Gatekeeping, Selection, and Welfare

Francesco Del Prato, Paolo Zacchia

January 2026

We study staged entry with costly gatekeeping in a differentiated-products economy: entrepreneurs observe noisy signals before paying a resource-intensive activation cost. Precision improves selection but requires more resources, reducing entry and variety: welfare need not rise with precision. Under CES preferences, the activation cutoff is efficient as profit displacement offsets the consumer-surplus gain from variety. Welfare losses arise from verification costs shrinking the feasible set of varieties, not from misaligned incentives. Because the market responds efficiently to any given regime, these losses cannot be corrected via Pigouvian taxes.

KEYWORDS: staged entry, gatekeeping, monopolistic competition, product variety, welfare.

JEL CLASSIFICATION: D24, D82, L11, L26.

Francesco Del Prato: *Department of Economics and Business Economics, Aarhus University*, francesco.delprato@econ.au.dk. Paolo Zacchia: *Department of Economics, Ca' Foscari University of Venice; CERGE-EI; CEPR; and IZA*, paolo.zacchia@unive.it. We thank Daisuke Adachi, Paolo Martellini and Jose P. Vasquez for helpful comments. This paper has previously circulated with the title "Frictions and Welfare in Monopolistic Competition."

1. Introduction

Firm entry often proceeds in stages. Regulators license firms before market access, banks must satisfy supervisory stress tests before expanding balance sheets, and drug developers must pass increasingly demanding trials before commercialization. At each gate, decision-makers act on imperfect information: they observe noisy indicators of future performance, while true productivity is revealed only after resources are committed (Jovanovic 1982).

This paper studies how such staged, signal-based entry reshapes selection and welfare in a standard differentiated-products environment. We embed staged entry with imperfect, signal-based gatekeeping into the closed-economy framework of Melitz (2003). Potential entrants first incur an experimentation cost to learn a noisy signal, then decide whether to activate and pay an activation cost, after which productivity is revealed and production occurs only if profitable.

To keep language compact, we use *gatekeeping* as shorthand for *pre-activation screening that is both informative and costly*. We refer to a *gatekeeping regime* as an information structure (signal precision) together with the associated per-activation resource requirement. We distinguish between raising *standards* (a higher cutoff) and improving *precision* (a better signal). While raising a standard implies a costless policy choice, improving precision absorbs real resources in verification and compliance. Consequently, tightening the regime does not mechanically raise the cutoff; instead, it alters equilibrium selection by changing the information structure and the cost of activation. This reshapes both the quality (selection) and quantity (variety) of firms, which jointly determine welfare.

Results. A natural presumption is that more informative screening should improve welfare by improving selection. We show this need not hold once we account for the resource cost of precision.

Comparing across regimes, higher verification precision improves average productivity through improved selection, but the anticipated cost of verification lowers the expected value of entry. This creates a chilling effect on experimentation: fewer entrepreneurs bother to pay the initial fixed cost to learn their signal. We derive a welfare decomposition showing that welfare is determined by the trade-off between average productivity (selection) and the mass of surviving varieties. Higher verification precision raises the former but shrinks the latter. Consequently, welfare is not monotone in precision: a "safer" market may be poorer because the high compliance burden discourages high-potential firms from ever emerging.

We distinguish this policy trade-off from the efficiency of the market's response. We show that, conditional on any given regime, the decentralized activation cutoff is constrained-efficient. In the CES benchmark, private entrants perfectly internalize the social value of activation taking the verification regime as given. Thus, welfare losses from strict verification regimes arise not from a private–social wedge along the activation margin, but from the choice of the regime itself: a more resource-intensive verification technology shrinks the feasible set of varieties even when market participants behave optimally.

Contributions. We make three main contributions.

First, we introduce a tractable model of staged entry where entrepreneurs learn imperfect information before paying a resource cost, and we provide a competitive intermediation microfoundation for this structure.

Second, in our maintained CES environment the activation margin is efficient: with constant markups, the consumer-surplus gain from an additional variety is offset by business stealing. As a result, conditional on any given regime the decentralized activation cutoff is constrained-efficient, which isolates the welfare consequences of changing verification regimes from the familiar entry wedge.

Third, we provide a sufficient condition for welfare to decline at high precision. We clarify when the selection gains from tighter screening are dominated by the contraction in variety, offering a check against the view that better information is always desirable.

Policy implications. More precise (and resource-intensive) gatekeeping does not guarantee welfare gains. Policies that tighten verification requirements have first-order welfare consequences for the *unseen* extensive margin—the firms that never apply. Because the market cutoff is efficient, the welfare loss is not a market failure to be corrected by taxes or subsidies, but a structural cost of the regulatory design itself.

Literature. Our benchmark is Dhingra and Morrow (2019), who establish Pareto-optimality of the closed Melitz economy. We show that adding staged entry preserves constrained efficiency along the activation cutoff, yet allows verification precision to depress welfare through the resource-variety channel. We study information and real resource costs of gatekeeping, abstracting from the borrowing constraints and heterogeneous markups emphasized in work on collateral-based misallocation (Midrigan and Xu 2014) and markup-driven reallocation (Baqaee and Farhi 2020).

Outline. The remainder of this paper is organized as follows. Section 2 presents the model and characterizes equilibrium. Section 3 analyzes welfare. Section 4 concludes. Derivations and proofs are in Appendix A.

2. The model

We extend the closed-economy framework of Melitz (2003) by introducing *staged entry with imperfect, signal-based gatekeeping*. Firm entry proceeds in two stages: entrepreneurs first pay an experimentation cost to learn a noisy signal about their productivity, then decide whether to incur an activation cost based on the signal. True productivity is revealed only after activation, at which point firms produce or exit.

2.1. Setup

We study a closed economy with monopolistic competition and CES preferences. The representative consumer's utility is $U^{\frac{\sigma-1}{\sigma}} = \int_{\omega\in\Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega$, where Ω represents the set of varieties available in equilibrium, $q(\omega)$ is consumption of variety $\omega \in \Omega$, and $\sigma > 1$ is the elasticity of substitution. Firms supply these varieties with heterogeneous productivities $\varphi(\omega) > 0$, which we treat as exogenous. Production requires a single factor input. It involves a fixed requirement $f > 0$ (in units of the input) and variable input usage proportional to output, yielding the total input demand function $l(q) = f + q/\varphi$ for a firm producing quantity q with productivity φ. The representative consumer supplies the input inelastically; we normalize total input to $L = 1$. The base cost per unit of the input is normalized to unity.

This economy inherits the standard properties of the monopolistic competition model by Dixit and Stiglitz (1977) as extended by Melitz (2003). In particular, each firm's optimal quantity scales with φ^{σ}, while revenues and profits scale with $\varphi^{\sigma-1}$. Hence, for any two firms with productivity φ_1 and φ_2, the ratio of their equilibrium revenues $r(\varphi)$ is $r(\varphi_1)/r(\varphi_2) = (\varphi_1/\varphi_2)^{\sigma-1}$. The productivity distribution is endogenously determined through competitive selection, but unlike in Melitz (2003), entry proceeds in two stages: entrepreneurs first pay an experimentation cost, then decide whether to incur an activation cost conditional on a noisy signal.

An entrepreneur is characterized by a pair (φ, θ), where $\varphi > 0$ is true productivity and $\theta > 0$ is a noisy *signal* about φ. Both are drawn from a joint distribution $G(\varphi, \theta)$ but are initially unobserved. Firm creation proceeds as follows. Entrepreneurs decide whether to attempt setting up a firm; doing so requires paying a one-time experimentation cost $f_n > 0$ to learn the signal θ. Conditional on θ, entrepreneurs decide whether to activate by paying the activation cost $f_b(\rho)$, where ρ indexes screening precision, after which true productivity φ is revealed. Activated firms then set prices and quantities and produce if profitable; operating firms remain in the market until forced to exit by an exogenous shock occurring with probability δ.

This staged entry structure features *imperfect, signal-based gatekeeping*: the activation decision is based on the noisy signal θ, with true productivity φ revealed only afterward.

Unlike extensions of Melitz that incorporate liquidity constraints for export market entry (Manova 2013; Chaney 2016), our structure affects the *domestic* entry margin through imperfect information rather than borrowing constraints.[1] This shapes firm selection in the left tail of the productivity distribution.

We conduct our analysis under three key assumptions.

ASSUMPTION 1 (Signal informativeness). *If* $\theta_1 > \theta_2$, *then the conditional distribution of productivity given the signal is shifted to the right:* $G(\varphi|\theta_1) \leq G(\varphi|\theta_2)$ *for all* $\varphi > 0$, *with strict inequality for some* φ.

Assumption 1 ensures that higher signals predict higher productivity in the sense of first-order stochastic dominance, which implies that expected profits are increasing in θ and delivers cutoff behavior at the activation stage.

We index the *screening precision* of the signal by a scalar $\rho \in (0,1)$: higher ρ corresponds to a more informative verification technology. We refer to a *gatekeeping regime* as the pair $(\rho, f_b(\rho))$. To capture that higher verification precision requires real resources (verification, documentation, data collection), we assume:

ASSUMPTION 2 (Costly gatekeeping). *The activation cost depends on screening precision and is given by* $f_b(\rho)$, *where* $f_b : (0,1) \to \mathbb{R}_+$ *is continuous and weakly increasing.*

Assumption 2 allows for general (possibly bounded) verification cost schedules. When it is useful to derive sharp statements about the high-precision limit, we impose additional structure on $f_b(\rho)$, including specifications in which costs become arbitrarily large as $\rho \uparrow 1$ (see Remark 1).

ASSUMPTION 3 (Log-normality). *For tractability, the joint distribution* $G(\varphi, \theta)$ *is bivariate log-normal with standard log-normal marginals. In this benchmark, screening precision is implemented as* $\rho = \mathbb{C}\text{orr}(\log\theta, \log\varphi) \in (0,1)$.

Assumption 3 is a parametric benchmark used for closed-form expressions and for the quantitative exercises; the qualitative logic of staged entry and cutoff behavior relies primarily on Assumptions 1 and 2. In the log-normal benchmark, Assumption 1 corresponds to $\rho \geq 0$, and we focus on the interior case $\rho \in (0,1)$.

2.2. Analysis

Once the set of firms that have paid both entry costs (f_n and $f_b(\rho)$) is determined, firm behavior proceeds as in the Melitz model. Operating firms face CES demand and produce

[1] The distinction between f_n and $f_b(\rho)$ can be interpreted as a staged cost structure: of the full Melitz entry cost f_e, entrepreneurs pay f_n upfront to learn their signal, then pay $f_b(\rho)$ conditional on passing the activation threshold.

under monopolistic competition; a firm with productivity φ earns per-period revenue $r(\varphi) = \sigma f(\varphi/\varphi^*)^{\sigma-1}$ and profit $\pi(\varphi) = r(\varphi)/\sigma - f = f[(\varphi/\varphi^*)^{\sigma-1} - 1]$, where φ^* is the zero-profit cutoff satisfying $\pi(\varphi^*) = 0$. Let $\delta \in (0,1)$ denote the effective discount rate, incorporating both the exogenous probability of firm exit and the consumer's time preference.[2] The equilibrium is characterized by two conditions governing the activation and experimentation decisions.

Let $\tilde{\pi}(\theta)$ denote the expected per-period profit from activating an entrepreneur with signal θ, incorporating the probability of exit after true productivity φ is revealed. [3] By Assumption 1, $\tilde{\pi}(\theta)$ is increasing in θ: higher signals predict higher expected profits. The equilibrium features an *activation cutoff* θ^*: entrepreneurs activate if and only if $\theta \geq \theta^*$. This cutoff is pinned down by the Activation Condition (AC):

$$\frac{\tilde{\pi}(\theta^*)}{\delta} = f_b(\rho). \tag{1}$$

The left-hand side is the expected lifetime value of profits from activating a marginal entrepreneur (with δ converting per-period profits to present value under exogenous exit); the right-hand side is the activation cost. A suitable interior θ^* exists and is unique under our assumptions.

The entrepreneurs' initial experimentation decision is governed by free entry. Let $C(\theta)$ denote the marginal cumulative distribution of the signal. Entrepreneurs enter experimentation until the expected payoff equals the experimentation cost f_n. Since activation occurs only for $\theta \geq \theta^*$, and the activation cost $f_b(\rho)$ is incurred with probability $1 - C(\theta^*)$, the Free Entry (FE) condition is:

$$\int_{\theta^*}^{\infty} \frac{\tilde{\pi}(\theta)}{\delta} dC(\theta) - \left[1 - C\left(\theta^*\right)\right] f_b(\rho) - f_n = 0. \tag{2}$$

The first term is expected lifetime profits conditional on activation; the second is the expected activation cost; the third is the experimentation cost. Together, the AC (1) and FE (2) conditions fully characterize equilibrium.

Our baseline formulation is reduced-form: entrepreneurs pay the activation cost as a real resource outlay when they choose to activate. In this setting, the same allocation can also be supported by equivalent microfoundations (e.g., competitive intermediaries who finance activation or perform verification in exchange for an actuarially fair claim contingent on θ). Because these microfoundations do not change the information structure or feasibility of activation, they redistribute surplus but do not affect the equilibrium cutoffs or aggregates, so we do not pursue them further. This is shown in Appendix A.2.

[2] Formally, in continuous time with flow discount rate r and Poisson exit rate λ, the effective discount rate is $\delta \equiv r + \lambda$.

[3] Formally, $\tilde{\pi}(\theta)$ accounts for endogenous survival through the operating cutoff φ^*; dividing by δ converts per-period profits into expected lifetime value under exogenous death.

When competitive intermediaries provide $f_b(\rho)$ in exchange for profit shares, zero-profit competition pins down the activation cutoff at exactly the same θ^* as in (1). Contract terms redistribute surplus but do not affect selection. We present staged entry as the baseline; Appendix A.2 provides the formal proof of allocation equivalence under finance.

To complete the analysis, we must characterize the function $\tilde{\pi}(\theta)$. Adapting the post-entry analysis from the Melitz model, for a given signal value θ, we have:

$$\tilde{\pi}(\theta) = \mathbb{E}_{\varphi|\theta}(\pi(\varphi)) = f\left\{\int_{\varphi^*}^{\infty}\left(\frac{\varphi}{\varphi^*}\right)^{\sigma-1} dG(\varphi|\theta) - \left[1 - G\left(\varphi^*|\theta\right)\right]\right\}, \tag{3}$$

where φ^* represents the productivity threshold below which firms find production unprofitable and exit in equilibrium. Equation (3) incorporates endogenous survival through the common cutoff φ^*, which induces a signal-dependent survival probability. Importantly, both the AC (1) and FE (2) conditions are implicitly functions of φ^* through equation (3). The equilibrium is fully determined by the threshold pair (θ^*, φ^*). The equilibrium geometry we use throughout is a feature of the log-normal benchmark: log-normality delivers a power-form AC locus and a single-peaked FE locus in the (θ^*, φ^*) plane (Appendix A.1). Assumption 1 ensures expected profits are increasing in θ (so the FE slope switches sign once), while Assumption 2 enters here only through the level of the activation cost $f_b(\rho)$ (its monotonicity in ρ matters for the comparative statics and welfare analysis below).

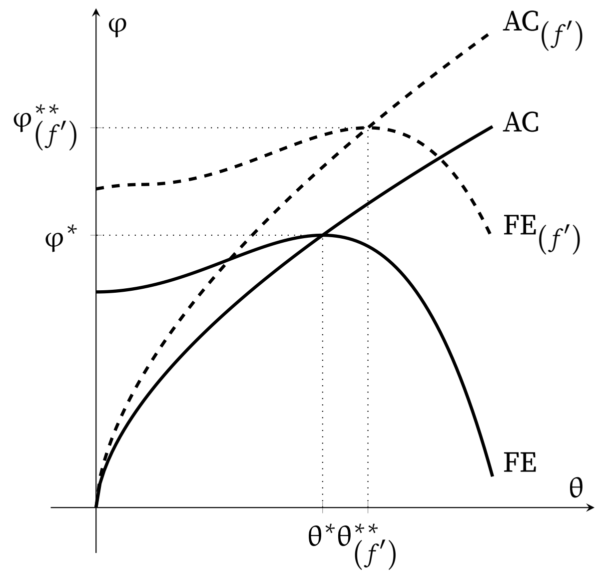
PROPOSITION 1. *Maintain Assumptions 1–3. Then an equilibrium pair* (θ^*, φ^*) *exists and is unique. In the log-normal benchmark, the equilibrium is characterized by the intersection of the AC curve,* $\varphi^* = A(\theta^*)^{\rho}$ *for some constant* $A > 0$*, and the FE locus, which is single-peaked in* θ^* *(equivalently, it has a unique maximizer and at most one stationary point). The intersection occurs precisely at this unique maximizer of the implicit function relating* φ^* *to* θ^* *traced out by the FE locus.*

The proof is in Appendix A.1.

Figure 1 illustrates the equilibrium as the intersection between the two solid curves representing the AC and FE conditions. The AC curve exhibits a monotonically increasing relationship because higher signal thresholds lead to higher productivity thresholds through improved firm selection.

The FE locus is single-peaked because raising the signal cutoff θ^* has two opposing effects in (2): it (i) excludes marginal activated types (lowering expected profits) and (ii) reduces the probability of paying the activation cost $f_b(\rho)$. When θ^* is low, the cost-saving effect dominates and the FE locus slopes upward; when θ^* is high, the foregone-profit effect dominates and the FE locus slopes downward. The peak occurs where these effects offset, equivalently where $\tilde{\pi}(\theta^*)/\delta = f_b(\rho)$, so in equilibrium the AC and FE curves intersect at the FE maximum.

FIGURE 1. Equilibrium of the model and comparative statics



Note: Solid lines: the equilibrium pair (θ^*, φ^*) solves AC (1) and FE (2). The AC locus slopes upward; the FE locus is single-peaked (upward sloping for $\theta < \theta^*$ and downward sloping for $\theta > \theta^*$). Dashed lines: a higher operating fixed cost ($f' > f$) shifts FE up and rotates AC left, raising both cutoffs.

Operating fixed-cost shocks. As a simple comparative statics exercise, increase the per-period operating fixed requirement from f to $f' > f$, holding $(f_n, f_b(\rho), \rho)$ fixed. This shifts the FE locus up and rotates the AC locus left, so the equilibrium features higher cutoffs θ^* and φ^* (dashed curves in Figure 1). Intuitively, a higher operating fixed requirement reduces profits at any given productivity, so sustaining entry requires stricter selection: a higher productivity cutoff, and (because signals predict productivity) a higher activation cutoff.

Limit cases. To clarify how staged entry nests the standard Melitz framework, Figure 2 (Panel A) summarizes two limit cases. Both are outside the maintained domain $\rho \in (0, 1)$, but they are useful for intuition.

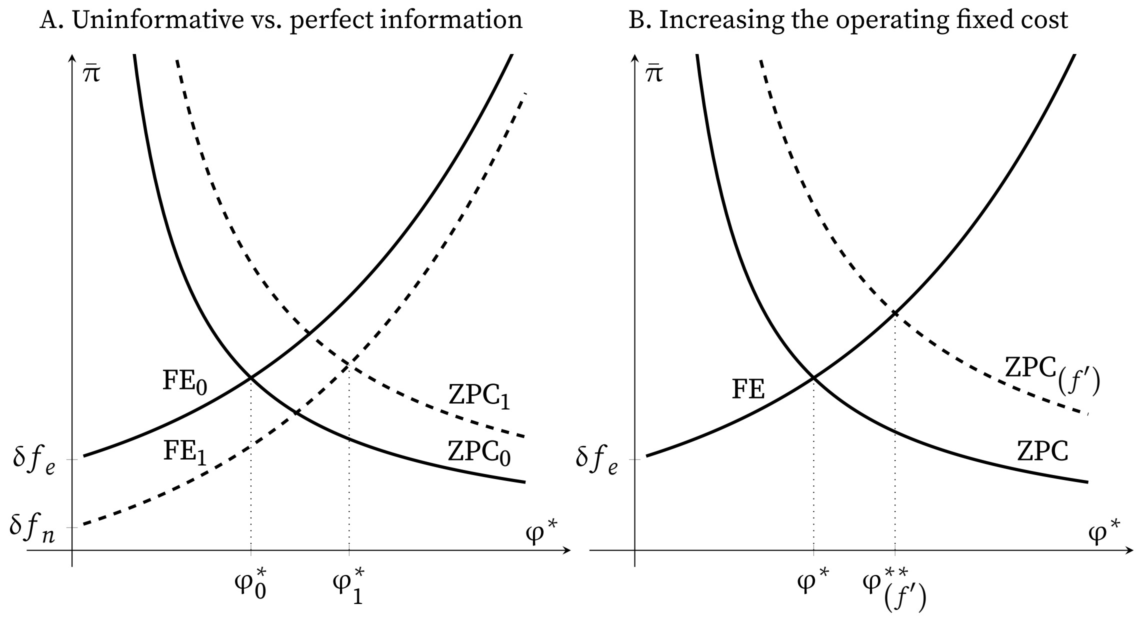
Uninformative signals ($\rho \to 0$). Signals contain no information, so $\theta^* \to 0$ and all experimenters activate. The total entry cost is paid by every entrant; in the $\rho \to 0$ limit this cost converges to $f_{e0} \equiv f_n + \underline{f}_b$, where $\underline{f}_b \equiv \lim_{\rho \downarrow 0} f_b(\rho)$. The FE condition reduces to:

$$\frac{\left(1 - G\left(\varphi_0^*\right)\right)}{\delta} \bar{\pi}_0 - f_{e0} = 0,$$

where $\bar{\pi}_0$ is expected profit conditional on survival and the ZPC is $\bar{\pi}_0 = f\, k(\varphi_0^*)$, with $k(\varphi^*) = [\widetilde{\varphi}(\varphi^*)/\varphi^*]^{\sigma-1} - 1$ as in Melitz (2003).

Perfect-information thought experiment ($\rho \to 1$ holding f_b fixed). Suppose instead that signals

FIGURE 2. Limit cases and fixed-cost intuition



Note: Panel A contrasts $\rho \to 0$ (solid) with a perfect-information thought experiment $\rho \to 1$ holding f_b fixed (dashed). Panel B shows the Melitz operating-fixed-cost shift: raising f to f' shifts the ZPC curve.

become perfectly informative while the activation cost is held fixed at some finite level $\bar{f}_b$. Then activation implements perfect selection, and the activation cost is equivalent to a higher effective operating fixed requirement $f + \delta \bar{f}_b$ in present-value terms. Free entry becomes:

$$\frac{\left(1 - G\left(\varphi_1^*\right)\right)}{\delta}\bar{\pi}_1 - f_n = 0,$$

with the ZPC given by $\bar{\pi}_1 = \left(f + \delta \bar{f}_b\right) k\left(\varphi_1^*\right)$. In this thought experiment, better information strengthens selection holding costs fixed, so $\varphi_1^* > \varphi_0^*$.

Panel B of Figure 2 gives the standard Melitz comparative statics: raising the operating fixed cost shifts the ZPC outward and increases the productivity cutoff. This corresponds to tighter selection in our two-stage model (higher φ^* and, by informativeness, higher θ^*), consistent with the joint-cutoff movement in Figure 1. Unlike the full gatekeeping exercise, however, Figure 2 is meant to isolate the selection channel; the welfare analysis below additionally incorporates the feasibility channel operating through the regime-dependent activation cost $f_b(\rho)$.

3. Welfare

This section analyzes how gatekeeping affects social welfare. We extend the standard Melitz welfare accounting to incorporate experimentation and activation costs, then proceed in two steps:

a. *Across-regime comparison*: We ask which gatekeeping regime $(\rho, f_b(\rho))$ maximizes welfare. This compares equilibria with different information structures and resource costs.
b. *Within-regime efficiency*: We ask whether, holding the regime fixed, the decentralized activation cutoff is efficient. This compares the market allocation to a constrained planner's solution.

As illustrated in Section 2 (Figure 2), changing the regime creates opposing forces. Higher verification precision improves selection (Panel A) but acts like a higher fixed cost that depresses variety (Panel B). The net welfare effect depends on which force dominates.

3.1. Welfare building blocks

Following Melitz (2003), social welfare is:[4]

$$W = \frac{\sigma - 1}{\sigma} M^{\frac{1}{\sigma-1}} \tilde{\varphi}, \tag{4}$$

where M represents the mass of operating firms in equilibrium and $\tilde{\varphi}$ denotes their "aggregate" productivity calculated as a generalized average of order $\sigma - 1$. The relative contributions of these two components depend on verification precision, indexed by ρ.

To evaluate equation (4) in steady state with staged entry, define the probabilities of passing the signal and productivity thresholds as $\mathcal{P}^*_\theta \equiv \Pr(\theta \geq \theta^*)$ and $\mathcal{P}^*_\varphi \equiv \Pr(\varphi \geq \varphi^*, \theta \geq \theta^*)$. In equilibrium, experimentation uses $L_n = M_e f_n$ and activation uses $L_b = \mathcal{P}^*_\theta M_e f_b(\rho)$, where M_e is the mass of entering firms. Rewriting welfare in terms of the experimentation margin M_e clarifies how the gatekeeping regime shapes variety: tighter verification can raise the per-experimenter resource cost and therefore reduce the mass of experiments that can be sustained in equilibrium. Combining (4) with the definition of $\tilde{\varphi}$ and the steady-state flow condition $\delta M = \mathcal{P}^*_\varphi M_e$ yields:

$$W(\rho)^{\sigma-1} = \left(\frac{\sigma - 1}{\sigma}\right)^{\sigma-1} \frac{M_e(\rho)}{\delta} \mathbb{E}\Big(\varphi^{\sigma-1} \mathbb{1}(\theta \geq \theta^*,\ \varphi \geq \varphi^*)\Big). \tag{5}$$

Here, the selection–variety tradeoff is transparent: gatekeeping affects welfare both through the "quality" of survivors (the selection channel, the expectation term), and the "quantity" of experiments (the variety channel, $M_e(\rho)$). Define the unconditional expected profits at the experimentation stage as $\breve{\pi}(\rho) \equiv \int_{\theta^*}^{\infty} \tilde{\pi}(\theta)\, dC(\theta)$. By free entry, $\breve{\pi}(\rho) = \delta\left(\mathcal{P}^*_\theta f_b(\rho) + f_n\right)$. Let $\bar{r}$ and $\bar{\pi}$ denote average flow revenue and profit per operating firm, respectively.

[4] As in the original monopolistic competition model by Dixit and Stiglitz (1977), social welfare equals the inverse of the price level.

Finally, the labor market clearing condition implies that resources per operating variety are given by:

$$\frac{L}{M} = \bar{r} - \bar{\pi} + \frac{\breve{\pi}(\rho)}{\mathcal{P}^*_\varphi}, \tag{6}$$

where $\breve{\pi}(\rho)/\mathcal{P}^*_\varphi$ is the amortized setup burden—i.e., the setup cost per experimenter scaled by the probability an experimenter becomes an operating firm. For fixed L, a higher setup burden reduces the sustainable mass of firms M and therefore reduces product variety.

3.2. Across-regime comparison: the cost of precision

We first examine how welfare varies with the gatekeeping regime indexed by ρ. Because equilibrium thresholds vary endogenously with ρ, the mapping from verification precision to welfare need not admit a closed-form expression. We therefore establish continuity and then show that welfare can strictly decline when verification becomes more precise, even if verification costs remain bounded.

PROPOSITION 2 (Welfare continuity). *Maintain Assumptions 1–3. The equilibrium welfare function* $W : (0,1) \to \mathbb{R}_+$ *is continuous in verification precision* ρ.

The proof is in Appendix A.3. Continuity ensures that welfare comparisons across regimes are well-defined and that interior optima, when they exist, can be characterized by standard methods.

PROPOSITION 3 (Welfare can decline with bounded verification costs). *Maintain Assumptions 1–3 and assume* $f_n > 0$. *For any* $0 < \rho_L < \rho_H < 1$, *there exists a bounded, continuous, weakly increasing cost schedule* $f_b : (0,1) \to \mathbb{R}_+$ *such that equilibrium welfare satisfies* $W(\rho_H) < W(\rho_L)$.

The proof is in Appendix A.4.

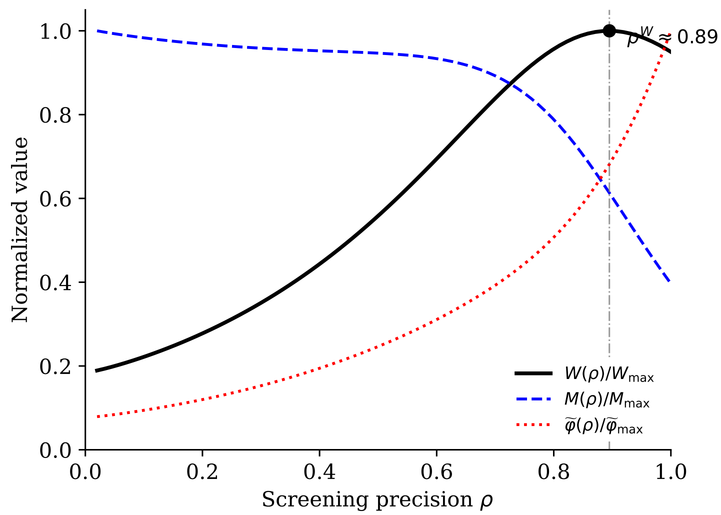
Proposition 3 formalizes the selection–variety tradeoff: higher precision improves selection but increases the resource burden per experiment; when the cost increase is sufficiently steep, the resulting loss of variety outweighs the productivity gains. The following corollary makes this tradeoff precise.

COROLLARY 1 (Local welfare condition). *Let* $S(\rho) \equiv \mathbb{E}_{\varphi^{\sigma-1}\mathbb{1}(\theta \geq \theta^*(\rho), \varphi \geq \varphi^*(\rho))}$ *denote the selection term and let* $B(\rho) \equiv \frac{L}{M_e(\rho)} = f_n + \mathcal{P}^*_\theta(\rho) f_b(\rho) + \frac{\mathcal{P}^*_\varphi(\rho)}{\delta}\big(\bar{r}(\rho) - \bar{\pi}(\rho)\big)$ *denote equilibrium resources per experimenter. If both are differentiable at* ρ, *then*

$$\frac{d}{d\rho} \log W(\rho) = \frac{1}{\sigma - 1}\left(\frac{S'(\rho)}{S(\rho)} - \frac{B'(\rho)}{B(\rho)}\right).$$

Welfare falls locally whenever the resources-per-experimenter elasticity exceeds the selection elasticity.

FIGURE 3. Gatekeeping and welfare.



Note: ρ indexes verification precision: higher ρ makes signals more informative but requires costlier verification. Solid black: welfare $W(\rho)$; dashed blue: product variety $M(\rho)$ (mass of operating firms); dotted red: average productivity $\widetilde{\varphi}(\rho)$. Activation cost: $f_b(\rho) = f_{b0}(1 + \kappa\rho^{\alpha})$ (bounded above by $f_{b0}(1 + \kappa)$). Parameters: $\sigma = 2$, $f = 0.15$, $f_n = 0.005$, $f_{b0} = 3$, $\kappa = 2$, $\alpha = 8$, $\delta = 0.1$. Each series is normalized by its maximum over the plotted grid; welfare is maximized at an interior $\rho^W \approx 0.89$.

The proof is in Appendix A.6.

REMARK 1 (Unbounded cost schedules). *If $f_b(\rho) \to \infty$ as $\rho \uparrow 1$, then $W(\rho) \to 0$ (Appendix A.5), guaranteeing an interior optimum. Proposition 3 shows this extreme assumption is unnecessary: the selection–variety tradeoff operates even under bounded costs.*

Figure 3 provides a quantitative illustration of Proposition 3 under a bounded cost schedule. This analysis distinguishes the choice of regime from the market allocation. While Dhingra and Morrow (2019) show that the Melitz economy is Pareto-optimal given fixed primitives, we show that changing the primitives themselves (information and resource costs) creates a non-monotonic welfare frontier. We plot equilibrium welfare, product variety (the mass of operating firms), and average productivity as functions of ρ over a grid that approaches one. In this calibration, higher screening precision raises average productivity through better selection but also increases verification costs, which reduces entry and product variety; the resulting tradeoff generates an interior welfare-maximizing level of verification precision.

The fact that welfare can decline with more precise (and costlier) verification raises a natural question: is the *laissez-faire* economy failing to balance these trade-offs optimally? To answer this, we now turn to the *within-regime* analysis, comparing the market allocation to that of a constrained planner.

3.3. Within-regime efficiency: the planner's problem

We consider a constrained planner who faces the same information and timing as the decentralized economy. The planner observes the signal θ before activation, but not true productivity φ and chooses an activation rule $m(\theta) \in [0, 1]$ to maximize welfare, subject to the resource constraint while taking the experimentation margin as given (equivalently, as pinned down by free entry for any activation rule).

This is a (constrained) *second-best* benchmark: the planner controls activation but takes the experimentation decision (and thus the entry externality) as given.[5]

The planner's first-order condition for activation depends on a marginal payoff kernel $\widetilde{G}(\theta)$, derived in Appendix A.7.1. Intuitively, $\widetilde{G}(\theta)$ measures the net social value of marginally increasing activation at signal θ: it aggregates the welfare consequences for variety (more operating firms), productivity (the expected quality of entrants conditional on θ), and resource use (the verification cost incurred). Only the sign of $\widetilde{G}(\theta)$ matters for cutoff characterization, so we treat the kernel up to a positive multiplicative normalization.

PROPOSITION 4 (Cutoff optimality under single-crossing). *Let $\widetilde{G}(\theta)$ denote the constrained planner's marginal payoff kernel. Assume* (SC)*: $\widetilde{G}(\theta)$ is weakly increasing in θ. Then there exists a cutoff θ^P such that an optimal activation rule is $m(\theta) = \mathbb{1}(\theta \geq \theta^P)$.*

The proof is in Appendix A.7.2. Appendix A.7.3 provides primitive sufficient conditions for (SC) under lognormality.

In general environments, (SC) is a sufficient regularity condition: $\widetilde{G}(\theta)$ aggregates the welfare consequences of activation for variety, productivity, and resource use, and need not be monotone in the signal. However, in the CES benchmark we do not need to impose (SC) separately.

COROLLARY 2 (CES makes single-crossing redundant). *Under CES, the planner's kernel $\widetilde{G}(\theta)$ can be normalized as $\widetilde{\pi}(\theta)/\delta - f_b(\rho)$. Since $\widetilde{\pi}(\theta)$ is weakly increasing in θ by Assumption 1,* (SC) *holds automatically.*

The proportionality arises because, under CES, the "appropriability" gain from an additional variety (consumer surplus that the entrant does not capture) is exactly offset by the "business stealing" effect (profit displaced from incumbents). As a result, private and social incentives are aligned on the activation margin, and the planner's marginal gain from activation is proportional to the entrant's private surplus.

This proportionality leads to the central efficiency result of this subsection:

LEMMA 1 (Activation-cutoff efficiency). *Under CES, the constrained-efficient activation rule coincides with the decentralized activation rule ($\theta^P = \theta^*$).*

[5] A first-best planner would also choose the experimentation intensity M_e. We focus on the constrained problem to isolate the efficiency of the activation margin.

The proof is in Appendix A.7.4.

An immediate policy implication is that no corrective intervention can improve on the market outcome at the activation margin:

COROLLARY 3 (No Pigouvian correction at the activation margin). *Fix a gatekeeping regime* $(\rho, f_b(\rho))$. *Consider a policy that adds a constant per-activation transfer s and uses an experimentation-stage transfer to preserve free entry. Then equilibrium welfare is maximized at* $s = 0$.

The proof is in Appendix A.9. In other words, welfare losses from tighter gatekeeping cannot be undone by Pigouvian intervention at the activation margin.

Lemma 1 is crucial for interpreting the welfare results in Section 3.2: if welfare declines with precision, the cause is not a market failure at the activation margin. Rather, the decline in welfare at high precision is a structural consequence of the regime choice itself. The resource intensity of precise gatekeeping shrinks the feasible set of varieties, even when market participants behave efficiently within each regime. This insight echoes and extends Dhingra and Morrow (2019), who show that the Melitz economy is Pareto-optimal given fixed primitives; we show that changing the primitives (verification precision and its resource cost) creates a non-monotonic welfare frontier even in the absence of any private–social wedge.

4. Conclusions

This paper studies selection and welfare in a differentiated-products economy with staged entry: entrepreneurs first experiment to obtain a noisy signal, then decide whether to activate under costly, signal-based gatekeeping. Tightening the gatekeeping regime improves selection and raises average productivity, but it also diverts real resources into verification and reduces the mass of surviving varieties. When the resource burden grows sufficiently fast with precision, the variety loss dominates the selection gain and welfare falls.

Crucially, this welfare loss is not a market failure. In the CES benchmark, the equilibrium activation cutoff is constrained-efficient conditional on any regime: private and social incentives are aligned, and no Pigouvian tax or subsidy at the activation margin can improve outcomes. The welfare effects of more precise verification arise entirely from comparing regimes that differ in information and resource intensity, not from a correctable distortion within any given regime.

This delivers a specific warning for regulation. Policymakers cannot evaluate gatekeeping solely by the performance of survivors. When compliance costs are high, an efficient market response is to reduce experimentation and variety; a “safer” market may therefore be a poorer one in welfare terms. Because the market allocation is already efficient at the activation margin, the appropriate policy response is not corrective taxation but a reconsideration of the verification regime itself.

Future research could extend the framework to open economies, where verification requirements may act as non-tariff barriers and reshape the selection–variety trade-off across countries.

References

Baqaee, David Rezza, and Emmanuel Farhi. 2020. "Productivity and misallocation in general equilibrium." *The Quarterly Journal of Economics* 135 (1): 105–163.

Chaney, Thomas. 2016. "Liquidity Constrained Exporters." *Journal of Economic Dynamics and Control* 72: 141–154.

Dhingra, Swati, and John Morrow. 2019. "Monopolistic Competition and Optimum Product Diversity under Firm Heterogeneity." *Journal of Political Economy* 127 (1): 196–232.

Dixit, Avinash K., and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67 (3): 297–308.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica: Journal of the econometric society*: 649–670.

Manova, Kalina. 2013. "Credit Constraints, Heterogeneous Firms, and International Trade." *The Review of Economic Studies* 80 (2): 711–744.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–1725.

Midrigan, Virgiliu, and Daniel Yi Xu. 2014. "Finance and misallocation: Evidence from plant-level data." *American Economic Review* 104 (2): 422–458.

Owen, Donald Bruce. 1980. "A table of normal integrals." *Communications in Statistics - Simulation and Computation* 9 (4): 389–419.

Appendix A. Additional model analysis

A.1. Equilibrium characterization

Maintain Assumption 3. To facilitate our analysis, we establish the auxiliary notation:

$$
\begin{aligned}
t &= \log \theta, \\
p &= \log \varphi, \\
h &= -\log \theta, \\
h' &= -\log \theta + \rho(\sigma - 1), \\
z &= \frac{\log \varphi - \rho \log \theta}{\sqrt{1-\rho^2}}.
\end{aligned}
$$

Asterisks denote transformations at threshold values (e.g., $t^* = \log \theta^*$). Let

$$
z^*(t) = \frac{p^* - \rho t}{\sqrt{1-\rho^2}}.
$$

Additionally, let $\phi(x)$ be the standard normal PDF and $\Phi(x)$ its cumulative distribution, both evaluated at x. We denote by $\Phi_\rho(x, y)$ the cumulative bivariate normal distribution with standard normal marginals and correlation parameter ρ, evaluated at (x, y).

Expression (3) can be elaborated as a function of any real t:

$$
\begin{aligned}
\frac{\widetilde{\pi}(e^t)}{f} &= \int_{p^*}^{\infty} \frac{e^{(\sigma-1)(p-p^*)}}{\sqrt{1-\rho^2}} \phi\left(\frac{p-\rho t}{\sqrt{1-\rho^2}}\right) dp - \left[1 - \Phi\left(\frac{p^* - \rho t}{\sqrt{1-\rho^2}}\right)\right] \\
&= \int_{z^*(t)}^{\infty} e^{(\sigma-1)\left(\sqrt{1-\rho^2}z + \rho t - p^*\right)} \phi(z)\, dz - \left[1 - \Phi\left(z^*(t)\right)\right] \\
&= e^{(\sigma-1)(\rho t - p^*) + \frac{1}{2}(\sigma-1)^2(1-\rho^2)} \int_{z^*(t)}^{\infty} \phi\left(z - (\sigma-1)\sqrt{1-\rho^2}\right) dz - \Phi\left(-z^*(t)\right) \\
&= e^{(\sigma-1)(\rho t - p^*) + \frac{1}{2}(\sigma-1)^2(1-\rho^2)} \Phi\left((\sigma-1)\sqrt{1-\rho^2} - z^*(t)\right) - \Phi\left(-z^*(t)\right) \\
&= e^{(\sigma-1)(\rho t - p^*) + \frac{1}{2}(\sigma-1)^2(1-\rho^2)} \Phi\left(\frac{\rho t - p^* + (\sigma-1)(1-\rho^2)}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{\rho t - p^*}{\sqrt{1-\rho^2}}\right).
\end{aligned}
$$

Therefore, the AC (1) reads:

$$
e^{(\sigma-1)(\rho t^* - p^*) + \frac{1}{2}(\sigma-1)^2(1-\rho^2)} \Phi\left(\frac{\rho t^* - p^* + (\sigma-1)(1-\rho^2)}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{\rho t^* - p^*}{\sqrt{1-\rho^2}}\right) - \frac{\delta f_b(\rho)}{f} = 0,
$$

and, for fixed ρ, it depends on (t^*, p^*) only through the index $x \equiv \rho t^* - p^*$ (with ρ entering as a parameter through the variance term $1 - \rho^2$). Hence the AC locus pins down x to a constant $x(\rho)$, implying a linear relation $p^* = \rho t^* - x(\rho)$. Let $a(\rho) \equiv -x(\rho)$ (equivalently,

$A(\rho) = e^{a(\rho)}$), so the AC curve can be written as $p^* = \rho t^* + a(\rho)$. Substituting into the right-hand side of AC yields a decreasing function of a that crosses zero if $\delta f_b(\rho)/f > 0$. Therefore, $a(\rho)$ (and hence $A(\rho)$) is unique, and it is both decreasing in $f_b(\rho)$ and increasing in f.

For the FE condition, the expected joint profit $\breve{\pi} \equiv \int_{\theta^*}^{\infty} \widetilde{\pi}(\theta) dC(\theta)$ can be expressed as a function of (t^*, p^*):

$$\begin{aligned}
\breve{\pi}\left(t^*, p^*\right) &= \int_{t^*}^{\infty} \widetilde{\pi}\left(e^t\right) \phi\left(t\right) dt \\
&= f \int_{t^*}^{\infty} e^{(\sigma-1)\left(\rho t - p^*\right) + \frac{1}{2}(\sigma-1)^2\left(1-\rho^2\right)} \Phi\left(\frac{\rho t - p^* + (\sigma-1)\left(1-\rho^2\right)}{\sqrt{1-\rho^2}}\right) \phi\left(t\right) dt \\
&\quad - f \int_{t^*}^{\infty} \Phi\left(\frac{\rho t - p^*}{\sqrt{1-\rho^2}}\right) \phi\left(t\right) dt \\
&= f e^{\frac{1}{2}(\sigma-1)^2 - (\sigma-1)p^*} \int_{-\infty}^{-t^* + \rho(\sigma-1)} \Phi\left(\frac{-\rho h' - p^* + (\sigma-1)}{\sqrt{1-\rho^2}}\right) \phi\left(h'\right) dh' \\
&\quad - f \int_{-\infty}^{-t^*} \Phi\left(\frac{-\rho h - p^*}{\sqrt{1-\rho^2}}\right) \phi\left(h\right) dh \\
&= f \left[e^{\frac{1}{2}(\sigma-1)^2 - (\sigma-1)p^*} \Phi_\rho\left(-p^* + \sigma - 1, -t^* + \rho\left(\sigma-1\right)\right) - \Phi_\rho\left(-p^*, -t^*\right) \right],
\end{aligned}$$

where the last line follows from the analysis of the standard normal cumulative distribution's moments as in Owen (1980). Thus, write the FE condition as follows:

$$\begin{aligned}
\mathcal{H}\left(p^*, t^*\right) = e^{\frac{1}{2}(\sigma-1)^2 - (\sigma-1)p^*} \Phi_\rho\left(-p^* + \sigma - 1, -t^* + \rho\left(\sigma-1\right)\right) - \\
- \Phi_\rho\left(-p^*, -t^*\right) - \frac{\delta f_b(\rho)}{f} \Phi\left(-t^*\right) - \frac{\delta f_n}{f} = 0.
\end{aligned}$$

Differentiating $\mathcal{H}(p^*, t^*)$ with respect to the log-productivity threshold p^* (holding t^* and ρ fixed) and using the standard derivative identities for the bivariate normal CDF yields: Let $k \equiv \sigma - 1$ and define $A(p^*) \equiv e^{\frac{1}{2}k^2 - kp^*}$. Write $x_1 \equiv -p^* + k$, $y_1 \equiv -t^* + \rho k$, and $x_0 \equiv -p^*$, $y_0 \equiv -t^*$, so

$$\mathcal{H}(p^*, t^*) = A(p^*)\, \Phi_\rho(x_1, y_1) - \Phi_\rho(x_0, y_0) - \frac{\delta f_b(\rho)}{f} \Phi(-t^*) - \frac{\delta f_n}{f}.$$

Since $A'(p^*) = -kA(p^*)$ and $dx_1/dp^* = dx_0/dp^* = -1$, we have

$$\frac{\partial \mathcal{H}}{\partial p^*} = -kA\, \Phi_\rho(x_1, y_1) - A\, \frac{\partial \Phi_\rho}{\partial x}(x_1, y_1) + \frac{\partial \Phi_\rho}{\partial x}(x_0, y_0).$$

Using $\frac{\partial \Phi_\rho}{\partial x}(x, y) = \phi(x)\Phi\left(\frac{y - \rho x}{\sqrt{1-\rho^2}}\right)$, note first that $y_1 - \rho x_1 = (-t^* + \rho k) - \rho(-p^* + k) =$

$-t^* + \rho p^* = y_0 - \rho x_0$. Second, since $x_1 = x_0 + k$ and $A = e^{\frac{1}{2}k^2 + kx_0}$ (because $x_0 = -p^*$), we have the exponential-tilting identity $A\,\phi(x_1) = \phi(x_0)$, so the last two terms cancel exactly.

$$\frac{\partial \mathcal{H}(p^*, t^*)}{\partial p^*} = -(\sigma - 1)\, e^{\frac{1}{2}(\sigma-1)^2 - (\sigma-1)p^*}\, \Phi_\rho\left(-p^* + \sigma - 1, -t^* + \rho(\sigma - 1)\right) < 0.$$

The inequality follows because $\sigma > 1$, the exponential term is positive, and $\Phi_\rho(\cdot, \cdot) \in (0, 1)$. Moreover, the derivative with respect to the log-signal threshold t^* is shown to be:

$$\frac{\partial \mathcal{H}(p^*, t^*)}{\partial t^*} = -\Bigg[e^{(\sigma-1)(\rho t^* - p^*) + \frac{1}{2}(\sigma-1)^2(1-\rho^2)}\, \Phi\left(\frac{\rho t^* - p^* + (\sigma - 1)(1 - \rho^2)}{\sqrt{1 - \rho^2}}\right) - \\ - \Phi\left(\frac{\rho t^* - p^*}{\sqrt{1 - \rho^2}}\right) - \frac{\delta f_b(\rho)}{f}\Bigg]\phi(t^*),$$

which is not a monotone function of t^*.

LEMMA A1 (Single-peakedness). *Fix p^*. The function $t^* \mapsto \mathcal{H}(p^*, t^*)$ has exactly one stationary point in t^*, at which it achieves its global maximum.*

PROOF. The expression in brackets in $\partial\mathcal{H}/\partial t^*$ equals $\tilde{\pi}(e^{t^*})/f - \delta f_b(\rho)/f$, which is zero precisely when the AC holds, i.e., at $t^* = (p^* - a(\rho))/\rho$. Moreover, $\tilde{\pi}(e^{t^*})$ is weakly increasing in t^* (equivalently in θ) by Assumption 1 (and in the log-normal benchmark this is implied by $\rho > 0$). Evaluating the sign of $\partial\mathcal{H}/\partial t^*$: for $t^* < (p^* - a(\rho))/\rho$, we have $\tilde{\pi}(e^{t^*}) < \delta f_b(\rho)$, so the bracket is negative and the derivative is positive (since it carries a leading minus sign and $\phi(t^*) > 0$). Symmetrically, for $t^* > (p^* - a(\rho))/\rho$, the derivative is negative. At the boundaries, $\lim_{t^* \to \pm\infty} \partial\mathcal{H}/\partial t^* = 0$ because $\phi(t^*) \to 0$. Hence $\mathcal{H}(p^*, t^*)$ rises to a unique maximum at $t^* = (p^* - a(\rho))/\rho$ and falls thereafter. □

These properties give rise to the pattern depicted in Figure 1. Furthermore, the line $p^* = \rho t^* + a(\rho)$ can intersect the FE-based implicit function of p^* in t^* only at its stationary point, since $a(\rho)$ is uniquely determined by the AC. A single stationary point guarantees a unique intersection, and thus a unique equilibrium.

A.2. Microfoundation for staged entry

The equilibrium allocation is fully characterized by the threshold pair (θ^*, φ^*): all entrepreneurs with signals $\theta \geq \theta^*$ are activated, all firms with productivity $\varphi \geq \varphi^*$ operate. The signal cutoff θ^* is determined by the AC (1) and FE (2) conditions, which depend only on expected profits $\tilde{\pi}(\theta)$, cost parameters $(f_n, f_b(\rho), \delta)$, and the joint distribution of (θ, φ).

We assume: (i) financiers observe the signal θ and can condition contracts on it; (ii) competition is type-by-type (no cross-subsidization across signals); and (iii) contracts are feasible in the sense that profit shares $b(\theta) \in [0,1]$ on the activated set. Feasibility holds because $b(\theta) = \delta f_b(\rho)/\tilde{\pi}(\theta)$ and $\tilde{\pi}(\theta) \geq \tilde{\pi}(\theta^*) = \delta f_b(\rho)$ for all $\theta \geq \theta^*$ (by Assumption 1), so $b(\theta) \leq 1$.

Under competitive, risk-neutral financiers, the zero-profit condition pins down the contract schedule: $b(\theta) = \delta f_b(\rho)/\tilde{\pi}(\theta)$ for all $\theta \geq \theta^*$. This schedule affects only the division of surplus between entrepreneurs and financiers—not the cutoff θ^*, the set of operating firms, or any aggregate outcome. Any alternative competitive contract form (upfront fees, actuarially fair loans, equity stakes) that satisfies zero expected profit at each θ yields the identical cutoff and allocation.

The key insight is that the equilibrium depends on two primitives: (i) the two-stage information structure (entrepreneurs observe θ before activation, φ after), and (ii) the resource costs $(f_n, f_b(\rho))$. This microfoundation provides one interpretation of the activation stage. Since competitive intermediation is allocation-equivalent to self-funding, the results hinge on staged entry and gatekeeping rather than on credit-market imperfections.

A.3. Welfare continuity

Proof of Proposition 2. We establish continuity of welfare in ρ by continuity of (i) the equilibrium thresholds and (ii) the mapping from thresholds to aggregates.

Thresholds. We exploit the equilibrium geometry from Appendix A.1 to avoid Jacobian regularity assumptions. Let $t^*(\rho) \equiv \log \theta^*(\rho)$ and $p^*(\rho) \equiv \log \varphi^*(\rho)$. Under Assumption 3, the AC depends on (t^*, p^*, ρ) only through the difference $a \equiv p^* - \rho t^*$. Let $k \equiv \sigma - 1$ and write $u \equiv p - \rho t$; then $u \mid t \sim N(0, 1-\rho^2)$ and, for any pair (t, p^*) with $a = p^* - \rho t$,

$$\frac{\tilde{\pi}(e^t)}{f} = \mathbb{E}\Big(\Big(e^{k(u-a)} - 1\Big)\, \mathbb{1}(u \geq a)\Big) = \int_a^\infty \Big(e^{k(u-a)} - 1\Big)\, \phi_{\sqrt{1-\rho^2}}(u)\, du,$$

where $\phi_{\sqrt{1-\rho^2}}$ is the $N(0, 1-\rho^2)$ density. By Leibniz' rule, the right-hand side is strictly decreasing in a:

$$\frac{\partial}{\partial a} \int_a^\infty \Big(e^{k(u-a)} - 1\Big)\, \phi_{\sqrt{1-\rho^2}}(u)\, du = -k \int_a^\infty e^{k(u-a)} \phi_{\sqrt{1-\rho^2}}(u)\, du < 0.$$

Since $f_b(\rho)$ is continuous and strictly positive, the AC equation $\tilde{\pi}(e^{t^*})/\delta = f_b(\rho)$ therefore pins down a unique $a(\rho)$ for each ρ, and $a(\rho)$ is continuous by continuity and strict monotonicity in a. Formally, define $F(a, \rho) \equiv \tilde{\pi}(e^t)/(\delta f) - f_b(\rho)/f$ using the representation above; F is continuous in (a, ρ) and strictly decreasing in a, so for any sequence $\rho_n \to \rho$ any limit point of $a(\rho_n)$ must solve $F(a, \rho) = 0$; uniqueness implies $a(\rho_n) \to a(\rho)$.

Given $a(\rho)$, consider the FE residual $\mathcal{H}(p, t; \rho)$ from Appendix A.1. Along the AC line $p = \rho t + a(\rho)$, Lemma A1 implies that t is the (unique) maximizer of $t' \mapsto \mathcal{H}(p, t'; \rho)$ given

p, hence $\partial\mathcal{H}/\partial t(\rho t + a(\rho), t; \rho) = 0$. Define the one-dimensional function

$$J(t;\rho) \equiv \mathcal{H}(\rho t + a(\rho), t; \rho).$$

Along the AC line $p = \rho t + a(\rho)$, the AC holds by construction for all t; equivalently, the bracket term in the expression for $\partial\mathcal{H}/\partial t$ is identically zero, so $\frac{\partial\mathcal{H}}{\partial t}(\rho t + a(\rho), t; \rho) = 0$ (as in Lemma A1). Hence the chain rule simplifies to

$$\frac{d}{dt}J(t;\rho) = \rho\,\frac{\partial\mathcal{H}}{\partial p}(\rho t + a(\rho), t; \rho).$$

Because $\partial\mathcal{H}/\partial p < 0$ everywhere (Appendix A.1), we obtain $\frac{d}{dt}J(t;\rho) < 0$, so $J(\cdot;\rho)$ is strictly decreasing in t. By Proposition 1, there is a unique $t^*(\rho)$ such that $J(t^*(\rho);\rho) = 0$, and $p^*(\rho) = \rho t^*(\rho) + a(\rho)$. Since $\mathcal{H}$ is continuous in (p, t, ρ) (it is built from normal CDFs/PDFs), and $a(\rho)$ is continuous, J is continuous in (t, ρ). Continuity and strict monotonicity in t imply that the unique root $t^*(\rho)$ varies continuously with ρ (an analogous limit-point argument applies), and thus so does $p^*(\rho)$.

Aggregates. Given $(p^*(\rho), t^*(\rho), \rho)$, the equilibrium tail probabilities and truncated moments under the bivariate normal are continuous in these arguments, so $\mathcal{P}^*_\theta(\rho)$, $\mathcal{P}^*_\varphi(\rho)$, $\widetilde{\varphi}(\rho)$, $M(\rho)$, and therefore $W(\rho)$ are continuous on $(0,1)$. □

A.4. Welfare decline under bounded costs

For the proof it is convenient to write $W(\rho; f_b)$ for equilibrium welfare at precision ρ when the activation cost is the *constant* level f_b (holding all other primitives fixed). Proposition 3 then follows by constructing a bounded, monotone schedule that maps ρ_L to a low cost and ρ_H to a high (but finite) cost.

LEMMA A2 (For fixed ρ, welfare vanishes as $f_b \to \infty$). *Maintain Assumptions 1–3 and assume* $f_n > 0$. *Fix any* $\rho \in (0,1)$ *and consider equilibria indexed by the activation-cost level* $f_b > 0$ *(holding all other primitives fixed). Then* $\lim_{f_b\to\infty} W(\rho; f_b) = 0$.

PROOF. Fix $\rho \in (0,1)$ and let $\{f_{b,n}\}_{n\geq 1}$ be any sequence with $f_{b,n} \to +\infty$. Let $(\theta^*_n, \varphi^*_n, M_{e,n})$ denote an equilibrium under $(\rho, f_b = f_{b,n})$, write $t^*_n \equiv \log\theta^*_n$, and let $\mathcal{P}^*_{\theta,n} \equiv \Pr(\theta \geq \theta^*_n) = \Phi(-t^*_n)$. From the resource constraint (6),

$$L \geq M_{e,n}\big(f_n + \mathcal{P}^*_{\theta,n} f_{b,n}\big) \geq M_{e,n} f_n,$$

so $M_{e,n} \leq L/f_n$ for all n.

Case 1: $\mathcal{P}^*_{\theta,n}$ *is bounded away from zero along a subsequence.* Suppose there exists a subsequence $\{n_j\}_{j\geq 1}$ and $\varepsilon > 0$ such that $\mathcal{P}^*_{\theta,n_j} \geq \varepsilon$ for all j. Since $f_{b,n_j} \to +\infty$, we have

$\mathcal{P}^*_{\theta,n_j} f_{b,n_j} \to +\infty$ and therefore

$$M_{e,n_j} \le \frac{L}{f_n + \mathcal{P}^*_{\theta,n_j} f_{b,n_j}} \xrightarrow[j\to\infty]{} 0.$$

Using (5) and $\mathbb{1}(\theta \ge \theta^*_n,\ \varphi \ge \varphi^*_n) \le 1$, we obtain

$$W(\rho; f_{b,n_j})^{\sigma-1} \le \left(\tfrac{\sigma-1}{\sigma}\right)^{\sigma-1} \frac{M_{e,n_j}}{\delta}\, \mathbb{E}\left(\varphi^{\sigma-1}\right).$$

Under Assumption 3, $\mathbb{E}\left(\varphi^{\sigma-1}\right) = \exp\left(\tfrac{1}{2}(\sigma-1)^2\right) < \infty$, so $W(\rho; f_{b,n_j}) \to 0$.

Case 2: $\mathcal{P}^*_{\theta,n} \to 0$ along a subsequence. Otherwise, there exists a subsequence $\{n_j\}_{j\ge1}$ such that $\mathcal{P}^*_{\theta,n_j} \to 0$, hence $t^*_{n_j} \to +\infty$. Under Assumption 3, with $(p,t) = (\log\varphi, \log\theta)$ jointly normal and $k \equiv \sigma - 1$, the exponential-tilting identity gives

$$\mathbb{E}\left(\varphi^k\, \mathbb{1}\left(\theta \ge \theta^*_{n_j}\right)\right) = \mathbb{E}\left(e^{kp}\, \mathbb{1}\left(t \ge t^*_{n_j}\right)\right) = \exp\left(\tfrac{1}{2}k^2\right)\Phi\left(-(t^*_{n_j} - \rho k)\right) \xrightarrow[j\to\infty]{} 0,$$

because $t^*_{n_j} \to +\infty$. Since $\{\varphi \ge \varphi^*_{n_j}, \theta \ge \theta^*_{n_j}\} \subseteq \{\theta \ge \theta^*_{n_j}\}$, we have

$$\mathbb{E}\left(\varphi^k\, \mathbb{1}\left(\varphi \ge \varphi^*_{n_j}, \theta \ge \theta^*_{n_j}\right)\right) \le \mathbb{E}\left(\varphi^k\, \mathbb{1}\left(\theta \ge \theta^*_{n_j}\right)\right),$$

and therefore, using (5) with $M_{e,n_j} \le L/f_n$, we obtain $W(\rho; f_{b,n_j}) \to 0$.

Conclusion. In either case, every subsequence of $\{W(\rho; f_{b,n})\}_{n\ge1}$ admits a further subsequence converging to 0, hence $W(\rho; f_{b,n}) \to 0$. Since the sequence $\{f_{b,n}\}$ was arbitrary, $\lim_{f_b\to\infty} W(\rho; f_b) = 0$. □

Proof of Proposition 3. Fix any $0 < \rho_L < \rho_H < 1$. Choose any finite $f_\ell > 0$ and let $W_L \equiv W(\rho_L; f_\ell)$ denote equilibrium welfare under the constant-cost regime $(\rho_L, f_b = f_\ell)$. Under the maintained assumptions (and equilibrium existence for each finite $f_b > 0$), we have $W_L > 0$.

Apply Lemma A2 at $\rho = \rho_H$ to select a finite $f_h > f_\ell$ such that

$$W(\rho_H; f_h) < W(\rho_L; f_\ell) = W_L.$$

Define the bounded, continuous, weakly increasing schedule $f_b : (0,1) \to \mathbb{R}_+$ by

$$f_b(\rho) = \begin{cases} f_\ell & \rho \le \rho_L, \\ f_\ell + (f_h - f_\ell)\dfrac{\rho - \rho_L}{\rho_H - \rho_L} & \rho \in [\rho_L, \rho_H], \\ f_h & \rho \ge \rho_H. \end{cases}$$

This schedule is bounded with $\sup_{\rho\in(0,1)} f_b(\rho) = f_h < \infty$, continuous, and weakly increasing. Moreover, by construction,

$$W(\rho_L) = W(\rho_L; f_\ell), \qquad W(\rho_H) = W(\rho_H; f_h),$$

so $W(\rho_H) < W(\rho_L)$. □

A.5. Unbounded cost schedules

Proof of the claims in Remark 1. Let $t^*(\rho) \equiv \log\theta^*(\rho)$ and $\mathcal{P}^*_\theta(\rho) = \Phi(-t^*(\rho))$. From the resource constraint (6), total labor used for experimentation and activation satisfies

$$L \geq M_e(\rho)\big(f_n + \mathcal{P}^*_\theta(\rho) f_b(\rho)\big) \geq M_e(\rho)\, f_n,$$

so $M_e(\rho) \leq L/f_n$ for all ρ.

Case 1: $\mathcal{P}^*_\theta(\rho)$ is bounded away from zero near $\rho = 1$. Suppose $\liminf_{\rho\uparrow 1} \mathcal{P}^*_\theta(\rho) > 0$. Then there exist $\varepsilon > 0$ and $\rho_0 \in (0,1)$ such that $\mathcal{P}^*_\theta(\rho) \geq \varepsilon$ for all $\rho \in (\rho_0, 1)$. Since $f_b(\rho) \to +\infty$ by hypothesis, it follows that $\mathcal{P}^*_\theta(\rho) f_b(\rho) \to +\infty$, and therefore

$$M_e(\rho) \leq \frac{L}{f_n + \mathcal{P}^*_\theta(\rho) f_b(\rho)} \xrightarrow[\rho\uparrow 1]{} 0.$$

Using (5) and $\mathbb{1}(\theta \geq \theta^*(\rho),\ \varphi \geq \varphi^*(\rho)) \leq 1$, we obtain

$$W(\rho)^{\sigma-1} \leq \left(\tfrac{\sigma-1}{\sigma}\right)^{\sigma-1} \frac{M_e(\rho)}{\delta}\, \mathbb{E}\big(\varphi^{\sigma-1}\big).$$

Moreover, $\mathbb{E}\big(\varphi^{\sigma-1}\big) = \exp\big(\tfrac{1}{2}(\sigma-1)^2\big) < \infty$ under Assumption 3, so $W(\rho) \to 0$ as $\rho \uparrow 1$.

Case 2: $\mathcal{P}^*_\theta(\rho) \to 0$ as $\rho \uparrow 1$. Otherwise, $\liminf_{\rho\uparrow 1} \mathcal{P}^*_\theta(\rho) = 0$. In particular, along any sequence $\rho_n \uparrow 1$ for which $\mathcal{P}^*_\theta(\rho_n) \to 0$, we have $t^*(\rho_n) \to +\infty$. Under Assumption 3, with $(p, t) = (\log\varphi, \log\theta)$ jointly normal and $k \equiv \sigma - 1$, the exponential-tilting identity implies

$$\mathbb{E}\Big(\varphi^k\, \mathbb{1}(\theta \geq \theta^*(\rho_n))\Big) = \mathbb{E}\Big(e^{kp}\, \mathbb{1}\big(t \geq t^*(\rho_n)\big)\Big) = \exp\Big(\tfrac{1}{2}k^2\Big)\, \Phi\big(-(t^*(\rho_n) - \rho_n k)\big) \xrightarrow[n\to\infty]{} 0,$$

because $t^*(\rho_n) \to +\infty$. Since $\{\varphi \geq \varphi^*(\rho_n), \theta \geq \theta^*(\rho_n)\} \subseteq \{\theta \geq \theta^*(\rho_n)\}$, we have

$$\mathbb{E}\Big(\varphi^k\, \mathbb{1}\big(\varphi \geq \varphi^*(\rho_n), \theta \geq \theta^*(\rho_n)\big)\Big) \leq \mathbb{E}\Big(\varphi^k\, \mathbb{1}\big(\theta \geq \theta^*(\rho_n)\big)\Big).$$

With $M_e(\rho_n) \leq L/f_n$, the same bound in (5) implies $W(\rho_n) \to 0$.

Conclusion. In either case, $W(\rho) \to 0$ as $\rho \uparrow 1$.

Finally, if welfare is not maximized at arbitrarily low precision, i.e., there exist $0 < \underline{\rho} < \bar{\rho} < 1$ such that $W(\bar{\rho}) > W(\underline{\rho})$, then by the limit $W(\rho) \to 0$ as $\rho \uparrow 1$ we can choose $\rho_1 \in (\bar{\rho}, 1)$

sufficiently close to one that $W(\rho_1) < W(\bar{\rho})$. Continuity of W (Proposition 2) implies that W attains a maximum on the compact interval $[\underline{\rho}, \rho_1] \subset (0,1)$, and the maximizer cannot be at $\underline{\rho}$ or at ρ_1; hence there exists at least one maximizer $\rho^W \in (0,1)$. □

A.6. Welfare elasticity condition

Start from (5) and define

$$S(\rho) \equiv \mathbb{E}_{\varphi^{\sigma-1} \mathbb{1}(\theta \geq \theta^*(\rho), \varphi \geq \varphi^*(\rho))}.$$

Using labor market clearing, total input used equals production input plus experimentation and activation input:

$$L = M(\rho)\big(\bar{r}(\rho) - \bar{\pi}(\rho)\big) + M_e(\rho) f_n + \mathcal{P}^*_\theta(\rho) M_e(\rho) f_b(\rho).$$

Using the steady-state flow condition $\delta M(\rho) = \mathcal{P}^*_\varphi(\rho) M_e(\rho)$, this implies

$$L = M_e(\rho)\left[f_n + \mathcal{P}^*_\theta(\rho) f_b(\rho) + \frac{\mathcal{P}^*_\varphi(\rho)}{\delta}\big(\bar{r}(\rho) - \bar{\pi}(\rho)\big)\right] \equiv M_e(\rho)\, B(\rho),$$

so $B(\rho) \equiv L/M_e(\rho)$, and hence

$$M_e(\rho) = \frac{L}{B(\rho)}.$$

Substituting into (5) yields

$$W(\rho)^{\sigma-1} = \left(\tfrac{\sigma-1}{\sigma}\right)^{\sigma-1} \frac{L}{\delta} \cdot \frac{S(\rho)}{B(\rho)}.$$

Taking logs and differentiating with respect to ρ gives Corollary 1.

A.7. Planner analysis and efficiency results

Following Section 3.3, this appendix derives the constrained planner's marginal payoff kernel $\widetilde{G}(\theta)$ and proves the efficiency results stated in the main text. The kernel measures the net social value of activating a marginal firm at signal θ, aggregating effects on variety, expected productivity, and verification costs. Under CES, single-crossing holds automatically (Corollary 2), and the constrained-efficient cutoff coincides with the decentralized cutoff (Lemma 1).

Because constant markups align private and social incentives at the activation margin, we do not need to set up a full constrained optimization problem with Lagrange multipliers. Instead, we derive the marginal payoff kernel directly via an accounting argument that exploits CES structure. The section proceeds as follows. First, we give a CES derivation of the marginal kernel $\widetilde{G}(\theta)$ that makes the constant-markup cancellation explicit. Second, we show that single-crossing of $\widetilde{G}(\theta)$ implies cutoff optimality (Proposition 4). Third, we provide an analytic lemma establishing monotonicity of key conditional objects under lognormality. Finally, we prove Lemma 1 and Corollary 2.

A.7.1. Kernel derivation

In the CES benchmark, the planner's marginal value of raising activation at signal θ can be expressed directly in terms of private surplus, without tracking how the operating cutoff φ^* adjusts with the activation rule. The argument is an accounting identity under constant markups.

With CES preferences and optimal prices $p(\varphi) = \frac{\sigma}{\sigma-1}\frac{1}{\varphi}$, welfare equals the inverse price index. Using $P^{1-\sigma} = \int_{\omega\in\Omega} p(\varphi(\omega))^{1-\sigma} d\omega$, we obtain

$$W^{\sigma-1} = P^{1-\sigma} = \left(\tfrac{\sigma-1}{\sigma}\right)^{\sigma-1} \int_{\omega\in\Omega} \varphi(\omega)^{\sigma-1}\, d\omega.$$

Thus, up to the constant $\left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1}$, welfare depends on the stock of operating firms only through their total "weight" $\int_{\omega\in\Omega} \varphi(\omega)^{\sigma-1} d\omega$.

Now consider a constrained planner who takes the mass of experimenters M_e as given and chooses an activation rule $m(\theta) \in [0, 1]$. This is a partial-derivative (within-regime) object: the experimentation margin is determined by free entry given the activation rule, but the activation cutoff can be assessed conditional on M_e. Increasing activation at a particular signal value θ by an infinitesimal amount adds an extra flow mass of activated entrepreneurs proportional to $M_e\, dC(\theta)$. In steady state, the induced stock effect scales by expected lifetime $1/\delta$. Conditional on signal θ, only firms with $\varphi \geq \varphi^*$ operate, so the expected contribution to the CES weight per activated entrepreneur is

$$\Psi(\theta) \equiv \mathbb{E}\Big(\varphi^{\sigma-1}\, \mathbb{1}\big(\varphi \geq \varphi^*\big) \mid \theta\Big).$$

Hence the marginal welfare benefit from raising activation at θ is proportional to $\Psi(\theta)/\delta$.

Under CES, revenue and productivity are proportional:

$$r(\varphi) = \sigma f \left(\frac{\varphi}{\varphi^*}\right)^{\sigma-1} = \frac{\sigma f}{(\varphi^*)^{\sigma-1}} \varphi^{\sigma-1}.$$

Let $\widetilde{r}(\theta) \equiv \mathbb{E}(r(\varphi)\, \mathbb{1}(\varphi \geq \varphi^*) \mid \theta)$. Because the proportionality factor $\sigma f/(\varphi^*)^{\sigma-1}$ does not depend on θ, we have $\widetilde{r}(\theta) \propto \Psi(\theta)$, and thus the marginal welfare benefit can equivalently be written as a positive multiple of $\widetilde{r}(\theta)/\delta$.

Activation itself uses $f_b(\rho)$ units of labor. In addition, when an activated firm draws φ and operates, it uses flow labor equal to its fixed plus variable input. Under CES pricing with markup $\mu = \frac{\sigma}{\sigma-1}$, variable labor satisfies $q(\varphi)/\varphi = r(\varphi)/\mu = \frac{\sigma-1}{\sigma} r(\varphi)$. Therefore total flow labor for an operating firm is

$$f + \frac{q(\varphi)}{\varphi} = f + \frac{\sigma - 1}{\sigma} r(\varphi) = r(\varphi) - \pi(\varphi),$$

where the last equality uses $\pi(\varphi) = r(\varphi)/\sigma - f$. Let $\widetilde{\pi}(\theta) \equiv \mathbb{E}(\pi(\varphi)\,\mathbb{1}(\varphi \geq \varphi^*) \mid \theta)$. Then the expected present-value production labor induced by an activation at θ equals $(\widetilde{r}(\theta) - \widetilde{\pi}(\theta))/\delta$.

Finally, let $\kappa > 0$ denote the positive conversion factor that puts the marginal welfare gain and the marginal labor cost on a common scale (equivalently, absorb the constant CES scaling and the planner's resource multiplier into κ). From the previous steps, the planner's marginal net gain from raising activation at signal θ has the same sign as

$$\frac{\widetilde{r}(\theta)}{\delta} - \frac{\widetilde{r}(\theta) - \widetilde{\pi}(\theta)}{\delta} - f_b(\rho) = \frac{\widetilde{\pi}(\theta)}{\delta} - f_b(\rho).$$

Since all suppressed factors in the proportionality are positive, we normalize the marginal payoff kernel as

$$\widetilde{G}(\theta) \equiv \frac{\widetilde{\pi}(\theta)}{\delta} - f_b(\rho).$$

A.7.2. Proof of Proposition 4

Proof of Proposition 4. Let $\widetilde{G}(\theta)$ denote the constrained planner's marginal payoff kernel from raising the activation probability at signal θ (defined up to a positive normalization). Since $M_e > 0$ and $dC(\theta) > 0$ for all $\theta > 0$, the marginal effect of raising $m(\theta)$ is proportional to $\widetilde{G}(\theta)$, so its sign is determined by the sign of $\widetilde{G}(\theta)$. Therefore an optimal rule can be chosen so that $m(\theta) = 1$ whenever $\widetilde{G}(\theta) > 0$ and $m(\theta) = 0$ whenever $\widetilde{G}(\theta) < 0$ (with any $m(\theta) \in [0,1]$ on $\{\theta : \widetilde{G}(\theta) = 0\}$). If $\widetilde{G}(\theta)$ is weakly increasing in θ (the single-crossing condition), then $\{\theta : \widetilde{G}(\theta) \geq 0\}$ is an interval of the form $[\theta^P, \infty)$, hence the optimal activation rule is a cutoff. □

A.7.3. Supporting results for single-crossing

LEMMA A3 (Monotonicity of building blocks under bivariate normal). *Maintain Assumption 3 (lognormal). Fix the cutoffs (t^*, p^*) and any $\rho \in (0,1)$. Then the following functions are weakly increasing in θ (equivalently in $t = \log\theta$): (i) $\Phi(-z^*(\log\theta)) = \Pr(p \geq p^* \mid \theta)$; (ii) $\Psi(\theta) = \mathbb{E}\big(\varphi^{\sigma-1}\,\mathbb{1}(\varphi \geq \varphi^*) \mid \theta\big)$; (iii) $\widetilde{\pi}(\theta) = \mathbb{E}(\pi(\varphi)\,\mathbb{1}(\varphi \geq \varphi^*) \mid \theta)$.*

PROOF. Under Assumption 3, conditional log productivity satisfies $p \mid t \sim N(\rho t, 1 - \rho^2)$. As t increases, the conditional mean shifts right while the variance is constant, so $p \mid t$ increases in the sense of first-order stochastic dominance. Hence, for any weakly increasing function $h(\cdot)$, $\mathbb{E}(h(p) \mid t)$ is weakly increasing in t. The three claims follow by taking $h(p) = \mathbb{1}(p \geq p^*)$, $h(p) = e^{(\sigma-1)p}\,\mathbb{1}(p \geq p^*)$, and $h(p) = \pi(e^p)\,\mathbb{1}(p \geq p^*)$, which is weakly increasing in p because profits $\pi(\varphi)$ are weakly increasing in productivity for $\varphi \geq \varphi^*$. □

A.7.4. Proof of Lemma 1

Proof of Lemma 1. Appendix A.7.1 shows that, under CES, the constrained planner's marginal payoff kernel can be normalized as $\tilde{\pi}(\theta)/\delta - f_b(\rho)$. Hence the constrained planner activates if and only if $\tilde{\pi}(\theta) \geq \delta f_b(\rho)$, which is exactly the decentralized activation condition (1). By Assumption 1, $\tilde{\pi}(\theta)$ is weakly increasing in θ, so the constrained-efficient rule is a cutoff at θ^*. Therefore $\theta^P = \theta^*$ and $W^P = W^D$. □

A.7.5. Proof of Corollary 2

Proof of Corollary 2. Appendix A.7.1 shows that, under CES, the planner's marginal net gain from activation at signal θ is proportional to $\tilde{\pi}(\theta)/\delta - f_b(\rho)$. Since $\tilde{\pi}(\theta)$ is weakly increasing in θ by Assumption 1, the single-crossing condition holds. □

A.8. Decentralizing cutoff rules

This appendix describes two non-distortionary ways to implement a cutoff activation rule. In the CES benchmark used in the paper, the constrained-efficient cutoff coincides with the decentralized cutoff (Lemma 1), so these instruments do not change the allocation; the constructions are useful for environments in which a cutoff wedge arises.

Fix a cutoff θ^P and consider the policy environments described below. For any cutoff rule $m(\theta) = \mathbb{1}\big(\theta \geq \theta^P\big)$, the steady-state allocation is pinned down by firms' operating rule ($\varphi \geq \varphi^*$), the steady-state flow condition $\delta M = \mathcal{P}^*_\varphi M_e$, and the resource constraint (6) together with the definition of $\tilde{\varphi}$. The remaining equilibrium object M_e is determined by free entry: entrepreneurs enter experimentation until their expected net payoff equals f_n. Thus, to implement the constrained planner allocation associated with cutoff θ^P, it suffices to implement the activation set $\{\theta \geq \theta^P\}$ and preserve the free-entry condition under that set.

Regulatory approval rule. Suppose a regulator enforces that activation is permitted if and only if $\theta \geq \theta^P$. Then the set of activated entrepreneurs is exactly $\{\theta \geq \theta^P\}$, so the induced activation rule coincides with the planner's rule $m(\theta) = \mathbb{1}\big(\theta \geq \theta^P\big)$. Given this activation set, each activated firm observes φ and produces if and only if $\varphi \geq \varphi^*$, so post-activation behavior and selection coincide with the model's operating stage. Finally, entrepreneurs' entry decision is governed by free entry under the same activation rule: the expected net value of experimentation equals $\int_{\theta \geq \theta^P} \big(\tilde{\pi}(\theta)/\delta - f_b(\rho)\big)\, dC(\theta)$, and in the planner allocation this equals f_n by the free-entry constraint. Hence, the equilibrium mass of experimenters under the approval rule coincides with the planner's M_e and the full constrained-planner allocation is implemented.

Activation transfer and entry fee. Now allow activation decisions to be private but introduce a constant per-activation transfer s (paid upon activation, possibly negative) and an entry fee τ (paid upon entering experimentation, possibly negative). Given any equilibrium objects, a type with signal θ activates if and only if the lifetime value of activation net of the transfer is non-negative:

$$\frac{\tilde{\pi}(\theta)}{\delta} - f_b(\rho) + s \geq 0.$$

Set $s \equiv f_b(\rho) - \tilde{\pi}(\theta^P)/\delta$. Then the activation condition becomes $\tilde{\pi}(\theta) \geq \tilde{\pi}(\theta^P)$, which holds if and only if $\theta \geq \theta^P$ because $\tilde{\pi}(\theta)$ is weakly increasing in θ under Assumption 1. Thus the transfer implements the cutoff rule θ^P.

Under this cutoff, the expected net value of entering experimentation (before paying f_n) equals

$$\int_{\theta \geq \theta^P} \left(\frac{\tilde{\pi}(\theta)}{\delta} - f_b(\rho) + s \right) dC(\theta) - \tau.$$

Choose $\tau \equiv \int_{\theta \geq \theta^P} s \, dC(\theta)$. Then the free-entry condition becomes

$$\int_{\theta \geq \theta^P} \left(\frac{\tilde{\pi}(\theta)}{\delta} - f_b(\rho) \right) dC(\theta) = f_n,$$

which is exactly the free-entry restriction satisfied by the constrained planner at cutoff θ^P. Therefore the equilibrium mass of experimenters is the same as in the planner allocation, and so are the induced objects $(M, M_e, \tilde{\varphi}, \varphi^*)$.

Budget balance. Total transfers to activated entrepreneurs equal $M_e \int_{\theta \geq \theta^P} s \, dC(\theta)$, while total entry fees equal $M_e \tau$. With $\tau = \int_{\theta \geq \theta^P} s \, dC(\theta)$ these are equal, so the policy is self-financed in expectation and requires no net lump-sum taxes.

A.9. Pigouvian irrelevance

PROOF. A per-activation transfer s changes the private activation condition to

$$\frac{\tilde{\pi}(\theta)}{\delta} - f_b(\rho) + s \geq 0,$$

which implements a cutoff rule $\theta(s)$ since $\tilde{\pi}(\theta)$ is increasing in θ. Choose the experimentation-stage transfer as in Appendix A.8 so that free entry holds under $\theta(s)$. By Lemma 1, the *laissez-faire* cutoff θ^* is constrained-efficient, so moving the cutoff cannot raise welfare. Hence welfare is maximized at $s = 0$, which reproduces $\theta(s) = \theta^*$. □