AI Safety and Competition*

Jay Pil Choi† Doh-Shin Jeon‡ Domenico Menicucci§

June 15, 2026

Abstract

This paper examines how competition affects the timing of AI deployment under safety risk. We show that competition can generate two distortions relative to joint–profit maximization: a race to the bottom and insufficient entry. A race to the bottom arises when first-mover advantages induce premature deployment and is more likely as technological correlation (homogenization) increases. Conversely, firms may delay entry to free-ride on rivals' experimentation, leading to insufficient entry. Even when private incentives under joint–profit maximization are aligned with social incentives, competition can still induce socially inefficient early deployment. We discuss policy implications for improving deployment timing.

JEL Codes: D4, L1, L5

Key Words: AI, Competition, Optimal Deployment Time, Race to the Bottom.

---

*We thank Jacques Crémer, David Gilo, Michael Impink, Chuqing Jin, Anton Korinek, Jorge Lemus, Massimo Motta, Sara Shahanaghi, Yongseok Shin, Alex Smolin, John Turner, Bryan Wilder, Liang Zhong, and the participants of the CEPR Paris Symposium 2025, the Digital Economics Conference 2026 (Toulouse), the Economics of AI conference 2026 (Barcelona), MaCCI 2025 (Mannheim), Paris Digital Conference 2025, IIOC 2026 (Boston) and the seminar at Korea University for valuable discussions and comments. Jeon acknowledges funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

†Department of Economics, Michigan State University, 220A Marshall-Adams Hall, East Lansing, MI 48824 -1038. E-mail: choijay@msu.edu

‡Toulouse School of Economics, University of Toulouse Capitole. E-mail: dohshin.jeon@tse-fr.eu

§Dipartimento di Scienze per l'Economia e l'Impresa, Università degli Studi di Firenze. E-mail: domenico.menicucci@unifi.it

1 Introduction

AI systems with human-level or superhuman intelligence offer substantial potential benefits, but also introduce significant and potentially far-reaching risks to society and humanity. The recent surge in generative AI—exemplified by technologies like ChatGPT—has sparked fierce competition among major tech companies (e.g., Microsoft, Google, Meta, Amazon) and startups (e.g., OpenAI, Anthropic, xAI). This intensifying race raises concerns that these firms might prioritize market leadership over AI safety, potentially leading to a "race to the bottom." For instance, Geoffrey Hinton, a pioneer of deep learning and a 2024 Novel Prize laureate, has expressed concern about increasingly powerful machines potentially surpassing human capabilities in ways that may not serve humanity's best interests.[1] Reflecting similar concerns, 33,707 individuals signed[2] an open letter posted by the Future of Life Institute on March 22, 2023, calling for all AI labs to "immediately pause for at least 6 months the training of AI systems more powerful than GPT-4." The goal was to develop and implement shared safety protocols for the design and development of advanced AI. Max Tegmark, a professor of physics and AI researcher at MIT, has also warned that the world is "witnessing a race to the bottom that must be stopped".[3]

There are three main categories of risks associated with general-purpose AI: malicious use risks, malfunction risks, and systemic risks (Bengio et al. (2025)). Malicious use risks arise when bad actors exploit general-purpose AI to harm individuals, organizations, or society. Such harms may include the creation of fake content, manipulation of public opinion, cyberattacks, and biological or chemical threats.

Even in the absence of malicious intent, significant risks can arise from the malfunctioning of general-purpose AI systems. These include reliability issues, bias, and loss of control. For example, AI systems may generate inaccurate or misleading outputs or amplify existing social and political biases. The loss of control refers to situations in which the objectives pursued by advanced AI systems diverge from those intended by their human designers, giving rise to alignment problems.[4] As AI systems become

[1] Sara Brown, "Why Neural Net Pioneeer Geoffrey Hinton is Sounding the Alarm on AI", Ideas Made to Matter, May 23, 2023.

[2] At the time of September 29, 2024.

[3] The Guardian, "Humanity at risk from AI 'race to the bottom', says tech expert", October, 26, 2023.

[4] There are two types of alignment problems (Gladstone AI (2023)). The first, known as the outer alignment problem, refers to the challenge of identifying objectives for an extremely competent optimizer that do not lead to dangerous or undesirable outcomes. This challenge arises because AI models can only be trained to optimize imperfect *proxies* for human goals. In contrast, the inner alignment problem concerns ensuring that AI systems reliably internalize the goals specified for them. This issue occurs because, under the current training paradigm, there is no surefire way to fully anticipate, control, or verify which goals an AI system will ultimately internalize. Anwar et al. (2024) provide 18 foundational challenges in assuring the alignment and safety of large language models.

more powerful, these concerns become more acute, as even small misalignments can lead to substantially different and potentially harmful outcomes. In extreme cases, alignment failures may render AI systems uncontrollable, leading them to act in ways that are adverse to human interests (Gladstone AI (2023)).

Finally, systemic risks arise from the widespread adoption of general-purpose AI. For instance, if organizations across critical sectors such as finance or healthcare rely on a small number of AI systems, a bug or vulnerability in one of these systems could trigger large-scale, simultaneous disruptions.

Although several national competition authorities and regulatory agencies (Autoridade da Concorrência (2023), Autoridade da Concorrência (2024), Autorité de Concurrence (2024), Competition & Markets Authority (2023), Competition & Markets Authority (2024), Japanese Fair Trade Commission (2024)) have published reports addressing competition issues in the market for foundation models, none has systematically examined the interplay between competition and AI safety. This oversight may reflect the view that AI safety falls outside the traditional scope of competition and consumer protection (Jeon (2025)). Alternatively, it may stem from intensifying international competition in AI, which leads national authorities to prioritize reducing entry barriers to remain competitive in the global race.

In this paper, we aim to understand the interplay between competition and AI safety—specifically, the conditions under which an AI race transforms into a race to the bottom. To this end, we consider a game in which each firm chooses the timing of AI deployment and characterize the conditions under which competitive pressures induce premature deployment.

According to Bommasani et al. (2022), the development of AI is characterized by increasing *emergence* and *homogenization*. Emergence describes the phenomenon where a system's behavior is implicitly induced rather than explicitly designed, fueling both scientific excitement and concerns over unforeseen consequences. Homogenization refers to the consolidation of methodologies used to build machine learning systems across various applications. Foundation models, such as those driving generative AI, have accelerated this homogenization to an unprecedented degree, with most state-of-the-art NLP (Natural Language Processing) models now adapted from a small set of such models. As a result, nearly all AI systems may inherit the same problematic biases originating from a few key foundation models.

Motivated by these stylized facts and additional ones presented in Section 1.1, we develop a simple two-period model in which two competing firms decide when to deploy their AI technology. To capture the emergent nature of AI, we assume there is a positive probability that a firm's AI technology is unsafe.[5] A safe technology

[5]See Section 1.1 for additional stylized facts regarding the safety issue of general-purpose AI

generates positive profits and consumer surplus, whereas an unsafe technology can trigger a catastrophic event that imposes substantial harm on society, with the firm bearing only a portion of this harm in the form of liability. If the technology is deployed in the first period, the market learns whether it is safe. We assume that deployment is reversible: if the technology is revealed to be unsafe in the first period, it can be withdrawn in the second period, generating no further harm. If the firm does not deploy in the first period, it can instead conduct beta testing, which provides imperfect signals about safety. To reflect the homogenization of AI systems, we assume that the risk profiles of the two competing technologies are positively correlated.

Finally, we incorporate a first-mover advantage: the firm that deploys its technology in the first period secures a larger market share compared to a competitor that enters in the second period. This advantage may arise from several sources, including data-feedback loops—where accumulated user data improves the model through retraining and generates data-driven network effects—intelligence feedback loops, and scaling laws.[6] For instance, Bengio et al. (2025) note that "market leaders benefit from self-reinforcing dynamics that reward winners".

We start by considering a monopoly benchmark and compare the private incentive to the social incentive in terms of the deployment timing of the AI technology. We find that the two are aligned if and only if the ratio of the harm externality (i.e., the gap between the societal harm and the firm's liability) over the private liability is equal to the ratio of the benefit externality (i.e., consumer surplus) over the profit.

We then turn to the main case of duopoly. We begin by comparing the equilibrium outcome with the joint-profit–maximizing outcome to delineate the effect of competition on the timing of AI deployment. Any discrepancy between the two can lead to a corresponding divergence between the market outcome and the welfare-maximizing outcome, even when the conditions for the alignment of private and social incentives hold in the monopoly case. For this analysis, we assume that the static payoff from deployment is negative, so that joint-profit maximization may, depending on parameter values, entail no firm entering in the first period.[7] We further restrict attention to cases in which beta testing is sufficiently informative such that, conditional on no first-period entry, the second-period entry decision is fully determined by the test outcome.

We begin by examining two extreme cases: perfect beta testing and perfect correla-

models. See also Bengio et al. (2025).

[6]See Section 1.1 for more detailed discussions of the first-mover advantage. See also Autoridade da Concorrência (2023), Competition & Markets Authority (2023) and Korinek and Vipra (2025) for a discussion of the first-mover advantage and other above-mentioned effects.

[7]Without this assumption, joint profit is always maximized when both firms enter in the first period.

tion. In the case of perfect beta testing, first-period deployment provides no additional information. Under our assumption that the static payoff from deployment is negative, joint-profit maximization therefore entails no firm entering in the first period. However, competition may induce both firms to enter in the first period, leading to a race to the bottom. This outcome is driven by what we term the dynamic gain from preemptive entry, arising from the first-mover advantage. This gain increases not only with the magnitude of the first-mover advantage but also with the likelihood that *both* firms' technologies are safe—a likelihood that rises with a higher degree of positive correlation between the technologies.

In the case of perfect correlation, one firm's experimentation through deployment generates a perfectly informative signal for its rival, introducing a new distortion relative to the joint-profit–maximizing outcome. A race to the bottom may arise when beta testing is sufficiently precise. However, when beta testing is relatively imprecise, the opposite distortion—insufficient entry—can occur: joint-profit maximization calls for both firms to enter in the first period, yet either no firm or only one firm enters. This outcome is driven by each firm's incentive to free ride on the rival's experimentation.

In the general case of imperfect correlation and imperfect beta testing, we identify three regimes characterized by discrepancies between the market outcome and the joint–profit–maximizing benchmark. For a given level of first-mover advantage, as the probability that the technology is unsafe increases, the equilibrium transitions from a race-to-the-bottom regime to an insufficient-entry regime, and eventually to a no-distortion regime in which risk is sufficiently high that no firm enters in the first period, regardless of whether firms compete or coordinate. The likelihood of a race to the bottom increases with both the strength of the first-mover advantage and the degree of correlation, while the effect of beta-testing precision is non-monotonic. A race to the bottom may arise either as a unique equilibrium or as one of multiple equilibria. In the latter case, we refer to it as a FOMO-driven race to the bottom, where FOMO denotes "fear of missing out."[8] We further show that when beta testing is more informative than the information generated by a rival's market experimentation, the rival's entry increases a firm's incentive to enter in the first period. This mechanism, through the FOMO-driven case, substantially expands the prevalence of race-to-the-bottom outcomes.

We find that the socially optimal outcome involves either both firms entering in

[8]FOMO is often invoked to explain the current boom in AI investment. According to Google Cloud chief Thomas Kurian, "There's been this sort of FOMO of, 'I need to be in generative AI for generative AI's sake.'" See "Google unveils enterprise AI tools, new AI chip" by Max A. Cherney, Reuters, August 29, 2023, at https://www.reuters.com/technology/google-unveils-enterprise-ai-tools-new-ai-chip-2023-08-29/. For another article, see "The AI investment paradox: Genuine transformation or FOMO at scale?", Paul Frenken, Mar 17, 2026, CIO, at https://www.cio.com/article/4145846/the-ai-investment-paradox-genuine-transformation-or-fomo-at-scale.html.

the first period or both postponing entry. This is because, conditional on one firm entering in the first period, the marginal value of additional experimentation from a second entrant remains positive. Focusing on the case in which the societal harm from catastrophic events is large, we show that welfare is maximized when no firm enters in the first period, provided that the weight on second-period payoffs is below a threshold greater than one. In this setting, aligning private incentives with the social optimum can be achieved in two steps. First, the social objective and firms' joint incentives can be aligned through appropriate liability design: as in the monopoly case, alignment requires that the ratio of external harm to private liability equals the ratio of external benefits to private profits. Second, individual incentives can be aligned with joint incentives by reducing first-mover advantages and limiting technological homogenization. However, policymakers may have limited instruments to directly influence these structural factors, particularly when the race to the bottom can be also driven by fear of missing out (FOMO). In such cases, direct regulatory interventions based on risk assessment—such as the EU AI Act—may be warranted to prevent premature deployment. More broadly, the market failure underlying the race to the bottom provides a rationale for AI risk regulation.

We also note that our framework applies more broadly to other emerging technologies, such as autonomous vehicles, where risks are not fully understood ex ante and can only be assessed through beta testing and real-world deployment. In such settings, risk profiles are likely to be correlated across firms, as their technologies rely on similar underlying fundamentals.

Our paper is closely related to Guerreiro, Rebelo and Teles (2026), who study the optimal regulation of AI when deployment generates both substantial benefits and potentially large societal risks. They focus on a setting with uncertainty and disagreement between developers and regulators regarding societal costs, as well as the possibility of learning about risks through beta testing and red teaming prior to deployment.[9] A key result is that standard Pigouvian taxation fails to achieve the first-best outcome because regulators cannot observe developers' beliefs, leading them to propose a two-stage policy governing experimentation and subsequent public release. Our paper is complementary to theirs but shifts attention from optimal regulation under asymmetric beliefs to the role of market structure. Whereas they analyze the problem in a monopoly-style environment, we emphasize how competition among AI developers changes deployment incentives and can itself generate distortions, such as a race to the bottom or insufficient entry, thereby creating a distinct rationale for policy intervention.

[9]Red teaming is a simulation process in which authorized professionals (the "red team") act as adversaries to identify and test vulnerabilities in a technology.

Jones (2024) analyzes the optimal use of AI within a growth model framework. While AI enhances economic growth, its advancements may also pose an "existential risk" leading to human extinction. He examines the conditions under which the rapid progress of AI should continue or be halted, approaching the issue from a collective societal perspective in a non-strategic model. He shows that the answer crucially depends on the curvature of the constant flow utility function of consumption. In contrast, our model takes a strategic approach, where the decision to deploy AI technology is driven by competition among AI developers.

Acemoglu and Lensman (2024) develop a dynamic multi-sector technology adoption model in which the risks associated with AI can be learned over time. They demonstrate that the socially optimal adoption strategy is gradual and follows a convex path. Gans (2025) examines the validity of recent proposals advocating for delaying AI adoption until its potential harms are fully understood. His analysis reveals that the optimality of such a delay depends on the learning mechanism and the reversibility of the associated harms.

Beyond the AI literature, our paper is also related to Bobtcheff, Levy and Mariotti (2025), who study preemption races in which firms compete to be the first to invest in a common-value risky project after receiving private signals. Their framework can be viewed as a polar case of our model in which risk is perfectly correlated across firms and the payoff structure reflects an extreme form of first-mover advantage—a winner-takes-all outcome. In addition, the focus of their analysis differs substantially from ours. They examine how the disclosure of private information can either mitigate or exacerbate inefficient preemption, but their model does not involve safety concerns or societal harm arising from unsafe technology. In contrast, our paper focuses on AI deployment decisions in the presence of safety risks and analyzes how the interplay between technological correlation and first-mover advantage shapes deployment timing when beta testing is available, showing that competition among AI developers can generate both a race to the bottom and insufficient entry.

The rest of the paper is organized as follows. Section 1.1 presents additional stylized facts on foundation models that motivate our modeling choices. Section 2 presents the basic model of a monopolistic AI developer and analyzes the optimal timing of technology deployment relative to the social optimum. Section 3 extends the monopoly model to a duopolistic setting, and Section 4 analyzes equilibrium outcomes relative to the joint–profit–maximizing benchmark. Section 5 provides a welfare analysis of duopolistic competition, and Section 6 discusses policy implications. Section 7 concludes. All detailed proofs are relegated to the Appendix.

1.1 Some stylized facts about foundation models

We below provide some stylized facts about foundation models which motivate our modeling choices.

First mover advantage

One important determinant of first-mover advantage in AI markets is the presence of *data feedback loops*, whereby increased user interaction generates data that improves model performance, which in turn attracts more users and further strengthens the cycle. While foundation models rely heavily on pre-training using large, often publicly available datasets, post-deployment interactions—such as user prompts, corrections, preferences, and usage patterns—provide valuable fine-tuning data, particularly for domain-specific tasks and long-tail queries. As models are increasingly deployed in real-world applications, this iterative learning process can generate cumulative advantages for early entrants. Several recent studies and policy reports highlight the importance of such feedback mechanisms.

For instance, a recent report by the Portuguese competition authority (Autoridade da Concorrência (2024)) emphasizes that model performance improves with exposure to a broader and richer set of inputs, especially long-tail prompts, which are difficult to capture ex ante. Similarly, the Competition & Markets Authority (2023) notes that firms with larger user bases are better positioned to collect and utilize feedback for subsequent model improvements. Moreover, Villalobos et al. (2022) suggest that high-quality training data may become increasingly scarce, further increasing the value of proprietary, user-generated data and reinforcing entry barriers.

Another source of first-mover advantage is the *intelligence feedback loop* that arises as foundation models are starting to play a significant role in cognitive work, especially in programming (Korinek and Vipra (2025)). If an AI lab has better internal foundation models available than its competitors, then it will be faster at making progress on the next model version. This phenomenon is already reflected in the current rapid pace of AI progress.[10] The intelligence feedback loop may enable the leading AI lab to further distance itself from competitors, separate itself further and further from the competition, developing a technology that is so advanced that it establishes market dominance (Korinek and Vipra (2025)).

In addition to data and intelligence feedback loops, *switching costs* provide another important source of first-mover advantage. As AI systems become embedded in workflows—particularly through personalized AI agents that learn users' preferences, habits, and contexts—switching to alternative providers becomes increasingly costly.

[10]For example, Google recently announced that more than 25 percent of the new code generated at the company comes from AI (Pichai (2024)).

These costs may arise from the loss of personalization, the need to retrain systems, and the disruption of complementary applications or ecosystems. Such switching frictions can lock in users and amplify the advantages of early deployment.

Taken together, data and intelligence feedback loops and switching costs provide strong economic foundations for first-mover advantage in AI markets, even if the strength of these mechanisms may vary across applications.[11]

Ultimately, the scale law in foundation models, meaning that both the costs of frontier models and their capabilities are increasing very rapidly, implies that the number of players that a market of a given size can support is shrinking fast, creating the first-mover advantage. The amount of computation used to train the most powerful AI models has increased by a factor of ten every year for the last ten years: today's most advanced AI models require five billion times the computing power of cutting-edge models from a decade ago (Bremmer and Suleyman (2023)).[12] Recent research suggests that rapidly scaling up models may remain physically feasible for at least several more years (Bengio et al. (2025)). In addition, the valuation history of leading private AI labs appears to follow the scale law, mirroring trends in model capabilities, with doubling times historically floating around the half-year mark.[13]

AI safety risk

The international AI safety report, prepared for the Paris AI Action Summit, by a panel nominated by the governments of 30 countries, as well as the UN, EU, and OECD, and chaired by Yoshua Bengio (Bengio et al. (2025)), provides state-of-the-art scientific evidence on the safety risks associated with general-purpose AI. The report highlights, among systemic risks, the threat of "market concentration and single points of failure," arising from the dominance of a small number of firms in the general-purpose AI market.

A single point of failure refers to a component whose malfunction can disrupt an entire system. As general-purpose AI models become widely deployed across critical sectors—including finance, healthcare, and cybersecurity—this risk becomes increasingly salient. These sectors are highly interdependent and play a central role in national security and economic stability, and they increasingly rely on AI for decision-making, resource optimization, automation, and threat detection. Because a small number of

[11] Hagiu and Wright (2025) argue that data feedback loops may be weaker in certain settings, particularly when data is non-exclusive or learning quickly saturates. See also Gans (2026) for a discussion of training data of AI as the source of market power. He argues that a key factor in determining the emergence and persistence of market power will be availability of data across firm boundaries.

[12] "Brain scale" models with more than 100 trillion parameters – roughly the number of synapses in the human brain – will be viable within five years (Bremmer and Suleyman (2023)).

[13] See the State of AI Report 2025 at https://www.stateof.ai/.

dominant AI models serve as the backbone for many such applications, vulnerabilities in these systems can propagate broadly. In particular, flaws, bugs, or biases in widely adopted models may trigger simultaneous disruptions across multiple industries. For example, a cybersecurity vulnerability in a dominant AI model could expose financial institutions, government agencies, and other critical infrastructures to coordinated attacks or system-wide failures.

Several technical features of general-purpose AI make risk management in this domain particularly difficult for multiple reasons. First, the range of possible uses and use contexts for general-purpose AI systems is unusually broad. Second, developers still understand little about how their general-purpose AI models operate.[14] Third, increasingly capable AI agents – general-purpose AI systems that can autonomously act, plan, and delegate to achieve goals – will likely present new, significant challenges for risk management.

In addition, economic and political factors can make risk management of general-purpose AI difficult. Both AI companies and governments frequently face intense competitive pressures, which may lead to the deprioritization of risk management. For companies, the drive to outpace competitors can incentivize them to allocate fewer resources and less time to addressing risks. Similarly, governments may underinvest in policies that support risk management when they perceive a trade-off between fostering international competitiveness and reducing potential risks.

2 The Monopoly Model

Consider a scenario involving a monopolistic developer of AI technology. The technology has the potential to generate substantial benefits but also entails risks that are not yet fully understood. These risks can only be fully revealed through actual deployment or partially assessed through internal beta testing before the technology is released to the public. The uncertainty stems from the emergent nature of AI systems, which involve complex and poorly understood interactions. For example, large language models contain trillions of parameters, and their bahavior is often characterized as a "black box," leading to unpredictable outcomes such as "hallucinations."

There are two possible states of nature: $\omega \in \Omega = \{S, N\}$. In state $\omega = S$, the AI technology is safe, and no catastrophic events with harmful effects occur. In this case, the technology generates a profit of B for the firm and a benefit of C for consumers. For simplicity, we normalize B to 1, so all benefits and costs are measured relative to the firm's private benefits in the safe state. In state $\omega = N$, the AI technology is

[14]The inner workings of general-purpose AI models are largely inscrutable, even to the model developers (Bengio et al. (2025)).

not safe, and catastrophic events may occur. If such an event takes place, the firm incurs a cost of L, while society bears a cost of D where $L(\leq D)$ represents the firm's liability. The difference $(D - L)$ represents the external harm imposed on society by the catastrophic event beyond the firm's liability. Since liability is just a monetary transfer, the total social harm is equal to D. The prior that the technology is not safe is given by $\lambda \in (0, 1)$.

To analyze the optimal timing of technology deployment by the firm and compare it to the socially optimal timing, we employ a simple two-period model. The firm can choose to deploy the technology in period 1, period 2, or decide not to introduce it at all. If the technology is not safe, we assume that a catastrophic event will inevitably occur in the period the technology is introduced.

If the firm deploys the technology in the first period and it proves to be safe, the firm continues deployment in the second period. However, if the technology is revealed to be unsafe and causes harm in the first period, we assume that the firm can cease deployment and exit the market in the second period, so that the damage is limited to the first period. In other words, in line with the terminology used by Acemoglu and Lensman (2024), we assume that the AI technology is *reversible*, meaning that it can be withdrawn without causing further damage.

If the firm decides to delay deployment, it can conduct beta testing in the first period to assess the safety of the technology. However, this testing is not fully informative. Specifically, if the technology is safe (i.e., $\omega = S$), beta testing will not yield any negative signal (i.e., the firm receives a null signal $\sigma = \phi$), meaning there are no false positives incorrectly indicating that the technology is unsafe. In contrast, if the technology is not safe (that is, $\omega = N$), the test will detect the unsafe state and generates a signal $\sigma = n$ with probability β, which represents the informativeness of the test. Thus, the signal structure follows a "no news is good news" pattern: the absence of a signal does not definitively confirm that the technology is safe, whereas receiving a signal $\sigma = n$ provides conclusive evidence that the technology is unsafe.

Mathematically, the signal structure of beta testing can be represented as follows:

$$\begin{aligned} \Pr(\sigma &= n|\omega = S) = 0, \\ \Pr(\sigma &= n|\omega = N) = \beta. \end{aligned}$$

The updated belief, conditional on the observed signal, is as follows:

$$\widetilde{\lambda}(\phi) = \Pr(\omega = N|\sigma = \phi) = \frac{\lambda(1-\beta)}{(1-\lambda)+\lambda(1-\beta)} = \left(\frac{1-\beta}{1-\beta\lambda}\right)\lambda < \lambda,$$

$$\widetilde{\lambda}(n) = \Pr(\omega = N|\sigma = n) = 1 > \lambda.$$

The main question we explore in this context is the optimal timing for the monopolist to deploy the technology and whether it aligns with the socially optimal timing. The monopolist faces the decision of whether to deploy the technology in the first period or delay deployment to the second period, using beta testing in the first period to assess potential risks and make a more informed decision.

Before analyzing the monopolist's optimal deployment timing in a dynamic model, we first examine the static one-period problem (without the opportunity for beta testing) as a benchmark. Then, the monopolist deploys the technology if and only if

$$E_s \equiv (1-\lambda)1 - \lambda L > 0,$$

that is, if $\lambda < \lambda_s^* \equiv \frac{1}{1+L}$, where E_s represents the incentive to enter in the static model.

2.1 The Monopolist's Optimal Deployment Timing

We now consider the monopolist's optimal entry timing in a two-period model. We denote $e_1 = 1$ if the monopolist enters the market in period 1 and $e_1 = 0$, otherwise.

When the technology is deployed in the first period without additional beta testing, the firm's expected payoff is

$$\begin{aligned} \Pi(e_1 &= 1) = [(1-\lambda)\cdot 1 - \lambda L] + \delta(1-\lambda) \\ &= (1-\lambda)\cdot(1+\delta) - \lambda L = (1+\delta)E_s + \delta I_d, \end{aligned}$$

where δ represents the relative weight of second-period payoffs compared to first-period payoffs and $I_d \equiv \lambda L$ captures the informational value obtained from deploying the technology; if the AI technology turns out to be unsafe (which occurs with probability λ), the monopolist has the option to exit the market in the second period and thereby avoid the liability cost L.[15]

The monopolist will have an incentive to enter the market in the first period only when $\Pi(e_1 = 1) > 0$, or $\lambda < \lambda_d^* \equiv \frac{1+\delta}{1+\delta+L}$. Note that $\lambda_d^* > \lambda_s^*$, meaning the monopolist may have incentives to enter the market in the first period even when it would not in the static model. This is due to the positive value of information. Since first-period deployment is akin to "experimentation," λ_d^* is also an increasing function of δ.

If the firm waits and conducts beta testing in the first period, it will deploy its technology in the second period only if it receives no negative signal. Therefore, its

[15] If the AI technology is *irreversible* and cannot be withdrawn from the market, I_d (the value of information term) will be zero.

expected payoff with beta testing in the first period is:

$$\Pi(e_1 = 0) = \max\{\delta E_2(\beta), 0\},$$

where $E_2(\beta) \equiv (1-\lambda) \cdot 1 - \lambda(1-\beta)L$ represents the incentive to enter in the second period after beta testing without any negative signals. Note that $E_2(\beta = 1) > 0$. More generally, we can rewrite $E_2(\beta)$ as

$$E_2(\beta) = E_s + I_\beta,$$

where $I_\beta \equiv \lambda\beta L (< I_d)$ represents the informational value of beta testing. It captures the expected savings in liability costs due to the information revealed by testing. Thus, the condition for $\Pi(e_1 = 0) > 0$ is equivalent to $E_2(\beta) = E_s + I_\beta > 0$, and is given by

$$\lambda < \lambda_2^*(\beta) \equiv \frac{1}{1 + (1-\beta)L} (> \lambda^s).$$

We now consider the case in which both entering in the first period and postponing entry to the second period yield strictly positive profits (i.e., $\lambda < \min\{\lambda_d^*, \lambda_2^*(\beta)\}$), and investigate the conditions under which the monopolist prefers to enter in the first period.

We have

$$\Pi(e_1 = 1) - \Pi(e_1 = 0) = \underbrace{(1-\lambda) \cdot 1 - \lambda L}_{=E^s} + \delta \underbrace{\lambda(1-\beta)L}_{=\Delta^I},$$

where $\Delta^I \equiv I_d - I_\beta$ represents the informational advantage of deployment relative to beta testing. Hence, if $E^s > 0$ (i.e., $\lambda < \lambda_s^*$), the monopolist strictly prefers to enter in the first period.

In general, the condition for $\Pi(e_1 = 1) - \Pi(e_1 = 0) > 0$ is given by

$$\lambda < \lambda_1^*(\beta) \equiv \frac{1}{1 + [1 - \delta(1-\beta)]\, L} (> \lambda_s^*).$$

Note that both $\lambda_1^*(\beta)$ and $\lambda_2^*(\beta)$ are larger than the static entry threshold λ_s^*. The relative magnitudes of $\lambda_1^*(\beta)$ and $\lambda_2^*(\beta)$ depend on the informativeness of beta testing. Observe that $\lambda_1^*(\beta)$ is decreasing in β while $\lambda_2^*(\beta)$ is increasing in β, which implies that as the informativeness of beta testing improves, delaying the deployment of the technology becomes more attractive. A direct comparison of them immediately yields the following result.

Lemma 1. *There exists a* $\beta^* \equiv \frac{\delta}{1+\delta}$ *such that (i)* $\lambda_1^*(\beta) < \lambda_d^* < \lambda_2^*(\beta)$ *if* $\beta > \beta^*$*, and (ii)* $\lambda_2^*(\beta) < \lambda_d^* < \lambda_1^*(\beta)$ *if* $\beta < \beta^*$*.*

From Lemma 1, consider first the case $\beta > \beta^*$. The inequality $\lambda_d^* < \lambda_2^*(\beta)$ implies that whenever entry in the first period yields a positive profit (i.e., when $\lambda < \lambda_d^*$), postponing entry also yields a positive profit. In this region, when both options are profitable, the monopolist prefers first-period entry if $\lambda < \lambda_1^*(\beta)$, and prefers to delay entry if $\lambda_1^*(\beta) < \lambda < \lambda_d$. For $\lambda_d < \lambda < \lambda_2^*(\beta)$, entry in the first period is not profitable, and the monopolist instead enters in the second period if the beta test yields no negative signal.

Next, consider $\beta < \beta^*$. The inequality $\lambda_2^*(\beta) < \lambda_d^*$ implies that whenever postponing entry yields a positive profit (i.e., when $\lambda < \lambda_2^*(\beta)$), first-period entry also does so. Moreover, since $\lambda_2^*(\beta) < \lambda_1^*(\beta)$, postponing entry is strictly dominated by entering in the first period. Therefore, in this case, the monopolist enters in the first period if and only if $\lambda < \lambda_d^*$.

In summary, the following proposition characterizes the monopolist's optimal entry decision.

Proposition 1. *There are two cases to consider depending on the precision of the beta testing signal.*
(i) $\beta > \beta^*$*: (a) If* $\lambda < \lambda_1^*(\beta)$*, the monopolist enters in the first period.*
(b) If $\lambda \in (\lambda_1^*(\beta), \lambda_2^*(\beta))$*, the monopolist engages in beta-testing and enters only when it receives no signal of unsafeness.*
(c) If $\lambda > \lambda_2(\beta)$*, it never deploys the technology.*
(ii) $\beta < \beta^*$*: The monopolist enters in the first period if* $\lambda < \lambda_d^*$*; otherwise, it does not enter. Thus, the availability of beta testing is irrelevant in the monopolist's entry decision.*

2.2 Socially Optimal Deployment Timing

To derive the socially optimal deployment timing, we can derive the corresponding social welfare associated with the deployment of the AI technology in the first period:

$$\begin{aligned} SW(e_1 &= 1) = [(1-\lambda)\cdot(1+C) - \lambda D] + \delta(1-\lambda)(1+C) \\ &= (1-\lambda)\cdot(1+\delta)(1+C) - \lambda D. \end{aligned}$$

Note that $SW(e_1 = 1) > 0$ if and only if $\lambda < \lambda_d^w \equiv \frac{(1+\delta)(1+C)}{(1+\delta)(1+C)+D}$.

Similarly, the social welfare with beta testing can be written as

$$SW(e_1 = 0) = \max\{\delta\left[(1-\lambda)\cdot(1+C) - \lambda(1-\beta)D\right], 0\}.$$

As in the analysis of monopolist's private incentives, we can derive the condition for $SW(e_1 = 0) > 0$, which is given by

$$\lambda < \lambda_2^w(\beta) \equiv \frac{(1+C)}{(1+C)+(1-\beta)D}.$$

A social planner would like to deploy the technology early when $SW(1) > SW(0)$. This condition can be written as

$$\lambda < \lambda_1^w(\beta) \equiv \frac{(1+C)}{(1+C)+[1-\delta(1-\beta)]\,D}.$$

As in the analysis of the private incentives, the relative rankings of $\lambda_1^w(\beta), \lambda_2^w(\beta)$, and $\lambda_d^w(\beta)$ depend on β.

Lemma 2. *There exists a $\beta^* \equiv \frac{\delta}{1+\delta}$ such that (i)$\lambda_1^w(\beta) < \lambda_d^w < \lambda_2^w(\beta)$ if $\beta > \beta^*$, and (ii) $\lambda_2^w(\beta) < \lambda_d^w < \lambda_1^w(\beta)$ if $\beta < \beta^*$.*

With the help of the above lemma, we can completely characterize the socially optimal decision for the monopolist in the following proposition.

Proposition 2. *There are two cases to consider depending on the precision of the beta testing signal.*
(i) $\beta > \beta^*$: (a) If $\lambda < \lambda_1^w(\beta)$, the monopolist entering in the first period is socially optimal.
(b) If $\lambda \in (\lambda_1^w(\beta), \lambda_2^w(\beta))$, the socially optimal outcome is for the monopolist to engage in beta-testing and to enter only when it receives no signal of unsafeness.
(c) If $\lambda > \lambda_2^w(\beta)$, it should never deploy the technology.
(ii) $\beta < \beta^*$: The monopolist entering in the first period is socially optimal if $\lambda < \lambda_d^w$. Otherwise, it should never deploy the technology. Thus, the availability of beta testing is irrelevant in the social planner's entry decision.

The inefficiency associated with the monopolist's decision compared to the social planner's can be assessed by comparing the threshold values for the monopolist ($\lambda_i^*(\beta)$) and the social planner's ($\lambda_i^w(\beta)$). The following proposition states that the comparison depends on the relative ratios of positive consumption externalities not appropriated by the monopolist and the damage not borne by the monopolist in the risk case.

Proposition 3. *$\lambda_i^*(\beta) \gtreqless \lambda_i^w(\beta)$ if and only if $\frac{D}{L} \gtreqless 1 + C$, where $i \in \{1, 2, d\}$.*

The private incentives to deploy the technology compared to social incentives change in a predictable way. The condition $\frac{D}{L} \gtreqless 1 + C$ is equivalent to

$$\frac{D-L}{L} \gtreqless \frac{(1+C)-1}{1},$$

where the L.H.S. represents the ratio of the harm externality to private liability, whereas the R.H.S. represents the ratio of the benefit externality to private benefit. Hence, aligning private incentives with the social optimum requires setting the liability level appropriately, namely, $L = D/(1 + C)$; the optimal liability increases with magnitude of potential harm from catastrophic events.

3 The Model of Duopolistic Competition

To analyze the effect of market competition on the AI deployment timing decision, we consider a situation in which two (symmetric) firms - firm 1 and firm 2 - compete to sell their AI technologies. Our focus is on the timing of the commercial release of each firm's technology. In this section, we describe the model.

A key feature of our analysis is the incorporation of potential *correlation* between the two technologies, driven by the *homogenization* of AI systems (Bommasani et al., 2022). As noted in the introduction, foundation models—such as generative AI—have led to an unprecedented level of homogenization, with most state-of-the-art natural language processing (NLP) models now being adapted from a small set of foundation models.

To capture the homogenization of AI systems, we assume a positive correlation in the risk profiles of the two AI technologies being developed. There are four possible states of nature regarding the safety of each firm's AI technology: $\omega = (\omega_1, \omega_2) \in \Omega = \{(S,S), (S,N)(N,S)(N,N)\}$, where ω_i denotes the safety status of firm i's AI technology with $i = 1, 2$. Let $\rho \in [0, 1]$ be the correlation parameter, with $\rho = 0$ representing independent risks and $\rho = 1$ representing perfectly correlated risk as two extreme cases. Under correlated risk, the prior probability distribution over the four states is specified as follows:

$$\begin{aligned}
\mu_{NN} &\equiv \Pr(N,N) = \lambda^2 + \rho\lambda(1-\lambda), \\
\mu_{NS} &\equiv \Pr(N,S) = \mu_{SN} \equiv \Pr(S,N) = \lambda(1-\lambda)(1-\rho), \\
\mu_{SS} &\equiv \Pr(S,S) = (1-\lambda)^2 + \rho\lambda(1-\lambda).
\end{aligned}$$

If a firm deploys its technology in the first period, the safety of that technology is fully revealed at the end of the period. If a technology is not deployed in the first period, the result of its beta test remains private information and is not observed by the other firm.[16] When neither firm deploys its technology in the first period, belief

[16]For an analysis of how information disclosure affects the timing of investment in preemption races involving a common-value risky project, see Bobtcheff, Levy and Mariotti (2025). For a model with public information revelation, see Figueroa, Guadalupi and Lemus (2026), who analyze innovation incentives in the presence of potential harms, where firms can engage in "Responsible Innovation" by

updating in the second period proceeds as in the monopoly model, since each firm's test result remains private.

Now consider the updating of beliefs about safety under an asymmetric entry configuration in the first period in which only one firm enters the market with firm j deploying its technology while firm $i(\neq j)$ does not. Then, firm i's posterior belief about the safety of its own technology depends on two pieces of information: (i) the publicly revealed safety status of firm j's technology, and (ii) the private test result of its own technology. Based on these, firm i's updated belief about the safety of its own technology can be derived as follows:

$$\begin{aligned}
\widetilde{\lambda}_i(\phi, S) &= \Pr(\omega_i = N | \sigma_i = \phi, \omega_j = S) \\
&= \frac{\lambda(1-\rho)(1-\beta)}{[(1-\lambda)+\rho\lambda]+\lambda(1-\rho)(1-\beta)} (< \widetilde{\lambda}(\phi) < \lambda), \\
\widetilde{\lambda}_i(\phi, N) &= \Pr(\omega_i = N | \sigma_i = \phi, \omega_j = N) \\
&= \frac{[\lambda+\rho(1-\lambda)](1-\beta)}{[(1-\lambda)(1-\rho)]+[\lambda+\rho(1-\lambda)](1-\beta)} (> \widetilde{\lambda}(\phi)), \\
\widetilde{\lambda}_i(n, N) &= \widetilde{\lambda}_i(n, S) = 1.
\end{aligned}$$

When $\sigma_i = \phi$ and $\omega_j = N$, with $j \neq i$, firm i receives conflicting signals. Its own test result (no negative signal) suggests that its technology is more likely to be safe than under the prior. However, learning that the other firm's technology is unsafe makes firm i more pessimistic about its own technology's safety due to the assumed positive correlation in risk. Which of these effects dominates depends on the informativeness of its beta testing (β), the degree of positive correlation (ρ), and the prior probability of risk (λ). Intuitively, the effect of firm i's own test result dominates the correlation effect—leading to a downward revision of the perceived risk (i.e., $\widetilde{\lambda}_i(\phi, N) < \lambda$)—if and only if β is sufficiently large relative to ρ. This is formalized in the following lemma.

Lemma 3. *$\widetilde{\lambda}_i(\phi, N) < \lambda$ if and only if $\beta > \frac{\rho}{\lambda+\rho(1-\lambda)}$.*

We next describe the payoff structure, which depends on the firms' entry configurations and the safety of each technology. In each period, if only one firm deploys its technology, it acts as a monopolist for that period. Its payoff is then the same as in the monopolistic case discussed in the previous section and depends on whether its technology is safe. If both firms deploy their technologies, we assume that each firm's payoff is proportional to its market share; that is, relative to the monopoly case, the only change is that the market is shared between the two firms. This assumption facilitates comparison with the monopoly benchmark and isolates the effect of competition on deployment timing.

proactively investigating and disclosing such risks.

More specifically, in the first period, if both firms enter, symmetry implies that they each capture half of the market. Accordingly, a firm earns a payoff of $\frac{1}{2}$ if its technology is safe, and $-\frac{L}{2}$ if it is unsafe. To delineate the effect of competition on the timing of AI deployment time, we thus abstract from potential product differentiation and price competition.[17]

Consider the second period. If a firm entered the market in the first period, it will continue to operate in the second period only if its technology is revealed to be safe. If both firms entered in the first period and both technologies are safe, they again share the market equally in the second period. However, if both firms entered but only one technology is safe, the firm with the safe technology becomes the monopolist in the second period.

Now consider an asymmetric entry scenario in which only one firm enters the market in the first period. If the entrant's technology turns out to be safe, it remains active in the second period. If the other firm enters in the second period, we assume the presence of a first-mover advantage: the incumbent retains a market share of $k \in (1/2, 1)$, while the entrant captures the remaining share $(1 - k)$. The term $k - \frac{1}{2} > 0$ measures the strength of the first-mover advantage, which may arise from data-feedback loops, intelligence feedback loops, and scaling laws, as discussed above. Moreover, since a firm that enters in the first period remains active in the second period only if its technology is safe, whereas a rival entering in the second period typically faces a risky technology (except in the polar cases of perfect beta testing or perfect correlation), $k - \frac{1}{2} > 0$ can also be interpreted as a premium for a safe technology.

Let $\pi(x, y)$ denote a firm's expected payoff when its first-period entry decision is x, and the other firm's entry decision is y, assuming that both firms behave optimally in the second period, where $(x, y) \in \{0, 1\}^2$ with 0 and 1 indicating no entry and entry, respectively. We denote the joint profit of the two firms by $\Pi(x, y)$. With symmetric firms, the following relationship holds:

$$\Pi(x, y) = \pi(x, y) + \pi(y, x).$$

Similarly, let $W(x, y)$ denote the social welfare associated with the entry configuration in which one firm chooses x and the other chooses y in the first period.

The market equilibrium for the first-period entry configuration can be analyzed

[17] In the Appendix, we extend our analysis by considering an alternative payoff structure in which competition reduces profits and increases consumer surplus, and we discuss the robustness of our results and how they may change under the alternative payoff structure. See Appendix A.14.

using the following entry game matrix:

Firm 1\Firm 2	E	NE
E	$\pi(1,1), \pi(1,1)$	$\pi(1,0), \pi(0,1)$
NE	$\pi(0,1), \pi(1,0)$	$\pi(0,0), \pi(0,0)$

(1)

The conditions under which both firms enter (i.e., (1,1)), only one firm enters (i.e., either (1,0) or (0,1)), or no firm enters (i.e., (0,0)) emerge as equilibrium entry configurations in the first period are as follows:

(1,1) Equilibrium: $\pi(1,1) > \pi(0,1)$

(1,0) or (0,1) Equilibrium: $\pi(1,0) > \pi(0,0)$ and $\pi(0,1) > \pi(1,1)$

(0,0) Equilibrium: $\pi(0,0) > \pi(1,0)$

To analyze a firm's incentive to enter given the rival's entry decision, we introduce the following notation:

$$\begin{aligned} \Psi^1 &\equiv \pi(1,1) - \pi(0,1), \\ \Psi^0 &\equiv \pi(1,0) - \pi(0,0), \end{aligned}$$

where Ψ^1 (respectively, Ψ^0) denotes the firm's incentives to enter in the first period when the other firm enters (respectively, does not enter). The expression $\Psi^1 \equiv \pi(1,1) - \pi(0,1)$ captures the incentive to enter in the first period given that the rival also enters, and can therefore be interpreted as an early-entry incentive driven by the fear of missing out (FOMO). By contrast,$\Psi^0 \equiv \pi(1,0) - \pi(0,0)$ represents the preemption incentive (PI), that is, the incentive to enter in the first period in order to preempt the rival.

In a competitive setting, there are two potential sources of learning about the riskiness of the developed AI technology: one is through beta testing, and the other is by observing the rival firm's outcome after its deployment, which can provide information about the safety of a firm's own technology due to correlation. We examine how these two learning mechanisms interact under competition and compare it with what happens when the firms coordinate on the first-period deployment decisions to maximize the joint profit.

For this purpose, we introduce an assumption which is a necessary condition for $\Pi(0,0) > \Pi(1,1)$ to occur for some parameters.

Assumption 1. *The static payoff from entry $E^s \equiv (1-\lambda) - \lambda L$ is strictly negative.*

If $E^s > 0$, $\Pi(1,1) > \Pi(0,0)$ always holds for any parameter values as (1,1) generates a positive profit in the first period and the highest possible joint profit in the second period due to the perfect information about (ω_1, ω_2). Therefore, the firms never have a joint incentive to delay their deployments.

We introduce another assumption which makes a firm's second-period entry decision depend on the outcome of its beta testing, conditional on no firm entering in the first period:

Assumption 2. *Conditional on no firm entering in the first period, there exists a unique equilibrium in the second period, in which a firm enters if and only if its signal is ϕ. Mathematically, this assumption is equivalent to*

$$\rho < \frac{(\beta\lambda + 1)(1 - \lambda - L\lambda + L\beta\lambda)}{\beta\lambda(1 + L - L\beta)(1 - \lambda)}. \tag{2}$$

Assumption 2 is relevant only under Assumption 1. If Assumption 1 is violated, condition (2) is always satisfied: when the static payoff is non-negative, it is clearly optimal for a firm that has not entered in the first period to enter in the second period, provided that beta testing reveals no negative signal. Under Assumption 1, the R.H.S. of (2) is strictly increasing in β, implying that Assumption 2 is more likely to hold as beta testing becomes more precise.[18] By contrast, if Assumption 2 is violated—that is, if beta testing has very low informational value—its availability is largely irrelevant for firms' deployment decisions.

Our objective is to analyze how competition affects the optimal timing of AI deployment under risk uncertainty. To isolate the pure effect of competition on firms' deployment decisions and the associated inefficiencies, we conduct two comparisons: first, between the market equilibrium and a coordinated outcome in which firms can coordinate on their first-period entry decisions; and second, between the market equilibrium and the socially optimal outcome.

Relative to each benchmark, we identify two types of inefficiencies induced by competition: "excessive entry" and "insufficient entry," defined as follows.

Definition 1. ***Excessive entry*** *with respect to joint profit maximization (respectively, social welfare maximization) arises when the number of firms entering in the first period exceeds the number that maximizes joint profits (respectively, social welfare).* ***Insufficient entry*** *with respect to joint profit maximization (respectively, social welfare maximization) arises when the number of firms entering in the first period is smaller than the number that maximizes joint profits (respectively, social welfare).*

[18] The R.H.S. of (2) is equal to 0 at $\beta_1 \equiv 1 - \frac{1-\lambda}{L\lambda}$ and is equal to 1 at $\beta_2 \equiv \sqrt{1 - \frac{1-\lambda}{L\lambda}}$. Hence in the (β, ρ) space, the R.H.S. is an increasing curve which cuts the horizontal line $\rho = 0$ at $\beta = \beta_1$ and the horizontal line $\rho = 1$ at $\beta = \beta_2$.

We define a race to the bottom as a particular case of excessive entry.

Definition 2. *A **race to the bottom** with respect to joint profit maximization (respectively, social welfare maximization) arises if both firms enter in the first period, even though joint profits (respectively, social welfare) would be the highest if neither firm entered in that period.*

4 Analysis of the Duopoly Model: Competition vs. No Competition

In this section, we analyze the model introduced in Section 3 to isolate the effects of competition on the timing of AI deployment. We begin with preliminary analysis in Section 4.1 that establishes results used later to compare the equilibrium outcome with the joint-profit–maximizing benchmark. To build intuition, we first examine two special cases—perfect beta testing ($\beta = 1$) in Section 4.2 and perfect correlation ($\rho = 1$) in Section 4.3 —before proceeding to the general case in Section 4.4.

4.1 Preliminary Analysis

As a preliminary step toward deriving the market equilibrium, we examine a firm's incentives to enter in the first period given the rival's entry decision—specifically, how a deviation from no entry to entry affects its own profit. Based on this analysis, we then compare $\Pi(1,1)$ and $\Pi(1,0)$, which facilitates the characterization of the joint-profit–maximizing outcome.

Let us first consider the case of symmetric entry configurations. The individual firm's profits when both firms enter and when neither firm enters in the first period are given, respectively, as follows.

$$\begin{aligned}
\pi(1,1) &= \mu_{SS}\left(\frac{1}{2}+\frac{1}{2}\delta\right)+\mu_{SN}\left(\frac{1}{2}+\delta\right)-\left(\mu_{NS}+\mu_{NN}\right)\frac{1}{2}L \\
&= \mu_{SS}\left(\frac{1}{2}+\frac{1}{2}\delta\right)+\mu_{SN}\left(\frac{1}{2}+\delta\right)-\frac{\lambda}{2}L, \\
\pi(0,0) &= \delta\left[\frac{1}{2}\mu_{SS}+\mu_{SN}\left((1-\beta)\frac{1}{2}+\beta\right)-\mu_{NS}(1-\beta)\frac{1}{2}L-\mu_{NN}(1-\beta)L\left((1-\beta)\frac{1}{2}+\beta\right)\right] \\
&= \delta\left[\frac{1}{2}\mu_{SS}+\mu_{SN}\left((1-\beta)\frac{1}{2}(1-L)+\beta\right)-\mu_{NN}(1-\beta)L\left((1-\beta)\frac{1}{2}+\beta\right)\right],
\end{aligned}$$

where the last equality comes from $\mu_{SN} = \mu_{NS}$.

To understand the expressions above, for example, consider firm i's second-period profit conditional on $(\omega_i, \omega_j) = (S, N)$. In this state of nature, the firm earns a

monopoly profit of 1 when $(x, y) = (1, 1)$, since the rival exits the market, whereas it obtains $\beta + (1 - \beta)/2$ when $(x, y) = (0, 0)$, as the rival enters with probability $1 - \beta$. Note that, in computing $\pi(0, 0)$, we assume that the second-period subgame admits a unique equilibrium in which each firm enters if and only if its signal is ϕ, as specified in Assumption 2. Furthermore, Assumption 2 implies $\pi(0, 0) > 0$.

Now consider the case of an asymmetric entry configuration in which only one firm (say, firm i) enters in the fist period. By the end of that period, firm j (with $j \neq i$) observes both ω_i and its own signal σ_j. If $\sigma_j = N$, firm j does not enter in the second period, since the signal provides conclusive evidence that $\omega_j = N$.

If $\sigma_j = \phi$ and $\omega_i = S$, the two signals are consistent, and firm j's updated belief about the riskiness of its technology is lower than without observing the rival's state $\omega_i = S$, that is, $\widetilde{\lambda}(\phi, S) < \widetilde{\lambda}(\phi)$. Under our Assumption 2, firm j therefore enters in the second period because

$$1 - \tilde{\lambda}(\phi, S) > \tilde{\lambda}(\phi, S)L, \tag{3}$$

as established in the next lemma.

Lemma 4. *Under Assumption 2,* $\tilde{\lambda}(\phi, S) < \lambda_s^* \equiv \frac{1}{1+L}$.

By contrast, if $\sigma_j = \phi$ and $\omega_i = N$, the two signals provide conflicting information about the underlying state. In this case, firm j enters if and only if its updated belief $\tilde{\lambda}(\phi, N)$ satisfies $1 - \tilde{\lambda}(\phi, N) > \tilde{\lambda}(\phi, N)L$, or equivalently

$$\tilde{\lambda}(\phi, N) < \lambda_s^* \equiv \frac{1}{1 + L}. \tag{4}$$

The validity of condition (4) depends on the relative magnitudes of β and ρ. For instance, condition (4) is satisfied when β is sufficiently high (relative to ρ), so that the firm's own signal is informative enough to outweigh the contradictory information conveyed by the other firm's safety record, implying that the firm's entry decision depends only on its own signal and there is no learning from the rival's experimentation. By contrast, the condition is violated when ρ is sufficiently high (relative to β), in which case the other firm's record becomes more informative about the safety of the firm's own technology. This generates an incentive to free-ride on the rival's experimentation by delaying one's entry. Note that (4) is satisfied in the case of perfect beta testing and is violated in the case of perfect correlation.

In what follows, we distinguish between the case in which (4) holds and the case in which it is violated. The individual profits for $(x, y) = (1, 0)$ and $(x, y) = (0, 1)$ are

given by

$$\begin{aligned}\pi(1,0) &= \mu_{SS}\left(1+\delta k\right)+\mu_{SN}\left(1+\delta\left((1-\beta)k+\beta\right)\right)-\lambda L,\\ \pi(0,1) &= \delta\left[\mu_{SS}(1-k)-\mu_{NS}(1-\beta)(1-k)L\right)\\ &+ \delta[\mu_{SN}-\mu_{NN}(1-\beta)L]\mathbf{1}_{[\tilde{\lambda}(\phi,N)<\lambda_s^*]}.\end{aligned}$$

Note that $\pi(1,0)$ is independent of whether (4) is satisfied, because this condition affects j's second-period entry only when $\omega_i = N$, in which case firm i is inactive in the second period, making the condition irrelevant.

Summarizing, we have

Lemma 5. *Suppose Assumption 2. We have*

$$\begin{aligned}\pi(1,1) &= \mu_{SS}\left(\frac{1}{2}+\frac{1}{2}\delta\right)+\mu_{SN}\left(\frac{1}{2}+\delta\right)-\frac{\lambda}{2}L,\\ \pi(0,1) &= \delta\left[\mu_{SS}(1-k)-\mu_{NS}(1-\beta)(1-k)L\right)+\delta[\mu_{SN}-\mu_{NN}(1-\beta)L]\mathbf{1}_{[\tilde{\lambda}(\phi,N)<\lambda_s^*]},\\ \pi(1,0) &= \mu_{SS}\left(1+\delta k\right)+\mu_{SN}\left(1+\delta\left((1-\beta)k+\beta\right)\right)-\lambda L,\\ \pi(0,0) &= \delta\left[\frac{1}{2}\mu_{SS}+\mu_{SN}\left((1-\beta)\frac{1}{2}(1-L)+\beta\right)-\mu_{NN}(1-\beta)L\left((1-\beta)\frac{1}{2}+\beta\right)\right].\end{aligned}$$

Define $E^p \equiv \delta\mu_{SS}(k-1/2)$, which represents the gain from preemptive entry under perfect beta testing (i.e., $\beta = 1$). Under perfect beta testing, if only one firm enters in the first period, it faces the rival's entry in the second period with probability μ_{SS} and enjoys the first-mover advantage of $(k-1/2)$. From Lemma 5, we obtain the following result on a firm's entry incentives conditional on the rival's decision.

Lemma 6. *Suppose Assumption 2. We have*

$$\begin{aligned}\Psi^1 &= \begin{cases}\frac{1}{2}E^s+E^p+\delta(1-\beta)L\left(\mu_{NS}(1-k)+\mu_{NN}\right), \textit{ if (4) holds},\\ \frac{1}{2}E^s+E^p+\delta\left(\mu_{SN}+\mu_{NS}(1-\beta)L(1-k)\right), \textit{otherwise},\end{cases}\\ \Psi^0 &= E^s+E^p+\delta\mu_{SN}(1-\beta)(k-\frac{1}{2})\\ &+ \delta(1-\beta)L\left[\mu_{SN}+\mu_{NN}\left((1-\beta)\frac{1}{2}+\beta\right)\right].\end{aligned}$$

Regarding Ψ^1, the first two terms are composed of $E^s/2+E^p$. When (4) holds, the last term of Ψ^1 represents the expected liability from entering in the second period when the rival has already entered in the first period. When (4) does not hold, observing the rival's failure induces the firm to exit the market in the second period. This affects Ψ^1 in two ways. First, the expected liability decreases because the firm

exits even if its own beta test yields the null signal ϕ. Second, the loss from the rival's preemptive entry increases. To see this, consider $(\omega_1, \omega_2) = (S, N)$ with firm 2 entering in the first period. If firm 1 also enters in the first period, it earns a profit of 1 in the second period. But if it waits, it observes the rival's failure and exits. This explains the additional term $\delta\mu_{SN}$.

Regarding Ψ^0, the first two terms are composed of $E^s + E^p$, while the last term represents the expected liability from entering in the second period when no firm entered in the first period. The third term, $\delta\mu_{SN}(1-\beta)(k-1/2)$, captures the additional gain from preemptive entry under imperfect beta testing: when $(\omega_1, \omega_2) = (S, N)$, a firm that enters alone in the first period faces the rival's entry in the second period with probability $(1-\beta)$ and thus enjoys the first-mover advantage $k - 1/2$.

From Lemma 6, if Assumption 1 does not hold, entry in the first period is a dominant strategy for each firm, as both $\Pi^1 > 0$ and $\Pi^0 > 0$.

Finally, we provide a useful result on the comparison between $\Pi(1,1)$ and $\Pi(1,0)$.

Lemma 7. *Under Assumption 2, we have*

$$\Pi(1,1) \geq \Pi(1,0)$$

with equality if and only if $\rho = 1$ *or* $\beta = 1$.

The result follows from the positive value of market experimentation. More specifically, experimenting with both technologies and experimenting with only one generate the same joint profit in the first period, whereas the second-period joint profit is higher in the former case than in the the latter. Equality arises only in the extreme case of perfect correlation—when the second experiment yields no additional information—or perfect beta testing—when experimentation provides no incremental value.

By the virtue of Lemma 7, it therefore suffices to compare $\Pi(1,1)$ and $\Pi(0,0)$ to identify the first-period entry configuration that maximizes the joint profit.[19] To analyze the firms' joint incentive to enter in the fist period, we introduce the following notation:

$$\Psi^J \equiv \Pi(1,1) - \Pi(0,0).$$

In Section 4.3 (respectively, Section 4.4), we fix the values of other parameters, study the market outcome in the (β, λ) space (respectively, the (β, ρ) space) and compare it with the joint–profit–maximizing outcome. Our analysis crucially depends on the shapes of the following four curves in the (β, λ) space (respectively, the (β, ρ) space): $\pi(0,0) = 0$, $\Psi^1 = 0$, $\Psi^0 = 0$, $\Psi^J = 0$. We label these four curves as the A2

[19] Since the joint-profit–maximizing entry configuration is symmetric, the firms do not need to make monetary transfers to achieve it; they simply need to coordinate on either $(x, y) = (0, 0)$ or $(x, y) = (1, 1)$.

curve, the FOMO curve, the PI curve, the JP curve, respectively, and assign a distinct color to each in the figures below.

- The curve defined by $\pi(0,0) = 0$ is referred to as the **A2** curve and is depicted in black.
- The curve defined by $\Psi^0 = 0$ is referred to as the **PI** (preemption incentive) curve and is depicted in orange.
- The curve defined by $\Psi^1 = 0$ is referred to as the **FOMO** (fear of missing out) curve and is depicted in red.
- The curve defined by $\Psi^J = 0$ is referred to as the **JP** (joint profit maximization) curve and is depicted in blue.

4.2 Special Case of Perfect Best Testing

We here study the case of $\beta = 1$. In this case, Assumption 2 and (4) are satisfied. This case is special because first-period experimentation does not generate any additional information beyond that obtained through beta testing.

From Lemma 5, the individual profit $\pi(x,y)$ with $(x,y) \in \{0,1\}^2$ is given as follows.

$$\begin{aligned}
\pi(1,1) &= \frac{1}{2}\left\{(1-\lambda)\left[1+\delta(1+\lambda(1-\rho))\right]-\lambda L\right\},\\
\pi(0,1) &= \delta\left\{\left[(1-\lambda)^2+\rho\lambda(1-\lambda)\right](1-k)+\lambda(1-\lambda)(1-\rho)\right\},\\
\pi(1,0) &= (1-\lambda)-\lambda L+\delta\left\{\left[(1-\lambda)^2+\rho\lambda(1-\lambda)\right]k+\lambda(1-\lambda)(1-\rho)\right\},\\
\pi(0,0) &= \delta\left\{\left[(1-\lambda)^2+\rho\lambda(1-\lambda)\right]\frac{1}{2}+\lambda(1-\lambda)(1-\rho)\right\}.
\end{aligned}$$

From Lemma 6, we have the following result on Ψ^i for $i = 0, 1$.

$$\begin{aligned}
\Psi^1 &= \frac{1}{2}E^s + E^p,\\
\Psi^0 &= E^s + E^p.
\end{aligned}$$

Note that under Assumption 1, a firm's incentive to enter in the first period is stronger when the rival also enters than when the rival delays, because the firms share the first-period loss when both enter.

Regarding the first-period entry decisions maximizing the joint profit, we find

$$\Psi^J \equiv \Pi(1,1) - \Pi(0,0) = \Pi(1,0) - \Pi(0,0) = E^s,$$

which is negative under Assumption 1. When $\beta = 1$, any firm which did not enter in the first period enters in the second period if the beta testing is successful. This means that the second-period joint profit is the same regardless of the number of the firms which enter in the first period. In addition, as long as at least one firm enters in the first period, the first period joint profit does not depend on the number of the firms which enter. Note that none of the joint profit differences $\Pi(1,1) - \Pi(0,0)$ and $\Pi(1,0) - \Pi(0,0)$ includes any value of experimentation as the experimentation does not generate any additional information relative to the perfect beta testing.

Therefore, we obtain:

Proposition 4. *Suppose Assumption 1 and $\beta = 1$. Then, joint-profit maximization requires no firm to enter in the first period. The equilibrium depends on the relative magnitudes of $|E^s|$ and E^p.*
(i) (race to the bottom) If $E^p > |E^s|$, it is a dominant strategy for each firm to enter in the first period, which generates a race to the bottom.
(ii) (FOMO-driven race to the bottom) If $|E^s| > E^p > \frac{|E^s|}{2}$, we have multiple equilibria: In one equilibrium both firms enter the market in the first period while in the other equilibrium no one enters in the first period. The first equilibrium is Pareto-dominated by the second one.
(iii) If $E^p < \frac{|E^s|}{2}$, it is a dominant strategy for each firm not to enter in the first period, which maximizes the joint profit.

The main point of the above proposition is that when $E^s < 0$ (hence, no entry in the first period maximizes the joint profit), if the dynamic gain from preemptive entry E^p is higher than $|E^s|$, each firm finds it a dominant strategy to enter in the first period, which generates a race to the bottom (with respect to the joint profit maximization).[20] Note that a race to the bottom can also arise when $|E^s| > E^p > \frac{|E^s|}{2}$ holds (see Proposition 4(ii)). This outcome is driven by FOMO, as a firm has a stronger incentive to enter in the first period when its rival does so.

From the fact that $E^p \equiv \delta(1-\lambda)(1-\lambda+\rho\lambda)(k-\frac{1}{2})$, we obtain the following corollary:

Corollary 1. *Suppose Assumption 1 and $\beta = 1$. As the positive correlation ρ increases or as the first-mover advantage $k - \frac{1}{2}$ increases, competition is more likely to generate a race to the bottom.*

The result that a larger first-mover advantage makes a race to the bottom more likely is intuitive. Greater homogenization of AI technologies—captured by a higher

[20] If $E^s > 0$, from $E^p > 0$, there is no distortion as $\Psi^i > 0$ for $i = 1, 0, J$.

degree of correlation—increases the probability of the state of nature (S, S), making it more likely that both firms compete in the second-period market. This, in turn, raises the gains from preemptive entry and thereby increases the likelihood of a race to the bottom.

4.3 The case of perfect correlation

We here study the case of perfect correlation, that is $\rho = 1$. Then $\mu_{SS} = 1 - \lambda$, $\mu_{SN} = \mu_{NS} = 0$, $\mu_{NN} = \lambda$, and (4) is not satisfied. In this section, we dispense with Assumption 1, as imposing it does not simplify the analysis.

Under perfect correlation ($\rho = 1$), Assumption 2 can be restated as follows.

Assumption 2 (when $\rho = 1$). $\lambda < \lambda_{A2} \equiv \frac{1}{1+(1-\beta^2)L}$.

The condition $\lambda < \lambda_{A2}$ is equivalent to $\pi(0,0) > 0$. Note that Assumption 2 holds when β is high.

By Lemma 5, the individual firm profit $\pi(x, y)$, with $(x, y) \in \{0, 1\}^2$, simplifies as follows:

$$\begin{aligned}
\pi(1,1) &= \frac{1}{2}\left[(1-\lambda)(1+\delta) - \lambda L\right] = \frac{1}{2}\left[E^s + \delta(1-\lambda)\right], \\
\pi(0,1) &= (1-\lambda)\delta(1-k), \\
\pi(1,0) &= (1-\lambda)(1+\delta k) - \lambda L = E^s + \delta(1-\lambda)k, \\
\pi(0,0) &= \delta\left(\frac{1}{2}(1-\lambda) - \frac{\lambda}{2}L\left(1-\beta^2\right)\right) = \delta\left[\frac{1}{2}E^s + \frac{\lambda}{2}L\beta^2\right],
\end{aligned}$$

where the term $\frac{\lambda}{2}L\beta^2$ in $\pi(0,0)$ represents the reduction in liability costs attributable to beta testing.

Regarding a firm's unilateral incentive to enter in the first period given the rival's entry decision—i.e., $\pi(1, y) - \pi(0, y)$ for $y \in \{0, 1\}$—we obtain the following result:

Lemma 8. *Suppose perfect correlation ($\rho = 1$).*
(i) $\Psi^1 = \frac{1}{2}E^s + E^p > 0$ *if and only if* $\lambda < \lambda_{FOMO} \equiv \frac{(2k-1)\delta+1}{(2k-1)\delta+1+L}$. *Note that* λ_{FOMO} *is independent of* β.
(ii) $\Psi^0 = E^s + E^p + \delta\frac{\lambda}{2}L\left(1-\beta^2\right) > 0$ *if and only if* $\lambda < \lambda_{PI} \equiv \frac{1+(k-\frac{1}{2})\delta}{1+(k-\frac{1}{2})\delta+\frac{1}{2}(2-\delta+\beta^2\delta)L}$.
(iii) There exists $\widehat{\beta} \in (\sqrt{\frac{\delta}{\delta+1}}, 1)$ *such that* $\lambda_{FOMO} \gtreqless \lambda_{PI}$ *iff* $\beta \gtreqless \widehat{\beta}$. *In addition,* $\max\{\lambda_{FOMO}, \lambda_{PI}\} < \lambda_{A2}$ *for* $\beta > \sqrt{\frac{\delta}{1+\delta}}$.

According to Lemma 8(i), Ψ^1 is equal to $\frac{1}{2}E^s + E^p$ as in the case of perfect beta testing. Ψ^1 is positive if and only if $\lambda < \lambda_{FOMO}$. Notice that λ_{FOMO} is independent of the precision of beta testing because beta testing is not used when at least one firm

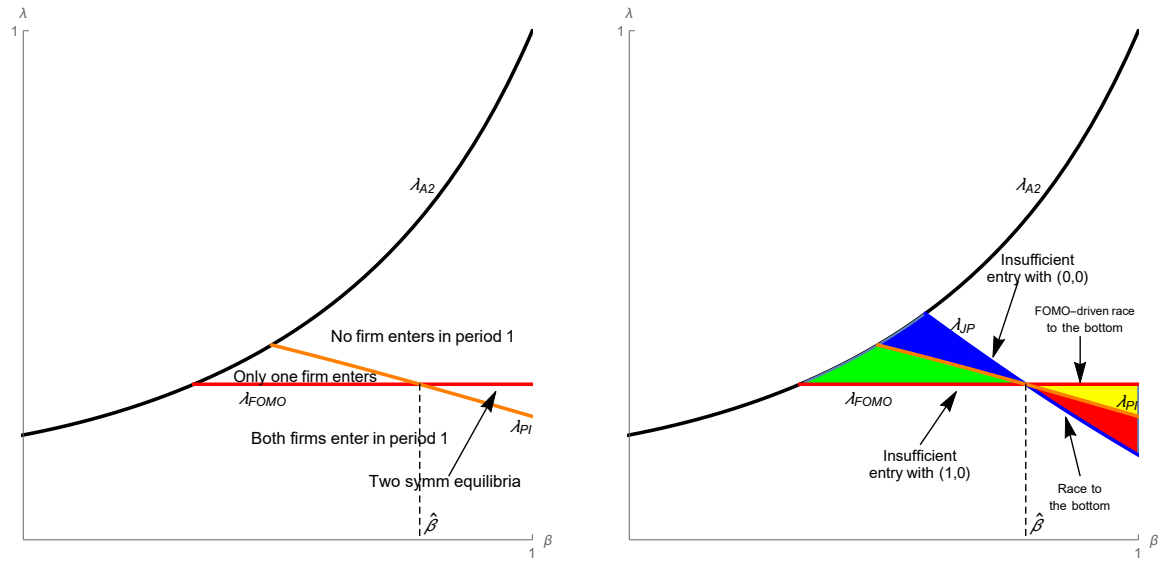
enters in period 1; then, from $\rho = 1$, observing safety or lack of safety of the rival's product is enough to get precise information about the safety of one's own product. If $\lambda > \lambda_{FOMO}$, then, given that the rival enters in the first period, a firm prefers to wait and free-ride on the rival's experimentation, as the rival's success or failure provides a perfect signal about the safety of its own technology.

According to Lemma 8(ii), we have

$$\Psi^0 = E^s + E^p + \delta \frac{\lambda}{2} L \left(1 - \beta^2\right).$$

The first two terms coincide with those under perfect beta testing. The last term arises under imperfect beta testing and captures the expected liability cost per firm when each firm enters in the second period upon receiving the signal ϕ. The expression Ψ^0 is positive if and only if $\lambda < \lambda_{PI}$.

According to Lemma 8(iii), there exists $\widehat{\beta} \in (\sqrt{\frac{\delta}{\delta+1}}, 1)$ such that $\lambda_{FOMO} \gtreqless \lambda_{PI}$ if and only if $\beta \gtreqless \widehat{\beta}$. In addition, $\max\{\lambda_{FOMO}, \lambda_{PI}\} < \lambda_{A2}$ for $\beta > \sqrt{\frac{\delta}{1+\delta}}$.[21]



(a) Equilibria (b) Distortions

Figure 1: Equilibria and Distortions in the Case of Perfect Correlation

Under Assumption 2, we have four possible cases regarding the signs of Ψ^1 and Ψ^0. The next proposition describes the equilibrium (or the equilibria) for each of the four cases (see also Figure 1(a)):

Proposition 5. *In the case of perfect correlation with $\rho = 1$, under Assumption 2, the structure of equilibria is as follows:*

[21] $\lambda_{FOMO} < \lambda_{A2}$ holds if and only if $\beta > \sqrt{\frac{(2k-1)\delta}{1+(2k-1)\delta}}$ and that $\lambda_{PI} < \lambda_{A2}$ holds if and only if $\beta > \sqrt{\frac{k\delta}{1+k\delta}}$, and notice that $\sqrt{\frac{(2k-1)\delta}{1+(2k-1)\delta}} < \sqrt{\frac{k\delta}{1+k\delta}} < \sqrt{\frac{\delta}{1+\delta}}$.

(i) Case 1: For $\lambda > \max\{\lambda_{FOMO}, \lambda_{PI}\}$, there exists a unique equilibrium in which no firm enters in the first period.

(ii) Case 2: For $\lambda \in (\lambda_{PI}, \lambda_{FOMO})$ (which requires $\beta > \widehat{\beta}$), there are two symmetric equilibria: either both firms enter in the first period or no firm enters in the first period.

(iii) Case 3: For $\lambda \in (\lambda_{FOMO}, \lambda_{PI})$ (which requires $\beta < \widehat{\beta}$), there are two asymmetric equilibria in pure strategies and, in both equilibria, only one firm enters in the first period.

(iv) Case 4: For $\lambda < \min\{\lambda_{FOMO}, \lambda_{PI}\}$, there exists a unique equilibrium in which both firms enter in the fist period.

Regarding the joint profit maximization, we have:

$$\Psi^J \equiv \Pi(1,1) - \Pi(0,0) = E^s(1-\delta) + \delta((1-\lambda) - \lambda L\beta^2),$$

where $\Psi^J > 0$ if and only if $\lambda < \frac{1}{1+(1-\delta+\beta^2\delta)L} \equiv \lambda_{JP}$. To provide intuition for the condition $\Psi^J > 0$, consider the case in which $\delta = 1$. Then, the key trade-off is between the gain from deploying a safe technology in period two following first-period experimentation (captured by $\delta(1-\lambda)$) and the savings in second-period liability costs due to beta testing in the absence of first-period experimentation (captured by $\delta\lambda L\beta^2$). Recall that $\Pi(1,1) = \Pi(1,0)$ in the case of perfect correlation, as shown in Lemma 7. We also note that $\lambda_{JP} < \lambda_{A2}$ if and only if $\beta > \sqrt{\frac{\delta}{1+\delta}}$.

We now compare the market equilibrium with the joint-profit–maximizing outcome. Proposition 5 shows that the market equilibrium is characterized by two threshold values of λ—λ_{PI} and λ_{FOMO}—whereas the joint-profit maximizing entry configuration is characterized by λ_{JP}. The following lemma on the relative magnitudes of these threshold values facilitate the comparison between the market equilibrium and the joint-profit maximizng entry configurations (see also Figure 1(b)).

Lemma 9. *There exists a unique $\widehat{\beta} \in (\sqrt{\frac{\delta}{\delta+1}}, 1)$ such that $\lambda_{PI}(\widehat{\beta}) = \lambda_{FOMO}(\widehat{\beta}) = \lambda_{JP}(\widehat{\beta})$.*
(i) If $\beta < \widehat{\beta}$, we have $\lambda_{FOMO} < \lambda_{PI} < \lambda_{JP}$.
(ii) If $\beta > \widehat{\beta}$, we have $\lambda_{FOMO} > \lambda_{PI} > \lambda_{JP}$.

Based on the relative ranking of the two threshold values of λ that govern each firm's unilateral entry incentives in the first period (λ_{PI} and λ_{FOMO}) and the threshold λ_{JP} that characterizes the joint profit maximizing entry configuration, we obtain the following proposition identifying the potential divergence of the market equilibrium from the joint-profit–maximizing outcome (see Figure 1(b)).

Proposition 6. *In the case of perfect correlation with $\rho = 1$, under Assumption 2, the market outcome maximizes the joint profit but for the following cases.*

(i) Race to the bottom

For $\lambda \in (\lambda_{JP}, \lambda_{FOMO})$, which requires $\beta > \widehat{\beta}$, the joint profit maximization requires no firm to enter in the first period but

(a) for $\lambda \in (\lambda_{JP}, \lambda_{PI})$, in the unique equilibrium, both firms enter in the first period, leading to a race to the bottom;

(b) for $\lambda \in (\lambda_{PI}, \lambda_{FOMO})$, there exists an equilibrium in which both firms enter in the first period, leading to a FOMO-driven race to the bottom.

(ii) Insufficient entry

For $\lambda \in (\lambda_{PI}, \lambda_{JP})$, which requires $\beta < \widehat{\beta}$, the social optimum calls for one or both firms to enter in the first period, whereas in the unique equilibrium no firm enters in the first period.

In addition to the race to the bottom—which arises when beta testing is sufficiently precise and the probability of risk is low—an opposite distortion of insufficient entry can emerge when beta testing is less informative and the probability of risk is high. This distortion is driven by firms' incentives to free-ride on rivals' experimentation. Note that for $\lambda \in (\lambda_{FOMO}, \lambda_{PI})$, any pure-strategy equilibrium involves only one firm entering in the first period. However, this does not generate inefficiency, since, under perfect correlation, $\Pi(1,0) = \Pi(1,1)$ as shown in Lemma 7.

4.4 The general case of imperfect correlation

We now analyze the general case with imperfect correlation and imperfect beta testing. In this setting, a complete characterization of the market equilibrium and its comparison to the joint–profit–maximizing outcome becomes considerably more complex, due to the presence of multiple parameters that jointly determine the equilibrium.

In the analysis below, we focus on how the key parameters—$(\lambda, \rho, \beta, k)$— affect the market equilibrium by setting $L = \delta = 1$. $\delta = 1$ means that payoffs in the two periods are weighted equally. In contrast, if $1/\delta$ is close to zero (i.e., δ is very large), welfare (and joint profit) maximization would require (1,1) since the cost of early experimentation becomes negligible. We do not focus on this scenario. Note that when $\delta = 1$, Assumption 1 is equivalent to $\lambda > 1/2$. $L = 1$ implies $L = B$ as we normalized B to 1. We also compare the market outcome with the joint–profit–maximizing benchmark. In the next sections, we compare the market outcome with the social-welfare–maximizing benchmark and discuss policy implications.

Applying $L = \delta = 1$ to the profit functions in Lemma 5 yields

$$\begin{aligned}
\pi(1,1) &= \mu_{SS}\left(\frac{1}{2}+\frac{1}{2}\right)+\mu_{SN}\left(\frac{1}{2}+1\right)-\frac{\lambda}{2},\\
\pi(0,1) &= \mu_{SS}(1-k)-\mu_{NS}(1-\beta)(1-k))\\
&+ [\mu_{SN}-\mu_{NN}(1-\beta)]\mathbf{1}_{[1-\tilde{\lambda}(\phi,N)>\tilde{\lambda}(\phi,N)]},\\
\pi(1,0) &= \mu_{SS}\left(1+k\right)+\mu_{SN}\left(1+\delta\left((1-\beta)k+\beta\right)\right)-\lambda,\\
\pi(0,0) &= \frac{1}{2}\mu_{SS}+\mu_{SN}\beta-\mu_{NN}(1-\beta)\left((1-\beta)\frac{1}{2}+\beta\right).
\end{aligned}$$

Similarly, applying $L = \delta = 1$ to the profit differences in Lemma 6 yields

$$\begin{aligned}
\Psi^1 &= \begin{cases} \frac{1}{2}E^s+E^p+(1-\beta)\left(\mu_{NS}(1-k)+\mu_{NN}\right), \text{if (4) holds,}\\ \frac{1}{2}E^s+E^p+\left(\mu_{SN}+\mu_{NS}(1-\beta)(1-k)\right), \text{otherwise,}\end{cases}\\
\Psi^0 &= E^s+E^p+\mu_{SN}(1-\beta)(k-\frac{1}{2})\\
&+ (1-\beta)\left[\mu_{SN}+\mu_{NN}\left((1-\beta)\frac{1}{2}+\beta\right)\right].
\end{aligned}$$

In our analysis, we fix a point $(\lambda, k) \in (1/2, 1)^2$ and study the market outcome in the (β, ρ) space (recall that $(\beta, \rho) \in [0, 1]^2$), comparing it with the joint–profit–maximizing outcome. Our analysis crucially depends on the shapes of the four curves in the (β, ρ) space: the A2 curve, the FOMO curve, the PI curve, the JP curve. Figure 2 illustrates these curves for a (λ, k) satisfying $\frac{\lambda}{2-2\lambda} > k$ and $\lambda < \frac{2}{3}$. The region to the right of the A2 curve (in black) satisfies Assumption 2. The FOMO curve (in red) corresponds to $\Psi^1 = 0$, which governs a firm's incentive to enter when its rival also enters. The PI curve (in orange) corresponds to $\Psi^0 = 0$, which governs a firm's incentive to enter when its rival does not enter: to the left (respectively, right) of the PI curve, a firm has an incentive to enter (respectively, wait) in the first period when its rival does not enter. The JP curve (in blue) represents $\Psi^J = 0$: to the left (respectively, right) of this curve, joint profits are maximized when both firms (respectively, no firm) enter in the first period. Finally, Figure 2 contains also a brown curve which traces the points (β, ρ) satisfying $\tilde{\lambda}(\phi, N) = \lambda_s^*$ such that (4) holds (does not hold) if (β, ρ) is below (above) the brown curve.

Note that the FOMO curve consists of two connected segments, depending on whether condition (4) is satisfied or not. When (4) holds, the curve is relatively vertical; when it is violated, the curve becomes relatively horizontal. Accordingly, we refer to these segments as the vertical FOMO curve and the horizontal FOMO curve, respectively. To the left (respectively, right) of the vertical FOMO curve, and below (respectively, above) the horizontal FOMO curve, a firm has an incentive to enter

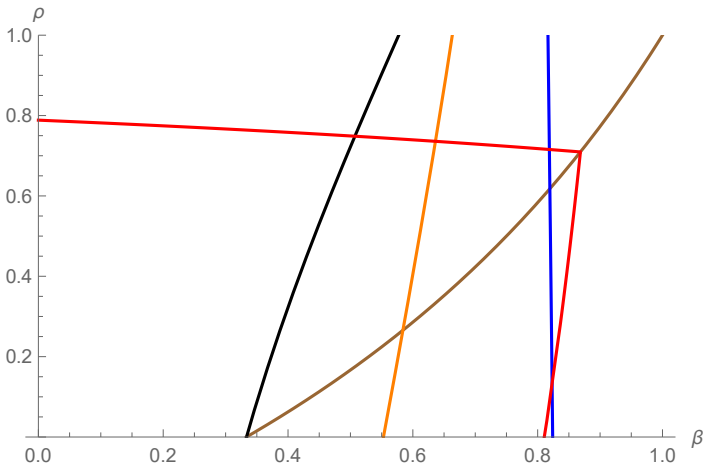


Figure 2: Illustration of different curves for a (λ, k) satisfying $\frac{\lambda}{2-2\lambda} > k$ and $\lambda < \frac{2}{3}$.

(respectively, wait) in the first period when its rival also enters.

It is useful to partition $(\lambda, k) \in (1/2, 1)^2$ into six regions, depending on whether each of the following curves —the PI curve, the vertical FOMO curve, and the horizontal FOMO curve —appears in the relevant (β, ρ) space (i.e., the area satisfying Assumption 2).[22]

- Case 1 : $k > \frac{2\lambda+\lambda^2-1}{2(1-\lambda)^2}$. In this case, $\Psi^1 > 0$ and $\Psi^0 > 0$.
- Case 2: $\frac{2\lambda+\lambda^2-1}{2(1-\lambda)^2} > k > \frac{\lambda^2}{2(1-\lambda)^2}$. In this case, $\Psi^1 > 0$ and the PI curve appears.
- Case 3: $\frac{\lambda^2}{2(1-\lambda)^2} > k > \frac{\lambda}{2-2\lambda}$. In this case, only the PI and the vertical FOMO curves appear.
- Case 4: $\frac{\lambda}{2-2\lambda} > k$ and $\lambda < \frac{2}{3}$. In this case, the PI, the vertical FOMO and the horizontal FOMO curves appear.
- Case 5: $2/3 \leq \lambda < \sqrt{3} - 1$. In this case, $\Psi^0 < 0$ and only the vertical FOMO and the horizontal FOMO curves appear.
- Case 6: $\sqrt{3} - 1 < \lambda < 1$. In this case, $\Psi^1 < 0$ and $\Psi^0 < 0$.

Conditional on that we select the Pareto-dominant equilibrium in the presence of multiple equilibria, we obtain the following results. There are three regimes:

(i) The race to the bottom regime: This regime is defined as Cases 1-3 (i.e., the upper-left region of Figure 3) as a race to the bottom inefficiency arises for each (λ, k) in the regime.

[22]For instance, the PI curve "does not appear" if the condition defining it is irrelevant—that is, if either $\Psi^0 > 0$ or $\Psi^0 < 0$ holds throughout the relevant (β, ρ) space. By contrast, the PI curve "appears" if both regions satisfying $\Psi^0 > 0$ and $\Psi^0 < 0$ coexist within the relevant (β, ρ) space.

(*ii*) The insufficient entry regime: This regime is defined as Cases 4-5 (i.e., the lower-left region of Figure 3) as insufficient entry arises for each (λ, k) in the regime.

(*iii*) The no distortion regime: This regime is defined as Case 6 (i.e., to the right of the vertical line in Figure 3) as there is no distortion for any (λ, k) in the regime.

By the occurrence of a race to the bottom or insufficient entry in the above definition of regimes, we mean that for each point $(\lambda, k) \in (1/2, 1)^2$, there exists a corresponding region in the (β, ρ) space where the respective distortion arises. In Case 6, the risk is sufficiently high that no firm enters in the first period, regardless of whether firms compete or coordinate.

In order to understand our result, fix $k \in (1/2, 1)$ and increase λ from a value close to 1/2. As (λ, k) moves across Cases 1 through 6, the nature of the distortion changes sequentially: first, a race to the bottom arises; next, insufficient entry emerges; and finally, no distortion occurs.

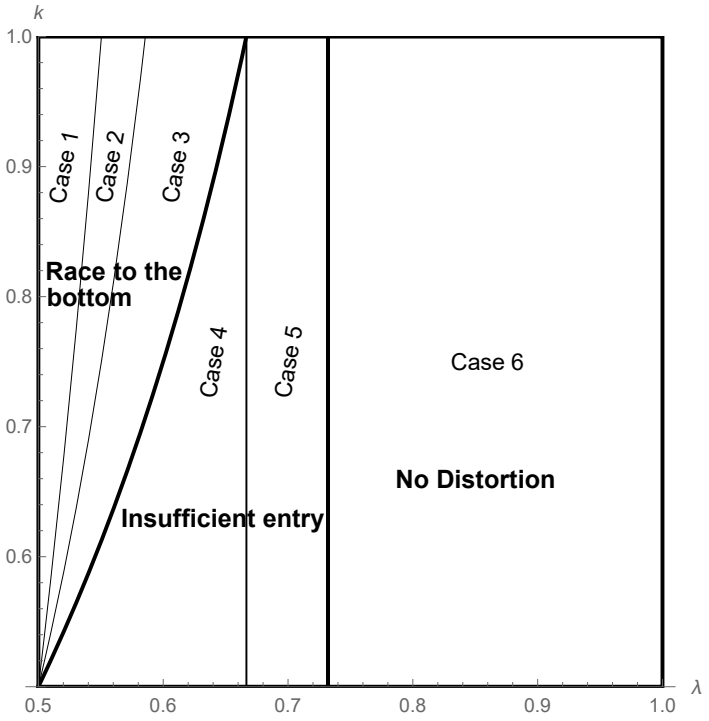


Figure 3: Partition of the set (λ, k)

The next lemma is about Assumption 2.

Lemma 10. *(Assumption 2)* $\pi(0, 0)$ *is independent of* k*, increases in* β*, decreases in* λ *and* ρ*. The A2 curve moves to the right in the space of* (β, ρ) *as* λ *increases.*

The next three lemmas provide useful properties of Ψ^1, Ψ^0 and Ψ^J, respectively.

Lemma 11. *(FOMO)* Ψ^1 *increases in* k*, decreases in* β *and* λ*.* Ψ^1 *increases (respectively, decreases) in* ρ *when (4) is satisfied (respectively, not satisfied).* $\Psi^1 > 0$ *for all* (β, ρ) *satisfying Assumption 2 in cases 1 and 2. For case 3,* $\Psi^1 > 0$ *for each* (β, ρ) *which violates (4). For cases 4 and 5, both the vertical and horizontal FOMO curves appear. As* λ *increases, the FOMO curve moves to the left in the space of* (β, ρ) *and in Case 6,* $\Psi^1 < 0$ *for all* (β, ρ) *satisfying Assumption 2.*

Lemma 12. *(PI)* Ψ^0 *increases in* k *and* ρ*, decreases in* β *and* λ*.* $\Psi^0 > 0$ *for all* (β, ρ) *satisfying Assumption 2 in Case 1. As* λ *increases, the PI curve moves to the left in the space of* (β, ρ) *so that in cases 5 and 6,* $\Psi^0 < 0$ *for all* (β, ρ) *satisfying Assumption 2.*

Lemma 13. *(Joint Profit) Under Assumption 2,* Ψ^J *is independent of* k*, decreases in* β *and* ρ *and* λ*. As* λ *increases, the JP curve moves to the left in the space of* (β, ρ) *so that in cases 1-3, the JP curve is on the right side of the A2 curve but in Case 6,* $\Psi^J < 0$ *for all* (β, ρ) *satisfying Assumption 2.*

The comparative statics of Ψ^1 and Ψ^0 with respect to k and λ are intuitive: both increase in k and decrease in λ. A larger first-mover advantage raises the incentive to enter in the first period, whereas a higher safety risk reduces it. As a result, for a given $\lambda \leq 2/3$, an increase in k makes a race to the bottom more likely. Conversely, for a given k, an increase in λ reduces the likelihood of a race to the bottom, as illustrated in Figure 3.

As Ψ^i decreases in λ (respectively, in β) for each $i = 1, 0, J$, there is an alignment between the individual incentive and the joint incentive regarding how an increase in λ (respectively, in β) affects first-period entries. As Ψ^i increases in k for each $i = 1, 0$ but Ψ^J is independent of k, there can be a conflict between the individual incentive and the joint incentive. When (4) is satisfied, as Ψ^i increases in ρ for each $i = 1, 0$ but Ψ^J decreases in ρ, there can be a conflict between the individual incentive and the joint incentive. We will show that this makes the race to the bottom more likely as ρ increases (see Proposition 9(iii)).

While there are many cases, Cases 1, 2, and 6 are relatively straightforward, and Case 5 is a subcase of Case 4. Accordingly, we focus on Cases 3 and 4; a complete analysis of the remaining cases is provided in the Appendix (see Appendix A.8 and A.9).

The key distinction between Case 3 and Case 4 is as follows. Conditional on (4) being violated, that is conditional on (β, ρ) above the brown curve, we have $\Psi^1 > 0$ throughout Case 3, whereas in Case 4, both $\Psi^1 > 0$ and $\Psi^1 < 0$ arise depending on the parameter values. Equivalently, the horizontal FOMO curve appears only in Case

4, while the vertical FOMO curve appears in both cases. This difference stems from the fact that Ψ^1 increases with k.

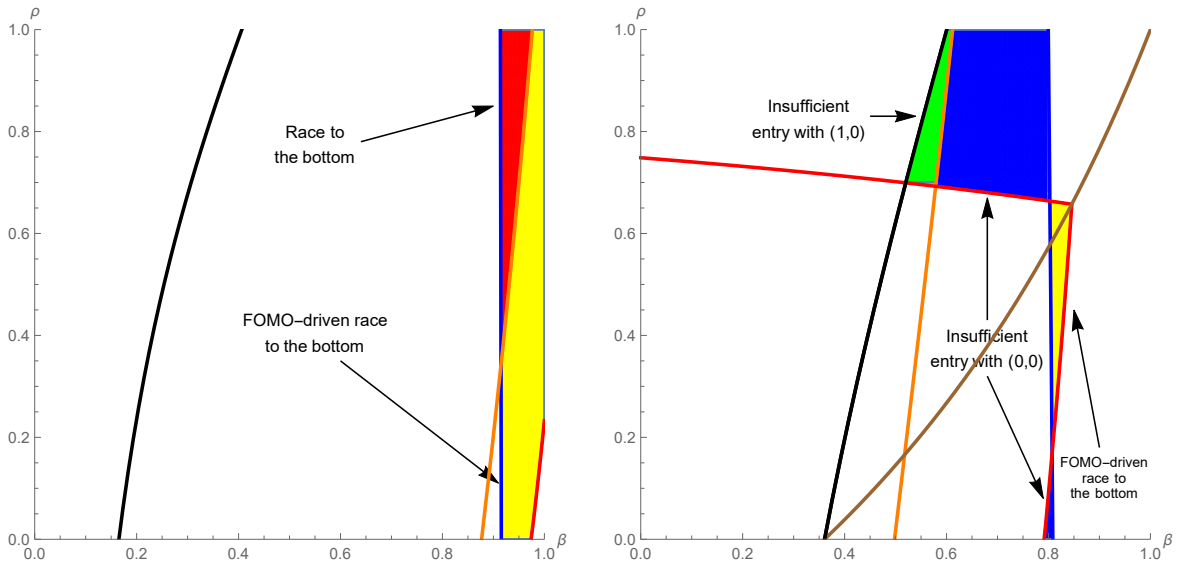
Comparing the PI curve with the FOMO curve, we obtain the following result (see Figure 4(a) for Case 3 and Figure 4(b) for Case 4):

Lemma 14. *(i) The PI curve lies to the left of the vertical FOMO curve. Formally, whenever (4) holds, $\Psi^0 > 0$ implies $\Psi^1 > 0$.*

(ii) In Case 4, the horizontal FOMO curve intersects the PI curve.

When (4) holds, the signal from beta testing is more informative than the information revealed by the rival's deployment. In this case, Lemma 14(i) implies that the rival's entry in the first period increases a firm's incentive to enter; that is, the FOMO effect dominates the PI effect in driving first-period entry. The intuition is that, under condition (4), the rival's entry has no bearing on the firm's second-period entry decision. As a result, a firm has no incentive to free-ride on the information generated by the rival's first-period entry, and the potential informational loss associated with FOMO—namely, the forgone opportunity to learn from the rival's first-period deployment—is eliminated.

As k decreases when moving from Case 3 to Case 4, Ψ^1 declines and may become negative for sufficiently high ρ, where (4) is violated. Since the PI curve does not depend on whether (4) holds, it is possible to have $\Psi^0 > 0 > \Psi^1$. Based on Lemma 14, we obtain the following equilibrium configurations in Cases 3-4 (see Figures 4(a) and (b)):



(a) Case 3: Race to the Bottom

(b) Case 4: Insufficient Entry

Figure 4: Case 3 and Case 4

Proposition 7. *Suppose that Assumptions 1 and 2 hold, and that $\delta = L = 1$.*

(i) In Case 3, (1,1) is the unique equilibrium to the left of the PI curve (in orange), while (0,0) is the unique equilibrium to the right of the FOMO curve (in red). There are multiple equilibria between the two curves with both (1,1) and (0,0) arising as equilibria.

(ii) In Case 4, (1,1) is the unique equilibrium in the region below the horizontal FOMO curve (in red) and to the left of the PI curve (in orange). In contrast, (1,0) is the unique pure-strategy equilibrium to the left of the PI curve and above the horizontal FOMO curve. In the region between the PI curve and the vertical FOMO curve, but below the horizontal FOMO curve, both (1,1) and (0,0) arise as equilibria. Finally, (0,0) is the unique equilibrium to the right of the curve defined by the pointwise maximum of the PI and FOMO curves.

In Case 3, to the left of the PI curve, entry in the first period is a dominant strategy, whereas to the right of the FOMO curve, waiting is a dominant strategy. This establishes Proposition 7(i). When we move to Case 4, this configuration is largely preserved, except in the region with sufficiently high ρ where (4) is violated. In that region, we have $\Psi^0 > 0 > \Psi^1$ to the left of the PI curve and above the horizontal FOMO curve. This yields (1,0) as the unique equilibrium in pure strategies.

We now turn to the comparison between the market outcome and the joint–profit–maximizing outcome. Comparing the JP curve with the PI and FOMO curves leads to the following lemma (see Figures 4(a) and (b)):

Lemma 15. *(i) In Case 3, the PI curve (in orange) lies to the right of the JP curve (in blue) at $\rho = 1$ and intersects it. The FOMO curve (in red) lies entirely to the right of the JP curve.*

(ii) In Case 4, the PI curve (in orange) lies entirely to the left of the JP curve (in blue). The horizontal FOMO curve intersects both the PI and JP curves. The intersection point between the vertical and horizontal FOMO curves is located to the right of the JP curve. The vertical FOMO curve may or may not intersect the JP curve.

In Case 3, Lemma 15(i) implies that, for sufficiently high ρ, the PI curve lies to the right of the JP curve. In this region, a race to the bottom inefficiency arises between the two curves: to the right of the JP curve, (0,0) maximizes joint profits, whereas to the left of the PI curve, (1,1) is the unique equilibrium. As k decreases, the PI curve shifts to the left (see Lemma 12), while the JP curve is independent of k (see Lemma 13). Consequently, in Case 4, the PI curve lies entirely to the left of the JP curve, eliminating the race-to-the-bottom region. Instead, two types of insufficient entry arise to the left of the JP curve and above the horizontal FOMO curve. In

this region, (1,1) maximizes joint profits, but the market equilibrium differs: (0,0) is the unique equilibrium between the PI and JP curves, while (1,0) is the unique pure-strategy equilibrium between the A2 (in black) and PI curves.

Summarizing, we have:

Proposition 8. *Suppose that Assumptions 1 and 2 hold, and that* $\delta = L = 1$.

(i) Case 3: For sufficiently high ρ*, the PI curve (in orange) lies to the right of the JP curve (in blue). In this region, a race-to-the-bottom inefficiency arises between the two curves. This region extends up to (and includes)* $\rho = 1$.

(ii) Case 4: For sufficiently high ρ*, the horizontal FOMO curve (in red) intersects both the PI and JP curves. In this case, two types of insufficient entry arise in the region to the left of the JP curve and above the horizontal FOMO curve. In this region, (1,1) maximizes joint profits, but the market equilibrium differs: (0,0) is the unique equilibrium between the PI and JP curves, while (1,0) is the unique pure-strategy equilibrium between the A2 (in black) and PI curves. This region extends up to (and includes)* $\rho = 1$.

Note that the region of the race to the bottom and the region of the insufficient entry described in the above proposition include $\rho = 1$, consistent with the results obtained in Section 4.3.[23]

Remark 1. *[FOMO-driven race to the bottom] Although we chose the Pareto-dominant equilibrium in this subsection, this was mainly to shorten the exposition. It is worthwhile to mention that when we include the FOMO-driven race to the bottom, this significantly expands the region in which the race to the bottom arises. Whenever the vertical FOMO curve appears (i.e., in cases 3-5), from Lemma 14(i), the PI curve lies to its left. Then, there are multiple equilibria of (1,1) and (0,0) between the PI curve and the vertical FOMO curve (and below the horizontal FOMO curve if it exists) and hence the FOMO-driven race to the bottom can arise. For example, the FOMO-driven race to the bottom can occur even when* $\rho = 0$[24] *or in the regime of insufficient entry.*[25]

Remark 2. *[Insufficient entry for zero correlation] In Case 4, in addition to the region of the insufficient entry described in Proposition 8, there can exist another region with insufficient entry where no firm enters whereas joint profit maximization requires both firms to enter. For instance, in Figure 4(b), this arises between the vertical FOMO curve and the JP curve for zero or small correlation. As the JP curve does not depends*

[23] In Cases 1 and 2, the region of the race to the bottom includes $\beta = 1$ consistent with the results obtained in Section 4.2.

[24] In Figure 4(a), it occurs between the pointwise maximum of the orange and the blue curves and the red curve.

[25] In Figure 4(b), it occurs between the blue curve and the red curve.

on k but the FOMO curves move to the right as k increases, the area is the most likely to exist for $k = 1/2$. As this insufficient entry occurs for zero correlation, it is not driven by the incentive to free-ride on rival's experimentation. It is driven by an incentive to free ride on the rival's exit: without first-move advantage and with zero correlation, a rival's entry in the first-period increases a firm's payoff from waiting as the rival will exit the market with probability $\lambda > 1/2$ and in that case the firm becomes a monopolist in the second period.

Finally, we conduct comparative statics. Regarding the insufficient entry which occurs in Cases 4-5, we focus on the area above the horizontal FOMO curve and to the left of the JP curve where (0,0) or (1,0) arises whereas joint profit maximization requires (1,1) but do not perform comparative static with respect to the A2 curve.[26]

Proposition 9. *(comparative static) Suppose Assumptions 1 and 2 and $\delta = L = 1$.*
(i) (regime change) For $\lambda < 2/3$, as k increases, we move from the insufficient entry regime to the race to the bottom regime. For $\lambda > 2/3$, a change in k has no effect on the regime.
(ii) (regime change) As λ increases, we move from the race to the bottom regime, to the insufficient entry regime and then to the no distortion regime.
(iii) (within a regime) Given any case in the race to the bottom regime, an increase in ρ makes the race to the bottom more likely. Given any case in the insufficient entry regime, an increase in ρ generally makes insufficient entry more likely, except when it crosses the JP curve.[27]
(iv) (within a regime) An increase in β has mostly a non-monotonic effect on both kinds of distortions but for Case 1 where an increase in β makes the race to the bottom more likely.

The results in Proposition 9(i) and (ii) were previously explained with the help of Figure 3. For Proposition 9(iii), the first part—concerning the race-to-the-bottom inefficiency—follows from the fact that Ψ^J decreases in ρ (Lemma 13), while Ψ^0 increases in ρ (Lemma 12). Since the race-to-the-bottom region lies between the JP and PI curves, these comparative statics imply that the region expands as ρ increases (see Figure 4(a)). The second part of Proposition 9(iii), on insufficient entry, arises because this inefficiency occurs above the horizontal FOMO curve (i.e., for sufficiently large ρ). The non-monotonicity is driven by two forces: Ψ^J decreases in both ρ and β (so the JP curve is negatively sloped) from Lemma 13, and insufficient entry occurs to the

[26] This is because that it does not make much sense to say that an increase in a parameter value makes a particular distortion more or less likely depending on the shape of the A2 curve.

[27] If it crosses the JP curve, an increase in ρ makes insufficient entry more (less) likely for small (large) ρ.

left of the JP curve (see Figure 4(b)). Finally, Proposition 9(iv) follows from the fact that these distortions typically arise for intermediate levels of β, except in Case 1.

5 Welfare Analysis of the Duopoly Model

We now turn to the welfare analysis. Let $S \equiv 1 + C$. Social welfare $W(x, y)$, for $(x, y) \in \{0, 1\}^2$, is given as follows:

$$\begin{aligned}
W(1,1) &= \mu_{SS}(1+\delta)S + \mu_{SN}(1+2\delta)S - \lambda D, \\
W(1,0) &= \begin{cases} \mu_{SS}(1+\delta)S + \mu_{SN}\left[S + \delta((1-\beta)k+\beta)S - (1-\beta)(1-k)D\right] \\ +\mu_{SN}(-D+\delta S) - \mu_{NN}D(1+\delta(1-\beta)) \text{ when (4) holds}, \\ \\ \mu_{SS}(1+\delta)S + \mu_{SN}\left[S + \delta((1-\beta)k+\beta)S - (1-\beta)(1-k)D\right] \\ -\lambda D \text{ when (4) does not hold}, \end{cases} \\
W(0,0) &= \delta\left[\mu_{SS}S + \mu_{SN}\left((1+\beta)S - (1-\beta)D\right) - \mu_{NN}(1-\beta)D\left((1-\beta)+2\beta\right)\right].
\end{aligned}$$

We first establish that, as in the case of joint-profit maximization in Lemma 7, social welfare is never maximized by an asymmetric entry configuration in which only one firm enters in the first period.

Lemma 16. *Under Assumption 2,*

$$W(1,1) \geq W(1,0)$$

with equality if and only if $\rho = 1$ *or* $\beta = 1$.

This result follows from the positive value of market experimentation. Experimenting with both technologies and experimenting with only one technology generate the same first-period welfare, while second-period welfare is weakly higher in the former case than in the latter. Equality holds only in the extreme cases of perfect correlation or perfect beta testing.

By Lemma 16, it suffices to compare $W(1,1)$ and $W(0,0)$ to characterize the socially optimal outcome. We have:

$$\begin{aligned}
W(1,1) - W(0,0) &= \mu_{SS}S + \mu_{SN}\left[(S-D) + \delta(1-\beta)(S+D)\right] \\
&\quad - \mu_{NN}\left[1 - \delta\left(1-\beta^2\right)\right]D, \\
\frac{\partial\left[W(1,1) - W(0,0)\right]}{\partial D} &= -\mu_{SN}\left[(1-\delta(1-\beta))\right] - \mu_{NN}\left[1-\delta\left(1-\beta^2\right)\right].
\end{aligned}$$

Therefore, we have

Proposition 10. *There exists $\bar{\delta} > 1$ such that, for $\delta < \bar{\delta}$ (respectively, $\delta > \bar{\delta}$, $W(1,1) - W(0,0)$ is strictly decreasing (respectively, increasing) with D, implying that $W(1,1) < W(0,0)$ (respectively, $W(1,1) > W(0,0)$) for sufficiently large D.*

When D is large—as in the case of catastrophic events—we have $W(1,1) < W(0,0)$ as long as δ is not too large. The intuition is as follows. Under $(1,1)$, damage can occur only in the first period, whereas under $(0,0)$, damage can occur only in the second period. As δ increases, second period losses receive greater weight, which reduces $W(0,0)$ relative to $W(1,1)$.[28]

Since we assume $\delta = 1$ in Section 4.4, Proposition 10 applies to that section for sufficiently large D: welfare maximization requires that no firm enters in the first period. Therefore, we have:

Corollary 2. *Suppose that $\delta = 1$ and that D is sufficiently large. Then, the race to the bottom with respect to joint-profit maximization implies the race to the bottom with respect to welfare maximization.*

6 Policy implications

To derive policy implications from our results, we focus on the case in which the expected damage from catastrophic events is sufficiently large and the two periods are similarly weighted, so that welfare maximization requires no firm to enter in the first period. In this setting, aligning firms' private incentives with the social optimum can be achieved in two steps.

First, alignment between private and social incentives can be achieved by appropriately setting the level of liability. In particular, in the monopoly case, the two coincide when $D/L = 1 + C(\equiv S)$, as shown in Proposition 3. This result extends to the duopoly setting when comparing joint profit maximization with welfare maximization. As the expected damage from catastrophic events increases, the optimal liability level should rise accordingly. Once liability is properly calibrated, joint–profit maximization implies that no firm enters in the first period.

Second, regulatory intervention is needed to eliminate the race to the bottom. Proposition 9 shows that such inefficiency becomes more likely as the first-mover advantage and the degree of correlation (i.e., technological homogenization) increase. Accordingly, policies that reduce first-mover advantages or limit homogenization in general-purpose AI systems are socially desirable. For example, public support for

[28] For instance, when $\delta = 2$ and $\beta = \frac{1}{2}$, $W(1,1) - W(0,0)$ increases in D, making $W(1,1) > W(0,0)$ for large D.

open-source or open-weight AI can mitigate first-mover advantages by lowering the cost of alternatives to proprietary systems.[29] Similarly, reducing barriers to AI research—particularly for academic institutions—by expanding access to compute and data (e.g., through initiatives such as the National AI Research Resource (NAIRR) in the United States) can foster diversity in approaches and reduce homogenization.

However, policymakers may have limited instruments to directly influence these structural drivers, especially when the FOMO-driven race to the bottom is taken into account, since fear of missing out (FOMO) substantially expands the range of parameter values under which such inefficiency arises. In such cases, direct regulatory interventions based on risk assessment—such as the EU AI Act—may be warranted to prevent premature deployment.

7 Concluding Remarks

This paper studies the optimal timing of AI deployment under risk uncertainty and examines how competition distorts firms' incentives relative to both joint–profit maximization and social welfare. By developing a simple yet flexible framework that incorporates imperfect information, endogenous learning through deployment and beta testing, and key features such as first-mover advantage and technological correlation, we identify fundamental trade-offs in firms' deployment decisions.

In particular, we show that first-mover advantages and technological homogenization in AI can generate a race to the bottom, implying that AI risk regulation may be socially desirable. However, international competition in AI makes such regulation difficult to implement. For instance, a key priority in the current U.S. administration's AI Action Plan is to "remove red tape and burdensome regulation." Similarly, the European Commission has proposed delaying the implementation of parts of the AI Act until 2027.

These developments point to a broader governance challenge concerning the direction of AI research. A small number of commercial actors engaged in intense competitive races may disproportionately shape the research agenda, while independent academic researchers exert little influence because of restricted access to compute and data.

Avoiding a global AI race to the bottom may therefore require international cooperation, particularly between the United States and China, much as the United States and the Soviet Union cooperated during the Cold War to manage the risks posed by

[29] Open-weight AI models are models whose trained neural network parameters (weights and biases) are publicly available, allowing developers to download, run, and fine-tune them. This can reduce vendor lock-in and enhance transparency.

nuclear weapons. In addition, to ensure that disinterested academic researchers can play a meaningful role in shaping the future of AI, broader access to compute resources and high-quality data is essential.

Although our model assumes homogeneous firms, in practice, major AI companies differ in how they approach safety risks. For example, Anthropic was founded by former OpenAI employees dissatisfied with OpenAI's handling of safety concerns and is known for its strong commitment to AI safety, including the use of techniques such as "Constitutional AI."[30] Nevertheless, according to a former AI safety researcher at Anthropic, there were persistent internal pressures to deprioritize safety considerations.[31]

While our analysis highlights the role of competition in shaping deployment timing, it abstracts from firms' endogenous investment in AI safety. In practice, firms allocate resources both to improving the performance of their AI systems and to mitigating safety risks. Extending the model to incorporate this trade-off would be a natural and important direction for future research. In particular, it would be valuable to compare firms' private incentives to invest in safety with the corresponding social incentives, and to examine how the intensity of competition affects this trade-off.

References

Acemoglu, Daron, and Todd Lensman. 2024. "Regulating Transformative Technologies." *American Economic Review: Insights*, 359–376.

Anwar, Usman, Abulhaor Saparov, Javier Rando, Daniel Paleka, Miles Turpin, Peter Hase, Ekdeep Singh Lubana, et al. 2024. "Foundational challenges in assuring alignment and safety of large language models." *arXiv preprint arXiv:2404.09932t.*

Autoridade da Concorrência. 2023. "Competition and Generative Artificial Intelligence."

Autoridade da Concorrência. 2024. "Competition and Generative AI: Zooming in on Data."

Autorité de Concurrence. 2024. "On the Competitive Functionning of the Generative Artificial Intelligence Sector."

Bengio, Yoshua, Bronwyn Fox, André Carlos Ponce de Leon Ferreira de Carvalho, Mona Nemer, Raquel Pezoa Rivera, Yi Zeng, Juha Heikkilä,

[30] Constitutional AI begins by providing an AI model with a written set of principles—a "constitution"—and instructing it to follow those principles as closely as possible. A second AI model is then used to evaluate how well the first model adheres to the constitution.

[31] In his resignation letter shared on X on February 9, 2026, Mrinank Sharma wrote: "Throughout my time here, I've repeatedly seen how hard it is to truly let our values govern our actions. I've seen this within myself, within the organization, where we constantly face pressures to set aside what matters most, and throughout broader society too."

et al. 2025. "International AI Safety Report: The International Scientific Report on the Safety of Advanced AI." AI Action Summit.

Bobtcheff, Catherine, Raphael Levy, and Thomas Mariotti. 2025. "Information Disclosure in Preemption Races: Blessings or (Winner's) Curse?" *RAND Journal of Economics*, 144–162.

Bommasani, Rishi, Drew A. Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S. Bernstein, et al. 2022. "On the Opportunities and Risks of Foundation Models." Center for Research on Foundation Models, Stanford Institute for Human-Centered Artificial Intelligence.

Bremmer, Ian, and Mustafa Suleyman. 2023. "The AI Power Paradox." *Foreign Affairs*.

Competition & Markets Authority. 2023. "AI Foundation Models: Initial Report."

Competition & Markets Authority. 2024. "AI Foundation Models: Update Paper."

Figueroa, Nicolas, Carla Guadalupi, and Jorge Lemus. 2026. "Regulation and Responsible Innovation." *Journal of Law, Economics, and Organization*.

Gans, Joshua. 2025. "How Learning about Harms Impacts the Optimal Rate of Artificial Intelligence Adoption." *Economic Policy*, 40(121): 199–219.

Gans, Joshua. 2026. "Market Power in Artificial Intelligence." *Annual Review of Economics*.

Gladstone AI. 2023. "Survey of AI Technologies and AI R&D Trajectories."

Guerreiro, Joao, Sergio Rebelo, and Pedro Teles. 2026. "Regulating Artificial Intelligence." NBER Working Paper #31921.

Hagiu, Andrei, and Julian Wright. 2025. "Artificial intelligence and Competition Policy." *International Journal of Industrial Organization*, 103: 103134.

Japanese Fair Trade Commission. 2024. "Generative AI and Competition."

Jeon, Doh-Shin. 2025. "Doh-Shin Jeon discussion of: AI stack competition." *Economic Policy*, 40: 257–260.

Jones, Charles I. 2024. "The AI Dilemma: Growth versus Existential Risk." *American Economic Review: Insights*, 6(4): 575–590.

Korinek, Anton, and Jai Vipra. 2025. "Concentrating Intelligence: Scaling and Market Structure in Artificial Intelligence." *Economic Policy*, 40: 225–256.

Pichai, Sundar. 2024. "Alphabet Q3 2024 Earnings Call."

Villalobos, Pablo, Jaime Sevilla, Lennart Heim, Tamay Besiroglu, Marius Hobbhahn, and Anson Ho. 2022. "Will We run Out of Data? Limits of LLM Scaling Based on Human-Generated Data." *arXiv:2211.04325*.

Appendix

Proofs for some results including those of Section 2 and Section 4.2 are omitted as they can be obtained in a straightforward way from the explanations given in the main text.

A.1 Proof of (2)

Since no firm has entered in period one, only beta testing provides information about the state of the world and firm 1 does not enter if $\sigma_1 = n$, firm 2 does not enter if $\sigma_2 = n$. Hence we need to determine firm 1's (firm 2's) equilibrium behavior when $\sigma_1 = \phi$ (when $\sigma_2 = \phi$).

We first inquire the existence of an equilibrium such that firm i enters if $\sigma_i = \phi$, for $i = 1, 2$.

An equilibrium such that firm i enters if and only if $\sigma_i = \phi$, for $i = 1, 2$, exists if and only if τ in (5) below is positive We consider firm 1 and determine $\Pr\{\omega_1, \omega_2, \sigma_2 | \sigma_1 = \phi\}$ for $\omega_1 \in \{S, N\}$, $\omega_2 \in \{S, N\}$, $\sigma_2 \in \{\phi, n\}$, for which it is useful to recall that $\Pr\{\sigma_1 = \phi\} = 1 - \lambda\beta$:

$$\Pr\{(\omega_1, \omega_2) = (S, S), \sigma_2 = \phi | \sigma_1 = \phi\} = \frac{\mu_{SS}}{1 - \lambda\beta}, \quad \Pr\{(\omega_1, \omega_2) = (S, S), \sigma_2 = n | \sigma_1 = \phi\} = 0,$$

$$\Pr\{(\omega_1, \omega_2) = (S, N), \sigma_2 = \phi | \sigma_1 = \phi\} = \frac{\mu_{SN}(1 - \beta)}{1 - \lambda\beta},$$

$$\Pr\{(\omega_1, \omega_2) = (S, N), \sigma_2 = n | \sigma_1 = \phi\} = \frac{\mu_{SN}\beta}{1 - \lambda\beta},$$

$$\Pr\{(\omega_1, \omega_2) = (N, S), \sigma_2 = \phi | \sigma_1 = \phi\} = \frac{\mu_{NS}(1 - \beta)}{1 - \lambda\beta}, \quad \Pr\{(\omega_1, \omega_2) = (N, S), \sigma_2 = n | \sigma_1 = \phi\} = 0,$$

$$\Pr\{(\omega_1, \omega_2) = (N, N), \sigma_2 = \phi | \sigma_1 = \phi\} = \frac{\mu_{NN}(1 - \beta)^2}{1 - \lambda\beta},$$

$$\Pr\{(\omega_1, \omega_2) = (N, N), \sigma_2 = n | \sigma_1 = \phi\} = \frac{\mu_{NN}(1 - \beta)\beta}{1 - \lambda\beta}.$$

The profit of firm 1 from entering given $\sigma_1 = \phi$ (and given that firm 2 enters if and only if $\sigma_2 = \phi$) is equal to $\frac{\delta}{1-\lambda\beta}$ times τ with

$$\tau = \frac{1}{2}\mu_{SS} + \mu_{SN}\left((1 - \beta)\frac{1}{2} + \beta\right) + \mu_{NS}(1 - \beta)\left(-\frac{1}{2}L\right) + \mu_{NN}(1 - \beta)^2\left(-\frac{1}{2}L\right) + \mu_{NN}(1 - \beta)\beta\left(-L\right). \tag{5}$$

Therefore there exists an equilibrium at stage two in which each firm i enters if and only if $\sigma_i = \phi$ if and only if τ in (5) is positive.

Existence of an equilibrium such that no firm ever enters occurs if and only if $1-\lambda-\lambda(1-\beta)L<0$ There exists such an equilibrium if and only if $1-\lambda-\lambda(1-\beta)L<0$, as under monopoly.

An equilibrium such that firm 1 enters if and only if $\sigma_1=\phi$**, firm 2 never enters exists if and only if** $1-\lambda-\lambda(1-\beta)L>0$ **and** $\tau<0$ If $1-\lambda-\lambda(1-\beta)L>0$ and $\tau<0$, then there exists an equilibrium in which firm 1 enters if and only if $\sigma_1=\phi$ (that is a best reply because $1-\lambda-\lambda(1-\beta)L>0$) and firm 2 never enters (that is a best reply because $\tau<0$), and there also exists an equilibrium with reversed roles for the firms. But these are asymmetric equilibria. The unique symmetric equilibrium, given that both firms stayed out in period 1, is such that each firm, conditional on receiving signal ϕ, enters with probability $\frac{1-\lambda-\lambda(1-\beta)L}{1-\lambda-\lambda(1-\beta)L-\tau}$, because the matrix of the game (actions are conditional on having received signal ϕ, utilities are not) is

Firm 1\Firm 2	E	NE
E	$\delta\tau,\delta\tau$	$\delta(1-\lambda-\lambda(1-\beta)L),0$
NE	$0,\delta(1-\lambda-\lambda(1-\beta)L)$	$0,0$

Hence the mixed strategy equilibrium yields profit zero to each firm.

Summary The inequality $\tau>0$ implies $1-\lambda-\lambda(1-\beta)L>0$. Hence Assumption 2 is satisfied if and only if $\tau>0$. Using the expressions of $\mu_{SS},\mu_{SN},\mu_{NS},\mu_{NN}$, we find that $\tau>0$ is equivalent to inequality (2).

A.2 Proof of Lemma 4

Condition (3), which is equivalent to $\tilde{\lambda}(\phi,S)<\lambda_s^*$, can be restated in terms of ρ as follows:

$$\rho>\frac{\lambda+L\lambda-L\beta\lambda-1}{\lambda\left(1+L-L\beta\right)}. \tag{6}$$

If Assumption 1 is violated, then $1-\lambda-\lambda L\geq 0$ and therefore the numerator of the R.H.S. in (6) is negative, implying that (6) is satisfied. If Assumption 1 holds, then from footnote 18 a necessary condition for Assumption 2 to be satisfied is $\beta>\beta_1\equiv 1-\frac{1-\lambda}{L\lambda}$. As the R.H.S. of (6) is equal to zero at $\beta=\beta_1\equiv 1-\frac{1-\lambda}{L\lambda}$ and the R.H.S. strictly decreases in β, it follows that Assumption 2 implies that (6) holds.

A.3 Proof of Lemma 7

Suppose that (4) is violated. Then we compare $\Pi(1,1)$ with $\Pi(1,0)$ and obtain

$$\Pi(1,1) - \Pi(1,0) = \mu_{SN}\delta\left[2 - (1-\beta)k - \beta + (1-\beta)(1-k)L\right],$$

which is positive unless $\rho = 1$ (notice that $\rho = 1$ makes $\tilde{\lambda}(\phi, N)$ equal to 1 and hence (4) not satisfied), as in such case $\mu_{SN} = 0$ and $\Pi(1,1) = \Pi(1,0)$.

Suppose that (4) is satisfied. Then comparing $\Pi(1,1)$ with $\Pi(1,0)$ reveals that

$$\Pi(1,1) - \Pi(1,0) = \mu_{SN}\delta(1-\beta)(1-k)\left(1+L\right) + \mu_{NN}\delta\left(1-\beta\right)L,$$

which is positive unless $\beta = 1$ (notice that $\beta = 1$ makes $\tilde{\lambda}(\phi, N)$ equal to zero and hence (4) satisfied).

A.4 Proof of Lemma 10

In order to see that $\pi(0,0)$ is increasing in β, notice that in state SS, β has no effect on the profit of firm 1; in state SN, an increase in β increases the probability that firm 2 stays out, which increases the profit of firm 1; in states $NS\&NN$ it increases the probability that firm 1 stays out, which yields firm 1 a profit of 0 instead of a negative profit.

In order to see that $\pi(0,0)$ is decreasing in ρ, notice that the equilibrium which is played in period two is such that the profit of firm 1 (to fix the ideas) is

- lower in state SS than in state SN because in the former state firm 2 enters for sure in period 2, but in the latter state there is a chance that firm 2 does not enter (firm 2 does not enter if its beta testing yields signal n), and then firm 1 is monopolist in period 2;
- lower in state NN than in state NS because firm 1 (if it enters) is going to lose money as $\omega_1 = N$, but in state NS it is certain that firm 2 will enter as well, which reduces firm 1's loss to $\frac{1}{2}L$ whereas in state NN there is a chance that firm 2 does not enter, and then firm 1 incurs the full loss L.

Hence an increase in ρ reallocates probability towards states in which firm 1 makes a lower profit, which makes $\pi(0,0)$ decreasing in ρ.

Since $\pi(0,0)$ is increasing in β and decreasing in ρ, it follows that the curve A2 has a positive slope in the space (β, ρ).

Finally, $\pi(0,0)$ is decreasing in λ because of the following: since $\lambda > \frac{1}{2}$, the product $\lambda(1-\lambda)$ is decreasing in λ, thus $\mu_{SS}, \mu_{SN}, \mu_{NS}$ are all decreasing in λ and μ_{NN} increases.

As a result, an increase in λ shifts probability towards state NN in which the profit of firm 1 is lower than in the other states.

Starting from any (β, ρ) such that $\pi(0,0) = 0$, an increase in λ makes $\pi(0,0)$ negative. Hence an increase in β is needed to restore the equality $\pi(0,0) = 0$, that is the A2 curve moves to the right as λ increases.

A.5 Proof of Lemma 11

The expression of $\pi(0,1)$ is determined by whether (4) is satisfied or not. The latter inequality is equivalent to[32]

$$\rho < \frac{1-\lambda-L\lambda+L\beta\lambda}{(1-\lambda)(1+L-L\beta)}. \tag{7}$$

When (7) is satisfied, we have that $\pi(0,1) = (1-k)(\mu_{SS} - \mu_{NS}(1-\beta)) + \mu_{NS} - \mu_{NN}(1-\beta)$. Hence Ψ^1 is decreasing in β since $\frac{\partial \pi(0,1)}{\partial \beta} = \mu_{SN}(1-k) + \mu_{NN} > 0$ and $\pi(1,1)$ does not depend on β. Moreover, $\frac{\partial \pi(1,1)}{\partial \rho} = -\lambda(1-\lambda)\frac{1}{2} > \frac{\partial \pi(0,1)}{\partial \rho} = \lambda(1-\lambda)(1-k-(1-\beta)(1-k)-1-(1-\beta))$, hence Ψ^1 is increasing in ρ. It follows that the vertical FOMO curve is increasing in the space (β, ρ). The latter can be seen also from the fact that $\Psi^1 > 0$ is equivalent to

$$\rho > \frac{2\lambda(1-k+k\lambda)\beta + \lambda^2 - 2\lambda - 2k(\lambda-1)(2\lambda-1)}{\lambda(1-\lambda)(4k-1) - 2\lambda(1-\lambda)k\beta}. \tag{8}$$

The denominator is positive ($\triangleq$). The derivative wrt β is $2\frac{3k\lambda+2k^2\lambda-\lambda-2k^2}{\lambda(1-4k+2k\beta)^2(1-\lambda)}$ ($\triangleq$), which is positive since $\lambda > \frac{1}{2}$, $k > \frac{1}{2}$ ($\triangleq$). At $\beta = 1$, the quotient in (8) is negative if and only if $k > \frac{\lambda^2}{2(1-\lambda)^2}$. Hence the vertical FOMO curve does not appear in the (β, ρ) space if $k > \frac{\lambda^2}{2(1-\lambda)^2}$ ($\triangleq$), that is $\Psi^1 > 0$ in Cases 1,2. Conversely, the vertical FOMO curve appears and is increasing if $k < \frac{\lambda^2}{2(1-\lambda)^2}$.

The effect of an increase in λ is seen from

$$\begin{aligned}
\frac{\partial \pi(1,1)}{\partial \lambda} &= (-2(1-\lambda) + \rho(1-2\lambda)) + (1-\rho)(1-2\lambda)\frac{3}{2} - \frac{1}{2}, \\
\frac{\partial \pi(0,1)}{\partial \lambda} &= (1-k)(-2(1-\lambda) + \rho(1-2\lambda) - (1-\rho)(1-\beta)(1-2\lambda)) \\
&+ (1-\rho)(1-2\lambda) - (2\lambda + \rho(1-2\lambda))(1-\beta).
\end{aligned}$$

[32] The inequality $\beta \geq \beta_1$, which is necessary for $\pi(0,0)$ to be positive, implies $\frac{1-\lambda-L\lambda+L\beta\lambda}{(1-\lambda)(1+L-L\beta)} > 0$.

We have

$$\begin{aligned}\frac{\partial\pi(1,1)}{\partial\lambda}-\frac{\partial\pi(0,1)}{\partial\lambda} &= (2k\lambda\rho-2k\lambda-k\rho+k-1)\,\beta+4\lambda-2\rho\,(2\lambda-1)\\ &- (k-1)\,(\rho\,(2\lambda-1)-2\lambda+(\rho-1)\,(2\lambda-1)+2)+\frac{1}{2}\,(\rho-1)\,(2\lambda-1)-\frac{5}{2}\end{aligned}$$

where $2k\lambda\rho-2k\lambda-k\rho+k-1<0$. Since $\beta\geq 2-\frac{1}{\lambda}$ to the right of the A2 curve, it follows that

$$\begin{aligned}\frac{\partial\pi(1,1)}{\partial\lambda}-\frac{\partial\pi(0,1)}{\partial\lambda} &< (2k\lambda\rho-2k\lambda-k\rho+k-1)\left(2-\frac{1}{\lambda}\right)+4\lambda-2\rho\left(2\lambda-\frac{1}{\lambda}\right)\\ &- (k-1)\,(\rho\,(2\lambda-1)-2\lambda+(\rho-1)\,(2\lambda-1)+2)+\frac{1}{2}\,(\rho-1)\,(2\lambda-1)-\frac{5}{2}\\ &= \frac{(2\lambda+2\rho-4\lambda\rho-2)\,k+(2\lambda^2\rho-2\lambda^2-\lambda\rho-2\lambda+2)}{2\lambda},\end{aligned}$$

which is negative at $k=\frac{1}{2}$ and at $k=1$.

When (7) is violated, we have that $\pi(0,1)=(1-k)\,(\mu_{SS}-\mu_{SN}(1-\beta))$ and $\pi(1,1)>\pi(0,1)$ is equivalent to

$$\rho<\frac{2k+4\lambda-6k\lambda-5\lambda^2+4k\lambda^2-2\lambda\,(1-\lambda)\,(1-k)\,\beta}{\lambda\,(1-\lambda)\,(5-4k)-2\lambda\,(1-\lambda)\,(1-k)\,\beta}. \tag{9}$$

The denominator is positive. The quotient is greater than 1 – independently of β – if and only if $k>\frac{\lambda}{2-2\lambda}$. Hence, in Cases 1, 2, 3 the horizontal FOMO curve does not appear in the (β,ρ) space. If instead $k<\frac{\lambda}{2-2\lambda}$, then the quotient is less than 1, and the curve appears in the (β,ρ) space; the derivative of the R.H.S. of (9) with respect to β is $\frac{2(1-k)(2k-\lambda-2k\lambda)}{\lambda(1-\lambda)(5-4k-2\beta+2k\beta)^2}$ ($\triangleq$), which is negative ($\triangleq$).

When (7) is violated, Ψ^1 is decreasing in β since $\pi(1,1)$ does not depend on β and $\frac{\partial\pi(0,1)}{\partial\beta}=(1-k)(1-\rho)\lambda(1-\lambda)>0$. Moreover, $\frac{\partial\pi(1,1)}{\partial\rho}=-\lambda(1-\lambda)\frac{1}{2}<0$, $\frac{\partial\pi(0,1)}{\partial\rho}=(1-k)(2-\beta)\lambda(1-\lambda)>0$, thus Ψ^1 is decreasing in ρ and the horizontal FOMO curve is decreasing in the (β,ρ) space.

The effect of an increase in λ on Ψ^1 is seen from $\frac{\partial\pi(1,1)}{\partial\lambda}=\lambda\,(\rho-1)-\frac{1}{2}\rho-1<\frac{\partial\pi(0,1)}{\partial\lambda}=(1-k)\,(-2(1-\lambda)+\rho(1-2\lambda)-(1-\rho)(1-\beta)(1-2\lambda))$, that is the adverse change in $\mu_{SS},\mu_{SN},\mu_{NS},\mu_{NN}$ determined by an increase in λ reduces $\pi(1,1)$ more than $\pi(0,1)$.

Therefore, independently of whether (7) holds or not, Ψ^1 decreases in λ and in β, thus the FOMO curve moves left as λ increases.

Comparing (7) and (9) reveals that $\frac{1-\lambda-L\lambda+L\beta\lambda}{(1-\lambda)(1+L-L\beta)}<\frac{2k+4\lambda-6k\lambda-5\lambda^2+4k\lambda^2-2\lambda(1-\lambda)(1-k)\beta}{\lambda(1-\lambda)(5-4k)-2\lambda(1-\lambda)(1-k)\beta}$ holds if and only if $\beta<\frac{k+3\lambda-4k\lambda-\sqrt{-2k\lambda+3\lambda^2-2k\lambda^2+k^2}}{2\lambda-2k\lambda}=\hat{\beta}$. Comparing (7) and (8) and reveals that $\frac{2\lambda(1-k+k\lambda)\beta+\lambda^2-2\lambda-2k(\lambda-1)(2\lambda-1)}{\lambda(1-\lambda)(4k-1)-2\lambda(1-\lambda)k\beta}<\frac{1-\lambda-L\lambda+L\beta\lambda}{(1-\lambda)(1+L-L\beta)}$ holds if and only if $\beta<\hat{\beta}$.

The crossing point is $(\hat{\beta}, \frac{1-\lambda-\lambda+\hat{\beta}\lambda}{(1-\lambda)(2-\hat{\beta})})$, and $\hat{\beta}$ must be larger than $2-\frac{1}{\lambda}$ for $(\hat{\beta}, \frac{1-\lambda-\lambda+\hat{\beta}\lambda}{(1-\lambda)(2-\hat{\beta})})$ to satisfy A2. This is equivalent to $2-\frac{1}{\lambda} \leq \frac{k+3\lambda-4k\lambda-\sqrt{-2k\lambda+3\lambda^2-2k\lambda^2+k^2}}{2\lambda-2k\lambda}$ and it boils down to $\lambda \leq \sqrt{3}-1$. It follows that $\Psi^1 < 0$ for each $\lambda > \sqrt{3}-1$.

A.6 Proof of Lemma 12

The difference $\Psi^0 = \pi(1,0) - \pi(0,0)$ is increasing in ρ because $\frac{\partial \pi(1,0)}{\partial \rho} = -\lambda(1-\lambda)(1-k)\beta > \frac{\partial \pi(0,0)}{\partial \rho} = -\lambda(1-\lambda)\frac{1}{2}\beta(2-\beta)$. Moreover, Ψ^0 is decreasing in β because $\frac{\partial \pi(1,0)}{\partial \beta} = \mu_{SN}(1-k)$ but $\frac{\partial \pi(0,0)}{\partial \beta} = \mu_{SN} +$ other positive terms. Hence the curve PI is increasing in the space (β, ρ). At $\beta = 1$, Ψ^0 is equal to $\left[\frac{\lambda^2\beta^2-\left(2k\lambda^2-2k\lambda\right)\beta-2k+2\lambda+2k\lambda-1}{\lambda(1-\lambda)(2k-\beta)\beta}\right]_{\beta=1} \triangleq \frac{2\lambda-2k+4k\lambda+\lambda^2-2k\lambda^2-1}{\lambda(1-\lambda)(2k-1)}$, which is negative if and only if $k > \frac{2\lambda+\lambda^2-1}{2(1-\lambda)^2}$ ($\triangleq$), that is in Case 1. Hence the PI curve appears if and only if $k < \frac{2\lambda+\lambda^2-1}{2(1-\lambda)^2}$, and in such case it is increasing. The expression of $\pi(1,0)$ is such that the derivative with respect to λ is $-2\beta(1-\rho)(1-k)\lambda-k+\beta-k\beta-\beta\rho+k\beta\rho-2$. The expression of $\pi(0,0)$ is such that the derivative with respect to λ is $-\beta(2)-\beta)(1-\rho)\lambda+\beta-\beta\rho+\frac{1}{2}\beta^2\rho-1$. The difference between the former and the latter is $\beta(1-\rho)(2k-\beta)\lambda+k\beta\rho-k\beta-\frac{1}{2}\beta^2\rho-k-1$, which is smaller than $\beta(1-\rho)(2k-\beta)+k\beta\rho-k\beta-\frac{1}{2}\beta^2\rho-k-1 = k\beta-k+\frac{1}{2}\beta^2\rho-\beta^2-1-k\beta\rho < 0$. Starting from any (β,ρ) such that $\Psi^0 = 0$, an increase in λ makes Ψ^0 negative, hence a decrease in β is needed to restore the equality $\pi(1,0) = \pi(0,0)$ and hence the PI curve moves to the left as λ increases.

About Cases 5 and 6 (i.e., when $\lambda > \frac{2}{3}$), the PI curve is to the left of the A2 curve. Precisely, the PI curve has equation

$$\rho = \frac{\lambda^2\beta^2 - (2k\lambda^2 - 2k\lambda)\beta - 2k + 2\lambda + 2k\lambda - 1}{\lambda(1-\lambda)(2k-\beta)\beta}$$

and $\frac{\lambda^2\beta^2-\left(2k\lambda^2-2k\lambda\right)\beta-2k+2\lambda+2k\lambda-1}{\lambda(1-\lambda)(2k-\beta)\beta} > \frac{(\beta\lambda+1)(1-2\lambda+\beta\lambda)}{\beta\lambda(1-\lambda)(2-\beta)}$ (the right hand side comes from (2)) holds for each β satisfying $\pi(0,0) > 0$ because $\frac{\lambda^2\beta^2-\left(2k\lambda^2-2k\lambda\right)\beta-2k+2\lambda+2k\lambda-1}{\lambda(1-\lambda)(2k-\beta)\beta} - \frac{(\beta\lambda+1)(1-2\lambda+\beta\lambda)}{\beta\lambda(1-\lambda)(2-\beta)} = 2\frac{\lambda(1-k)\beta^2+(k-2\lambda-k\lambda+1)\beta+(2\lambda-3k+4k\lambda-1)}{\beta\lambda(1-\lambda)(2-\beta)(2k-\beta)}$, which we next prove is positive. Precisely, $\lambda(1-k) > 0$ and the numerator is minimized at $\beta = 1$ if $k \geq \frac{1}{3\lambda-1}$, is minimized at $\beta = \frac{2\lambda+k\lambda-1-k}{2\lambda(1-k)}$ if $k < \frac{1}{3\lambda-1}$. In the first case, the minimum value of the numerator is $3k(\lambda-\frac{2}{3}) > 0$. In second case, the value of the numerator at the minimum β is $\frac{(14\lambda-17\lambda^2-1)k^2+(4\lambda^2-2\lambda-2)k+4\lambda^2-1}{4\lambda(1-k)}$ and the numerator is concave in k, positive at $k = 1/2$ and at $k = \frac{1}{3\lambda-1}$.

A.7 Proof of Lemma 13

The difference $\Psi^{JP} = \pi(1,1) - \pi(0,0)$ does not depend on k. It is decreasing in β because $\pi(1,1)$ does not depend on β and $\pi(0,0)$ is increasing in β. It is decreasing in ρ because $\pi(1,1)$ decreases in ρ more than $\pi(0,0)$ does: $\frac{\partial \pi(1,1)}{\partial \rho} = -\lambda(1-\lambda)\frac{1}{2} < \frac{\partial \pi(0,0)}{\partial \rho} = -\lambda(1-\lambda)\frac{1}{2}\beta(2-\beta)$; hence in the space (β,ρ) the JP curve is decreasing. Moreover, $\Psi^{JP} = -\frac{1}{2}(1-\beta)^2(1-\rho)\lambda^2 - \left(\beta + \frac{1}{2}\rho(\beta-1)^2\right)\lambda + \frac{1}{2}$ is decreasing in λ essentially because under $(1,1)$ there is more difference in the payoffs among different states of the world than under $(0,0)$, and the adverse change in $\mu_{SS}, \mu_{SN}, \mu_{NS}, \mu_{NN}$ determined by an increase in λ reduces $\pi(1,1)$ more than $\pi(0,0)$. Starting from any (β,ρ) such that $\Psi^{JP} = 0$, an increase in λ makes Ψ^{JP} negative, hence a decrease in β is needed to restore the equality $\Psi^{JP} = 0$ and the JP curve moves to the left as λ increases.

Each point (β,ρ) on the JP curve satisfies $\beta \geq \sqrt{\frac{1-\lambda}{\lambda}}$, and each (β,ρ) on the A2 curve satisfies $\beta \leq \sqrt{\frac{2\lambda-1}{\lambda}}$. For $\lambda < \frac{2}{3}$ we have that $\sqrt{\frac{2\lambda-1}{\lambda}} < \sqrt{\frac{1-\lambda}{\lambda}}$, hence the A2 curve is to the left of the JP curve in Cases 1, 2, 3. For Case 6, the JP curve lies to the left of the A2 curve as each (β,ρ) on the JP curve satisfies $\beta < 1 - \frac{1}{\lambda} + \sqrt{\frac{2-2\lambda}{\lambda^2}}$, and each (β,ρ) on the A2 curve satisfies $\beta > 2 - \frac{1}{\lambda}$, and now we prove that $1 - \frac{1}{\lambda} + \sqrt{\frac{2-2\lambda}{\lambda^2}} < 2 - \frac{1}{\lambda}$. Indeed, $2 - \frac{1}{\lambda} - \left(1 - \frac{1}{\lambda} + \sqrt{\frac{2-2\lambda}{\lambda^2}}\right) = 1 - \sqrt{\frac{2-2\lambda}{\lambda^2}}$ and $1 - \sqrt{\frac{2-2\lambda}{\lambda^2}} > 0$ is equivalent to $\lambda^2 \geq 2 - 2\lambda$, or $\lambda \geq \sqrt{3} - 1$.

The inequality $\Psi^{JP} > 0$ reduces to

$$\rho > \frac{-\lambda^2\beta^2 - 2\lambda(1-\lambda)\beta + 1 - \lambda^2}{\lambda(1-\beta)^2(1-\lambda)}. \tag{10}$$

The derivative of the R.H.S. of the inequality is $\frac{2(1-\lambda-\beta\lambda)}{\lambda(1-\beta)^3(1-\lambda)}$ ($\triangleq$), which is negative if $\beta > \frac{1-\lambda}{\lambda}$. The quotient $\frac{-\lambda^2\beta^2-2\lambda(1-\lambda)\beta+1-\lambda^2}{\lambda(1-\beta)^2(1-\lambda)}$ is greater than 1 if and only if $\beta^2 < \frac{1-\lambda}{\lambda}$ ($\triangleq$), hence the JP curve appears only if $\beta \geq \sqrt{\frac{1-\lambda}{\lambda}}$, which implies that the derivative is negative. The quotient is zero for $\beta > 1 - \frac{1}{\lambda} + \sqrt{\frac{2-2\lambda}{\lambda^2}}$ ($\triangleq$), hence the curve appears only for β between $\sqrt{\frac{1-\lambda}{\lambda}}$ and $1 - \frac{1}{\lambda} + \sqrt{\frac{2-2\lambda}{\lambda^2}}$. That is a very small space, implying that the curve is almost vertical.

A.8 Proof for Case 1, 2 and 6 of Section 4.4

From Lemmas 11-12, we know that $\Psi^1 > 0$ for all (β,ρ) in Cases 1 and 2, which implies that (1,1) is always an equilibrium in these cases. This, together with $\Psi^0 > 0$ for all (β,ρ) in Case 1 implies that (1,1) is the unique equilibrium in Case 1. In Case 2, the PI curve appears and hence (1,1) is the unique equilibrium to the left of the

curve and there are multile equilibria of (1,1) and (0,0) to the right of the curve. In Case 6, as both $\Psi^1 < 0$ and $\Psi^0 < 0$ hold for all (β, ρ) satisfying A2, (0,0) is the unique equilibrium. Summarizing,

Corollary 3. *(i) In Case 1,* $(1,1)$ *is the unique equilibrium.*
(ii) In Case 2, for all (β, ρ) *satisfying A2,* $(1,1)$ *is the unique equilibrium to the left of the PI curve and both (1,1) and (0,0) are equilibria to the right of it.*
(iii) In Case 6, $(0,0)$ *is the unique equilibrium.*

From Lemma 13, in Case 6, $\Psi^J < 0$ holds for all (β, ρ) satisfying A2. Hence, there is no distortion in this case.

In Case 1, the JP curve exists in the (β, ρ) space. Hence, there is no distortion to the left of the curve but there is a race to the bottom to its right. In Case 2, we find that the PI curve is located always to the right side of the JP curve for ρ =1 and either the whole PI curve is located to the right of the JP curve or the former intersects once the latter. This implies that between the JP curve and the PI curve, there is a race to the bottom and that there is a FOMO-driven race to the bottom to the right of the pointwise maximum of the PI curve and the JP curve. Summarizing, we have (see also Figure 5)

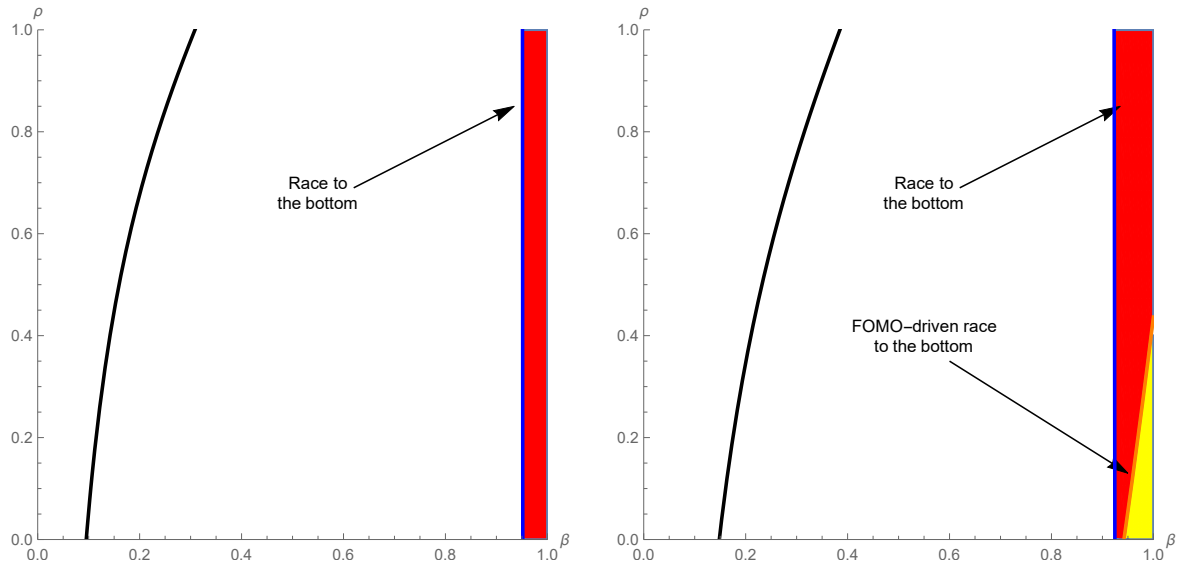
Proposition 11. *(i) In Case 1, there is a race to the bottom to the right of the JP curve.*
(ii) In Case 2, there is a race to the bottom to between the JP curve and the PI curve and there is a FOMO-driven race to the bottom to the right of the pointwise maximum of the PI curve and the JP curve.
(iii) In Case 6, there is no distortion.

A.9 Proof for Case 5 of Section 4.4

In Case 5, as the PI curve is located entirely to the left of the A2 curve, Proposition 7(ii)(b) applies. Hence, in the region between the A2 curve and the vertical FOMO curve, but below the horizontal FOMO curve, both (1,1) and (0,0) arise as equilibria and (0,0) is the unique equilibrium to the right of the curve defined by the pointwise maximum of the PI and FOMO curves. Therefore, we have (see also Figure 6)

Proposition 12. *In Case 5, there is an insufficient entry in the region between the A2 curve and the JP curve and above the horizontal FOMO curve: the joint profit maximization requires both firms to enter but (0,0) is the unique equilibrium.*

We note that there can be a region of the FOMO-dirven race to the bottom between the JP curve and the FOMO curve (see Figure 6).



(a) Case 1 (b) Case 2

Figure 5: Cases 1 and 2

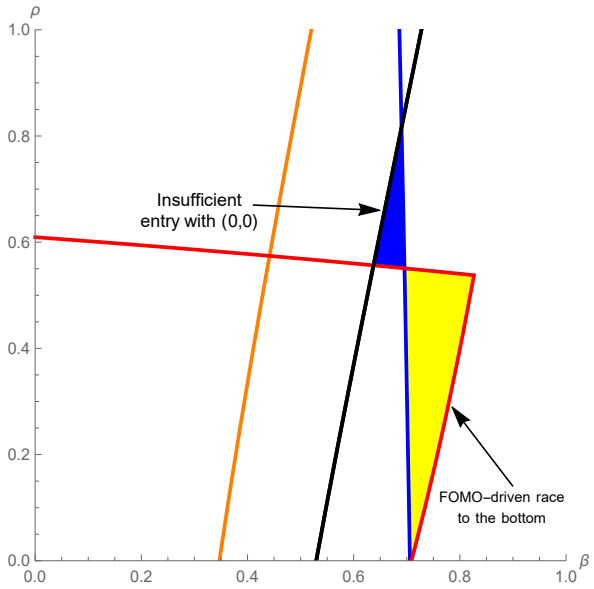


Figure 6: Case 5

A.10 Proof for Lemma 14

(i) The inequality $\Psi^0 > 0$ is equivalent to $\rho > \frac{\lambda^2\beta^2-\left(2k\lambda^2-2k\lambda\right)\beta-2k+2\lambda+2k\lambda-1}{\lambda(1-\lambda)(2k-\beta)\beta}$. The inequality $\Psi^1 > 0$ is equivalent to $\rho > \frac{2\lambda(1-k+k\lambda)\beta+\lambda^2-2\lambda-2k(\lambda-1)(2\lambda-1)}{\lambda(1-\lambda)(4k-1)-2\lambda(1-\lambda)k\beta}$, of which the right hand side is positive if and only if $\beta > \frac{2\lambda-\lambda^2+2k(\lambda-1)(2\lambda-1)}{2\lambda(1-k+k\lambda)}$. We below prove that the difference between $\frac{\lambda^2\beta^2-\left(2k\lambda^2-2k\lambda\right)\beta-2k+2\lambda+2k\lambda-1}{\lambda(1-\lambda)(2k-\beta)\beta}$ and $\frac{2\lambda-\lambda^2+2k(\lambda-1)(2\lambda-1)}{2\lambda(1-k+k\lambda)}$ is positive for $\beta > \frac{2\lambda-\lambda^2+2k(\lambda-1)(2\lambda-1)}{2\lambda(1-k+k\lambda)}$.

The difference is equal to

$$\frac{2\lambda\left(1-k\right)\beta^3+\left(2k\lambda-2\lambda-2k\right)\beta^2+2k\left(1-\lambda\right)\left(4k+1\right)\beta+\left(4k-1\right)\left(-1-2k+2\lambda+2k\lambda\right)}{\beta\lambda\left(4k-2k\beta-1\right)\left(2k-\beta\right)\left(1-\lambda\right)}$$

At $\beta=\frac{2\lambda-\lambda^2+2k(\lambda-1)(2\lambda-1)}{2\lambda(1-k+k\lambda)}$, the numerator of the fraction above is

$$\frac{3k\lambda+2k^2\lambda-\lambda-2k^2}{4}\,\frac{\begin{array}{c}\left(64\lambda^3-16\lambda^4-96\lambda^2+64\lambda-16\right)k^3+\left(20\lambda^4-72\lambda^3+108\lambda^2-80\lambda+24\right)k^2\\+\left(32\lambda^3-8\lambda^4-24\lambda^2\right)k+\lambda^4-4\lambda^3+4\lambda^2+8\lambda-4\end{array}}{\lambda\left(1-k+k\lambda\right)^3}$$

and it is positive. At $\beta=1$, it is $\left(2\lambda-1\right)\left(2k-1\right)>0$. The derivative of the numerator is $6\lambda(1-k)\beta^2+2(2k\lambda-2\lambda-2k)\beta+2k(1-\lambda)(4k+1)$ and at $\beta=1$ it has value $\frac{8}{3}k^2-\frac{14}{3}k+\frac{4}{3}<0$. Hence the derivative is either first positive and then negative, or constantly negative. In either case the numerator is constantly positive since it is positive at the extremes.

(ii) The PI curve hits the horizontal curve $\rho=0$ at $\beta_1=\frac{\sqrt{2k-2\lambda+k^2\lambda^2-2k\lambda-2k^2\lambda+k^2+1}-k+k\lambda}{\lambda}$, hits the horizontal curve $\rho=1$ at $\beta_2=\frac{1}{\lambda}\sqrt{\lambda+2k\lambda-2\lambda^2-2k\lambda^2}$. It is simple to see that $\beta_1>0$ since $k>\frac{1}{2}$, $\beta_2<1$ since $k<\frac{\lambda}{2-2\lambda}$.

Recall that the horizontal FOMO curve lies below the horizontal curve $\rho=1$ for each β since $k<\frac{\lambda}{2-2\lambda}$. From the expression of the vertical and horizontal FOMO curves, we find that in Case 4 they intersect at $\hat{\beta}=\frac{1}{2\lambda-2k\lambda}\left(k+3\lambda-4k\lambda-\sqrt{-2k\lambda+3\lambda^2-2k\lambda^2+k^2}\right)$ In order to prove that the horizontal FOMO intersects the PI curve, it suffices to prove that $\beta_2<\hat{\beta}$. This inequality is equivalent to

$$\frac{1}{2\lambda-2k\lambda}\left(k+3\lambda-4k\lambda-\sqrt{-2k\lambda+3\lambda^2-2k\lambda^2+k^2}\right)>\frac{1}{\lambda}\sqrt{\lambda+2k\lambda-2\lambda^2-2k\lambda^2}$$

$$\begin{aligned}
&\iff\quad k+3\lambda-4k\lambda>\sqrt{-2k\lambda+3\lambda^2-2k\lambda^2+k^2}+(2-2k)\sqrt{\lambda+2k\lambda-2\lambda^2-2k\lambda^2}\\
&\iff\quad \lambda\left(2k+7\lambda-8k\lambda-4k^2\lambda+4k^2-2\right)>2\sqrt{(\lambda+2k\lambda-2\lambda^2-2k\lambda^2)(-2k\lambda+3\lambda^2-2k\lambda^2+k^2)}\\
&\iff\quad Z(\lambda,k)\equiv\left(16k^4+64k^3-8k^2-104k+73\right)\lambda^3+\left(40k^2-80k^3-32k^4+28k-40\right)\lambda^2\\
&+\quad \left(16k^4+24k^3+12k^2+4\right)\lambda-\left(8k^3+4k^2\right)>0.
\end{aligned}$$

Since we are considering Case 4, we need to prove $Z(\lambda,k)>0$ for each $(k,\lambda)\in(\frac{1}{2},1)\times(\frac{2k}{2k+1},\frac{2}{3})$. First notice that $Z(\frac{2k}{2k+1},k)=4k\left(2k-1\right)\frac{29k-42k^2+16k^3-2}{(2k+1)^3}$, which is positive for each $k\in(\frac{1}{2},1)$. Then notice that $\frac{\partial Z}{\partial\lambda}=\left(16k^4+64k^3-8k^2-104k+73\right)3\lambda^2+$ $2\left(40k^2-80k^3-32k^4+28k-40\right)\lambda+16k^4+24k^3+12k^2+4$ and $\frac{\partial^2 Z}{\partial\lambda^2}$ $=\left(16k^4+64k^3-8k^2-104k+73\right)6\lambda-64k^4-160k^3+80k^2+56k-80$. We have that $16k^4+64k^3-8k^2-104k+73>0$ for each $k\in(\frac{1}{2},1)$, and it is immedi-

ate that $\frac{\partial Z}{\partial \lambda}$ is minimized with respect to λ at $\lambda = \frac{-28k-40k^2+80k^3+32k^4+40}{-312k-24k^2+192k^3+48k^4+219}$. Since $\frac{\partial Z}{\partial \lambda}(\frac{-28k-40k^2+80k^3+32k^4+40}{-312k-24k^2+192k^3+48k^4+219}, k)$ coincides with $\frac{-256k^8-896k^7+960k^6+4928k^5-3760k^4-6360k^3+4948k^2+992k-724}{3(73-104k-8k^2+64k^3+16k^4)}$, which is positive for each $k \in (\frac{1}{2}, \frac{17}{20}]$, it follows that Z is increasing in λ if $k \in (\frac{1}{2}, \frac{17}{20}]$, thus $Z(\lambda, k) > 0$ for each $\lambda \in (\frac{2k}{2k+1}, \frac{2}{3})$ holds if $k \in (\frac{1}{2}, \frac{17}{20}]$. When $k \in (\frac{17}{20}, 1)$, we notice that $\frac{\partial^2 Z}{\partial \lambda^2} < (16k^4 + 64k^3 - 8k^2 - 104k + 73)\, 6(\frac{2}{3}) - 64k^4 - 160k^3 + 80k^2 + 56k - 80 = 96k^3 + 48k^2 - 360k + 212$, which is negative for each $k \in (\frac{17}{20}, 1)$; hence Z is concave in λ for $\lambda \in (\frac{2k}{2k+1}, \frac{2}{3})$. Since we know that $Z(\frac{2k}{2k+1}, k) > 0$, it suffices to verify that $Z(\frac{2}{3}, k) = \frac{32}{27}k^4 - \frac{232}{27}k^3 + \frac{524}{27}k^2 - \frac{496}{27}k + \frac{176}{27}$, is positive for each $k \in (\frac{17}{20}, 1)$ to conclude that $Z(\lambda, k) > 0$ for each $\lambda \in (\frac{2k}{2k+1}, \frac{2}{3})$ when $k \in (\frac{17}{20}, 1)$.

A.11 Proof of Lemma 15

(i) When $k = \frac{\lambda}{2-2\lambda}$, the PI curve is the most to the left and it is given by $\rho = (\beta\lambda + 1)\frac{\lambda+\beta\lambda-1}{\beta\lambda(\lambda+\beta\lambda-\beta)}$, which takes the value of $\rho = 1$ at $\beta \triangleq \sqrt{\frac{1-\lambda}{\lambda}}$, and we also know that the JP curve also contains $(\rho = 1, \beta = \sqrt{\frac{1-\lambda}{\lambda}})$. Hence the two curves intersect at $\beta = \sqrt{\frac{1-\lambda}{\lambda}}$, that is PI curve is completely to the left of JP curve. But for a greater k, that is between $\frac{\lambda}{2-2\lambda}$ and $\frac{\lambda^2}{2(1-\lambda)^2}$, the PI curve moves to the right and the two curves intersect properly. At $\rho = 0$ we know that the PI curve is to the left of the JP curve for $k = \frac{\lambda^2}{2(1-\lambda)^2}$. Therefore; for lower k it is even more to the left of the JP curve. As a result, for each k between $\frac{\lambda}{2-2\lambda}$ and $\frac{\lambda^2}{2(1-\lambda)^2}$, they intersect.

The JP curve reaches $\rho = 0$ at $\beta \triangleq 1 - \frac{1}{\lambda} + \sqrt{\frac{2-2\lambda}{\lambda^2}}$ and the vertical FOMO curve reaches $\rho = 0$ at $\beta \triangleq \frac{2\lambda-\lambda^2+2k(\lambda-1)(2\lambda-1)}{2\lambda(1-k+k\lambda)}$. Moreover, the vertical FOMO curve is the most to the left when $k = \frac{\lambda}{2-2\lambda}$ and then we find

$$\begin{aligned}
&\left[\frac{2\lambda - \lambda^2 + 2k\,(\lambda-1)\,(2\lambda-1)}{2\lambda\,(1-k+k\lambda)} - \left(1 - \frac{1}{\lambda} + \sqrt{\frac{2-2\lambda}{\lambda^2}}\right)\right]_{k=\frac{\lambda}{2-2\lambda}} \\
\triangleq\ & \frac{2\,(1-\lambda^2)}{\lambda\,(2-\lambda)} - \sqrt{\frac{2-2\lambda}{\lambda^2}} = \frac{\sqrt{2}}{\lambda}\left(\frac{\sqrt{2}\,(1-\lambda)\,(1+\lambda)}{2-\lambda} - \sqrt{1-\lambda}\right) \\
=\ & \frac{\sqrt{2}\sqrt{1-\lambda}}{\lambda}\left(\frac{\sqrt{2}\sqrt{1-\lambda}\,(1+\lambda)}{2-\lambda} - 1\right)
\end{aligned}$$

which is positive if and only if $\left(\frac{2(1-\lambda^2)}{\lambda(2-\lambda)}\right)^2 > \frac{2-2\lambda}{\lambda^2}$, which is equivalent to $\left(\frac{2(1-\lambda^2)}{\lambda(2-\lambda)}\right)^2 - \frac{2-2\lambda}{\lambda^2} = 2\,(1-\lambda)\,(2\lambda-1)\,\frac{2-2\lambda-\lambda^2}{\lambda^2(2-\lambda)^2} > 0$ for each $\lambda \in (\frac{1}{2}, \frac{2}{3})$. Finally, recall that the vertical FOMO curve is increasing and the JP curve is decreasing. Hence, the FOMO curve is completely to the right of the JP curve.

(ii) It suffices to prove that the intersection point between the vertical and hori-

zontal FOMO curves is located to the right of the JP curve. The intersection point is $(\hat{\beta}, \frac{1-\lambda-\lambda+\hat{\beta}\lambda}{(1-\lambda)(2-\hat{\beta})})$, with $\hat{\beta} = \frac{k+3\lambda-4k\lambda-\sqrt{-2k\lambda+3\lambda^2-2k\lambda^2+k^2}}{2\lambda-2k\lambda}$. The difference $\frac{1-\lambda-\lambda+\beta\lambda}{(1-\lambda)(2-\beta)} - \frac{-\lambda^2\beta^2-2\lambda(1-\lambda)\beta+1-\lambda^2}{\lambda(1-\beta)^2(1-\lambda)}$ (where the second term is the R.H.S. of (10)) is equal to $\frac{\beta+\lambda+2\beta\lambda-\beta^2\lambda-2}{\lambda(2-\beta)(1-\beta)^2(1-\lambda)}$. At $\beta = \hat{\beta}$ the numerator is equal to

$$\frac{1}{2}(2k-1)\frac{-k+\lambda+k\lambda-2\lambda^2-\lambda\sqrt{-2k\lambda+3\lambda^2-2k\lambda^2+k^2}+k\lambda^2+\sqrt{-2k\lambda+3\lambda^2-2k\lambda^2+k^2}}{\lambda(k-1)^2}$$

and this is positive since $(1-\lambda)\sqrt{-2k\lambda+3\lambda^2-2k\lambda^2+k^2} > k-\lambda-k\lambda+2\lambda^2-k\lambda^2$ is equivalent to $(1-\lambda)^2(-2k\lambda+3\lambda^2-2k\lambda^2+k^2) - (k-\lambda-k\lambda+2\lambda^2-k\lambda^2)^2 > 0$, which is equivalent to $\lambda^2(2-2\lambda-\lambda^2)(k-1)^2 > 0$, where the sign is from $\lambda < \frac{2}{3}$.

A.12 Proof of Lemma 16

Suppose that (4) is violated. Then we have

$$\begin{aligned} W(1,1) &= \mu_{SS}(1+\delta)S + \mu_{SN}(1+2\delta)S - \lambda D, \\ W(1,0) &= \mu_{SS}(1+\delta)S + \mu_{SN}[S+\delta((1-\beta)k+\beta)S - (1-\beta)(1-k)D)] - \lambda D. \end{aligned}$$

Therefore, we obtain

$$W(1,1) - W(1,0) = \mu_{SN}\delta(2S - (1-\beta)kS - \beta S + (1-\beta)(1-k)D)$$

which is positive unless $\rho = 1$, as in such case $\mu_{SN} = 0$ and $W(1,1) = W(1,0)$.

Suppose that (4) is satisfied. Then we have

$$\begin{aligned} W(1,1) &= \mu_{SS}(1+\delta)S + \mu_{SN}(1+2\delta)S - \lambda D, \\ W(1,0) &= \mu_{SS}(1+\delta)S + \mu_{SN}[S+\delta((1-\beta)k+\beta)S - (1-\beta)(1-k)D)] \\ &+ \mu_{SN}(-D+\delta S) - \mu_{NN}D(1+\delta(1-\beta)). \end{aligned}$$

Therefore, we obtain

$$W(1,1) - W(1,0) = \mu_{SN}\delta(S - (1-\beta)kS - \beta S + (1-\beta)(1-k)D) + \mu_{NN}\delta(1-\beta)D,$$

which is positive unless $\beta = 1$.

A.13 Proof of Proposition 10

$\frac{\partial[W(1,1)-W(0,0)]}{\partial D} \gtreqless 0$ if and only if $\delta \gtreqless \frac{\mu_{SN}+\mu_{NN}}{\mu_{SN}(1-\beta)+\mu_{NN}(1-\beta^2)}$. We have

$$\frac{\mu_{SN}+\mu_{NN}}{\mu_{SN}(1-\beta)+\mu_{NN}(1-\beta^2)} > \frac{1}{(1-\beta^2)} > 1.$$

Therefore, we can set $\bar{\delta}$ equal to $\frac{\mu_{SN}+\mu_{NN}}{\mu_{SN}(1-\beta)+\mu_{NN}(1-\beta^2)}$.

A.14 Extension: When Competition Reduces Profit and Increases Consumer Surplus

In the baseline model, we assumed that the only effect of competition is that competing firms share the market. We now relax this assumption by allowing competition to reduce profits while increasing consumer surplus, with the increase in consumer surplus exceeding the reduction in profits.

Note that this new assumption does not affect Lemma 16, since it raises $W(1,1)$ relative to $W(1,0)$. In addition, Proposition 10 continues to hold for $\delta = 1$ so that provided that D is sufficiently large, no first-period entry remains socially optimal.

Accordingly, we focus on how this new assumption affects the comparison between the market outcome and the joint–profit–maximizing outcome. More precisely, we assume that competition reduces profit per consumer from 1 to $(1-\Delta) < 1$ whenever two firms compete in a given period. We examine how a small increase in Δ from zero affects the equilibrium comparison.

Under this approach, Lemma 7 continues to hold except in the polar cases of perfect correlation or perfect beta testing. Our analysis focuses on Cases 3 and 4 of Section 4.4.

Individual profits are now given as follows as a function of first-period entry profiles:

$$\begin{aligned}
\pi(1,1) &= \mu_{SS}\left(\frac{1}{2}+\frac{1}{2}\right)(1-\Delta)+\mu_{SN}\left(\frac{1}{2}(1-\Delta)+1\right)-\frac{\lambda}{2},\\
\pi(0,1) &= \mu_{SS}(1-k)(1-\Delta)-\mu_{NS}(1-\beta)(1-k))\\
&+ [\mu_{SN}-\mu_{NN}(1-\beta)]\mathbf{1}_{[1-\tilde{\lambda}(\phi,N)>\tilde{\lambda}(\phi,N)]},\\
\pi(1,0) &= \mu_{SS}\left(1+k(1-\Delta)\right)+\mu_{SN}\left(1+((1-\beta)k(1-\Delta)+\beta)\right)-\lambda,\\
\pi(0,0) &= \frac{1}{2}(1-\Delta)\mu_{SS}+\mu_{SN}(\beta+(1-\beta)(1-\Delta)\frac{1}{2})\\
&- \mu_{NS}(1-\beta)\frac{1}{2}-\mu_{NN}(1-\beta)\left((1-\beta)\frac{1}{2}+\beta\right).
\end{aligned}$$

Therefore, we have:

$$\begin{aligned}
\Psi^J &= \mu_{SS}(1-\Delta) + \mu_{SN}(1+\frac{1}{2}(1-\Delta)) - \frac{\lambda}{2} \\
&\quad - \left(\frac{1}{2}(1-\Delta)\mu_{SS} + \mu_{SN}(\beta + (1-\beta)(1-\Delta)\frac{1}{2}) - \mu_{SN}\frac{1}{2}(1-\beta) - \mu_{NN}(\frac{1}{2} - \frac{1}{2}\beta^2)\right), \\
\Psi^0 &= \mu_{SS}(1+k(1-\Delta)) + \mu_{SN}(1+((1-\beta)k(1-\Delta)+\beta)) - \lambda \\
&\quad - \left(\frac{1}{2}(1-\Delta)\mu_{SS} + \mu_{SN}(\beta + (1-\beta)(1-\Delta)\frac{1}{2}) - \mu_{SN}\frac{1}{2}(1-\beta) - \mu_{NN}(\frac{1}{2} - \frac{1}{2}\beta^2)\right), \\
\Psi^1 &= \mu_{SS}(1-\Delta) + \mu_{SN}(1+\frac{1}{2}(1-\Delta)) - \frac{\lambda}{2} \\
&\quad - ((1-\Delta)(1-k)\mu_{SS} - \mu_{SN}(1-\beta)(1-k) + (\mu - (1-\beta)\mu_{NN})).
\end{aligned}$$

From the above expressions, it is straightforward to see that Ψ^J, Ψ^0 and Ψ^1 are all strictly decreasing in Δ. Since Ψ^J, Ψ^0, Ψ^1 are also decreasing in β from Lemmas 11-13, it follows that an increase in Δ from 0 to a positive value shifts the JP, PI, FOMO curves to the left in the (β, ρ) space.

Let $\left(\frac{d\beta}{d\Delta}\right)^J$ denote the derivative of β along the JP curve with respect to Δ: this is a measure of how much the JP curve shifts to the left as Δ increases. Similarly, $\left(\frac{d\beta}{d\Delta}\right)^0$ and $\left(\frac{d\beta}{d\Delta}\right)^1$ measure the corresponding shifts of the PI and FOMO curves. We obtain

$$\begin{aligned}
\left(\frac{d\beta}{d\Delta}\right)^J &= -\frac{\mu_{SS} + \beta\mu_{SN}}{2\mu_{SN} + \Delta\mu + 2\beta\mu_{NN}} < 0, \\
\left(\frac{d\beta}{d\Delta}\right)^0 &= \frac{-(\mu_{SN} + \mu_{SS} - \beta\mu_{SN})(2k-1)}{2k\mu_{SN} + \Delta\mu_{SN} + 2\beta\mu_{NN} - 2k\Delta\mu_{SN}} < 0, \\
\left(\frac{d\beta}{d\Delta}\right)^1 &= -\frac{\mu_{SN} + 2k\mu_{SS}}{2(\mu_{SN} + \mu_{NN} - k\mu_{SN})} < 0.
\end{aligned}$$

Consider first Case 3, where we have $\lambda \in (\frac{1}{2}, \frac{2}{3})$. Note that for (β, ρ) to lie to the right of the JP curve, it is necessary that $\beta \geq \sqrt{\frac{1-\lambda}{\lambda}}$, which implies $\beta > \frac{1}{2}\sqrt{2}$.

We first show that $\left(\frac{d\beta}{d\Delta}\right)^J < \left(\frac{d\beta}{d\Delta}\right)^0$ when $\Delta = 0$:

$$\begin{aligned}
&\left(\frac{d\beta}{d\Delta}\right)^0 - \left(\frac{d\beta}{d\Delta}\right)^J \\
&= \frac{-(\mu_{SN} + \mu_{SS} - \beta\mu_{SN})(2k-1)}{2k\mu_{SN} + 2\beta\mu_{NN}} - \left(-\frac{\mu_{SS} + \beta\mu_{SN}}{2\mu_{SN} + 2\beta\mu_{NN}}\right) \\
&= \frac{A}{(2\mu_{SN} + 2\beta\mu_{NN})(2k\mu_{SN} + 2\beta\mu_{NN})},
\end{aligned}$$

where

$$\begin{aligned} A &= 4k\mu_{SN}\mu_{NN}\beta^2 + 2\left(\mu_{SN}\mu_{NN} - \mu_{SN}^2 + 2\mu_{NN}\mu_{SS} + 3k\mu_{SN}^2 - 2k\mu_{SN}\mu_{NN} - 2k\mu_{NN}\mu_{SS}\right)\beta \\ &+ 2\mu_{SN}^2 + 2\mu_{SN}\mu_{SS} - 4k\mu_{SN}^2 - 2k\mu_{SN}\mu_{SS}. \end{aligned}$$

The numerator A is equal to $(\mu_{SN} + 2\beta\mu_{NN})(\mu_{SS} + \beta\mu_{SN}) > 0$ when $k = \frac{1}{2}$, and to $2(2\beta - 1)(\mu_{SN} + \beta\mu_{NN})\mu > 0$ when $k = 1$. Since A is linear in k, it follows that $A > 0$ for all $k \in (\frac{1}{2}, 1)$. Therefore, $\left(\frac{d\beta}{d\Delta}\right)^J < \left(\frac{d\beta}{d\Delta}\right)^0$, which implies that, as Δ increases from 0 to a small positive value, the JP curve shifts to the left more than the PI curve.

We now prove $\left(\frac{d\beta}{d\Delta}\right)^1 < \left(\frac{d\beta}{d\Delta}\right)^J$:

$$\begin{aligned} &\left(\frac{d\beta}{d\Delta}\right)^J - \left(\frac{d\beta}{d\Delta}\right)^1 \\ = &-\frac{\mu_{SS} + \beta\mu_{SN}}{2\mu_{SN} + 2\beta\mu_{NN}} - \left(-\frac{\mu_{SN} + 2k\mu_{SS}}{2(\mu_{SN} + \mu_{NN} - k\mu_{SN})}\right) \\ = &\frac{(2\beta\mu_{SN}^2 + 6\mu_{SS}\mu_{SN} + 4\beta\mu_{NN}\mu_{SS})k + 2\mu_{SN}^2 - 2\mu_{SN}\mu_{SS} - 2\mu_{NN}\mu_{SS} - 2\beta\mu_{SN}^2}{2(\mu_{SN} + \mu_{NN} - k\mu_{SN})(2\mu_{SN} + 2\beta\mu_{NN})} \end{aligned}$$

When $k = 1$, the numerator is equal to $4\mu_{NN}\mu_{SS}\beta + 2\mu_{SN}^2 + 4\mu_{SS}\mu_{SN} - 2\mu_{NN}\mu_{SS}$, which is greater than $4\mu_{NN}\mu_{SS}\frac{1}{2} + 2\mu_{SN}^2 + 4\mu_{SS}\mu_{SN} - 2\mu_{NN}\mu_{SS} = 2\mu_{SN}(\mu_{SN} + 2\mu_{SS}) > 0$. Since $k \geq \frac{\lambda}{2(1-\lambda)}$ in Case 3, we note that at $k = \frac{\lambda}{2(1-\lambda)}$ the numerator is equal to λ times

$$-\lambda(7\lambda - 2 - 5\lambda^2)(1-\beta)\rho^2 + \lambda(6\beta + 19\lambda - 10\lambda^2 - 14\beta\lambda + 10\beta\lambda^2 - 9)\rho + (7\lambda - 2 - 5\lambda^2)(1 - \lambda + \beta\lambda), \tag{11}$$

which is obtained by replacing $\mu_{SS}, \mu_{SN}, \mu_{NN}$ with their respective expressions. Since $\lambda \in (\frac{1}{2}, \frac{2}{3}]$, (11) is concave in ρ. At $\rho = 0$, it equals $\lambda(5\lambda - 2)(1 - \lambda)(1 - \lambda + \beta\lambda) > 0$ for all $\lambda \in (\frac{1}{2}, \frac{2}{3}]$. At $\rho = 1$, it equals $2\lambda(\lambda + \beta\lambda - 1) \geq 2\lambda\left(\lambda + \sqrt{\frac{1-\lambda}{\lambda}}\lambda - 1\right) = 2\lambda\sqrt{1-\lambda}\left(\sqrt{\lambda} - \sqrt{1-\lambda}\right) > 0$ for all $\lambda \in (\frac{1}{2}, \frac{2}{3}]$, where the first inequality follows from $\beta \geq \sqrt{\frac{1-\lambda}{\lambda}}$. Therefore (11) is positive for all $\rho \in (0, 1)$, implying that the numerator is positive for all $k \in (\frac{\lambda}{2(1-\lambda)}, 1)$. As a result, $\left(\frac{d\beta}{d\Delta}\right)^1 < \left(\frac{d\beta}{d\Delta}\right)^J$ in Case 3. Combined with $\left(\frac{d\beta}{d\Delta}\right)^0 > \left(\frac{d\beta}{d\Delta}\right)^J$, this implies $\left(\frac{d\beta}{d\Delta}\right)^0 > \left(\frac{d\beta}{d\Delta}\right)^1$.

Therefore, in Case 3, we have the following results:

- Since the race to the bottom occurs between the JP curve and the PI curve, the inequality $\left(\frac{d\beta}{d\Delta}\right)^J < \left(\frac{d\beta}{d\Delta}\right)^0$ implies that the reduction in profit from competition expands the area of the race to the bottom.

- Since the FOMO-driven race to the bottom arises between the pointwise maximum of the JP and PI curves and the pointwise minimum of the FOMO curve

and the vertical line at $\beta = 1$, the ordering $\left(\frac{d\beta}{d\Delta}\right)^1 < \left(\frac{d\beta}{d\Delta}\right)^J < \left(\frac{d\beta}{d\Delta}\right)^0$ implies that the reduction in profit from competition expands the area of the FOMO-driven race to the bottom between the pointwise maximum of the JP and the PI curves and the vertical line, while shrinking the area between the former and the FOMO curve.

Consider now Case 4, where $\lambda \in (\frac{1}{2}, \frac{2}{3})$. We obtain the following results:

- Regarding the region of insufficient entry with (0,0) for high ρ, since $\left(\frac{d\beta}{d\Delta}\right)^J < \left(\frac{d\beta}{d\Delta}\right)^0 < 0$, the reduction in profits caused by competition shifts the JP curve to the left more than the PI curve, thereby shrinking the region horizontally. However, because the horizontal FOMO curve moves downward,[33] the region expands vertically.

- Regarding the region of insufficient entry with (1,0), the PI curve shifts to the left while the horizontal FOMO curve moves downward. As a result, the reduction in profits from competition shrinks the region horizontally (for a given A2 curve) but expands it vertically.[34]

- Regarding the region of the FOMO-driven race to the bottom, the vertical FOMO curve shifts to the left more than the JP curve. Consequently, the reduction in profits from competition shrinks this region.

[33] This occurs because an increase in Δ reduces Ψ^1, and Ψ^1 is decreasing in ρ when (4) is violated: see Lemma 11. In particular, Ψ^1 decreases because the term $\mu_{SN}(1+\frac{1}{2}(1-\Delta))$ falls, and the reduction in $\mu_{SS}(1-\Delta)$ is larger than that in $\mu_{SS}(1-\Delta)(1-k)$, since $1-k<1$.

[34] Since the A2 curve itself moves to the right, this further shrinks the region.