

Corporate Mission in the Technology Race: Can Responsible AI Compete?

Martin Oehmke

Marcus Opp

September 14, 2026

Abstract

AI providers are engaged in an unprecedented technology race. Can responsible AI survive in this race? In our framework, two AI providers invest in model capability and then set prices and guardrails. Guardrails curb harmful use but lower users' willingness to pay. Each firm's corporate mandate determines how it weighs harm prevention against profits. With model capabilities fixed, competition does not erode responsible conduct: guardrails depend only on a firm's mandate, and even purely profit-driven founders adopt public-benefit mandates to soften price competition. A technology race overturns this conclusion. A responsible mandate weakens a firm's position in the race, and its rival responds by expanding capability. When capability investment is sufficiently elastic, even mission-driven founders understate or abandon their mission. Market forces alone do not deliver responsible AI.

1 Introduction

Artificial intelligence is becoming a general-purpose technology. It is also a dual-use technology: the same model that identifies candidate molecules for cancer treatment can, in principle, help synthesize toxic compounds, and the same model that finds and patches security flaws in a bank’s software can write the malicious code that exploits them. AI providers can curb misuse through “guardrails,” such as usage policies, refusal rules, and deployment gates. However, guardrails cannot perfectly separate harmful from legitimate use, so they also degrade the product for legitimate users: responsible AI appears to come at the expense of competitiveness.

This trade-off plays out in a market with two defining features. First, AI providers are engaged in an intense technology race. Firms compete for users through prices and model capability, and in 2025 alone, the five largest hyperscalers spent close to half a trillion dollars on capital expenditure, largely AI infrastructure, four times their rate of spending when GPT-4 was released in 2023 (Epoch AI, 2026). Second, unlike firms in most industries, several leading AI companies were founded with charters that formally rank a safety mission above profits, through nonprofit, capped-profit, or public-benefit structures. These mandates are in flux, and the response to their revision shows that they matter: OpenAI’s plan to end nonprofit control of its business was abandoned after public criticism, a review by the Delaware and California attorneys general, and a lawsuit by co-founder Elon Musk.¹ This raises the question of whether responsible AI can survive market competition, or whether the technology race forces even mission-driven founders to set up profit-maximizing firms.

To answer this question, we develop an industrial organization model of competition between two AI providers. We first show that, when technology is given, competition does not erode responsible conduct. A firm’s optimal guardrail depends only on its own mandate and not on the intensity of competition, consistent with the irrelevance result of Dewatripont and Tirole (2024). Moreover, adopting a public-benefit charter is financially profitable because it commits the firm to price less aggressively, which induces its rival to follow suit. Therefore, even purely profit-driven founders adopt public-benefit mandates,

¹ OpenAI was founded as a nonprofit and from 2019 operated through a capped-profit subsidiary under the nonprofit’s control (OpenAI, 2019). Its 2025 recapitalization turned the subsidiary into a public benefit corporation with ordinary equity but left the nonprofit in control, with the power to appoint the board and to halt model releases (OpenAI, 2025, Delaware Department of Justice, 2025, California Department of Justice, 2025). Anthropic is a public benefit corporation whose board majority is appointed by a trust of financially disinterested trustees (Anthropic, 2023, 2026). xAI dropped its benefit-corporation status in 2024 (CNBC, 2025), and OpenAI and Anthropic filed for initial public offerings in 2026 (TechCrunch, 2026).

echoing the delegation logic of [Fershtman and Judd \(1987\)](#). Taken together, these results suggest that responsible firms can survive and that competition may even encourage public-benefit mandates.

Our main result is that a technology race, arguably the defining feature of AI markets, overturns this conclusion. When firms invest in model capability, a responsible mandate weakens a firm’s position in the race. Its rival responds by expanding capability, which works against the gain from softer price competition. If investment in model capability is sufficiently elastic, responsible mandates erode: even founders who genuinely care about harm write charters that understate their mission weight, the weight they place on the harm their firm causes, and founders whose mission weight falls below a threshold write pure profit charters. Market forces alone therefore do not deliver responsible AI.

We develop these results in a parsimonious model with two AI providers that interact in two stages. In the first stage, each firm invests in model capability. Capability raises users’ willingness to pay (by shifting the demand intercepts in the linear differentiated duopoly of [Singh and Vives \(1984\)](#)), but more capable models also enable more harm. In the second stage, firms simultaneously set prices and guardrail intensities. Users differ in how much harm their use of the technology generates. A firm’s guardrail reduces the harm its users cause but also lowers their willingness to pay. Each firm operates under a corporate mandate, the weight its charter places on the harm the firm causes, which determines both its prices and its guardrails. A founder’s mission weight is the weight the founder places on this harm, which may differ from the corporate mandate written into the charter. In the final part of the analysis, we add a prior stage in which founders choose the corporate charter. Charters may deviate from founders’ preferences in either direction: a mission-driven founder can adopt a pure profit charter, and a profit-driven founder can adopt a public-benefit charter.

As a benchmark, we first study a setting with fixed technologies and fixed mandates. We show that firms choose guardrails to minimize their harm-adjusted unit cost, which consists of the internalized harm per unit sold and the reduction in willingness to pay caused by the guardrail. This choice depends only on the firm’s own mandate; it is independent of the rival’s guardrail and of the intensity of competition. Guardrails matter nonetheless: by reducing the harm a responsible firm internalizes, they lower its effective cost and allow it to remain active where it would otherwise be priced out of the market.

We then show that, once guardrails are chosen optimally, a firm’s capability and mandate enter equilibrium margins, quantities, and payoffs only through a single statistic, the firm’s *harm-discounted technology*: model capability multiplied by a mandate-dependent

discount factor that lies between zero and one and falls as the mandate becomes more stringent. Market structure is then determined by the ratio of the firms' harm-discounted technologies and the substitutability of their models. When harm-discounted technologies are sufficiently similar and models are not too substitutable, both firms are active and compete as differentiated duopolists. When one firm is sufficiently more productive in harm-discounted terms, it alone serves the market, either because its monopoly price already deters entry or because it prices to exclude its rival, as in [Dixit \(1979\)](#). The central implication of this analysis is that a responsible mandate acts like a tax on technology: a responsible firm competes as if it were less capable than it is or, equivalently, as if its marginal cost exceeded its financial cost.

We next endogenize technology for given mandates. Because a mandate acts like a tax on a firm's technology, it lowers the return to investing in model capability. Under symmetric mandates, both firms therefore choose lower model capability than they would as profit-driven firms. In this case, mandates slow the technology race, and a duopoly equilibrium exists unless model substitutability exceeds a threshold. Under asymmetric mandates, the laxer firm invests more aggressively, so that technology choice amplifies differences in mandates. The threshold for model substitutability below which both firms can remain active is lower when mandates are more asymmetric. Beyond this threshold, the market tips: the laxer firm invests its way into an uncontested monopoly, and the responsible firm rationally stays out. In an intermediate region, the duopoly and monopoly equilibria coexist.

Two features of the technology stage are worth highlighting. First, because payoffs depend on mandates only through harm-discounted technologies, model substitutability and the ratio of the mandate discount factors jointly characterize the set of equilibria. Second, guardrails cannot remove the handicap of responsible AI, but they reduce it: with guardrails, a given difference in mandates maps into a smaller difference in discount factors, so that responsible firms remain in the race under mandate gaps that would otherwise tip the market.

We then turn to the choice of mandates. As a benchmark, we hold technology fixed and consider founders who care only about profits. Here a charter plays the role of a delegation contract in the sense of [Fershtman and Judd \(1987\)](#). A firm with a more responsible mandate charges a higher price because harm internalization effectively raises its marginal cost. Because *prices are strategic complements* under Bertrand competition, the rival follows suit. A marginal mandate distorts the firm's own price only to second order but relaxes the rival's price to first order, so the commitment pays even unilaterally.²

²In contrast, under Cournot competition, as in [Bond and Levit \(2025\)](#), quantities are strategic

In equilibrium, both firms adopt a public-benefit mandate, and mandates are strategic complements. Under price competition alone, responsible mandates thus arise without any prosocial preferences.

Once technology is chosen, this logic fails. As shown above, a mandate weakens the firm in the technology race. Because *technology investments are strategic substitutes*, the rival responds by investing more to steal business. If model capability responds sufficiently strongly to changes in its marginal value, this capability response dominates the gain from softer price competition, and a mandate lowers the firm’s financial value.³ The condition applies with particular force to AI firms, where capability scales with compute, data, and algorithms. In our baseline model with linear demand, the capability response always dominates the price response in the duopoly region, so that each firm’s financial value is strictly decreasing in its own mandate. For profit-driven founders, pure profit charters are then the unique equilibrium, even though both firms would be more valuable under a common public-benefit mandate. The same logic extends to mission-driven founders. Their charters remain pure profit as long as their mission weight lies below a threshold; above the threshold, their charters understate it. In contrast, a monopolist, sheltered from the technology race, chooses a charter that reflects the founder’s preferences.

Our results have implications for policy. If the technology race eliminates responsible mandates even when founders would prefer to be responsible, the safety commitments of AI providers cannot be expected to withstand competition, and safety instruments such as guardrails, kill switches, or deployment controls may need to be mandated by regulation. The EU AI Act takes this approach for frontier models: providers of general-purpose models above a training-compute threshold must evaluate their models, including through adversarial testing, and mitigate systemic risks.⁴ Our analysis supports this view: a uniform mandate imposed on both firms slows the technology race but leaves the duopoly viable, whereas the same mandate adopted by one firm alone can tip the market. The analysis also identifies what eases the tension: the mission weight below

substitutes: one firm’s harm internalization, which raises its price and lowers its output, invites the other firm to expand its output and lower its price.

³We show that this condition, and the decomposition of a mandate’s value into the rival’s price and capability responses, hold in a general class of models in which firms compete in quality-adjusted offers (Armstrong and Vickers, 2001); our baseline model with linear demand is a special case.

⁴Regulation (EU) 2024/1689 (the Artificial Intelligence Act), Chapter V (Articles 51–56). Models trained with more than 10^{25} floating-point operations are presumed to present systemic risk (Article 51); their providers must evaluate the model, including through adversarial testing, assess and mitigate systemic risks, report serious incidents, and ensure cybersecurity (Article 55), in addition to the documentation, copyright, and training-data obligations that apply to all providers of general-purpose AI models (Article 53). These obligations apply from 2 August 2025.

which mission-driven founders adopt pure profit charters falls with the effectiveness of guardrails, so that improvements in guardrail technology make it less costly for founders to express their mission in the charter. Finally, since the technology race is global while regulation is national, the logic that sets AI firms against one another may also apply to jurisdictions that compete for them. We leave the analysis of regulatory competition to future work.

Related literature. Our paper contributes to the literature on competition and ethical behavior by firms. [Dewatripont and Tirole \(2024\)](#) consider product market competition and show that, under flexible prices, intense competition does not crowd out consequentialist ethics. Their result holds exactly at the guardrail margin of our model. In our setting, erosion instead operates through two margins outside its premises: technology is a sunk investment, so that a mandate shifts the firm’s position in the race and cannot be undone by repricing; and the mandate is a charter, a commitment whose expression competition governs. More broadly, our treatment of the corporate mandate as an equilibrium object relates to the literature on socially responsible firms and investors, including [Bénabou and Tirole \(2010\)](#), [Hart and Zingales \(2017\)](#), [Oehmke and Opp \(2025\)](#), [Bond and Levit \(2025\)](#), and [Bisceglia, Piccolo, and Schneemeier \(2026\)](#).

Methodologically, our model builds on the industrial organization literature. The product market is the linear differentiated duopoly of [Singh and Vives \(1984\)](#) and [Shubik and Levitan \(1980\)](#); the characterization of market structure, with monopoly pricing, limit pricing, and duopoly as the relative strength of the two firms varies, follows [Dixit \(1979\)](#); the investment stage draws on the analysis of innovation incentives under competition in [Vives \(2008\)](#); and the general treatment in Section 4 uses the utility-space formulation of [Armstrong and Vickers \(2001\)](#), which nests the discrete-choice systems of [Anderson, de Palma, and Thisse \(1992\)](#). Relative to this literature, our model contains two key innovations: harm generated by users, which places the externality inside the technology race, and endogenous public-benefit mandates. For the latter we rely on the literature on strategic delegation following [Vickers \(1985\)](#), [Fershtman and Judd \(1987\)](#), [Sklivas \(1987\)](#), and [Bond and Levit \(2025\)](#), in which owners commit to objectives that deviate from profit maximization in order to influence rivals. In our setting the corporate charter plays the role of the incentive contract, and the sign of the delegation gain depends on whether firms compete only in prices or also in a technology race.

Finally, our paper contributes to the emerging economic literature on artificial intelligence. [Acemoglu and Lensman \(2024\)](#) study the optimal adoption of a transformative technology under uncertainty about its risks, and [Jones \(2024\)](#) analyzes the trade-off be-

tween growth and existential risk. These papers treat the technology as given and focus on the planner’s problem. Closer to our question, [Armstrong, Bostrom, and Shulman \(2016\)](#) model an AI race in which competing developers reduce safety to win; in their setting the safety choice is a direct decision, whereas in ours responsibility is a corporate commitment whose survival is decided by capability investment. We study the market conduct of the firms that develop and deploy AI models, and how competition shapes their capability, guardrails, and corporate governance.

Paper outline. The paper proceeds as follows. Section 2 introduces the model and characterizes the product market and guardrail equilibrium as well as the technology race for given mandates. Section 3 makes mandates endogenous and contains our main results on whether responsible AI can compete. Section 4 establishes the mechanism beyond the linear specification. Section 5 concludes. All proofs are in Appendix A.

2 Model

We consider an economy with two competing AI providers. At date 0 (the first stage), each AI firm $j \in \{1, 2\}$ chooses a technology level $\tau_j \in \mathbb{R}^+$ at strictly convex investment cost $I(\tau_j)$, which we interpret as model capability. Technology is the only long-run commitment. At date 1 (the second stage), the two firms *simultaneously* choose prices p_j and (costless) guardrail intensities $g_j \in [0, 1]$ that limit the harmful use of AI. The interpretation is that guardrails are short-term policy instruments (system prompts, usage policies, refusal rules, deployment gates) that can be adjusted at the same frequency as prices, whereas technology takes time to build.

Users are indexed by a one-dimensional *exposure* type $x \in [0, 1]$, the weight of a user’s use case on the dual-use component of the technology. Exposure is distributed with density f and unit mass; our benchmark specification is uniform, $f \equiv 1$, though all results below depend on the exposure distribution only through its mean $\mathbb{E}(x)$. A user of exposure x served by firm j causes harm of $\tau_j x(1 - g_j)$ per unit of consumption: more capable models enable more harm and tighter guardrails deliver less harmful capability. Guardrails degrade the product uniformly for all users, so that their collateral cost is the forgone willingness to pay of legitimate users, while the harm they prevent is concentrated in high-exposure use.

In each of those exposure segments x , demand is generated by a representative con-

sumer,

$$q(\mathbf{p}; \mathbf{a}(x)) \equiv \arg \max_{q \in \mathbb{R}_+^2} (a_1(x) - p_1) q_1 + (a_2(x) - p_2) q_2 - \frac{\beta(\lambda)}{2} (q_1^2 + 2\lambda q_1 q_2 + q_2^2), \quad (1)$$

where we assume

$$a_j(x) = \tau_j (1 - \kappa(g_j)). \quad (2)$$

This specification captures the idea that technology is an attribute that increases willingness to pay whereas guardrails generate a nuisance to users. For simplicity, we assume that this nuisance is uniform across segments. The nuisance function κ is strictly convex and satisfies $\kappa(0) = \kappa'(0) = 0$ and $\lim_{g \rightarrow 1} \kappa'(g) = \infty$: successive guardrails cut ever deeper into legitimate use.

Whenever both firms sell positive quantities, the consumer problem (1) generates a linear (inverse) demand system

$$p_j = a_j(x) - \beta(\lambda) (q_j(x) + \lambda q_{-j}(x)), \quad (3)$$

where $\lambda \in [0, 1)$ governs substitutability and $\beta(\lambda) > 0$ the common slope of demand in all segments (nesting, for a constant $b > 0$, the Singh–Vives (1984) case $\beta = b$ and the Shubik–Levitan case $\beta = 2b/(1 + \lambda)$). The corner cases of (1) are characterized in Appendix A.

Firm objective. The firm cannot price-discriminate across exposures, so it sets a single price p_j and faces total demand $Q_j = \int_0^1 q_j(x) f(x) dx$. At date 1, the product market payoff consists of revenue and harm internalization,

$$\Pi_j^1(\mathbf{p}, \mathbf{g} | \boldsymbol{\tau}) \equiv p_j Q_j - \gamma_j H_j, \quad (4)$$

where $\gamma_j \in [0, 1]$ denotes firm j 's degree of internalization of the harm it causes,

$$H_j \equiv \tau_j (1 - g_j) \int_0^1 x q_j(x) f(x) dx. \quad (5)$$

For now, we think of γ_j as predetermined by the corporate charter. The case $\gamma_j = 0$ is profit maximization.

The date-0 objective is the date-1 product market payoff net of the investment cost,

$$\Pi_j^0 = \Pi_j^1(\mathbf{p}, \mathbf{g} | \boldsymbol{\tau}) - I(\tau_j).$$

2.1 Product market and guardrail equilibrium

Problem 1 (Bertrand pricing-and-guardrail game $\mathcal{G}(\tau)$) *Given τ , firms $j \in \{1, 2\}$ simultaneously choose $(p_j, g_j) \in \mathbb{R}_+ \times [0, 1]$, where we delete the weakly dominated strategies that price below the internalized cost of harm, $p_j < \gamma_j \tau_j (1 - g_j) \mathbb{E}(x)$.⁵ An equilibrium is a strategy profile $(\mathbf{p}^*, \mathbf{g}^*)$ of undominated strategies such that*

$$(p_j^*, g_j^*) \in \arg \max_{(p_j, g_j)} \Pi_j^1(p_j, p_{-j}^*, g_j, g_{-j}^* | \tau)$$

for both firms. Firm j is called active if and only if $Q_j > 0$.

It is useful to map our non-standard model (with user-generated harm and guardrails) into a classical IO model. Since intercepts (2) do not depend on x , all exposure segments purchase the same quantity, $q_j(x) \equiv q_j = Q_j$, so that the product market payoff (4) becomes

$$\Pi_j^1 = [p_j - \gamma_j \tau_j (1 - g_j) \mathbb{E}(x)] q_j;$$

internalized harm acts as a constant marginal cost. A guardrail lowers this cost but, by (2), destroys willingness to pay $\tau_j \kappa(g_j)$ on every unit sold, a productivity cost that the firm bears through the price it can charge at any given user surplus $a_j - p_j$. The harm-adjusted unit cost of operating is thus

$$c_j(g_j) := \tau_j \gamma_j (1 - g_j) \mathbb{E}(x) + \tau_j \kappa(g_j), \quad (6)$$

which consists of the per unit cost of internalized harm (at the chosen guardrails) and the per unit productivity cost of guardrails. As is intuitive, the optimal guardrail choice minimizes this harm-adjusted unit cost by equating the marginal productivity loss with the marginal reduction in the internalized externality.

Lemma 1 (Guardrail choice and harm-discounted technology) *In every equilibrium of $\mathcal{G}(\tau)$, the guardrail of an active firm j minimizes its harm-adjusted unit cost (6): g_j^* uniquely solves*

$$\kappa'(g_j^*) = \gamma_j \mathbb{E}(x),$$

⁵ Deleting below-cost pricing rules out equilibria in which an excluded firm depresses the active firm's price with below-cost offers that never result in sales (see the proof of Lemma 2). Pricing exactly *at* the internalized cost must be retained: it is the boundary threat that supports the limit-pricing equilibrium and, since it is itself weakly dominated (by any small markup), requiring fully undominated strategies would leave the limit-pricing regime with no equilibrium, the familiar open-set problem of Bertrand games.

independently of the rival's strategy, the own price, technologies, and λ , and g_j^* is increasing in γ_j . Given $\mathbf{g}^*$, the date-1 game is a differentiated Bertrand duopoly with quality τ_j and minimized unit cost $c_j(g_j^*)$, and equilibrium margins, quantities, and payoffs depend on (τ_j, γ_j) only through the harm-discounted technology $\delta(\gamma_j) \tau_j$ where

$$\delta(\gamma_j) := \frac{\tau_j - c_j(g_j^*)}{\tau_j} = 1 - \kappa(g_j^*) - \gamma_j(1 - g_j^*) \mathbb{E}(x) \quad (7)$$

Internalization thus acts exactly like a technology handicap, which optimally chosen guardrails mitigate, i.e., $c_j(0) > c_j(g_j^*) > 0$ whenever $\gamma_j > 0$. We refer to $\delta(\gamma) \in (0, 1]$ as the *mandate discount factor*. By the envelope theorem, $\delta'(\gamma) = -(1 - g^*(\gamma)) \mathbb{E}(x) < 0$, with $\delta(0) = 1$ and $\delta(\gamma) \geq 1 - \gamma \mathbb{E}(x)$: internalization taxes technology, and the gap to $1 - \gamma \mathbb{E}(x)$ is the value of the guardrail option.

Note also what Lemma 1 rules out: the guardrail margin is untouched by competition. Conditional on its charter, how much harm a firm abates per unit of delivered capability, and with it $\delta(\gamma)$, depends neither on λ nor on the rival. This is an instance of the irrelevance result of Dewatripont and Tirole (2024), whose premises hold exactly at this margin: the cost of moral behavior scales with output, harm is internalized in proportion to own sales, and prices are flexible. In our model, competition will reach harm through the margins outside that logic: the technology race (Section 2.2) and the choice of charters (Section 3).

We label firms in this section such that firm 1 is the (weakly) stronger firm in harm-discounted terms, $\delta(\gamma_1) \tau_1 \geq \delta(\gamma_2) \tau_2$. With the relative technology $\frac{\tau_2}{\tau_1}$, the relative harm-discounted technology is then the product of the discount ratio and the technology ratio, $\frac{\delta(\gamma_2) \tau_2}{\delta(\gamma_1) \tau_1} \in [0, 1]$. As the next lemma shows, the market structure is determined by this ratio, jointly with the substitutability λ .

Lemma 2 (Date-1 equilibrium) *Fix $\boldsymbol{\tau}$ with $\delta(\gamma_1) \tau_1 > 0$. A pure-strategy equilibrium of $\mathcal{G}(\boldsymbol{\tau})$ exists for all parameters, equilibrium payoffs and the guardrails, margins, and quantities of active firms are unique, and two regimes obtain:*

(i) **Entry deterrence** $\frac{\delta(\gamma_2) \tau_2}{\delta(\gamma_1) \tau_1} < \frac{\lambda}{2 - \lambda^2}$: only firm 1 is active and $\Pi_2^1 = 0$, with two pricing subcases:

(a) Monopoly pricing $\frac{\delta(\gamma_2) \tau_2}{\delta(\gamma_1) \tau_1} \leq \lambda/2$: firm 1 charges its unconstrained monopoly price, $p_1 = \gamma_1 \tau_1 (1 - g_1^*) \mathbb{E}(x) + \frac{\delta(\gamma_1) \tau_1}{2}$, so that

$$\Pi_1^1 = \frac{[\delta(\gamma_1) \tau_1]^2}{4\beta(\lambda)}.$$

- (b) Limit pricing $\frac{\lambda}{2} < \frac{\delta(\gamma_2)\tau_2}{\delta(\gamma_1)\tau_1} < \frac{\lambda}{2-\lambda^2}$: firm 1 prices to deter firm-2 entry, with margin $p_1 - \gamma_1\tau_1(1 - g_1^*) \mathbb{E}(x) = \delta(\gamma_1)\tau_1 - \delta(\gamma_2)\tau_2/\lambda$, and firm 2 prices at its internalized cost, so that

$$\Pi_1^1 = \left(\delta(\gamma_1)\tau_1 - \frac{\delta(\gamma_2)\tau_2}{\lambda} \right) \frac{\delta(\gamma_2)\tau_2}{\lambda\beta(\lambda)}.$$

- (ii) **Duopoly** $\frac{\delta(\gamma_2)\tau_2}{\delta(\gamma_1)\tau_1} \geq \lambda/(2 - \lambda^2)$: both firms are active, with

$$p_j = \gamma_j\tau_j(1 - g_j^*) \mathbb{E}(x) + \frac{(2 - \lambda^2)\delta(\gamma_j)\tau_j - \lambda\delta(\gamma_{-j})\tau_{-j}}{4 - \lambda^2},$$

$$\Pi_j^1 = \frac{[(2 - \lambda^2)\delta(\gamma_j)\tau_j - \lambda\delta(\gamma_{-j})\tau_{-j}]^2}{(4 - \lambda^2)^2(1 - \lambda^2)\beta(\lambda)}.$$

In regimes (i.b) and (ii), equilibrium strategies are unique; in regime (i.a), the excluded firm's (payoff-irrelevant) strategy (p_2, g_2) is indeterminate among undominated strategies.

Intuitively, whether there is space for one or two firms on the product market depends on their substitutability λ and their relative harm-discounted technology $\frac{\delta(\gamma_2)\tau_2}{\delta(\gamma_1)\tau_1}$. When firm 1 is sufficiently more productive, charging the unconstrained monopoly price alone keeps firm 2 out. In subcase (b) exclusion requires an active limit price. Dixit (1979) formalizes this distinction in a closely related linear Cournot duopoly with fixed entry costs. Note also that equilibrium payoffs are homogeneous of degree two in the harm-discounted technologies: scaling (τ_1, τ_2) by $s > 0$ leaves guardrails and the regime unchanged and scales margins and quantities by s and payoffs by s^2 . The market structure thus depends on technologies only through the ratio τ_2/τ_1 , while their scale only multiplies profits.

2.2 Technology choice

We now turn to technology choice. Since marginal operating payoffs are linear in own technology, a cubic cost delivers interior solutions in closed form: we set

$$I(\tau_j) = \frac{k\tau_j^3}{3}.$$

By Lemma 1, technology enters the product market only through the harm-discounted technology $\delta(\gamma_j)\tau_j$. The cost scale $k > 0$ and the demand slope $\beta(\lambda)$ enter all date-0 objects only through multiplicative scales, so that the structure of equilibrium depends only on λ and on the mandates through the discount ratio $\delta(\gamma_2)/\delta(\gamma_1) \in [0, 1]$. The latter

ratio captures the *effective mandate symmetry* (where firm 2 is weakly more responsible).

Since an excluded firm earns zero operating payoff, no firm ever invests into the entry-deterrence region of Lemma 2: at date 0, each firm chooses either $\tau_j = 0$ or a technology at which it is active (see the proof of Corollary 3). Date-0 equilibrium outcomes are therefore of two types, *monopoly* (one firm out, the other at its unconstrained investment level) and *both-active* equilibria.

In the both-active region, firm j 's first-order condition equates the marginal investment cost to the marginal value of technology:

$$k\tau_j^2 = \delta(\gamma_j) \left(1 - \frac{\lambda^2}{4 - \lambda^2}\right) q_j, \quad j = 1, 2. \quad (8)$$

The three factors have a transparent interpretation. A unit of technology delivers $\delta(\gamma_j)$ units of harm-discounted quality (the mandate discount). At the pricing equilibrium, the direct value of a unit of harm-discounted quality equals output q_j ; since quality and cost reduction coincide in $\delta(\gamma_j) \tau_j$, this is the classic observation that the incentive to innovate is larger when the firm produces a larger output (cf. Vives, 2008). The factor $1 - \lambda^2/(4 - \lambda^2) \in (2/3, 1]$, finally, is the strategic devaluation of investment under price competition: a unit of own quality induces the rival to cut its margin by $\lambda/(4 - \lambda^2)$, which claws back the fraction $\lambda^2/(4 - \lambda^2)$ of the output value: nothing at $\lambda = 0$, one third as $\lambda \rightarrow 1$. Substituting the equilibrium quantities of Lemma 2 and taking ratios, the relative technology $\nu := \tau_2/\tau_1$ solves the cubic

$$\lambda(\nu^3 - 1) - (2 - \lambda^2) \left[\frac{\delta(\gamma_1)}{\delta(\gamma_2)} \nu^2 - \frac{\delta(\gamma_2)}{\delta(\gamma_1)} \nu \right] = 0. \quad (9)$$

Proposition 1 (Both-active equilibrium) *Let $\gamma_2 \geq \gamma_1$. A both-active equilibrium exists if and only if substitutability is below a cutoff*

$$\lambda \leq \bar{\lambda} \left(\frac{\delta(\gamma_2)}{\delta(\gamma_1)} \right), \quad (10)$$

where the cutoff $\bar{\lambda}(\eta)$ is the unique solution in $(0, \bar{\lambda}(1)]$ of $9\lambda(T^2 - 3\lambda^2)/(2T^3) = \eta^3$ with $T := 2 - \lambda^2$, strictly increasing in the effective mandate symmetry $\eta = \frac{\delta(\gamma_2)}{\delta(\gamma_1)}$, with $\bar{\lambda}(1) = (\sqrt{17} - 3)/2 \approx 0.56$; at the cutoff the responsible firm exactly breaks even. The both-active equilibrium is unique. Relative technologies solve (9). Under symmetric mandates, equilibrium technology investments are equal. Otherwise, the laxer firm 1 leads the

technology race

$$\frac{\tau_1^*}{\tau_2^*} \geq \left(\frac{\delta(\gamma_1)}{\delta(\gamma_2)} \right)^2 > 1.$$

Intuitively, both firms can remain active if and only if product substitutability is sufficiently low, so that both firms can remain profitable after accounting for the cost of investment. Asymmetric internalization causes the laxer firm to dominate the technology race in a duopoly and reduces the threshold beyond which only one firm can remain profitable. Uniqueness rules out, in particular, asymmetric both-active equilibria under symmetric mandates.⁶ Note also that limit pricing, subcase (i.b) of Lemma 2, never occurs on the date-0 equilibrium path: excluding an active rival presupposes that the rival has invested, but a firm that never sells does not invest. On path, entry is either blockaded outright or both firms are active.

A second cutoff governs when the market can instead sustain a blockaded monopoly.

Proposition 2 (Blockaded monopoly) *Let $\gamma_2 \geq \gamma_1$. A blockaded monopoly, in which one firm alone invests at its unconstrained level $\delta(\gamma_j)^2 / (2k\beta(\lambda))$ with no strategic over-investment, exists for the laxer firm 1 if and only if substitutability is above a second cutoff,*

$$\lambda \geq \underline{\lambda} \left(\frac{\delta(\gamma_2)}{\delta(\gamma_1)} \right), \quad (11)$$

where the cutoff $\underline{\lambda}(\eta)$ is the unique solution in $(0, \underline{\lambda}(1)]$ of $f_b(\lambda) := 9\lambda(4 - \lambda^2)(T^2 - \lambda^2)/(8T^3) = \eta^3$, strictly increasing in $\eta = \delta(\gamma_2)/\delta(\gamma_1)$, with $\underline{\lambda}(1) \approx 0.449$, and it lies strictly below the both-active cutoff, $\underline{\lambda}(\eta) < \bar{\lambda}(\eta)$. The responsible firm 2 can be a blockaded monopolist only under nearly symmetric mandates, $\delta(\gamma_1)/\delta(\gamma_2) \lesssim 1.14$, and its region is a strict subset of the laxer firm's: whenever the responsible firm can survive alone, so can the laxer firm, but not conversely.⁷

The tipping condition (10) thus demarcates where the duopoly *can* survive, and (11) where a monopoly *can* persist: blockade requires enough substitutability for the

⁶ At $\gamma_1 = \gamma_2$ the cubic (9) factors as $(\nu - 1)(\lambda\nu^2 - (2 - \lambda^2 - \lambda)\nu + \lambda) = 0$. The roots ν and $1/\nu$ of the quadratic factor form an asymmetric pair of stationary points of the first-order conditions, along which $\tau_1 + \tau_2$ is pinned at a constant; they are not equilibria: at each, the smaller firm's harm-discounted ratio falls below $3\lambda/(2 - \lambda^2)$, so that its operating profit cannot cover its investment cost (see the proof of Corollary 3) and exit is a profitable deviation.

⁷ Blockade means that the rival's best entry response against the monopolist earns at most zero. The responsible firm's monopoly is a consequence, by continuity, of the labeling symmetry at $\gamma_1 = \gamma_2$. Its exact region is characterized in the proof: since f_b peaks at ≈ 1.47 at $\lambda \approx 0.76$, the responsible firm's monopoly exists if and only if $(\delta(\gamma_1)/\delta(\gamma_2))^3 \leq f_b(\lambda)$ for $\lambda \leq \lambda^* \approx 0.765$ and $(\delta(\gamma_1)/\delta(\gamma_2))^3 \leq 9/(8\lambda)$ beyond; the region closes at high substitutability, except under nearly identical mandates, because the laxer rival would profitably leapfrog the incumbent.

incumbent's unconstrained investment alone to deter entry. Together, the two cutoffs organize the equilibrium set.

Corollary 1 (Market structure) *As substitutability rises, keeping the degree of mandate symmetry $\delta(\gamma_2)/\delta(\gamma_1)$ fixed, the market passes through three regions:*

- (i) Low substitutability ($\lambda < \underline{\lambda}$): *the unique equilibrium has both firms active.*
- (ii) Intermediate substitutability ($\underline{\lambda} \leq \lambda \leq \bar{\lambda}$): *the both-active equilibrium and the blockaded monopoly coexist; which one prevails is decided by entry expectations alone.*
- (iii) High substitutability ($\lambda > \bar{\lambda}$): *only one firm serves the market, and absent nearly symmetric mandates it is the laxer one.*

Holding substitutability fixed instead, the same two cutoffs read as bounds on the mandate gap.

Corollary 2 (Mandate gaps) *Fix $\lambda < \underline{\lambda}(1) \approx 0.449$. As the mandate gap widens, with the discount ratio $\delta(\gamma_2)/\delta(\gamma_1)$ falling from one, the market passes through the same three regions in mirror image:*

- (i) Similar mandates ($\delta(\gamma_2)/\delta(\gamma_1) > \underline{\lambda}^{-1}(\lambda)$): *the unique equilibrium has both firms active.*
- (ii) Intermediate gaps ($\bar{\lambda}^{-1}(\lambda) \leq \delta(\gamma_2)/\delta(\gamma_1) \leq \underline{\lambda}^{-1}(\lambda)$): *the both-active equilibrium and the blockaded monopoly coexist.*
- (iii) Large gaps ($\delta(\gamma_2)/\delta(\gamma_1) < \bar{\lambda}^{-1}(\lambda)$): *only the laxer firm serves the market.*

For $\lambda \geq \underline{\lambda}(1)$, region (i) is empty, since even identical mandates support the blockaded monopoly, and beyond $\bar{\lambda}$ only a monopoly remains (the laxer firm's or, under nearly symmetric mandates, possibly the responsible firm's).

Figure 1 assembles these results into a single picture. Both cutoffs depend on mandates only through the effective mandate symmetry $\delta(\gamma_2)/\delta(\gamma_1)$, which equals one when mandates are equally stringent in effective terms and falls toward zero as they diverge. Because date-0 profits depend on mandates only through the harm-discounted technologies $\delta(\gamma_j)\tau_j$, this symmetry and the substitutability λ are jointly a sufficient statistic for the equilibrium set: mandate levels scale investments and profits but never the market

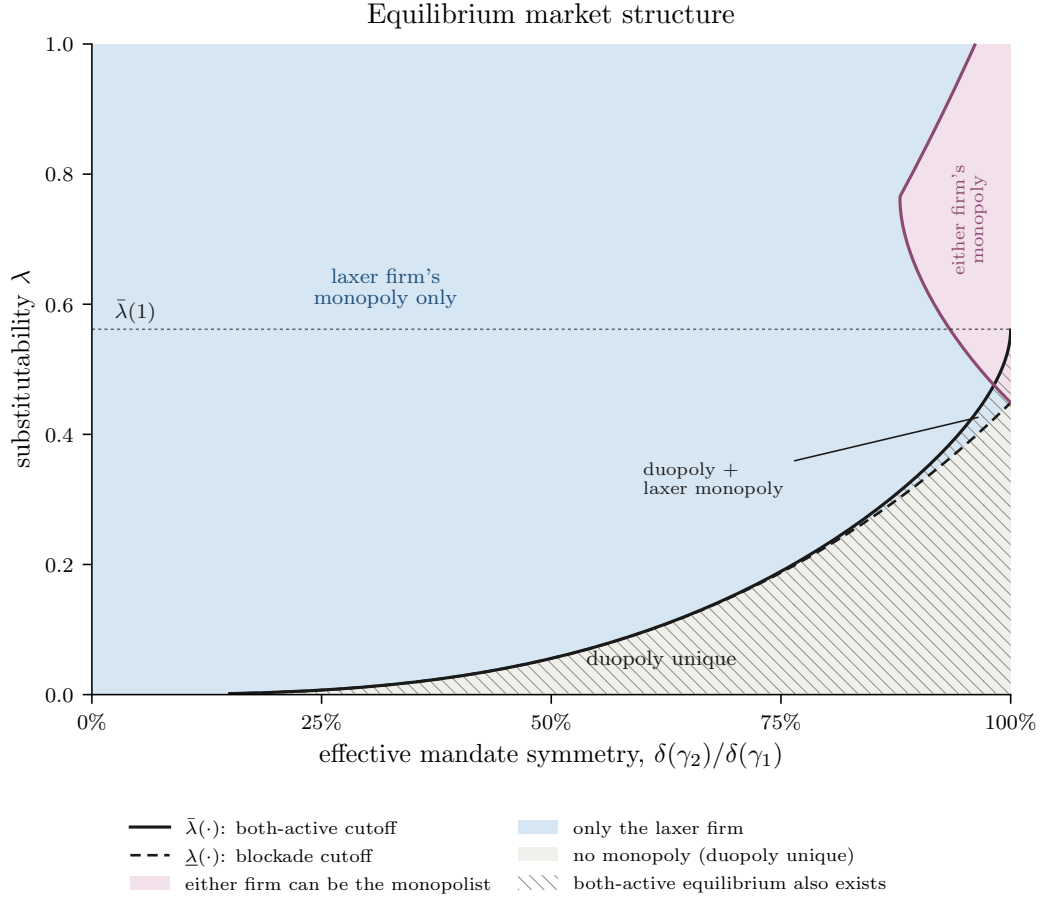


Figure 1. Equilibrium market structure. For every pair of effective mandate symmetry $\delta(\gamma_2)/\delta(\gamma_1)$ and substitutability λ , the figure shows which date-0 equilibria exist (firm 1 is the weakly laxer firm, $\gamma_1 \leq \gamma_2$). The solid curve is the both-active cutoff $\bar{\lambda}(\cdot)$ of Proposition 1: a both-active equilibrium exists in the hatched region below it. The dashed curve is the blockade cutoff $\underline{\lambda}(\cdot)$ of Proposition 2: the laxer firm's blockaded monopoly is an equilibrium everywhere above it. In the wedge at the top right, mandates are similar enough that the responsible firm can be the blockaded monopolist as well; the wedge closes at high substitutability, where the laxer rival would profitably leapfrog a responsible incumbent. Because the two axes form a sufficient statistic for the date-0 game, the figure involves no parameterization: it is a complete characterization of the activity regimes.

structure, so the figure is a complete characterization of activity regimes rather than a numerical illustration. Reading up a vertical line traces Corollary 1 (the unique duopoly, then coexistence between the two cutoffs, then monopoly alone), and reading leftward along a horizontal line traces Corollary 2. The wedge at the top right is where the identity of the blockaded monopolist is indeterminate. Its boundary is non-monotone in λ : a responsible incumbent must deter both entry margins of the laxer rival, entry alongside the incumbent, which higher substitutability discourages, and leapfrogging past it, which higher substitutability encourages. The responsible firm's monopoly therefore appears

at intermediate substitutability and closes again at high substitutability.

The technology stage yields three conclusions. Even though guardrails are chosen optimally at date 1, softening the mandate handicap from $1 - \gamma \mathbb{E}(x)$ to $\delta(\gamma)$, internalization still operates as a proportional tax on technology. Competition does not attenuate mandate asymmetries but amplifies them—at least cubically, and without bound as λ approaches $\bar{\lambda}$ —and for moderate substitutability the market tips: the least-internalizing firm invests its way into an uncontested monopoly, while the responsible firm rationally stays out. What guardrails deliver is compression of the discount ratio itself: since $|\delta'(\gamma)| = (1 - g^*) \mathbb{E}(x) < \mathbb{E}(x)$ whenever $g^* > 0$, a given difference in mandates maps into a smaller difference in discount factors than without guardrails, relaxing the tipping condition (10) in mandate space or, equivalently, raising both cutoffs $\underline{\lambda}$ and $\bar{\lambda}$.

We close the technology analysis with the symmetric benchmark $\gamma_1 = \gamma_2 = \gamma$, the natural case of an industry-wide standard, whether imposed by a regulator or adopted in identical corporate charters. Symmetry keeps the duopoly viable: by Proposition 1, the both-active equilibrium exists at every common mandate whenever $\lambda < \bar{\lambda}(1)$ (for $\lambda \geq \underline{\lambda}(1)$ it coexists with a blockaded monopoly), so the mandate shifts outcomes along the intensive margin.

Corollary 3 (Symmetric mandates) *Let $\gamma_1 = \gamma_2 = \gamma$: for every $\lambda < \bar{\lambda}(1)$, the unique both-active equilibrium is symmetric, with*

$$\tau_s = \frac{\delta(\gamma)^2}{2k\beta(\lambda)} \cdot \left(1 - \frac{\lambda^2}{4 - \lambda^2}\right) \cdot \frac{1}{\left(1 - \frac{\lambda}{2}\right)(1 + \lambda)}, \quad (12)$$

$$\Pi_j^0 = \frac{\delta(\gamma)^2 \tau_s^2}{(4 - \lambda^2)^2 (1 - \lambda^2) \beta(\lambda)} \cdot \frac{(2 - \lambda^2 - \lambda)(2 - \lambda^2 - 3\lambda)}{3}. \quad (13)$$

Π_j^0 is positive for $\lambda < \bar{\lambda}$ and equals zero exactly at $\lambda = \bar{\lambda}(1)$: the both-active equilibrium survives precisely as long as the duopoly recovers its investment cost. For $\lambda > \bar{\lambda}$, the unique equilibrium (up to labeling) is a monopoly with blockaded entry at

$$\tau_1 = \frac{\delta(\gamma)^2}{2k\beta(\lambda)}. \quad (14)$$

The three factors of τ_s , (12), mirror (8): the monopoly investment (14), the strategic devaluation of investment, and the ratio of equilibrium output per unit of technology in the symmetric duopoly relative to monopoly, bounded above by 1: with independent products ($\lambda = 0$) each duopolist invests exactly like a monopolist, and competition depresses investment below the monopoly level for every $\lambda \in (0, \bar{\lambda})$. Mandates slow the technology race, $\tau_s \propto \delta(\gamma)^2$, the more so without guardrails.

3 The Choice of Mandates

So far, mandates were fixed characteristics of the two firms. We now make them choices. Each founder starts out with true preferences and selects a corporate charter, a commitment to the mandate γ_j under which the firm will operate, before technologies and prices are chosen. The charter may deviate from those preferences in either direction: a mission-driven founder can adopt a charter that understates or abandons the mission, and a profit-driven founder can adopt a public-benefit charter. We ask which mandates are chosen in equilibrium.

3.1 Profit-driven founders

We begin with the starkest case: both founders are ex ante profit maximizers, so that a chartered mandate can only be a strategic commitment. We proceed in two steps: first with technologies exogenously fixed, so that charters play out in the pricing game alone, and then with technology chosen.

Suppose first, then, that technologies are exogenously fixed at a common level $\tau_1 = \tau_2 = \tau > 0$, and write V_j for firm j 's expected financial value. Because internalized harm involves no cash outflow, financial value differs from the mandate objective: in the symmetric pricing equilibrium without guardrails ($g \equiv 0$),

$$V_j = p_j Q_j = \frac{(1 - \gamma \mathbb{E}(x)) (1 - \lambda + \gamma \mathbb{E}(x)) \tau^2}{(2 - \lambda)^2 (1 + \lambda) \beta(\lambda)}, \quad (15)$$

maximal at the interior mandate $\gamma \mathbb{E}(x) = \lambda/2$: duopoly prices lie below the revenue-maximizing level, so a common mandate softens price competition, which is the strategic-delegation logic of [Fershtman and Judd \(1987\)](#), with the mandate in the role of the managerial incentive contract. If charters could be coordinated, profit-driven founders would thus adopt public-benefit mandates. Would they adopt them unilaterally? A charter is exactly a delegation contract: it commits the firm to price as if its marginal cost were higher. Because prices are strategic complements, such a commitment pays even *unilaterally*: at $\gamma_1 = \gamma_2 = 0$,

$$\frac{dV_j}{d\gamma_j} = \frac{\lambda^2 \mathbb{E}(x) \tau m_j}{(4 - \lambda^2) (1 - \lambda^2) \beta(\lambda)} > 0,$$

where m_j is the common equilibrium margin: a marginal mandate distorts the own price only to second order but relaxes the rival's price to first order. Pure profit charters therefore cannot survive, and the fixed-technology charter game can be solved in closed

form.

Proposition 3 (Strategic mandates under fixed technology) *Let both founders maximize expected financial value, let technologies be exogenously fixed at a common level $\tau > 0$, let $g \equiv 0$, let $0 < \lambda \leq 1/2$, and let $\mathbb{E}(x) \geq \lambda^2/4$ (so that the charter cap $\gamma_j \leq 1$ is slack). Against any rival charter that leaves both firms active,⁸ the best response is unique, carries a strictly positive mandate, and is linear and strictly increasing in the rival's mandate,*

$$\delta(\gamma_j) = \frac{4 - \lambda^2}{4} + \frac{\lambda^3}{4(2 - \lambda^2)} \delta(\gamma_{-j}) : \quad (16)$$

mandates are strategic complements, with slope $d\gamma_j/d\gamma_{-j} = \lambda^3/[4(2 - \lambda^2)] \in (0, 1)$. The charter game has a unique equilibrium. It is symmetric and interior: both firms adopt the public-benefit mandate

$$\gamma^F \mathbb{E}(x) = \frac{\lambda^2(1 - \lambda)(2 + \lambda)}{8 - 4\lambda^2 - \lambda^3} = \frac{\lambda^2}{4} + O(\lambda^3), \quad (17)$$

strictly increasing in λ on the stated range and strictly below the collusive pricing optimum $\lambda/2$ of (15).

Under fixed technology, then, public-benefit mandates without mission-driven preferences are not merely sustainable but are the unique equilibrium outcome: each firm adopts a public-benefit charter for purely anticompetitive reasons, and strategic complementarity makes the two charters mutually reinforcing. The equilibrium mandate is nonetheless modest: second order in substitutability, against the first-order collusive optimum $\lambda/2$, since unilateral delegation internalizes only a small part of the collective gain. (With guardrails, the positive corner derivative is unchanged, since $g^*(0) = 0$ and $\kappa(g^*)$ is of second order; the linear best response (16) is specific to $g \equiv 0$.)

Does the delegation logic survive once technology is chosen? The *collective* motive does, and it strengthens. A common mandate now also de-escalates the technology race: at $\gamma = 0$ firms overinvest relative to the joint financial optimum (business stealing), so the induced reduction of both technologies, (12), saves investment cost to first order; across λ , the date-0 gain is 1.5–2 times the fixed-technology gain. The technology response cuts both ways, since lower equilibrium technologies scale down revenue ($\tau_s \propto \delta(\gamma)^2$, so revenue carries $\delta(\gamma)^4$), which is why the collusive optimum falls from $\lambda/2$ to

$$\gamma^* \mathbb{E}(x) = \frac{\lambda(6 + \lambda)}{2(10 + 3\lambda - 2\lambda^2)} < \frac{\lambda}{2}, \quad (18)$$

⁸ For $\mathbb{E}(x) \leq 5/7$ every charter pair leaves both firms active on the stated range, since $1 - \mathbb{E}(x) \geq 2/7 \geq \lambda/(2 - \lambda^2)$ for $\lambda \leq 1/2$; the qualifier is then vacuous.

still strictly positive (Appendix A.5).

Whether the *unilateral* logic survives is a race between two responses to a marginal mandate. We decompose its value at $\gamma_1 = \gamma_2 = 0$ into the fixed-technology delegation gain $G > 0$ (the display preceding Proposition 3) and the loss L from the rival's investment response; the own investment response drops out by the envelope theorem. The general answer turns on the *elasticity of capability supply*. For this result only, we generalize the investment cost to $I(\tau) = k\tau^n/n$ with $n > 2$, so that equilibrium capability responds to a proportional shift in its marginal value with elasticity $1/(n-2)$; the maintained specification is $n = 3$.

Proposition 4 (When the race dominates) *Let $g \equiv 0$ and $I(\tau) = k\tau^n/n$ with $n > 2$, and decompose the marginal value of a mandate at $\gamma_1 = \gamma_2 = 0$, evaluated at the symmetric interior investment equilibrium,⁹ as $dV_j/d\gamma_j = G + L$, with $G > 0$ the fixed-technology delegation gain and L the value of the rival's investment response. Investments are strategic substitutes, so $L < 0$, and as $\lambda \rightarrow 0$,*

$$\frac{L}{G} \rightarrow -\frac{n}{(n-2)^2} : \quad (19)$$

for weak substitutability, a marginal mandate destroys financial value if $n < 4$ and does not if $n > 4$: the elasticity of capability supply, $1/(n-2)$, must exceed $1/2$.

Proposition 4 turns the reversal into a statement about technology, and thereby delimits its scope. The two polar cases frame this section: as $n \rightarrow \infty$, capability supply becomes inelastic, the race response vanishes, and the fixed-technology world of Proposition 3, where strategic mandates thrive, reemerges; as n falls toward 2, the race response grows without bound (for $n < 3$ the symmetric investment equilibrium loses stability before $\bar{\lambda}(1)$, at the root of $2 - \lambda^2 = n\lambda/(n-2)$). The threshold $n < 4$ is conservative: $|L/G|$ increases in λ , the boundary widens from 4 to ≈ 5.1 at $\lambda = 0.45$, and on the maintained specification $n = 3$ the ratio grows from 3 at weak substitutability without bound as $\lambda \rightarrow \bar{\lambda}(1)$.¹⁰ The condition has a clear economic content: competition punishes mandates through the rival's *expansion*, so it does so only where capability can be expanded at gently rising marginal cost. Scalable technologies—capability built from compute, data, and algorithms, as in the AI race—are the elastic case; mature capacity

⁹ For $n = 3$ this is the equilibrium of Corollary 3; for other n a regular symmetric interior equilibrium exists for small λ by continuity from the decoupled case $\lambda = 0$.

¹⁰ Numerical statements in the text and in the proofs are verified in the replication code accompanying the paper (`robustness_race.py` for the elasticity comparisons and `mandate_game.py` for the charter game).

with steeply rising marginal cost—cement, say—is the inelastic one, where the pricing layer dominates and public-benefit charters survive as ordinary strategic commitments. Competition suppresses mandates only for technologies that scale. Section 4 establishes the decomposition and the elasticity threshold without the linear specification.

On the maintained specification the reversal is not merely marginal but global:

Proposition 5 (No mandates without mission) *Let both founders maximize expected financial value and let $\lambda < \bar{\lambda}(1)$. For any rival charter $\gamma_{-j} \in [0, 1]$, firm j 's financial value is strictly decreasing in its own mandate γ_j wherever firm j is active.¹¹ The unique equilibrium of the charter game is $\gamma_1 = \gamma_2 = 0$: even though both firms' financial value at the common mandate γ^* of (18) strictly exceeds its value under pure profit charters (for $g \equiv 0$; with guardrails, a small common mandate raises both values), no positive mandate survives unilateral deviation: the charter game is a prisoner's dilemma.*

The proof (Appendix A.9) rests on an envelope argument that isolates why a mandate cannot pay. At $\gamma_j = 0$ the charter objective coincides with financial value, so firm j 's own investment and pricing are already financially optimal, and a marginal mandate affects value only through the rival's responses. These pull in opposite directions. In the pricing subgame, prices are strategic complements: a commitment to price softly induces the rival to price softly as well, the delegation logic behind the collective gain of the fixed-technology benchmark. In the technology subgame, however, investments are strategic substitutes: a firm with a stricter mandate invites the rival to invest more, and the induced technology gap erodes margins. The proof shows that the substitutes response dominates the complements response at every point of the both-active region, not only at $\gamma_j = 0$, so that value falls monotonically in the own mandate all the way to the corner. The same amplification that tips markets under asymmetric mandates makes a unilateral mandate self-defeating here: whatever a mandate saves in price competition, it more than concedes in the technology race (Proposition 4: at $n = 3$, at least three times over). For $\lambda \geq \bar{\lambda}(1)$ the conclusion is immediate: the continuation is a blockaded monopoly under any charters, and a monopolist's financial value, too, falls in its own mandate. Anticompetitive motives alone therefore cannot explain public-benefit charters once technology is at stake; mandates must reflect genuine preferences, to which we turn next.

¹¹ Where the both-active equilibrium coexists with a blockaded monopoly, we select the both-active equilibrium (under a blockaded monopoly the excluded firm earns zero); for $\lambda < \underline{\lambda}(1)$ the continuation is unique, so no selection is involved. The same selection applies in Proposition 6.

3.2 Mission-driven founders

Suppose instead that both founders genuinely care about the harm their technologies cause: founder j 's payoff is $U_j = V_j - \theta H_j$, where V_j is financial value, $H_j := (1 - g_j) \mathbb{E}(x) \tau_j Q_j$ is firm j 's expected harm, and $\theta \geq 0$ is the common mission weight. The charter remains a free commitment. Because the mandate objective deducts internalized harm that involves no cash outflow, it satisfies the identity $\Pi_j^0 = V_j - \gamma_j H_j$: a charter with mandate γ_j is precisely a commitment to weigh harm at γ_j in place of θ . The question is whether chartered weights reflect true ones.

Proposition 6 (Mission suppression) *Let both founders share the mission weight $\theta \geq 0$ and let $\lambda < \bar{\lambda}(1)$.*

- (i) *A blockaded monopolist's optimal charter is its true weight, $\gamma = \theta$ (or the cap 1 if $\theta > 1$).*
- (ii) *In the duopoly, pure profit charters $\gamma_1 = \gamma_2 = 0$ constitute an equilibrium (and then the unique equilibrium of the charter game) if and only if $\theta \leq \min \{\bar{\theta}(\lambda), \theta_x(\lambda)\}$,¹² where*

$$\bar{\theta}(\lambda) = \frac{\lambda^2 (\lambda^2 + 3\lambda + 4)}{\mathbb{E}(x) (4 - \lambda^2) \left[\frac{2 - \lambda^2 - 3\lambda}{\kappa''(0)} + 10 - 5\lambda^2 - 6\lambda \right]}, \quad \theta_x(\lambda) = \frac{2 - \lambda^2 - 3\lambda}{3\mathbb{E}(x) (4 - \lambda^2)}, \quad (20)$$

with $1/\kappa''(0) := 0$ in the absence of guardrails.

- (iii) *Without guardrails ($g \equiv 0$), let $\bar{\theta}(\lambda) \leq \theta_x(\lambda)$ and $\bar{\theta}(\lambda) < \theta \leq 1 + (1 - \mathbb{E}(x)) \bar{\theta}(\lambda)$ (so that $\gamma^* \leq 1$ below). Provided a unilateral switch to the pure profit charter leaves the market untipped, $\lambda \leq \bar{\lambda}(\delta(\gamma^*))$, the charter game has a unique symmetric equilibrium (its global unilateral optimality is verified numerically, see the proof): both firms adopt the interior mandate*

$$\gamma^*(\theta) = \frac{\theta - \bar{\theta}(\lambda)}{1 - \mathbb{E}(x) \bar{\theta}(\lambda)} \in (0, \theta), \quad \text{with} \quad \theta - \gamma^*(\theta) = \bar{\theta}(\lambda) \delta(\gamma^*). \quad (21)$$

- (iv) *The strategic threshold $\bar{\theta}$ is strictly increasing in substitutability and strictly decreasing in the effectiveness of guardrails, with $\bar{\theta} \rightarrow 0$ as $\kappa''(0) \rightarrow 0$; the participation*

¹² Necessity of $\theta \leq \bar{\theta}$ follows from the corner derivative, and necessity of $\theta \leq \theta_x$ from the option to secure zero by self-exclusion, which requires that a charter $\gamma \leq 1$ can exclude the firm, $\delta(1)^3 < 9\lambda(T^2 - 3\lambda^2)/(2T^3)$; at the benchmark $\mathbb{E}(x) = 1/2$ this holds throughout the range where θ_x binds, with or without guardrails. Sufficiency, and the uniqueness of the equilibrium, are verified numerically (see the proof).

bound θ_x is strictly decreasing in substitutability, independent of guardrails, and vanishes at $\bar{\lambda}(1)$.

Part (i) is the benchmark: sheltered from competition, mission is chartered truthfully. Given any behavior of the rival, the founder's value of a charter γ changes at the rate $(\gamma - \theta) H'(\gamma)$, which singles out $\gamma = \theta$. In the duopoly, by contrast, mission is suppressed below the strategic threshold $\bar{\theta}$: even genuinely mission-driven founders write pure profit charters, because the same strategic wedge as in Proposition 5 (the numerator of $\bar{\theta}$) must first be overcome before any of the mission is worth expressing. The denominator is the marginal harm reduction a small mandate buys, and it has two margins: the scale margin (the mandate shrinks the technology race) and the guardrail margin $(2 - \lambda^2 - 3\lambda) / \kappa''(0)$, which is larger the cheaper guardrails are at the margin. Effective guardrails thus lower the cost of expressing mission and thereby *lower* the strategic threshold. The participation bound θ_x is of a different nature: it is the mission weight at which the zero-charter duopoly's true value $\Pi_j^0 - (\theta - 0) H_j$ hits zero, and beyond it a mission-driven founder prefers exiting the market to operating without a mandate, so pure profit charters fail by exit rather than by marginal deviation. The two bounds cross at $\lambda \approx 0.44$ without guardrails (≈ 0.45 with the illustration nuisance function): the binding threshold is hump-shaped in substitutability, rising while the strategic bound binds and then falling to zero at $\bar{\lambda}(1)$, where even an unchartered duopoly can no longer carry a mission. Close to the tipping point, mission-driven founders either exit or operate as monopolists.

Above the threshold, mission finally surfaces in charters, but never in full. In the interior equilibrium (21), charters understate mission by exactly $\bar{\theta} \delta(\gamma^*)$: the same strategic wedge that suppresses small missions also discounts large ones, so chartered weights fall short of true weights by up to the strategic threshold itself. Two features of the solution stand out. First, charters *over*-respond at the margin: $d\gamma^*/d\theta = 1 / (1 - \mathbb{E}(x) \bar{\theta}) > 1$, so observed charters vary more than the underlying missions: competition suppresses weak missions entirely and amplifies differences among those that are expressed. Second, the equilibrium is self-limiting. A higher mission weight cuts the equilibrium discount factor, $\delta(\gamma^*) = (1 - \mathbb{E}(x) \theta) / (1 - \mathbb{E}(x) \bar{\theta})$, and once λ exceeds $\bar{\lambda}(\delta(\gamma^*))$ a defector to the pure profit charter can tip the market. The defection pays whenever the resulting monopoly is worth more to a mission-driven founder than the interior duopoly, which is the case unless substitutability is low (in our computations, for $\lambda \gtrsim 0.3$; at lower λ the monopoly's harm outweighs its profit and the interior equilibrium survives up to the charter cap). Where it pays, the symmetric equilibrium unravels exactly as asymmetric mandates unravel the duopoly in Section 2.2. Beyond that point (or beyond the crossing, where $\theta_x < \bar{\theta}$ and exit binds first) no symmetric charter profile survives: a high mission weight endogenously re-

creates the fragility that Figure 1 attaches to asymmetric mandates, and mission-driven founders again either exit or operate as monopolists.

Relative to Dewatripont and Tirole (2024), who show that product market competition alone does not erode ethical conduct, two features of our setting generate the erosion of mission. First, technology is an investment, its cost independent of output and sunk before prices: a mandate cannot be neutralized by repricing but shifts the firm’s position in the race, and the rival’s investment response amplifies the shift. The market-share and reduced-stakes effects that cancel exactly in their setting therefore do not cancel here, and the strategic cost of expressing mission rises with substitutability relative to the harm it saves, which is why $\bar{\theta}$ increases in λ . Second, the mandate is a commitment, so competition governs not only how a given weight on harm is acted upon but whether it is expressed at all: truthfully by a monopolist, suppressed or understated in the duopoly. The erosion operates through none of the channels by which competition can degrade ethics in their framework: prices are flexible throughout, consumers do not reward harm, firms with suppressed missions earn strictly positive profits for $\theta < \theta_x$ (suppression is a strategic choice, not a viability constraint), and no replacement logic is involved, since each firm internalizes only its own harm.

4 Robustness

The analysis so far relies on a linear demand system and a cubic investment cost. This section shows that the mechanism behind the mandate results of Section 3—delegation gains at fixed technology, overturned by the rival’s capability response whenever capability supply is sufficiently elastic—depends on neither. We recast the model in the class in which firms compete in utility space (Armstrong and Vickers, 2001): all instruments reach consumers through a one-dimensional quality-adjusted offer. Throughout the section we focus on the duopoly configuration, with both firms active before and after marginal deviations, the case in which the interesting forces operate; regime changes (blockade, exit) are inherently parametric and were characterized in Sections 2–3.

There are two firms. At date 0, firm j acquires capability (technology) $\tau_j \geq 0$ at strictly increasing, strictly convex cost $I(\tau_j)$. At date 1, firms set prices. Capability delivers quality $A(\tau_j)$ and per-unit expected harm $h(\tau_j)\mathbb{E}(x)$, with A and h strictly increasing, so that a charter γ_j acts as the as-if unit cost $\gamma_j\mathbb{E}(x)h(\tau_j)$, exactly as in Section 2 (we set $g \equiv 0$; optimally chosen guardrails are second order at the corner

studied below). Demand depends only on the two *quality-adjusted offers*,

$$z_j := A(\tau_j) - p_j, \quad q_j = D_j(z_j, z_{-j}),$$

with D_j smooth, strictly increasing in the own offer and strictly decreasing in the rival's. Firm j 's date-1 charter objective is $(p_j - \gamma_j \mathbb{E}(x) h(\tau_j)) D_j$, its financial value is $V_j = p_j D_j - I(\tau_j)$, and its behavior depends on charter and capability only through the *efficiency index*

$$e_j := A(\tau_j) - \gamma_j \mathbb{E}(x) h(\tau_j),$$

since the charter objective equals $(e_j - z_j) D_j(z_j, z_{-j})$ up to the sunk investment cost: capability raises efficiency, the charter taxes it.

The model of Section 2 is the member of this class with $A(\tau) = h(\tau) = \tau$, $I(\tau) = k\tau^3/3$, and the linear demand system generated by (1),

$$D_j(z_j, z_{-j}) = \frac{z_j - \lambda z_{-j}}{\beta(\lambda)(1 - \lambda^2)},$$

in which case the efficiency index is precisely the harm-discounted technology of Lemma 1, $e_j = (1 - \gamma_j \mathbb{E}(x)) \tau_j = \delta(\gamma_j) \tau_j$. Discrete-choice systems are equally nested: multinomial logit demand with mean utility $A(\tau_j) - p_j$ (Anderson et al., 1992) is in the class, as is any quality-augmented Hotelling specification. We maintain throughout: (i) the date-1 pricing game has a unique, stable equilibrium with both firms active at all relevant charters; (ii) the date-0 game has a symmetric equilibrium at $\gamma_1 = \gamma_2 = 0$, interior and regular; (iii) prices are strategic complements and the own-cost passthrough at date 1 lies in $(0, 1)$; (iv) $\partial^2 D_j / \partial z_j \partial z_{-j} \leq 0$, as in the linear system; (v) at $\gamma_1 = \gamma_2 = 0$, a marginal mandate lowers the marginal value of the firm's own capability and raises the rival's.¹³

Proposition 7 (Mandates and the rival's offer) *Let $\gamma_1 = \gamma_2 = 0$. Then*

$$\frac{dV_j}{d\gamma_j} = V_z \left[\frac{\partial z_{-j}}{\partial \gamma_j} \Big|_{\tau} + \frac{\partial z_{-j}}{\partial \tau_{-j}} \frac{d\tau_{-j}^*}{d\gamma_j} \right], \quad V_z := \frac{\partial V_j}{\partial z_{-j}} < 0, \quad (22)$$

all other effects vanishing by the envelope theorem. The first bracket term is the rival's price accommodation at fixed technologies: it is negative, and $G := V_z \partial z_{-j} / \partial \gamma_j|_{\tau} > 0$ is the delegation gain of Section 3.1. The second is the rival's capability response: under strategic substitutes at the investment stage it is positive, and $L := V_z (\partial z_{-j} / \partial \tau_{-j}) (d\tau_{-j}^ / d\gamma_j) < 0$*

¹³ Condition (v) is the reduced-form statement that a mandate shifts the capability race in the rival's favor; it holds in the linear system at every λ and is implied by (iii)–(iv) when the cross-curvature terms of demand are small.

0. *Pure profit charters are locally optimal if and only if a marginal mandate weakly improves the rival's quality-adjusted offer, that is, if and only if the capability response dominates the price accommodation.*

Proposition 7 is the general form of the comparison in Section 3.1: under delegation, a mandate makes the rival's offer *worse*, because prices accommodate; in the race, it makes the rival's offer *better*, because capability expands. Which effect dominates depends on an elasticity, which we now define without taking limits. At the symmetric equilibrium, let $F_o := \partial^2 \Pi_j / \partial \tau_j^2 < 0$ denote the equilibrium curvature of the investment problem (the only object in the entire comparison that contains I'') and define the *supply elasticity of capability*,

$$\varepsilon := \frac{I'(\tau)}{\tau |F_o|},$$

the percentage response of a firm's capability to a one-percent shift in its marginal value, holding the rival's actions fixed. The investment technology matters for the comparison of G and L only through ε : its level I' pins the equilibrium point, and its curvature enters nowhere but F_o . (In the linear model at weak overlap, $\varepsilon = 1/(n-2)$: the maintained cubic has unit-elastic capability supply.)

Proposition 8 (Elasticity threshold) *(i) Fix the demand system and the symmetric equilibrium point. Among investment technologies consistent with that point, $|L/G|$ is strictly increasing in ε , vanishing in the inelastic limit $\varepsilon \rightarrow 0$ (the fixed-technology world of Proposition 3, where delegation prevails) and diverging as ε approaches the stability bound of the investment game (where $|F_o|$ falls to the cross-derivative $|F_c|$ of the investment condition). There is therefore a unique threshold ε^* such that pure profit charters are locally optimal if and only if $\varepsilon \geq \varepsilon^*$.*

(ii) Along a family of demand systems in which the cross-sensitivities D_2 and D_{12} vanish at a common rate while the own derivatives converge to a monopoly benchmark $D(z)$ (the weak-overlap limit, the general counterpart of $\lambda \rightarrow 0$), the elasticity converges to its primitive form, $\varepsilon \rightarrow [\tau I''/I' - \tau \mathcal{M}'/\mathcal{M}]^{-1}$ with $\mathcal{M} = \mu D_1 A'$ the marginal value of capability, and

$$\left| \frac{L}{G} \right| \rightarrow \varphi_1 \varepsilon + \varphi_2 \varepsilon^2, \quad (23)$$

where $\varphi_1, \varphi_2 > 0$ are passthrough statistics of the benchmark, given explicitly in Appendix A.11, so that $\varepsilon^* = [\sqrt{\varphi_1^2 + 4\varphi_2} - \varphi_1]/(2\varphi_2)$ in closed form. In the linear model $\varphi_1 = 1$ and $\varphi_2 = 2$, so $\varepsilon^* = 1/2$: Proposition 4 is the special case $\varepsilon = 1/(n-2)$.

In general form, then, our main result takes the following form: a charter is a tax on the firm’s efficiency index, and competition eliminates it exactly when the rival responds to the tax by expanding capability instead of accommodating in price, which happens, at any strength of competition, when capability supply is sufficiently elastic. The linear term of (23) is the rival’s expansion against the mandate itself; the quadratic term is its expansion against the firm’s own induced retreat. Neither proposition uses the linear system: the arguments involve only the offer structure, envelope conditions, and implicit differentiation. Both parts are also verified computationally, part (i) exactly on the linear and logit systems at strong coupling and part (ii) on non-linear demand families with concave quality and convex harm maps.

5 Conclusion

Can responsible AI survive market competition? Several leading AI providers were founded with charters that rank a safety mission above profits, and whether such mandates can withstand an intense technology race is a central question for the governance and regulation of AI. This paper develops an industrial organization model to address this question, in which two AI providers invest in model capability, compete in prices, and choose guardrails that limit harmful use, and in which founders choose the corporate mandate under which their firm operates.

We draw three main conclusions. First, competition in the product market alone does not erode responsible conduct. A firm’s guardrail depends only on its own mandate, and because prices are strategic complements, a public-benefit mandate softens price competition and raises financial value, so that even profit-driven founders adopt one. Second, a technology race overturns this conclusion. A mandate acts like a tax on technology, so that a responsible firm invests less and its rival invests more; competition amplifies differences in mandates and, for sufficiently substitutable models, tips the market toward the laxer firm. When investment in capability is sufficiently elastic, as it is when capability is built from compute, data, and algorithms, profit-driven founders adopt pure profit charters and mission-driven founders understate or abandon their mission, whereas a monopolist charters its mission truthfully. Third, guardrails reduce the handicap of responsible AI without removing it: they compress the differences in mandate discount factors that drive tipping and lower the threshold below which mission-driven founders abandon their mission. Market forces alone therefore do not deliver responsible AI, and safety instruments may need to be mandated by regulation, as the EU AI Act does for frontier models. A mandate imposed uniformly on all firms slows the technology race

but leaves the duopoly viable, whereas the same mandate adopted unilaterally can tip the market.

Our framework is deliberately parsimonious, and several extensions are worth pursuing. The mechanism behind our results, delegation gains at fixed technology overturned by the rival’s capability response, holds in a general class of models in which firms compete in quality-adjusted offers, so that the analysis applies beyond AI to any industry in which capability scales with investment and the product can cause harm. Within AI, a welfare analysis of policy instruments such as guardrail mandates, liability for AI harms, and charter regulation is a natural next step, as is the study of founders with heterogeneous missions and of entry. Finally, since the technology race is global while regulation is national, the forces that set AI providers against one another may also set jurisdictions against one another; we leave the analysis of regulatory competition to future work.

References

- Daron Acemoglu and Todd Lensman. Regulating transformative technologies. *American Economic Review: Insights*, 6(3):359–376, 2024.
- Simon P. Anderson, André de Palma, and Jacques-François Thisse. *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge, MA, 1992.
- Anthropic. The long-term benefit trust. September 19, 2023. URL <https://www.anthropic.com/news/the-long-term-benefit-trust>.
- Anthropic. Anthropic’s long-term benefit trust appoints Vas Narasimhan to board of directors. April 14, 2026. URL <https://www.anthropic.com/news/narasimhan-board>.
- Mark Armstrong and John Vickers. Competitive price discrimination. *RAND Journal of Economics*, 32(4):579–605, 2001.
- Stuart Armstrong, Nick Bostrom, and Carl Shulman. Racing to the precipice: A model of artificial intelligence development. *AI & Society*, 31(2):201–206, 2016.
- Roland Bénabou and Jean Tirole. Individual and corporate social responsibility. *Economica*, 77(305):1–19, 2010.
- Michele Bisceglia, Alessio Piccolo, and Jan Schneemeier. Externalities of responsible investments. *Journal of Finance*, 2026. forthcoming.

- Philip Bond and Doron Levit. ESG: A panacea for market power? *Journal of Financial Economics*, 165:103991, 2025.
- California Department of Justice. Attorney general Bonta issues statement on OpenAI’s recapitalization plan. Press release, October 28, 2025. URL <https://oag.ca.gov/news/press-releases/attorney-general-bonta-issues-statement-openai%E2%80%99s-recapitalization-plan>.
- CNBC. Elon Musk’s xAI secretly dropped its benefit corporation status while fighting OpenAI. August 25, 2025. URL <https://www.cnbc.com/2025/08/25/elon-musk-xai-dropped-public-benefit-corp-status-while-fighting-openai.html>.
- Delaware Department of Justice. AG Jennings completes review of OpenAI recapitalization. Press release, October 28, 2025. URL <https://news.delaware.gov/2025/10/28/ag-jennings-completes-review-of-openai-recapitalization/>.
- Mathias Dewatripont and Jean Tirole. The morality of markets. *Journal of Political Economy*, 132(8):2655–2694, 2024.
- Avinash Dixit. A model of duopoly suggesting a theory of entry barriers. *Bell Journal of Economics*, 10(1):20–32, 1979.
- Epoch AI. Hyperscaler capex has quadrupled since GPT-4’s release. Data Insight (Isabel Juniewicz), February 26, 2026. URL <https://epoch.ai/data-insights/hyperscaler-capex-trend>.
- Chaim Fershtman and Kenneth L. Judd. Equilibrium incentives in oligopoly. *American Economic Review*, 77(5):927–940, 1987.
- Oliver Hart and Luigi Zingales. Companies should maximize shareholder welfare not market value. *Journal of Law, Finance, and Accounting*, 2(2):247–275, 2017.
- Charles I. Jones. The A.I. dilemma: Growth versus existential risk. *American Economic Review: Insights*, 6(4):575–590, 2024.
- Martin Oehmke and Marcus M. Opp. A theory of socially responsible investment. *Review of Economic Studies*, 92(2):1193–1225, 2025.
- OpenAI. OpenAI LP. March 11, 2019. URL <https://openai.com/index/openai-lp/>.

- OpenAI. Built to benefit everyone. October 28, 2025. URL <https://openai.com/index/built-to-benefit-everyone/>.
- Martin Shubik and Richard Levitan. *Market Structure and Behavior*. Harvard University Press, Cambridge, MA, 1980.
- Nirvikar Singh and Xavier Vives. Price and quantity competition in a differentiated duopoly. *RAND Journal of Economics*, 15(4):546–554, 1984.
- Steven D. Sklivas. The strategic choice of managerial incentives. *RAND Journal of Economics*, 18(3):452–458, 1987.
- TechCrunch. OpenAI files confidentially for IPO, following Anthropic. June 8, 2026. URL <https://techcrunch.com/2026/06/08/following-anthropic-openai-files-confidentially-for-ipo/>.
- John Vickers. Delegation and the theory of the firm. *Economic Journal*, 95(Supplement): 138–147, 1985.
- Xavier Vives. Innovation and competitive pressure. *Journal of Industrial Economics*, 56(3):419–469, 2008.

A Proofs

A.1 Segment demand

Fix a segment with intercepts (a_1, a_2) and prices $\mathbf{p}$, and let $z_j := a_j - p_j$. The Kuhn–Tucker conditions of (1) yield three demand configurations:

$$\begin{aligned} \text{both active } (z_1 > \lambda z_2, z_2 > \lambda z_1) : \quad & q_j = \frac{z_j - \lambda z_{-j}}{\beta(\lambda)(1 - \lambda^2)}, \\ \text{only } j \text{ active } (z_j > 0, z_{-j} \leq \lambda z_j) : \quad & q_j = \frac{z_j}{\beta(\lambda)}, \quad q_{-j} = 0, \\ \text{none active } (z_1 \leq 0, z_2 \leq 0) : \quad & q_1 = q_2 = 0. \end{aligned} \tag{24}$$

Demand is continuous, piecewise linear, and concave in the own price on the region where $q_j > 0$ (the own-price slope steepens from $1/\beta$ to $1/[\beta(1 - \lambda^2)]$ as the rival becomes active).

A.2 Proof of Lemma 1 (Guardrail choice and harm-discounted technology)

Proof: Since intercepts (2) do not depend on x , all segments face the same surplus $z_j := a_j - p_j = \tau_j(1 - \kappa(g_j)) - p_j$, so segment demands coincide, $q_j(x) \equiv q_j$ and $Q_j = q_j$, with $q_j = q_j(z_1, z_2)$ given by the demand blocks (24). Harm (5) aggregates to $H_j = \tau_j(1 - g_j)q_j \int_0^1 x f(x) dx = \tau_j(1 - g_j)\mathbb{E}(x)q_j$, so that, substituting $p_j = \tau_j(1 - \kappa(g_j)) - z_j$ and using (6),

$$\Pi_j^1 = (p_j - \gamma_j \tau_j(1 - g_j)\mathbb{E}(x))q_j = [(\tau_j - c_j(g_j)) - z_j]q_j(z_1, z_2).$$

In the strategy variables (z_j, g_j) , the own guardrail enters firm j 's payoff only through $-c_j(g_j)$ and enters neither demand nor the rival's payoff. Consider an equilibrium strategy of an active firm ($q_j > 0$) with g_j not a minimizer of $c_j(\cdot)$. The compensated deviation to (z_j, g_j^*) —lowering the price one-for-one with the willingness to pay destroyed by the guardrail, which leaves z_j and hence all quantities unchanged—strictly raises the margin. It is also feasible: an active firm's equilibrium margin is nonnegative (exclusion guarantees a payoff of zero), so $z_j \leq \tau_j - c_j(g_j)$ and the compensated price satisfies $p_j' = \tau_j(1 - \kappa(g_j^*)) - z_j \geq \gamma_j \tau_j(1 - g_j^*)\mathbb{E}(x) \geq 0$. Hence the guardrail of an active firm minimizes $c_j(\cdot)$. Since $c_j'(g) = \tau_j(\kappa'(g) - \gamma_j \mathbb{E}(x))$ is strictly increasing with $c_j'(0) = -\tau_j \gamma_j \mathbb{E}(x) \leq 0$ and $\lim_{g \rightarrow 1} c_j'(g) = \infty$, the minimizer g_j^* is unique, solves $\kappa'(g_j^*) = \gamma_j \mathbb{E}(x)$ (interior for $\gamma_j > 0$, $g_j^* = 0$ for $\gamma_j = 0$), and depends neither on the rival's strategy, nor on z_j (equivalently p_j), τ , or λ ; implicit differentiation gives $dg_j^*/d\gamma_j = \mathbb{E}(x)/\kappa''(g_j^*) > 0$. At $\mathbf{g}^*$, firm j chooses z_j (equivalently p_j) to maximize $[\tau_j - c_j(g_j^*) - z_j]q_j(z_1, z_2)$: a differentiated Bertrand duopoly in which the primitives (τ_j, γ_j) enter both firms' problems only through $\tau_j - c_j(g_j^*)$, given (λ, β) . ■

A.3 Proof of Lemma 2 (Date-1 equilibrium)

Proof: Within this proof only, abbreviate $w_j := \tau_j - c_j(g_j^*) = \delta(\gamma_j) \tau_j$ and, for a strategy (p_j, g_j) , let $z_j := \tau_j(1 - \kappa(g_j)) - p_j$ denote the surplus offered. By the proof of Lemma 1, the payoff is $[(\tau_j - c_j(g_j)) - z_j] q_j(z_1, z_2)$, and at $g_j = g_j^*$ the margin over internalized cost equals $w_j - z_j$.

Dominance. A strategy with $p_j < \gamma_j \tau_j(1 - g_j) \mathbb{E}(x)$, i.e., $z_j > \tau_j - c_j(g_j)$, yields $\Pi_j^1 \leq 0$, with equality if and only if $q_j = 0$, while (p_j, g_j) with $z_j = \tau_j - c_j(g_j)$ yields exactly 0 against any rival strategy; this establishes the weak dominance invoked in Problem 1. Every undominated strategy thus offers $z_j \leq \tau_j - c_j(g_j) \leq w_j$.

Existence. By Lemma 1, it suffices to construct an equilibrium of the reduced game in which firms play g_j^* and choose $z_j \in [0, w_j]$ with payoff $(w_j - z_j) q_j(z_1, z_2)$ (offers $z_j < 0$ are equivalent to $z_j = 0$). Given z_{-j} , own demand is continuous and piecewise linear in z_j , with slope falling from $1/[\beta(1 - \lambda^2)]$ to $1/\beta$ at the kink $z_j = z_{-j}/\lambda$ where the rival exits; the payoff is therefore continuous and quasi-concave in z_j (zero up to $z_j = \lambda z_{-j}$, concave beyond: at the kink the payoff derivative drops, since the margin is positive and the demand slope falls). Existence follows from Debreu–Glicksberg on $[0, w_1] \times [0, w_2]$. Any reduced-game equilibrium extends to $\mathcal{G}(\tau)$: a deviation (p'_j, g'_j) offering z'_j earns $[(\tau_j - c_j(g'_j)) - z'_j] q_j(z'_j, z_{-j}) \leq (w_j - z'_j) q_j(z'_j, z_{-j})$, which is no improvement.

Regime (ii): duopoly. The interior first-order conditions are $w_j - z_j = \beta(1 - \lambda^2) q_j$, i.e., $z_j - \lambda z_{-j} = w_j - z_j$. This is a linear system in (z_1, z_2) with dominant diagonal; its unique solution is the margin displayed in the lemma. Both margins are nonnegative iff $(2 - \lambda^2) w_2 \geq \lambda w_1$, i.e., $w_2/w_1 \geq \lambda/(2 - \lambda^2)$, and then both interiority conditions $z_j > \lambda z_{-j}$ hold. Payoffs follow from $\Pi_j^1 = (w_j - z_j) q_j$. No exclusionary equilibrium exists in this regime: excluding firm 2 requires $z_1 \geq w_2/\lambda$ (otherwise firm 2 enters profitably with an offer in $(\lambda z_1, w_2)$), and at $z_1 = w_2/\lambda$ the shared-branch derivative of firm 1's payoff is proportional to $w_1 - w_2(2 - \lambda^2)/\lambda \leq 0$, so firm 1 gains by lowering z_1 .

Regimes (i.a) and (i.b): entry deterrence. Suppose firm 2 plays an undominated posture offering $z_2 \leq w_2$, and let $w_2 < \lambda w_1/(2 - \lambda^2)$. On the exclusive branch $z_1 \geq z_2/\lambda$, firm 1's payoff $(w_1 - z_1) z_1/\beta$ is single-peaked at $z_1 = w_1/2$; on the shared branch, its derivative at the kink $z_1 = z_2/\lambda$ is proportional to $w_1 - z_2(2 - \lambda^2)/\lambda > 0$ for all $z_2 \leq w_2$. Firm 1's best response is therefore exclusionary, $z_1 = \max\{w_1/2, z_2/\lambda\}$, and firm 2 is inactive in any equilibrium of this regime, with $\Pi_2^1 = 0$.

(i.a) If $w_2/w_1 \leq \lambda/2$, then $z_2 \leq w_2 \leq \lambda w_1/2$: the blockade holds with slack at firm 1's unconstrained monopoly point $z_1 = w_1/2$, i.e., $p_1 = \gamma_1 \tau_1(1 - g_1^*) \mathbb{E}(x) + w_1/2$, whatever undominated posture (p_2, g_2) firm 2 takes; the latter is payoff-irrelevant.

(i.b) If $\lambda/2 < w_2/w_1 < \lambda/(2 - \lambda^2)$, we claim $z_2 = w_2$ in any equilibrium, which forces the cost-minimizing posture $g_2 = g_2^*$ and $p_2 = \gamma_2 \tau_2(1 - g_2^*) \mathbb{E}(x)$ (the unique undominated strategy attaining $z_2 = w_2$). Suppose instead $z_2 < w_2$. Firm 1's best response gives $\lambda z_1 = \max\{\lambda w_1/2, z_2\} < w_2$ (using $w_2 > \lambda w_1/2$), so firm 2 gains $(w_2 - z'_2)(z'_2 - \lambda z_1)/[\beta(1 - \lambda^2)] > 0$ by deviating to g_2^* and some $z'_2 \in (\lambda z_1, w_2)$, a contradiction. Given $z_2 = w_2$, firm 1's best response is the kink $z_1 = w_2/\lambda$ (its margin $w_1 - w_2/\lambda$ falls short of the monopoly margin $w_1/2$ precisely because $w_2 > \lambda w_1/2$), and firm 2 is indifferent: it is exactly at its entry margin with zero markup, and any deviation yields at most 0.

Global uniqueness of outcomes. The cases (i.a), (i.b), and (ii) partition $w_2/w_1 \in [0, 1]$, within each regime the equilibrium conditions pin down guardrails, margins, and quantities uniquely, and no other configuration survives: an equilibrium with only firm 2 active would require $z_1 \leq \lambda z_2$ at firm 1's most competitive undominated offer $z_1 = w_1$, i.e., $w_1 \leq \lambda z_2 \leq \lambda w_2$, contradicting the labeling $w_1 \geq w_2$ and $w_1 > 0$ (for $\lambda < 1$); an equilibrium with no firm active is ruled out since firm 1 can always secure a strictly positive payoff (at least its limit-pricing payoff). ■

A.4 Proof of Corollary 3 and derivation of (8)–(9)

Proof: *On-path regimes.* By Lemma 2, an excluded firm's operating payoff is zero, so any $\tau_j > 0$ at which firm j never sells yields $\Pi_j^0 = -k\tau_j^3/3 < 0$ and is strictly dominated by $\tau_j = 0$: no firm invests into the entry-deterrence region. Date-0 outcomes therefore have either $\tau_2 = 0$ (monopoly, in the blockaded subcase (i.a) of Lemma 2) or both firms active, proving the assertions preceding (8).

Marginal value of technology. Write $m_j := p_j - \gamma_j \tau_j (1 - g_j^*) \mathbb{E}(x)$ for the duopoly margin of Lemma 2, so that $\Pi_j^1 = m_j^2 / [(1 - \lambda^2)\beta]$ and, using (7),

$$\frac{\partial m_j}{\partial \tau_j} = \frac{(2 - \lambda^2) \delta(\gamma_j)}{4 - \lambda^2}.$$

Hence, using $q_j = m_j / [(1 - \lambda^2)\beta]$,

$$\frac{d\Pi_j^1}{d\tau_j} = 2q_j \frac{(2 - \lambda^2) \delta(\gamma_j)}{4 - \lambda^2} = \delta(\gamma_j) \left(1 - \frac{\lambda^2}{4 - \lambda^2}\right) q_j,$$

and equating to $k\tau_j^2$ gives (8). Substituting the margin formula turns (8) into the closed-form system

$$k\tau_j^2 = \frac{2(2 - \lambda^2) \delta(\gamma_j) [(2 - \lambda^2) \delta(\gamma_j) \tau_j - \lambda \delta(\gamma_{-j}) \tau_{-j}]}{(4 - \lambda^2)^2 (1 - \lambda^2) \beta(\lambda)}, \quad j = 1, 2;$$

dividing the two conditions and writing $\nu = \tau_2/\tau_1$ yields (9).

Stationary values. Substituting the first-order condition into $\Pi_j^0 = \Pi_j^1 - k\tau_j^3/3$ shows that at any stationary point in the duopoly regime,

$$\Pi_j^0 = \frac{m_j}{(4 - \lambda^2)(1 - \lambda^2)\beta(\lambda)} \left[\frac{2 - \lambda^2}{3} \delta(\gamma_j) \tau_j - \lambda \delta(\gamma_{-j}) \tau_{-j} \right] : \quad (25)$$

with cubic investment costs, a stationary firm covers its investment cost precisely if its harm-discounted technology is at least $3\lambda/(2 - \lambda^2)$ times the rival's.

Symmetric mandates. At $\gamma_1 = \gamma_2 = \gamma$, (9) factors as $(\nu - 1)(\lambda\nu^2 - (2 - \lambda^2 - \lambda)\nu + \lambda) = 0$. Subtracting the two closed-form conditions shows that either $\tau_1 = \tau_2$ or $\tau_1 + \tau_2$ is pinned at a constant, which reproduces the roots of the quadratic factor. Solving the symmetric condition gives τ_s as displayed (using $4 - \lambda^2 = (2 - \lambda)(2 + \lambda)$, $1 - \lambda^2 = (1 - \lambda)(1 + \lambda)$ and $2 - \lambda^2 - \lambda = (1 - \lambda)(2 + \lambda)$); substituting into $\Pi_j^0 = \Pi_j^1 - k\tau_s^3/3$ and using (8) at the symmet-

ric point delivers the stated value, which is proportional to $(2 - \lambda^2 - \lambda)(2 - \lambda^2 - 3\lambda)$ and hence positive precisely for $\lambda < \bar{\lambda}$, the specialization of (25) to the symmetric case. The second-order condition along (8) requires $k\tau_j > (2 - \lambda^2)^2 \delta(\gamma_j)^2 / [(4 - \lambda^2)^2 (1 - \lambda^2)\beta]$; since the first-order conditions imply $k\tau_s = 2(2 - \lambda^2)(2 - \lambda^2 - \lambda) \delta(\gamma)^2 / [(4 - \lambda^2)^2 (1 - \lambda^2)\beta]$, it holds if and only if $2 - \lambda^2 > 2\lambda$, which is implied by $\lambda < \bar{\lambda}$. The quadratic-factor roots are not equilibria: the quadratic is negative at $\nu = 3\lambda / (2 - \lambda^2)$ (setting $y := 3\lambda / (2 - \lambda^2) < 1$, its value there is proportional to $(y + 2)(y - 1) < 0$), so its smaller root lies below $3\lambda / (2 - \lambda^2)$ and the smaller firm's stationary value (25) is negative: exit is a profitable deviation (likewise for the mirror root with the labels reversed).

Monopoly for $\lambda \geq \bar{\lambda}$. Both stationary-value requirements $\Pi_j^0 \geq 0$ demand $3\lambda / (2 - \lambda^2) \leq \nu \leq (2 - \lambda^2) / (3\lambda)$, an empty interval for $\lambda > \bar{\lambda}$: no both-active equilibrium exists. The monopoly candidate $\tau_1 = \delta(\gamma)^2 / (2k\beta)$ (from $\partial[\delta(\gamma)^2 \tau_1^2 / (4\beta)] / \partial \tau_1 = k\tau_1^2$) is an equilibrium because entry is blockaded: firm 2's value is nonpositive on every branch of its entry problem against $w_1 := \delta(\gamma)\tau_1$. On the branches where the entrant would itself exclude the incumbent, the value is decreasing or negative throughout: the entrant's unconstrained monopoly peak lies inside its monopoly branch only if $\delta(\gamma_2)^3 / \delta(\gamma_1)^3 > 2/\lambda$, which is impossible, and the maximal value on the limit-pricing stretch, attained at $\tau^* = [\delta(\gamma_2)w_1 / (\lambda\beta k)]^{1/2}$, is nonpositive if and only if $\delta(\gamma_2)^3 / \delta(\gamma_1)^3 \leq 9/(8\lambda)$, which holds for all $\lambda < 9/8$. On the duopoly branch, the stationary points solve firm 2's closed-form condition above (a quadratic in τ_2 given τ_1), and by (25) entry is profitable only if its larger root reaches $\delta(\gamma_2)\tau_2 / w_1 \geq 3\lambda / (2 - \lambda^2)$. Evaluating the quadratic there shows that this fails whenever $9\lambda(4 - \lambda^2) \left[(2 - \lambda^2)^2 - \lambda^2 \right] \geq 8(2 - \lambda^2)^3$, which covers $\lambda \in [\bar{\lambda}, 0.924]$ (a nonnegative value of the upward quadratic at $3\lambda / (2 - \lambda^2)$ places both roots below that point because its vertex, at $\delta(\gamma_2)\tau_2 / w_1 = 9\lambda\eta^3 / [4(2 - \lambda^2)f_b(\lambda)]$ with $\eta = \delta(\gamma_2) / \delta(\gamma_1)$ and f_b the bound of Proposition 2, lies below $3\lambda / (2 - \lambda^2)$ whenever $\eta^3 \leq f_b(\lambda)$; and if the quadratic has no real root, the branch value is decreasing throughout and maximal at the lower end of the branch, where the entrant is excluded and earns $-k\tau_2^3 / 3 < 0$); for $\lambda \geq 0.917$, where $(2 - \lambda^2)(4 - \lambda^2) \leq 4\lambda$, the quadratic has no root inside the duopoly branch at all, so the branch value increases into its upper boundary and is bounded by the (nonpositive) limit-pricing value. Together the two conditions cover all $\lambda \in [\bar{\lambda}, 1)$. ■

A.5 Derivation of the collusive benchmarks (15) and (18)

Proof: In the symmetric pricing equilibrium at common technology τ , the margins of Lemma 2 give

$$p_j = \frac{\tau(1 - \lambda + \gamma\mathbb{E}(x))}{2 - \lambda}, \quad Q_j = \frac{\tau(1 - \gamma\mathbb{E}(x))}{(2 - \lambda)(1 + \lambda)\beta},$$

hence (15); the product $(1 - \gamma\mathbb{E}(x))(1 - \lambda + \gamma\mathbb{E}(x))$ is concave with maximizer $\gamma\mathbb{E}(x) = \lambda/2$. For the date-0 statement, substituting τ_s from Corollary 3 makes date-0 financial profits proportional to

$$(1 - \gamma\mathbb{E}(x))^5 \left[\frac{1 - \lambda + \gamma\mathbb{E}(x)}{(2 - \lambda)^2(1 + \lambda)} - \frac{2(2 - \lambda^2)(1 - \gamma\mathbb{E}(x))}{3(4 - \lambda^2)(2 - \lambda)(1 + \lambda)} \right],$$

whose first-order condition is linear in $\gamma \mathbb{E}(x)$, reducing to

$$4 - 5\lambda + 6\gamma \mathbb{E}(x) = \frac{4(2 - \lambda^2)}{2 + \lambda} (1 - \gamma \mathbb{E}(x))$$

and solving to (18); the bound $\gamma^* \mathbb{E}(x) < \lambda/2$ follows from $10 + 3\lambda - 2\lambda^2 > 6 + \lambda$. For the amplification statement, write $V(\gamma)$ for date-0 financial profit with technologies at their equilibrium level $\tau_s(\gamma)$. At $\gamma = 0$ the investment objective coincides with financial profit, so the investment first-order condition implies, by the envelope theorem,

$$V'(0) = \left. \frac{\partial (p_j Q_j)}{\partial \gamma} \right|_{\tau \text{ fixed}} + \frac{\partial (p_j Q_j)}{\partial \tau_{-j}} \tau'_s(0),$$

where the second term is strictly positive: the business-stealing spillover is negative and mandates reduce equilibrium investment, $\tau'_s < 0$. The date-0 gain therefore exceeds the fixed-technology gain locally at $\gamma = 0$; globally on $[0, \gamma^*]$ the ranking is verified numerically, with the date-0 gain 1.5–2 times the fixed-technology gain. ■

A.6 Proof of Propositions 1 and 2 and Corollaries 1–2

Proof: Within this proof only, abbreviate $T := 2 - \lambda^2$, $\eta := \delta(\gamma_2)/\delta(\gamma_1) \in (0, 1]$, and $\rho := \eta\nu$ (the relative harm-discounted technology), and let $A := 2T/[(4 - \lambda^2)^2(1 - \lambda^2)\beta(\lambda)]$, so that the closed-form system of Appendix A.4 reads $k\tau_j^2 = A\delta(\gamma_j)[T\delta(\gamma_j)\tau_j - \lambda\delta(\gamma_{-j})\tau_{-j}]$.

First-order-condition identities. Dividing each condition by τ_j ,

$$k\tau_1 = A\delta(\gamma_1)^2(T - \lambda\rho), \quad k\tau_2 = A\delta(\gamma_2)^2(T - \lambda/\rho). \quad (26)$$

Taking the ratio and using $\rho = \eta\nu$ eliminates the technologies: any interior stationary profile satisfies

$$\Phi(\rho) := \rho^2(T - \lambda\rho) - \eta^3(T\rho - \lambda) = 0, \quad (27)$$

which is (9) rewritten in the harm-discounted ratio.

Viability window. By (25), a both-active equilibrium requires both firms to cover their investment costs,

$$\frac{3\lambda}{T} \leq \rho \leq \frac{T}{3\lambda},$$

a nonempty window precisely for $\lambda \leq \bar{\lambda}$. By (26), the second-order conditions $k\tau_j \geq (T/2)A\delta(\gamma_j)^2$ read $T - \lambda\rho \geq T/2$ and $T - \lambda/\rho \geq T/2$, i.e., $2\lambda/T \leq \rho \leq T/(2\lambda)$: they hold strictly throughout the window.

Existence and uniqueness of the viable root. On the window, $\Phi''(\rho) = 2T - 6\lambda\rho \geq 0$: Φ is convex. At the endpoints,

$$\Phi\left(\frac{3\lambda}{T}\right) = 2\lambda\left(\frac{9\lambda(T^2 - 3\lambda^2)}{2T^3} - \eta^3\right), \quad \Phi\left(\frac{T}{3\lambda}\right) = \frac{2T^3 - 9\lambda\eta^3(T^2 - 3\lambda^2)}{27\lambda^2} > 0,$$

where positivity of the right endpoint follows from $\eta^3 \leq 1$ and $9\lambda(T^2 - 3\lambda^2) \leq 2T^3$ for $\lambda \leq \bar{\lambda}$. A convex function that is positive at the right endpoint has a root in

the window if and only if it is nonpositive at the left endpoint, and then exactly one; hence a viable stationary profile exists if and only if $\eta^3 \geq 9\lambda(T^2 - 3\lambda^2)/(2T^3)$, and it is unique. The threshold is strictly increasing in λ , since $d[\lambda(T^2 - 3\lambda^2)/T^3]/d\lambda = (T + 2\lambda^2)(T^2 - 9\lambda^2)/T^4 > 0$ for $\lambda < \bar{\lambda}$; it equals 0 at $\lambda = 0$ and 1 at $\lambda = \bar{\lambda}$. Hence, for fixed mandates, a viable profile exists if and only if $\lambda \leq \bar{\lambda}(\eta)$, where the cutoff $\bar{\lambda}(\eta)$ solves $9\lambda(T^2 - 3\lambda^2)/(2T^3) = \eta^3$: it is strictly increasing in η , with $\bar{\lambda}(1) = \bar{\lambda}$, which establishes (10). At the cutoff, the root is the left endpoint $\rho^* = 3\lambda/T$, where (25) gives firm 2 exactly zero: the responsible firm breaks even. At $\eta = 1$, $\Phi(1) = (T - \lambda)(1 - \eta^3) = 0$ delivers $\rho^* = 1$: technologies are equal.

The viable root is an equilibrium. At the root ρ^* , both second-order conditions hold strictly (above). On its duopoly branch, firm j 's value has derivative $2T\delta(\gamma_j)m_j/[(4 - \lambda^2)(1 - \lambda^2)\beta] - k\tau_j^2$, a downward parabola in τ_j (since m_j is linear in τ_j), so the value has a local minimum followed by a local maximum; the stationary point of the window is the latter, and it is the best choice on the branch because the value at the lower end of the branch, where firm j is just excluded, is $-k\tau_j^3/3 < 0 \leq$ its stationary value; the dead-zone branch is strictly negative. It remains to rule out deviations onto the exclusion branches. A deviating firm j would need $\tau_j \geq Tw_{-j}/[\lambda\delta(\gamma_j)]$ with $w_{-j} := \delta(\gamma_{-j})\tau_{-j}$; along the limit-pricing stretch the deviation value has slope $\delta(\gamma_j)w_{-j}/(\lambda\beta) - k\tau_j^2$, which is negative at the branch boundary and beyond whenever $k\beta T^2 w_{-j} \geq \lambda\delta(\gamma_j)^3$, and the monopoly stretch is then decreasing as well (its peak requirement is weaker by the factor $T^2/4 \leq 1$). For firm 2, using (26) and $\rho^* \leq 1$ (which follows from $\Phi(1) = (T - \lambda)(1 - \eta^3) \geq 0$), the condition reduces to $2T^3(T - \lambda) \geq \lambda(4 - \lambda^2)(T^2 - \lambda^2)$, which holds on $[0, \bar{\lambda}]$ since $(4 - \lambda^2)(T^2 - \lambda^2) = (2 - \lambda^2 - \lambda)(2 - \lambda^2 + \lambda)(4 - \lambda^2)$ and $2T^3 \geq \lambda(4 - \lambda^2)(T + \lambda)$ there; for firm 1, using $\rho^* \geq 3\lambda/T$ and $\eta^3 \geq 9\lambda(T^2 - 3\lambda^2)/(2T^3)$, it reduces to $6T(T^2 - 3\lambda^2) \geq (4 - \lambda^2)(T^2 - \lambda^2)$, i.e., $32 - 84\lambda^2 + 45\lambda^4 - 5\lambda^6 \geq 0$, which is decreasing in λ^2 and positive at $\bar{\lambda}^2$. Hence the viable root is the unique both-active equilibrium.

Amplification bounds. Since $\Phi \leq 0$ between the left endpoint and ρ^* , any point of the window at which $\Phi \geq 0$ lies weakly above ρ^* . Now

$$\Phi(\eta) = \eta^2 T (1 - \eta^2) \geq 0, \quad \Phi(\eta^3) = \lambda \eta^3 (1 - \eta^6) \geq 0,$$

and both η and η^3 lie in the window whenever the equilibrium exists (as $\eta^3 \geq 9\lambda(T^2 - 3\lambda^2)/(2T^3) > 3\lambda/T$ for $\lambda < \bar{\lambda}$). Hence $\rho^* \leq \eta^3 \leq \eta$ and $\tau_2/\tau_1 = \rho^*/\eta \leq \eta^2$, with equality if and only if $\lambda = 0$ or $\eta = 1$ (the displays are strict for $\lambda > 0$, $\eta < 1$). At $\lambda = 0$ the system (26) decouples into $\tau_j = \delta(\gamma_j)^2/(2k\beta)$, delivering the equalities.

Monotonicity. At the viable root, $\Phi'(\rho^*) > 0$ (convexity and the endpoint signs), so the implicit function theorem gives $d\rho^*/d\eta = 3\eta^2(T\rho - \lambda)/\Phi'(\rho^*) > 0$: the equilibrium ratio varies continuously and monotonically with the discount ratio.

Blockaded monopoly. The laxer firm's monopoly $\tau_1 = \delta(\gamma_1)^2/(2k\beta)$, $\tau_2 = 0$ is an equilibrium if and only if firm 2's entry is blockaded (τ_1 is firm 1's unique best response to $\tau_2 = 0$); by the argument in Appendix A.4, for $\lambda \leq 0.917$ this holds if and only if $\eta^3 \leq 9\lambda(4 - \lambda^2)(T^2 - \lambda^2)/(8T^3)$, while beyond 0.917 it holds outright (the entrant's stationary point exits the duopoly branch). The right side is strictly increasing for $\lambda \leq 0.76$, since its derivative is proportional to $(4 - \lambda^2)(8 - 14\lambda^2 + \lambda^4 - \lambda^6)$, and $8 - 14u + u^2 - u^3$

is strictly decreasing in $u = \lambda^2$ with root $u \approx 0.58$. It equals one at $\lambda \approx 0.449$ and exceeds one on $(0.449, 0.924)$, with maximum ≈ 1.47 at $\lambda \approx 0.76$. Hence blockade holds precisely for $\lambda \geq \underline{\lambda}(\eta)$, where $\underline{\lambda}(\eta)$ solves $9\lambda(4 - \lambda^2)(T^2 - \lambda^2)/(8T^3) = \eta^3$: it is strictly increasing in η with $\underline{\lambda}(1) \approx 0.449$, establishing (11). Moreover, $9\lambda(T^2 - 3\lambda^2)/(2T^3) \leq 9\lambda(4 - \lambda^2)(T^2 - \lambda^2)/(8T^3)$ reduces to $8 + \lambda^2 \geq T^2$, which always holds: the two thresholds are ordered, and hence $\underline{\lambda}(\eta) < \bar{\lambda}(\eta)$.

Responsible-firm monopoly. By relabeling, a blockaded monopoly of firm 2 is an equilibrium if and only if firm 1's entry is blockaded: for $\lambda \leq \lambda^*$ this reads $(1/\eta)^3 \leq 9\lambda(4 - \lambda^2)(T^2 - \lambda^2)/(8T^3)$, while beyond λ^* the binding requirement is that the entrant's exclusion branches be unprofitable, $(1/\eta)^3 \leq 9/(8\lambda)$; the switch point $\lambda^* \approx 0.765$ equates the two bounds. Unlike a symmetric entrant, whose stationary point stays interior up to $\lambda \approx 0.917$, a laxer entrant's leapfrogging deviation binds first. Since the bound's maximum is ≈ 1.47 , existence requires $1/\eta \lesssim 1.14$ and $\lambda \geq \underline{\lambda}(1/\eta)$; and since $\underline{\lambda}(\eta) \leq \underline{\lambda}(1) \leq \underline{\lambda}(1/\eta)$, the responsible firm's region is a strict subset of the laxer firm's. At high λ the region closes unless η is close to one: $9/(8\lambda)$ falls below $(1/\eta)^3$, and the laxer rival's exclusion-branch entry (leapfrogging) becomes profitable.

Corollary 1. For $\lambda < \underline{\lambda}(\eta)$ the both-active equilibrium exists and is the unique equilibrium: the laxer firm's monopoly fails by the blockade characterization, a monopoly of the responsible firm requires $\lambda \geq \underline{\lambda}(1/\eta) \geq \underline{\lambda}(\eta) > \lambda$, and the no-investment profile is not an equilibrium (firm 1's best response to $\tau_2 = 0$ is $\tau_1 > 0$). For $\underline{\lambda}(\eta) \leq \lambda \leq \bar{\lambda}(\eta)$ both characterizations hold: the duopoly and the blockaded monopoly coexist. For $\lambda > \bar{\lambda}(\eta)$ no both-active equilibrium exists, so every equilibrium is a blockaded monopoly; by the previous paragraph, the monopolist is the laxer firm unless mandates are nearly symmetric. Corollary 2 is the same partition read at fixed λ : since both cutoffs are strictly increasing, $\lambda \leq \bar{\lambda}(x)$ if and only if $x \geq \bar{\lambda}^{-1}(\lambda)$, and likewise for $\underline{\lambda}$; region (i) is nonempty precisely when $\underline{\lambda}^{-1}(\lambda) < 1$, that is, for $\lambda < \underline{\lambda}(1)$, where a responsible-firm monopoly is impossible and the three regions exhaust the equilibrium set. ■

A.7 Proof of Proposition 3 (Fixed technology)

Proof: Normalize $\beta = 1$ and $\tau = 1$ (all payoffs below scale by τ^2/β , which affects no best response), and write $\delta_j = 1 - \gamma_j \mathbb{E}(x) \in [1 - \mathbb{E}(x), 1]$ for the mandate discount factor without guardrails.

The both-active problem. By Lemma 2(ii), when both firms are active, firm j 's margin and quantity are $m_j = (T\delta_j - \lambda\delta_{-j})/(4 - \lambda^2)$ and $Q_j = m_j/(1 - \lambda^2)$ with $T := 2 - \lambda^2$, and since internalized harm involves no cash outflow, financial value is

$$(1 - \lambda^2) V_j = m_j^2 + (1 - \delta_j) m_j.$$

Writing $t := T/(4 - \lambda^2) \in (0, 1)$ for $dm_j/d\delta_j$, the second derivative is $2t(t - 1) < 0$: the both-active value is strictly concave in the own discount factor. The first-order condition $t(2m_j + 1 - \delta_j) = m_j$ rearranges to $\lambda^2 m_j = T(1 - \delta_j)$, and substituting m_j yields the

linear best response (16), with slope $\lambda^3/(4T) \in (0, 1)$ in δ and hence in γ . Moreover

$$1 - \delta(\gamma_j) = \frac{\lambda^2 (T - \lambda \delta(\gamma_{-j}))}{4T} \in \left(0, \frac{\lambda^2}{4}\right],$$

so every such best response carries a strictly positive mandate with $\gamma_j \mathbb{E}(x) \leq \lambda^2/4$ and satisfies $\delta(\gamma_j) \geq (4 - \lambda^2)/4$. (Differentiating at $\gamma_1 = \gamma_2 = 0$ and restoring the scales τ, β gives the display preceding the proposition.)

No equilibrium excludes a firm. Suppose firm j were inactive in a candidate equilibrium, earning 0. Deviating to $\gamma_j = 0$ makes it the (weakly) laxest firm, hence active by Lemma 2: if the continuation is both-active, its margin $(T - \lambda \delta_{-j}) / (4 - \lambda^2)$ is strictly positive; otherwise firm j blockades the rival and earns the monopoly or limit-pricing profit of Lemma 2(i), strictly positive since $\delta_{-j} < \lambda/T < \lambda$. The deviation is strictly profitable, a contradiction.

Unique equilibrium. In any equilibrium both firms are therefore active, and each charter must solve the rival-given best-response problem. On the both-active range the problem is strictly concave with interior optimum (16); the corner $\gamma_j = 0$ is excluded by the strictly positive derivative there; self-exclusion yields 0; and blockading the rival is infeasible, because it requires $\delta_{-j} < \lambda/T$ while the rival's own best response satisfies $\delta_{-j} \geq (4 - \lambda^2)/4 > \lambda/T$; the latter inequality is $8 - 4\lambda - 6\lambda^2 + \lambda^4 > 0$, which holds for $\lambda \leq 1/2$ since $4\lambda \leq 2$ and $6\lambda^2 \leq 3/2$. Hence both first-order conditions (16) hold; subtracting them gives $(\delta_1 - \delta_2) [1 + \lambda^3/(4T)] = 0$, so the equilibrium is symmetric, and solving the fixed point,

$$\delta^F = \frac{T(4 - \lambda^2)}{4T - \lambda^3}, \quad 1 - \delta^F = \frac{2\lambda^2 - \lambda^3 - \lambda^4}{8 - 4\lambda^2 - \lambda^3} = \frac{\lambda^2(1 - \lambda)(2 + \lambda)}{8 - 4\lambda^2 - \lambda^3},$$

which is (17). (The charter cap $\gamma_j \leq 1$ is slack: since $\gamma^F \mathbb{E}(x) \leq \lambda^2/4$, it suffices that $\mathbb{E}(x) \geq \lambda^2/4$, as assumed; the same holds for the interior mandates $\lambda/2$ and (18).)

Properties. $\gamma^F \mathbb{E}(x) < \lambda/2$ reduces to $8 - 4\lambda + 2\lambda^2 + \lambda^3 > 0$. Monotonicity on the stated range follows from the logarithmic derivative

$$\frac{d \ln [\gamma^F \mathbb{E}(x)]}{d\lambda} = \frac{2}{\lambda} - \frac{1}{1 - \lambda} + \frac{1}{2 + \lambda} + \frac{8\lambda + 3\lambda^2}{8 - 4\lambda^2 - \lambda^3} > \frac{2}{\lambda} - \frac{1}{1 - \lambda} > 0 \text{ for } \lambda < \frac{2}{3}.$$

Finally, $\gamma^F \mathbb{E}(x) = (\lambda^2/4)(2 - \lambda - \lambda^2) / (2 - \lambda^2 - \lambda^3/4) = \lambda^2/4 + O(\lambda^3)$. ■

A.8 Proof of Proposition 4

Proof: Normalize $k = \beta = 1$. With $g \equiv 0$, date-1 margins are $m_j = (T\delta_j\tau_j - \lambda\delta_{-j}\tau_{-j}) / (4 - \lambda^2)$ with $\delta_j = 1 - \gamma_j \mathbb{E}(x)$, and the date-0 charter objective is $\Pi_j = m_j^2 / (1 - \lambda^2) - \tau_j^n / n$, with investment condition

$$F_j := \frac{2m_j T \delta_j}{(4 - \lambda^2)(1 - \lambda^2)} - \tau_j^{n-1} = 0.$$

At $\gamma_1 = \gamma_2 = 0$ the charter objective coincides with financial value, so the own-investment response drops out by the envelope theorem and $dV_1/d\gamma_1 = G + L$, with $L = (\partial V_1 / \partial \tau_2) (d\tau_2^* / d\gamma_1)$

and $\partial V_1/\partial \tau_2 = -2\lambda m_1/[(4 - \lambda^2)(1 - \lambda^2)] < 0$. Total differentiation of $F_1 = F_2 = 0$ at the symmetric point gives

$$\frac{d\tau_2^*}{d\gamma_1} = \frac{F_c F_{1\gamma} - F_o F_{2\gamma}}{F_o^2 - F_c^2},$$

where, writing $P := 2T/[(4 - \lambda^2)(1 - \lambda^2)]$ and evaluating at the symmetric corner equilibrium (τ, τ) ,

$$F_o = \frac{PT}{4 - \lambda^2} - (n - 1)\tau^{n-2}, \quad F_c = -\frac{P\lambda}{4 - \lambda^2}, \quad F_{1\gamma} = -P\mathbb{E}(x) \left(\frac{T\tau}{4 - \lambda^2} + m \right), \quad F_{2\gamma} = \frac{P\lambda\mathbb{E}(x)\tau}{4 - \lambda^2}.$$

Here $F_c < 0$ is strategic substitutability, $F_o < 0$ is the second-order condition for the investment equilibrium, and $F_{1\gamma} < 0 < F_{2\gamma}$; hence $d\tau_2^*/d\gamma_1 > 0$ and $L < 0$. (The sign uses stability, $F_o^2 > F_c^2$, which by the symmetric first-order condition holds if and only if $n > 2T/(T - \lambda)$; at $n = 3$ this is exactly $\lambda < \bar{\lambda}(1)$, and for $n < 3$ it fails before $\bar{\lambda}(1)$, at the root of $2 - \lambda^2 = n\lambda/(n - 2)$.) As $\lambda \rightarrow 0$, the symmetric first-order condition gives $\tau^{n-2} \rightarrow 1/2$ and $m \rightarrow \tau/2$, so $F_o \rightarrow -(n - 2)/2$, $F_c \rightarrow -\lambda/4$, $F_{1\gamma} \rightarrow -\mathbb{E}(x)\tau$, $F_{2\gamma} \rightarrow \lambda\mathbb{E}(x)\tau/4$, and therefore

$$\frac{d\tau_2^*}{d\gamma_1} \rightarrow \frac{\lambda\mathbb{E}(x)\tau n}{2(n - 2)^2}, \quad L \rightarrow -\frac{\lambda^2\mathbb{E}(x)\tau^2 n}{8(n - 2)^2}, \quad G \rightarrow \frac{\lambda^2\mathbb{E}(x)\tau^2}{8},$$

which is (19); $G + L < 0$ if and only if $n/(n - 2)^2 > 1$, i.e. $(n - 1)(n - 4) < 0$. (With a concave capability-quality map $a = \tau^\alpha$, the same computation yields the limit $-\alpha(n + 1 - \alpha)/(n - 2\alpha)^2$. The monotonicity of $|L/G|$ in λ , the widening boundary, and the second-order failure for n close to 2 at high λ are verified numerically against the exact implicit system, which also confirms the decomposition by finite differences on the full two-stage equilibrium and replicates the pattern in a logit demand system with an outside good.) ■

A.9 Proof of Proposition 5

Proof: Normalize $k = \beta = 1$ (both enter all date-0 values only as positive multiplicative scales) and write $\delta_j := \delta(\gamma_j) \in (0, 1]$ for the mandate discount factor (with optimally chosen guardrails or, as a special case, with $g \equiv 0$) and $a_j := \gamma_j(1 - g_j^*)\mathbb{E}(x) \geq 0$ for the cash add-back rate, so that $1 - \delta_j = a_j + \kappa(g_j^*)$ and, without guardrails, $a_j = 1 - \delta_j = \gamma_j\mathbb{E}(x)$. Two facts about the primitives are used below. First, $\delta'_j(\gamma_j) = -(1 - g_j^*)\mathbb{E}(x) < 0$ by the envelope theorem. Second, $a'_j \leq -\delta'_j$: the gap is $d\kappa(g_j^*)/d\gamma_j \geq 0$, because g^* is nondecreasing in γ (for $\gamma_2 > \gamma_1$, adding the two revealed-preference inequalities from the guardrail problem gives $(\gamma_2 - \gamma_1)(g_2^* - g_1^*)\mathbb{E}(x) \geq 0$) and κ is increasing. We show that firm j 's financial value is strictly increasing in δ_j on the both-active region, strictly decreasing in γ_j under a blockaded monopoly, and jumps upward at every regime boundary crossed as γ_j falls; together these imply that the unique best response to any rival charter is $\gamma_j = 0$.

Values on the equilibrium path. In the both-active equilibrium, let $\rho := \delta_{-j}\tau_{-j}/(\delta_j\tau_j)$ denote the rival-to-own ratio of harm-discounted technologies, and let $s := T - \lambda\rho$ with

$T = 2 - \lambda^2$. Firm j 's first-order condition (8) pins its margin at $m_j = \tau_j^2 (4 - \lambda^2) (1 - \lambda^2) / (2T\delta_j)$, and the proof of Proposition 1 gives $\tau_j = A\delta_j^2 s$ with $A = 2T / [(4 - \lambda^2)^2 (1 - \lambda^2)]$ and the window $\rho \in [3\lambda/T, T/(3\lambda)]$. Substituting both into financial value $V_j = (m_j + a_j \tau_j) Q_j - \tau_j^3/3$, with $Q_j = m_j / (1 - \lambda^2)$, yields the exact decomposition

$$V_j = \delta_j^6 \hat{\Pi}(\rho) + a_j \delta_j^5 \hat{G}(\rho), \quad \hat{\Pi} = \frac{Cs^3(T - 3\lambda\rho)}{3}, \quad \hat{G} = C(4 - \lambda^2)s^3,$$

with $C := A^2 / [(4 - \lambda^2)^2 (1 - \lambda^2)] > 0$: the rival influences firm j 's value only through ρ . On the window, $s > 0$, $\hat{G} > 0$, and $\hat{\Pi} \geq 0$, with equality exactly at the tipping edge $\rho = T/(3\lambda)$.

The rival's response. The ratio ρ solves $\Phi(\rho) = \rho^2(T - \lambda\rho) - \eta^3(T\rho - \lambda) = 0$ with $\eta = \delta_{-j}/\delta_j$, so that $\eta^3(T\rho - \lambda) = \rho^2 s$ on the path. At the unique root, $\Phi'(\rho) = \rho^2 \Delta / (T\rho - \lambda)$ with

$$\Delta := T^2 + 3\lambda^2 - 2T\lambda(\rho + \rho^{-1}) \geq \frac{(T - 3\lambda)(T + 3\lambda)}{3} > 0:$$

the sum $\rho + \rho^{-1}$ is maximal at the reciprocal window endpoints, where the bound follows, and $T > 3\lambda$ is equivalent to $\lambda < \bar{\lambda}(1)$; this is where that hypothesis enters. (The root is unique in the window also when $\eta > 1$, i.e., when firm j is the responsible firm: Φ and the window are invariant under relabeling, and the endpoint signs follow from the existence condition written as $\eta^3 \leq 2T^3/[9\lambda(T^2 - 3\lambda^2)]$.) Since η falls as δ_j rises, total differentiation of Φ gives $\delta_j d\rho/d\delta_j = -3\rho^2 s / \Phi'(\rho) = -3s(T\rho - \lambda) / \Delta =: -K < 0$: cutting the own mandate strengthens the own technology position.

Monotonicity on the both-active region. Differentiating V_j along the equilibrium path and using $a'_j \leq -\delta'_j$ together with $\hat{G} > 0$ and $\delta'_j < 0$,

$$\begin{aligned} \frac{dV_j}{d\gamma_j} &= \delta'_j \left[\delta_j^5 (6\hat{\Pi} - K\hat{\Pi}') + a_j \delta_j^4 (5\hat{G} - K\hat{G}') \right] + a'_j \delta_j^5 \hat{G} \\ &\leq -|\delta'_j| \delta_j^4 \left[\delta_j (6\hat{\Pi} - \hat{G} - K\hat{\Pi}') + a_j (5\hat{G} - K\hat{G}') \right], \end{aligned}$$

so it suffices that both coefficients are positive. The second is immediate: $\hat{G}' = -3C\lambda(4 - \lambda^2)s^2 < 0$ and $K > 0$, so $5\hat{G} - K\hat{G}' > 0$. The first factors as

$$6\hat{\Pi} - \hat{G} - K\hat{\Pi}' = \frac{Cs^3\lambda^2(\lambda + 2\rho)[2T - \lambda\rho(1 + \lambda^2)]}{\rho\Delta} > 0,$$

strictly positive on the window because $\lambda\rho \leq T/3$ implies $\lambda\rho(1 + \lambda^2) < 2T$. Hence $dV_j/d\gamma_j < 0$ throughout the both-active region: financial value strictly falls in the own mandate, with or without guardrails, whatever the rival's charter.

Blockaded monopoly and regime changes. A blockaded monopolist's financial value is $V^m(\gamma) = \delta^6/48 + a\delta^5/8$, and the same bound gives $dV^m/d\gamma \leq \delta'(\delta^5/8 + 5a\delta^4/8) + |\delta'| \delta^5/8 = -\frac{5}{8}|\delta'|a\delta^4 < 0$ for $\gamma > 0$ (without guardrails, $dV^m/d\gamma = -\frac{5}{8}\gamma\mathbb{E}(x)^2\delta^4$): a monopolist, too, strictly prefers the zero mandate. Finally, lowering γ_j can only shift the selected continuation in firm j 's favor: from the rival's blockaded monopoly (value 0) into

the both-active region (value ≥ 0), or from the both-active region into firm j 's blockaded monopoly, whose value exceeds the adjacent duopoly value because a rival with positive technology strictly reduces firm j 's attainable cash flow at every own technology. Hence V_j is nonincreasing in γ_j globally, and strictly decreasing wherever firm j is active, so the unique best response to any rival charter is $\gamma_j = 0$ and $(0, 0)$ is the unique equilibrium. The prisoner's-dilemma statement follows from the collusive benchmark (18) of Appendix A.5: the common financial value at $\gamma^* > 0$ strictly exceeds its level at $\gamma = 0$. (With guardrails, the marginal value of a small common mandate at $\gamma = 0$ coincides with its $g \equiv 0$ counterpart, since $g^*(0) = 0$ and $\kappa(g^*)$ is of second order, so the common gain remains strictly positive.) ■

A.10 Proof of Proposition 6

Proof: Throughout, we use the objects of Appendix A.9 and the identity $\Pi_j^0 = V_j - \gamma_j H_j$, which holds because the mandate objective deducts $\gamma_j (1 - g_j) \mathbb{E}(x) \tau_j Q_j = \gamma_j H_j$ relative to cash.

(i) *Monopoly.* Given any rival behavior, the firm's conduct under charter γ maximizes $V - \gamma H$; write $M(\gamma)$ for the maximized value. The founder receives $W(\gamma) = M(\gamma) + (\gamma - \theta) H(\gamma)$, and the envelope theorem gives $M' = -H$, so $W'(\gamma) = (\gamma - \theta) H'(\gamma)$. For the blockaded monopolist, $H(\gamma) = (1 - g^*) \mathbb{E}(x) \delta^5/8$ is strictly decreasing (each factor is), so W increases below θ and decreases above: the unique optimum is $\gamma = \theta$.

(ii) *The two bounds.* On the both-active path, $Q_j = m_j/(1 - \lambda^2)$ and the first-order condition give firm j 's harm the compact form $H_j = (1 - g_j^*) \mathbb{E}(x) \delta_j^5 \hat{G}(\rho)$. At $\gamma_1 = \gamma_2 = 0$ (where $\rho = 1$, $\delta_j = 1$, $a_j = 0$, $a'_j = \mathbb{E}(x)$, $\delta'_j = -\mathbb{E}(x)$, and $\kappa(g^*)$ is of second order), Appendix A.9 gives $dV_j/d\gamma_j = -\mathbb{E}(x) [6\hat{\Pi} - \hat{G} - K\hat{\Pi}']$, and differentiating H_j ,

$$\frac{dH_j}{d\gamma_j} = -\mathbb{E}(x) \left[g'_0 \hat{G} + \mathbb{E}(x) (5\hat{G} - K\hat{G}') \right], \quad g'_0 := \frac{dg^*}{d\gamma} \Big|_0 = \frac{\mathbb{E}(x)}{\kappa''(0)},$$

where g'_0 follows from the guardrail first-order condition $\kappa'(g^*) = \gamma \mathbb{E}(x)$. Hence $dU_j/d\gamma_j \leq 0$ at the corner if and only if $\theta \leq \bar{\theta} := [6\hat{\Pi} - \hat{G} - K\hat{\Pi}'] / [g'_0 \hat{G} + \mathbb{E}(x) (5\hat{G} - K\hat{G}')]]$, all objects evaluated at $\rho = 1$. Substituting the factorization of Appendix A.9 at $\rho = 1$, where $s = T - \lambda = (1 - \lambda)(2 + \lambda)$, $\Delta = (T - \lambda)(T - 3\lambda)$, $K = 3(T - \lambda)/(T - 3\lambda)$, and $2T - \lambda(1 + \lambda^2) = (1 - \lambda)(\lambda^2 + 3\lambda + 4)$, yields the closed form in (20). For the participation bound, $U_j(0, 0) = \hat{\Pi}(1) - \theta \mathbb{E}(x) \hat{G}(1)$, and $\hat{\Pi}(1)/[\mathbb{E}(x) \hat{G}(1)] = (T - 3\lambda)/[3\mathbb{E}(x)(4 - \lambda^2)] = \theta_x$; a firm can secure exactly zero by chartering a mandate high enough to be excluded (the rival then serves the market as the laxer blockaded monopolist), which the cap $\gamma \leq 1$ permits whenever $\delta(1)^3 < 9\lambda(T^2 - 3\lambda^2)/(2T^3)$, i.e., throughout the range where θ_x binds at the benchmark $\mathbb{E}(x) = 1/2$, with or without the illustration nuisance function; so $\theta \leq \theta_x$ is necessary, and $\theta \leq \bar{\theta}$ is necessary by the corner derivative. For sufficiency, note the deviation gain $U_j(\gamma, 0) - U_j(0, 0)$ is strictly increasing in θ (its θ -derivative is $H_j(0, 0) - H_j(\gamma, 0) > 0$, since $H_j = (1 - g_j^*) \mathbb{E}(x) \delta_j^5 \hat{G}$ falls in the own mandate factor by factor), so it suffices to rule out profitable deviations at $\theta = \min\{\bar{\theta}, \theta_x\}$; we verify this numerically on the window.

Uniqueness. The same verification shows that, on a fine grid of rival charters, the best response to every $\gamma_{-j} > 0$ lies strictly below γ_{-j} (and the best response to 0 is 0), so the best-response map has no positive fixed point: an equilibrium with $\gamma_1 = BR(\gamma_2) > 0$ and $\gamma_2 = BR(\gamma_1) > 0$ would require $\gamma_1 < \gamma_2$ and $\gamma_2 < \gamma_1$. Profiles in which one firm is excluded are ruled out analytically: the excluded firm earns 0, while deviating to $\gamma = 0$ either preserves the both-active continuation—in which case, as the laxest firm, its position satisfies $\rho \leq 1$ (since $\Phi(1) = (T - \lambda)(1 - \eta^3) \geq 0$ and $\Phi' > 0$ at the root) and its payoff is $\hat{\Pi}(\rho) - \theta \mathbb{E}(x) \hat{G}(\rho) = \mathbb{E}(x) \hat{G}(\rho) [\theta_x(\rho) - \theta]$ with $\theta_x(\rho) := (T - 3\lambda\rho) / [3\mathbb{E}(x)(4 - \lambda^2)] \geq \theta_x$, strictly positive for $\theta < \theta_x$ —or tips the market in its favor, yielding the monopoly payoff $1/48 - \theta \mathbb{E}(x) / 8$, positive on the same range because $\theta_x \leq 1/[6\mathbb{E}(x)]$. Hence $(0, 0)$ is the unique equilibrium whenever it is one.

(iii) *The interior symmetric equilibrium.* Without guardrails, $\delta_j = 1 - \gamma_j \mathbb{E}(x)$, $a_j = 1 - \delta_j$, $a'_j = -\delta'_j = \mathbb{E}(x)$, and $H_j = \mathbb{E}(x) \delta_j^5 \hat{G}(\rho)$. At *any* symmetric profile $\gamma_1 = \gamma_2 = \gamma$ we have $\rho = 1$ regardless of the common discount $\delta = 1 - \gamma \mathbb{E}(x)$, so $\hat{\Pi}, \hat{G}, K$ take their $\rho = 1$ values, which are functions of λ alone. Writing $E_1 := 6\hat{\Pi} - \hat{G} - K\hat{\Pi}'$ and $E_0 := 5\hat{G} - K\hat{G}'$ at $\rho = 1$ (so that $\bar{\theta} = E_1 / [\mathbb{E}(x) E_0]$ by part (ii)), the derivatives of part (ii) generalize from $\delta = 1$ to any common δ :

$$\frac{dV_j}{d\gamma_j} = -\mathbb{E}(x) \delta^4 [\delta E_1 + (1 - \delta) E_0], \quad \frac{dH_j}{d\gamma_j} = -\mathbb{E}(x)^2 \delta^4 E_0.$$

Hence $dU_j/d\gamma_j = -\mathbb{E}(x)^2 \delta^4 E_0 [(1 - \gamma \mathbb{E}(x)) \bar{\theta} + \gamma - \theta]$: the first-order condition is *linear* in γ ,

$$(1 - \gamma \mathbb{E}(x)) \bar{\theta} + \gamma = \theta \quad \Longleftrightarrow \quad \gamma = \gamma^*(\theta) = \frac{\theta - \bar{\theta}}{1 - \mathbb{E}(x) \bar{\theta}},$$

and since the bracket increases in γ , the diagonal derivative is positive below γ^* and negative above: no symmetric profile other than (γ^*, γ^*) , in particular neither corner, can be an equilibrium, and γ^* is interior exactly for $\bar{\theta} < \theta \leq 1 + (1 - \mathbb{E}(x)) \bar{\theta}$. Direct substitution gives

$$\delta(\gamma^*) = \frac{1 - \mathbb{E}(x) \theta}{1 - \mathbb{E}(x) \bar{\theta}}, \quad \theta - \gamma^* = \bar{\theta} \delta(\gamma^*).$$

Three requirements make (γ^*, γ^*) an equilibrium. *No tipping:* a deviation to $\gamma = 0$ faces the ratio $\eta = \delta(\gamma^*)$, and the both-active continuation survives it if and only if $\delta(\gamma^*)^3 \geq 9\lambda(T^2 - 3\lambda^2)/(2T^3)$, i.e. $\lambda \leq \bar{\lambda}(\delta(\gamma^*))$. Writing $\bar{\eta}(\lambda)$ for the cube root of the boundary value, this reads

$$\theta \leq \theta_K := \bar{\theta} + \frac{(1 - \mathbb{E}(x) \bar{\theta})(1 - \bar{\eta}(\lambda))}{\mathbb{E}(x)};$$

when $\bar{\eta}(\lambda) \geq 1 - \mathbb{E}(x)$ (at $\mathbb{E}(x) = 1/2$, for $\lambda \geq 0.06$), this bound also enforces the charter cap $\gamma^* \leq 1$. *Participation:* substituting $\delta(\gamma^*)$ into $U_j = \delta^6 \hat{\Pi}(1) + (1 - \delta) \delta^5 \hat{G}(1) - \theta \mathbb{E}(x) \delta^5 \hat{G}(1)$ and using $\hat{\Pi}(1) = \theta_x \mathbb{E}(x) \hat{G}(1)$ together with $(1 - \delta) / \mathbb{E}(x) = \gamma^* = \theta - \bar{\theta} \delta$,

the payoff collapses to

$$U_j(\gamma^*, \gamma^*) = \mathbb{E}(x) \hat{G}(1) \delta(\gamma^*)^6 (\theta_x - \bar{\theta}),$$

nonnegative precisely on the stated region $\bar{\theta} \leq \theta_x$, independently of θ : as mission deepens, the larger harm deduction is exactly offset by the de-escalated technology race. In particular each firm prefers (γ^*, γ^*) to securing 0 by self-exclusion. *Global unilateral optimality*: the first-order condition is necessary only; that γ^* is the global best response to the rival at γ^* (across regime changes, including the tipping deviation) we verify numerically throughout the validity region. Beyond θ_K the pure profit deviation tips the market, but it pays only if the defector's monopoly value $1/48 - \theta \mathbb{E}(x)/8$ exceeds $U_j(\gamma^*, \gamma^*)$; where it does (for $\lambda \gtrsim 0.3$ at the benchmark), the best response drops toward zero just beyond θ_K , whereas for lower λ the interior equilibrium survives up to the charter cap.

(iv) *Comparative statics*. In (20), the numerator of $\bar{\theta}$ is strictly increasing in λ while both bracketed denominator terms are positive and strictly decreasing, so $\bar{\theta}$ strictly increases in λ ; $1/\kappa''(0)$ multiplies the positive term $2 - \lambda^2 - 3\lambda$, so $\bar{\theta}$ strictly decreases in guardrail effectiveness and vanishes as $\kappa''(0) \rightarrow 0$. For θ_x , the derivative of $(2 - \lambda^2 - 3\lambda)/(4 - \lambda^2)$ is proportional to $-(3\lambda^2 + 4\lambda + 12) < 0$, and $2 - \lambda^2 - 3\lambda = 0$ exactly at $\lambda = \bar{\lambda}(1)$. ■

A.11 Proofs for Section 4

Proof of Proposition 7: Write $\hat{V}_j(\gamma_j; \tau_j, \tau_{-j})$ for firm j 's financial value at the date-1 pricing equilibrium, and note three facts. First, γ_j does not enter cash directly: internalized harm involves no cash outflow, so the charter matters only through behavior. Second, at $\gamma_1 = \gamma_2 = 0$ the charter objective coincides with cash, so firm j 's own price and own capability each satisfy the cash first-order conditions; since the date-0 objective is the date-1 continuation value, the capability condition already internalizes the effect of τ_j on the rival's date-1 offer, and by the envelope theorem the induced movements of p_j and τ_j have no first-order effect on V_j . Third, $V_j = p_j D_j(z_j, z_{-j}) - I(\tau_j)$ depends on the rival's actions only through the rival's offer z_{-j} . Hence

$$\frac{dV_j}{d\gamma_j} = \left. \frac{\partial \hat{V}_j}{\partial \gamma_j} \right|_{\tau} + \frac{\partial \hat{V}_j}{\partial \tau_{-j}} \frac{d\tau_{-j}^*}{d\gamma_j},$$

and each term reduces to a movement of z_{-j} weighted by $V_z = p_j \partial D_j / \partial z_{-j} < 0$: at fixed technologies the only payoff-relevant movement is the rival's price, $\partial z_{-j} / \partial \gamma_j|_{\tau} = -\partial p_{-j} / \partial \gamma_j|_{\tau}$ (the date-1 equilibrium response, including the rival's reaction to the own price change), which is negative when prices are strategic complements (my as-if cost rises, my price rises, the rival's follows); and $\partial \hat{V}_j / \partial \tau_{-j} = V_z \partial z_{-j} / \partial \tau_{-j}$ with $\partial z_{-j} / \partial \tau_{-j} > 0$ (quality net of its price passthrough). This is (22); the sign statement follows from $V_z < 0$. ■

Proof of Proposition 8: *Part (i)*. Totally differentiating the two investment condi-

tions $F_j := \partial \Pi_j / \partial \tau_j = 0$ at the symmetric corner equilibrium gives the rival's response

$$\frac{d\tau_{-j}^*}{d\gamma_j} = \frac{F_c F_{1\gamma} - F_o F_{2\gamma}}{F_o^2 - F_c^2},$$

where $F_o < 0$ is the own curvature, with $|F_o| = I' / (\varepsilon \tau)$ by the definition of ε , while $F_c = \partial^2 \Pi_j / \partial \tau_j \partial \tau_{-j} < 0$ (strategic substitutes) and $F_{1\gamma} < 0 < F_{2\gamma}$ (maintained condition (v)) are demand-side objects containing no I'' . Writing $x := |F_o|$, $a := F_c F_{1\gamma} > 0$ and $b := F_{2\gamma} > 0$, the response equals $(a + bx) / (x^2 - F_c^2)$ on the stable range $x > |F_c|$, with derivative

$$-\frac{bx^2 + 2ax + bF_c^2}{(x^2 - F_c^2)^2} < 0 :$$

the response rises strictly as x falls (that is, as ε rises) from 0 at $x = \infty$ and without bound as $x \downarrow |F_c|$. Since neither G nor the offer weights in (22) contain I'' , $|L/G|$ inherits exactly these properties and crosses 1 at a unique ε^* .

Part (ii). Let subscripts on D denote partial derivatives with respect to the own and rival offer, in that order. At the monopoly benchmark $D(z)$ with margin μ and pricing condition $D = \mu D_1$, the passthrough of *any* efficiency change into the offer is

$$\rho := \frac{D_1}{2D_1 - \mu D_{11}} \in (0, 1) :$$

the price responds to unit cost with ρ and to quality with $1 - \rho$, so $dz_j = \rho de_j$. The marginal value of capability is $\mathcal{M}(\tau) = \mu D_1 A'$, the benchmark investment condition is $\mathcal{M} = I'$, and regularity gives the curvature $\left| \partial^2 \hat{V} / \partial \tau^2 \right| = \mathcal{M} / (\varepsilon \tau)$ by the definition of ε .

Own responses. A marginal charter taxes efficiency directly, $de_j = -\mathbb{E}(x) h d\gamma_j$, and induces the capability retreat $d\tau_j = -\omega d\gamma_j$, where differentiating the investment condition in γ and using $(1 - \rho) D_1 + \rho \mu D_{11} = \rho D_1$ (an identity implied by the definition of ρ) yields

$$\omega = \frac{\varepsilon \tau \mathbb{E}(x)}{\mu} \left[\rho h + \frac{h' D}{D_1 A'} \right].$$

The firm's total efficiency loss per unit of mandate is therefore $\mathbb{E}(x) h + A' \omega$, and its offer falls by ρ times that loss.

Rival responses. Both cross-responses are proportional to the vanishing cross-sensitivities of demand, which vanish at a common rate by assumption, so their ratio is finite. The price accommodation per unit of the neighbor's offer deterioration is $s_p := (\mu D_{12} - D_2) / (2D_1 - \mu D_{11}) > 0$, so $dp_{-j}|_\tau = s_p \rho \mathbb{E}(x) h d\gamma_j$. The capability response follows from differentiating the rival's investment condition in the neighbor's offer:

$$d\tau_{-j} = s_\tau \rho [\mathbb{E}(x) h + A' \omega] d\gamma_j, \quad s_\tau = \frac{\varepsilon \tau}{\mu} \left[\frac{s_p (1 - \rho)}{\rho} - \frac{\mu D_{12}}{D_1} \right],$$

and moves the rival's offer by $\rho A' d\tau_{-j}$. Assembling,

$$\left| \frac{L}{G} \right| = \frac{\rho A' s_\tau [\mathbb{E}(x) h + A' \omega]}{s_p \mathbb{E}(x) h} = \varphi_1 \varepsilon + \varphi_2 \varepsilon^2,$$

with

$$\varphi_1 = \frac{A' \tau}{\mu} [(1 - \rho) - \rho \chi], \quad \chi := \lim \frac{\mu D_{12}}{s_p D_1} \leq 0, \quad \varphi_2 = \varphi_1 \cdot \frac{A' \tau}{\mu h} \left[\rho h + \frac{h' D}{D_1 A'} \right],$$

both strictly positive under the maintained assumptions. Since $\varphi_1 \varepsilon + \varphi_2 \varepsilon^2$ is strictly increasing from 0, it crosses 1 at the unique ε^* stated in the proposition. In the linear system, D_1 is constant and $D_{11} = D_{12} = 0$, so $\rho = 1/2$ and $\chi = 0$; with $A = h = \text{id}$, the monopoly benchmark has $\mu = D = \tau/2$, whence $\varphi_1 = 1$ and $\varphi_2 = 2$ and $\varepsilon^* = 1/2$. The formulas are verified computationally on non-linear families (demand curvature $D_{11} \neq 0$, concave quality, convex harm). ■