eta})]\,f(c)\,g(\eta), && \forall (c,\eta) \in [c_L, c_H] \times [\tilde{\eta}, 1], \\
&\phi(c) = -\lambda_1 \int_{\tilde{\eta}}^{1} \mu(c,y)\,\mathrm{d}y + [v(\tilde{\eta}) - v_E]\,f(c), && \forall c \in [c_L, c_H], \\
&\sigma(c) = v_E\,[c\,f(c) + F(c)] + (\lambda_0 + \lambda_1\,c)\int_{\tilde{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta \\
&\qquad + \lambda_1 \int_{c_L}^{c}\int_{\tilde{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y + \int_{c_L}^{c} \phi(y)\,\mathrm{d}y + c\,\phi(c), && \forall c \in \mathcal{C}_1, \\
&\beta(c) = -v_E\,[c\,f(c) + F(c)] - (\lambda_0 + \lambda_1\,c)\int_{\tilde{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta \\
&\qquad - \lambda_1 \int_{c_L}^{c}\int_{\tilde{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - \int_{c_L}^{c} \phi(y)\,\mathrm{d}y - c\,\phi(c), && \forall c \in \mathcal{C}_2,
\end{aligned}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are subsets of $[c_L, c_H]$ satisfying: (i) $\mathcal{C}_1 \cup \mathcal{C}_2 = [c_L, c_H]$; and (ii) the measure of $\mathcal{C}_1 \cap \mathcal{C}_2$ is 0. These subsets will be specified later.

The nonnegativity of κ, ν, and μ follows from the definition of $\tilde{\eta}$ and the monotonicity of v directly.

The nonnegativity of ϕ follows from the second part of Lemma 2.

The expressions for σ and β can be simplified as

$$\sigma(c) = v(\tilde{\eta})\ [(c - \underline{c}(\tilde{\eta}))\, f(c) + F(c)] + \lambda_0\, f(c) \int_{\tilde{\eta}}^{1} v(\eta)\, \mathrm{d}G(\eta), \quad \forall c \in \mathcal{C}_1,$$

and

$$\beta(c) = -v(\tilde{\eta})\ [(c - \underline{c}(\tilde{\eta}))\, f(c) + F(c)] - \lambda_0\, f(c) \int_{\tilde{\eta}}^{1} v(\eta)\, \mathrm{d}G(\eta), \quad \forall c \in \mathcal{C}_2.$$

To prove the nonnegativity of σ, taking derivative yields

$$\sigma'(c) = v(\tilde{\eta})\ [(c - \underline{c}(\tilde{\eta}))\, f'(c) + 2\, f(c)] + f'(c) \underbrace{\lambda_0 \int_{\tilde{\eta}}^{1} v(\eta)\, \mathrm{d}G(\eta)}_{= -v(\tilde{\eta})\, \frac{F(\underline{c}(\tilde{\eta}))}{f(\underline{c}(\tilde{\eta}))} \geq 0} \leq 0, \quad \forall c \leq \underline{c}(\tilde{\eta}).$$

Meanwhile, we have

$$\sigma(\underline{c}(\tilde{\eta})) = \left[v(\tilde{\eta})\, \frac{F(\underline{c}(\tilde{\eta}))}{f(\underline{c}(\tilde{\eta}))} + \lambda_0 \int_{\tilde{\eta}}^{1} v(\eta)\, \mathrm{d}G(\eta) \right]\ f(\underline{c}(\tilde{\eta})) = 0.$$

In the first two possible cases, we have:

1. if $\underline{c}(\tilde{\eta}) < c_L$, we know that β is nonnegative on $[c_L, c_H]$. In this case, we have $\mathcal{C}_1 = \emptyset$ and $\mathcal{C}_2 = [c_L, c_H]$; and
2. if $c_L \leq \underline{c}(\tilde{\eta}) \leq c_H$, we know that σ and β are nonnegative on $[c_L, \underline{c}(\tilde{\eta})]$ and $[\underline{c}(\tilde{\eta}), c_H]$, respectively. In this case, we have $\mathcal{C}_1 = [c_L, \underline{c}(\tilde{\eta})]$ and $\mathcal{C}_2 = [\underline{c}(\tilde{\eta}), c_H]$.

Therefore, the nonnegativity of σ and β holds automatically.

Because

$$-\kappa - \int_{c_L}^{c_H} \phi(c)\, \mathrm{d}c - \lambda_1 \int_{c_L}^{c_H} \int_{\tilde{\eta}}^{1} \mu(c, \eta)\, \mathrm{d}\eta\, \mathrm{d}c = v_E,$$

$\forall (c, \eta) \in [c_L, c_H] \times [0, \tilde{\eta}]$:

$$\phi(c)\, g(\eta) - \nu(c, \eta) + \lambda_1\, g(\eta) \int_{\tilde{\eta}}^{1} \mu(c, z)\, \mathrm{d}z = [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall (c, \eta) \in [c_L, c_H] \times [\tilde{\eta}, 1]$:

$$\phi(c)\, g(\eta) + \mu(c, \eta) + \lambda_1\, g(\eta) \int_{\tilde{\eta}}^{1} \mu(c, z)\, \mathrm{d}z = [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall c \in \mathcal{C}_1$:

$$-\int_{c_L}^{c} \phi(y)\,\mathrm{d}y - c\,\phi(c) + \sigma(c) - (\lambda_0 + \lambda_1\,c)\,\int_{\tilde{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1\,\int_{c_L}^{c}\int_{\tilde{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y = v_E\,[c\,f(c) + F(c)],$$

$\forall c \in \mathcal{C}_2$:

$$-\int_{c_L}^{c} \phi(y)\,\mathrm{d}y - c\,\phi(c) - \beta(c) - (\lambda_0 + \lambda_1\,c)\,\int_{\tilde{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1\,\int_{c_L}^{c}\int_{\tilde{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y = v_E\,[c\,f(c) + F(c)],$$

we have

$$\begin{aligned}
&\kappa\,[-\Pi(c_H)] + \int_{\mathcal{C}_1} \sigma(c)\,\tau(c)\,\mathrm{d}c + \int_{\mathcal{C}_2} \beta(c)\,\left[-\tau(c)\right]\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \phi(c)\,\left[-\Pi(c_H) - \int_{c}^{c_H} \tau(y)\,\mathrm{d}y - c\,\tau(c) + \int_{0}^{1} \alpha_2(c,\eta)\,\mathrm{d}G(\eta)\right]\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{0}^{\tilde{\eta}} \nu(c,\eta)\,\left[-\alpha_2(c,\eta)\right]\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\tilde{\eta}}^{1} \mu(c,\eta)\,\left[\alpha_2(c,\eta) - \lambda_0\,\tau(c) - \lambda_1\,\left[\Pi(c_H) + \int_{c}^{c_H} \tau(y)\,\mathrm{d}y + c\,\tau(c) - \int_{0}^{1} \alpha_2(c,z)\,\mathrm{d}G(z)\right]\right]\mathrm{d}\eta\,\mathrm{d}c \\
=\ &\left[-\kappa - \int_{c_L}^{c_H} \phi(c)\,\mathrm{d}c - \lambda_1\,\int_{c_L}^{c_H}\int_{\tilde{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\right]\,\Pi(c_H) \\
&+ \int_{\mathcal{C}_1} \left[\sigma(c) - \int_{c_L}^{c} \phi(y)\,\mathrm{d}y - c\,\phi(c) - (\lambda_0 + \lambda_1\,c)\,\int_{\tilde{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1\,\int_{c_L}^{c}\int_{\tilde{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\,\tau(c)\,\mathrm{d}c \\
&+ \int_{\mathcal{C}_2} \left[-\beta(c) - \int_{c_L}^{c} \phi(y)\,\mathrm{d}y - c\,\phi(c) - (\lambda_0 + \lambda_1\,c)\,\int_{\tilde{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1\,\int_{c_L}^{c}\int_{\tilde{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\,\tau(c)\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{0}^{\tilde{\eta}} \left[\phi(c)\,g(\eta) - \nu(c,\eta) + \lambda_1\,g(\eta)\,\int_{\tilde{\eta}}^{1} \mu(c,z)\,\mathrm{d}z\right]\,\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\tilde{\eta}}^{1} \left[\phi(c)\,g(\eta) + \mu(c,\eta) + \lambda_1\,g(\eta)\,\int_{\tilde{\eta}}^{1} \mu(c,z)\,\mathrm{d}z\right]\,\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
=\ &v_E\,\Pi(c_H) + \int_{c_L}^{c_H}\int_{0}^{1} \left[v(\eta) - v_E\right]\,\alpha_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) + v_E\,\int_{c_L}^{c_H} \left[c + \frac{F(c)}{f(c)}\right]\,\tau(c)\,f(c)\,\mathrm{d}c.
\end{aligned}$$

Therefore, the objective function is bounded by

$$\int_{\mathcal{C}_1} \sigma(c)\,\mathrm{d}c.$$

Substituting our candidate solution into the objective function achieves the upper bound in

the first two cases, which follows from:

1. if $\underline{c}(\tilde{\eta}) \leq c_L$, we have $c^* = c_L$ and the measure of $\mathcal{C}_1$ is 0. The government's expected payoff generated by our candidate policy and its upper bound are both 0;

2. if $c_L < \underline{c}(\tilde{\eta}) \leq c_H$, we have

$$\begin{aligned}
&\int_{c_L}^{\underline{c}(\tilde{\eta})} \sigma(c)\,\mathrm{d}c - \lambda_0\,F(\underline{c}(\tilde{\eta}))\,\int_{\tilde{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta) \\
&= \int_{c_L}^{\underline{c}(\tilde{\eta})} \left[v(\tilde{\eta})\,\left[[c-\underline{c}(\tilde{\eta})]\,f(c)+F(c)\right] + \lambda_0\,f(c)\,\int_{\tilde{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta)\right]\mathrm{d}c - \lambda_0\,F(\underline{c}(\tilde{\eta}))\,\int_{\tilde{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta) \\
&= v(\tilde{\eta})\,\int_{c_L}^{\underline{c}(\tilde{\eta})} \left[[c-\underline{c}(\tilde{\eta})]\,f(c)+F(c)\right]\mathrm{d}c \\
&= 0.
\end{aligned}$$

To verify the optimality in case (2)(c), define

$$\begin{aligned}
&\kappa = -v(\eta^*), \\
&\nu(c,\eta) = [v(\eta^*)-v(\eta)]\;f(c)\,g(\eta), && \forall (c,\eta)\in[c_L,c_H]\times[0,\eta^*], \\
&\mu(c,\eta) = [v(\eta)-v(\eta^*)]\;f(c)\,g(\eta), && \forall (c,\eta)\in[c_L,c_H]\times[\eta^*,1], \\
&\phi(c) = -\lambda_1\int_{\eta^*}^{1}\mu(c,y)\,\mathrm{d}y + [v(\eta^*)-v_E]\;f(c), && \forall c\in[c_L,c_H], \\
&\sigma(c) = v_E\,[c\,f(c)+F(c)] + (\lambda_0+\lambda_1\,c)\int_{\eta^*}^{1}\mu(c,\eta)\,\mathrm{d}\eta \\
&\qquad + \lambda_1\int_{c_L}^{c}\int_{\eta^*}^{1}\mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y + \int_{c_L}^{c}\phi(y)\,\mathrm{d}y + c\,\phi(c), && \forall c\in[c_L,c_H],
\end{aligned}$$

where $\eta^* \equiv \sup\{\eta\in[0,1] : \underline{c}(\eta)\geq c_H\}$. Because $\underline{c}(\tilde{\eta}) > c_H > \underline{c}(\bar{\eta})$, we know that $\tilde{\eta} < \eta^* < \bar{\eta}$. The nonnegativity of κ, ν, and μ follows from the corresponding definitions directly. Because

$$\phi(c) \geq -\lambda_1\int_{\tilde{\eta}}^{1}[v(y)-v(\tilde{\eta})]\;f(c)\,g(y)\,\mathrm{d}y + [v(\tilde{\eta})-v_E]\;f(c) \geq 0, \quad \forall c\in[c_L,c_H],$$

the desired nonnegativity condition holds automatically. Finally, we have

$$\sigma(c) = v(\eta^*)\;[(c-c_H)\,f(c)+F(c)] + \lambda_0\,f(c)\int_{\eta^*}^{1} v(\eta)\,\mathrm{d}G(\eta), \quad \forall c\in[c_L,c_H].$$

The monotonicity of σ still holds, and we have

$$\begin{aligned}\sigma(c_H) &= \left[v(\eta^*)\,\frac{F(c_H)}{f(c_H)} + \lambda_0 \int_{\eta^*}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] f(c_H)\\ &\geq \left[v(\tilde{\eta})\,\frac{F(\underline{c}(\tilde{\eta}))}{f(\underline{c}(\tilde{\eta}))} + \lambda_0 \int_{\tilde{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] f(c_H)\\ &= 0,\end{aligned}$$

where the inequality follows from the monotonicity property of v and the fact that $\eta^* > \tilde{\eta}$.

Because

$$\begin{aligned}&\kappa\,[-\Pi(c_H)] + \int_{c_L}^{c_H} \sigma(c)\,\tau(c)\,\mathrm{d}c\\ &+ \int_{c_L}^{c_H} \phi(c)\,\left[-\Pi(c_H) - \int_{c}^{c_H} \tau(y)\,\mathrm{d}y - c\,\tau(c) + \int_{0}^{1} \alpha_2(c,\eta)\,\mathrm{d}G(\eta)\right]\mathrm{d}c\\ &+ \int_{c_L}^{c_H}\int_{0}^{\eta^*} \nu(c,\eta)\,\left[-\alpha_2(c,\eta)\right]\mathrm{d}\eta\,\mathrm{d}c\\ &+ \int_{c_L}^{c_H}\int_{\eta^*}^{1} \mu(c,\eta)\,\left[\alpha_2(c,\eta) - \lambda_0\,\tau(c) - \lambda_1\,\left[\Pi(c_H) + \int_{c}^{c_H} \tau(y)\,\mathrm{d}y + c\,\tau(c) - \int_{0}^{1} \alpha_2(c,z)\,\mathrm{d}G(z)\right]\right]\mathrm{d}\eta\,\mathrm{d}c\\ =\ &\left[-\kappa - \int_{c_L}^{c_H} \phi(c)\,\mathrm{d}c - \lambda_1 \int_{c_L}^{c_H}\int_{\eta^*}^{1} \mu(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\right]\,\Pi(c_H)\\ &+ \int_{c_L}^{c_H}\left[\sigma(c) - \int_{c_L}^{c} \phi(y)\,\mathrm{d}y - c\,\phi(c) - (\lambda_0 + \lambda_1\,c)\int_{\eta^*}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1 \int_{c_L}^{c}\int_{\eta^*}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\,\tau(c)\,\mathrm{d}c\\ &+ \int_{c_L}^{c_H}\int_{0}^{\eta^*}\left[\phi(c)\,g(\eta) - \nu(c,\eta) + \lambda_1\,g(\eta)\int_{\eta^*}^{1} \mu(c,z)\,\mathrm{d}z\right]\,\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\\ &+ \int_{c_L}^{c_H}\int_{\eta^*}^{1}\left[\phi(c)\,g(\eta) + \mu(c,\eta) + \lambda_1\,g(\eta)\int_{\eta^*}^{1} \mu(c,z)\,\mathrm{d}z\right]\,\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\\ =\ &v_E\,\Pi(c_H) + \int_{c_L}^{c_H}\int_{0}^{1}\left[v(\eta) - v_E\right]\,\alpha_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) + v_E\int_{c_L}^{c_H}\left[c + \frac{F(c)}{f(c)}\right]\,\tau(c)\,f(c)\,\mathrm{d}c,\end{aligned}$$

the objective function is bounded by

$$\int_{c_L}^{c_H} \sigma(c)\,\mathrm{d}c.$$

Substituting our candidate solution into the objective function achieves the upper bound,

which follows from

$$\begin{aligned}
&\int_{c_L}^{c_H} \sigma(c)\,\mathrm{d}c - \lambda_0 \int_{\eta^*}^{1} v(\eta)\,\mathrm{d}G(\eta) \\
&= \int_{c_L}^{c_H} \left[v(\eta^*)\,\left[(c-c_H)\,f(c)+F(c)\right] + f(c)\,\lambda_0 \int_{\eta^*}^{1} v(\eta)\,\mathrm{d}G(\eta) \right] \mathrm{d}c - \lambda_0 \int_{\eta^*}^{1} v(\eta)\,\mathrm{d}G(\eta) \\
&= v(\eta^*) \int_{c_L}^{c_H} \left[(c-c_H)\,f(c)+F(c)\right] \mathrm{d}c \\
&= 0.
\end{aligned}$$

It remains to prove the last part of the proposition, where the optimal mechanism is to grant full conditional approval to firms whose types are below the threshold c^*. In the first two subcases, we have $\bar{c}(\underline{\eta}) < c_H \leq \bar{c}(0)$, implying that $\underline{\eta} = \eta^*$.

Define

$$\begin{aligned}
&\kappa = -v(\underline{\eta}), \\
&\nu(c,\eta) = \left[v(\underline{\eta}) - v(\eta)\right]\, f(c)\,g(\eta), && \forall (c,\eta) \in [c_L,c_H]\times[0,\underline{\eta}], \\
&\mu(c,\eta) = \left[v(\eta) - v(\underline{\eta})\right]\, f(c)\,g(\eta), && \forall (c,\eta) \in [c_L,c_H]\times[\underline{\eta},1], \\
&\psi(c) = \lambda_1 \int_{\underline{\eta}}^{1} \mu(c,y)\,\mathrm{d}y - \left[v(\underline{\eta}) - v_E\right]\, f(c), && \forall c \in [c_L,c_H], \\
&\sigma(c) = v_E\left[c\,f(c)+F(c)\right] + (\lambda_0+\lambda_1\,c) \int_{\underline{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta \\
&\qquad + \lambda_1 \int_{c_L}^{c} \int_{\underline{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - \int_{c_L}^{c} \psi(y)\,\mathrm{d}y - (c-1)\,\psi(c), && \forall c \in \mathcal{C}_1, \\
&\beta(c) = -v_E\left[c\,f(c)+F(c)\right] - (\lambda_0+\lambda_1\,c) \int_{\underline{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta \\
&\qquad - \lambda_1 \int_{c_L}^{c} \int_{\underline{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y + \int_{c_L}^{c} \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c), && \forall c \in \mathcal{C}_2,
\end{aligned}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are subsets of $[c_L,c_H]$ satisfying: (i) $\mathcal{C}_1 \cup \mathcal{C}_2 = [c_L,c_H]$; and (ii) the measure of $\mathcal{C}_1 \cap \mathcal{C}_2$ is 0. These subsets will be specified later.

The nonnegativity of κ, ν, and μ follows from the definition of $\underline{\eta}$ and the monotonicity of v directly.

The nonnegativity of ψ follows from the last part of Lemma 2.

The expressions of σ and β can be rewritten as

$$\sigma(c) = v(\underline{\eta})\ \left[(c - \bar{c}(\underline{\eta}))\, f(c) + F(c)\right] + f(c)\ \left[v_E + (\lambda_0 + \lambda_1)\ \int_{\underline{\eta}}^{1} v(y)\, \mathrm{d}G(y)\right]$$

and

$$\beta(c) = -v(\underline{\eta})\ \left[(c - \bar{c}(\underline{\eta}))\, f(c) + F(c)\right] - f(c)\ \left[v_E + (\lambda_0 + \lambda_1)\ \int_{\underline{\eta}}^{1} v(y)\, \mathrm{d}G(y)\right].$$

Differentiating σ yields

$$\sigma'(c) = v(\underline{\eta})\ \left[(c - \bar{c}(\underline{\eta}))\, f'(c) + 2\, f(c)\right] + f'(c)\ \underbrace{\left[v_E + (\lambda_0 + \lambda_1)\ \int_{\underline{\eta}}^{1} v(y)\, \mathrm{d}G(y)\right]}_{=-v(\underline{\eta})\, \frac{F(\bar{c}(\underline{\eta}))}{f(\bar{c}(\underline{\eta}))}} \leq 0, \quad \forall c \leq \bar{c}(\underline{\eta}).$$

Meanwhile, we have

$$\sigma(\bar{c}(\underline{\eta})) = \left[v(\underline{\eta})\, \frac{F(\bar{c}(\underline{\eta}))}{f(\bar{c}(\underline{\eta}))} + \left[v_E + (\lambda_0 + \lambda_1)\ \int_{\underline{\eta}}^{1} v(y)\, \mathrm{d}G(y)\right]\right]\ f(\bar{c}(\underline{\eta})) = 0.$$

In the first two possible cases, we have:

1. if $\bar{c}(\underline{\eta}) < c_L$, we know that β is nonnegative on $[c_L, c_H]$. In this case, we have $\mathcal{C}_1 = \emptyset$ and $\mathcal{C}_2 = [c_L, c_H]$;

2. if $c_L \leq \bar{c}(\underline{\eta}) \leq c_H$, we know that σ and β are nonnegative on $[c_L, \bar{c}(\underline{\eta})]$ and $[\bar{c}(\underline{\eta}), c_H]$, respectively. In this case, we have $\mathcal{C}_1 = [c_L, \bar{c}(\underline{\eta})]$ and $\mathcal{C}_2 = [\bar{c}(\underline{\eta}), c_H]$.

Therefore, the nonnegativity of σ and β holds automatically.

Because

$$-\kappa + \int_{c_L}^{c_H} \psi(c)\, \mathrm{d}c - \lambda_1\ \int_{c_L}^{c_H} \int_{\underline{\eta}}^{1} \mu(c, \eta)\, \mathrm{d}\eta\, \mathrm{d}c = v_E,$$

$\forall (c, \eta) \in [c_L, c_H] \times [0, \underline{\eta}]$:

$$-\psi(c)\, g(\eta) - \nu(c, \eta) + \lambda_1\, g(\eta)\ \int_{\underline{\eta}}^{1} \mu(c, y)\, \mathrm{d}y = [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\underline{\eta}, 1]$:

$$-\psi(c)\, g(\eta) + \mu(c,\eta) + \lambda_1\, g(\eta) \int_{\underline{\eta}}^{1} \mu(c,y)\,\mathrm{d}y = [v(\eta) - v_E]\; g(\eta)\, f(c),$$

$\forall c \in \mathcal{C}_1$:

$$\int_{c_L}^{c} \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) + \sigma(c) - (\lambda_0 + \lambda_1\, c) \int_{\underline{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1 \int_{c_L}^{c}\int_{\underline{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y = v_E\,[c\, f(c) + F(c)],$$

$\forall c \in \mathcal{C}_2$:

$$\int_{c_L}^{c} \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) - \beta(c) - (\lambda_0 + \lambda_1\, c) \int_{\underline{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1 \int_{c_L}^{c}\int_{\underline{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y = v_E\,[c\, f(c) + F(c)],$$

we have

$$\begin{aligned}
&\kappa\,[-\Pi(c_H)] + \int_{\mathcal{C}_1} \sigma(c)\,\tau(c)\,\mathrm{d}c + \int_{\mathcal{C}_2} \beta(c)\;[-\tau(c)]\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \psi(c) \left[\Pi(c_H) + \int_{c}^{c_H} \tau(y)\,\mathrm{d}y + c\,\tau(c) - \int_0^1 \alpha_2(c,\eta)\,\mathrm{d}G(\eta) - \tau(c)\right] \mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_0^{\underline{\eta}} \nu(c,\eta)\;[-\alpha_2(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_{\underline{\eta}}^{1} \mu(c,\eta) \left[\alpha_2(c,\eta) - \lambda_0\,\tau(c) - \lambda_1 \left[\Pi(c_H) + \int_c^{c_H} \tau(y)\,\mathrm{d}y + c\,\tau(c) - \int_0^1 \alpha_2(c,z)\,\mathrm{d}G(z)\right]\right] \mathrm{d}\eta\,\mathrm{d}c \\
= &\left[-\kappa + \int_{c_L}^{c_H} \psi(c)\,\mathrm{d}c - \lambda_1 \int_{c_L}^{c_H}\int_{\underline{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\right] \Pi(c_H) \\
&+ \int_{\mathcal{C}_1} \left[\sigma(c) + \int_{c_L}^{c} \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) - (\lambda_0 + \lambda_1\, c) \int_{\underline{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1 \int_{c_L}^{c}\int_{\underline{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right] \tau(c)\,\mathrm{d}c \\
&+ \int_{\mathcal{C}_2} \left[-\beta(c) + \int_{c_L}^{c} \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) - (\lambda_0 + \lambda_1\, c) \int_{\underline{\eta}}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \lambda_1 \int_{c_L}^{c}\int_{\underline{\eta}}^{1} \mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right] \tau(c)\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_0^{\underline{\eta}} \left[-\psi(c)\, g(\eta) - \nu(c,\eta) + \lambda_1\, g(\eta) \int_0^1 \mu(c,z)\,\mathrm{d}z\right] \alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_{\underline{\eta}}^{1} \left[-\psi(c)\, g(\eta) + \mu(c,\eta) + \lambda_1\, g(\eta) \int_0^1 \mu(c,z)\,\mathrm{d}z\right] \alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
= &\, v_E\,\Pi(c_H) + \int_{c_L}^{c_H} \int_0^1 [v(\eta) - v_E]\;\alpha_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) + v_E \int_{c_L}^{c_H} \left[c + \frac{F(c)}{f(c)}\right] \tau(c)\, f(c)\;\mathrm{d}c.
\end{aligned}$$

Therefore, the objective function is bounded by

$$\int_{\mathcal{C}_1} \sigma(c)\,\mathrm{d}c.$$

Substituting our candidate solution into the objective function achieves the upper bound in both cases, which follows from

1. if $\bar{c}(\underline{\eta}) < c_L$, we have $c^* = c_L$ and the measure of $\mathcal{C}_1$ is 0. The government's expected payoff generated by our candidate policy and its upper bound are both 0;

2. if $c_L \leq \bar{c}(\underline{\eta}) \leq c_H$, we have

$$\begin{aligned}
&\int_{c_L}^{\bar{c}(\underline{\eta})} \sigma(c)\,\mathrm{d}c - \left[v_E + (\lambda_0+\lambda_1)\int_{\underline{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] F(\bar{c}(\underline{\eta})) \\
&= \int_{c_L}^{\bar{c}(\underline{\eta})} \left[v(\underline{\eta})\,\left[(c-\bar{c}(\underline{\eta}))\,f(c)+F(c)\right] + f(c)\,\left[v_E + (\lambda_0+\lambda_1)\int_{\underline{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta)\right]\right]\mathrm{d}c \\
&\quad - \left[v_E + (\lambda_0+\lambda_1)\int_{\underline{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] F(\bar{c}(\underline{\eta})) \\
&= v(\underline{\eta})\int_{c_L}^{\bar{c}(\underline{\eta})} \left[(c-\bar{c}(\underline{\eta}))\,f(c)+F(c)\right]\mathrm{d}c \\
&= 0.
\end{aligned}$$

To prove the optimality in case (3)(c), define

$$\begin{aligned}
&\kappa = -v(\eta^*), \\
&\nu(c,\eta) = [v(\eta^*) - v(\eta)]\; f(c)\,g(\eta), && \forall (c,\eta) \in [c_L, c_H]\times[0,\eta^*], \\
&\mu(c,\eta) = [v(\eta) - v(\eta^*)]\; f(c)\,g(\eta), && \forall (c,\eta) \in [c_L, c_H]\times[\eta^*,1], \\
&\psi(c) = \lambda_1 \int_{\eta^*}^{1} \mu(c,y)\,\mathrm{d}y - [v(\eta^*) - v_E]\; f(c), && \forall c \in [c_L, c_H], \\
&\sigma(c) = v_E\,[c\,f(c)+F(c)] + (\lambda_0 + \lambda_1\,c)\int_{\eta^*}^{1}\mu(c,\eta)\,\mathrm{d}\eta \\
&\qquad + \lambda_1 \int_{c_L}^{c}\int_{\eta^*}^{1}\mu(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - \int_{c_L}^{c}\psi(y)\,\mathrm{d}y - (c-1)\,\psi(c), && \forall c \in [c_L, c_H].
\end{aligned}$$

The monotonicity of $\bar{c}$ implies that $\underline{\eta} \leq \eta^* < \bar{\eta}$. The nonnegativity of κ, ν, and μ follows from their definitions directly.

When $h(\bar{\eta}) > 0$, we have

$$\psi(c) = h(\eta^*)\, f(c) \geq h(\bar{\eta})\, f(c) > 0.$$

When $\check{c} > \bar{c}(\check{\eta})$, by Lemma 2, we know that ψ is nonnegative if and only if $\eta^* \leq \check{\eta}$, which follows from

$$\bar{c}(\check{\eta}) < \check{c} \leq c_H \quad \Rightarrow \quad \check{\eta} > \eta^*.$$

Similar to the previous case, the expression of σ can be simplified as

$$\sigma(c) = v(\eta^*)\, \left[(c - c_H)\, f(c) + F(c)\right] + f(c)\, \left[v_E + (\lambda_0 + \lambda_1) \int_{\eta^*}^{1} v(y)\, \mathrm{d}G(y)\right], \quad \forall c \in [c_L, c_H].$$

The monotonicity of σ still holds, and we have

$$\begin{aligned}
\sigma(c_H) &= v(\eta^*) + f(c_H)\, \left[v_E + (\lambda_0 + \lambda_1) \int_{\eta^*}^{1} v(y)\, \mathrm{d}G(y)\right] \\
&\geq \left[v(\underline{\eta})\, \frac{F(\bar{c}(\underline{\eta}))}{f(\bar{c}(\underline{\eta}))} + \left[v_E + (\lambda_0 + \lambda_1) \int_{\underline{\eta}}^{1} v(y)\, \mathrm{d}G(y)\right]\right]\, f(c_H) \\
&\geq 0,
\end{aligned}$$

implying the nonnegativity of σ. Because

$$\begin{aligned}
&\kappa\, [-\Pi(c_H)] + \int_{c_L}^{c_H} \sigma(c)\, \tau(c)\, \mathrm{d}c \\
&+ \int_{c_L}^{c_H} \psi(c)\, \left[\Pi(c_H) + \int_{c}^{c_H} \tau(y)\, \mathrm{d}y + c\, \tau(c) - \int_{0}^{1} \alpha_2(c, \eta)\, \mathrm{d}G(\eta) - \tau(c)\right] \mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_{0}^{\eta^*} \nu(c, \eta)\, \left[-\alpha_2(c, \eta)\right] \mathrm{d}\eta\, \mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_{\eta^*}^{1} \mu(c, \eta)\, \left[\alpha_2(c, \eta) - \lambda_0\, \tau(c) - \lambda_1\, \left[\Pi(c_H) + \int_{c}^{c_H} \tau(y)\, \mathrm{d}y + c\, \tau(c) - \int_{0}^{1} \alpha_2(c, z)\, \mathrm{d}G(z)\right]\right] \mathrm{d}\eta\, \mathrm{d}c \\
= &\left[-\kappa + \int_{c_L}^{c_H} \psi(c)\, \mathrm{d}c - \lambda_1 \int_{c_L}^{c_H} \int_{\eta^*}^{1} \mu(c, \eta)\, \mathrm{d}\eta\, \mathrm{d}c\right]\, \Pi(c_H) \\
&+ \int_{c_L}^{c_H} \left[\sigma(c) + \int_{c_L}^{c} \psi(y)\, \mathrm{d}y + (c - 1)\, \psi(c) - (\lambda_0 + \lambda_1\, c) \int_{\eta^*}^{1} \mu(c, \eta)\, \mathrm{d}\eta - \lambda_1 \int_{c_L}^{c} \int_{\eta^*}^{1} \mu(y, \eta)\, \mathrm{d}\eta\, \mathrm{d}y\right]\, \tau(c)\, \mathrm{d}c
\end{aligned}$$

$$
\begin{aligned}
&+ \int_{c_L}^{c_H} \int_{0}^{\eta^*} \left[-\psi(c)\, g(\eta) - \nu(c,\eta) + \lambda_1\, g(\eta) \int_{\eta^*}^{1} \mu(c,z)\, \mathrm{d}z \right] \alpha_2(c,\eta)\, \mathrm{d}\eta\, \mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_{\eta^*}^{1} \left[-\psi(c)\, g(\eta) + \mu(c,\eta) + \lambda_1\, g(\eta) \int_{\eta^*}^{1} \mu(c,z)\, \mathrm{d}z \right] \alpha_2(c,\eta)\, \mathrm{d}\eta\, \mathrm{d}c \\
&= v_E\, \Pi(c_H) + \int_{c_L}^{c_H} \int_{0}^{1} \left[v(\eta) - v_E \right] \alpha_2(c,\eta)\, \mathrm{d}G(\eta)\, \mathrm{d}F(c) + v_E \int_{c_L}^{c_H} \left[c + \frac{F(c)}{f(c)} \right] \tau(c)\, f(c)\ \mathrm{d}c,
\end{aligned}
$$

the objective function is bounded by

$$
\int_{c_L}^{c_H} \sigma(c)\, \mathrm{d}c.
$$

Substituting our candidate solution into the objective function achieves the upper bound, which follows from

$$
\begin{aligned}
&\int_{c_L}^{c_H} \sigma(c)\, \mathrm{d}c - \left[v_E + (\lambda_0 + \lambda_1) \int_{\eta^*}^{1} v(\eta)\, \mathrm{d}G(\eta) \right] \\
&= \int_{c_L}^{c_H} \left[v(\eta^*)\ \left[(c - c_H)\, f(c) + F(c) \right] + f(c) \left[v_E + (\lambda_0 + \lambda_1) \int_{\eta^*}^{1} v(\eta)\, \mathrm{d}G(\eta) \right] \right] \mathrm{d}c \\
&\quad - \left[v_E + (\lambda_0 + \lambda_1) \int_{\eta^*}^{1} v(\eta)\, \mathrm{d}G(\eta) \right] \\
&= v(\eta^*) \int_{c_L}^{c_H} \left[(c - c_H)\, f(c) + F(c) \right] \mathrm{d}c \\
&= 0.
\end{aligned}
$$

This completes the proof. □

Proof. Proof of Proposition 4. When the testing cost is high and the government induces testing, the firm's individual rationality constraint (IR) must be binding. Otherwise, the government could strictly increase its expected payoff by reducing the costly incentives offered to the firm. That is, the government's first-best optimization problem can be written as follows:

$$
\begin{aligned}
&\max_{\eta}\ v_E\, \tilde{a}_1(\eta, c) + (\lambda_0 + \lambda_1\, \tilde{a}_1(\eta, c)) \int_{\eta}^{1} v(y)\, \mathrm{d}G(y), \\
&\text{subject to}\ \ 0 \leq \tilde{a}_1(\eta, c) \leq 1,
\end{aligned}
$$

where $\tilde{a}_1$ is defined by (31). We refer to the above optimization problem without the feasi-

bility constraint on a_1 as the "relaxed problem".

Since the firm's revenue across both periods is bounded above by $\bar{c}(0) = 1 + \lambda_0 + \lambda_1$, the government cannot induce testing when the testing cost exceeds this value. In the remainder of the proof, we focus on the cases where $c \leq \bar{c}(0)$.

As established in Section 4.1, the derivative of the objective function with respect to η has the same sign as the function h defined in (33). If the drug is promising ($h(\bar{\eta}) \geq 0$) and $c > 1 + (\lambda_0 + \lambda_1)\,[1 - G(\bar{\eta})]$, since $\tilde{a}_1(\bar{\eta}, c) > 1$, the optimal solution to the relaxed problem is infeasible. Because η is increasing in a_1, it is optimal to set $a_1 = 1$, and the corresponding value of η is determined such that the individual rationality constraint (IR) binds.

Next, we consider the case in which the drug is not promising ($h(\bar{\eta}) \leq 0$) and $c > \lambda_0\,[1 - G(\bar{\eta})]$. Recall that $\check{\eta}$ is the unique root of the function h on $(0, \bar{\eta}]$. If $\tilde{a}_1(\check{\eta}, c) \in [0, 1]$, then the optimal solution to the relaxed problem is feasible, and hence also optimal for the original problem. Instead, if $\tilde{a}_1(\check{\eta}, c) < 0$, then any feasible value of η must lie above $\check{\eta}$. Since η is increasing in a_1 and the objective function is decreasing in η for all $\eta \in [\check{\eta}, \bar{\eta}]$, it is optimal to set $a_1 = 0$, which uniquely determines the final approval threshold $\eta^* = G^{-1}\left(1 - \frac{c}{\lambda_0}\right)$. Finally, if $\tilde{a}_1(\check{\eta}, c) > 1$, then any feasible value of η must lie below $\check{\eta}$. Since η is increasing in a_1 and the objective function is increasing in η for all $\eta \in [0, \check{\eta}]$, it is optimal to set $a_1 = 1$, and the corresponding final approval threshold is given by $\eta^* = G^{-1}\left(1 - \frac{c-1}{\lambda_0+\lambda_1}\right)$.

When the government's expected payoff from inducing testing ($\tau = 1$) is negative, it is optimal to forgo testing by setting $\tau = 0$.

This completes the proof. □

11.4 Proofs for Section 5

Proof. Proof of Proposition 5. To prove the desired result, it suffices to show that for any feasible solution to (13), we can construct a feasible solution to (48) that yields the same objective function value.

Let $\mathcal{M} \equiv \{\tau, a_1, a_2\}$ denote a feasible solution to the optimization problem (13). By construction, it is straightforward to verify that $\bar{\mathcal{M}} \equiv \{\tau, \alpha_1, \alpha_2, \alpha_3\}$, where

$$\begin{aligned}
&\alpha_1(c) = \tau(c)\, a_1(c), && \forall c \in [c_L, c_H],\\
&\alpha_2(c, \eta) = \tau(c)\, a_2(c, \eta), && \forall (c, \eta) \in [c_L, c_H] \times [0, 1],\\
&\alpha_3(c, \eta) = \tau(c)\, a_1(c)\, a_2(c, \eta), && \forall (c, \eta) \in [c_L, c_H] \times [0, 1],
\end{aligned}$$

is feasible to the upper-bound problem (48) and yields the same objective function value.

This completes the proof. □

Similar to Lemma 3, after eliminating α_1 using the envelope theorem, the government's upper-bound problem can be formulated as follows:

$$\begin{aligned}\max_{\{\tau,\hat{\alpha}_2,\hat{\alpha}_3,\Pi(c_H)\}\in\Omega'_R} \quad & v_E\,\Pi(c_H)+v_E\int_{c_L}^{c_H}\left[c+\frac{F(c)}{f(c)}\right]\tau(c)\,f(c)\,\mathrm{d}c \\ & +\int_{c_L}^{c_H}\int_0^1\left[\lambda_0\,[v(\eta)-v_E]+\Delta\,[v(\eta)+1]\right]\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) \\ & +\lambda_1\int_{c_L}^{c_H}\int_0^1[v(\eta)-v_E]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c),\end{aligned} \tag{68}$$

where the feasible set Ω'_R is determined by (2), (45), (47), (62), (64), and for any $(c,\eta)\in[c_L,c_H]\times[0,1]$,

$$0\le\Pi(c_H)+\int_c^{c_H}\tau(y)\,\mathrm{d}y+c\,\tau(c)-\lambda_0\int_0^1\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta)-\lambda_1\int_0^1\hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta)\le\tau(c), \tag{69}$$

and

$$\hat{\alpha}_3(c,\eta)\le\Pi(c_H)+\int_c^{c_H}\tau(y)\,\mathrm{d}y+c\,\tau(c)-\lambda_0\int_0^1\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta)-\lambda_1\int_0^1\hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta). \tag{70}$$

With a slight abuse of notation, in this subsection, let

$$\bar{c}(\eta_1,\eta_2)\equiv 1+\lambda_0\,[1-G(\eta_1)]+\lambda_1\,[1-G(\eta_2)]\,,\quad\forall\eta_1,\eta_2\in[0,1]. \tag{71}$$

By Lemma 2(1), we know that there exists a unique $\check{\eta}_2\in[0,\bar{\eta}]$ such that $h(\check{\eta}_2)=0$ if $h(\bar{\eta})\le 0$. Let

$$\Psi_1(\eta_1,\eta_2)\equiv(\lambda_0+\Delta)\,v(\eta_1)+\Delta-\lambda_0\,v(\eta_2) \tag{72}$$

and

$$\Psi_2(\eta_1,\eta_2)\equiv(\lambda_0+\lambda_1+\Delta)\,[v(\eta_2)-v(\eta_1)]+\frac{\lambda_1}{\lambda_0}\,\Delta\,[v(\eta_2)+1]\,. \tag{73}$$

In order to prove Proposition 6, we need the following technical results.

Lemma 5. *When $h(\bar{\eta})\le 0$, we have the following results.*

(1) If $v(\check{\eta}_2)+1\ge 0$, there exists a unique $\check{\eta}_1\in[0,\check{\eta}_2]$ such that $\Psi_1(\check{\eta}_1,\check{\eta}_2)=0$.

(2) If $v(\check{\eta}_2)+1<0$, there uniquely exist $(\mathring{\eta}_1,\mathring{\eta}_2)$ such that $0\le\mathring{\eta}_1\le\mathring{\eta}_2\le\tilde{\eta}$, $\Psi_2(\mathring{\eta}_1,\mathring{\eta}_2)=0$,

and $\iota(\mathring{\eta}_1, \mathring{\eta}_2) = 0$*, where*

$$\eta^\dagger \equiv \inf \{\eta \in [0,1] : v(\eta) + 1 \geq 0\} \tag{74}$$

and

$$\begin{aligned}\iota(\eta_1, \eta_2) \equiv\ & (\lambda_0 + \lambda_1 + \Delta) \int_{\eta_1}^{\eta_2} [v(\eta_2) - v(z)]\ g(z)\,\mathrm{d}z - \lambda_1 \int_{\eta_2}^{1} [v(z) - v(\eta_2)]\ g(z)\,\mathrm{d}z \\ & + \frac{\Delta}{\lambda_0}\ [v(\eta_2) + 1]\ [1 + \lambda_1\ [1 - G(\eta_1)]] + v(\eta_2) - v_E. \end{aligned} \tag{75}$$

Proof. Proof of Lemma 5. When $\check{\eta}_2$ exists and $v(\check{\eta}_2) + 1 \geq 0$, we have

$$\Psi_1(0, \check{\eta}_2) = \lambda_0\ [v(0) - v(\check{\eta}_2)] + \Delta\ [v(0) + 1] \leq 0,$$

and

$$\Psi_1(\check{\eta}_2, \check{\eta}_2) = \Delta\ [v(\check{\eta}_2) + 1] \geq 0.$$

The unique existence of $\check{\eta}_1$ follows directly from the monotonicity of Ψ_1 in η_1.

If $v(\check{\eta}_2) + 1 < 0$, define

$$\begin{aligned}\Xi(\eta_2) \equiv\ & -\lambda_0\, \iota(\eta_1(\eta_2), \eta_2) \\ =\ & \lambda_0 \left[-(\lambda_0 + \lambda_1 + \Delta) \int_{\eta_1(\eta_2)}^{\eta_2} [v(\eta_2) - v(z)]\ g(z)\,\mathrm{d}z + \lambda_1 \int_{\eta_2}^{1} [v(z) - v(\eta_2)]\ g(z)\,\mathrm{d}z\right] \\ & - \Delta\ [v(\eta_2) + 1]\ [1 + \lambda_1\ [1 - G(\eta_1(\eta_2))]] - \lambda_0\ [v(\eta_2) - v_E]\,, \end{aligned} \tag{76}$$

where $\eta_1(\eta_2)$ is determined by

$$\Psi_2(\eta_1(\eta_2), \eta_2) = 0 \quad \Leftrightarrow \quad v(\eta_1(\eta_2)) = v(\eta_2) + \frac{\lambda_1\, \Delta\ [v(\eta_2) + 1]}{\lambda_0\, (\lambda_0 + \lambda_1 + \Delta)}.$$

Note that $\eta_1(\eta_2)$ is increasing.

The monotonicity of v implies that $\check{\eta}_2 < \eta^\dagger$. We have $\eta_1(\eta^\dagger) = \eta^\dagger$, implying that

$$\begin{aligned}\Xi(\eta^\dagger) &\equiv \lambda_0 \left[\lambda_1 \int_{\eta^\dagger}^{1} \left[v(z) - v(\eta^\dagger)\right]\ g(z)\,\mathrm{d}z - v(\eta^\dagger) + v_E\right] \\ &\leq \lambda_0 \left[\lambda_1 \int_{\check{\eta}_2}^{1} [v(z) - v(\check{\eta}_2)]\ g(z)\,\mathrm{d}z - v(\check{\eta}_2) + v_E\right] \end{aligned}$$

$$= 0,$$

where the inequality follows from the monotonicity of the function h. To prove the desired result, it suffices to show that $\Xi(\check{\eta}_2) \geq 0$ and Ξ is non-increasing.

Taking derivative yields

$$\begin{aligned}
\Xi'(\eta_2) = & -\lambda_0 (\lambda_0 + \lambda_1 + \Delta) \int_{\eta_1(\eta_2)}^{\eta_2} v'(\eta_2)\, g(z)\, \mathrm{d}z + \lambda_0 (\lambda_0 + \lambda_1 + \Delta)\, [v(\eta_2) - v(\eta_1(\eta_2))]\, g(\eta_1(\eta_2))\, \eta_1'(\eta_2) \\
& - \lambda_0 \lambda_1 \int_{\eta_2}^{1} v'(\eta_2)\, g(z)\, \mathrm{d}z - \Delta\, v'(\eta_2)\, [1 + \lambda_1\, (1 - G(\eta_1(\eta_2)))] \\
& + \Delta\, [v(\eta_2) + 1]\, \lambda_1\, g(\eta_1(\eta_2))\, \eta_1'(\eta_2) - \lambda_0\, v'(\eta_2) \\
= & \Big\{ -\lambda_0 (\lambda_0 + \lambda_1 + \Delta)\, [G(\eta_2) - G(\eta_1(\eta_2))] - \lambda_0 \lambda_1\, [1 - G(\eta_2)] \\
& \quad -\lambda_0 - \Delta\, [1 + \lambda_1\, (1 - G(\eta_1(\eta_2)))] \Big\}\, v'(\eta_2) \\
\leq & \, 0,
\end{aligned}$$

where the second equality follows from the definition of $\eta_1(\eta_2)$, i.e., $\Psi_2(\eta_1(\eta_2), \eta_2) = 0$.

Since $h(\check{\eta}_2) = 0$, the desired condition $\Xi(\check{\eta}_2) \geq 0$ reduces to $\tilde{\Xi}(\check{\eta}_2) \geq 0$, where

$$\tilde{\Xi}(\eta) \equiv -\lambda_0 (\lambda_0 + \lambda_1 + \Delta) \int_{\eta_1(\eta)}^{\eta} [v(\eta) - v(z)]\, g(z)\, \mathrm{d}z - \Delta\, [v(\eta) + 1]\, [1 + \lambda_1\, [1 - G(\eta_1(\eta))]]\,.$$

Taking derivative yields

$$\begin{aligned}
\tilde{\Xi}'(\eta) = & -\lambda_0 (\lambda_0 + \lambda_1 + \Delta) \left[\int_{\eta_1(\eta)}^{\eta} v'(\eta)\, g(z)\, \mathrm{d}z - [v(\eta) - v(\eta_1(\eta))]\, g(\eta_1(\eta))\, \eta_1'(\eta) \right] \\
& - \Delta\, v'(\eta)\, [1 + \lambda_1\, [1 - G(\eta_1(\eta))]] + \Delta\, [v(\eta) + 1]\, \lambda_1\, g(\eta_1(\eta))\, \eta_1'(\eta) \\
= & -\{\lambda_0 (\lambda_0 + \lambda_1 + \Delta)\, [G(\eta) - G(\eta_1(\eta))] + \Delta\, [1 + \lambda_1\, [1 - G(\eta_1(\eta))]]\}\, v'(\eta) \\
\leq & \, 0.
\end{aligned}$$

Therefore, the desired result is implied by

$$\tilde{\Xi}(\check{\eta}_2) \geq \tilde{\Xi}(\eta^\dagger) = 0.$$

This completes the proof. □

Let

$$\check{c}_1 \equiv \sup\left\{c \in [c_L, c_H] : (\lambda_0 + \Delta) \int_{\check{\eta}_1}^{1} [v(\eta) - v(\check{\eta}_1)]\ g(\eta)\,\mathrm{d}\eta + v(\check{\eta}_2)\ \left[c + \frac{F(c)}{f(c)}\right] \geq 0\right\} \quad (77)$$

and

$$\check{c}_2 \equiv \sup\left\{c \in [c_L, c_H] : (\lambda_0 + \Delta) \int_{\mathring{\eta}_2}^{1} [v(\eta) - v(\mathring{\eta}_2)]\ g(\eta)\,\mathrm{d}\eta + \left[c + \frac{F(c)}{f(c)}\right]\ \left[v(\mathring{\eta}_2) + \frac{\Delta}{\lambda_0}\ [v(\mathring{\eta}_2) + 1]\right] \geq 0\right\}. \quad (78)$$

By Lemma 5, we know that $\check{c}_1$ and $\check{c}_2$ are well-defined when $v(\check{\eta}_2) + 1 \geq 0$ and $v(\check{\eta}_2) + 1 < 0$, respectively, with the existence of $\check{\eta}_2$ requiring $h(\bar{\eta}) \leq 0$.

Lemma 6. *When $h(\bar{\eta}) \leq 0$, we have the following results.*

(1) If $\check{c}_2 < \underline{c}(\mathring{\eta}_2)$, there exists a unique $\tilde{\eta}_2 \in [\mathring{\eta}_2, \bar{\eta}']$ such that

$$(\lambda_0 + \Delta) \int_{\tilde{\eta}_2}^{1} [v(\eta) - v(\tilde{\eta}_2)]\ g(\eta)\,\mathrm{d}\eta + \left[\underline{c}(\tilde{\eta}_2) + \frac{F(\underline{c}(\tilde{\eta}_2))}{f(\underline{c}(\tilde{\eta}_2))}\right]\ \left[v(\tilde{\eta}_2) + \frac{\Delta}{\lambda_0}\ [v(\tilde{\eta}_2) + 1]\right] = 0, \quad (79)$$

where $\bar{\eta}'$ is defined by (55). If $v(\tilde{\eta}_2) + 1 \leq 0$, there exists $\tilde{\eta}_1 \leq \tilde{\eta}_2$ such that $\Psi_2(\tilde{\eta}_1, \tilde{\eta}_2) = 0$. In this case, we have $\iota(\tilde{\eta}_1, \tilde{\eta}_2) \geq 0$, where ι is defined by (75).

(2) If (i) $v(\tilde{\eta}_2) + 1 > 0$, there exists a unique $\acute{\eta}_1 \in [\mathring{\eta}_2, \bar{\eta}']$, or if (ii) $\check{c}_1 < \underline{c}(\check{\eta}_1)$, there exists a unique $\acute{\eta}_1 \in [\check{\eta}_1, \bar{\eta}']$, such that

$$v(\acute{\eta}_2)\ \left[\underline{c}(\acute{\eta}_1) + \frac{F(\underline{c}(\acute{\eta}_1))}{f(\underline{c}(\acute{\eta}_1))}\right] + (\lambda_0 + \Delta) \int_{\acute{\eta}_1}^{1} [v(\eta) - v(\acute{\eta}_1)]\ g(\eta)\,\mathrm{d}\eta = 0 \quad (80)$$

and $\Psi_1(\acute{\eta}_1, \acute{\eta}_2) = 0$, respectively. Moreover, we have $h(\acute{\eta}_2) \leq 0$.

Proof. Proof of Lemma 6. To prove the first part, define

$$h_2(\eta) \equiv (\lambda_0 + \Delta) \int_{\eta}^{1} [v(z) - v(\eta)]\ g(z)\,\mathrm{d}z + \left[\underline{c}(\eta) + \frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))}\right]\ \left[v(\eta) + \frac{\Delta}{\lambda_0}\ [v(\eta) + 1]\right]. \quad (81)$$

Taking derivative yields

$$\begin{aligned} h_2'(\eta) = & -(\lambda_0 + \Delta)\, v'(\eta)\ [1 - G(\eta)] + \frac{\mathrm{d}}{\mathrm{d}\eta}\left[\underline{c}(\eta) + \frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))}\right]\ \left[v(\eta) + \frac{\Delta}{\lambda_0}\ [v(\eta) + 1]\right] \\ & + \left[\underline{c}(\eta) + \frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))}\right]\ \left(\frac{\Delta}{\lambda_0} + 1\right)\ v'(\eta) \end{aligned}$$

$$= \underbrace{\frac{\mathrm{d}}{\mathrm{d}\eta}\left[\underline{c}(\eta) + \frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))}\right]}_{\leq 0}\left[v(\eta) + \frac{\Delta}{\lambda_0}\left[v(\eta)+1\right]\right] + \left(\frac{\Delta}{\lambda_0}+1\right)\frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))}\,v'(\eta),$$

which is nonnegative if $\eta \leq \bar{\eta}'$. We have

$$h_2(\bar{\eta}') = (\lambda_0+\Delta)\int_{\bar{\eta}'}^{1}\left[v(z)-v(\bar{\eta}')\right]\,g(z)\,\mathrm{d}z \geq 0$$

and

$$h_2(\mathring{\eta}_2) = (\lambda_0+\Delta)\int_{\mathring{\eta}_2}^{1}\left[v(z)-v(\mathring{\eta}_2)\right]\,g(z)\,\mathrm{d}z + \left[\underline{c}(\mathring{\eta}_2) + \frac{F(\underline{c}(\mathring{\eta}_2))}{f(\underline{c}(\mathring{\eta}_2))}\right]\left[v(\mathring{\eta}_2) + \frac{\Delta}{\lambda_0}\left[v(\mathring{\eta}_2)+1\right]\right] < 0,$$

where the inequality follows from $\check{c}_2 < \underline{c}(\mathring{\eta}_2)$, implying the unique existence of $\tilde{\eta}_2 \in [\mathring{\eta}_2, \bar{\eta}']$. The existence of $\tilde{\eta}_1 \in [0, \tilde{\eta}_2]$ is ensured by $v(\tilde{\eta}_2)+1 \leq 0$. It remains to verify that $\iota(\tilde{\eta}_1, \tilde{\eta}_2) \geq 0$, which follows from

$$\iota(\tilde{\eta}_1, \tilde{\eta}_2) = -\frac{\Xi(\tilde{\eta}_2)}{\lambda_0} \geq -\frac{\Xi(\mathring{\eta}_2)}{\lambda_0} = 0.$$

To prove the second part, we first verify that $v_E \leq 0$, which follows from

$$h(\bar{\eta}) \leq 0 \quad \Rightarrow \quad \lambda_1 \underbrace{\int_{\bar{\eta}}^{1} v(y)\,\mathrm{d}G(y)}_{>0} + v_E \leq 0 \quad \Rightarrow \quad v_E < 0.$$

Define

$$h_1(\eta) \equiv \frac{1}{\lambda_0}\left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\left[\underline{c}(\eta) + \frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))}\right] + (\lambda_0+\Delta)\int_{\eta}^{1}\left[v(z)-v(\eta)\right]\,g(z)\,\mathrm{d}z. \quad (82)$$

Taking derivative yields

$$\begin{aligned} h_1'(\eta) = {} & \frac{1}{\lambda_0}(\lambda_0+\Delta)\,v'(\eta)\left[\underline{c}(\eta) + \frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))}\right] + \frac{1}{\lambda_0}\left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\underbrace{\frac{\mathrm{d}}{\mathrm{d}\eta}\left[\underline{c}(\eta) + \frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))}\right]}_{\leq 0} \\ & - (\lambda_0+\Delta)\,v'(\eta)\left[1-G(\eta)\right] \\ = {} & \frac{1}{\lambda_0}\left\{(\lambda_0+\Delta)\,v'(\eta)\frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))} + \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\frac{\mathrm{d}}{\mathrm{d}\eta}\left[\underline{c}(\eta) + \frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))}\right]\right\}. \end{aligned}$$

We know that function h_1 is increasing for any $\eta \in [0, \bar{\eta}']$.

To prove the unique existence of $\acute{\eta}_1$, it suffices to show that, in cases (i) and (ii), we

have $h_1(\mathring{\eta}_2) \le 0 \le h_1(\bar{\eta}')$ and $h_1(\check{\eta}_1) \le 0 \le h_1(\bar{\eta}')$, respectively. The nonnegative condition follows from

$$h_1(\bar{\eta}') = (\lambda_0 + \Delta) \int_{\bar{\eta}'}^1 [v(z) - v(\bar{\eta}')]\ g(z)\,\mathrm{d}z \ge 0.$$

It remains to prove the other desired inequality condition.

- When $v(\tilde{\eta}_2) + 1 > 0$, we have

$$\begin{aligned} h_1(\mathring{\eta}_2) = &\left[v(\mathring{\eta}_2) + \frac{\Delta}{\lambda_0}\ [v(\mathring{\eta}_2) + 1]\right] \left[\underline{c}(\mathring{\eta}_2) + \frac{F(\underline{c}(\mathring{\eta}_2))}{f(\underline{c}(\mathring{\eta}_2))}\right] + (\lambda_0 + \Delta) \int_{\mathring{\eta}_2}^1 [v(z) - v(\mathring{\eta}_2)]\ g(z)\,\mathrm{d}z \\ &< 0, \end{aligned}$$

  where the inequality follows from $\check{c}_2 < \underline{c}(\mathring{\eta}_2)$.

- When $v(\check{\eta}_2) + 1 \ge 0$ and $\check{c}_1 < \underline{c}(\check{\eta}_1)$, we have

$$\begin{aligned} h_1(\check{\eta}_1) = &\left[v(\check{\eta}_1) + \frac{\Delta}{\lambda_0}\ [v(\check{\eta}_1) + 1]\right] \left[\underline{c}(\check{\eta}_1) + \frac{F(\underline{c}(\check{\eta}_1))}{f(\underline{c}(\check{\eta}_1))}\right] + (\lambda_0 + \Delta) \int_{\check{\eta}_1}^1 [v(z) - v(\check{\eta}_1)]\ g(z)\,\mathrm{d}z \\ &< \left[v(\check{\eta}_1) + \frac{\Delta}{\lambda_0}\ [v(\check{\eta}_1) + 1]\right] \left[\underline{c}(\check{\eta}_1) + \frac{F(\underline{c}(\check{\eta}_1))}{f(\underline{c}(\check{\eta}_1))}\right] - v(\check{\eta}_2) \left[\underline{c}(\check{\eta}_1) + \frac{F(\underline{c}(\check{\eta}_1))}{f(\underline{c}(\check{\eta}_1))}\right] \\ &= 0, \end{aligned}$$

  where the inequality follows from $\check{c}_1 < \underline{c}(\check{\eta}_1)$.

Therefore, the given conditions guarantee the unique existence of $\acute{\eta}_1$.

It remains to verify that $h(\acute{\eta}_2) \le 0$, which follows from

$$\begin{aligned} &(i)\ \acute{\eta}_2 \ge \acute{\eta}_1 \ge \mathring{\eta}_2 \ge \check{\eta}_2 \quad \Rightarrow \quad h(\acute{\eta}_2) \le 0, \text{ and} \\ &(ii)\ \acute{\eta}_1 \ge \check{\eta}_1 \quad \Rightarrow \quad \acute{\eta}_2 \ge \check{\eta}_2 \quad \Rightarrow \quad h(\acute{\eta}_2) \le 0, \end{aligned}$$

where $\mathring{\eta}_2 \ge \check{\eta}_2$ follows from the proof of Lemma 5.

This completes the proof. □

Lemma 7. *If $\check{c}_2 > \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1)$, we have $t_1(\eta^\dagger) \ge 0$, where $\eta^\dagger$ is defined by (74),*

$$\begin{aligned} t_1(\eta_2) \equiv\ & v(\eta_2) \left[\bar{c}(\eta_1(\eta_2), \eta_2) + \frac{F(\bar{c}(\eta_1(\eta_2), \eta_2))}{f(\bar{c}(\eta_1(\eta_2), \eta_2))}\right] + (\lambda_0 + \Delta) \int_{\eta_1(\eta_2)}^1 [v(z) - v(\eta_1(\eta_2))]\ g(z)\,\mathrm{d}z \\ &+ \lambda_1 \int_{\eta_2}^1 [v(z) - v(\eta_2)]\ g(z)\,\mathrm{d}z - v(\eta_2) + v_E \end{aligned} \tag{83}$$

and function $\eta_1(\eta_2)$ is determined by $\Psi_1(\eta_1(\eta_2),\eta_2)=0$.

Proof. Proof of Lemma 7. The given condition $\check{c}_2 > \bar{c}(\mathring{\eta}_1,\mathring{\eta}_1)$ can be rewritten as

$$(\lambda_0+\Delta)\int_{\mathring{\eta}_2}^{1}[v(\eta)-v(\mathring{\eta}_2)]\;g(\eta)\,\mathrm{d}\eta+\left[\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1)+\frac{F(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}{f(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}\right]\left[v(\mathring{\eta}_2)+\frac{\Delta}{\lambda_0}\;[v(\mathring{\eta}_2)+1]\right]>0.$$

Recall that the two thresholds $\mathring{\eta}_1$ and $\mathring{\eta}_2$ satisfy $\iota(\mathring{\eta}_1,\mathring{\eta}_2)=0$, we have

$$\begin{aligned}
&(\lambda_0+\Delta)\int_{\mathring{\eta}_2}^{1}[v(\eta)-v(\mathring{\eta}_2)]\;g(\eta)\,\mathrm{d}\eta+\left[\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1)+\frac{F(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}{f(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}\right]\left[v(\mathring{\eta}_2)+\frac{\Delta}{\lambda_0}\;[v(\mathring{\eta}_2)+1]\right]\\
&=(\lambda_0+\Delta)\int_{\mathring{\eta}_2}^{1}[v(\eta)-v(\mathring{\eta}_2)]\;g(\eta)\,\mathrm{d}\eta+\left[\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1)+\frac{F(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}{f(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}\right]\;v(\mathring{\eta}_2)\\
&\quad+\frac{\Delta}{\lambda_0}\left[\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1)+\frac{F(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}{f(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}\right]\;[v(\mathring{\eta}_2)+1]\\
&\le(\lambda_0+\Delta)\int_{\mathring{\eta}_2}^{1}[v(\eta)-v(\mathring{\eta}_2)]\;g(\eta)\,\mathrm{d}\eta+\left[\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1)+\frac{F(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}{f(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}\right]\;v(\mathring{\eta}_2)\\
&\quad+\frac{\Delta}{\lambda_0}\;[1+\lambda_1\;[1-G(\mathring{\eta}_1)]]\;[v(\mathring{\eta}_2)+1]+\Delta\;[1-G(\mathring{\eta}_1)]\;[v(\mathring{\eta}_2)+1]\\
&=(\lambda_0+\lambda_1+\Delta)\int_{\mathring{\eta}_1}^{1}[v(\eta)-v(\mathring{\eta}_2)]\;g(\eta)\,\mathrm{d}\eta+\left[\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1)+\frac{F(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}{f(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}\right]\;v(\mathring{\eta}_2)-v(\mathring{\eta}_2)+v_E\\
&\quad+\Delta\;[1-G(\mathring{\eta}_1)]\;[v(\mathring{\eta}_2)+1]\\
&\le(\lambda_0+\lambda_1+\Delta)\int_{\mathring{\eta}_2}^{1}[v(\eta)-v(\mathring{\eta}_2)]\;g(\eta)\,\mathrm{d}\eta+\left[\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1)+\frac{F(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}{f(\bar{c}(\mathring{\eta}_1,\mathring{\eta}_1))}\right]\;v(\mathring{\eta}_2)-v(\mathring{\eta}_2)+v_E\\
&\quad+\Delta\;[1-G(\mathring{\eta}_2)]\;[v(\mathring{\eta}_2)+1]\,,
\end{aligned}$$

where the first inequality and the second equality follow from $v(\mathring{\eta}_2)+1\le v(\eta^\dagger)+1=0$ and $\iota(\mathring{\eta}_1,\mathring{\eta}_2)=0$, respectively.

Define

$$\begin{aligned}
\Upsilon(\eta_2)\equiv{}&(\lambda_0+\lambda_1+\Delta)\int_{\eta_2}^{1}[v(\eta)-v(\eta_2)]\;g(\eta)\,\mathrm{d}\eta+\left[\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2))+\frac{F(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}\right]\;v(\eta_2)\\
&-v(\eta_2)+v_E+\Delta\;[1-G(\eta_2)]\;[v(\eta_2)+1]\,,\quad\forall\eta_2\in[0,\eta^\dagger],
\end{aligned}$$

where $\eta_1(\eta_2)$ is determined by $\Psi_2(\eta_1(\eta_2),\eta_2)=0$. Taking derivative yields

$$\Upsilon'(\eta_2)=-(\lambda_0+\lambda_1+\Delta)\,v'(\eta_2)\;[1-G(\eta_2)]+\left[\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2))+\frac{F(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}\right]\;v'(\eta_2)$$

$$
\begin{aligned}
&+ \frac{\mathrm{d}}{\mathrm{d}\eta_2}\left[\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)) + \frac{F(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}\right] v(\eta_2) - v'(\eta_2) \\
&+ \Delta\,\left[-\left[v(\eta_2)+1\right]\,g(\eta_2) + \left[1-G(\eta_2)\right]\,v'(\eta_2)\right] \\
&\geq \frac{\mathrm{d}}{\mathrm{d}\eta_2}\left[\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)) + \frac{F(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}\right] v(\eta_2) - \Delta\,\left[v(\eta_2)+1\right]\,g(\eta_2) + \frac{F(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}\,v' \\
&\geq 0,
\end{aligned}
$$

where the first inequality follows from $\eta_1(\eta_2) \leq \eta_2$ for any $\eta_2 \in [0, \eta^\dagger]$, as established in Lemma 5(2). Therefore, we have

$$
t_1(\eta^\dagger) = \Upsilon(\eta^\dagger) \geq \Upsilon(\mathring{\eta}_2) > 0,
$$

which completes the proof. □

Lemma 8. *When one of the following conditions holds, (i) $h(\bar{\eta}) \leq 0$, $\check{c}_2 > \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1)$, (ii) $h(\bar{\eta}) \leq 0$, $\check{c}_1 > \bar{c}(\check{\eta}_1, \check{\eta}_2)$, or (iii) $h(\bar{\eta}) > 0$, we have the following results.*

(1) If $t_1(\eta^\dagger) \leq 0$, there uniquely exist $\eta^\dagger \leq \dot{\eta}_1 \leq \dot{\eta}_2 \leq \bar{\eta}$ such that $t_1(\dot{\eta}_2) = 0$, where function t_1 is defined by (83), $\dot{\eta}_1 = \eta_1(\dot{\eta}_2)$, and function $\eta_1(\eta_2)$ is determined by $\Psi_1(\eta_1(\eta_2), \eta_2) = 0$. Moreover, we have $h(\dot{\eta}_2) \geq 0$.

(2) If $t_2(\breve{\eta}) \leq 0 \leq t_1(\eta^\dagger) = t_2(\eta^\dagger)$, there uniquely exist $\breve{\eta} \leq \grave{\eta}_1 \leq \grave{\eta}_2 \leq \eta^\dagger$ such that $t_2(\grave{\eta}_2) = 0$, where $\grave{\eta}_1 = \eta_1(\grave{\eta}_2)$, $\breve{\eta}$ is determined by $\eta_1(\breve{\eta}) = 0$,

$$
\begin{aligned}
t_2(\eta_2) \equiv\ & \left[v(\eta_2) + \frac{\Delta}{\lambda_0}\,\left[v(\eta_2)+1\right]\right]\,\left[\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)) + \frac{F(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}\right] \\
&+ (\lambda_0 + \Delta)\int_{\eta_2}^{1}\left[v(\eta) - v(\eta_2)\right]\,g(\eta)\,\mathrm{d}\eta - \iota(\eta_1(\eta_2),\eta_2),
\end{aligned} \tag{84}
$$

and function $\eta_1(\eta_2)$ is determined by $\Psi_2(\eta_1(\eta_2), \eta_2) = 0$. Moreover, we have $\iota(\grave{\eta}_1, \grave{\eta}_2) \leq 0$, where ι is defined by (75).

Proof. Proof of Lemma 8. To prove the result in case (1), taking derivative yields

$$
\begin{aligned}
t_1'(\eta_2) =\ & v'(\eta_2)\,\left[\bar{c}(\eta_1(\eta_2),\eta_2) + \frac{F(\bar{c}(\eta_1(\eta_2),\eta_2))}{f(\bar{c}(\eta_1(\eta_2),\eta_2))}\right] + v(\eta_2)\,\frac{\mathrm{d}}{\mathrm{d}\eta_2}\left(\bar{c}(\eta_1(\eta_2),\eta_2) + \frac{F(\bar{c}(\eta_1(\eta_2),\eta_2))}{f(\bar{c}(\eta_1(\eta_2),\eta_2))}\right) \\
&- (\lambda_0+\Delta)\,v'(\eta_1(\eta_2))\,\left[1-G(\eta_1(\eta_2))\right]\,\eta_1'(\eta_2) - \lambda_1\,v'(\eta_2)\,\left[1-G(\eta_2)\right] - v'(\eta_2) \\
=\ & v'(\eta_2)\,\frac{F(\bar{c}(\eta_1(\eta_2),\eta_2))}{f(\bar{c}(\eta_1(\eta_2),\eta_2))} + v(\eta_2)\,\frac{\mathrm{d}}{\mathrm{d}\eta_2}\left(\bar{c}(\eta_1(\eta_2),\eta_2) + \frac{F(\bar{c}(\eta_1(\eta_2),\eta_2))}{f(\bar{c}(\eta_1(\eta_2),\eta_2))}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ \left[\lambda_0\, v'(\eta_2) - (\lambda_0+\Delta)\, v'(\eta_1(\eta_2))\, \eta_1'(\eta_2)\right]\, \left[1 - G(\eta_1(\eta_2))\right] \\
&= v'(\eta_2)\, \frac{F(\bar{c}(\eta_1(\eta_2),\eta_2))}{f(\bar{c}(\eta_1(\eta_2),\eta_2))} + v(\eta_2)\, \frac{\mathrm{d}}{\mathrm{d}\eta_2}\left(\bar{c}(\eta_1(\eta_2),\eta_2) + \frac{F(\bar{c}(\eta_1(\eta_2),\eta_2))}{f(\bar{c}(\eta_1(\eta_2),\eta_2))}\right),
\end{aligned}
$$

which is positive for any $\eta_2 \in [0,\bar{\eta}]$. The unique existence of $\dot{\eta}_2$ is ensured by the monotonicity of t_1, $t_1(\eta^\dagger) \leq 0$, and the following:

- In Lemma 7, we have proved that given $t_1(\eta^\dagger) \leq 0$, it is impossible to have $\check{c}_2 > \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1)$.
- If $h(\bar{\eta}) \leq 0$ and $\check{c}_1 > \bar{c}(\check{\eta}_1, \check{\eta}_2)$, we have

  $$
  \begin{aligned}
  t_1(\check{\eta}_2) &= v(\check{\eta}_2)\, \left[\bar{c}(\check{\eta}_1,\check{\eta}_2) + \frac{F(\bar{c}(\check{\eta}_1,\check{\eta}_2))}{f(\bar{c}(\check{\eta}_1,\check{\eta}_2))}\right] + (\lambda_0+\Delta) \int_{\check{\eta}_1}^1 \left[v(z) - v(\check{\eta}_1)\right]\, g(z)\, \mathrm{d}z \\
  &\quad + \lambda_1 \int_{\check{\eta}_2}^1 \left[v(z) - v(\check{\eta}_2)\right]\, g(z)\, \mathrm{d}z - v(\check{\eta}_2) + v_E \\
  &= v(\check{\eta}_2)\, \left[\bar{c}(\check{\eta}_1,\check{\eta}_2) + \frac{F(\bar{c}(\check{\eta}_1,\check{\eta}_2))}{f(\bar{c}(\check{\eta}_1,\check{\eta}_2))}\right] + (\lambda_0+\Delta) \int_{\check{\eta}_1}^1 \left[v(z) - v(\check{\eta}_1)\right]\, g(z)\, \mathrm{d}z \\
  &\geq 0,
  \end{aligned}
  $$

  where the inequality follows from $\check{c}_1 > \check{c}(\check{\eta}_1, \check{\eta}_2)$.
- If $h(\bar{\eta}) > 0$, or equivalently,

  $$
  \lambda_1 \int_{\bar{\eta}}^1 v(z)\, g(z)\, \mathrm{d}z + v_E > 0,
  $$

  we have

  $$
  t_1(\bar{\eta}) = (\lambda_0+\Delta) \int_{\eta_1(\bar{\eta})}^1 \left[v(z) - v(\eta_1(\bar{\eta}))\right]\, g(z)\, \mathrm{d}z + \underbrace{\lambda_1 \int_{\bar{\eta}}^1 v(z)\, g(z)\, \mathrm{d}z + v_E}_{>0} > 0.
  $$

The remaining condition $h(\dot{\eta}_2) \geq 0$ follows from either $\dot{\eta}_2 \leq \check{\eta}_2$ or $h(\bar{\eta}) > 0$.

To prove the second part, differentiating t_2 yields

$$
\begin{aligned}
t_2'(\eta_2) = &\left(\frac{\Delta}{\lambda_0} + 1\right)\, v'(\eta_2)\, \left[\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)) + \frac{F(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}\right] \\
&+ \left[v(\eta_2) + \frac{\Delta}{\lambda_0}\, \left[v(\eta_2)+1\right]\right]\, \frac{\mathrm{d}}{\mathrm{d}\eta_2}\left[\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)) + \frac{F(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2),\eta_1(\eta_2)))}\right]
\end{aligned}
$$

$$
\begin{aligned}
& - (\lambda_0 + \Delta)\, v'(\eta_2)\, [1 - G(\eta_2)] + \frac{\Xi'(\eta_2)}{\lambda_0} \\
= & \left(\frac{\Delta}{\lambda_0} + 1\right) v'(\eta_2) \left[1 + (\lambda_0 + \lambda_1)\, [1 - G(\eta_1(\eta_2))] - \lambda_0\, [1 - G(\eta_2)] + \frac{F(\bar{c}(\eta_1(\eta_2), \eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2), \eta_1(\eta_2)))}\right] \\
& + \left[v(\eta_2) + \frac{\Delta}{\lambda_0}\, [v(\eta_2) + 1]\right] \frac{\mathrm{d}}{\mathrm{d}\eta_2} \left[\bar{c}(\eta_1(\eta_2), \eta_1(\eta_2)) + \frac{F(\bar{c}(\eta_1(\eta_2), \eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2), \eta_1(\eta_2)))}\right] \\
& - \left\{(\lambda_0 + \lambda_1 + \Delta)\, [G(\eta_2) - G(\eta_1(\eta_2))] + \lambda_1\, [1 - G(\eta_2)] + \frac{\Delta}{\lambda_0}\, [1 + \lambda_1\, [1 - G(\eta_1(\eta_2))]] + 1\right\} v'(\eta_2) \\
= & \left[v(\eta_2) + \frac{\Delta}{\lambda_0}\, [v(\eta_2) + 1]\right] \frac{\mathrm{d}}{\mathrm{d}\eta_2} \left[\bar{c}(\eta_1(\eta_2), \eta_1(\eta_2)) + \frac{F(\bar{c}(\eta_1(\eta_2), \eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2), \eta_1(\eta_2)))}\right] \\
& + \left(\frac{\Delta}{\lambda_0} + 1\right) v'(\eta_2)\, \frac{F(\bar{c}(\eta_1(\eta_2), \eta_1(\eta_2)))}{f(\bar{c}(\eta_1(\eta_2), \eta_1(\eta_2)))}.
\end{aligned}
$$

For any $\eta_2 \in [0, \bar{\eta}']$, we have $(\lambda_0 + \Delta)\, v(\eta_2) + \Delta \leq 0$, implying that $t_2'(\eta_2) \geq 0$. The unique existence of $\grave{\eta}_2 \in [\breve{\eta}, \eta^\dagger]$ is ensured by $t_2(\breve{\eta}) \leq 0 \leq t_2(\eta^\dagger)$ and the monotonicity of t_2.

By definition of Ξ in Lemma 5, to prove the non-positivity of $\iota(\grave{\eta}_1, \grave{\eta}_2)$, it is equivalent to showing that $\Xi(\grave{\eta}_2) \geq 0$. We prove the desired result in the following cases:

- If $h(\bar{\eta}) \leq 0$ and $\breve{c}_1 > \bar{c}(\breve{\eta}_1, \breve{\eta}_2)$, the desired result follows from

$$
\iota(\grave{\eta}_1, \grave{\eta}_2) = -\frac{\Xi(\grave{\eta}_2)}{\lambda_0} \leq -\frac{\Xi(\eta^\dagger)}{\lambda_0} \leq -\frac{\Xi(\breve{\eta}_2)}{\lambda_0} \leq 0,
$$

  where the three inequalities follow from $\grave{\eta}_2 \leq \eta^\dagger$, $\eta^\dagger \leq \breve{\eta}_2$, and $\Xi(\breve{\eta}_2) \geq 0$, respectively;

- If $h(\bar{\eta}) \leq 0$ and $\breve{c}_2 > \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1)$, since $\Xi(\mathring{\eta}_2) = 0$, it suffices to verfiy that $\grave{\eta}_2 \leq \mathring{\eta}_2$, or equivalently, $t_2(\mathring{\eta}_2) \geq 0$. The desired result follows from

$$
\begin{aligned}
t_2(\mathring{\eta}_2) = & \left[v(\mathring{\eta}_2) + \frac{\Delta}{\lambda_0}\, [v(\mathring{\eta}_2) + 1]\right] \left[\bar{c}(\mathring{\eta}_1, \mathring{\eta}_1) + \frac{F(\bar{c}(\mathring{\eta}_1, \mathring{\eta}_1))}{f(\bar{c}(\mathring{\eta}_1, \mathring{\eta}_1))}\right] \\
& + (\lambda_0 + \Delta) \int_{\mathring{\eta}_2}^1 [v(\eta) - v(\mathring{\eta}_2)]\, g(\eta)\, \mathrm{d}\eta - \underbrace{\iota(\mathring{\eta}_1, \mathring{\eta}_2)}_{=0} \\
\geq & \, 0,
\end{aligned}
$$

  where the inequality is implied by $\breve{c}_2 > \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1)$.

- If $h(\bar{\eta}) > 0$, the desired result follows from

$$
\iota(\grave{\eta}_1, \grave{\eta}_2) = -\frac{\Xi(\grave{\eta}_2)}{\lambda_0} \leq -\frac{\Xi(\eta^\dagger)}{\lambda_0} \leq 0,
$$

where the second inequality is implied by $\Xi(\eta^\dagger) = \lambda_0\, h(\eta^\dagger) \geq \lambda_0\, h(\bar{\eta}) \geq 0$.

This completes the proof. □

The following proposition summarizes the values of $\hat{c}$, $\hat{\alpha}$, $\hat{\eta}_1$, and $\hat{\eta}_2$ in all possible cases. We assume $t_2(\check{\eta}) \leq 0$ to exclude scenarios where the optimal $\hat{\eta}_1 = 0$, meaning final approval is granted regardless of the revealed efficacy level. Such cases imply that testing is not utilized, which is unrealistic in practice.

Proposition 8. *The values of $\hat{c}$, $\hat{\alpha}$, $\hat{\eta}_1$, and $\hat{\eta}_2$, depend on model parameters in the following cases.*

(1) If $\underline{c}(\check{\eta}_1) \leq \check{c}_1 \leq \bar{c}(\check{\eta}_1, \check{\eta}_2)$, the optimal policy follows case 1, as defined in (49). We have $\hat{c} = \check{c}_1$, $\hat{\alpha} = \frac{\hat{c} - \lambda_0\,[1-G(\hat{\eta}_1)]}{1+\lambda_1\,[1-G(\hat{\eta}_2)]}$, $\hat{\eta}_1 = \check{\eta}_1$, and $\hat{\eta}_2 = \check{\eta}_2$.

(2) If $\underline{c}(\mathring{\eta}_2) \leq \check{c}_2 \leq \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1)$, the optimal policy follows case 2, as defined in (50). We have $\hat{c} = \check{c}_2$, $\hat{\alpha} = \frac{\hat{c} - \lambda_0\,[1-G(\hat{\eta}_2)]}{1+\lambda_0\,[G(\hat{\eta}_2)-G(\hat{\eta}_1)]+\lambda_1\,[1-G(\hat{\eta}_1)]}$, $\hat{\eta}_1 = \mathring{\eta}_1$, and $\hat{\eta}_2 = \mathring{\eta}_2$.

(3) If $\check{c}_1 < \underline{c}(\check{\eta}_1)$ or $\check{c}_2 < \underline{c}(\mathring{\eta}_2)$, the government never grants conditional approval, i.e., $\hat{\alpha} = 0$, implying that $\hat{\alpha}_3(c, \eta) = 0$ for any $(c, \eta) \in [c_L, c_H] \times [0, 1]$. The thresholds $\hat{c}$, $\hat{\eta}_1$, and $\hat{\eta}_2$ are determined as follows.

(a) When $\check{c}_2 < \underline{c}(\mathring{\eta}_2)$ and $v(\tilde{\eta}_2) + 1 \leq 0$, the optimal policy follows case 2, as defined in (50),

(i) if $\underline{c}(\tilde{\eta}_2) < c_L$, we have $\hat{c} = c_L$ and $\hat{\eta}_2 = 1$;

(ii) if $c_L \leq \underline{c}(\tilde{\eta}_2) < c_H$, we have $\hat{c} = \underline{c}(\tilde{\eta}_2)$ and $\hat{\eta}_2 = \tilde{\eta}_2$;

(iii) if $\underline{c}(\bar{\eta}') < c_H \leq \underline{c}(\tilde{\eta}_2)$, we have $\hat{c} = c_H$ and $\hat{\eta}_2 = \sup\{\eta \in [0, 1] : \underline{c}(\eta) \geq c_H\}$; and

(iv) if $c_H \leq \underline{c}(\bar{\eta}')$, we have $\hat{c} = c_H$ and $\hat{\eta}_2 = \bar{\eta}'$.

The other threshold $\hat{\eta}_1$ can take any values between 0 and $\hat{\eta}_2$.

(b) When (i) $\check{c}_2 < \underline{c}(\mathring{\eta}_2)$ and $v(\tilde{\eta}_2) + 1 > 0$, or (ii) $\check{c}_1 < \underline{c}(\check{\eta}_1)$, the optimal policy follows case 1, as defined in (49),

(i) if $\underline{c}(\acute{\eta}_1) < c_L$, we have $\hat{c} = c_L$ and $\hat{\eta}_1 = 1$;

(ii) if $c_L \leq \underline{c}(\acute{\eta}_1) < c_H$, we have $\hat{c} = \underline{c}(\acute{\eta}_1)$ and $\hat{\eta}_1 = \acute{\eta}_1$;

(iii) if $\underline{c}(\bar{\eta}') < c_H \leq \underline{c}(\acute{\eta}_1)$, we have $\hat{c} = c_H$ and $\hat{\eta}_1 = \sup\{\eta \in [0, 1] : \underline{c}(\eta) \geq c_H\}$; and

(iv) if $c_H \leq \underline{c}(\bar{\eta}')$*, we have* $\hat{c} = c_H$ *and* $\hat{\eta}_1 = \bar{\eta}'$*.*

The other threshold $\hat{\eta}_2$ *can take any values between* $\hat{\eta}_1$ *and* 1*.*

(4) If $h(\bar{\eta}) > 0$*, or* $\check{c}_1 > \bar{c}(\check{\eta}_1, \check{\eta}_2)$*, or* $\check{c}_2 > \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1)$*, the government grants full conditional approval to firms whose types are below the threshold, i.e.,* $\hat{\alpha} = 1$*. The thresholds* $\hat{c}$*,* $\hat{\eta}_1$*, and* $\hat{\eta}_2$ *are determined as follows.*

(a) When $t_1(\eta^\dagger) \leq 0$*, the optimal policy follows case 1, as defined in* (49)*,*

(i) if $\bar{c}(\dot{\eta}_1, \dot{\eta}_2) < c_L$*, we have* $\hat{c} = c_L$ *and* $\hat{\eta}_1 = \hat{\eta}_2 = 1$*;*

(ii) if $c_L \leq \bar{c}(\dot{\eta}_1, \dot{\eta}_2) < c_H$*, we have* $\hat{c} = \bar{c}(\dot{\eta}_1, \dot{\eta}_2)$*,* $\hat{\eta}_1 = \dot{\eta}_1$*, and* $\hat{\eta}_2 = \dot{\eta}_2$*;*

(iii) if $\bar{c}(\bar{\eta}', \bar{\eta}) < c_H \leq \bar{c}(\dot{\eta}_1, \dot{\eta}_2)$*, we have* $\hat{c} = c_H$ *and* $(\hat{\eta}_1, \hat{\eta}_2)$ *such that* $\bar{c}(\hat{\eta}_1, \hat{\eta}_2) = c_H$ *and* $\Psi_1(\hat{\eta}_1, \hat{\eta}_2) = 0$*; and*

(iv) if $c_H \leq \bar{c}(\bar{\eta}', \bar{\eta})$*, we have* $\hat{c} = c_H$*,* $\hat{\eta}_1 = \bar{\eta}'$*, and* $\hat{\eta}_2 = \bar{\eta}$*.*

(b) When $0 \leq t_1(\eta^\dagger) = t_2(\eta^\dagger)$*, for cases (i)–(iii), the optimal policy follows case 2, as defined in* (50)*; and for cases (iv) and (v), the optimal policy follows case 1, as defined in* (49)*,*

(i) if $\bar{c}(\grave{\eta}_1, \grave{\eta}_1) < c_L$*, we have* $\hat{c} = c_L$ *and* $\hat{\eta}_1 = 1$*;*

(ii) if $c_L \leq \bar{c}(\grave{\eta}_1, \grave{\eta}_1) < c_H$*, we have* $\hat{c} = \bar{c}(\grave{\eta}_1, \grave{\eta}_1)$ *and* $\hat{\eta}_1 = \grave{\eta}_1$*;*

(iii) if $\bar{c}(\eta^\dagger, \eta^\dagger) \leq c_H \leq \bar{c}(\grave{\eta}_1, \grave{\eta}_1)$*, we have* $\hat{c} = c_H$ *and* $\hat{\eta}_1 = \sup\{\eta \in [0,1] : \bar{c}(\eta, \eta) \geq c_H\}$*;*

(iv) if $\bar{c}(\bar{\eta}', \bar{\eta}) < c_H \leq \bar{c}(\eta^\dagger, \eta^\dagger)$*, we have* $\hat{c} = c_H$ *and* $(\hat{\eta}_1, \hat{\eta}_2)$ *such that* $\bar{c}(\hat{\eta}_1, \hat{\eta}_2) = c_H$ *and* $\Psi_1(\hat{\eta}_1, \hat{\eta}_2) = 0$*; and*

(v) if $c_H \leq \bar{c}(\bar{\eta}', \bar{\eta})$*, we have* $\hat{c} = c_H$*,* $\hat{\eta}_1 = \bar{\eta}'$*, and* $\hat{\eta}_2 = \bar{\eta}$*.*

For cases (i)–(iii), the other threshold $\hat{\eta}_2$ *can take any values between* $\hat{\eta}_1$ *and* 1*.*

Proof. Proof of Proposition 8. The mechanism is a "take-it-or-leave-it" offer, and a firm of threshold type $\hat{c} < c_H$'s expected payoff is always 0. When $\hat{c} = c_H$, the mechanism is pooling and clearly feasible. Moreover, by Lemmas 5–8, we have $\hat{\eta}_1 \leq \hat{\eta}_2$, ensuring the feasibility of our candidate solutions in all cases.

We start with proving the optimality in case (1). Define

$$\kappa \equiv -v(\check{\eta}_2),$$

$$\chi(c,\eta) \equiv (\lambda_0 + \Delta)\,[v(\check{\eta}_1) - v(\eta)]\,g(\eta)\,f(c), \quad \forall (c,\eta) \in [c_L, c_H] \times [0, \check{\eta}_1],$$

$$\omega(c,\eta) \equiv \begin{cases} \{(\lambda_0 + \lambda_1 + \Delta)\,[v(\check{\eta}_1) - v(\eta)] + \lambda_1\,[v(\check{\eta}_2) - v(\check{\eta}_1)]\}\,g(\eta)\,f(c), & \forall (c,\eta) \in [c_L, c_H] \times [0, \check{\eta}_1], \\ \lambda_1\,[v(\check{\eta}_2) - v(\eta)]\,g(\eta)\,f(c), & \forall (c,\eta) \in [c_L, c_H] \times [\check{\eta}_1, \check{\eta}_2], \end{cases}$$

$$\mu(c,\eta) \equiv (\lambda_0 + \Delta)\ [v(\eta) - v(\check{\eta}_1)]\ g(\eta)\, f(c), \quad \forall (c,\eta) \in [c_L, c_H] \times [\check{\eta}_1, 1],$$
$$\xi(c,\eta) \equiv \lambda_1\ [v(\eta) - v(\check{\eta}_2)]\ g(\eta)\, f(c), \quad \forall (c,\eta) \in [c_L, c_H] \times [\check{\eta}_2, 1],$$
$$\sigma(c) \equiv \left\{ (\lambda_0 + \Delta) \int_{\check{\eta}_1}^{1} [v(\eta) - v(\check{\eta}_1)]\ g(\eta)\, \mathrm{d}\eta + v(\check{\eta}_2)\ \left[c + \frac{F(c)}{f(c)} \right] \right\} f(c), \quad \forall c \in [c_L, \hat{c}], \text{ and}$$
$$\beta(c) \equiv -\sigma(c), \quad \forall c \in [\hat{c}, c_H].$$

The nonnegativity of κ, χ, the second part of ω, μ, and ξ follows directly from their respective definitions, while the nonnegativity of the first part of ω follows from its monotonicity in η. The nonnegativity of σ and β is ensured by the monotonicity of $\sigma(c)/f(c)$ and $\beta(c)/f(c)$ (Assumption 1), along with the definition of the testing threshold $\hat{c} = \check{c}_1$. Because

$$-\kappa - \int_{c_L}^{c_H} \int_{\check{\eta}_2}^{1} \xi(c,\eta)\, \mathrm{d}\eta\, \mathrm{d}c = v_E,$$

$\forall (c,\eta) \in [c_L, c_H] \times [0, \check{\eta}_1]$:

$$-\chi(c,\eta) + \lambda_0\, g(\eta) \int_{\check{\eta}_2}^{1} \xi(c,z)\, \mathrm{d}z = [\lambda_0\ [v(\eta) - v_E] + \Delta\ [v(\eta) + 1]]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\check{\eta}_1, 1]$:

$$\mu(c,\eta) + \lambda_0\, g(\eta) \int_{\check{\eta}_2}^{1} \xi(c,z)\, \mathrm{d}z = [\lambda_0\ [v(\eta) - v_E] + \Delta\ [v(\eta) + 1]]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [0, \check{\eta}_1]$:

$$-\omega(c,\eta) + \chi(c,\eta) + \lambda_1\, g(\eta) \int_{\check{\eta}_2}^{1} \xi(c,z)\, \mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\check{\eta}_1, \check{\eta}_2]$:

$$-\omega(c,\eta) + \lambda_1\, g(\eta) \int_{\check{\eta}_2}^{1} \xi(c,z)\, \mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\check{\eta}_2, 1]$:

$$\xi(c,\eta) + \lambda_1\, g(\eta) \int_{\check{\eta}_2}^{1} \xi(c,z)\, \mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall c \in [c_L, \hat{c}]$:

$$\sigma(c) - \int_{\check{\eta}_1}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c} \int_{\check{\eta}_2}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - c \int_{\check{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta = v_E\,[c\,f(c) + F(c)],$$

$\forall c \in [\hat{c}, c_H]$:

$$-\beta(c) - \int_{\check{\eta}_1}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c} \int_{\check{\eta}_2}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - c \int_{\check{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta = v_E\,[c\,f(c) + F(c)],$$

we have

$$\begin{aligned}
&\kappa\,[-\Pi(c_H)] + \int_{c_L}^{\hat{c}} \sigma(c)\,\tau(c)\,\mathrm{d}c + \int_{\hat{c}}^{c_H} \beta(c)\,[-\tau(c)]\,\mathrm{d}c + \int_{c_L}^{c_H} \int_{\check{\eta}_1}^{1} \mu(c,\eta)\,[\hat{\alpha}_2(c,\eta) - \tau(c)]\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \int_{0}^{\check{\eta}_2} \omega(c,\eta)\,[-\hat{\alpha}_3(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c + \int_{c_L}^{c_H} \int_{0}^{\check{\eta}_1} \chi(c,\eta)\,[\hat{\alpha}_3(c,\eta) - \hat{\alpha}_2(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \int_{\check{\eta}_2}^{1} \xi(c,\eta)\,\left[\hat{\alpha}_3(c,\eta) - \Pi(c_H) - \int_{c}^{c_H} \tau(y)\,\mathrm{d}y - c\,\tau(c) + \int_{0}^{1} [\lambda_0\,\hat{\alpha}_2(c,z) + \lambda_1\,\hat{\alpha}_3(c,z)]\,\mathrm{d}G(z)\right] \mathrm{d}\eta\,\mathrm{d}c \\
= &\left[-\kappa - \int_{c_L}^{c_H} \int_{\check{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\right]\,\Pi(c_H) \\
&\quad + \int_{c_L}^{c_H} \int_{0}^{\check{\eta}_1} \left[-\chi(c,\eta) + \lambda_0\,g(\eta) \int_{\check{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z\right]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \int_{\check{\eta}_1}^{1} \left[\mu(c,\eta) + \lambda_0\,g(\eta) \int_{\check{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z\right]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \int_{0}^{\check{\eta}_1} \left[-\omega(c,\eta) + \chi(c,\eta) + \lambda_1\,g(\eta) \int_{\check{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z\right]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \int_{\check{\eta}_1}^{\check{\eta}_2} \left[-\omega(c,\eta) + \lambda_1\,g(\eta) \int_{\check{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z\right]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \int_{\check{\eta}_2}^{1} \left[\xi(c,\eta) + \lambda_1\,g(\eta) \int_{\check{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z\right]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \left[\sigma(c) - \int_{\check{\eta}_1}^{1} \mu(c,\eta)\,\mathrm{d}\eta - c \int_{\check{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c} \int_{\check{\eta}_2}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\,\tau(c)\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \left[-\beta(c) - \int_{\check{\eta}_1}^{1} \mu(c,\eta)\,\mathrm{d}\eta - c \int_{\check{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c} \int_{\check{\eta}_2}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\,\tau(c)\,\mathrm{d}c \\
= &\, v_E\,\Pi(c_H) + v_E \int_{c_L}^{c_H} \left[c + \frac{F(c)}{f(c)}\right]\,\tau(c)\,f(c)\,\mathrm{d}c + \int_{c_L}^{c_H} \int_{0}^{1} [\lambda_0\,[v(\eta) - v_E] + \Delta\,[v(\eta) + 1]]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) \\
&\quad + \lambda_1 \int_{c_L}^{c_H} \int_{0}^{1} [v(\eta) - v_E]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c).
\end{aligned}$$

Therefore, the objective function is bounded by $\int_{c_L}^{\hat{c}} \sigma(c)\,\mathrm{d}c$, which is attained by our candi-

date primal solution, i.e.,

$$
\begin{aligned}
&\int_{c_L}^{\hat{c}} \sigma(c)\,\mathrm{d}c - \left[v_E\,\hat{\alpha} + \lambda_0 \int_{\check{\eta}_1}^{1} v(\eta)\,\mathrm{d}G(\eta) + \lambda_1\,\hat{\alpha} \int_{\check{\eta}_2}^{1} v(\eta)\,\mathrm{d}G(\eta) + \Delta \int_{\check{\eta}_1}^{1} [v(\eta)+1]\,\mathrm{d}G(\eta)\right] F(\hat{c}) \\
&= (\lambda_0+\Delta)\,F(\hat{c}) \int_{\check{\eta}_1}^{1} v(\eta)\,g(\eta)\,\mathrm{d}\eta - (\lambda_0+\Delta)\,F(\hat{c})\,v(\check{\eta}_1)\,[1-G(\check{\eta}_1)] + \hat{c}\,F(\hat{c})\,v(\check{\eta}_2) \\
&\quad - \left[v_E\,\hat{\alpha} + \lambda_0 \int_{\check{\eta}_1}^{1} v(\eta)\,\mathrm{d}G(\eta) + \lambda_1\,\hat{\alpha} \int_{\check{\eta}_2}^{1} v(\eta)\,\mathrm{d}G(\eta) + \Delta \int_{\check{\eta}_1}^{1} [v(\eta)+1]\,\mathrm{d}G(\eta)\right] F(\hat{c}) \\
&= -(\lambda_0+\Delta)\,F(\hat{c})\,v(\check{\eta}_1)\,[1-G(\check{\eta}_1)] + \hat{c}\,F(\hat{c})\,v(\check{\eta}_2) \\
&\quad - \left[v_E\,\hat{\alpha} + \underbrace{\lambda_1\,\hat{\alpha} \int_{\check{\eta}_2}^{1} v(\eta)\,\mathrm{d}G(\eta)}_{=\hat{\alpha}\,[v(\check{\eta}_2)-v_E+\lambda_1\,v(\check{\eta}_2)\,[1-G(\check{\eta}_2)]]} + \Delta\,[1-G(\check{\eta}_1)]\right] F(\hat{c}) \\
&= -(\lambda_0+\Delta)\,F(\hat{c})\,v(\check{\eta}_1)\,[1-G(\check{\eta}_1)] + \hat{c}\,F(\hat{c})\,v(\check{\eta}_2) - [[\hat{c}-\lambda_0\,[1-G(\check{\eta}_1)]]\,v(\check{\eta}_2) + \Delta\,[1-G(\check{\eta}_1)]]\,F(\hat{c}) \\
&= [-(\lambda_0+\Delta)\,v(\check{\eta}_1) + \lambda_0\,v(\check{\eta}_2) - \Delta]\,[1-G(\check{\eta}_1)]\,F(\hat{c}) \\
&= 0.
\end{aligned}
$$

To prove the optimality in case (2), define

$$
\begin{aligned}
&\kappa \equiv -\frac{1}{\lambda_0}\,[(\lambda_0+\Delta)\,v(\mathring{\eta}_2)+\Delta]\,, \\
&\chi(c,\eta) \equiv (\lambda_0+\Delta)\,[v(\mathring{\eta}_2)-v(\eta)]\,g(\eta)\,f(c), \quad \forall (c,\eta) \in [c_L,c_H]\times[0,\mathring{\eta}_2], \\
&\mu(c,\eta) \equiv (\lambda_0+\Delta)\,[v(\eta)-v(\mathring{\eta}_2)]\,g(\eta)\,f(c), \quad \forall (c,\eta) \in [c_L,c_H]\times[\mathring{\eta}_2,1], \\
&\omega(c,\eta) \equiv \left\{(\lambda_0+\lambda_1+\Delta)\,[v(\mathring{\eta}_2)-v(\eta)] + \frac{\lambda_1}{\lambda_0}\,\Delta\,[v(\mathring{\eta}_2)+1]\right\} g(\eta)\,f(c), \quad \forall (c,\eta) \in [c_L,c_H]\times[0,\mathring{\eta}_1], \\
&\xi(c,\eta) \equiv \begin{cases} -\left\{(\lambda_0+\lambda_1+\Delta)\,[v(\mathring{\eta}_2)-v(\eta)] + \frac{\lambda_1}{\lambda_0}\,\Delta\,[v(\mathring{\eta}_2)+1]\right\} g(\eta)\,f(c), & \forall (c,\eta) \in [c_L,c_H]\times[\mathring{\eta}_1,\mathring{\eta}_2], \\ \left\{\lambda_1\,[v(\eta)-v(\mathring{\eta}_2)] - \frac{\lambda_1}{\lambda_0}\,\Delta\,[v(\mathring{\eta}_2)+1]\right\} g(\eta)\,f(c), & \forall (c,\eta) \in [c_L,c_H]\times[\mathring{\eta}_2,1], \end{cases} \\
&\sigma(c) \equiv \left\{(\lambda_0+\Delta)\int_{\mathring{\eta}_2}^{1} [v(\eta)-v(\mathring{\eta}_2)]\,g(\eta)\,\mathrm{d}\eta + \left[c+\frac{F(c)}{f(c)}\right]\left[v(\mathring{\eta}_2)+\frac{\Delta}{\lambda_0}\,[v(\mathring{\eta}_2)+1]\right]\right\} f(c), \quad \forall c \in [c_L,\hat{c}], \\
&\beta(c) \equiv -\sigma(c), \quad \forall c \in [\hat{c},c_H].
\end{aligned}
$$

The nonnegativity of κ follows from $\mathring{\eta}_2 \leq \eta^{\dagger}$, as established in Lemma 5(2). The nonnegativity of χ and μ follows from their definitions directly. The nonnegativity of ω and ξ follows from the relationship between $\mathring{\eta}_1$ and $\mathring{\eta}_2$ (i.e., $\Psi_2(\mathring{\eta}_1,\mathring{\eta}_2)=0$, implying $\omega(c,\mathring{\eta}_1)=\xi(c,\mathring{\eta}_1)=0$

for any $c \in [c_L, c_H]$) and their monotonicity in η. The nonnegativity of σ and β is ensured by the monotonicity of $\sigma(c)/f(c)$ and $\beta(c)/f(c)$, along with the definition of the testing threshold $\hat{c} = \check{c}_2$. Because

$$-\kappa - \int_{c_L}^{c_H} \int_{\mathring{\eta}_1}^{1} \xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c = v_E,$$

$\forall (c,\eta) \in [c_L, c_H] \times [0, \mathring{\eta}_2]$:

$$-\chi(c,\eta) + \lambda_0\, g(\eta) \int_{\mathring{\eta}_1}^{1} \xi(c,z)\,\mathrm{d}z = [\lambda_0\ [v(\eta) - v_E] + \Delta\ [v(\eta) + 1]]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\mathring{\eta}_2, 1]$:

$$\mu(c,\eta) + \lambda_0\, g(\eta) \int_{\mathring{\eta}_1}^{1} \xi(c,z)\,\mathrm{d}z = [\lambda_0\ [v(\eta) - v_E] + \Delta\ [v(\eta) + 1]]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [0, \mathring{\eta}_1]$:

$$-\omega(c,\eta) + \chi(c,\eta) + \lambda_1\, g(\eta) \int_{\mathring{\eta}_1}^{1} \xi(c,z)\,\mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\mathring{\eta}_1, \mathring{\eta}_2]$:

$$\chi(c,\eta) + \xi(c,\eta) + \lambda_1\, g(\eta) \int_{\mathring{\eta}_1}^{1} \xi(c,z)\,\mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\mathring{\eta}_2, 1]$:

$$\xi(c,\eta) + \lambda_1\, g(\eta) \int_{\mathring{\eta}_1}^{1} \xi(c,z)\,\mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall c \in [c_L, \hat{c}]$:

$$\sigma(c) - \int_{\mathring{\eta}_2}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c} \int_{\mathring{\eta}_1}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - c \int_{\mathring{\eta}_1}^{1} \xi(c,\eta)\,\mathrm{d}\eta = v_E\,[c\, f(c) + F(c)],$$

$\forall c \in [\hat{c}, c_H]$:

$$-\beta(c) - \int_{\mathring{\eta}_2}^{1} \mu(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c} \int_{\mathring{\eta}_1}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - c \int_{\mathring{\eta}_1}^{1} \xi(c,\eta)\,\mathrm{d}\eta = v_E\,[c\, f(c) + F(c)],$$

we have

$$
\begin{aligned}
&\kappa\,[-\Pi(c_H)] + \int_{c_L}^{\hat{c}} \sigma(c)\,\tau(c)\,\mathrm{d}c + \int_{\hat{c}}^{c_H} \beta(c)\,[-\tau(c)]\,\mathrm{d}c + \int_{c_L}^{c_H}\int_{\mathring{\eta}_2}^{1} \mu(c,\eta)\,[\hat{\alpha}_2(c,\eta) - \tau(c)]\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H}\int_{0}^{\mathring{\eta}_1} \omega(c,\eta)\,[-\hat{\alpha}_3(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c + \int_{c_L}^{c_H}\int_{0}^{\mathring{\eta}_2} \chi(c,\eta)\,[\hat{\alpha}_3(c,\eta) - \hat{\alpha}_2(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H}\int_{\mathring{\eta}_1}^{1} \xi(c,\eta)\,\left[\hat{\alpha}_3(c,\eta) - \Pi(c_H) - \int_{c}^{c_H} \tau(y)\,\mathrm{d}y - c\,\tau(c) + \int_{0}^{1} [\lambda_0\,\hat{\alpha}_2(c,z) + \lambda_1\,\hat{\alpha}_3(c,z)]\,\mathrm{d}G(z)\right]\mathrm{d}\eta\,\mathrm{d}c \\
=\;& \left[-\kappa - \int_{c_L}^{c_H}\int_{\mathring{\eta}_1}^{1} \xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\right]\,\Pi(c_H) \\
&\quad + \int_{c_L}^{c_H}\int_{0}^{\mathring{\eta}_2} \left[-\chi(c,\eta) + \lambda_0\,g(\eta)\int_{\mathring{\eta}_1}^{1} \xi(c,z)\,\mathrm{d}z\right]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H}\int_{\mathring{\eta}_2}^{1} \left[\mu(c,\eta) + \lambda_0\,g(\eta)\int_{\mathring{\eta}_1}^{1} \xi(c,z)\,\mathrm{d}z\right]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H}\int_{0}^{\mathring{\eta}_1} \left[-\omega(c,\eta) + \chi(c,\eta) + \lambda_1\,g(\eta)\int_{\mathring{\eta}_1}^{1} \xi(c,z)\,\mathrm{d}z\right]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H}\int_{\mathring{\eta}_1}^{\mathring{\eta}_2} \left[\chi(c,\eta) + \xi(c,\eta) + \lambda_1\,g(\eta)\int_{\mathring{\eta}_1}^{1} \xi(c,z)\,\mathrm{d}z\right]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H}\int_{\mathring{\eta}_2}^{1} \left[\xi(c,\eta) + \lambda_1\,g(\eta)\int_{\mathring{\eta}_1}^{1} \xi(c,z)\,\mathrm{d}z\right]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \left[\sigma(c) - \int_{\mathring{\eta}_2}^{1} \mu(c,\eta)\,\mathrm{d}\eta - c\int_{\mathring{\eta}_1}^{1} \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c}\int_{\mathring{\eta}_1}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\,\tau(c)\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \left[-\beta(c) - \int_{\mathring{\eta}_2}^{1} \mu(c,\eta)\,\mathrm{d}\eta - c\int_{\mathring{\eta}_1}^{1} \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c}\int_{\mathring{\eta}_1}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\,\tau(c)\,\mathrm{d}c \\
=\;& v_E\,\Pi(c_H) + v_E\int_{c_L}^{c_H} \left[c + \frac{F(c)}{f(c)}\right]\,\tau(c)\,f(c)\,\mathrm{d}c + \int_{c_L}^{c_H}\int_{0}^{1} [\lambda_0\,[v(\eta) - v_E] + \Delta\,[v(\eta) + 1]]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) \\
&\quad + \lambda_1\int_{c_L}^{c_H}\int_{0}^{1} [v(\eta) - v_E]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c).
\end{aligned}
$$

Therefore, the objective function is bounded by $\int_{c_L}^{\hat{c}} \sigma(c)\,\mathrm{d}c$, which is attained by our candidate primal solution, i.e.,

$$
\begin{aligned}
&\int_{c_L}^{\hat{c}} \sigma(c)\,\mathrm{d}c - \left[v_E\,\hat{\alpha} + \hat{\alpha}\int_{\mathring{\eta}_1}^{\mathring{\eta}_2} [(\lambda_0 + \Delta)\,v(\eta) + \Delta]\,\mathrm{d}G(\eta)\right. \\
&\qquad \left. + \int_{\mathring{\eta}_2}^{1} [(\lambda_0 + \Delta)\,v(\eta) + \Delta]\,\mathrm{d}G(\eta) + \lambda_1\,\hat{\alpha}\int_{\mathring{\eta}_1}^{1} v(\eta)\,\mathrm{d}G(\eta)\right]\,F(\hat{c}) \\
=\;& (\lambda_0 + \Delta)\,F(\hat{c})\int_{\mathring{\eta}_2}^{1} [v(\eta) - v(\mathring{\eta}_2)]\,g(\eta)\,\mathrm{d}\eta + \hat{c}\,F(\hat{c})\left[v(\mathring{\eta}_2) + \frac{\Delta}{\lambda_0}\,[v(\mathring{\eta}_2) + 1]\right]
\end{aligned}
$$

$$
\begin{aligned}
&- \left[v_E\,\hat{\alpha} + \hat{\alpha} \int_{\mathring{\eta}_1}^{\mathring{\eta}_2} \left[(\lambda_0 + \Delta)\, v(\eta) + \Delta \right] \mathrm{d}G(\eta) \right. \\
&\qquad \left. + \int_{\mathring{\eta}_2}^{1} \left[(\lambda_0 + \Delta)\, v(\eta) + \Delta \right] \mathrm{d}G(\eta) + \lambda_1\, \hat{\alpha} \int_{\mathring{\eta}_1}^{1} v(\eta)\, \mathrm{d}G(\eta) \right] F(\hat{c}) \\
= &\left\{ -(\lambda_0 + \Delta)\, v(\mathring{\eta}_2)\, \left[1 - G(\mathring{\eta}_2) \right] + \hat{c} \left[v(\mathring{\eta}_2) + \frac{\Delta}{\lambda_0}\, \left[v(\mathring{\eta}_2) + 1 \right] \right] \right. \\
&\left. - \left[v_E\,\hat{\alpha} + \hat{\alpha} \int_{\mathring{\eta}_1}^{\mathring{\eta}_2} \left[(\lambda_0 + \Delta)\, v(\eta) + \Delta \right] \mathrm{d}G(\eta) + \Delta\, \left[1 - G(\mathring{\eta}_2) \right] + \lambda_1\, \hat{\alpha} \int_{\mathring{\eta}_1}^{1} v(\eta)\, \mathrm{d}G(\eta) \right] \right\} F(\hat{c}),
\end{aligned}
$$

where $\hat{\alpha} = \frac{\hat{c} - \lambda_0 \left[1 - G(\hat{\eta}_2) \right]}{1 + \lambda_0 \left[G(\hat{\eta}_2) - G(\hat{\eta}_1) \right] + \lambda_1 \left[1 - G(\hat{\eta}_1) \right]}$ represents the scale of conditional approval. Because

$$
\begin{aligned}
\lambda_1 \int_{\mathring{\eta}_1}^{1} v(\eta)\, \mathrm{d}G(\eta) &= \lambda_1 \left[\int_{\mathring{\eta}_1}^{\mathring{\eta}_2} v(\eta)\, \mathrm{d}G(\eta) + \int_{\mathring{\eta}_2}^{1} v(\eta)\, \mathrm{d}G(\eta) \right] \\
&= \lambda_1 \int_{\mathring{\eta}_1}^{\mathring{\eta}_2} v(\eta)\, \mathrm{d}G(\eta) + \left[v(\mathring{\eta}_2) - v_E \right] + \frac{\Delta}{\lambda_0}\, \left[v(\mathring{\eta}_2) + 1 \right] \left[1 + \lambda_1\, \left[1 - G(\mathring{\eta}_1) \right] \right] \\
&\quad + (\lambda_0 + \lambda_1 + \Delta) \int_{\mathring{\eta}_1}^{\mathring{\eta}_2} \left[v(\mathring{\eta}_2) - v(z) \right]\, g(z)\, \mathrm{d}z + \lambda_1\, v(\mathring{\eta}_2)\, \left[1 - G(\mathring{\eta}_2) \right],
\end{aligned}
$$

where the second equality follows from $\iota(\mathring{\eta}_1, \mathring{\eta}_2) = 0$, and

$$
\hat{c} = \hat{\alpha}\, \left[\lambda_0\, \left[G(\mathring{\eta}_2) - G(\mathring{\eta}_1) \right] + \lambda_1\, \left[1 - G(\mathring{\eta}_1) \right] + 1 \right] + \lambda_0\, \left[1 - G(\mathring{\eta}_2) \right],
$$

the duality gap boils down to

$$
\begin{aligned}
&- (\lambda_0 + \Delta)\, v(\mathring{\eta}_2)\, \left[1 - G(\mathring{\eta}_2) \right] - \Delta\, \left[1 - G(\mathring{\eta}_2) + \hat{\alpha}\, \left[G(\mathring{\eta}_2) - G(\mathring{\eta}_1) \right] \right] \\
&+ \left[\hat{\alpha}\, \left[\lambda_0\, \left[G(\mathring{\eta}_2) - G(\mathring{\eta}_1) \right] + \lambda_1\, \left[1 - G(\mathring{\eta}_1) \right] \right] + \lambda_0\, \left[1 - G(\mathring{\eta}_2) \right] \right]\, v(\mathring{\eta}_2) \\
&+ \left[\hat{\alpha}\, \lambda_0\, \left[G(\mathring{\eta}_2) - G(\mathring{\eta}_1) \right] + \lambda_0\, \left[1 - G(\mathring{\eta}_2) \right] \right]\, \frac{\Delta}{\lambda_0}\, \left[v(\mathring{\eta}_2) + 1 \right] \\
&- \hat{\alpha}\, \left[(\lambda_0 + \lambda_1 + \Delta)\, v(\mathring{\eta}_2)\, \left[G(\mathring{\eta}_2) - G(\mathring{\eta}_1) \right] + \lambda_1\, v(\mathring{\eta}_2)\, \left[1 - G(\mathring{\eta}_2) \right] \right] \\
= &\, 0.
\end{aligned}
$$

To prove the optimality of cases (3)(a)(i) and (3)(a)(ii), define

$$
\begin{aligned}
&\kappa \equiv -\frac{1}{\lambda_0}\, \left[(\lambda_0 + \Delta)\, v(\tilde{\eta}_2) + \Delta \right], \\
&\chi(c, \eta) \equiv (\lambda_0 + \Delta)\, \left[v(\tilde{\eta}_2) - v(\eta) \right]\, g(\eta)\, f(c), \quad \forall (c, \eta) \in [c_L, c_H] \times [0, \tilde{\eta}_2], \\
&\mu(c, \eta) \equiv (\lambda_0 + \Delta)\, \left[v(\eta) - v(\tilde{\eta}_2) \right]\, g(\eta)\, f(c), \quad \forall (c, \eta) \in [c_L, c_H] \times [\tilde{\eta}_2, 1],
\end{aligned}
$$

$$\omega(c,\eta) \equiv \left\{(\lambda_0+\lambda_1+\Delta)\ [v(\tilde{\eta}_2)-v(\eta)]+\frac{\lambda_1}{\lambda_0}\Delta\ [v(\tilde{\eta}_2)+1]\right\} g(\eta)\, f(c), \quad \forall (c,\eta)\in[c_L,c_H]\times[0,\tilde{\eta}_1],$$

$$\xi(c,\eta) \equiv \begin{cases} -\left\{(\lambda_0+\lambda_1+\Delta)\ [v(\tilde{\eta}_2)-v(\eta)]+\frac{\lambda_1}{\lambda_0}\Delta\ [v(\tilde{\eta}_2)+1]\right\} g(\eta)\, f(c), & \forall (c,\eta)\in[c_L,c_H]\times[\tilde{\eta}_1,\tilde{\eta}_2], \\ \left\{\lambda_1\ [v(\eta)-v(\tilde{\eta}_2)]-\frac{\lambda_1}{\lambda_0}\Delta\ [v(\tilde{\eta}_2)+1]\right\} g(\eta)\, f(c), & \forall (c,\eta)\in[c_L,c_H]\times[\tilde{\eta}_2,1], \end{cases}$$

$$\phi(c) \equiv \iota(\tilde{\eta}_1,\tilde{\eta}_2)\, f(c), \quad \forall c\in[c_L,c_H],$$

$$\sigma(c) \equiv \left\{(\lambda_0+\Delta)\int_{\tilde{\eta}_2}^{1}[v(\eta)-v(\tilde{\eta}_2)]\ g(\eta)\,\mathrm{d}\eta+\left[c+\frac{F(c)}{f(c)}\right]\left[v(\tilde{\eta}_2)+\frac{\Delta}{\lambda_0}\ [v(\tilde{\eta}_2)+1]\right]\right\} f(c), \quad \forall c\in\mathcal{C}_1,$$

$$\beta(c) \equiv -\sigma(c), \quad \forall c\in\mathcal{C}_2,$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are subsets of $[c_L,c_H]$ satisfying: (i) $\mathcal{C}_1\cup\mathcal{C}_2=[c_L,c_H]$; and (ii) the measure of $\mathcal{C}_1\cap\mathcal{C}_2$ is 0. These subsets will be specified later.

The nonnegativity of κ is implied by $\tilde{\eta}_2\leq\bar{\eta}'$, as established in Lemma 6(1). The nonnegativity of χ and μ follows from their definitions directly. The nonnegativity of ω and ξ follows from the relationship between $\tilde{\eta}_1$ and $\tilde{\eta}_2$ (i.e., $\Psi_2(\tilde{\eta}_1,\tilde{\eta}_2)=0$, implying $\omega(c,\tilde{\eta}_1)=\xi(c,\tilde{\eta}_1)=0$ for any $c\in[c_L,c_H]$) and their monotonicity in η. The nonnegativity of ϕ follows from the fact that $\iota(\tilde{\eta}_1,\tilde{\eta}_2)\geq 0$, as established in Lemma 6(1).

In the first two possible cases, we have:

1. if $\underline{c}(\tilde{\eta}_2)<c_L$, we know that β is nonnegative on $[c_L,c_H]$. In this case, we have $\mathcal{C}_1=\emptyset$ and $\mathcal{C}_2=[c_L,c_H]$; and

2. if $c_L\leq\underline{c}(\tilde{\eta}_2)\leq c_H$, we know that σ and β are nonnegative on $[c_L,\underline{c}(\tilde{\eta}_2)]$ and $[\underline{c}(\tilde{\eta}_2),c_H]$, respectively. In this case, we have $\mathcal{C}_1=[c_L,\underline{c}(\tilde{\eta}_2)]$ and $\mathcal{C}_2=[\underline{c}(\tilde{\eta}_2),c_H]$.

Therefore, the nonnegativity of σ and β holds automatically.

Because

$$-\kappa-\int_{c_L}^{c_H}\phi(c)\,\mathrm{d}c-\int_{c_L}^{c_H}\int_{\tilde{\eta}_1}^{1}\xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c=v_E,$$

$\forall(c,\eta)\in[c_L,c_H]\times[0,\tilde{\eta}_2]$:

$$\lambda_0\,\phi(c)\,g(\eta)-\chi(c,\eta)+\lambda_0\,g(\eta)\int_{\tilde{\eta}_1}^{1}\xi(c,z)\,\mathrm{d}z=[\lambda_0\ [v(\eta)-v_E]+\Delta\ [v(\eta)+1]]\ g(\eta)\,f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\tilde{\eta}_2, 1]$:

$$\lambda_0\, \phi(c)\, g(\eta) + \mu(c,\eta) + \lambda_0\, g(\eta) \int_{\tilde{\eta}_1}^{1} \xi(c,z)\, \mathrm{d}z = [\lambda_0\ [v(\eta) - v_E] + \Delta\ [v(\eta) + 1]]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [0, \tilde{\eta}_1]$:

$$\lambda_1\, \phi(c)\, g(\eta) - \omega(c,\eta) + \chi(c,\eta) + \lambda_1\, g(\eta) \int_{\tilde{\eta}_1}^{1} \xi(c,z)\, \mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\tilde{\eta}_1, \tilde{\eta}_2]$:

$$\lambda_1\, \phi(c)\, g(\eta) + \chi(c,\eta) + \xi(c,\eta) + \lambda_1\, g(\eta) \int_{\tilde{\eta}_1}^{1} \xi(c,z)\, \mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\tilde{\eta}_2, 1]$:

$$\lambda_1\, \phi(c)\, g(\eta) + \xi(c,\eta) + \lambda_1\, g(\eta) \int_{\tilde{\eta}_1}^{1} \xi(c,z)\, \mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\, f(c),$$

$\forall c \in \mathcal{C}_1$:

$$\begin{aligned} &- \int_{c_L}^{c} \phi(y)\, \mathrm{d}y - c\, \phi(c) + \sigma(c) - c \int_{\tilde{\eta}_1}^{1} \xi(c,\eta)\, \mathrm{d}\eta \\ &\qquad - \int_{c_L}^{c} \int_{\tilde{\eta}_1}^{1} \xi(y,\eta)\, \mathrm{d}\eta\, \mathrm{d}y - \int_{\tilde{\eta}_2}^{1} \mu(c,\eta)\, \mathrm{d}\eta = v_E\, [c\, f(c) + F(c)], \end{aligned}$$

$\forall c \in \mathcal{C}_2$:

$$\begin{aligned} &- \int_{c_L}^{c} \phi(y)\, \mathrm{d}y - c\, \phi(c) - \beta(c) - c \int_{\tilde{\eta}_1}^{1} \xi(c,\eta)\, \mathrm{d}\eta \\ &\qquad - \int_{c_L}^{c} \int_{\tilde{\eta}_1}^{1} \xi(y,\eta)\, \mathrm{d}\eta\, \mathrm{d}y - \int_{\tilde{\eta}_2}^{1} \mu(c,\eta)\, \mathrm{d}\eta = v_E\, [c\, f(c) + F(c)], \end{aligned}$$

we have

$$\begin{aligned} &\kappa\ [-\Pi(c_H)] + \int_{\mathcal{C}_1} \sigma(c)\, \tau(c)\, \mathrm{d}c + \int_{\mathcal{C}_2} \beta(c)\ [-\tau(c)]\, \mathrm{d}c + \int_{c_L}^{c_H} \int_{\tilde{\eta}_2}^{1} \mu(c,\eta)\ [\hat{\alpha}_2(c,\eta) - \tau(c)]\, \mathrm{d}\eta\, \mathrm{d}c \\ &+ \int_{c_L}^{c_H} \int_{0}^{\tilde{\eta}_1} \omega(c,\eta)\ [-\hat{\alpha}_3(c,\eta)]\, \mathrm{d}\eta\, \mathrm{d}c + \int_{c_L}^{c_H} \int_{0}^{\tilde{\eta}_2} \chi(c,\eta)\ [\hat{\alpha}_3(c,\eta) - \hat{\alpha}_2(c,\eta)]\, \mathrm{d}\eta\, \mathrm{d}c \end{aligned}$$

$$
\begin{aligned}
&+ \int_{c_L}^{c_H} \int_{\tilde{\eta}_1}^{1} \xi(c,\eta) \left[\hat{\alpha}_3(c,\eta) - \Pi(c_H) - \int_c^{c_H} \tau(y)\,\mathrm{d}y - c\,\tau(c) + \int_0^1 \left[\lambda_0\,\hat{\alpha}_2(c,z) + \lambda_1\,\hat{\alpha}_3(c,z)\right] \mathrm{d}G(z) \right] \mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \phi(c) \left[-\Pi(c_H) - \int_c^{c_H} \tau(y)\,\mathrm{d}y - c\,\tau(c) + \lambda_0 \int_0^1 \hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta) + \lambda_1 \int_0^1 \hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta) \right] \mathrm{d}c \\
= &\left[-\kappa - \int_{c_L}^{c_H} \phi(c)\,\mathrm{d}c - \int_{c_L}^{c_H} \int_{\tilde{\eta}_1}^{1} \xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \right] \Pi(c_H) \\
&+ \int_{c_L}^{c_H} \int_0^{\tilde{\eta}_2} \left[\lambda_0\,\phi(c)\,g(\eta) - \chi(c,\eta) + \lambda_0\,g(\eta) \int_{\tilde{\eta}_1}^1 \xi(c,z)\,\mathrm{d}z \right] \hat{\alpha}_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_{\tilde{\eta}_2}^1 \left[\lambda_0\,\phi(c)\,g(\eta) + \mu(c,\eta) + \lambda_0\,g(\eta) \int_{\tilde{\eta}_1}^1 \xi(c,z)\,\mathrm{d}z \right] \hat{\alpha}_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_0^{\tilde{\eta}_1} \left[\lambda_1\,\phi(c)\,g(\eta) - \omega(c,\eta) + \chi(c,\eta) + \lambda_1\,g(\eta) \int_{\tilde{\eta}_1}^1 \xi(c,z)\,\mathrm{d}z \right] \hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_{\tilde{\eta}_1}^{\tilde{\eta}_2} \left[\lambda_1\,\phi(c)\,g(\eta) + \chi(c,\eta) + \xi(c,\eta) + \lambda_1\,g(\eta) \int_{\tilde{\eta}_1}^1 \xi(c,z)\,\mathrm{d}z \right] \hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \int_{\tilde{\eta}_2}^1 \left[\lambda_1\,\phi(c)\,g(\eta) + \xi(c,\eta) + \lambda_1\,g(\eta) \int_{\tilde{\eta}_1}^1 \xi(c,z)\,\mathrm{d}z \right] \hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{\mathcal{C}_1} \left[-\int_{c_L}^{c} \phi(y)\,\mathrm{d}y - c\,\phi(c) + \sigma(c) - \int_{\tilde{\eta}_2}^1 \mu(c,\eta)\,\mathrm{d}\eta - c \int_{\tilde{\eta}_1}^1 \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c} \int_{\tilde{\eta}_1}^1 \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y \right] \tau(c)\,\mathrm{d}c \\
&+ \int_{\mathcal{C}_2} \left[-\int_{c_L}^{c} \phi(y)\,\mathrm{d}y - c\,\phi(c) - \beta(c) - \int_{\tilde{\eta}_2}^1 \mu(c,\eta)\,\mathrm{d}\eta - c \int_{\tilde{\eta}_1}^1 \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c} \int_{\tilde{\eta}_1}^1 \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y \right] \tau(c)\,\mathrm{d}c \\
= &\, v_E\,\Pi(c_H) + v_E \int_{c_L}^{c_H} \left[c + \frac{F(c)}{f(c)} \right] \tau(c)\,f(c)\,\mathrm{d}c + \int_{c_L}^{c_H} \int_0^1 \left[\lambda_0\,\left[v(\eta) - v_E\right] + \Delta\,\left[v(\eta) + 1\right] \right] \hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) \\
&+ \lambda_1 \int_{c_L}^{c_H} \int_0^1 \left[v(\eta) - v_E\right]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c).
\end{aligned}
$$

Therefore, the objective function is bounded by $\int_{\mathcal{C}_1} \sigma(c)\,\mathrm{d}c$. Substituting our candidate solution into the objective function achieves the upper bound in the first two cases:

1. If $\underline{c}(\tilde{\eta}_2) \leq c_L$, we have $\hat{c} = c_L$ and the measure of $\mathcal{C}_1$ is 0. The government's expected payoff from our candidate policy and its upper bound are both 0.

2. If $c_L < \underline{c}(\tilde{\eta}_2) \leq c_H$, we have

$$
\begin{aligned}
&\int_{c_L}^{\underline{c}(\tilde{\eta}_2)} \sigma(c)\,\mathrm{d}c - F(\underline{c}(\tilde{\eta}_2)) \int_{\tilde{\eta}_2}^1 \left[(\lambda_0 + \Delta)\,v(\eta) + \Delta \right] \mathrm{d}G(\eta) \\
= &\int_{c_L}^{\underline{c}(\tilde{\eta}_2)} \left\{ (\lambda_0 + \Delta) \int_{\tilde{\eta}_2}^1 \left[v(\eta) - v(\tilde{\eta}_2)\right]\,g(\eta)\,\mathrm{d}\eta + \left[c + \frac{F(c)}{f(c)} \right] \left[v(\tilde{\eta}_2) + \frac{\Delta}{\lambda_0}\,\left[v(\tilde{\eta}_2) + 1\right] \right] \right\} f(c)\,\mathrm{d}c \\
&- F(\underline{c}(\tilde{\eta}_2)) \int_{\tilde{\eta}_2}^1 \left[(\lambda_0 + \Delta)\,v(\eta) + \Delta \right] \mathrm{d}G(\eta)
\end{aligned}
$$

$$= \left\{-(\lambda_0+\Delta)\, v(\tilde{\eta}_2)\, [1-G(\tilde{\eta}_2)] + \underline{c}(\tilde{\eta}_2)\, \left[v(\tilde{\eta}_2) + \frac{\Delta}{\lambda_0}\, [v(\tilde{\eta}_2)+1]\right] - \Delta\, [1-G(\tilde{\eta}_2)]\right\} F(\underline{c}(\tilde{\eta}_2))$$
$$= 0,$$

establishing the optimality.

In case (3)(a)(iii), we set $\mathcal{C}_1 = [c_L, c_H]$ and $\mathcal{C}_2 = \emptyset$, and replace the thresholds $\tilde{\eta}_1$ and $\tilde{\eta}_2$ with $\eta_1(\breve{\eta}_2)$ and $\breve{\eta}_2$, respectively, where $\breve{\eta}_2$ is determined by $\underline{c}(\breve{\eta}_2) = c_H$, and $\eta_1(\eta_2)$ is obtained by solving $\Psi_2(\eta_1(\eta_2), \eta_2) = 0$. In this case, the nonnegativity of κ, χ, μ, ω, and ξ follows from their definitions directly. The desired nonnegativity of ϕ follows from

$$\iota(\breve{\eta}_1, \breve{\eta}_2) = -\frac{\Xi(\breve{\eta}_2)}{\lambda_0} \geq -\frac{\Xi(\tilde{\eta}_2)}{\lambda_0} = \iota(\tilde{\eta}_1, \tilde{\eta}_2) \geq 0,$$

where function Ξ is defined by (76). The first inequality follows from the monotonicity of Ξ and the fact that $\breve{\eta}_2 \geq \tilde{\eta}_2$. Due to the monotonicity of σ, its desired nonnegativity condition boils down to $\sigma(c_H) \geq 0$, which follows from

$$\sigma(c_H) = h_2(\breve{\eta}_2)\, f(c_H) \geq h_2(\tilde{\eta}_2)\, f(c_H) = 0,$$

where the function h_2 is defined in (81), and the inequality follows from the monotonicity of h_2, as established in Lemma 6. Because all the multipliers remain nonnegative, the objective function is bounded by $\int_{c_L}^{c_H} \sigma(c)\, \mathrm{d}c$, which is attained by our candidate solution, i.e.,

$$\begin{aligned}
&\int_{c_L}^{c_H} \sigma(c)\, \mathrm{d}c - \int_{\breve{\eta}_2}^{1} [(\lambda_0+\Delta)\, v(\eta) + \Delta]\, \mathrm{d}G(\eta) \\
=\ &\int_{c_L}^{c_H} \left\{(\lambda_0+\Delta) \int_{\breve{\eta}_2}^{1} [v(\eta) - v(\breve{\eta}_2)]\, g(\eta)\, \mathrm{d}\eta + \left[c + \frac{F(c)}{f(c)}\right] \left[v(\breve{\eta}_2) + \frac{\Delta}{\lambda_0}\, [v(\breve{\eta}_2)+1]\right]\right\} f(c)\, \mathrm{d}c \\
&- \int_{\breve{\eta}_2}^{1} [(\lambda_0+\Delta)\, v(\eta) + \Delta]\, \mathrm{d}G(\eta) \\
=\ &-(\lambda_0+\Delta)\, v(\breve{\eta}_2)\, [1-G(\breve{\eta}_2)] + c_H \left[v(\breve{\eta}_2) + \frac{\Delta}{\lambda_0}\, [v(\breve{\eta}_2)+1]\right] - \Delta\, [1-G(\breve{\eta}_2)] \\
=\ &0,
\end{aligned}$$

where the last step follows from $c_H = \lambda_0\, [1-G(\breve{\eta}_2)]$.

In case (3)(a)(iv), note that the integrand of the objective function (43) is bounded above

by

$$
\begin{aligned}
& v_E\,\alpha_1(c) + \int_0^1 \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta) + \lambda_1 \int_0^1 v(\eta)\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta) \\
\leq\ & v_E\,\alpha_1(c) + \int_{\bar{\eta}'}^1 \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta) + \lambda_1 \int_{\bar{\eta}}^1 v(\eta)\,\alpha_1(c)\,\mathrm{d}G(\eta) \\
=\ & \left[v_E + \lambda_1 \int_{\bar{\eta}}^1 v(\eta)\,\mathrm{d}G(\eta)\right]\,\alpha_1(c) + \int_{\bar{\eta}'}^1 \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta) \\
\leq\ & \int_{\bar{\eta}'}^1 \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta),
\end{aligned}
$$

where the first inequality follows from the definitions of $\bar{\eta}$ and $\bar{\eta}'$ and the upper bound of $\hat{\alpha}_3(c,\eta)$, while the second inequality follows from the given condition $h(\bar{\eta}) \leq 0$. This upper bound is attained by committing to no conditional approval and adopting the myopic threshold ($\hat{\eta}_2 = \bar{\eta}'$). When $c_H \leq \underline{c}(\bar{\eta}')$, this policy enables the government to induce firms of all types to participate. Therefore, the mechanism is optimal under the given conditions.

To prove the optimality of cases (3)(b)(i) and (3)(b)(ii), define

$$
\begin{aligned}
& \kappa \equiv -v(\acute{\eta}_2), \\
& \nu(c,\eta) \equiv (\lambda_0+\Delta)\,\left[v(\acute{\eta}_1) - v(\eta)\right]\,g(\eta)\,f(c), \quad \forall (c,\eta) \in [c_L,c_H]\times[0,\acute{\eta}_1], \\
& \mu(c,\eta) \equiv (\lambda_0+\Delta)\,\left[v(\eta) - v(\acute{\eta}_1)\right]\,g(\eta)\,f(c), \quad \forall (c,\eta) \in [c_L,c_H]\times[\acute{\eta}_1,1], \\
& \omega(c,\eta) \equiv \lambda_1\,\left[v(\acute{\eta}_2) - v(\eta)\right]\,g(\eta)\,f(c), \quad \forall (c,\eta) \in [c_L,c_H]\times[0,\acute{\eta}_2], \\
& \xi(c,\eta) \equiv \lambda_1\,\left[v(\eta) - v(\acute{\eta}_2)\right]\,g(\eta)\,f(c), \quad \forall (c,\eta) \in [c_L,c_H]\times[\acute{\eta}_2,1], \\
& \phi(c) \equiv \left\{-\lambda_1 \int_{\acute{\eta}_2}^1 \left[v(z) - v(\acute{\eta}_2)\right]\,g(z)\,\mathrm{d}z + v(\acute{\eta}_2) - v_E\right\}\,f(c), \quad \forall c \in [c_L,c_H], \\
& \sigma(c) \equiv \left\{v(\acute{\eta}_2)\,\left[c + \frac{F(c)}{f(c)}\right] + (\lambda_0+\Delta)\int_{\acute{\eta}_1}^1 \left[v(\eta) - v(\acute{\eta}_1)\right]\,g(\eta)\,\mathrm{d}\eta\right\}\,f(c), \quad \forall c \in \mathcal{C}_1, \\
& \beta(c) \equiv -\sigma(c), \quad \forall c \in \mathcal{C}_2.
\end{aligned}
$$

The nonnegativity of κ follows from the fact that $\acute{\eta}_2 \leq \bar{\eta}' \leq \bar{\eta}$. The nonnegativity of ν, μ, ω, and ξ follows form their respective definitions directly. The nonnegativity of ϕ follows from Lemma 6(2).

In the first two possible cases, we have:

1. if $\underline{c}(\acute{\eta}_1) < c_L$, we know that β is nonnegative on $[c_L,c_H]$. In this case, we have $\mathcal{C}_1 = \emptyset$ and $\mathcal{C}_2 = [c_L,c_H]$; and

2. if $c_L \leq \underline{c}(\acute{\eta}_1) \leq c_H$, we know that σ and β are nonnegative on $[c_L, \underline{c}(\acute{\eta}_1)]$ and $[\underline{c}(\acute{\eta}_1), c_H]$, respectively. In this case, we have $\mathcal{C}_1 = [c_L, \underline{c}(\acute{\eta}_1)]$ and $\mathcal{C}_2 = [\underline{c}(\acute{\eta}_1), c_H]$.

Therefore, the nonnegativity of σ and β holds automatically.

Because

$$-\kappa - \int_{c_L}^{c_H} \phi(c)\,\mathrm{d}c - \int_{c_L}^{c_H} \int_{\acute{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c = v_E,$$

$\forall (c,\eta) \in [c_L, c_H] \times [0, \acute{\eta}_1]$:

$$\lambda_0\,\phi(c)\,g(\eta) - \nu(c,\eta) + \lambda_0\,g(\eta) \int_{\acute{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z = [\lambda_0\,[v(\eta) - v_E] + \Delta\,[v(\eta) + 1]]\,g(\eta)\,f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\acute{\eta}_1, 1]$:

$$\lambda_0\,\phi(c)\,g(\eta) + \mu(c,\eta) + \lambda_0\,g(\eta) \int_{\acute{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z = [\lambda_0\,[v(\eta) - v_E] + \Delta\,[v(\eta) + 1]]\,g(\eta)\,f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [0, \acute{\eta}_2]$:

$$\lambda_1\,\phi(c)\,g(\eta) - \omega(c,\eta) + \lambda_1\,g(\eta) \int_{\acute{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z = \lambda_1\,[v(\eta) - v_E]\,g(\eta)\,f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\acute{\eta}_2, 1]$:

$$\lambda_1\,\phi(c)\,g(\eta) + \xi(c,\eta) + \lambda_1\,g(\eta) \int_{\acute{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z = \lambda_1\,[v(\eta) - v_E]\,g(\eta)\,f(c),$$

$\forall c \in \mathcal{C}_1$:

$$\begin{aligned} &- \int_{c_L}^{c} \phi(y)\,\mathrm{d}y - c\,\phi(c) + \sigma(c) - c \int_{\acute{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta \\ &\qquad - \int_{c_L}^{c} \int_{\acute{\eta}_2}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - \int_{\acute{\eta}_1}^{1} \mu(c,\eta)\,\mathrm{d}\eta = v_E\,[c\,f(c) + F(c)], \end{aligned}$$

$\forall c \in \mathcal{C}_2$:

$$- \int_{c_L}^{c} \phi(y)\,\mathrm{d}y - c\,\phi(c) - \beta(c) - c \int_{\acute{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta$$

$$-\int_{c_L}^{c}\int_{\acute{\eta}_2}^{1}\xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y-\int_{\acute{\eta}_1}^{1}\mu(c,\eta)\,\mathrm{d}\eta=v_E\,[c\,f(c)+F(c)],$$

we have

$$\begin{aligned}
&\kappa\,[-\Pi(c_H)]+\int_{\mathcal{C}_1}\sigma(c)\,\tau(c)\,\mathrm{d}c+\int_{\mathcal{C}_2}\beta(c)\,[-\tau(c)]\,\mathrm{d}c+\int_{c_L}^{c_H}\int_{0}^{\acute{\eta}_1}\nu(c,\eta)\,[-\hat{\alpha}_2(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c\\
&+\int_{c_L}^{c_H}\int_{\acute{\eta}_1}^{1}\mu(c,\eta)\,[\hat{\alpha}_2(c,\eta)-\tau(c)]\,\mathrm{d}\eta\,\mathrm{d}c+\int_{c_L}^{c_H}\int_{0}^{\acute{\eta}_2}\omega(c,\eta)\,[-\hat{\alpha}_3(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c\\
&+\int_{c_L}^{c_H}\int_{\acute{\eta}_2}^{1}\xi(c,\eta)\,\left[\hat{\alpha}_3(c,\eta)-\Pi(c_H)-\int_{c}^{c_H}\tau(y)\,\mathrm{d}y-c\,\tau(c)+\int_{0}^{1}[\lambda_0\,\hat{\alpha}_2(c,z)+\lambda_1\,\hat{\alpha}_3(c,z)]\,\mathrm{d}G(z)\right]\mathrm{d}\eta\,\mathrm{d}c\\
&+\int_{c_L}^{c_H}\phi(c)\,\left[-\Pi(c_H)-\int_{c}^{c_H}\tau(y)\,\mathrm{d}y-c\,\tau(c)+\lambda_0\int_{0}^{1}\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta)+\lambda_1\int_{0}^{1}\hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta)\right]\mathrm{d}c\\
=\;&\left[-\kappa-\int_{c_L}^{c_H}\phi(c)\,\mathrm{d}c-\int_{c_L}^{c_H}\int_{\acute{\eta}_2}^{1}\xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\right]\Pi(c_H)\\
&+\int_{c_L}^{c_H}\int_{0}^{\acute{\eta}_1}\left[\lambda_0\,\phi(c)\,g(\eta)-\nu(c,\eta)+\lambda_0\,g(\eta)\int_{\acute{\eta}_2}^{1}\xi(c,z)\,\mathrm{d}z\right]\hat{\alpha}_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\\
&+\int_{c_L}^{c_H}\int_{\acute{\eta}_1}^{1}\left[\lambda_0\,\phi(c)\,g(\eta)+\mu(c,\eta)+\lambda_0\,g(\eta)\int_{\acute{\eta}_2}^{1}\xi(c,z)\,\mathrm{d}z\right]\hat{\alpha}_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\\
&+\int_{c_L}^{c_H}\int_{0}^{\acute{\eta}_2}\left[\lambda_1\,\phi(c)\,g(\eta)-\omega(c,\eta)+\lambda_1\,g(\eta)\int_{\acute{\eta}_2}^{1}\xi(c,z)\,\mathrm{d}z\right]\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\\
&+\int_{c_L}^{c_H}\int_{\acute{\eta}_2}^{1}\left[\lambda_1\,\phi(c)\,g(\eta)+\xi(c,\eta)+\lambda_1\,g(\eta)\int_{\acute{\eta}_1}^{1}\xi(c,z)\,\mathrm{d}z\right]\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\\
&+\int_{\mathcal{C}_1}\left[-\int_{c_L}^{c}\phi(y)\,\mathrm{d}y-c\,\phi(c)+\sigma(c)-\int_{\acute{\eta}_1}^{1}\mu(c,\eta)\,\mathrm{d}\eta-c\int_{\acute{\eta}_2}^{1}\xi(c,\eta)\,\mathrm{d}\eta-\int_{c_L}^{c}\int_{\acute{\eta}_2}^{1}\xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\tau(c)\,\mathrm{d}c\\
&+\int_{\mathcal{C}_2}\left[-\int_{c_L}^{c}\phi(y)\,\mathrm{d}y-c\,\phi(c)-\beta(c)-\int_{\acute{\eta}_1}^{1}\mu(c,\eta)\,\mathrm{d}\eta-c\int_{\acute{\eta}_2}^{1}\xi(c,\eta)\,\mathrm{d}\eta-\int_{c_L}^{c}\int_{\acute{\eta}_2}^{1}\xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\tau(c)\,\mathrm{d}c\\
=\;&v_E\,\Pi(c_H)+v_E\int_{c_L}^{c_H}\left[c+\frac{F(c)}{f(c)}\right]\tau(c)\,f(c)\,\mathrm{d}c+\int_{c_L}^{c_H}\int_{0}^{1}[\lambda_0\,[v(\eta)-v_E]+\Delta\,[v(\eta)+1]]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c)\\
&+\lambda_1\int_{c_L}^{c_H}\int_{0}^{1}[v(\eta)-v_E]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c).
\end{aligned}$$

Therefore, the objective function is bounded by $\int_{\mathcal{C}_1}\sigma(c)\,\mathrm{d}c$. Substituting our candidate solution into the objective function achieves the upper bound in the first two cases:

1. If $\underline{c}(\acute{\eta}_1)\leq c_L$, we have $\hat{c}=c_L$ and the measure of $\mathcal{C}_1$ is 0. The government's expected payoff from our candidate policy and its upper bound are both 0.

2. If $c_L < \underline{c}(\acute{\eta}_1) \leq c_H$, we have

$$\begin{aligned}
&\int_{c_L}^{\underline{c}(\acute{\eta}_1)} \sigma(c)\,\mathrm{d}c - F(\underline{c}(\acute{\eta}_1)) \int_{\acute{\eta}_1}^{1} \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\mathrm{d}G(\eta) \\
=\ &\int_{c_L}^{\underline{c}(\acute{\eta}_1)} \left\{ v(\acute{\eta}_2)\,\left[c+\frac{F(c)}{f(c)}\right] + (\lambda_0+\Delta)\int_{\acute{\eta}_1}^{1}\left[v(\eta)-v(\acute{\eta}_1)\right]\,g(\eta)\,\mathrm{d}\eta \right\} f(c)\,\mathrm{d}c \\
&- F(\underline{c}(\acute{\eta}_1)) \int_{\acute{\eta}_1}^{1} \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\mathrm{d}G(\eta) \\
=\ &\left[\lambda_0\,v(\acute{\eta}_2) - (\lambda_0+\Delta)\,v(\acute{\eta}_1) - \Delta\right]\,\left[1-G(\acute{\eta}_1)\right]\,F(\underline{c}(\acute{\eta}_1)) \\
=\ &0,
\end{aligned}$$

where the last step follows from $\Psi_1(\acute{\eta}_1, \acute{\eta}_2) = 0$.

In case (3)(b)(iii), we set $\mathcal{C}_1 = [c_L, c_H]$ and $\mathcal{C}_2 = \emptyset$, and replace the thresholds $\acute{\eta}_1$ and $\acute{\eta}_2$ with $\breve{\eta}_1$ and $\eta_2(\breve{\eta}_1)$, respectively, where $\breve{\eta}_1$ is determined by $\underline{c}(\breve{\eta}_1) = c_H$, and $\eta_2(\eta_1)$ is obtained by solving $\Psi_1(\eta_1, \eta_2(\eta_1)) = 0$. In this case, the nonnegativity of κ, ν, μ, ω, and ξ follows from their definitions directly. The desired nonnegativity of ϕ follows from

$$h(\breve{\eta}_2) \leq h(\acute{\eta}_2) \leq 0.$$

The nonnegativity of σ follows from its monotonicity and

$$\sigma(c_H) = h_1(\breve{\eta}_1)\,f(c_H) \geq h_1(\acute{\eta}_1)\,f(c_H) = 0,$$

where the function h_1 is defined in (82), and the inequality follows from the monotonicity of function h_1, as established in Lemma 6. Because all the multipliers remain nonnegative, the objective function is bounded by $\int_{c_L}^{c_H} \sigma(c)\,\mathrm{d}c$, which is attained by our candidate solution, i.e.,

$$\begin{aligned}
&\int_{c_L}^{c_H} \sigma(c)\,\mathrm{d}c - \int_{\breve{\eta}_1}^{1} \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\mathrm{d}G(\eta) \\
=\ &\int_{c_L}^{c_H} \left\{ (\lambda_0+\Delta)\int_{\breve{\eta}_1}^{1}\left[v(\eta)-v(\breve{\eta}_1)\right]\,g(\eta)\,\mathrm{d}\eta + \left[c+\frac{F(c)}{f(c)}\right]\,\left[v(\breve{\eta}_1)+\frac{\Delta}{\lambda_0}\,\left[v(\breve{\eta}_1)+1\right]\right] \right\} f(c)\,\mathrm{d}c \\
&- \int_{\breve{\eta}_1}^{1} \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\mathrm{d}G(\eta) \\
=\ &-(\lambda_0+\Delta)\,v(\breve{\eta}_1)\,\left[1-G(\breve{\eta}_1)\right] + c_H\,\left[v(\breve{\eta}_1)+\frac{\Delta}{\lambda_0}\,\left[v(\breve{\eta}_1)+1\right]\right] - \Delta\,\left[1-G(\breve{\eta}_1)\right]
\end{aligned}$$

$= 0,$

where the last step follows from $c_H = \lambda_0 \left[1 - G(\breve{\eta}_1)\right]$.

In case (3)(b)(iv), similar to case (3)(a)(iv), committing to no conditional approval and adopting the myopic threshold ($\hat{\eta}_1 = \bar{\eta}'$) enables the government to induce all firms to participate. Therefore, this mechanism is optimal under the given conditions.

It remains to prove the optimality of the last case. For (4)(a)(i) and (4)(a)(ii), define

$$
\begin{aligned}
&\kappa \equiv -v(\dot{\eta}_2), \\
&\nu(c,\eta) \equiv (\lambda_0 + \Delta)\, [v(\dot{\eta}_1) - v(\eta)]\, g(\eta)\, f(c), \quad \forall (c,\eta) \in [c_L, c_H] \times [0, \dot{\eta}_1], \\
&\mu(c,\eta) \equiv (\lambda_0 + \Delta)\, [v(\eta) - v(\dot{\eta}_1)]\, g(\eta)\, f(c), \quad \forall (c,\eta) \in [c_L, c_H] \times [\dot{\eta}_1, 1], \\
&\omega(c,\eta) \equiv \lambda_1\, [v(\dot{\eta}_2) - v(\eta)]\, g(\eta)\, f(c), \quad \forall (c,\eta) \in [c_L, c_H] \times [0, \dot{\eta}_2], \\
&\xi(c,\eta) \equiv \lambda_1\, [v(\eta) - v(\dot{\eta}_2)]\, g(\eta)\, f(c), \quad \forall (c,\eta) \in [c_L, c_H] \times [\dot{\eta}_2, 1], \\
&\psi(c) \equiv \left\{ \lambda_1 \int_{\dot{\eta}_2}^{1} [v(z) - v(\dot{\eta}_2)]\, g(z)\, \mathrm{d}z - v(\dot{\eta}_2) + v_E \right\} f(c), \\
&\sigma(c) \equiv \left\{ v(\dot{\eta}_2) \left[c + \frac{F(c)}{f(c)} \right] + (\lambda_0 + \Delta) \int_{\dot{\eta}_1}^{1} [v(\eta) - v(\dot{\eta}_1)]\, g(\eta)\, \mathrm{d}\eta \right. \\
&\qquad \left. + \lambda_1 \int_{\dot{\eta}_2}^{1} [v(z) - v(\dot{\eta}_2)]\, g(z)\, \mathrm{d}z - v(\dot{\eta}_2) + v_E \right\} f(c), \quad \forall c \in \mathcal{C}_1, \\
&\beta(c) \equiv -\sigma(c), \quad \forall c \in \mathcal{C}_2.
\end{aligned}
$$

The nonnegativity of κ follows from the fact that $\dot{\eta}_2 \leq \bar{\eta}$. The nonnegativity of ν, μ, ω, and ξ follows form their respective definitions directly. The nonnegativity of ψ follows from Lemma 8(1).

In the first two possible cases, we have:

1. if $\bar{c}(\dot{\eta}_1, \dot{\eta}_2) < c_L$, we know that β is nonnegative on $[c_L, c_H]$. In this case, we have $\mathcal{C}_1 = \emptyset$ and $\mathcal{C}_2 = [c_L, c_H]$; and

2. if $c_L \leq \bar{c}(\dot{\eta}_1, \dot{\eta}_2) \leq c_H$, we know that σ and β are nonnegative on $[c_L, \bar{c}(\dot{\eta}_1, \dot{\eta}_2)]$ and $[\bar{c}(\dot{\eta}_1, \dot{\eta}_2), c_H]$, respectively. In this case, we have $\mathcal{C}_1 = [c_L, \bar{c}(\dot{\eta}_1, \dot{\eta}_2)]$ and $\mathcal{C}_2 = [\bar{c}(\dot{\eta}_1, \dot{\eta}_2), c_H]$.

Therefore, the nonnegativity of σ and β holds automatically.

Because

$$-\kappa + \int_{c_L}^{c_H} \psi(c)\,\mathrm{d}c - \int_{c_L}^{c_H} \int_{\dot{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c = v_E,$$

$\forall (c,\eta) \in [c_L, c_H] \times [0, \dot{\eta}_1]$:

$$-\lambda_0\,\psi(c)\,g(\eta) - \nu(c,\eta) + \lambda_0\,g(\eta) \int_{\dot{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z = [\lambda_0\ [v(\eta) - v_E] + \Delta\ [v(\eta) + 1]]\ g(\eta)\,f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\dot{\eta}_1, 1]$:

$$-\lambda_0\,\psi(c)\,g(\eta) + \mu(c,\eta) + \lambda_0\,g(\eta) \int_{\dot{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z = [\lambda_0\ [v(\eta) - v_E] + \Delta\ [v(\eta) + 1]]\ g(\eta)\,f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [0, \dot{\eta}_2]$:

$$-\lambda_1\,\psi(c)\,g(\eta) - \omega(c,\eta) + \lambda_1\,g(\eta) \int_{\dot{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\,f(c),$$

$\forall (c,\eta) \in [c_L, c_H] \times [\dot{\eta}_2, 1]$:

$$-\lambda_1\,\psi(c)\,g(\eta) + \xi(c,\eta) + \lambda_1\,g(\eta) \int_{\dot{\eta}_2}^{1} \xi(c,z)\,\mathrm{d}z = \lambda_1\ [v(\eta) - v_E]\ g(\eta)\,f(c),$$

$\forall c \in [c_L, \hat{c}]$:

$$\int_{c_L}^{c} \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) + \sigma(c) - c \int_{\dot{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c} \int_{\dot{\eta}_2}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - \int_{\dot{\eta}_1}^{1} \mu(c,\eta)\,\mathrm{d}\eta = v_E\,[c\,f(c) + F(c)],$$

$\forall c \in [\hat{c}, c_H]$:

$$\int_{c_L}^{c} \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) - \beta(c) - c \int_{\dot{\eta}_2}^{1} \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^{c} \int_{\dot{\eta}_2}^{1} \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - \int_{\dot{\eta}_1}^{1} \mu(c,\eta)\,\mathrm{d}\eta = v_E\,[c\,f(c) + F(c)],$$

we have

$$
\begin{aligned}
&\kappa\,[-\Pi(c_H)] + \int_{\mathcal{C}_1} \sigma(c)\,\tau(c)\,\mathrm{d}c + \int_{\mathcal{C}_2} \beta(c)\,[-\tau(c)]\,\mathrm{d}c + \int_{c_L}^{c_H}\int_0^{\dot\eta_1} \nu(c,\eta)\,[-\hat\alpha_2(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\dot\eta_1}^1 \mu(c,\eta)\,[\hat\alpha_2(c,\eta)-\tau(c)]\,\mathrm{d}\eta\,\mathrm{d}c + \int_{c_L}^{c_H}\int_0^{\dot\eta_2} \omega(c,\eta)\,[-\hat\alpha_3(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\dot\eta_2}^1 \xi(c,\eta)\left[\hat\alpha_3(c,\eta)-\Pi(c_H)-\int_c^{c_H}\tau(y)\,\mathrm{d}y - c\,\tau(c) + \int_0^1 [\lambda_0\,\hat\alpha_2(c,z)+\lambda_1\,\hat\alpha_3(c,z)]\,\mathrm{d}G(z)\right]\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \psi(c)\left[\Pi(c_H) + \int_c^{c_H}\tau(y)\,\mathrm{d}y + c\,\tau(c) - \lambda_0\int_0^1 \hat\alpha_2(c,\eta)\,\mathrm{d}G(\eta) - \lambda_1\int_0^1 \hat\alpha_3(c,\eta)\,\mathrm{d}G(\eta) - \tau(c)\right]\mathrm{d}c \\
=\ &\left[-\kappa + \int_{c_L}^{c_H}\psi(c)\,\mathrm{d}c - \int_{c_L}^{c_H}\int_{\dot\eta_2}^1 \xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\right]\Pi(c_H) \\
&+ \int_{c_L}^{c_H}\int_0^{\dot\eta_1}\left[-\lambda_0\,\psi(c)\,g(\eta) - \nu(c,\eta) + \lambda_0\,g(\eta)\int_{\dot\eta_2}^1 \xi(c,z)\,\mathrm{d}z\right]\hat\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\dot\eta_1}^1\left[-\lambda_0\,\psi(c)\,g(\eta) + \mu(c,\eta) + \lambda_0\,g(\eta)\int_{\dot\eta_2}^1 \xi(c,z)\,\mathrm{d}z\right]\hat\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_0^{\dot\eta_2}\left[-\lambda_1\,\psi(c)\,g(\eta) - \omega(c,\eta) + \lambda_1\,g(\eta)\int_{\dot\eta_2}^1 \xi(c,z)\,\mathrm{d}z\right]\hat\alpha_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\dot\eta_2}^1\left[-\lambda_1\,\psi(c)\,g(\eta) + \xi(c,\eta) + \lambda_1\,g(\eta)\int_{\dot\eta_1}^1 \xi(c,z)\,\mathrm{d}z\right]\hat\alpha_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{\mathcal{C}_1}\left[\int_{c_L}^c \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) + \sigma(c) - \int_{\dot\eta_1}^1 \mu(c,\eta)\,\mathrm{d}\eta - c\int_{\dot\eta_2}^1 \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^c\int_{\dot\eta_2}^1 \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\tau(c)\,\mathrm{d}c \\
&+ \int_{\mathcal{C}_2}\left[\int_{c_L}^c \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) - \beta(c) - \int_{\dot\eta_1}^1 \mu(c,\eta)\,\mathrm{d}\eta - c\int_{\dot\eta_2}^1 \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^c\int_{\dot\eta_2}^1 \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\tau(c)\,\mathrm{d}c \\
=\ &v_E\,\Pi(c_H) + v_E\int_{c_L}^{c_H}\left[c + \frac{F(c)}{f(c)}\right]\tau(c)\,f(c)\,\mathrm{d}c + \int_{c_L}^{c_H}\int_0^1 [\lambda_0\,[v(\eta)-v_E] + \Delta\,[v(\eta)+1]]\,\hat\alpha_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) \\
&+ \lambda_1\int_{c_L}^{c_H}\int_0^1 [v(\eta)-v_E]\,\hat\alpha_3(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c).
\end{aligned}
$$

Therefore, the objective function is bounded by $\int_{\mathcal{C}_1}\sigma(c)\,\mathrm{d}c$. Substituting our candidate solution into the objective function achieves the upper bound in the first two cases, which follows from:

1. if $\bar c(\dot\eta_1,\dot\eta_2) \le c_L$, we have $\hat c = c_L$ and the measure of $\mathcal{C}_1$ is 0. The government's expected payoff from our candidate policy and its upper bound are both 0;

2. if $c_L < \bar c(\dot\eta_1,\dot\eta_2) \le c_H$, the government's expected payoff from our candidate policy

attains this upper bound, i.e.,

$$\begin{aligned}
&\int_{c_L}^{\bar{c}(\dot{\eta}_1,\dot{\eta}_2)} \sigma(c)\,\mathrm{d}c - \left[v_E + \int_{\dot{\eta}_1}^{1} \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\mathrm{d}G(\eta) + \lambda_1 \int_{\dot{\eta}_2}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] F(\bar{c}(\dot{\eta}_1,\dot{\eta}_2)) \\
&= \{v(\dot{\eta}_2)\,[1+\lambda_0\,[1-G(\dot{\eta}_1)]+\lambda_1\,[1-G(\dot{\eta}_2)]] - \Delta\,[1-G(\dot{\eta}_1)] \\
&\quad -(\lambda_0+\Delta)\,v(\dot{\eta}_1)\,[1-G(\dot{\eta}_1)] - \lambda_1\,v(\dot{\eta}_2)\,[1-G(\dot{\eta}_2)] - v(\dot{\eta}_2)\}\,F(\bar{c}(\dot{\eta}_1,\dot{\eta}_2)) \\
&= [\lambda_0\,v(\dot{\eta}_2) - \Delta - (\lambda_0+\Delta)\,v(\dot{\eta}_1)]\,[1-G(\dot{\eta}_1)]\,F(\bar{c}(\dot{\eta}_1,\dot{\eta}_2)) \\
&= 0,
\end{aligned}$$

where the last step follows from $\Psi_1(\dot{\eta}_1,\dot{\eta}_2)=0$.

In case (4)(a)(iii), we set $\mathcal{C}_1=[c_L,c_H]$ and $\mathcal{C}_2=\emptyset$, and replace the thresholds $\dot{\eta}_1$ and $\dot{\eta}_2$ with $\breve{\eta}_1$ and $\breve{\eta}_2$, respectively, where the two thresholds are determined by $\bar{c}(\breve{\eta}_1,\breve{\eta}_2)=c_H$ and $\Psi_1(\breve{\eta}_1,\breve{\eta}_2)=0$. In this case, the nonnegativity of κ, ν, μ, ω, and ξ follows from their definitions directly. The nonnegativity of ψ follows as: if $h(\bar{\eta})\le 0$, we have $h(\breve{\eta}_2)\ge h(\breve{\eta}_2)=0$; otherwise, we have $h(\breve{\eta})\ge h(\bar{\eta})>0$. The nonnegativity of σ follows from the fact that for any $c\in[c_L,c_H]$, we have

$$\sigma(c)\ge\sigma(c_H)=t_1(\breve{\eta}_2)\,f(c_H)\ge t_1(\dot{\eta}_2)\,f(c_H)=0,$$

where the function t_1 is defined in (83), and the inequality follows from the monotonicity of t_1, as established in Lemma 8. Because all the multipliers remain nonnegative, the objective function is bounded by $\int_{c_L}^{c_H}\sigma(c)\,\mathrm{d}c$, which is attained by our candidate solution, i.e.,

$$\begin{aligned}
&\int_{c_L}^{c_H} \sigma(c)\,\mathrm{d}c - \left[v_E + \int_{\breve{\eta}_1}^{1} \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\mathrm{d}G(\eta) + \lambda_1 \int_{\breve{\eta}_2}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] F(c_H) \\
&= c_H\,v(\breve{\eta}_2) + (\lambda_0+\Delta)\int_{\breve{\eta}_1}^{1}[v(\eta)-v(\breve{\eta}_1)]\,g(\eta)\,\mathrm{d}\eta + \lambda_1\int_{\breve{\eta}_2}^{1}[v(z)-v(\breve{\eta}_2)]\,g(z)\,\mathrm{d}z - v(\breve{\eta}_2) + v_E \\
&\quad - \left[v_E + \int_{\breve{\eta}_1}^{1} \left[(\lambda_0+\Delta)\,v(\eta)+\Delta\right]\mathrm{d}G(\eta) + \lambda_1 \int_{\breve{\eta}_2}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] \\
&= c_H\,v(\breve{\eta}_2) - (\lambda_0+\Delta)\,v(\breve{\eta}_1)\,[1-G(\breve{\eta}_1)] - \lambda_1\,v(\breve{\eta}_2)\,[1-G(\breve{\eta}_2)] - v(\breve{\eta}_2) - \Delta\,[1-G(\breve{\eta}_1)] \\
&= [c_H - \lambda_1\,[1-G(\breve{\eta}_2)]-1]\left[v(\breve{\eta}_1)+\frac{\Delta}{\lambda_0}\,[v(\breve{\eta}_1)+1]\right] - (\lambda_0+\Delta)\,v(\breve{\eta}_1)\,[1-G(\breve{\eta}_1)] - \Delta\,[1-G(\breve{\eta}_1)] \\
&= 0,
\end{aligned}$$

where the last equality follows from

$$c_H = 1 + \lambda_0 \left[1 - G(\breve{\eta}_1)\right] + \lambda_1 \left[1 - G(\breve{\eta}_2)\right].$$

In case (4)(a)(iv), we replace the two thresholds $\dot{\eta}_1$ and $\dot{\eta}_2$ with $\bar{\eta}'$ and $\bar{\eta}$, respectively. The desired nonnegativity condition for ψ boils down to

$$\lambda_1 \int_{\bar{\eta}}^{1} v(z)\,\mathrm{d}G(z) + v_E \geq 0 \quad \Leftrightarrow \quad h(\bar{\eta}) \geq 0.$$

Suppose $h(\bar{\eta}) < 0$. We have

$$\check{c}_1 > \bar{c}(\breve{\eta}_1, \breve{\eta}_2) \geq 1 + \lambda_0 \left[1 - G(\bar{\eta}')\right] + \lambda_1 \left[1 - G(\bar{\eta})\right]$$

or

$$\check{c}_2 > \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1) \geq 1 + (\lambda_0 + \lambda_1) \left[1 - G(\eta^{\dagger})\right] \geq 1 + (\lambda_0 + \lambda_1) \left[1 - G(\bar{\eta}')\right],$$

which contradicts the condition $c_H \leq \bar{c}(\bar{\eta}', \bar{\eta})$. Therefore, we know that $h(\bar{\eta}) > 0$, guaranteeing the nonnegativity of ψ and σ, where the latter one follows from

$$\sigma(c) = \left\{ (\lambda_0 + \Delta) \int_{\bar{\eta}'}^{1} \left[v(\eta) - v(\bar{\eta}')\right] g(\eta)\,\mathrm{d}\eta + \underbrace{\lambda_1 \int_{\bar{\eta}}^{1} v(z)\,g(z)\,\mathrm{d}z + v_E}_{=h(\bar{\eta})\geq 0} \right\} f(c) \geq 0.$$

The candidate primal solution attains the upper bound $\int_{c_L}^{c_H} \sigma(c)\,\mathrm{d}c$, i.e.,

$$\begin{aligned}
&\int_{c_L}^{c_H} \sigma(c)\,\mathrm{d}c - \left[v_E + \int_{\bar{\eta}'}^{1} \left[(\lambda_0 + \Delta)\,v(\eta) + \Delta\right] \mathrm{d}G(\eta) + \lambda_1 \int_{\bar{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] F(c_H) \\
&= (\lambda_0 + \Delta) \int_{\bar{\eta}'}^{1} \left[v(\eta) - v(\bar{\eta}')\right] g(\eta)\,\mathrm{d}\eta + \lambda_1 \int_{\bar{\eta}}^{1} v(z)\,g(z)\,\mathrm{d}z + v_E \\
&\quad - \left[v_E + \int_{\bar{\eta}'}^{1} \left[(\lambda_0 + \Delta)\,v(\eta) + \Delta\right] \mathrm{d}G(\eta) + \lambda_1 \int_{\bar{\eta}}^{1} v(\eta)\,\mathrm{d}G(\eta)\right] \\
&= -\left[(\lambda_0 + \Delta)\,v(\bar{\eta}') + \Delta\right] \left[1 - G(\bar{\eta}')\right] \\
&= 0,
\end{aligned}$$

where the last step follows from the definition of $\bar{\eta}'$.

It remains to prove the optimality in case (4)(b). For the first two subcases, define

$$\kappa \equiv -\frac{1}{\lambda_0}\left[(\lambda_0+\Delta)\, v(\grave{\eta}_2)+\Delta\right],$$
$$\chi(c,\eta) \equiv (\lambda_0+\Delta)\left[v(\grave{\eta}_2)-v(\eta)\right] g(\eta)\, f(c), \quad \forall (c,\eta)\in[c_L,c_H]\times[0,\grave{\eta}_2],$$
$$\mu(c,\eta) = (\lambda_0+\Delta)\left[v(\eta)-v(\grave{\eta}_2)\right] g(\eta)\, f(c), \quad \forall (c,\eta)\in[c_L,c_H]\times[\grave{\eta}_2,1],$$
$$\omega(c,\eta) = \left\{(\lambda_0+\lambda_1+\Delta)\left[v(\grave{\eta}_2)-v(\eta)\right]+\frac{\lambda_1}{\lambda_0}\Delta\left[v(\grave{\eta}_2)+1\right]\right\} g(\eta)\, f(c), \quad \forall (c,\eta)\in[c_L,c_H]\times[0,\grave{\eta}_1],$$
$$\xi(c,\eta) \equiv \begin{cases} -\left\{(\lambda_0+\lambda_1+\Delta)\left[v(\grave{\eta}_2)-v(\eta)\right]+\frac{\lambda_1}{\lambda_0}\Delta\left[v(\grave{\eta}_2)+1\right]\right\} g(\eta)\, f(c), & \forall (c,\eta)\in[c_L,c_H]\times[\grave{\eta}_1,\grave{\eta}_2], \\ \left\{\lambda_1\left[v(\eta)-v(\grave{\eta}_2)\right]-\frac{\lambda_1}{\lambda_0}\Delta\left[v(\grave{\eta}_2)+1\right]\right\} g(\eta)\, f(c), & \forall (c,\eta)\in[c_L,c_H]\times[\grave{\eta}_2,1], \end{cases}$$
$$\psi(c) \equiv -\iota(\grave{\eta}_1,\grave{\eta}_2)\, f(c), \quad \forall c\in[c_L,c_H],$$
$$\sigma(c) \equiv \left\{\left[v(\grave{\eta}_2)+\frac{\Delta}{\lambda_0}\left[v(\grave{\eta}_2)+1\right]\right]\left[c+\frac{F(c)}{f(c)}\right] + (\lambda_0+\Delta)\int_{\grave{\eta}_2}^{1}\left[v(\eta)-v(\grave{\eta}_2)\right] g(\eta)\,\mathrm{d}\eta+\frac{\psi(c)}{f(c)}\right\} f(c), \quad \forall c\in\mathcal{C}_1,$$
$$\beta(c) \equiv -\sigma(c), \quad \forall c\in\mathcal{C}_2.$$

The nonnegativity of κ follows from the fact that $\grave{\eta}_2\le\eta^\dagger\le\bar{\eta}'$. The nonnegativity of χ and μ follows from their respective definitions directly. The nonnegativity of ω and ξ follows from the relationship between $\grave{\eta}_1$ and $\grave{\eta}_2$ (i.e., $\Psi_2(\grave{\eta}_1,\grave{\eta}_2)=0$, implying $\omega(c,\grave{\eta}_1)=\xi(c,\grave{\eta}_1)=0$ for any $c\in[c_L,c_H]$) and their monotonicity in η. The nonnegativity of ψ follows from Lemma 8(2).

In the first two possible cases, we have:

1. if $\bar{c}(\grave{\eta}_1,\grave{\eta}_1)<c_L$, we know that β is nonnegative on $[c_L,c_H]$. In this case, we have $\mathcal{C}_1=\emptyset$ and $\mathcal{C}_2=[c_L,c_H]$; and

2. if $c_L\le\bar{c}(\grave{\eta}_1,\grave{\eta}_1)\le c_H$, we know that σ and β are nonnegative on $[c_L,\bar{c}(\grave{\eta}_1,\grave{\eta}_1)]$ and $[\bar{c}(\grave{\eta}_1,\grave{\eta}_1),c_H]$, respectively. In this case, we have $\mathcal{C}_1=[c_L,\bar{c}(\grave{\eta}_1,\grave{\eta}_1)]$ and $\mathcal{C}_2=[\bar{c}(\grave{\eta}_1,\grave{\eta}_1),c_H]$.

Therefore, the nonnegativity of σ and β holds automatically.

Because

$$-\kappa+\int_{c_L}^{c_H}\psi(c)\,\mathrm{d}c-\int_{c_L}^{c_H}\int_{\grave{\eta}_1}^{1}\xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c=v_E,$$

$\forall(c,\eta) \in [c_L, c_H] \times [0, \grave{\eta}_2]$:

$$-\lambda_0\,\psi(c)\,g(\eta) - \chi(c,\eta) + \lambda_0\,g(\eta)\int_{\grave{\eta}_1}^{1}\xi(c,z)\,\mathrm{d}z = [\lambda_0\,[v(\eta) - v_E] + \Delta\,[v(\eta)+1]]\,g(\eta)\,f(c),$$

$\forall(c,\eta) \in [c_L, c_H] \times [\grave{\eta}_2, 1]$:

$$-\lambda_0\,\psi(c)\,g(\eta) + \mu(c,\eta) + \lambda_0\,g(\eta)\int_{\grave{\eta}_1}^{1}\xi(c,z)\,\mathrm{d}z = [\lambda_0\,[v(\eta) - v_E] + \Delta\,[v(\eta)+1]]\,g(\eta)\,f(c),$$

$\forall(c,\eta) \in [c_L, c_H] \times [0, \grave{\eta}_1]$:

$$-\lambda_1\,\psi(c)\,g(\eta) - \omega(c,\eta) + \chi(c,\eta) + \lambda_1\,g(\eta)\int_{\grave{\eta}_1}^{1}\xi(c,z)\,\mathrm{d}z = \lambda_1\,[v(\eta) - v_E]\,g(\eta)\,f(c),$$

$\forall(c,\eta) \in [c_L, c_H] \times [\grave{\eta}_1, \grave{\eta}_2]$:

$$-\lambda_1\,\psi(c)\,g(\eta) + \chi(c,\eta) + \xi(c,\eta) + \lambda_1\,g(\eta)\int_{\grave{\eta}_1}^{1}\xi(c,z)\,\mathrm{d}z = \lambda_1\,[v(\eta) - v_E]\,g(\eta)\,f(c),$$

$\forall(c,\eta) \in [c_L, c_H] \times [\grave{\eta}_2, 1]$:

$$-\lambda_1\,\psi(c)\,g(\eta) + \xi(c,\eta) + \lambda_1\,g(\eta)\int_{\grave{\eta}_1}^{1}\xi(c,z)\,\mathrm{d}z = \lambda_1\,[v(\eta) - v_E]\,g(\eta)\,f(c),$$

$\forall c \in [c_L, \hat{c}]$:

$$\begin{aligned}&\int_{c_L}^{c}\psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) + \sigma(c) - \int_{\grave{\eta}_2}^{1}\mu(c,\eta)\,\mathrm{d}\eta\\ &\quad - \int_{c_L}^{c}\int_{\grave{\eta}_1}^{1}\xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - c\int_{\grave{\eta}_1}^{1}\xi(c,\eta)\,\mathrm{d}\eta = v_E\,[c\,f(c) + F(c)],\end{aligned}$$

$\forall c \in [\hat{c}, c_H]$:

$$\begin{aligned}&\int_{c_L}^{c}\psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) - \beta(c) - \int_{\grave{\eta}_2}^{1}\mu(c,\eta)\,\mathrm{d}\eta\\ &\quad - \int_{c_L}^{c}\int_{\grave{\eta}_1}^{1}\xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y - c\int_{\grave{\eta}_1}^{1}\xi(c,\eta)\,\mathrm{d}\eta = v_E\,[c\,f(c) + F(c)],\end{aligned}$$

we have

$$
\begin{aligned}
&\kappa\,[-\Pi(c_H)] + \int_{\mathcal{C}_1} \sigma(c)\,\tau(c)\,\mathrm{d}c + \int_{\mathcal{C}_2} \beta(c)\,[-\tau(c)]\,\mathrm{d}c + \int_{c_L}^{c_H}\int_0^{\grave{\eta}_2} \chi(c,\eta)\,[\hat{\alpha}_3(c,\eta) - \hat{\alpha}_2(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\grave{\eta}_2}^1 \mu(c,\eta)\,[\hat{\alpha}_2(c,\eta) - \tau(c)]\,\mathrm{d}\eta\,\mathrm{d}c + \int_{c_L}^{c_H}\int_0^{\grave{\eta}_1} \omega(c,\eta)\,[-\hat{\alpha}_3(c,\eta)]\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\grave{\eta}_1}^1 \xi(c,\eta)\left[\hat{\alpha}_3(c,\eta) - \Pi(c_H) - \int_c^{c_H} \tau(y)\,\mathrm{d}y - c\,\tau(c) + \int_0^1 [\lambda_0\,\hat{\alpha}_2(c,z) + \lambda_1\,\hat{\alpha}_3(c,z)]\,\mathrm{d}G(z)\right]\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H} \psi(c)\left[\Pi(c_H) + \int_c^{c_H} \tau(y)\,\mathrm{d}y + c\,\tau(c) - \lambda_0\int_0^1 \hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta) - \lambda_1\int_0^1 \hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta) - \tau(c)\right]\mathrm{d}c \\
=\ &\left[-\kappa + \int_{c_L}^{c_H} \psi(c)\,\mathrm{d}c - \int_{c_L}^{c_H}\int_{\grave{\eta}_1}^1 \xi(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c\right]\Pi(c_H) \\
&+ \int_{c_L}^{c_H}\int_0^{\grave{\eta}_2}\left[-\lambda_0\,\psi(c)\,g(\eta) - \chi(c,\eta) + \lambda_0\,g(\eta)\int_{\grave{\eta}_1}^1 \xi(c,z)\,\mathrm{d}z\right]\hat{\alpha}_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\grave{\eta}_2}^1\left[-\lambda_0\,\psi(c)\,g(\eta) + \mu(c,\eta) + \lambda_0\,g(\eta)\int_{\grave{\eta}_1}^1 \xi(c,z)\,\mathrm{d}z\right]\hat{\alpha}_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_0^{\grave{\eta}_1}\left[-\lambda_1\,\psi(c)\,g(\eta) - \omega(c,\eta) + \chi(c,\eta) + \lambda_1\,g(\eta)\int_{\grave{\eta}_1}^1 \xi(c,z)\,\mathrm{d}z\right]\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\grave{\eta}_1}^{\grave{\eta}_2}\left[-\lambda_1\,\psi(c)\,g(\eta) + \chi(c,\eta) + \xi(c,\eta) + \lambda_1\,g(\eta)\int_{\grave{\eta}_1}^1 \xi(c,z)\,\mathrm{d}z\right]\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{c_L}^{c_H}\int_{\grave{\eta}_2}^1\left[-\lambda_1\,\psi(c)\,g(\eta) + \xi(c,\eta) + \lambda_1\,g(\eta)\int_{\grave{\eta}_1}^1 \xi(c,z)\,\mathrm{d}z\right]\hat{\alpha}_3(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&+ \int_{\mathcal{C}_1}\left[\int_{c_L}^c \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) + \sigma(c) - \int_{\grave{\eta}_2}^1 \mu(c,\eta)\,\mathrm{d}\eta - c\int_{\grave{\eta}_1}^1 \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^c\int_{\grave{\eta}_1}^1 \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\tau(c)\,\mathrm{d}c \\
&+ \int_{\mathcal{C}_2}\left[\int_{c_L}^c \psi(y)\,\mathrm{d}y + (c-1)\,\psi(c) - \beta(c) - \int_{\grave{\eta}_2}^1 \mu(c,\eta)\,\mathrm{d}\eta - c\int_{\grave{\eta}_1}^1 \xi(c,\eta)\,\mathrm{d}\eta - \int_{c_L}^c\int_{\grave{\eta}_1}^1 \xi(y,\eta)\,\mathrm{d}\eta\,\mathrm{d}y\right]\tau(c)\,\mathrm{d}c \\
=\ &v_E\,\Pi(c_H) + v_E\int_{c_L}^{c_H}\left[c + \frac{F(c)}{f(c)}\right]\tau(c)\,f(c)\,\mathrm{d}c + \int_{c_L}^{c_H}\int_0^1 [\lambda_0\,[v(\eta) - v_E] + \Delta\,[v(\eta)+1]]\,\hat{\alpha}_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c) \\
&+ \lambda_1\int_{c_L}^{c_H}\int_0^1 [v(\eta) - v_E]\,\hat{\alpha}_3(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c).
\end{aligned}
$$

Therefore, the objective function is bounded by $\int_{\mathcal{C}_1} \sigma(c)\,\mathrm{d}c$. Substituting our candidate solution into the objective function achieves the upper bound in the first two cases, which follows from:

1. if $\bar{c}(\grave{\eta}_1, \grave{\eta}_1) \leq c_L$, we have $\hat{c} = c_L$ and the measure of $\mathcal{C}_1$ is 0. The government's expected payoff from our candidate policy and its upper bound are both 0;

2. if $c_L < \bar{c}(\grave{\eta}_1, \grave{\eta}_1) \leq c_H$, we have

$$\begin{aligned}
&\int_{c_L}^{\bar{c}(\grave{\eta}_1,\grave{\eta}_1)} \sigma(c)\,\mathrm{d}c - \left[v_E + \int_{\grave{\eta}_1}^{1} \left[(\lambda_0+\lambda_1+\Delta)\,v(\eta)+\Delta\right]\mathrm{d}G(\eta)\right] F(\bar{c}(\grave{\eta}_1,\grave{\eta}_1)) \\
&= \left\{\bar{c}(\grave{\eta}_1,\grave{\eta}_1)\,\left[v(\grave{\eta}_2)+\frac{\Delta}{\lambda_0}\,[v(\grave{\eta}_2)+1]\right] + (\lambda_0+\Delta)\int_{\grave{\eta}_2}^{1}[v(\eta)-v(\grave{\eta}_2)]\;g(\eta)\,\mathrm{d}\eta\right. \\
&\quad \left. -\iota(\grave{\eta}_1,\grave{\eta}_2) - \left[v_E + \int_{\grave{\eta}_1}^{1}\left[(\lambda_0+\lambda_1+\Delta)\,v(\eta)+\Delta\right]\mathrm{d}G(\eta)\right]\right\} F(\bar{c}(\grave{\eta}_1,\grave{\eta}_1)) \\
&= 0.
\end{aligned}$$

In case (4)(b)(iii), we set $\mathcal{C}_1 = [c_L, c_H]$ and $\mathcal{C}_2 = \emptyset$, and replace the thresholds $\acute{\eta}_1$ and $\acute{\eta}_2$ with $\breve{\eta}_1$ and $\breve{\eta}_2$, respectively, where the two thresholds are determined by $\bar{c}(\breve{\eta}_1, \breve{\eta}_1) = c_H$ and $\Psi_2(\breve{\eta}_1, \breve{\eta}_2) = 0$. In this case, the nonnegativity of κ, χ, μ, ω, and ξ follows from their definitions directly. Given the monotonicity of the function Ξ (defined by (76)), the desired nonnegativity of ψ follows from the following argument:

- If $h(\bar{\eta}) \leq 0$ and $\check{c}_1 > \bar{c}(\check{\eta}_1, \check{\eta}_2)$, we have

$$\iota(\breve{\eta}_1, \breve{\eta}_2) = -\frac{\Xi(\breve{\eta}_2)}{\lambda_0} \leq -\frac{\Xi(\check{\eta}_2)}{\lambda_0} \leq= 0.$$

  The inequalities above follow from the fact that $\breve{\eta}_2 \leq \check{\eta}_2$ and $\Xi(\check{\eta}_2) \geq 0$, as established in Lemma 5.

- If $h(\bar{\eta}) \leq 0$ and $\check{c}_2 > \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1)$, we have

$$\iota(\breve{\eta}_1, \breve{\eta}_2) = -\frac{\Xi(\breve{\eta}_2)}{\lambda_0} \leq -\frac{\Xi(\mathring{\eta}_2)}{\lambda_0} = \iota(\mathring{\eta}_1, \mathring{\eta}_2) = 0.$$

  The inequalities above follow from the fact that $\breve{\eta}_2 \leq \mathring{\eta}_2$.

- If $h(\bar{\eta}) > 0$, we have

$$\iota(\breve{\eta}_1, \breve{\eta}_2) = -\frac{\Xi(\breve{\eta}_2)}{\lambda_0} \leq -\frac{\Xi(\eta^\dagger)}{\lambda_0} \leq 0.$$

  The inequalities above follow from the fact that $\breve{\eta}_2 \leq \eta^\dagger$ and $\Xi(\eta^\dagger) = \lambda_0\, h(\eta^\dagger) \geq \lambda_0\, h(\bar{\eta}) > 0$.

The nonnegativity of σ follows from the fact that for any $c \in [c_L, c_H]$, we have

$$\sigma(c) \geq \sigma(c_H) = t_2(\breve{\eta}_2)\, f(c_H) \geq t_2(\grave{\eta}_2)\, f(c_H) = 0,$$

where the function t_2 is defined in (84), and the inequality follows from the monotonicity of t_2, as established in Lemma 8. Because all the multipliers remain nonnegative, the objective function is bounded by $\int_{c_L}^{c_H} \sigma(c)\,\mathrm{d}c$, which is attained by our candidate solution, i.e.,

$$\begin{aligned}
&\int_{c_L}^{c_H} \sigma(c)\,\mathrm{d}c - \left[v_E + \int_{\breve{\eta}_1}^{1} \left[\left(\lambda_0 + \lambda_1 + \Delta\right) v(\eta) + \Delta\right] \mathrm{d}G(\eta)\right] \\
&= \frac{\Delta}{\lambda_0}\left[v(\breve{\eta}_2) + 1\right]\left[c_H - \left[1 + \lambda_1\left[1 - G(\breve{\eta}_1)\right]\right]\right] + \left[c_H - \left(\lambda_0 + \lambda_1 + \Delta\right)\left[1 - G(\breve{\eta}_1)\right] - 1\right] v(\breve{\eta}_2) \\
&\quad - \Delta\left[1 - G(\breve{\eta}_1)\right] \\
&= \Delta\left[v(\breve{\eta}_2) + 1\right]\left[1 - G(\breve{\eta}_1)\right] - \Delta\left[1 - G(\breve{\eta}_1)\right] v(\breve{\eta}_2) - \Delta\left[1 - G(\breve{\eta}_1)\right] \\
&= 0,
\end{aligned}$$

where the second equality follows from

$$1 + \left(\lambda_0 + \lambda_1\right)\left[1 - G(\breve{\eta}_1)\right] = c_H.$$

When $h(\bar{\eta}) \leq 0$ and $\check{c}_2 > \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1)$, we have

$$c_H \geq \check{c}_2 > \bar{c}(\mathring{\eta}_1, \mathring{\eta}_1) \geq \bar{c}(\eta^\dagger, \eta^\dagger).$$

Therefore, it remains to prove cases (4)(b)(iv) and (4)(b)(v) for either (1) $h(\bar{\eta}) \leq 0$ and $\check{c}_1 > \bar{c}(\breve{\eta}_1, \breve{\eta}_2)$, or (2) $h(\bar{\eta}) > 0$.

The proof of case (4)(b)(iv) closely follows that of case (4)(a)(iv). The nonnegativity of ψ follows as: if $h(\bar{\eta}) \leq 0$ and $\check{c}_1 > \bar{c}(\breve{\eta}_1, \breve{\eta}_2)$, we have $h(\breve{\eta}_2) \geq h(\eta^\dagger) \geq h(\breve{\eta}_2) = 0$; otherwise, we have $h(\breve{\eta}) \geq h(\bar{\eta}) > 0$. The nonnegativity of σ follows from the fact that for any $c \in [c_L, c_H]$, we have

$$\sigma(c) \geq \sigma(c_H) = t_1(\breve{\eta}_2)\, f(c_H) \geq t_1(\eta^\dagger)\, f(c_H) \geq 0.$$

Since all the multipliers are nonnegative, optimality is established.

The proof of case (4)(b)(v) is identical to that of case (4)(a)(iv).

This completes the proof. □

Proof. Proof of Corollary 3. When $\Delta = 0$, we have

$$\Psi_1(\eta_1, \eta_2) = 0 \text{ or } \Psi_2(\eta_1, \eta_2) = 0 \quad \Rightarrow \quad \eta_1 = \eta_2.$$

Therefore, the optimal solution to the upper-bound problem (48) is reduced to

$$\hat{\alpha}_2(c,\eta) = \begin{cases} 1, & \forall (c,\eta) \in [c_L, \hat{c}) \times [\hat{\eta}, 1], \\ 0, & \text{otherwise}, \end{cases} \qquad \hat{\alpha}_3(c,\eta) = \begin{cases} \hat{\alpha}, & \forall (c,\eta) \in [c_L, \hat{c}) \times [\hat{\eta}, 1], \\ 0, & \text{otherwise}. \end{cases}$$

Given any feasible mechanism $\{\hat{\tau}(c), \hat{\alpha}_1(c), \hat{\alpha}_2(c,\eta), \hat{\alpha}_3(c,\eta)\}$, the following mechanism $\{\hat{\tau}(c), a_1(c), a_2(c,\eta)\}$ is feasible for the original problem (13) and yields the same objective function value, where a_1 and a_2 are constructed as follows:

$$a_1(c) = \begin{cases} \frac{\hat{\alpha}_1(c)}{\hat{\tau}(c)}, & \text{if } \hat{\tau}(c) > 0, \\ 0, & \text{if } \hat{\tau}(c) = 0, \end{cases} \qquad a_2(c,\eta) = \begin{cases} \frac{\hat{\alpha}_2(c,\eta)}{\hat{\tau}(c)}, & \text{if } \hat{\tau}(c) > 0, \\ 0, & \text{if } \hat{\tau}(c) = 0. \end{cases}$$

This completes the proof. □

Proof. Proof of Proposition 7. By construction, we have

$$\int_0^1 \hat{\tau}(c)\, [\lambda_0 + \lambda_1\, a_1'(c)]\; a_2'(c,\eta)\, \mathrm{d}G(\eta) = \int_0^1 [\lambda_0\, \hat{\alpha}_2(c,\eta) + \lambda_1\, \hat{\alpha}_3(c,\eta)]\, \mathrm{d}G(\eta), \quad \forall c \in [c_L, c_H].$$

Hence, under the following mechanism

$$\mathcal{M}_1 \equiv \left\{ \hat{\tau}(c) = \mathbb{1}_{c<\hat{c}},\; a_1'(c) = \hat{\alpha} \cdot \mathbb{1}_{c<\hat{c}},\; a_2'(c,\eta) = \mathbb{1}_{(c,\eta)\in[c_L,\hat{c}]\times[\eta_1,1]} \right\},$$

constraints (IC) and (IR) hold automatically. The remaining constraints (2), (3), and (4) hold trivially by their definitions. When $\eta_1 > \eta_2$, by definition of η_2, we have

$$\int_{\eta_2}^1 [(\lambda_0 + \lambda_1\, \hat{\alpha} + \Delta)\, v(\eta) + \Delta]\, \mathrm{d}G(\eta) \geq \int_{\eta_1}^1 [(\lambda_0 + \lambda_1\, \hat{\alpha} + \Delta)\, v(\eta) + \Delta]\, \mathrm{d}G(\eta).$$

Furthermore, the testing threshold satisfies

$$\hat{c} \leq \hat{\alpha} + (\lambda_0 + \lambda_1\, \hat{\alpha})\; [1 - G(\eta_1)] < \hat{\alpha} + (\lambda_0 + \lambda_1\, \hat{\alpha})\; [1 - G(\eta_2)] = c_2,$$

implying that the government is better off under mechanism

$$\mathcal{M}_2 \equiv \left\{ \tau''(c) = \mathbb{1}_{c\leq c_2},\; a_1''(c) = \hat{\alpha} \cdot \mathbb{1}_{c\leq c_2},\; a_2''(c,\eta) = \mathbb{1}_{(c,\eta)\in[c_L,c_2]\times[\eta_2,1]} \right\}.$$

It is straightforward to verify the feasibility of $\mathcal{M}_2$. Because U_R^* is the optimal value of the

objective function, the feasibility property implies that

$$\tilde{U}_R^* \leq U_R^*.$$

This completes the proof. □

11.5 Proofs for Section 5: either conditional approval or leniency unavailable

In this section, we provide the optimal mechanisms in closed form when either condition approval ($a_1(c) = 0$ for any $c \in [c_L, c_H]$) or leniency ($a_2(c, \eta) = \mathbb{1}_{\eta \in [\bar{\eta}', 1]}$ for any $(c, \eta) \in [c_L, c_H] \times [0, 1]$) is unavailable, for completeness.

With a slight abuse of notation, in this section, we introduce a new decision variable

$$\alpha_2(c, \eta) \equiv \tau(c)\, a_2(c, \eta), \quad \forall (c, \eta) \in [c_L, c_H] \times [0, 1].$$

Both parties' payoff functions boil down to

$$\hat{\Pi}(c, c') = -c\, \tau(c') + \lambda_0 \int_0^1 \alpha_2(c', \eta)\, \mathrm{d}G(\eta) \tag{85}$$

and

$$U_R(\tau, \alpha_2) = \int_{c_L}^{c_H} \int_0^1 \left[(\lambda_0 + \Delta)\, v(\eta) + \Delta\right]\, \alpha_2(c, \eta)\, \mathrm{d}G(\eta)\, \mathrm{d}F(c). \tag{86}$$

We redefine the government's utility function as

$$\tilde{v}(\eta) \equiv (\lambda_0 + \Delta)\, v(\eta) + \Delta, \quad \forall \eta \in [0, 1],$$

and denote $\tilde{v}(0)$ and $\mathbb{E}[\tilde{v}(\eta)]$ by $\tilde{v}_L$ and $\tilde{v}_E$, respectively. With a slight abuse of notation, in this section, we define

$$\bar{\eta} \equiv \inf \left\{\eta \in [0, 1] : \tilde{v}(\eta) \geq 0\right\}.$$

The government's problem can be formulated as follows:

$$\max_{\{\tau, \alpha_2\} \in \Omega_1'} U_R(\tau, \alpha_2), \tag{87}$$

where the feasible set Ω_1' is determined by (IC), (IR), (2), and

$$0 \leq \alpha_2(c,\eta) \leq \tau(c), \quad \forall (c,\eta) \in [c_L, c_H] \times [0,1]. \tag{88}$$

The optimal solution to this restricted problem remains a "take-it-or-leave-it" offer, as stated in the following proposition.

Proposition 9. *Consider the optimization problem* (87) *under Assumption 1. There exists a cost threshold* $\hat{c}' \in [c_L, c_H]$ *and a final approval threshold* $\hat{\eta}' \in [0,1]$ *such that it is optimal to request testing and grant final approval as follows:*

$$\tau^*(c) = \begin{cases} 1, & \forall c \in [c_L, \hat{c}'), \\ 0, & \forall c \in [\hat{c}', c_H], \end{cases} \qquad \alpha_2^*(c,\eta) = \begin{cases} 1, & \forall (c,\eta) \in [c_L, \hat{c}') \times [\hat{\eta}', 1], \\ 0, & \textit{otherwise.} \end{cases}$$

The values of $\hat{c}'$ *and* $\hat{\eta}'$ *depend on model parameters in the following cases.*

(1) If $\tilde{v}_L \frac{F(\underline{c}(0))}{f(\underline{c}(0))} + \lambda_0 \tilde{v}_E > 0$, *we have* $\hat{c}' = c_H$ *and* $\hat{\eta}' = \min\{\underline{c}^{-1}(c_H), \bar{\eta}\}$.

(2) If $\tilde{v}_L \frac{F(\underline{c}(0))}{f(\underline{c}(0))} + \lambda_0 \tilde{v}_E \leq 0$, *there exists a unique* $\tilde{\eta} \in [0, \bar{\eta}]$ *such that*

$$\tilde{v}(\tilde{\eta}) \frac{F(\underline{c}(\tilde{\eta}))}{f(\underline{c}(\tilde{\eta}))} + \lambda_0 \int_{\tilde{\eta}}^1 \tilde{v}(y)\, g(y)\, \mathrm{d}y = 0.$$

(a) If $c_L > \underline{c}(\tilde{\eta})$, *we have* $\hat{c}' = c_L$ *and* $\hat{\eta}' = 1$.

(b) If $c_L \leq \underline{c}(\tilde{\eta}) \leq c_H$, *we have* $\hat{c}' = \underline{c}(\tilde{\eta})$ *and* $\hat{\eta}' = \tilde{\eta}$.

(c) If $\underline{c}(\bar{\eta}) < c_H < \underline{c}(\tilde{\eta})$, *we have* $\hat{c}' = c_H$ *and* $\hat{\eta}' = \underline{c}^{-1}(c_H)$.

(d) If $c_H \leq \underline{c}(\bar{\eta})$, *we have* $\hat{c}' = c_H$ *and* $\hat{\eta}' = \bar{\eta}$.

Proof. Proof of Proposition 9. It is straightforward to verify that our candidate solutions in all the cases are feasible.

When $c_H \leq \underline{c}(\bar{\eta})$, our candidate solution attains the upper bound of the objective function $\int_{\bar{\eta}}^1 \tilde{v}(\eta)\, \mathrm{d}G(\eta)$, with all constraints holding automatically. In the remainder of the proof, we focus on scenarios where $c_H > \underline{c}(\bar{\eta})$.

Invoking envelope theorem, the incentive compatibility constraint can be replaced by (62) and

$$c\,\tau(c) - \lambda_0 \int_0^1 \alpha_2(c,\eta)\, \mathrm{d}G(\eta) + \Pi(c_H) + \int_c^{c_H} \tau(y)\, \mathrm{d}y = 0, \quad \forall c \in [c_L, c_H]. \tag{89}$$

When $c_H > \underline{c}(\bar{\eta})$, to prove the optimality in cases 1 and 2(3), define

$$
\begin{aligned}
&\kappa = -\tilde{v}(\underline{c}^{-1}(c_H)),\\
&\nu(c,\eta) = \lambda_0 \left[\tilde{v}(\underline{c}^{-1}(c_H)) - \tilde{v}(\eta)\right] f(c)\, g(\eta), && \forall (c,\eta) \in [c_L, c_H] \times [0, \underline{c}^{-1}(c_H)],\\
&\mu(c,\eta) = \lambda_0 \left[\tilde{v}(\eta) - \tilde{v}(\underline{c}^{-1}(c_H))\right] f(c)\, g(\eta), && \forall (c,\eta) \in [c_L, c_H] \times [\underline{c}^{-1}(c_H), 1],\\
&\phi(c) = -\tilde{v}(\underline{c}^{-1}(c_H))\, f(c), && \forall c \in [c_L, c_H],\\
&\sigma(c) = \left[\tilde{v}(\underline{c}^{-1}(c_H)) \left[(c - c_H) + \frac{F(c)}{f(c)}\right] + \lambda_0 \int_{\underline{c}^{-1}(c_H)}^{1} \tilde{v}(y)\, g(y)\, \mathrm{d}y\right] f(c), && \forall c \in [c_L, c_H].
\end{aligned}
$$

The monotonicity of $\underline{c}$ implies that

$$
c_H > \underline{c}(\bar{\eta}) \quad \Leftrightarrow \quad \underline{c}^{-1}(c_H) < \bar{\eta},
$$

guaranteeing the nonnegativity of κ. The nonnegativity of ν and μ follows from their definitions directly. Because $\sigma(c)/f(c)$ is decreasing, it suffices to verify the nonnegativity of $\sigma(c_H)$. When $\tilde{v}_L \frac{F(\underline{c}(0))}{f(\underline{c}(0))} + \lambda_0 \tilde{v}_E > 0$, we have

$$
\sigma(c_H) = \underbrace{\left[\tilde{v}(\underline{c}^{-1}(c_H)) \frac{1}{f(c_H)} + \lambda_0 \int_{\underline{c}^{-1}(c_H)}^{1} \tilde{v}(y)\, g(y)\, \mathrm{d}y\right]}_{>0} f(c_H) \geq 0.
$$

When $\tilde{v}_L \frac{F(\underline{c}(0))}{f(\underline{c}(0))} + \lambda_0 \tilde{v}_E \leq 0$ and $c_H < \underline{c}(\tilde{\eta})$, we have $\underline{c}^{-1}(c_H) > \tilde{\eta}$, and hence,

$$
\begin{aligned}
\sigma(c_H) = &\left[\tilde{v}(\underline{c}^{-1}(c_H)) \frac{F(c_H)}{f(c_H)} + \lambda_0 \int_{\underline{c}^{-1}(c_H)}^{1} \tilde{v}(y)\, g(y)\, \mathrm{d}y\right] f(c_H)\\
\geq &\left[\tilde{v}(\tilde{\eta}) \frac{F(\underline{c}(\tilde{\eta}))}{f(\underline{c}(\tilde{\eta}))} + \lambda_0 \int_{\tilde{\eta}}^{1} \tilde{v}(y)\, g(y)\, \mathrm{d}y\right] f(c_H)\\
= &\, 0,
\end{aligned}
$$

where the inequality follows from the monotonicity of function

$$
\Upsilon(\eta) \equiv \tilde{v}(\eta) \frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))} + \lambda_0 \int_{\eta}^{1} \tilde{v}(y)\, g(y)\, \mathrm{d}y,
$$

i.e.,

$$\Upsilon'(\eta) = \tilde{v}'(\eta)\,\frac{F(\underline{c}(\eta))}{f(\underline{c}(\eta))} - \lambda_0\,\tilde{v}(\eta)\,g(\eta)\,\underbrace{\frac{\mathrm{d}}{\mathrm{d}c}\left[c + \frac{F(c)}{f(c)}\right]}_{\geq 0 \text{ by Assumption 1}}\Bigg|_{c=\underline{c}(\eta)} \geq 0.$$

Because

$$\begin{aligned}
&\kappa\,[-\Pi(c_H)] + \int_{c_L}^{c_H} \phi(c)\,\left[c\,\tau(c) - \lambda_0\int_0^1 \alpha_2(c,\eta)\,\mathrm{d}G(\eta) + \Pi(c_H) + \int_c^{c_H}\tau(y)\,\mathrm{d}y\right]\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H}\sigma(c)\,\tau(c)\,\mathrm{d}c - \int_{c_L}^{c_H}\int_0^{\underline{c}^{-1}(c_H)}\nu(c,\eta)\,\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c + \int_{c_L}^{c_H}\int_{\underline{c}^{-1}(c_H)}^1 \mu(c,\eta)\,\left[\alpha_2(c,\eta) - \tau(c)\right]\mathrm{d}\eta\,\mathrm{d}c \\
=\;& \left[-\kappa + \int_{c_L}^{c_H}\phi(c)\,\mathrm{d}c\right]\,\Pi(c_H) + \int_{c_L}^{c_H}\int_0^{\underline{c}^{-1}(c_H)}\left[-\lambda_0\,\phi(c)\,g(\eta) - \nu(c,\eta)\right]\,\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H}\int_{\underline{c}^{-1}(c_H)}^1\left[-\lambda_0\,\phi(c)\,g(\eta) + \mu(c,\eta)\right]\,\alpha_2(c,\eta)\,\mathrm{d}\eta\,\mathrm{d}c \\
&\quad + \int_{c_L}^{c_H}\left[\sigma(c) + c\,\phi(c) + \int_{c_L}^{c}\phi(y)\,\mathrm{d}y - \int_0^1\mu(c,\eta)\,\mathrm{d}\eta\right]\,\tau(c)\,\mathrm{d}c \\
=\;& \lambda_0\int_{c_L}^{c_H}\int_0^1 \tilde{v}(\eta)\,\alpha_2(c,\eta)\,\mathrm{d}G(\eta)\,\mathrm{d}F(c),
\end{aligned}$$

the objective function is bounded by $\int_{c_L}^{c_H}\sigma(c)\,\mathrm{d}c$. Substituting the decision variables with our candidate solution attains this upper bound, which follows from

$$\begin{aligned}
&\int_{c_L}^{c_H}\sigma(c)\,\mathrm{d}c - \lambda_0\int_{\underline{c}^{-1}(c_H)}^1\tilde{v}(\eta)\,\mathrm{d}G(\eta) \\
=\;& \int_{c_L}^{c_H}\left[\tilde{v}(\underline{c}^{-1}(c_H))\,\left[(c - c_H) + \frac{F(c)}{f(c)}\right] + \lambda_0\int_{\underline{c}^{-1}(c_H)}^1\tilde{v}(y)\,g(y)\,\mathrm{d}y\right]\,f(c)\,\mathrm{d}c - \lambda_0\int_{\underline{c}^{-1}(c_H)}^1\tilde{v}(\eta)\,\mathrm{d}G(\eta) \\
=\;& \tilde{v}(\underline{c}^{-1}(c_H))\int_{c_L}^{c_H}\left[(c - c_H)\,f(c) + F(c)\right]\mathrm{d}c \\
=\;& 0.
\end{aligned}$$

To prove the optimality in the remaining two subcases ((2)(a) and (2)(b)), define

$$\begin{aligned}
&\kappa = -\tilde{v}(\tilde{\eta}), \\
&\nu(c,\eta) = \lambda_0\,\left[\tilde{v}(\tilde{\eta}) - \tilde{v}(\eta)\right]\,f(c)\,g(\eta), && \forall (c,\eta)\in[c_L,c_H]\times[0,\tilde{\eta}], \\
&\mu(c,\eta) = \lambda_0\,\left[\tilde{v}(\eta) - \tilde{v}(\tilde{\eta})\right]\,f(c)\,g(\eta), && \forall (c,\eta)\in[c_L,c_H]\times[\tilde{\eta},1],
\end{aligned}$$

$$\phi(c) = -\tilde{v}(\tilde{\eta})\, f(c), \qquad \forall c \in [c_L, c_H],$$
$$\sigma(c) = \left[\tilde{v}(\tilde{\eta}) \left[[c - \underline{c}(\tilde{\eta})] + \frac{F(c)}{f(c)}\right] + \lambda_0 \int_{\tilde{\eta}}^{1} \tilde{v}(y)\, g(y)\, \mathrm{d}y\right] f(c), \qquad \forall c \in \mathcal{C}_1,$$
$$\beta(c) = -\left[\tilde{v}(\tilde{\eta}) \left[[c - \underline{c}(\tilde{\eta})] + \frac{F(c)}{f(c)}\right] + \lambda_0 \int_{\tilde{\eta}}^{1} \tilde{v}(y)\, g(y)\, \mathrm{d}y\right] f(c), \quad \forall c \in \mathcal{C}_2.$$

When $c_L > \underline{c}(\tilde{\eta})$, we know that β is nonnegative on $[c_L, c_H]$. In this case, we have $\mathcal{C}_1 = \emptyset$ and $\mathcal{C}_2 = [c_L, c_H]$. When $c_L \leq \underline{c}(\tilde{\eta}) \leq c_H$, we know that σ and β are nonnegative on $[c_L, \underline{c}(\tilde{\eta})]$ and $[\underline{c}(\tilde{\eta}), c_H]$, respectively. In this case, we have $\mathcal{C}_1 = [c_L, \underline{c}(\tilde{\eta})]$ and $\mathcal{C}_2 = [\underline{c}(\tilde{\eta}), c_H]$.This guarantees the nonnegativity of all the multipliers (except ϕ, which corresponds to an equality constraint).

Because

$$\begin{aligned}
&\kappa\, [-\Pi(c_H)] + \int_{c_L}^{c_H} \phi(c) \left[c\,\tau(c) - \lambda_0 \int_0^1 \alpha_2(c,\eta)\, \mathrm{d}G(\eta) + \Pi(c_H) + \int_c^{c_H} \tau(y)\, \mathrm{d}y\right] \mathrm{d}c + \int_{\mathcal{C}_1} \sigma(c)\, \tau(c)\, \mathrm{d}c \\
&\quad - \int_{\mathcal{C}_2} \beta(c)\, \tau(c)\, \mathrm{d}c - \int_{c_L}^{c_H} \int_0^{\tilde{\eta}} \nu(c,\eta)\, \alpha_2(c,\eta)\, \mathrm{d}\eta\, \mathrm{d}c + \int_{c_L}^{c_H} \int_{\tilde{\eta}}^1 \mu(c,\eta)\, [\alpha_2(c,\eta) - \tau(c)]\, \mathrm{d}\eta\, \mathrm{d}c \\
= &\left[-\kappa + \int_{c_L}^{c_H} \phi(c)\, \mathrm{d}c\right] \Pi(c_H) + \int_{c_L}^{c_H} \int_0^{\tilde{\eta}} [-\lambda_0\, \phi(c)\, g(\eta) - \nu(c,\eta)]\, \alpha_2(c,\eta)\, \mathrm{d}\eta\, \mathrm{d}c \\
&\quad + \int_{c_L}^{c_H} \int_{\tilde{\eta}}^1 [-\lambda_0\, \phi(c)\, g(\eta) + \mu(c,\eta)]\, \alpha_2(c,\eta)\, \mathrm{d}\eta\, \mathrm{d}c \\
&\quad + \int_{\mathcal{C}_1} \left[\sigma(c) + c\,\phi(c) + \int_{c_L}^{c} \phi(y)\, \mathrm{d}y - \int_0^1 \mu(c,\eta)\, \mathrm{d}\eta\right] \tau(c)\, \mathrm{d}c \\
&\quad + \int_{\mathcal{C}_2} \left[-\beta(c) + c\,\phi(c) + \int_{c_L}^{c} \phi(y)\, \mathrm{d}y - \int_0^1 \mu(c,\eta)\, \mathrm{d}\eta\right] \tau(c)\, \mathrm{d}c \\
= &\lambda_0 \int_{c_L}^{c_H} \int_0^1 \tilde{v}(\eta)\, \alpha_2(c,\eta)\, \mathrm{d}G(\eta)\, \mathrm{d}F(c),
\end{aligned}$$

the objective function is bounded by $\int_{\mathcal{C}_1} \sigma(c)\, \mathrm{d}c$.

1. In case (2)(a), the government's payoff generated by our candidate solution and its upper bound are both 0.

2. In case (2)(b), we have

$$\int_{c_L}^{\underline{c}(\tilde{\eta})} \sigma(c)\, \mathrm{d}c - \lambda_0\, F(\underline{c}(\tilde{\eta})) \int_{\tilde{\eta}}^1 \tilde{v}(\eta)\, \alpha_2(c,\eta)\, \mathrm{d}G(\eta)$$

$$
\begin{aligned}
&= \int_{c_L}^{\underline{c}(\tilde{\eta})} \left[\tilde{v}(\tilde{\eta}) \left[[c - \underline{c}(\tilde{\eta})] + \frac{F(c)}{f(c)} \right] + \lambda_0 \int_{\tilde{\eta}}^{1} \tilde{v}(y)\, g(y)\, \mathrm{d}y \right] f(c)\, \mathrm{d}c \\
&\quad - \lambda_0\, F(\underline{c}(\tilde{\eta})) \int_{\tilde{\eta}}^{1} \tilde{v}(\eta)\, \alpha_2(c, \eta)\, \mathrm{d}G(\eta) \\
&= \tilde{v}(\tilde{\eta}) \int_{c_L}^{\underline{c}(\tilde{\eta})} [[c - \underline{c}(\tilde{\eta})]\, f(c) + F(c)]\, \mathrm{d}c \\
&= 0,
\end{aligned}
$$

establishing the optimality.

This completes the proof. □

When leniency is unavailable, in which case we assume that the government will adopt the myopic threshold $\bar{\eta}'$ (defined in (55)), both parties' payoff functions are reduced to

$$
\hat{\Pi}(c, c') = \alpha_1(c') - c\, \tau(c') + [\lambda_0\, \tau(c') + \lambda_1\, \alpha_1(c')]\ \bar{G}(\bar{\eta}')
$$

and

$$
U_R = \int_{c_L}^{c_H} \left[v_E\, \alpha_1(c) + [\lambda_0\, \tau(c) + \lambda_1\, \alpha_1(c)] \underbrace{\int_{\bar{\eta}'}^{1} v(\eta)\, \mathrm{d}G(\eta)}_{\equiv z_1} + \tau(c) \underbrace{\Delta \int_{\bar{\eta}'}^{1} [v(\eta) + 1]\, \mathrm{d}G(\eta)}_{\equiv z_2} \right] \mathrm{d}F(c),
$$

where $\bar{G}$ denotes the complementary cumulative distribution function of G. By envelope theorem, we have

$$
\alpha_1(c) = \frac{1}{1 + \lambda_1\, \bar{G}(\bar{\eta}')} \left[\Pi(c_H) + \int_{c}^{c_H} \tau(y)\, \mathrm{d}y + (c - \lambda_0\, \bar{G}(\bar{\eta}'))\, \tau(c) \right].
$$

After eliminating α_1, the optimization problem can be reformulated as follows:

$$
\max_{\{\tau, \alpha_1\} \in \Omega_2'} \ \frac{v_E + \lambda_1\, z_1}{1 + \lambda_1\, \bar{G}(\bar{\eta}')}\, \Pi(c_H) + \int_{c_L}^{c_H} w(c)\, \tau(c)\, \mathrm{d}F(c), \tag{90}
$$

where the virtual valuation function w is given by

$$
w(c) = \frac{v_E + \lambda_1\, z_1}{1 + \lambda_1\, \bar{G}(\bar{\eta}')} \left[c + \frac{F(c)}{f(c)} - \lambda_0\, \bar{G}(\bar{\eta}') \right] + \lambda_0\, z_1 + z_2,
$$

and the feasible set Ω_2' is determined by (2), (62), (64) and

$$0 \leq \frac{1}{1+\lambda_1 \bar{G}(\bar{\eta}')} \left[\Pi(c_H) + \int_c^{c_H} \tau(y)\, \mathrm{d}y + (c - \lambda_0 \bar{G}(\bar{\eta}'))\, \tau(c) \right] \leq \tau(c), \quad \forall (c,\eta) \in [c_L, c_H] \times [0,1]. \tag{91}$$

Assumption 1 guarantees that the virtual valuation $w(c)$ is non-increasing in c when $v_E + \lambda_1 z_1$ is negative, and non-decreasing otherwise. Moreover, w changes its sign at most once as c increases in $[c_L, c_H]$. Therefore, when $v_E + \lambda_1 z_1 < 0$, w is nonnegative if $c < c_0$, and non-positive if $c_0 < c$, where the threshold c_0 is defined by

$$c_0 \equiv \sup \{ c \in [c_L, c_H] : w(c) \geq 0 \}. \tag{92}$$

To characterize the optimal mechanism for the government's problem without leniency (90), we make the following assumption.

Assumption 2. *The density function $f(c)$ is differentiable and non-increasing.*

Assumption 2 implies that the virtual valuation function $c + F(c)/f(c)$ is increasing.

The optimal mechanism is summarized in the following proposition.

Proposition 10. *Under Assumption 2, the optimal solution α_1^* and τ^* for the optimization problem* (90) *is as follows:*

(1) if $v_E + \lambda_1 z_1 \geq 0$, then $\alpha_1^*(c) = 1$ and $\tau^*(c) = 1$ for any $c \in [c_L, c_H]$;

(2) if $v_E + \lambda_1 z_1 < 0$ and $c_H \leq \lambda_0 \bar{G}(\bar{\eta}')$, then $\alpha_1^*(c) = 0$ and $\tau^*(c) = 1$ for any $c \in [c_L, c_H]$;

(3) if $v_E + \lambda_1 z_1 < 0$, $c_H > \lambda_0 \bar{G}(\bar{\eta}')$, and $c_0 \geq \lambda_0 \bar{G}(\bar{\eta}')$, then

$$\alpha_1^*(c) = \begin{cases} \dfrac{c_0 - \lambda_0 \bar{G}(\bar{\eta}')}{1+\lambda_1 \bar{G}(\bar{\eta}')}, & \forall c \in [c_L, c_0], \\ 0, & \forall c \in (c_0, c_H], \end{cases} \quad \tau^*(c) = \begin{cases} 1, & \forall c \in [c_L, c_0], \\ 0, & \forall c \in (c_0, c_H]; \end{cases}$$

(4) if $v_E + \lambda_1 z_1 < 0$, $c_H > \lambda_0 \bar{G}(\bar{\eta}')$, and $c_0 < \bar{G}(\bar{\eta}')$, then

$$\alpha_1^*(c) = 0, \ \forall c \in [c_L, c_H], \quad \tau^*(c) = \begin{cases} 1, & \forall c \in [c_L, \lambda_0 \bar{G}(\bar{\eta}')], \\ 0, & \forall c \in (\lambda_0 \bar{G}(\bar{\eta}'), c_H], \end{cases}$$

where c_0 is defined by (92).

Proof. Proof of Proposition 10. In case (1), the coefficient of α_1 in the integrand of the objective function is nonnegative, i.e., $v_E+\lambda_1 z_1 \geq 0$. Therefore, it is optimal to set $\alpha_1^*(c) = 1$ for any $c \in [c_L, c_H]$. In this case, the government can induce firms of all types to participate, and hence, this policy is optimal.

In case (2), the coefficient of α_1 in the integrand of the objective function is negative, i.e., $v_E + \lambda_1 z_1 < 0$. Therefore, ignoring all constraints, it is optimal to set $\alpha_1^*(c) = 0$ for any $c \in [c_L, c_H]$. In this case, the government can still induce firms of all types to participate ($c_H \leq \lambda_0 \bar{G}(\bar{\eta}')$), and hence, this policy is optimal.

In case (3), we first ignore constraints (62) and (91), and solving the relaxed problem point-wise yields

$$\tau(c) = \begin{cases} 1, & \forall c \in [c_L, c_0], \\ 0, & \forall c \in (c_0, c_H], \end{cases}$$

satisfying (62) automatically. Plugging the expression of τ into (63) yields

$$\alpha_1(c) = \begin{cases} \dfrac{c_0 - \lambda_0 \bar{G}(\bar{\eta}')}{1 + \lambda_1 \bar{G}(\bar{\eta}')}, & \forall c \in [c_L, c_0], \\ 0, & \forall c \in (c_0, c_H]. \end{cases}$$

When $c_0 \geq \lambda_0 \bar{G}(\bar{\eta}')$, the feasibility constraint (91) holds automatically, and thus the solution above is feasible.

In case (4), it is straightforward to verify that our candidate solution is feasible. When $c_L \leq c_0 < \lambda_0 \bar{G}(\bar{\eta}')$, define

$$\begin{aligned}
\phi &= -\frac{1}{c_2 - c_1} \int_{c_L}^{c_1} w(y)\, f(y)\, \mathrm{d}y - \frac{v_E + \lambda_1 z_1}{1 + \lambda_1 \bar{G}(\bar{\eta}')}, \\
\nu(c) &= \frac{1}{(c_2 - c)^2} \left[\int_{c_L}^{c} w(y)\, f(y)\, \mathrm{d}y + (c_2 - c)\, w(c)\, f(c) \right], && \forall c \in [c_L, c_1], \\
\sigma(c) &= \frac{1}{c_2 - c_1} \int_{c_L}^{c_1} w(y)\, f(y)\, \mathrm{d}y + w(c)\, f(c), && \forall c \in [c_1, c_2], \text{ and} \\
\mu(c) &= -\frac{1}{c_2 - c_1} \int_{c_L}^{c_1} w(y)\, f(y)\, \mathrm{d}y - w(c)\, f(c), && \forall c \in [c_2, c_H],
\end{aligned}$$

where $c_2 = \lambda_0 \bar{G}(\bar{\eta}')$ and c_1 is determined by

$$\int_{c_L}^{c_1} w(y)\, f(y)\, \mathrm{d}y + (c_2 - c_1)\, w(c_2)\, f(c_2) = 0.$$

We first verify the monotonicity of $w(c)\,f(c)$ on $[c_L, c_2]$. By definition, we have

$$w(c)\,f(c) = \frac{1}{1+\lambda_1\,\bar{G}(\bar{\eta}')}\left[\left[\left(\lambda_0\,z_1+z_2\right)\left(1+\lambda_1\,\bar{G}(\bar{\eta}')\right)+\underbrace{\left(v_E+\lambda_1\,z_1\right)\left(c-\lambda_0\,\bar{G}(\bar{\eta}')\right)}_{\text{nonnegative and decreasing in } c}\right]\,f(c)+\underbrace{\left(v_E+\lambda_1\,z_1\right)}_{<0}\,F(c)\right].$$

Under Assumption 2, we know that $w(c)\,f(c)$ is decreasing in c.

Then we prove the unique existence of c_1 on $[c_L, c_2]$. Let

$$\psi(c) \equiv \int_{c_L}^{c} w(y)\,f(y)\,\mathrm{d}y + (c_2-c)\,w(c_2)\,f(c_2), \quad \forall c \in [c_L, c_2].$$

Because

$$\begin{aligned}
&\psi'(c) = w(c)\,f(c) - w(c_2)\,f(c_2) \geq 0,\\
&\psi(c_L) = (c_2-c_L)\,w(c_2)\,f(c_2) < 0, \text{ and}\\
&\psi(c_2) = \int_{c_L}^{c_2} w(y)\,f(y)\,\mathrm{d}y > 0,
\end{aligned}$$

the unique existence of c_1 is guaranteed. The first inequality follows from the monotonicity of $w(c)\,f(c)$, and the last inequality follows from the fact that under the mechanism granting no conditional approval, the government's expected payoff is strictly positive.

Next, we will prove that ϕ, ν, σ, and μ are all nonnegative. By definition of c_1, we have

$$\phi = \left(\lambda_0\,z_1+z_2\right)f(c_2) - \frac{\overbrace{v_E+\lambda_1\,z_1}^{<0}}{1+\lambda_1\,\bar{G}(\bar{\eta}')}\,\left[1-F(c_2)\right] \geq 0.$$

The nonnegativity of ν boils down to

$$h(c) \equiv \int_{c_L}^{c} w(y)\,f(y)\,\mathrm{d}y + (c_2-c)\,w(c)\,f(c) \geq 0, \quad \forall c \in [c_L, c_1].$$

Taking derivative yields

$$h'(c) = (c_2-c)\,\frac{\mathrm{d}}{\mathrm{d}c}\left(w(c)\,f(c)\right) \leq 0, \quad \forall c \in [c_L, c_1],$$

where the inequality follows from the monotonicity of $w(c)\,f(c)$. Therefore, the desired

property is reduced to

$$h(c_1) = \int_{c_L}^{c_1} w(y)\, f(y)\, \mathrm{d}y + (c_2 - c_1)\, w(c_1)\, f(c_1) \geq 0,$$

which follows from

$$\int_{c_L}^{c_1} w(y)\, f(y)\, \mathrm{d}y + (c_2 - c_1)\, w(c_1)\, f(c_1) \geq \int_{c_L}^{c_1} w(y)\, f(y)\, \mathrm{d}y + (c_2 - c_1)\, w(c_2)\, f(c_2) = 0.$$

For any $c \in [c_1, c_2]$ and $c \in [c_2, c_H]$, we have

$$\sigma'(c) = \frac{\mathrm{d}}{\mathrm{d}c}\left(w(c)\, f(c)\right) \leq 0$$

and

$$\mu(c) \geq \mu(c_2),$$

respectively. The first and second inequalities follow from the monotonicity of $w(c)\, f(c)$ on $c \in [c_L, c_2]$ and

$$\begin{aligned}
\mu(c) \geq \mu(c_2) \;&\Leftrightarrow\; w(c)\, f(c) \leq w(c_2)\, f(c_2) \\
&\Leftrightarrow\; \left[(\lambda_0\, z_1 + z_2)\,(1 + \lambda_1\, \bar{G}(\bar{\eta}')) + (v_E + \lambda_1\, z_1)\,(c - \lambda_0\, \bar{G}(\bar{\eta}'))\right]\, f(c) + (v_E + \lambda_1\, z_1)\, F(c) \\
&\leq \left[(\lambda_0\, z_1 + z_2)\,(1 + \lambda_1\, \bar{G}(\bar{\eta}')) + (v_E + \lambda_1\, z_1)\,(c_2 - \lambda_0\, \bar{G}(\bar{\eta}'))\right]\, f(c_2) + (v_E + \lambda_1\, z_1)\, F(c_2)
\end{aligned}$$

for any $c \in [c_2, c_H]$, where the last step follows from

$$v_E + \lambda_1\, z_1 < 0, \quad f(c) \leq f(c_2), \quad \text{and} \quad c - \lambda_0\, \bar{G}(\bar{\eta}') > 0 = c_2 - \lambda_0\, \bar{G}(\bar{\eta}').$$

The desired nonnegativity conditions are automatically implied by $\sigma(c_2) = \mu(c_2) = 0$.

Because

$$\int_{c_L}^{c} \nu(y) = \frac{1}{c_2 - c} \int_{c_L}^{c} w(y)\, f(y)\, \mathrm{d}y, \quad \forall c \in [c_L, c_1],$$

we have

$$\begin{aligned}
&-\phi - \int_{c_L}^{c_1} \nu(y)\, \mathrm{d}y = \frac{v_E + \lambda_1\, z_1}{1 + \lambda_1\, \bar{G}(\bar{\eta}')}, \\
&(c_2 - c)\, \nu(c) - \int_{c_L}^{c} \nu(y)\, \mathrm{d}y = w(c)\, f(c), \qquad\qquad \forall c \in [c_L, c_1],
\end{aligned}$$

$$\sigma(c) - \int_{c_L}^{c_1} \nu(y)\,\mathrm{d}y = w(c)\,f(c), \qquad \forall c \in [c_1, c_2], \text{ and}$$
$$-\mu(c) - \int_{c_L}^{c_1} \nu(y)\,\mathrm{d}y = w(c)\,f(c), \qquad \forall c \in [c_2, c_H].$$

Therefore,

$$\begin{aligned}
&\phi\,[-\Pi(c_H)] - \int_{c_L}^{c_1} \left[\Pi(c_H) + \int_{c}^{c_H} \tau(y)\,\mathrm{d}y + (c - \lambda_0\,\bar{G}(\bar{\eta}'))\,\tau(c)\right] \nu(c)\,\mathrm{d}c \\
&\quad + \int_{c_1}^{c_2} \tau(c)\,\sigma(c)\,\mathrm{d}c - \int_{c_2}^{c_H} \tau(c)\,\mu(c)\,\mathrm{d}c \\
=&\left[-\phi - \int_{c_L}^{c_1} \nu(c)\,\mathrm{d}c\right] \Pi(c_H) + \int_{c_L}^{c_1} \left[(c_2 - c)\,\nu(c) - \int_{c_L}^{c} \nu(y)\,\mathrm{d}y\right] \tau(c)\,\mathrm{d}c \\
&\quad + \int_{c_1}^{c_2} \left[\sigma(c) - \int_{c_L}^{c_1} \nu(y)\,\mathrm{d}y\right] \tau(c)\,\mathrm{d}c + \int_{c_2}^{c_H} \left[-\mu(c) - \int_{c_L}^{c_1} \nu(y)\,\mathrm{d}y\right] \tau(c)\,\mathrm{d}c \\
=&\frac{v_E + \lambda_1\,z_1}{1 + \lambda_1\,\bar{G}(\bar{\eta}')}\,\Pi(c_H) + \int_{c_L}^{c_H} w(c)\,\tau(c)\;\mathrm{d}F(c).
\end{aligned}$$

Therefore, the objective function is bounded by $\int_{c_1}^{c_2} \sigma(c)\,\mathrm{d}c$. Substituting our candidate primal solution into the objective function yields $\int_{c_L}^{c_2} w(y)\,f(y)\,\mathrm{d}y$, which is the same value as the upper bound, i.e.,

$$\begin{aligned}
\int_{c_1}^{c_2} \sigma(y)\,\mathrm{d}y &= \int_{c_1}^{c_2} \left[\frac{1}{c_2 - c_1} \int_{c_L}^{c_1} w(z)\,f(z)\,\mathrm{d}z + w(y)\,f(y)\right] \mathrm{d}y \\
&= \int_{c_L}^{c_1} w(y)\,f(y)\,\mathrm{d}y + \int_{c_1}^{c_2} w(y)\,f(y)\,\mathrm{d}y \\
&= \int_{c_L}^{c_2} w(y)\,f(y)\,\mathrm{d}y.
\end{aligned}$$

This establishes the optimality of our candidate solution.

When $c_0 = -\infty$, the virtual valuation function w is always non-positive on $[c_L, c_H]$. Moreover, the coefficient of $\Pi(c_H)$ is also negative; hence, it is evident that the optimal mechanism is to not request testing and to grant no conditional approval.

This completes the proof. □

References

Armstrong A (2024) Preventing "3 billion visits" to FDA a year with CBER's Peter Marks. *Fierce Biotech* URL `https://www.fiercebiotech.com/biotech/preventing-3-billion-visits-fda-year-cbers-peter-marks`, accessed 2024-07-02.

Baron DP, Myerson RB (1982) Regulating a monopolist with unknown costs. *Econometrica* 911–930.

Berndt ER, Glennerster R, Kremer MR, Lee J, Levine R, Weizsäcker G, Williams H (2007) Advance market commitments for vaccines against neglected diseases: Estimating costs and effectiveness. *Health Economics* 16(5):491–511.

Bognar K, Romley JA, Bae JP, Murray J, Chou JW, Lakdawalla DN (2017) The role of imperfect surrogate endpoint information in drug approval and reimbursement decisions. *Journal of Health Economics* 51:1–12.

Budish E, Roin BN, Williams H (2015) Do firms underinvest in long-term research? Evidence from cancer clinical trials. *American Economic Review* 105(7):2044–2085.

Carpenter D, Ting MM (2007) Regulatory errors with endogenous agendas. *American Journal of Political Science* 51(4):835–852.

Centers for Medicare & Medicaid Services (2022) CMS finalizes Medicare coverage policy for monoclonal antibodies directed against amyloid for the treatment of Alzheimer's disease. URL `https://www.cms.gov/newsroom/press-releases/cms-finalizes-medicare-coverage-policy-monoclonal-antibodies-directed-against-amyloid-treatment`.

Centers for Medicare & Medicaid Services (2024) Medicare & Medicaid spending by drug. URL `https://data.cms.gov/summary-statistics-on-use-and-payments/medicare-medicaid-spending-by-drug`.

Chandra A, Kao J, Miller KL, Stern AD (2024) Regulatory incentives for innovation: The FDA's breakthrough therapy designation. *The Review of Economics and Statistics* 1–46.

Collard-Wexler A, Grennan M, Steck A (2024) Regulating new product testing: The FDA vs. the invisible hand. *Working Paper* 1–32.

DiMasi JA, Grabowski HG, Hansen RW (2016) Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics* 47:20–33.

Dubois P, Moisson PH, Tirole J (2022) The economics of transferable patent extensions. *Toulouse School of Economics Working Paper* 1377:1–55.

Food and Drug Administration (2025) Advancing health through innovation: New drug therapy approvals 2024. URL `https://www.fda.gov/drugs/novel-drug-approvals-fda/novel-drug-approvals-2024`.

Frank RG, Shahzad M, Emanuel EJ (2022) Accelerated approval of cancer drugs: No economic reward for drug makers that conduct confirmatory trials. *Health Affairs* 41(9):1273–1280.

Gans JS, Ridley DB (2013) Innovation incentives under transferable fast-track regulatory review. *Journal of Industrial Economics* 61(3):789–816.

Grennan M, Town RJ (2020) Regulating innovation with uncertain quality: Information, risk, and access in medical devices. *American Economic Review* 110(1):120–161.

Henry E, Loseto M, Ottaviani M (2022) Regulation with experimentation: Ex ante approval, ex post withdrawal, and liability. *Management Science* 68(7):5330–5347.

Henry E, Ottaviani M (2019) Research and the approval process: The organization of persuasion. *American Economic Review* 109(3):911–955.

Jena AB, Zhang J, Lakdawalla DN (2017) The trade-off between speed and safety in drug approvals. *JAMA Oncology* 3(11):1465–1466.

Kaplan S (2016) 'Tough as nails': Storm swirls around FDA drug cop who approved controversial drug. `https://www.statnews.com/2016/09/20/janet-woodcock-sarepta-fda/`, accessed: 2025-12-16.

Kong E, Zhao O (2024) Market exclusivity and innovation: Evidence from antibiotics. `https://papers.ssrn.com/abstract=4676384`, SSRN Working Paper.

Kremer M, Levin J, Snyder CM (2022) Designing advance market commitments for new vaccines. *Management Science* 68(7):4786–4814.

Light DW, Warburton RN (2005) Extraordinary claims require extraordinary evidence. *Journal of Health Economics* 24(5):1030–1033.

Liu ITT, Kesselheim AS, Cliff ERS (2024) Clinical benefit and regulatory outcomes of cancer drugs receiving accelerated approval. *Journal of the American Medical Association* 331(17):1471–1479.

McClellan A (2022) Experimentation and approval mechanisms. *Econometrica* 90(5):2215–2247.

Myerson RB (1981) Optimal auction design. *Mathematics of Operations Research* 6(1):58–73.

Orlov D, Skrzypacz A, Zryumov P (2020) Persuading the principal to wait. *Journal of Political Economy* 128(7):2542–2578.

Peltzman S (1973) An evaluation of consumer protection legislation: The 1962 drug amendments. *Journal of Political Economy* 81(5):1049–1091.

Philipson TJ, Sun E (2008) Is the Food and Drug Administration safe and effective? *Journal of Economic Perspectives* 22(1):85–102.

Ridley DB, Lee CY (2020) Does Medicare reimbursement drive up drug launch prices? *The Review of Economics and Statistics* 102(5):980–993.

Ridley DB, Sun P, Xu C (2025) Optimal push, pull, and failure funding for global health, available at SSRN: `https://ssrn.com/abstract=5266214`.

Robey S, David FS (2016) Drug launch curves in the modern era. *Nature Reviews Drug Discovery* 16(1):13–14.

Snyder CM, Hoyt K, Gouglas D (2023) An optimal mechanism to fund the development of vaccines against emerging epidemics. *Journal of Health Economics* 91:102795.

The Economist (2025) America's drug regulator is in turmoil. `https://www.economist.com/united-states/2025/08/10/americas-drug-regulator-is-in-turmoil`, accessed August 11, 2025.

United Press International (1988) Police arrest aids protesters blocking access to fda offices. *Los Angeles Times* URL `https://www.latimes.com/archives/la-xpm-1988-10-11-mn-3909-story.html`.

US Congress (2023) Consolidated appropriations act, 2023. URL `https://www.congress.gov/117/bills/hr2617/BILLS-117hr2617enr.pdf`, pub. L. No. 117-328, div. FF, title III, §3210(d)(D), 136 Stat. 4459.

US Food and Drug Administration (2024) 21 cfr § 314.50 — content and format of an nda. https://www.ecfr.gov/current/title-21/chapter-I/subchapter-D/part-314/subpart-B/section-314.50, electronic Code of Federal Regulations; accessed: 2025-02-16.

US Food and Drug Administration (2025) FAQs: Commissioner's national priority voucher pilot program. https://www.fda.gov/news-events/press-announcements/faqs-commissioners-national-priority-voucher-pilot-program, accessed: 2025-12-14.

Wouters OJ, Kesselheim AS, Kuha J, Luyten J (2024) Sales revenues for new therapeutic agents approved by the United States Food and Drug Administration from 1995 to 2014. *Value in Health* 27(10):1373–1381.

Wouters OJ, McKee M, Luyten J (2020) Estimated research and development investment needed to bring a new medicine to market, 2009-2018. *Journal of the American Medical Association* 323(9):844–853.

Xu L, Zhao H, Petruzzi NC (2021) Inducing compliance with postmarket studies for drugs under FDA's accelerated approval pathway. *Manufacturing & Service Operations Management* 23(1):170–190.

Zhou EW, Jackson MJ, Ledley FD (2023) Spending on phased clinical development of approved drugs by the US National Institutes of Health compared with industry. *JAMA Health Forum* 4(7):e231921–e231921.