

Strategic Licensing and Innovation Incentives in Oligopoly

Byung-Cheol Kim*

November 4, 2025

Abstract

This paper analyzes how a research lab should choose the breadth of technology licensing when potential licensees can also innovate. In an oligopoly with endogenous R&D effort, broader licensing spreads cost reductions but weakens incentives to invest. The lab's optimal breadth balances rent extraction and innovation erosion, producing an interior solution. Under nonlinear demand, correlation across firms' R&D outcomes breaks neutrality and depresses aggregate effort. Social welfare is single-peaked in licensing breadth, implying that private licensing is too narrow relative to the social optimum.

Keywords: Technology licensing; Innovation incentives; Oligopoly; Correlated innovation; Welfare.

JEL Codes: L13; O31; L24; D43.

*Department of Economics, Finance, and Legal Studies, University of Alabama. Email: bkim34@ua.edu.

1 Introduction

How broadly should an innovator license a cost-saving technology when potential licensees can also engage in their own R&D? This question lies at the intersection of two classic literatures in industrial organization—the theory of innovation incentives and the theory of technology licensing—but their integration has remained surprisingly limited. The innovation literature has long emphasized how competition shapes incentives to invest in cost-reducing innovation. In the seminal work of Arrow (1962), the replacement effect implies that monopolists invest less in innovation than potential entrants, since successful discovery replaces their own rents. Subsequent studies formalized this idea and examined how market structure affects R&D rivalry: Gilbert and Newbery (1982) showed that incumbents may invest preemptively to maintain dominance, whereas Reinganum (1983, 1989) found that entrants often have stronger incentives when innovations are drastic.

Parallel to this tradition, the licensing literature has explored how exclusive versus non-exclusive access to a new technology shapes competition and diffusion. Pioneering models by Katz and Shapiro (1985) and Kamien and Tauman (1986) examined the optimal design of fixed-fee and royalty contracts in Cournot markets, while Gallini (1984) and Farrell and Gallini (1988) demonstrated that incumbents may strategically license to rivals—either to deter duplicative R&D or to commit to lower future prices and encourage adoption. Despite their conceptual overlap, these two branches have evolved largely in isolation: most innovation models treat licensing as exogenous, while licensing models assume innovation effort is given or binary. In practice, however, the two margins are deeply intertwined. Firms often decide simultaneously whether to license and how much to invest in innovation within the same strategic environment.

Empirically, the breadth of licensing—the number of licensees granted access to a new technology—varies widely across industries and over time. Pharmaceuticals and proprietary software frequently rely on exclusive or narrowly licensed processes to sustain high returns,

while semiconductors, green technologies, and interoperability standards diffuse through broader or even open licensing. These differences reflect not only appropriability conditions but also the correlation structure of innovation efforts. Firms often share research infrastructures, scientific inputs, or data resources that make their R&D outcomes correlated. Such interdependence creates an additional layer of strategic complexity: the value of holding a license depends not only on how many firms receive it, but also on how likely non-licensees are to achieve similar innovations through correlated research. Understanding how licensing breadth and correlated innovation interact is essential for designing innovation policy in markets where competition, cooperation, and duplication coexist.

This paper develops a unified framework that endogenizes both *licensing breadth* and *innovation effort* within an oligopolistic industry. An upstream research lab discovers a process innovation that lowers marginal cost from c_0 to $c_1 < c_0$ and chooses how many downstream producers to license, at what price, and under what contract form. Firms that do not license may invest in their own R&D with success probability $p(x_i)$ determined by effort x_i and cost $K(x_i)$, and successes may be correlated through a common technological shock or network structure. Market competition is Cournot in the baseline, but the analysis accommodates more general conduct through a reduced-form parameter. In this setting, broader licensing deterministically diffuses the innovation but reduces the residual prize for non-licensees, weakening their incentive to invest. The interaction between these forces determines equilibrium effort, the innovator's optimal licensing breadth, and the resulting welfare distribution.

The analysis yields several insights. First, equilibrium R&D effort declines with licensing breadth. As more firms obtain the cost-reducing technology deterministically, the marginal return to innovation falls, reducing private investment. Second, the innovator's licensing profit is single-peaked: exclusive licensing yields high rents per firm but limits diffusion, while excessive breadth intensifies competition and erodes revenue. Hence, the optimal breadth balances rent extraction and diffusion. Third, social welfare is also single-peaked

but maximized at broader diffusion than the private optimum. The innovator ignores consumer surplus gains from diffusion, leading to systematic under-diffusion relative to the social benchmark. This welfare gap is most pronounced when innovations are moderate in size or when competition is intense. Finally, when innovation outcomes are correlated, the model identifies a *neutrality-break channel*: under linear demand and Cournot competition, correlation affects welfare but not incentives. However, once market curvature or competition departs from Cournot, this knife-edge breaks: positive dependence compresses payoff variance and reduces equilibrium effort. In correlated research environments, diffusion thus entails a higher incentive cost.

The framework connects and extends several strands of industrial organization. It links innovation races (Cardon and Sasaki, 1998) with strategic licensing (Gallini, 1984; Farrell and Gallini, 1988), allowing licensing breadth and R&D investment to be jointly determined. This feedback is largely absent in prior work. The model also relates to theories of cooperative R&D and research joint ventures (D’Aspremont and Jacquemin, 1988a; Katz, 1986; Kamien et al., 1992), where firms internalize spillovers. Here, cooperation is endogenous—captured by licensing breadth—and partial sharing may dominate both full duplication and full joint ventures.

Moreover, the framework intersects with the literature on weak patents and enforcement risk (Shapiro, 2003; Farrell and Shapiro, 2008). Recent work by Hovenkamp et al. (2025) shows how patentees may strategically delay engagement—“lurking”—to extract greater rents ex post, even when developers are forward-looking. Such delayed enforcement can reduce R&D effort despite rational anticipation. Similarly, Lemus and Temnyalov (2017) analyze how patent privateering reshapes innovation incentives through litigation threats. These models complement ours by showing how contract incompleteness and endogenous enforcement distort innovation and diffusion outcomes. While they focus on enforcement ex post, our model centers on licensing breadth ex ante, but both reveal how strategic delay or exclusion misaligns private and social diffusion.

The analysis also speaks to growing empirical evidence on diffusion, licensing, and standardization. Broad licensing—such as the non-exclusive licensing of the Cohen-Boyer patents in biotechnology—can enhance follow-on innovation (Feldman, 2007). Conversely, uncertainty in patent validity or delayed assertion can suppress R&D, even in rational settings (Bessen and Maskin, 2009; Cohen et al., 2019). By modeling both breadth and correlation, our framework provides a theoretical lens to interpret when narrow diffusion inhibits innovation and when broader access may improve welfare without eroding incentives.

In summary, this paper develops a theory of licensing breadth and innovation incentives in oligopoly, identifies a novel neutrality-break channel linking correlation and market structure, and provides a welfare mapping that clarifies when private diffusion is excessive or insufficient. The results offer guidance for competition and patent policy: they suggest when exclusivity sustains incentives and when openness enhances efficiency.

Roadmap. Section 2 presents the model. Sections 3 and 4 develop the core equilibrium analysis, focusing on innovation incentives and optimal licensing breadth. Section 5 turns to welfare implications. Section 6 examines the role of correlation in innovation outcomes, and Section 7 considers extensions. The conclusion discusses policy implications.

2 The Model

There are n potential producers indexed by $i \in \{1, \dots, n\}$ and one independent research lab. Baseline marginal cost is $c_0 > 0$. The lab holds a process innovation that delivers a lower marginal cost $c_1 < c_0$.

Demand is inverse $P(Q; z)$ with $Q = \sum_i q_i$ and $P_Q < 0$, where z is a demand shifter. Product-market conduct is summarized by a parameter $\lambda \in [0, 1]$ via the reduced-form first-order condition

$$P(Q; z) + \lambda q_i P_Q(Q; z) = MC_i(q_i) \in \{c_0, c_1\},$$

with $\lambda = 1$ (Cournot), $\lambda = 0$ (Bertrand/price-taking), and intermediate λ for conjectures/partial collusion. We take constant marginal costs as the baseline (process innovation only).

Each firm can obtain the low cost c_1 in one of two mutually exclusive ways:

- (a) License from the lab (deterministic). Accepting a license deterministically grants c_1 (effective cost c_1 under fixed fee F , or $c_1 + r$ under royalty r). A licensed firm exerts no further R&D (Lemma 2).
- (b) Self-innovate (risky). A non-licensee may invest $x_i \geq 0$ at cost $K(x_i)$ with $K' > 0$, $K'' > 0$, achieving success with probability $p(x_i)$ where $p' > 0$, $p'' < 0$. Success yields c_1 ; failure leaves the firm at c_0 .

The lab can offer fixed fees, royalties, or two-part tariffs $\mathcal{T} \in \{F, r, (F, r)\}$ (possibly menus), and chooses a *licensing breadth* $m \in \{0, \dots, n-1\}$: in total, $m+1$ firms receive the technology deterministically (exclusive licensing is $m = 0$).

Timing and Information.

1. **Licensing stage.** The lab announces (m, \mathcal{T}) and prices. Firms decide whether to take a license. Licensees obtain c_1 deterministically and exit the R&D decision.
2. **Innovation stage.** Each *non-licensee* i chooses $x_i \geq 0$ at cost $K(x_i)$; success occurs with probability $p(x_i)$. (Correlation across non-licensees is introduced in Section 6.)
3. **Realization & product market.** The set of low-cost firms is realized (all licensees plus successful non-licensees). Firms then compete given realized costs under conduct parameter λ .
4. **Payoffs.** Fees/royalties and R&D costs are paid; profits and surplus are realized.

Given these primitives and timing, equilibrium outcomes depend on the lab's licensing breadth m and non-licensees' effort choices. Section 3 defines the benchmark payoffs from licensing versus self-innovation and develops tools used throughout the analysis.

3 Preliminaries

3.1 Outside Options and Benchmark Payoffs

Fix n and licensing breadth m (so $m+1$ firms hold c_1 deterministically). Let $x^*(n, m)$ denote the symmetric equilibrium effort among non-licensees and $p^*(n, m) = p(x^*(n, m))$.

Licensee's expected profit. If firm i accepts a license, it obtains c_1 deterministically and exerts no R&D. Anticipating that m other firms are licensed and that the remaining $n - m - 1$ firms succeed independently with probability $p^*(n, m)$, the number of additional successful non-licensees is $M \sim \text{Bin}(n - m - 1, p^*)$. The licensee's expected gross product-market profit (before fees/royalties) is

$$\pi_i^L(m, \mathcal{T}) = \mathbb{E}_M[\pi_i^*(c_1; n, m+1+M, \lambda)], \quad (1)$$

evaluated at effective cost c_1 (fixed fee) or $c_1 + r$ (royalty).

Outside option (decline license). If firm i declines the license, it can either attempt its own innovation or remain at the baseline cost c_0 . When self-innovating, the firm chooses effort $x_i \geq 0$ to maximize its expected payoff, taking as given that m competitors are licensed and that the other $n - m - 2$ non-licensees exert the same equilibrium effort $x^*(n, m)$. Let M denote the number of *other* successful innovators among these $n - m - 2$ firms (excluding

i), with $M \sim \text{Bin}(n - m - 2, p^*(n, m))$. The outside option is therefore

$$\pi_i^O(m) = \max_{x_i \geq 0} \left\{ \underbrace{p(x_i) \mathbb{E}_M \left[\pi_i^*(c_1; n, m + 1 + M, \lambda) - \pi_i^*(c_0; n, m + 1 + M, \lambda) \right]}_{\text{expected incremental gain if innovation succeeds}} \right. \\ \left. \underbrace{- K(x_i)}_{\text{innovation cost}} \right\} \vee \underbrace{\mathbb{E}_M \left[\pi_i^*(c_0; n, m + 1 + M, \lambda) \right]}_{\text{fallback payoff with no R\&D attempt}}. \quad (2)$$

Participation and extraction. A license is individually rational iff

$$F(m) \leq \pi_i^L(m, \mathcal{T}) - \pi_i^O(m). \quad (3)$$

Under symmetric pricing with full extraction, the lab sets the fee (and royalty, if any) to bind (3). With breadth m , expected lab profit is

$$\Pi_L(m) = (m + 1) \left[\pi_i^L(m, \mathcal{T}) - \pi_i^O(m) \right], \quad (4)$$

and the lab chooses $m^* \in \arg \max_m \Pi_L(m)$ anticipating $x^*(n, m)$.

The next subsections develop general tools—an envelope identity and the mutual exclusivity of licensing and R&D—that will be used repeatedly to evaluate π_i^L , π_i^O , and the lab’s profit function $\Pi_L(m)$ in later sections.

Innovation payoffs.— We begin with a general expression for the value of process innovation. Fix the number of firms n and the conduct parameter λ , and let all rivals have marginal cost c_0 . When firm i reduces its marginal cost from c_0 to $c_1 < c_0$, the total change in equilibrium profit is given by the following envelope identity.

Lemma 1 (Envelope identity for process innovation). *The exact profit gain from a cost*

reduction from c_0 to c_1 is

$$PI(n; c_0 \rightarrow c_1) = \int_{c_1}^{c_0} q_i^*(n; c, c_0, \lambda) dc, \quad (5)$$

where $q_i^*(\cdot)$ denotes firm i 's equilibrium output when its own cost is c and all rivals' costs remain c_0 . This expression holds for any downward-sloping demand $P(Q)$ and any conduct parameter λ .

Intuitively, innovation payoff equals the area under the firm's equilibrium output schedule as its cost declines—an envelope property that does not depend on the specific form of competition. A small local cost improvement therefore yields a payoff proportional to equilibrium output:

Corollary 1 (Local approximation). *For a small cost reduction $\Delta c = c_0 - c_1$,*

$$PI(n) \approx q_i^*(n; c_0, c_0, \lambda) \Delta c = \frac{Q^*(n; \lambda)}{n} \Delta c.$$

Hence, local innovation incentives are proportional to the firm's equilibrium market share.

This observation implies that any factor expanding a firm's equilibrium output—such as softer competition or broader licensing—will locally scale up its incentive to innovate.

Licensing versus own R&D.— We next clarify how licensing interacts with firms' own research efforts. Suppose a firm can either license an existing technology or attempt to innovate on its own. When licensing deterministically delivers the low cost $c_1 < c_0$, further self-innovation is dominated, as formalized below.

Lemma 2 (Mutual exclusivity of licensing and R&D). *Suppose (i) licensing deterministically grants the low marginal cost $c_1 < c_0$; (ii) each firm's innovation effort x_i affects only its own success probability $p(x_i)$; (iii) no superior technology is available in the period; and (iv) the*

cost $K(x_i)$ is strictly increasing for $x_i > 0$. Then, conditional on accepting a license, firm i optimally sets $x_i = 0$.

Proof. Given a license, $c_i = c_1$ regardless of x_i . Under (ii)–(iii), effort no longer affects payoffs, while $K(x_i)$ is strictly increasing by (iv). The unique optimum is therefore $x_i = 0$. □

This logic extends immediately to more general contracts: as long as the effective marginal cost is fixed by the licensing arrangement, additional R&D effort yields no benefit.

Corollary 2 (Applicability under royalties or two-part tariffs). *Lemma 2 holds under any contract form (F, r) that fixes the licensee’s effective marginal cost at c_1 (or $c_1 + r$ under a royalty).*

Taken together, these results imply a simple partition of firm behavior: firms either license and forgo own R&D, or they reject the license and invest in innovation. Once a superior technology is secured deterministically, any additional research is strictly wasteful. In equilibrium, therefore, only non-licensees exert positive R&D effort, while licensees rely entirely on the licensed technology.

4 Equilibrium Effort and Licensing Breadth

We now characterize firms’ equilibrium R&D effort, the lab’s choice of licensing breadth, and the resulting equilibrium configuration, building on the envelope result from Lemma 1.

4.1 Effort Choice Given (n, m)

Fix the total number of firms n and let licensing breadth m denote the number of licensees, so that $m + 1$ firms hold the advanced technology with certainty. Each of the remaining $n - m - 1$ non-licensees chooses R&D effort $x_i \geq 0$ to maximize expected profit, taking as

given that (i) m competitors are licensed deterministically, and (ii) the other non-licensees choose the same equilibrium effort $x^*(n, m)$.

Let M denote the number of successful innovators among the remaining $n - m - 2$ non-licensees. Under independence, $M \sim \text{Bin}(n - m - 2, p^*(n, m))$ with success probability $p^*(n, m) = p(x^*(n, m))$. Given this, each non-licensee solves

$$\max_{x_i \geq 0} p(x_i) \mathbb{E}_M \left[\pi_i^*(c_1; n, m + 1 + M, \lambda) - \pi_i^*(c_0; n, m + 1 + M, \lambda) \right] - K(x_i). \quad (6)$$

The first-order condition for an interior optimum is

$$p_x(x^*(n, m)) \mathbb{E}_M \left[\pi_i^*(c_1; n, m + 1 + M, \lambda) - \pi_i^*(c_0; n, m + 1 + M, \lambda) \right] = K'(x^*(n, m)). \quad (7)$$

Under $p_{xx} < 0$ and $K'' > 0$, a unique symmetric equilibrium effort $x^*(n, m)$ exists. Equation (7) highlights that the marginal benefit of effort equals the expected incremental profit from obtaining the innovation, weighted by the marginal success probability.

Intuitively, a broader license reduces the residual incentive to innovate: as more firms obtain c_1 deterministically, the probability that a non-licensee's R&D success matters declines, depressing equilibrium effort.

4.2 The Lab's Optimal Licensing Breadth

Given $x^*(n, m)$, the lab's expected profit from licensing to $m+1$ firms is

$$\Pi_L(m) = (m + 1) \left[\pi_i^L(m, \mathcal{T}) - \pi_i^O(m) \right], \quad (8)$$

where $\pi_i^L(m, \mathcal{T})$ and $\pi_i^O(m)$ denote a licensee's and outsider's equilibrium payoffs, respectively, as defined in Section 2.5. The optimal licensing breadth solves

$$m^* \in \arg \max_{m \in \{0, \dots, n-1\}} \Pi_L(m),$$

with first-order condition

$$\frac{\partial \Pi_L(m)}{\partial m} = (m+1) \frac{\partial}{\partial m} [\pi_i^L(m, \mathcal{T}) - \pi_i^O(m)] + [\pi_i^L(m, \mathcal{T}) - \pi_i^O(m)] = 0.$$

Equilibrium summary.— The overall equilibrium is characterized by the tuple

$$\{x^*(n, m), p^*(n, m), m^*, F^*(m^*)\},$$

where (x^*, p^*) satisfy (7), and (m^*, F^*) jointly maximize the lab's objective (8) subject to participation constraints (3). Licensed firms exert no R&D effort, by Lemma 2, while non-licensees invest according to (7). The resulting allocation determines both the intensity and breadth of innovation diffusion in equilibrium.

4.3 Comparative Statics

We establish comparative statics that do not rely on linear demand or Cournot per se, but only on standard shape and monotonicity conditions. The linear–Cournot benchmark with closed forms is relegated to Appendix B.

Environment and assumptions. Let $P(Q)$ be twice continuously differentiable with $P'(Q) < 0$ and $P''(Q) \leq 0$. Marginal costs are constant at $c_1 < c_0$. Industry conduct is summarized by a parameter λ ; for any (n, k, λ) , let $q_i^*(c; n, k, \lambda)$ and $\pi_i^*(c; n, k, \lambda)$ denote the symmetric equilibrium quantity and profit of a firm with marginal cost $c \in \{c_0, c_1\}$ when k

firms in the market have cost c_1 and $n - k$ have cost c_0 . We impose the following regularities, which hold in standard symmetric quantity games and in many conduct specifications:

- (A1) *Uniqueness and continuity.* A unique symmetric equilibrium exists for all (n, k, λ) and depends continuously on primitives.
- (A2) *Monotone rivalry.* $q_i^*(c; n, k, \lambda)$ is (weakly) decreasing in the number of efficient rivals k and in industry size n .
- (A3) *Profit order and single crossing.* For any (n, k, λ) , $\pi_i^*(c_1; n, k, \lambda) \geq \pi_i^*(c_0; n, k, \lambda)$, and the incremental profit $\Delta\pi(n, k) := \pi_i^*(c_1; n, k, \lambda) - \pi_i^*(c_0; n, k, \lambda)$ is (weakly) decreasing in k .¹
- (A4) *Effort technology.* $p(x)$ is increasing and concave ($p_x > 0$, $p_{xx} < 0$); $K(x)$ is convex and strictly increasing ($K'(x) > 0$, $K''(x) \geq 0$).
- (A5) *Independence across non-licensees.* Conditional on $x^*(n, m)$, non-licensee successes are i.i.d. across firms.

Given licensing breadth m (so $m+1$ firms hold cost c_1 deterministically), a representative non-licensee chooses $x_i \geq 0$ to maximize (6). Let $x^*(n, m)$ denote the resulting symmetric equilibrium among non-licensees, and let $p^*(n, m) = p(x^*(n, m))$.

The next results characterize how the diffusion of an innovation through licensing and the scale of industry competition shape firms' equilibrium innovation effort.

Proposition 1 (Broader licensing reduces R&D effort). *Under (A1)–(A5), the symmetric equilibrium effort $x^*(n, m)$ of non-licensees is (weakly) decreasing in the licensing breadth m .*

Broader diffusion reduces the marginal value of being a successful innovator. Let $\mathcal{V}(n, m, p^*) = \mathbb{E}_M[\Delta\pi(n, m + M)]$ denote the expected incremental prize, where M is the number of other

¹In Cournot with $P'' \leq 0$, (A3) follows from the envelope property and the fact that an additional efficient rival reduces the price–cost wedge faced by each firm.

successful non-licensees. Because $\Delta\pi$ falls with $m + M$ and the distribution of M first-order stochastically increases in m , \mathcal{V} declines with m . Since the first-order condition $p_x(x^*)\mathcal{V} = K'(x^*)$ equates the expected marginal gain and cost, a fall in \mathcal{V} shifts effort downward. Hence, deterministic diffusion—by reducing the residual market value of discovery—depresses outsiders’ R&D incentives.

Proposition 2 (Non-monotone effect of competition scale). *Suppose in addition that $\mathcal{V}(n, m, p)$ is continuous in n and satisfies: (i) $\partial\mathcal{V}/\partial n > 0$ for $n \in [2, \underline{n}]$, and (ii) $\partial\mathcal{V}/\partial n < 0$ for $n \in (\bar{n}, \infty)$, for some $\underline{n} \leq \bar{n}$. Then there exists $\bar{n} \in [\underline{n}, \bar{n}]$ such that $x^*(n, m)$ increases in n for $n < \bar{n}$ and decreases for $n > \bar{n}$.*

When the industry is small, additional rivals intensify the “business-stealing” value of winning the innovation race, making success more lucrative and raising effort. Once the industry becomes large, however, each firm’s success becomes less pivotal and profit shares thin out. This tension produces an inverted-U relation between industry size and R&D effort—akin to the Schumpeterian pattern documented in empirical studies of innovation and competition.

The licensor’s problem. Having characterized outsiders’ behavior, we turn to the licensor’s decision over how broadly to disseminate the technology. The laboratory faces a trade-off between diffusion rents and cannibalization among licensees: each additional licensee expands total royalties but erodes each license’s margin.

Proposition 3 (Optimal licensing breadth and its determinants). *Let per-license profit $f(m) = \pi_i^L(m) - \pi_i^O(m)$ be strictly decreasing and weakly concave in m . Then:*

- (i) (Single-peaked profit) *The lab’s total profit $\Pi_L(m) = (m + 1)f(m)$ is single-peaked; the optimal breadth $m^* \in \arg \max_m \Pi_L(m)$ is unique unless $f'(m)$ never changes sign.*
- (ii) (Comparative statics) *If $f(m)$ decreases pointwise when (a) the number of competitors n rises or conduct softens, or (b) the cost gap $\Delta = c_0 - c_1$ widens, then m^* (weakly)*

decreases. Conversely, if demand becomes locally less elastic—raising margins—then $f(m)$ increases and m^ (weakly) rises.*

Because $f'(m) < 0$ and $f''(m) \leq 0$, total profit $g(m) = (m + 1)f(m)$ has a derivative $g'(m) = f(m) + (m + 1)f'(m)$ that falls monotonically with m , ensuring a unique interior optimum. Higher rivalry or a wider efficiency gap makes each license's incremental value shrink faster, shifting $g'(m)$ down and moving the optimum toward exclusivity. Less elastic demand cushions this cannibalization, supporting broader diffusion. Overall, the optimal licensing breadth balances revenue expansion against intensified product-market competition.

Discussion and relation to prior work. These comparative statics unify two classic insights. First, the negative relation between diffusion breadth and effort in Proposition 1 echoes the strategic-deterrence logic of Gallini (1984): broad licensing or disclosure softens outsiders' R&D incentives by shrinking the incremental rent to discovery. Second, the laboratory's incentive to commit to limited diffusion in Proposition 3 parallels Farrell and Gallini (1988), where incumbents voluntarily invite competition to alleviate dynamic inconsistency yet stop short of complete openness. In our framework, both effects arise endogenously from the same primitives—how market rivalry and technological substitution shape the expected prize $\mathcal{V}(n, m, p)$. By embedding these forces in a unified conduct-based model, we connect the strategic-licensing tradition (e.g., Kamien and Tauman, 1986; D'Aspremont and Jacquemin, 1988b) with the modern view of innovation races (Cardon and Sasaki, 1998; Bessen and Maskin, 2009), highlighting how licensing breadth can serve as a commitment device that governs both diffusion and innovation incentives.

5 Welfare Analysis

We conclude by analyzing how the diffusion of innovation through licensing and self-innovation affects total welfare. The analysis highlights a fundamental trade-off between *static efficiency*, which favors broad diffusion, and *dynamic incentives*, which rely on the rents preserved by limited licensing.

5.1 Welfare Components

Let $W(m)$ denote expected welfare when the lab licenses to $m+1$ firms and the remaining $n-m-1$ firms invest with equilibrium effort $x^*(n, m)$. For any realization in which k firms obtain the low cost c_1 , aggregate output and profits follow the equilibrium $\{Q^*(n, k, \lambda), \pi_i^*(c_j; n, k, \lambda)\}$. Expected welfare is

$$W(m) = \mathbb{E}_M \left[CS(Q^*(n, m+1+M, \lambda)) + \sum_i \pi_i^*(n, m+1+M, \lambda) \right] - \sum_{i \in N \setminus L} K(x^*(n, m)), \quad (9)$$

where $CS(Q)$ is consumer surplus at output Q , and M is the number of successful innovators among non-licensees. The first term captures the static gains from lower costs and greater output, while the last subtracts total R&D costs. Together, they summarize the economy's central trade-off between diffusion and incentives.

5.2 Welfare and Efficiency

Licensing breadth m shapes welfare through two opposing forces. A larger m enhances productive efficiency by diffusing the superior technology to more firms, but also weakens incentives for self-innovation as the expected return to R&D declines. The lab's private objective captures only the revenue side of this trade-off: it expands licensing until the marginal license just offsets the cannibalization of existing profits. Social welfare, by contrast, also values the consumer-surplus gains from broader diffusion. Let m^{SW} maximize $W(m)$,

and m^* the lab’s private optimum.

Proposition 4 (Under-diffusion of licensing breadth). *Suppose Assumptions (A1)–(A5) hold and that both $\Pi_L(m)$ and $W(m)$ are single-peaked in m . Then $m^{SW} \geq m^*$, with strict inequality whenever the marginal consumer-surplus gain is positive at the private optimum, $CS'(m^*) > 0$.*

Corollary 3 (Magnitude of the welfare gap). *In the Linear-Cournot Benchmark (Appendix B), $CS'(m) > 0$ for all $m < m^{SW}$, so $m^{SW} > m^*$. The gap $m^{SW} - m^*$ widens when (i) the innovation size Δ is moderate, so diffusion yields large consumer gains while private cannibalization remains salient, or (ii) competition is intense (large n), which amplifies diffusion benefits relative to private returns.*

The welfare function combines three forces: (i) static efficiency from broader adoption ($\partial CS/\partial m > 0$); (ii) business-stealing, which lowers producer surplus ($\partial \pi_i^*/\partial m < 0$); and (iii) dynamic incentives, as $x^*(n, m)$ declines with m by Proposition 1. For small m , static gains dominate and $W(m)$ rises with diffusion; for large m , dynamic losses prevail, yielding a single-peaked welfare curve with interior maximizer m^{SW} . Because the lab internalizes only licensing revenues, not consumer surplus, it typically stops expanding diffusion too early—hence under-diffusion relative to the social optimum.

Comparative statics and implications. Competition and innovation interact in non-monotonic ways. As the number of firms n rises, prices fall and consumer surplus grows, but innovation rents shrink, producing an “inverted-U” relationship between competition and welfare. Larger cost reductions Δ or less elastic demand both magnify the diffusion-incentive trade-off: when cannibalization is strong or R&D costs are highly convex, the planner prefers narrower diffusion; when spillovers are strong or innovation is inexpensive, broader diffusion is socially optimal. All of these mechanisms are verified in the Linear-Cournot Benchmark, where closed forms confirm that both $W(m)$ and $\Pi_L(m)$ are single-peaked and that

$m^{SW} > m^*$ whenever consumer-surplus gains dominate private cannibalization.

6 Correlated Innovation and Incentive Neutrality

We next examine how correlation across firms' innovation outcomes affects equilibrium effort and welfare. The analysis reveals a sharp neutrality-break pattern: under linear demand and Cournot competition, correlation has no effect on R&D incentives; but once demand curvature or nonlinear conduct is introduced, correlation strictly depresses effort by compressing the variance of innovation payoffs.

6.1 Model and Benchmark Neutrality

Let $(S_1, \dots, S_{n'})$ denote the binary innovation outcomes of the $n' = n - m - 1$ non-licensed firms, where $S_i = 1$ if firm i 's R&D succeeds and $S_i = 0$ otherwise. Each firm exerts effort $x_i \geq 0$ that determines its marginal success probability $p_i = p(x_i)$, and these outcomes are jointly distributed according to a one-parameter copula C_θ with dependence parameter $\theta \in [0, 1]$. A higher θ represents stronger positive dependence—e.g., greater correlation in a Gaussian copula or a higher probability of joint success under a common shock. Given licensing breadth m , let $M = \sum_{j \neq i} S_j$ denote the number of successful innovators other than firm i . The expected incremental prize for firm i is then

$$\mathbb{E}_M^\theta[\Delta\pi(n, m; M)] = \sum_M \Pr_\theta(M) \Delta\pi(n, m; M),$$

and the symmetric equilibrium effort $x^*(n, m; \theta)$ satisfies

$$p_x(x^*(n, m; \theta)) \mathbb{E}_M^\theta[\Delta\pi(n, m; M)] = K'(x^*(n, m; \theta)).$$

In the linear-Cournot benchmark with constant marginal costs, homogeneous firms, and

uniform cost reduction Δ , the incremental profit $\Delta\pi(n, m; M)$ is *affine* in M . Hence only the mean $\mathbb{E}[M]$ matters for expected profits, and $\mathbb{E}[M]$ is invariant to the dependence parameter θ . Equilibrium effort therefore satisfies $\partial x^*/\partial\theta = 0$: correlation among innovation outcomes is completely neutral. This neutrality holds irrespective of the copula specification as long as firms are risk-neutral and market payoffs are linear in aggregate success. It provides a convenient benchmark—a knife-edge case where the market structure and technology eliminate all covariance effects.

6.2 Correlation Effects under Curvature and Conduct

Once demand departs from linearity or market conduct deviates from Cournot, this neutrality collapses. Under concave-down demand, the profit difference between a successful and an unsuccessful innovator falls at an increasing rate as more rivals succeed. Correlation compresses the distribution of M , reducing the effective variance of innovation payoffs and therefore lowering the marginal benefit of effort. The following result formalizes this comparative static.

Proposition 5 (Correlation Effects on Incentives and Welfare). *Let the dependence among firms' innovation outcomes be governed by a copula parameter $\theta \in [0, 1]$, such that $\text{Var}(M)$ increases in θ . Then:*

- (i) (*Effort Neutrality*) *In the linear-Cournot benchmark with homogeneous firms and constant marginal cost gaps, equilibrium effort $x^*(n, m; \theta)$ is independent of θ :*

$$\frac{\partial x^*(n, m; \theta)}{\partial \theta} = 0.$$

- (ii) (*Welfare Decrease*) *Under the same linear-Cournot assumptions, expected total welfare is strictly decreasing in θ :*

$$\frac{\partial \mathbb{E}[W(n, m; \theta)]}{\partial \theta} < 0.$$

(iii) (*Neutrality Break*) If the profit gap $\Delta\pi(n, m; M)$ is nonlinear in M —for instance, due to nonlinear demand, strategic conduct, royalties, or risk aversion—then higher θ reduces equilibrium effort:

$$\frac{\partial x^*(n, m; \theta)}{\partial \theta} < 0.$$

Proposition 5 highlights a stark contrast between incentives and welfare under correlation. In the linear–Cournot benchmark, positive dependence among innovation outcomes does not affect firms’ equilibrium R&D effort—since the expected prize depends only on the mean number of successful rivals, which remains fixed regardless of the correlation structure. However, correlation still worsens welfare by increasing the variance of diffusion, which lowers expected consumer surplus and profits due to concavity in the number of low-cost firms.

This neutrality is fragile. Once the environment departs from linearity—due to concave demand, non–Cournot conduct, per–unit royalties, or heterogeneous innovation costs—the incremental profit $\Delta\pi(n, m; M)$ becomes nonlinear in M , and correlation compresses the distribution of innovation payoffs. This lowers the marginal return to effort, and firms optimally reduce R&D investment.

Correlated innovation arises naturally when firms share research infrastructure, face common technological risks, or coordinate through consortia. While such interdependence may facilitate diffusion or reduce redundant investment, it also undermines incentives and reduces expected welfare in all but knife–edge linear environments. These results suggest that regulators and policymakers must account for covariance effects when designing R&D subsidies, licensing mandates, or competition policy—especially in sectors where innovation paths are tightly linked.

7 Extensions

We now explore two extensions that enrich the baseline model and link it to broader themes in industrial organization and contract theory. The first introduces heterogeneity across firms and shows how the licensing problem can be interpreted as a screening mechanism. The second endogenizes market size, showing how diffusion policy interacts with entry and competition. Formal derivations and proofs are provided in Online Appendices E and F.

7.1 Heterogeneous Firms and Nonlinear Licensing Contracts

Suppose firms differ in their innovation efficiency parameter k_i , which affects either the cost or success probability of R&D effort. The laboratory can offer a menu of licensing contracts $\{(F_i, r_i)\}$, where F_i is an up-front fee and r_i a royalty rate, allowing firms to self-select between licensing and independent innovation.

Mechanism design interpretation. Let $\pi^L(k_i)$ and $\pi^O(k_i)$ denote the expected payoffs from licensing and from self-innovation, respectively, before fixed transfers. Assume a single-crossing condition in (F_i, k_i) ,

$$\frac{\partial^2(\pi^L - \pi^O)}{\partial F \partial k_i} > 0,$$

which ensures that the relative attractiveness of self-innovation increases with efficiency.

The equilibrium allocation (L, O) then admits a cutoff \bar{k} such that

$$\text{license if } k_i > \bar{k}, \quad \text{self-innovate if } k_i < \bar{k}.$$

The laboratory's problem becomes a standard screening design: choose (F_i, r_i) to maximize expected licensing revenue subject to incentive-compatibility and participation constraints for all k_i . Efficient firms (low k_i) self-innovate to avoid information rents and royalty distor-

tion, while inefficient firms (high k_i) license the technology.

Implications. Heterogeneity transforms the diffusion problem into an endogenous partition of the industry. The equilibrium licensing breadth is no longer a scalar choice m but a function of the type distribution $G(k)$. Broader heterogeneity (a thicker upper tail of inefficient firms) expands the set of licensees and shifts the aggregate diffusion margin outward. By contrast, when firms are uniformly efficient, the screening mechanism collapses to the homogeneous benchmark in which licensing and R&D are mutually exclusive discrete choices. Online Appendix D provides a closed-form solution for the cutoff \bar{k} in the linear-Cournot environment and shows that the optimal contract involves a two-part tariff: a fixed fee plus a modest royalty that trades off rent extraction and incentive preservation.

7.2 Endogenous Entry and Market Size

We now endogenize industry size by allowing potential entrants to decide whether to enter before the licensing stage. Each entrant pays a fixed cost f and then chooses whether to license or self-innovate. Let n^* satisfy the free-entry condition

$$\pi^{\text{ex ante}}(n^*) = f,$$

where $\pi^{\text{ex ante}}$ is expected profit from entry after optimizing over the licensing decision.

Entry-diffusion interaction. Licensing breadth influences entry incentives in two opposing ways. A broader license (m larger) increases the expected number of low-cost firms, which expands total output and reduces price—raising consumer surplus but diluting per-firm profits. For small or incremental innovations (low ψ), this dilution effect is mild and licensing encourages entry, yielding a larger equilibrium market. For drastic innovations (high ψ), the dilution effect dominates and exclusive licensing ($m^* = 0$) may be necessary to maintain entry incentives. The laboratory thus faces a new form of trade-off: diffusion not only affects

innovation incentives but also shapes market structure through entry.

Comparative statics. When fixed costs f are low or demand is elastic, equilibrium entry n^* rises with licensing breadth, and the social planner prefers broader diffusion to encourage participation and competition. When f is high or the innovation is drastic, n^* declines with m , and exclusive licensing becomes privately and socially optimal. In this sense, licensing breadth acts as an endogenous barrier to entry: restrictive licensing protects incumbents' innovation rents, whereas liberal licensing supports a thicker, more competitive industry. Online Appendix E formalizes these results for the linear–Cournot case, deriving threshold conditions in (ψ, f, n) for which the entry effect switches sign.

8 Conclusion

This paper developed a unified framework for how a research lab should choose the breadth of technology licensing when downstream firms can also invest in their own innovation. Licensing breadth and innovation incentives are shown to be mutually dependent: broader licensing accelerates diffusion and output expansion but erodes the expected prize for self-innovation, weakening private R&D effort.

The analysis yields three central insights. First, the lab's optimal licensing breadth is interior—balancing rent extraction against incentive preservation—and typically falls short of the socially efficient level. Second, introducing correlation across innovation outcomes reveals a sharp neutrality–break result: under nonlinear demand or non–Cournot conduct, correlation compresses payoff variance and reduces aggregate effort. Third, the lab's licensing strategy determines the distribution of marginal costs in the industry, shaping competition and welfare in equilibrium.

These results have direct policy implications. In competitive or high–spillover environments, innovation is likely to be under–diffused relative to the social optimum. Policies that broaden access to frontier technologies—such as cross–licensing, open standards, or diffusion

subsidies—can raise welfare without necessarily reducing private returns. The framework also provides a tractable benchmark for studying asymmetric firms and entry dynamics in future work.

Empirically, our results speak to ongoing debates on patent pools, FRAND commitments, and licensing practices in concentrated innovation markets. By integrating diffusion, incentives, and market structure in a single framework, the paper aims to offer a foundation for both empirical testing and policy design in innovation-intensive industries.

References

Kenneth J. Arrow. Economic welfare and the allocation of resources for invention. In R. Nelson, editor, *The Rate and Direction of Inventive Activity*, pages 609–626. Princeton University Press, 1962.

James Bessen and Eric Maskin. Sequential innovation, patents, and imitation. *RAND Journal of Economics*, 40(4):611–635, 2009.

James H. Cardon and Dan Sasaki. Preemptive search and R&D clustering. *RAND Journal of Economics*, 29(2):324–338, 1998.

Lauren Cohen, Umit Gurun, and Scott Duke Kominers. Patent trolls: Evidence from targeted firms. *Management Science*, 65(12):5461–5486, 2019.

Claude D’Aspremont and Alexis Jacquemin. Cooperative and noncooperative R&D in duopoly with spillovers. *American Economic Review*, 78(5):1133–1137, 1988a.

Claude D’Aspremont and Alexis Jacquemin. Cooperative and noncooperative R&D in duopoly with spillovers. *American Economic Review*, 78(5):1133–1137, 1988b.

Joseph Farrell and Nancy T. Gallini. Second-sourcing as a commitment: Monopoly incentives to attract competition. *Quarterly Journal of Economics*, 103(4):673–694, 1988.

- Joseph Farrell and Carl Shapiro. How strong are weak patents? *American Economic Review*, 98(4):1347–1369, 2008.
- Robin Feldman. The role of science in law. *Harvard Journal of Law and Technology*, 20(2): 321–381, 2007.
- Nancy T. Gallini. Deterrence by market sharing: A strategic incentive for licensing. *American Economic Review*, 74(5):931–941, 1984.
- Richard J. Gilbert and David M. G. Newbery. Preemptive patenting and the persistence of monopoly. *American Economic Review*, 72(3):514–526, 1982.
- Erik Hovenkamp, Jorge Lemus, and John L. Turner. Product development with lurking patentees. *RAND Journal of Economics*, 56(2):163–183, Summer 2025.
- Morton I. Kamien and Yair Tauman. Fees versus royalties and the private value of a patent. *Quarterly Journal of Economics*, 101(3):471–491, 1986.
- Morton I. Kamien, Erik Muller, and Israel Zang. Research joint ventures and R&D cartels. *American Economic Review*, 82(5):1293–1306, 1992.
- Michael L. Katz. An analysis of cooperative research and development. *RAND Journal of Economics*, 17(4):527–543, 1986.
- Michael L. Katz and Carl Shapiro. On the licensing of innovations. *RAND Journal of Economics*, 16(4):504–520, 1985.
- Jorge Lemus and Emil Temnyalov. Patent privateering, litigation, and R&D incentives. *RAND Journal of Economics*, 48(4):1004–1026, 2017.
- Jennifer F. Reinganum. Uncertain innovation and the persistence of monopoly. *Economics Letters*, 11(2):243–248, 1983.

Jennifer F. Reinganum. The timing of innovation: Research, development, and diffusion. *Journal of Economic Literature*, 27(3):940–968, 1989.

Carl Shapiro. Antitrust limits to patent settlements. *RAND Journal of Economics*, 34(2): 391–411, 2003.

Appendices

A Mathematical Proofs

Proof of Lemma 1. Fix the number of firms n and the conduct parameter λ . Let all rivals $j \neq i$ have marginal cost c_0 , and let firm i 's marginal cost be $c \in [c_1, c_0]$. Given output vector $q = (q_i, q_{-i})$ and cost c , firm i 's (one-period) profit is

$$\pi_i(q_i, q_{-i}; c) = (P(Q) - c) q_i, \quad Q \equiv \sum_{k=1}^n q_k,$$

with $P(Q)$ any downward-sloping inverse demand.

For fixed (n, c_0, λ) and own cost c , let $q_i^*(n; c, c_0, \lambda)$ denote firm i 's equilibrium output when its cost is c and all rivals have cost c_0 . Let $q_{-i}^*(n; c, c_0, \lambda)$ denote the corresponding equilibrium outputs of the other firms, and define the induced equilibrium profit of firm i as

$$\Pi_i(c) \equiv \pi_i(q_i^*(n; c, c_0, \lambda), q_{-i}^*(n; c, c_0, \lambda); c).$$

By assumption on conduct, for each fixed (n, c_0, λ) firm i chooses q_i to maximize its own profit given its (perhaps distorted) perception of how Q responds to q_i . Thus $q_i^*(n; c, c_0, \lambda)$ solves the first-order condition

$$\left. \frac{\partial \pi_i}{\partial q_i}(q_i, q_{-i}; c) \right|_{q_i=q_i^*(\cdot)} = 0,$$

with the effect of λ and of rivals' behavior embedded in the equilibrium correspondence $q^*(\cdot)$. Under the usual regularity conditions (interior solution and differentiability), the envelope theorem then gives

$$\frac{d\Pi_i(c)}{dc} = \frac{\partial \pi_i}{\partial c}(q_i^*(n; c, c_0, \lambda), q_{-i}^*(n; c, c_0, \lambda); c),$$

since the indirect effect of c through q_i^* vanishes by the first-order condition.

But $\pi_i(q_i, q_{-i}; c) = (P(Q) - c)q_i$, so the partial derivative with respect to c is simply

$$\frac{\partial \pi_i}{\partial c}(q_i, q_{-i}; c) = -q_i.$$

Evaluating at the equilibrium outputs gives

$$\frac{d\Pi_i(c)}{dc} = -q_i^*(n; c, c_0, \lambda).$$

Now define the exact gain in profit from a cost reduction from c_0 to $c_1 < c_0$ as

$$PI(n; c_0 \rightarrow c_1) \equiv \Pi_i(c_1) - \Pi_i(c_0).$$

Integrating the derivative of $\Pi_i(c)$ along this cost path yields

$$\Pi_i(c_1) - \Pi_i(c_0) = \int_{c_0}^{c_1} \frac{d\Pi_i(c)}{dc} dc = - \int_{c_0}^{c_1} q_i^*(n; c, c_0, \lambda) dc = \int_{c_1}^{c_0} q_i^*(n; c, c_0, \lambda) dc.$$

Rewriting this expression using the notation in the lemma gives

$$PI(n; c_0 \rightarrow c_1) = \int_{c_1}^{c_0} q_i^*(n; c, c_0, \lambda) dc,$$

which is exactly (5). This derivation does not depend on the particular form of $P(\cdot)$ beyond downward sloping, nor on the specific conduct parameter λ , except through its effect on the equilibrium output $q_i^*(\cdot)$. \square

Proof of Proposition 1. Fix (n, m) and write the representative non-licensee's problem as

$$\max_{x \geq 0} p(x) \mathcal{V}(n, m, p(x)) - K(x),$$

where $\mathcal{V}(n, m, p) := \mathbb{E}_M[\Delta\pi(n, m + M)]$ is the expected incremental prize and M is the number of *other* successful non-licensees.

Step 1 (Comparative statics of the prize in m). Let $N(n, m)$ denote the number of non-licensees. When the focal firm chooses x , the number of *other* non-licensees equals $N(n, m) - 1$, and $M \sim \text{Binomial}(N(n, m) - 1, p)$ with $p = p(x)$ by (A5). Consider increasing m to $m + 1$. Then the new number of *other* non-licensees falls by one, so $M' \sim \text{Binomial}(N(n, m) - 2, p)$. Define

$$K := m + M, \quad K' := (m + 1) + M'.$$

Using the binomial decomposition $M = \text{Bernoulli}(p) + M'$ (in distribution) with M' independent of the Bernoulli draw, we get

$$K' = m + 1 + M' = m + M + (1 - \text{Bernoulli}(p)) \geq K$$

almost surely. Hence K' first-order stochastically dominates K . Since $\Delta\pi(n, k)$ is (weakly) decreasing in k by (A3), it follows by monotonicity of expectation that $\mathcal{V}(n, m, p)$ is (weakly) decreasing in m for every fixed p .

Step 2 (FOC and implicit function). At an interior optimum $x^*(n, m)$ the first-order condition is

$$F(x, m) := K'(x) - p_x(x) \mathcal{V}(n, m, p(x)) = 0. \quad (10)$$

By (A4) and standard regularity (concavity of the objective in x), the second-order condition holds:

$$\partial_x F(x, m) = K''(x) - \left[p_{xx}(x) \mathcal{V} + p_x(x)^2 \partial_p \mathcal{V}(n, m, p(x)) \right] > 0$$

at $x = x^*(n, m)$. (This inequality is a standard sufficient condition for a unique best response and is subsumed by (A1).)

Step 3 (Sign of $\partial x^/\partial m$).* By the implicit function theorem,

$$\frac{\partial x^*}{\partial m} = -\frac{\partial_m F(x^*, m)}{\partial_x F(x^*, m)} = -\frac{-p_x(x^*) \partial_m \mathcal{V}(n, m, p(x^*))}{\partial_x F(x^*, m)}.$$

From Step 1, $\partial_m \mathcal{V} \leq 0$; from (A4), $p_x > 0$; and from the SOC, $\partial_x F > 0$. Therefore $\partial x^*/\partial m \leq 0$. \square

Proof of Proposition 3. Let $g(m) := (m + 1)f(m)$, where $f'(m) < 0$ and $f''(m) \leq 0$ by assumption.

Part (i): single-peakedness and uniqueness. Compute

$$g'(m) = f(m) + (m + 1)f'(m), \quad g''(m) = 2f'(m) + (m + 1)f''(m).$$

Since $f'(m) < 0$ and $f''(m) \leq 0$ for all m , we have $g''(m) < 0$ for all m . Hence g is strictly concave on its domain, so it has a unique maximizer whenever an interior solution exists (and a unique best corner otherwise). This yields single-peakedness and (generically) a unique m^* .

Part (ii): comparative statics. Let θ be a parameter that shifts f downward pointwise, i.e., for each m , $\partial f(m; \theta)/\partial \theta < 0$. Examples in the text include $\theta \in \{n, \Delta\}$ (larger n or larger Δ) and conversely a downward shift in the opposite direction for locally less elastic demand (interpreted as $\partial f/\partial \theta > 0$ for a parameter θ indexing inelasticity). Then $g(m; \theta) = (m + 1)f(m; \theta)$ is strictly concave in m for each θ and satisfies $\partial g/\partial \theta < 0$ pointwise.

By strict concavity, the first-order condition $g'(m^*; \theta) = 0$ characterizes the unique maximizer. Differentiating implicitly,

$$\frac{\partial m^*}{\partial \theta} = -\frac{\partial g'(m^*; \theta)/\partial \theta}{\partial g'(m^*; \theta)/\partial m} = -\frac{\partial}{\partial \theta} \left[f(m^*; \theta) + (m^* + 1)f'(m^*; \theta) \right] \bigg/ g''(m^*; \theta).$$

Since $g''(m^*; \theta) < 0$ (strict concavity), the sign of $\partial m^*/\partial \theta$ matches the sign of $\partial g'(m^*; \theta)/\partial \theta$.

A pointwise downward shift in f implies $\partial f/\partial\theta < 0$ and, by weak concavity of f , does not offset through f' so as to invert the sign. Therefore $\partial g'/\partial\theta < 0$ and thus $\partial m^*/\partial\theta \leq 0$: the argmax moves (weakly) left when f shifts down. Conversely, a pointwise upward shift in f (e.g., locally less elastic demand that raises margins) yields $\partial m^*/\partial\theta \geq 0$.

Formally, because g is strictly concave and shifts monotonically in the single-crossing sense with θ , the monotone comparative statics of argmax for concave problems applies (a simple derivative argument above suffices), establishing the stated directions. \square

Proof of Proposition 2. Retain the notation of (10). For an interior solution, the implicit derivative of $x^*(n, m)$ with respect to n is

$$\frac{\partial x^*}{\partial n} = -\frac{\partial_n F(x^*, n)}{\partial_x F(x^*, n)} = -\frac{-p_x(x^*) \partial_n \mathcal{V}(n, m, p(x^*))}{\partial_x F(x^*, n)}.$$

By (A4), $p_x > 0$; by the SOC, $\partial_x F > 0$ as in the previous proof. Hence the sign of $\partial x^*/\partial n$ coincides with the sign of $\partial_n \mathcal{V}$ evaluated at equilibrium.

By hypothesis, $\mathcal{V}(n, m, p)$ is continuous in n , strictly increasing on $[2, \underline{n})$ and strictly decreasing on (\bar{n}, ∞) . Therefore there exists at least one $\bar{n} \in [\underline{n}, \bar{n}]$ at which $\partial_n \mathcal{V}(\bar{n}, m, p) = 0$ (intermediate value), with $\partial_n \mathcal{V} > 0$ for $n < \bar{n}$ and $\partial_n \mathcal{V} < 0$ for $n > \bar{n}$. The derivative formula then implies $\partial x^*/\partial n > 0$ for $n < \bar{n}$ and $\partial x^*/\partial n < 0$ for $n > \bar{n}$. \square

Proof of Proposition 3. Let $g(m) := (m + 1)f(m)$, where $f'(m) < 0$ and $f''(m) \leq 0$ by assumption.

Part (i): single-peakedness and uniqueness. Compute

$$g'(m) = f(m) + (m + 1)f'(m), \quad g''(m) = 2f'(m) + (m + 1)f''(m).$$

Since $f'(m) < 0$ and $f''(m) \leq 0$ for all m , we have $g''(m) < 0$ for all m . Hence g is strictly concave on its domain, so it has a unique maximizer whenever an interior solution exists (and

a unique best corner otherwise). This yields single-peakedness and (generically) a unique m^* .

Part (ii): comparative statics. Let θ be a parameter that shifts f downward pointwise, i.e., for each m , $\partial f(m; \theta)/\partial \theta < 0$. Examples in the text include $\theta \in \{n, \Delta\}$ (larger n or larger Δ) and conversely a downward shift in the opposite direction for locally less elastic demand (interpreted as $\partial f/\partial \theta > 0$ for a parameter θ indexing inelasticity). Then $g(m; \theta) = (m + 1)f(m; \theta)$ is strictly concave in m for each θ and satisfies $\partial g/\partial \theta < 0$ pointwise.

By strict concavity, the first-order condition $g'(m^*; \theta) = 0$ characterizes the unique maximizer. Differentiating implicitly,

$$\frac{\partial m^*}{\partial \theta} = -\frac{\partial g'(m^*; \theta)/\partial \theta}{\partial g'(m^*; \theta)/\partial m} = -\frac{\partial}{\partial \theta} \left[f(m^*; \theta) + (m^* + 1)f'(m^*; \theta) \right] / g''(m^*; \theta).$$

Since $g''(m^*; \theta) < 0$ (strict concavity), the sign of $\partial m^*/\partial \theta$ matches the sign of $\partial g'(m^*; \theta)/\partial \theta$. A pointwise downward shift in f implies $\partial f/\partial \theta < 0$ and, by weak concavity of f , does not offset through f' so as to invert the sign. Therefore $\partial g'/\partial \theta < 0$ and thus $\partial m^*/\partial \theta \leq 0$: the argmax moves (weakly) left when f shifts down. Conversely, a pointwise upward shift in f (e.g., locally less elastic demand that raises margins) yields $\partial m^*/\partial \theta \geq 0$.

Formally, because g is strictly concave and shifts monotonically in the single-crossing sense with θ , the monotone comparative statics of argmax for concave problems applies (a simple derivative argument above suffices), establishing the stated directions. \square

Proof of Proposition 4. Fix n and the conduct parameter λ . Under Assumptions (A1)–(A5), equilibrium welfare can be decomposed as

$$W(m) = \Pi_L(m) + CS(m) + K,$$

where $CS(m)$ is consumer surplus and K collects all components of welfare that do not

depend on the licensing breadth m .² Hence

$$W'(m) = \Pi'_L(m) + CS'(m) \quad \text{whenever these derivatives exist.} \quad (11)$$

By definition, the privately optimal breadth m^* maximizes $\Pi_L(m)$. Given single-peakedness of $\Pi_L(m)$, any interior optimum m^* satisfies

$$\Pi'_L(m^*) = 0, \quad (12)$$

and $\Pi'_L(m) > 0$ for $m < m^*$, $\Pi'_L(m) < 0$ for $m > m^*$.

Similarly, m^{SW} maximizes $W(m)$. Single-peakedness of $W(m)$ implies that $W'(m) > 0$ for $m < m^{SW}$, $W'(m) < 0$ for $m > m^{SW}$, and any interior maximizer satisfies $W'(m^{SW}) = 0$.

Now evaluate (11) at the private optimum m^* . Using (12),

$$W'(m^*) = \Pi'_L(m^*) + CS'(m^*) = 0 + CS'(m^*) = CS'(m^*).$$

If $CS'(m^*) \geq 0$, then $W'(m^*) \geq 0$, so W is (weakly) increasing at m^* . By single-peakedness of $W(m)$, this can only happen if m^* lies weakly to the *left* of the welfare maximizer, i.e.,

$$m^{SW} \geq m^*.$$

Moreover, if $CS'(m^*) > 0$, then $W'(m^*) > 0$, so m^* cannot itself be a maximizer of $W(\cdot)$. Again by single-peakedness, the unique peak of W must lie strictly to the right, implying

$$m^{SW} > m^*.$$

Finally, if m^* or m^{SW} is at a boundary of the feasible interval for m , the same logic

²Intuitively, (A1)–(A5) ensure that, conditional on equilibrium effort and market structure, the only part of producer surplus that varies with m is the lab's profit.

applies using the appropriate one–sided derivatives: a positive marginal consumer–surplus gain at m^* still implies that moving to a (weakly) higher m increases welfare, so the welfare maximizer cannot lie below m^* .

Hence $m^{SW} \geq m^*$, with strict inequality whenever $CS'(m^*) > 0$. □

Proof. Fix (n, m) and let $n' := n - m - 1$ denote the number of non–licensed firms who choose R&D effort. For each non–licensee i , let $S_i \in \{0, 1\}$ denote success, with $\Pr(S_i = 1) = p(x)$ where $p' > 0$, $p'' < 0$. Let $M = \sum_{i=1}^{n'} S_i$ be the number of successful non–licensees. The copula parameter $\theta \in [0, 1]$ governs the dependence across the $\{S_i\}$'s such that $\mathbb{E}_\theta[M] = n'p(x)$ for all θ , while $\text{Var}_\theta(M)$ increases in θ .

Let $\Delta\pi(n, m; M)$ denote the profit advantage of a successful non–licensee relative to remaining at cost c_0 when M other non–licensees succeed. The expected R&D payoff of a representative firm is

$$V(x; \theta) = p(x) \mathbb{E}_\theta[\Delta\pi(n, m; M)] - k(x),$$

where $k(x)$ is the convex cost of effort. The symmetric equilibrium $x^*(n, m; \theta)$ satisfies

$$p'(x^*) \mathbb{E}_\theta[\Delta\pi(n, m; M)] - k'(x^*) = 0. \tag{13}$$

(i) Effort Neutrality. Under the linear–Cournot benchmark (linear demand, homogeneous firms, constant cost reduction Δc), the stage–2 equilibrium is linear in costs. The incremental profit gap between a low– and high–cost firm is therefore affine in the number of low–cost firms $K = \bar{K} + M$:

$$\Delta\pi(n, m; M) = \alpha_0 + \alpha_1 M, \quad \alpha_0, \alpha_1 > 0.$$

Hence $\mathbb{E}_\theta[\Delta\pi(n, m; M)] = \alpha_0 + \alpha_1 \mathbb{E}_\theta[M] = \alpha_0 + \alpha_1 n' p(x)$, which is independent of θ because $\mathbb{E}_\theta[M]$ depends only on the marginals. From (13), x^* depends only on this expectation,

implying

$$\frac{\partial x^*(n, m; \theta)}{\partial \theta} = 0.$$

(ii) Welfare Decrease. Let $\widetilde{W}(K)$ denote total (static) welfare in the product market when $K = \bar{K} + M$ firms have low cost c_1 . In linear Cournot, $\widetilde{W}(K)$ is strictly concave in K :

$$\widetilde{W}(K) = A + BK - CK^2, \quad C > 0.$$

Because the copula affects only the dispersion of M (and hence of K), a higher θ generates a mean-preserving spread in K while keeping $\mathbb{E}_\theta[K]$ fixed. By Jensen's inequality,

$$\mathbb{E}_{\theta'}[\widetilde{W}(K)] < \mathbb{E}_\theta[\widetilde{W}(K)] \quad \text{for } \theta' > \theta,$$

so expected welfare is strictly decreasing in θ :

$$\frac{\partial \mathbb{E}[W(n, m; \theta)]}{\partial \theta} < 0.$$

(iii) Neutrality Break. Suppose now that $\Delta\pi(n, m; M)$ is *nonlinear* in M (e.g. due to nonlinear demand, strategic conduct, royalties, or risk aversion) and that it is strictly concave in M . Then a mean-preserving spread of M lowers the expected profit gap:

$$\frac{\partial}{\partial \theta} \mathbb{E}_\theta[\Delta\pi(n, m; M)] < 0.$$

Differentiating the FOC (13) and applying the Implicit Function Theorem,

$$\frac{\partial x^*(n, m; \theta)}{\partial \theta} = -\frac{\frac{\partial F}{\partial \theta}}{\frac{\partial F}{\partial x}} = -\frac{p'(x^*) \frac{\partial}{\partial \theta} \mathbb{E}_\theta[\Delta\pi]}{p''(x^*) \mathbb{E}_\theta[\Delta\pi] - k''(x^*)}.$$

Because $p' > 0$, $p'' < 0$, $k'' \geq 0$, and $\frac{\partial}{\partial \theta} \mathbb{E}_\theta[\Delta\pi] < 0$, the denominator is negative and the numerator positive, yielding

$$\frac{\partial x^*(n, m; \theta)}{\partial \theta} < 0.$$

Hence higher correlation (larger θ) depresses equilibrium effort.

Combining the three parts proves Proposition 5. □

B Equilibrium Analysis: Linear-Cournot Benchmark

This appendix provides a full derivation and illustration of the model under linear demand and Cournot competition. The purpose is twofold: (i) to verify that the general assumptions stated in Section 4.3 hold in a familiar environment, and (ii) to show in closed form how the comparative statics in Propositions 1–3 arise from explicit primitives.

B.1 Setup

Let $P(Q) = a - Q$ (normalizing the slope to one) and $C_i(q_i) = c_i q_i$, with $c_1 = c_0 - \Delta$, $\Delta > 0$. The total number of firms is n , and k denotes the number of low-cost firms (c_1). As in the main text, m denotes licensing breadth, so that $m + 1$ firms are deterministic licensees and the remaining $n - m - 1$ firms undertake stochastic innovation.

Market-stage profits. When k firms have cost c_1 and $n - k$ have c_0 , the Cournot equilibrium profits are

$$\pi_L(n, k) = \frac{N_L(n, k)^2}{(n + 1)^2}, \quad N_L(n, k) := a - 2c_1 + (n - k)\Delta,$$

for a low-cost firm and

$$\pi_H(n, k) = \frac{N_H(n, k)^2}{(n + 1)^2}, \quad N_H(n, k) := a + (n - 3)c_0 - k\Delta,$$

for a high-cost firm. These expressions follow from the standard Cournot outcome $q_i = (a - (n + 1)c_i + \sum_{j \neq i} c_j)/(n + 1)$ and $\pi_i = q_i^2$.

B.2 Incremental Prize and Expected Value

Fix m . A focal non-licensee faces M additional successful innovators among the other $n - m - 2$ non-licensees. If it succeeds, there are $k_s = m + 1 + M$ low-cost firms and it earns $\pi_L(n, k_s)$; if it fails, there are $k_f = m + M$ and it earns $\pi_H(n, k_f)$. Hence

$$\Delta\pi(n, m; M) = \pi_L(n, m + 1 + M) - \pi_H(n, m + M) = \frac{N_s(M)^2 - N_f(M)^2}{(n + 1)^2}, \quad (14)$$

where

$$N_s(M) = \alpha_s - \beta M, \quad N_f(M) = \alpha_f - \beta M,$$

with identical slope $\beta = \Delta$ and

$$\alpha_s = a - 2c_0 + (n - m + 1)\Delta, \quad \alpha_f = a + (n - 2)c_0 - m\Delta.$$

Let $M \sim \text{Bin}(n - m - 2, p^*)$ denote the number of other successful innovators, where $p^* = p(x^*(n, m))$ is the equilibrium success probability. Because N_s and N_f share the same linear slope, the quadratic terms in M cancel, yielding

$$\mathbb{E}_M[\Delta\pi(n, m; M)] = \frac{(\alpha_s - \alpha_f)(\alpha_s + \alpha_f - 2\mu\Delta)}{(n + 1)^2}, \quad (15)$$

where $\mu = \mathbb{E}[M] = (n - m - 2)p^*$. This cancellation is special to linear demand and constant marginal cost; in nonlinear settings, cross terms would remain and require numerical integration.

B.3 Effort Condition and Parametric Example

Substituting (15) into the first-order condition (7) from the main text gives

$$p_x(x^*(n, m)) \cdot \frac{(\alpha_s - \alpha_f) \left(\alpha_s + \alpha_f - 2(n - m - 2)p^*(n, m) \Delta \right)}{(n + 1)^2} = K'(x^*(n, m)). \quad (16)$$

This identity holds for any increasing concave $p(\cdot)$ and convex $K(\cdot)$.

Illustration. For $p(x) = 1 - e^{-x/k}$ and $K(x) = \frac{1}{2}\gamma x^2$, (16) becomes

$$\frac{e^{-x^*/k}}{k} \cdot \frac{(\alpha_s - \alpha_f) \left(\alpha_s + \alpha_f - 2(n - m - 2)(1 - e^{-x^*/k}) \Delta \right)}{(n + 1)^2} = \gamma x^*,$$

which has a unique interior solution since the LHS is decreasing and the RHS increasing in x^* .

B.4 Comparative Statics

The closed forms derived above allow us to verify the general results of Section 4.3 explicitly within the linear–Cournot benchmark.

(i) Effect of licensing breadth m . From the first-order condition (16), the marginal cost of effort $K'(x^*)$ increases with x^* , while the marginal benefit (the left-hand side) falls as licensing breadth m expands. The decline operates through two channels. First, the term $\alpha_s + \alpha_f = 2a + (n - 5)c_0 + (n + 1 - 2m)\Delta$ decreases linearly with m ($\partial/\partial m = -2\Delta < 0$), capturing the direct erosion of market rents when more firms hold the superior technology deterministically. Second, the expected number of other innovators $\mu = (n - m - 2)p^*$ also falls with m , further reducing the expected incremental prize from success. Since both effects lower the left-hand side of (16) at any given x^* , equilibrium effort must decline to restore equality. Hence, in line with Proposition 1, broader deterministic diffusion weakens

non-licensees' incentives to innovate.

(ii) Effect of industry size n . Differentiating (16) with respect to n reveals three countervailing forces familiar from standard Cournot analysis. A larger n dilutes rents via the scale factor $1/(n+1)^2$ (a negative effect), but also raises the potential gain from moving from high to low cost—the “business-stealing” effect through $\alpha_s + \alpha_f$ (a positive effect). At the same time, the expected number of other successes $\mu = (n - m - 2)p^*$ grows with n , which reduces the marginal impact of an individual's success (a negative “crowding” effect). For small n , the business-stealing term dominates and effort rises with entry; for large n , dilution and crowding dominate and effort falls. The combined pattern yields a threshold \bar{n} at which $x^*(n, m)$ peaks, consistent with Proposition 2.

(iii) Shape of the lab's profit function. Under linear demand, per-license profits $\pi_i^L(m) - \pi_i^O(m)$ decline in m because additional licensees intensify competition. Moreover, the rate of decline steepens with m : each new license exerts stronger cannibalization on existing ones than the last. This curvature ensures that the lab's total licensing profit,

$$\Pi_L(m) = (m + 1)[\pi_i^L(m) - \pi_i^O(m)],$$

is strictly concave in m , confirming the single-peaked structure described in Proposition 3. The interior optimum m^* balances the average per-license margin with the marginal cannibalization effect of adding one more license.

(iv) Comparative statics of the optimal breadth m^* . Combining the above relationships with (15) yields the same qualitative comparative statics as in Proposition 3. An increase in competition intensity (n) or a larger cost gap Δ steepens the cannibalization curve, reducing the optimal breadth m^* . By contrast, when demand is less elastic (a higher $a - c_0$ margin), per-license profits erode more slowly with m , shifting the optimum outward.

In sum, the linear–Cournot benchmark reproduces all general patterns of the main text: broader licensing depresses innovation effort, industry expansion has a non-monotonic effect on incentives, and the lab’s optimal licensing breadth is single-peaked and contracts with stronger competition or a larger technological advantage. The explicit formulas here make these mechanisms transparent.

C Welfare in the Linear–Cournot Benchmark

We specialize the welfare analysis to the linear–Cournot environment with $P(Q) = a - Q$ and $C_i(q_i) = c_i q_i$, where $c_1 = c_0 - \Delta$ with $\Delta > 0$. As in Appendix B, let k be the number of low–cost firms (c_1) that are active in the market stage, and let licensing breadth be m (so $m+1$ licensees hold c_1 deterministically and the remaining $n - m - 1$ non–licensees exert equilibrium effort $x^*(n, m)$ with success probability $p^*(n, m) = p(x^*(n, m))$).

C.1 Welfare by configuration and in expectation

Given k low–cost firms, the Cournot outcome implies

$$Q(n, k) = \frac{na - (n+1)[kc_1 + (n-k)c_0]}{n+1} = \frac{na}{n+1} - nc_0 + k\Delta =: B + k\Delta,$$

with $B := \frac{na}{n+1} - nc_0$. Consumer surplus is $CS(Q) = \frac{1}{2}Q^2$. Total producer surplus is

$$\Pi(n, k) = k\pi_L(n, k) + (n-k)\pi_H(n, k) = \frac{kN_L(k)^2 + (n-k)N_H(k)^2}{(n+1)^2},$$

where, from Appendix B,

$$N_L(k) = A_L - \Delta k, \quad A_L := a - 2c_1 + n\Delta = a - 2c_0 + (n+2)\Delta,$$

$$N_H(k) = A_H - \Delta k, \quad A_H := a + (n-3)c_0.$$

Hence $\Pi(n, k)$ is a quadratic in k :

$$\Pi(n, k) = \frac{\alpha_2 k^2 + \alpha_1 k + \alpha_0}{(n+1)^2}, \quad \begin{cases} \alpha_2 = 2\Delta(n-1)c_0 - (n+4)\Delta^2, \\ \alpha_1 = (A_L - A_H)(A_L + A_H) - 2A_H\Delta n, \\ \alpha_0 = n A_H^2. \end{cases} \quad (17)$$

Ex post welfare given k is

$$W(n, k) = \frac{1}{2}Q(n, k)^2 + \Pi(n, k) - (n-k)K(x^*(n, m)).$$

Ex ante (given breadth m), the number of additional successes among the non-licenseses is $M \sim \text{Bin}(t, p^*)$ with $t := n - m - 1$ and $p^* := p^*(n, m)$. Thus $k = m + 1 + M$ so that

$$\mathbb{E}[k] =: E_k = m + 1 + t p^*, \quad \text{Var}(k) =: V_k = t p^*(1 - p^*).$$

Using $Q = B + \Delta k$, we obtain closed forms for the expected components:

$$\mathbb{E}[CS] = \frac{1}{2} \mathbb{E}[(B + \Delta k)^2] = \frac{1}{2} [(B + \Delta E_k)^2 + \Delta^2 V_k], \quad (18)$$

$$\mathbb{E}[\Pi] = \frac{\alpha_2 (V_k + E_k^2) + \alpha_1 E_k + \alpha_0}{(n+1)^2}, \quad (19)$$

$$\mathbb{E}[(n-k)K(x^*)] = (n - E_k) K(x^*(n, m)) = t(1 - p^*) K(x^*(n, m)). \quad (20)$$

Therefore

$$W(m) = \frac{1}{2} [(B + \Delta E_k)^2 + \Delta^2 V_k] + \frac{\alpha_2 (V_k + E_k^2) + \alpha_1 E_k + \alpha_0}{(n+1)^2} - t(1 - p^*) K(x^*(n, m)), \quad (21)$$

with $E_k = m + 1 + t p^*$, $V_k = t p^*(1 - p^*)$, and $t = n - m - 1$.

C.2 Decomposition of dW/dm and interpretation

Differentiating (21) with respect to m yields a transparent decomposition:

$$\frac{dW}{dm} = \underbrace{\left(\partial_{E_k} W\right) \frac{dE_k}{dm} + \left(\partial_{V_k} W\right) \frac{dV_k}{dm}}_{\text{Static + producer surplus}} - \underbrace{\frac{d}{dm} \left[t(1-p^*)K(x^*) \right]}_{\text{Dynamic R\&D term}},$$

where

$$\frac{dE_k}{dm} = 1 - p^* + t \frac{dp^*}{dm}, \quad \frac{dV_k}{dm} = -p^*(1-p^*) + t(1-2p^*) \frac{dp^*}{dm},$$

and $\frac{dp^*}{dm} = p_x(x^*) \frac{dx^*}{dm} < 0$ by Proposition 1 (broader deterministic diffusion depresses effort).

The *direct* static effect of a broader license is the conversion of one non-licensee into a deterministic adopter (the $1-p^*$ term in dE_k/dm), which raises output and consumer surplus (via (18)) and shifts producer surplus (via (19)). Two *indirect* effects counteract this gain: (i) p^* falls with m , reducing E_k (fewer endogenous successes) and (ii) variance V_k shrinks (fewer draws), which diminishes the curvature-driven gains in (18)–(19). The dynamic term in (20) falls with m through the shrinkage of t but rises through the drop in p^* and the induced decline in x^* ; on net, broader diffusion weakens innovation incentives.

For small m , the direct static gain dominates and $dW/dm > 0$. As m grows, the fall in p^* and x^* strengthens the negative dynamic term, and dW/dm becomes negative. Hence $W(m)$ is single-peaked with a unique maximizer m^{SW} balancing diffusion (static gains) against incentives (dynamic losses).

The lab internalizes only licensing revenues, not CS . In this benchmark, the planner's $W(m)$ (which includes (18)) rises faster in m than the lab's $\Pi_L(m)$ when competition is strong (large n) or the cost gap Δ is large. Consequently $m^{SW} > m^*$ in those regions: private diffusion is insufficient relative to the social optimum.

Online Appendices (Not for Publication)

“Strategic Licensing and Innovation Incentives in Oligopoly”

D Heterogeneous Firms and Nonlinear Licensing Contracts

D.1 Setup with Heterogeneity

Let firms differ in their innovation efficiency parameter $k_i > 0$, which scales their R&D cost or success probability. For convenience, write the success probability as

$$p(x_i; k_i) = 1 - \exp(-x_i/k_i), \quad K(x_i; k_i) = \frac{1}{2}\gamma x_i^2/k_i,$$

so larger k_i means a more efficient innovator (higher success for a given effort). The density of k_i on support $[\underline{k}, \bar{k}]$ is $f(k)$ with CDF $F(k)$.

At the licensing stage, the lab offers a menu of contracts

$$\mathcal{T} = \{(F(k), r(k))\}_{k \in [\underline{k}, \bar{k}]},$$

where F is an upfront fee and r a per-unit royalty. Firms simultaneously decide whether to accept the contract intended for their type or reject all contracts and self-innovate.

Licensee and outside options. For type k , the gross licensee profit (before paying F) is

$$\pi^L(k; r) = \mathbb{E}_M[\pi_L(n, m + 1 + M; c_1 + r)],$$

and the outside option from self-innovation is

$$\pi^O(k) = \max_{x_i \geq 0} \left\{ p(x_i; k) \mathbb{E}_M[\pi_L(n, m + 1 + M; c_1) - \pi_H(n, m + M; c_0)] - K(x_i; k) \right\}.$$

Denote the maximized value by $\hat{\pi}^O(k)$. Under the parametric functions above, the first-order condition for self-innovation is

$$p_x(x^*; k) \Delta \bar{\pi} = K_x(x^*; k), \quad \text{where } \Delta \bar{\pi} = \mathbb{E}_M[\pi_L - \pi_H].$$

Solving yields $x^*(k) = k \ln(1 + \Delta \bar{\pi} k / \gamma)$ and

$$\hat{\pi}^O(k) = \frac{1}{2} \frac{k(\Delta \bar{\pi})^2}{\gamma} + O(k^2), \quad (\text{D.1})$$

so $\hat{\pi}^O(k)$ is strictly increasing and strictly convex in k .

D.2 Incentive and Participation Constraints

A type- k firm that accepts contract (F, r) earns

$$U(k) = \pi^L(k; r) - F.$$

The lab chooses $(F(k), r(k))$ to maximize

$$\Pi_L = \int_{\underline{k}}^{\bar{k}} (F(k) + r(k)q^L(k; r))f(k) dk, \quad (\text{D.2})$$

subject to:

$$U(k) \geq \hat{\pi}^O(k), \quad (\text{participation}) \quad (\text{D.3})$$

$$U(k) \geq \pi^L(k; r(\tilde{k})) - F(\tilde{k}), \quad \forall \tilde{k}, \quad (\text{incentive compatibility}) \quad (\text{D.4})$$

where $q^L(k; r)$ is expected licensed output.

Single-crossing property. We assume

$$\frac{\partial^2(\pi^L - \hat{\pi}^O)}{\partial k \partial r} > 0,$$

so that more efficient firms value lower royalties relatively more. This ensures that IC can be handled by standard monotonicity methods.

D.3 Characterization

Differentiating (D.4) and applying the envelope theorem yields

$$U'(k) = \frac{\partial}{\partial k} [\pi^L(k; r(k)) - \hat{\pi}^O(k)].$$

Integrating from \underline{k} and imposing participation for the lowest type, $U(\underline{k}) = 0$, gives

$$U(k) = \int_{\underline{k}}^k \frac{\partial}{\partial t} [\pi^L(t; r(t)) - \hat{\pi}^O(t)] dt.$$

Substituting into the lab's objective (D.2) and integrating by parts gives the standard mechanism-design representation:

$$\Pi_L = \int_{\underline{k}}^{\bar{k}} \left\{ [\pi^L(k; r(k)) - \hat{\pi}^O(k)] - \frac{1 - F(k)}{f(k)} \frac{\partial}{\partial k} [\pi^L(k; r(k)) - \hat{\pi}^O(k)] \right\} f(k) dk. \quad (\text{D.5})$$

The lab chooses $r(k)$ to maximize the integrand pointwise, subject to monotonicity of allocations.

Proposition D.1 (Self-selection structure). *Under the single-crossing property and differentiability, there exists a cutoff $\bar{k} \in [\underline{k}, \bar{k}]$ such that*

$$k < \bar{k} \Rightarrow \text{self-innovation}, \quad k > \bar{k} \Rightarrow \text{license contract } (F(k), r(k)).$$

Efficient innovators (large k) prefer to self-innovate, while less efficient firms license the technology.

Sketch. Monotonicity of $r(k)$ and the envelope result imply that $U'(k) > 0$ if and only if $\frac{\partial}{\partial k} [\pi^L(k; r(k)) - \hat{\pi}^O(k)] > 0$. Given (D.1), $\hat{\pi}^O(k)$ increases faster in k than $\pi^L(k; r)$, implying a single crossing point \bar{k} . □

The cutoff \bar{k} partitions the industry into two segments: innovative firms that prefer self-development and imitative firms that license. The lab behaves as a nonlinear seller, screening firms by their innovation efficiency. As heterogeneity widens (larger dispersion of k_i), the licensing share $1 - F(\bar{k})$ increases and the lab's revenue rises.

D.4 Comparative Statics

- **Heterogeneity:** $\partial\bar{k}/\partial\text{Var}(k) < 0$. Greater dispersion enlarges the set of licensees.
- **Innovation size:** $\partial\bar{k}/\partial\psi > 0$ (bigger cost gap ψ strengthens incentives for efficient firms to self-innovate).
- **Market competitiveness:** $\partial\bar{k}/\partial n < 0$ (more competition reduces $\pi^L - \pi^O$, expanding licensing).

A regulator observing heterogeneity in firms' innovation efficiencies can infer welfare effects from the location of \bar{k} : a high \bar{k} (many self-innovators) implies duplication of R&D; a low \bar{k} (many licensees) implies potential under-investment. Subsidies or differentiated licensing fees can move \bar{k} toward the welfare-maximizing composition.

E Endogenous Entry

E.1 Setup with Entry Stage

Before the licensing and innovation stages, a continuum of potential firms decides whether to enter the industry. Entry requires paying a fixed cost $f > 0$. After entry, the game proceeds exactly as described in the main model: the lab announces licensing breadth m , non-licensees exert effort x_i , and product-market competition ensues.

Let $\Pi(n, m)$ denote the expected per-firm *gross profit* (before paying f) when n firms are active and licensing breadth is m :

$$\Pi(n, m) = \mathbb{E}_M[\pi_i^*(n, m + 1 + M)] - \mathbb{1}_{\{i \notin L\}} K(x^*(n, m)),$$

where $\mathbb{E}_M[\cdot]$ averages over the distribution of additional successful innovators M among non-licensees.

E.2 Free-entry condition

Free entry determines the equilibrium number of firms n^* by

$$\Pi(n^*, m^*) = f, \tag{E.1}$$

where (m^*, x^*) are the equilibrium licensing breadth and effort characterized in Section 4. No potential entrant expects positive net profit at (n^*, m^*) , and existing firms earn zero rents net of the entry cost.

The comparative statics of n^* are derived from differentiating (E.1):

$$\frac{dn^*}{dz} = -\frac{\partial\Pi/\partial z}{\partial\Pi/\partial n}, \quad z \in \{m, \psi, f\}. \tag{E.2}$$

Since $\partial\Pi/\partial n < 0$ under standard Cournot competition, the sign of dn^*/dz is opposite to the sign of $\partial\Pi/\partial z$.

E.3 Comparative statics and interpretation

Licensing breadth m . An increase in m raises the number of deterministic low-cost firms and reduces individual profits through stronger competition: $\partial\Pi/\partial m < 0$. Hence

$$\frac{dn^*}{dm} > 0.$$

Broader licensing (greater diffusion) encourages entry by lowering rival profits and attracting new firms, consistent with pro-competitive diffusion.

Innovation size ψ . A larger cost reduction ψ increases post-innovation profits but reduces per-firm profits under competition (business-stealing effect). The net effect on entry is

ambiguous:

$$\text{if } \psi \text{ small, } \frac{dn^*}{d\psi} > 0; \quad \text{if } \psi \text{ large, } \frac{dn^*}{d\psi} < 0.$$

Small innovations support broader participation; drastic ones favor concentration.

Entry cost f . From (E.2), $dn^*/df < 0$ trivially: higher fixed costs reduce equilibrium entry.

Key insight. Licensing breadth m^* and market structure n^* are jointly determined. Combining (E.1) and the lab's objective $\max_m \Pi_L(m)$ yields an equilibrium pair (n^*, m^*) satisfying

$$\begin{cases} \Pi(n^*, m^*) = f, & \text{(entry equilibrium)} \\ \frac{\partial \Pi_L(m^*)}{\partial m} = 0, & \text{(licensing optimum).} \end{cases}$$

The two conditions jointly define a *licensing–entry equilibrium*. Comparative statics follow directly:

Proposition E.2 (Licensing–entry comparative statics). *In the linear–Cournot benchmark with fixed cost f :*

1. $\frac{\partial m^*}{\partial f} < 0$: higher entry costs (smaller n^*) reduce the lab's optimal licensing breadth.
2. $\frac{\partial m^*}{\partial \psi} < 0$: larger innovation sizes (drastic innovations) favor exclusivity.
3. $\frac{\partial n^*}{\partial m} > 0$: broader diffusion encourages entry, holding ψ and f fixed.

E.4 Welfare and Policy Implications

Welfare now depends on (n^*, m^*) jointly:

$$W^* = W(n^*, m^*) = \mathbb{E}_M \left[CS(Q^*(n^*, m^*+1+M)) + \sum_i \pi_i^*(n^*, m^*+1+M) \right] - \sum_{i \notin L} K(x^*(n^*, m^*)).$$

The socially optimal configuration (n^{SW}, m^{SW}) maximizes $W(n, m)$. Comparison with Proposition E.2 yields:

- Private equilibrium (n^*, m^*) typically involves too many firms (excess entry) and too few licenses (under-diffusion), as the lab does not internalize the entry externality on consumer surplus.
- Entry regulation (adjusting f) and diffusion policy (setting a minimum licensing breadth \underline{m}) are substitutes: raising f or enforcing larger \underline{m} both mitigate overentry.

Summary. Endogenous entry introduces a second margin of market selection. Broader licensing encourages entry by reducing incumbents' rents, while drastic innovations concentrate the market. The equilibrium pair (n^*, m^*) reflects a balance between diffusion incentives and entry competition; socially optimal policy combines diffusion and entry regulation to correct the joint distortion.