Runs on Artificial Intelligence Platforms

Balbinder Singh Gill*

Abstract

An AI platform produces its quality from participation withdrawable on demand. Because its most informative contributors are its most mobile, a confidence shock strips the best data first, so quality falls faster than headcount. A global-games refinement gives a unique equilibrium with an exact average-quality threshold, an endogenous run probability, and a panic component priced when runs fall in high-marginal-utility states. Three results depart from the network-effects-as-moats benchmark: fragility is set by the sensitivity of quality to participation, not its level; it is hump-shaped in network intensity; and platform equity carries a left-tail disaster component scaled by the panic width.

Keywords: Artificial intelligence platforms; runs and coordination failure; global games; network effects; systemic risk; platform-run risk premium.

JEL classification: G01, G12, D82, D85, L86, C72.

*Stevens Institute of Technology. Email: balbindersgill@gmail.com. An Internet Appendix, included at the end of this document, supplements the main text: it develops the continuous-time version of the model, in which a run is a first-passage event and the price of platform survival has a closed-form term structure; it presents three extensions that sharpen the main results and delimit the model's scope (a strategic platform response during a run, an endogenous stochastic discount factor, and N-platform contagion), together with robustness results for correlated signals and non-monotone informativeness and a calibrated illustration; it maps the paper's central predictions to empirical proxies; and it collects all proofs.

1 Introduction

How safe are the dominant firms of the artificial intelligence economy, and how should their risk be priced? A small number of platforms now sit beneath a widening share of economic activity (software development, customer service, analytics, research, content production) and capital markets capitalize their ecosystem rents at trillions of dollars. The real economy that depends on them, and the equity that prices them, both treat that dominance as durable. Whether it actually is, and how the risk of its failure should be priced, is the subject of this paper.

The frameworks in which the question is usually asked do not address it. The finance of digital infrastructure focuses on stablecoins, money-market funds, and other claim-like instruments in which the productive layer is taken as given (Kacperczyk and Schnabl, 2013; Schroth et al., 2014; Falato et al., 2021; Gill and John, 2026a,c; Gill, 2025); the economics of data treats data as a nonrival input to growth and firm size (Jones and Tonetti, 2020; Farboodi and Veldkamp, 2021), not as the substance run on. Neither captures the feature that distinguishes an AI platform: its asset (model quality, ecosystem, capability) is produced in real time by the very participation that can be withdrawn on demand. The finance literature on payoff complementarities and fragility (Diamond and Dybvig, 1983; Goldstein and Pauzner, 2005; Chen et al., 2010; He and Xiong, 2012; Goldstein et al., 2017) contains the right machinery for that feature, but has been applied to financial claims, not to the production of the asset itself.

The conventional wisdom, inherited from the economics of network goods, says dominance is stable. Network effects are read as moats: once an installed base is built, switching costs lock users in and entrench the incumbent, so the largest platform is the safest and dominance self-reinforcing (Katz and Shapiro, 1985). On this view the only serious risk to a dominant AI platform is that a better technology eventually displaces it, a gradual, fundamentals-driven decline, not a sudden break, and capital markets capitalize ecosystem rents as durable assets. The gap between this view and the alternative we develop is the empirical stake of the paper.

The framework these valuations rest on inverts under the conditions that characterize an AI platform. Quality is not a stock sitting on a balance sheet; it is produced, in real time, out of participation that users can withdraw on demand (their data, their revenue, their developer effort), and the participation each user can withdraw enters quality together with everyone else's. When quality rises with participation and participation is observable only through noisy signals, the decision to stay is a strategic complement to others' decisions to stay, so the platform's value is a coordination equilibrium in the same sense a bank's solvency is. Embedding the platform in a global game, we prove a unique threshold equilibrium whose cutoff is an explicit average-quality fundamental, $\theta^* = \kappa - \Delta\bar{H}$, exact for every noise level (Theorem 2). Here κ is the participation hurdle, Δ is the complementarity strength (how much quality rises from no to full participation) and $\bar{H} \in (0, 1)$ is the average curvature of the adoption curve. The indeterminate "belief" of the textbook account is thus replaced by a determinate run probability. The very complementarity that the network-goods literature reads as a moat is, under incomplete information, the amplifier of a run on the platform. It does not entrench dominance; it creates a knife-edge between dominance

and collapse, of width $\Delta(1-\bar{H})$ in the fundamental, the band of fundamentals that would survive full participation yet collapse on a coordination failure (Figure 1).

The reversal sharpens once the complementarity is microfounded rather than assumed (Section 4). Quality is the output of a statistical learning technology whose input is user-generated data, and data is heterogeneous in informativeness. Because the most informative contributors (sophisticated, multi-homing, low-lock-in users) are also the most mobile, a confidence shock strips out the best data first, and quality falls faster than headcount. This *data-composition* channel has no analogue in deposit runs, where claims are homogeneous and the asset is fixed, or in headcount network effects, where only the count matters. The failure mode is correspondingly more severe than a relabeled bank run: the platform loses the *informative* users first, so the asset that produces value collapses before the user base does.

The interaction of these forces yields two further results. First, holding peak quality fixed, a learning technology with weaker diminishing returns to data (the improvement AI labs race for) is strictly more run-prone, because each marginal increment of data is more pivotal: capability is bought with fragility (Proposition 7). Second, the multi-homing that diversifies any single firm's exposure to platform-run risk is exactly the channel that synchronizes runs across platforms: the informative users are the footloose multi-homers, so the same instrument that reduces idiosyncratic platform risk raises systemic platform risk (Theorem 3, Corollary 6). Neither result follows from mapping the bank-run framework onto platforms; both arise from the interaction of the model's own components.

Further implications follow. Fragility is set by the *sensitivity* of quality to participation, $\Delta = q(\theta,1)-q(\theta,0)$, not by its level: two platforms with identical capability can differ sharply in fragility (Proposition 6). Fragility is hump-shaped in network intensity once data and compute channels saturate, so the very largest platforms can be safer than mid-sized ones, reversing the reversal at the top end (Section 9, Figure 2). The panic component of run risk, though a pure coordination artifact, is priced and undiversifiable within an ecosystem (Proposition 13); the diversification a single firm engineers across platforms is what makes the system more correlated. In a dynamic setting the run becomes a first-passage event with a closed-form term structure of survival probabilities (developed in the Internet Appendix). The policy instruments that could arrest a panic each target a different externality at a different cost. Portability and interoperability cut a run's severity but raise its frequency and erode the rents that finance frontier investment. A public compute-and-data backstop can remove the bad equilibrium at the price of dulling private resilience, a digital too-big-to-fail. Designating the most-depended-upon platforms as systemically important digital infrastructure addresses under-provision head-on while risking entrenchment. The model identifies which externality each instrument reaches, and what it sacrifices.

The paper's contribution is to place AI platforms within finance's theory of fragile institutions and to identify what makes their fragility unlike the bank-run case the theory was built for: the asset being run on is produced by the run-able flow. The argument is mechanical, not analogical, each formal claim is verified from primitives. The model is built on the architecture of the modern

Table 1: Three runs, one mechanism.

	Bank run	**Stablecoin run**	**AI-platform run**
What is withdrawn	Deposits / money	Token redemptions	Trust, data, developer effort, adoption, complements
Source of fragility	Maturity mismatch	Reserve uncertainty	Retention–quality feedback (Δ)
What is lost	Liquidity	The peg	Data, revenue, quality, complements
Threshold object	Withdrawal cutoff	Redemption cutoff	Average-quality threshold
Backstop	Lender of last resort	Reserve rescue	Compute/data backstop, portability, SIDI designation

theory of coordination failures (Diamond and Dybvig, 1983; Carlsson and van Damme, 1993; Morris and Shin, 1998, 2003; Goldstein and Pauzner, 2005; He and Xiong, 2012), and engages with the finance literatures on payoff complementarities and runs, non-bank fragility, disaster-risk asset pricing, intangibles and data, and financial-network contagion, as developed in the paragraphs that follow. Table 1 places the platform case beside its two financial precedents, the bank run and the stablecoin run. Reading down the rows, all three are coordination failures: each turns on whether participants expect others to stay, and each has a fundamental threshold below which withdrawal becomes self-fulfilling. What differs is the object. In a bank run depositors withdraw money and the asset that backs it is fixed; in a stablecoin run holders redeem tokens against a fixed reserve; in a platform run participants withdraw the data, effort, and adoption that *produce* the asset itself, so the quality being run on degrades as the run proceeds. The platform's threshold is therefore an average-quality cutoff rather than a liquidity or peg cutoff, its source of fragility is the retention–quality feedback Δ rather than maturity mismatch or reserve uncertainty, and its backstops act on compute, data, and portability rather than on a lender or reserve of last resort. The feature that sets the platform apart is visible in the first row: what is withdrawn is what produces value.

The paper sits at the intersection of four literatures, and its contribution is sharpest stated against each in turn. The equilibrium technology is the global-games selection of Carlsson and van Damme (1993) and Morris and Shin (1998, 2003), applied to runs as in Diamond and Dybvig (1983) and refined by Goldstein and Pauzner (2005), Rochet and Vives (2004), and He and Xiong (2012); the empirical link from payoff complementarities to fragility is established in finance by Chen et al. (2010), Goldstein et al. (2017), and, in the recent COVID-era stress, Falato et al. (2021), with non-bank financial runs documented in Kacperczyk and Schnabl (2013) and Schroth et al. (2014) and run dynamics in stablecoins and other digital-asset protocols studied in Gill and John (2026a), Gill and John (2026c), Gill (2025), and Gill and John (2026b). In all of these the object withdrawn is a financial claim against a fixed asset and claimants are homogeneous, so a run concerns the sequencing of identical claims under a liquidity or solvency constraint. Our point of departure is that the object withdrawn is a production input (data, effort, adoption) so the asset's quality is endogenous to participation: what is withdrawn is what produces value. Fragility is then

governed by the participation-sensitivity Δ and the curvature $\bar{H}$ of the quality-from-participation map, objects with no analogue in the financial-runs literature. Where that literature carries run logic across financial claims, asset-backed commercial paper, money-market funds, bond funds, we carry it to non-financial digital infrastructure and price the resulting exposure of the real firms that depend on the platform.

On the data side, Jones and Tonetti (2020) and Farboodi and Veldkamp (2021) establish data as a nonrival input with increasing but bounded returns, and Crouzet et al. (2022) develop the implications of non-rivalry and limited excludability for intangible-capital pricing and firm scale, the closest finance analogue to platform quality produced by participation. We adopt the bounded-returns foundation for the learning technology but invert its use: the same data feedback that raises efficiency becomes, under incomplete information and demandable participation, a source of fragility; and we add adverse selection in exit, which this literature does not contain. The asset-pricing layer connects to the rare-disaster tradition of Barro (2006), Gabaix (2012), and Wachter (2013) and to structural credit risk (Merton, 1974; Black and Cox, 1976) through a first-passage representation, but the disaster is a coordination event and the barrier is a coordination threshold rather than a solvency boundary; the panic component is priced and undiversifiable within an ecosystem. The cross-platform contagion is a financial-networks result (Allen and Gale, 2000; Acemoglu et al., 2015; Elliott et al., 2014), but the linkage is the composition of the data that produces quality, not balance-sheet exposures; the downstream-sector cascades use a production-network logic (Acemoglu et al., 2012) but the upstream shock is a participation-driven quality collapse rather than a sectoral productivity shock. The platform and network-effects tradition (Katz and Shapiro, 1985) supplies the complementarity but reads it as a moat; we show the same force is a run amplifier.

The remainder of the paper is organized as follows. Section 2 motivates the mechanism, grounds it in documented episodes (Section 2.6), and maps premises to assumptions. Section 3 specifies primitives. Section 4 derives the quality function from a learning technology and introduces the data-composition mechanism. Section 5 solves the common-knowledge benchmark, including the interior equilibrium and its instability. Section 6 proves the uniqueness theorem and characterizes the run probability and triggers. Section 7 endogenizes integration and severity. Section 8 states the five main propositions. Section 9 endogenizes the complementarity, derives the reinforcement matrix, and establishes the stalling and reversal results. Section 10 specifies resilience and the downstream sector. Section 11 develops asset pricing. Section 12 extends the model to two platforms and derives cross-platform contagion through shared contributors. Section 13 covers welfare and policy, Section 14 distills the five central predictions, and Section 15 concludes. The Internet Appendix develops the continuous-time version of the model, in which the run is a first-passage event with a closed-form term structure of survival probabilities. It then presents three extensions that sharpen the main results and delimit the model's scope, a mapping from the central predictions to empirical proxies, and all proofs.

2 The Economic Mechanism

Before formalizing, we motivate the central mechanism and explain why it is specific to AI platforms rather than a relabeling of bank-run theory or of standard network effects. The argument has three institutional premises, a contrast with classical network goods, and an account of why the relevant object is a discontinuous *run*. We then map each premise to a modeling assumption, so that the assumptions are earned rather than assumed for convenience.

2.1 What we mean by an AI platform

The model applies to any digital infrastructure whose delivered quality is produced from current user participation that can be withdrawn on demand. In the AI economy, this includes three concrete classes of firm. First, *frontier foundation-model providers*: firms that train and serve large general-purpose models (text, image, audio, code) and whose usage data, human-feedback loops, and revenue flows fund the next generation of training. Second, *participation-produced knowledge platforms*: platforms whose value is the corpus of user-contributed knowledge, tools, or code that an AI layer surfaces or augments (developer Q&A platforms, AI-assisted reference and creative-content sites, marketplaces of fine-tuned models). Third, *AI-application ecosystems*: vertical platforms that combine a model layer with a developer ecosystem of agents, plug-ins, and integrations, where the developer side is itself a participation flow. What unites the three classes is that quality is the output of an estimation or aggregation technology operating on inputs that participants supply and can withdraw, not a stock of weights or a fixed asset. What is out of scope is any digital service in which the production technology has no live-participation input (a pure software product whose quality does not respond to current usage), and any financial claim against an AI firm in which the run is on the claim rather than on the production input. The data-composition mechanism (Section 4) is what distinguishes the three classes above from a generic network-goods environment.

2.2 Three premises

Premise 1: quality is produced from live participation, not stored. A frontier model's delivered quality is not a fixed asset sitting on a balance sheet; it is regenerated continuously out of the flow of current usage. Interaction logs and human feedback are the inputs to fine-tuning and alignment; subscription and API revenue are the inputs to the compute and talent that train the next checkpoint; third-party tools, plug-ins, and integrations (built and maintained by developers who are themselves participants) are what make the platform usable for the marginal adopter. Each input is a *flow* supplied voluntarily and contemporaneously. This dependence is the decisive structural difference from a bank. A bank's assets are largely fixed when a run begins, so a run is a liquidity event against a given asset. Here the "assets" (model quality, ecosystem, capability) are *produced by the very participation that can be withdrawn*, so a withdrawal degrades the asset's quality directly and immediately. The object of withdrawal is a production input, not a claim.

Premise 2: withdrawal is fast and demandable. The participation that produces quality can

be withdrawn on short notice and at low coordination cost. An enterprise can re-point an API integration to a rival in days; a developer can stop maintaining a plug-in immediately; a user can cease the interactions that generate training data with no notice at all. Unlike physical-capital relationships, there is no long lead time that smooths exit. Demandability is what makes the strategic problem a *run* rather than a gradual reallocation: exit is a discrete, near-immediate action, and its payoff depends on how many others exit at the same time.

Premise 3: triggers are commonly observed. The events that move beliefs about a platform's viability are public and salient: an outage, a benchmark loss to a rival, a security or privacy incident, a regulatory action, a price increase, a flagship-product discontinuation. Common observation is what allows beliefs to move together, which (Section 6) is precisely the channel through which a low-fundamental-content event can coordinate a collapse.

2.3 Why this is not ordinary network effects

The classical network-effects literature (Katz and Shapiro, 1985) treats the installed base as a *stock* that, once built, entrenches dominance: switching costs and compatibility lock users in, and the incumbent is stable precisely because the base is sticky. The AI-platform mechanism inverts this. Because quality is produced from a *flow* of live participation (Premise 1) that is demandable (Premise 2), the "installed base" is not a stock that must be slowly eroded but a flow that can evaporate as fast as beliefs change. The same complementarity that the classical literature reads as a moat (quality rising in participation) is here a transmission mechanism for collapse: it makes each participant's value depend on others' continued presence, so a coordinated withdrawal is self-validating. Network effects do not buy stability; they buy a knife-edge between dominance and collapse whose width we characterize exactly (Theorem 1, Eq. 15). For this reason the model's fragility statistic is the participation *sensitivity* of quality, $\Delta = q(\theta, 1) - q(\theta, 0)$, not its level.

2.4 Why a run, not a decline

Three features make the outcome discontinuous rather than a smooth deterioration. First, strategic complementarity (quality increasing in participation) makes individual best responses increasing in aggregate participation, so small belief changes are amplified rather than damped. Second, the existence of dominance regions (Assumption 3) means the economy has two stable basins, thriving and collapse, separated by an unstable interior (Theorem 1); the dynamics move *away* from the interior toward a corner. Third, the global-games refinement (Section 6) selects between the basins as a function of the fundamental crossing a sharp threshold θ^*, so as fundamentals or beliefs deteriorate the platform does not glide down a demand curve; it sits at near-full participation until the threshold is breached and then collapses. The continuous-time version (Section IA.1) makes this literal: the platform survives until the fundamental first crosses θ^*, at which point the regime switches discontinuously. Gradual decline is possible only in the knife-edge interior that the analysis shows is never selected.

2.5 From premises to assumptions

The formal model is the minimal structure that encodes the three premises. Premise 1 (quality produced from participation) is the complementarity $q(\theta, n) = \theta + \Delta h(n)$ with $\Delta > 0$ and h increasing (Assumption 2), and is microfounded channel-by-channel in Section 9.1, where Δ is itself the equilibrium of the data–compute–ecosystem system rather than a free parameter. Premise 2 (demandability) is the binary, simultaneous-move action $a_i \in \{S, E\}$ with an outside option net of a switching cost (Section 3.1); the switching cost is endogenized in Section 7. Premise 3 (common observation) is the information structure of Section 6: private signals deliver a unique threshold, while the public-signal extension (Proposition 2) shows how common observation restores the coordinating power of low-content events. The dominance regions (Assumption 3) require only that fundamentals can be good enough to support the platform unilaterally or bad enough to doom it, an unobjectionable richness condition. No assumption is doing covert work: the substantive results below turn on the *shape* and *equilibrium determination* of these objects, not on their mere existence.

2.6 The mechanism in practice

The model is not a thought experiment in search of an application; its objects already correspond to documented episodes in the AI-platform economy. Three are worth stating, because each pins down a distinct piece of the model and shows that the relevant phenomena are observed rather than hypothesized.

A confidence-driven run with no change in fundamentals. In November 2023 the governance crisis at a leading AI platform, a board's abrupt dismissal of its chief executive, triggered, within a single weekend, contemporaneous reports that more than one hundred enterprise customers approached a direct rival, with others contacting further competitors and cloud providers. Nothing about the underlying model had changed; what changed was the commonly observed signal and the belief about whether others would stay. One departing customer's stated reason (that the platform had been chosen for its *perceived stability*) is exactly the logic of the model: the asset being withdrawn is confidence, the trigger is a coordinating signal rather than a fundamental shock (Definition 4), and the withdrawal is demandable and near-immediate. The episode is a clean instance of a panic run on the coordination region, and the migration dynamics it set off have continued, with capital and enterprise demand reallocating across platforms on governance and policy news rather than on measured quality.

A coordinating signal that reprices the moat. In late January 2025 the release of a low-cost open competitor was a commonly observed, low-fundamental-content event. It nonetheless erased roughly $589 billion of a dominant supplier's market value on a single day, the largest single-day market-capitalization loss in U.S. stock-market history. The episode forced a broad re-examination of whether incumbents' data and compute advantages were durable moats or fragile complementar-

ities.[1] Once a dominant position is understood as confidence-sensitive, small revisions to assumed growth and discount rates translate into large swings in value. This is the asset-pricing content of the model: platform equity carries a left-tail disaster component tied to the panic width, and the market reprices it when a salient signal shifts beliefs about the moat (Section 11, Remark 9).

A run on a participation-produced asset, with composition shifting. The decline of the dominant programmer question-and-answer platform after the diffusion of AI coding assistants is a near-literal instance of the model's central premise. The platform's value (a stock of community knowledge) was produced entirely by its contributors; once a superior outside option appeared, contributors left, the content thinned, and the platform became less useful, accelerating exit in a self-undermining loop. Monthly contributions fell by roughly nine-tenths from their 2020 peak (Stack Overflow's own public Data Explorer), and the *composition* of what remained shifted, not merely its volume: independent contributor-level studies have documented a systematic change in the difficulty and length of remaining questions consistent with newer, marginal contributors withdrawing first and leaving a smaller, harder residual. This is the data-composition channel of Section 4: quality is produced from participation, the marginal departures are not representative, and the asset degrades faster than headcount.

These episodes do more than motivate. They discipline the modeling choices: the trigger taxonomy must include coordinating signals with little fundamental content (it does, Definition 4); the asset-pricing layer must deliver a confidence-driven disaster component rather than ordinary cash-flow risk (it does, Section 11); and the quality technology must make the produced good depend on the composition, not merely the count, of participants (it does, Section 4). The practical relevance of the theory is therefore not an afterthought to the formalism but the reason its specific objects are the right ones.

3 The Model

We now translate the three premises of Section 2 into the simplest environment that can carry them, introducing each object by its economic meaning before its symbol. The whole model fits in one sentence: a large population of users each decide whether to keep using a platform whose quality rises with how many others keep using it, and each stays only if that quality is worth more than the best alternative net of the cost of switching. The pieces that follow are exactly the content of that sentence, the population and its decision, the technology that turns participation into quality, the payoff from staying versus leaving, and the regions of fundamentals in which the outcome is, or is not, a foregone conclusion.

[1] The market-value figure is widely reported in financial press and corresponds to the close-to-close move on January 27, 2025; the Stack Overflow contribution figures below are from the platform's own public Data Explorer. These contemporaneous data points are cited for dating and orders of magnitude, not as evidence.

3.1 Participants, actions, and information

Consider a single date $t = 0$ (Sections 10 and IA.1 add earlier and continuous dates). The platform serves a large population of participants: users, developers, and enterprise clients, treated symmetrically until Section IA.2. We model the population as a continuum, indexed by $i \in [0, 1]$ with Lebesgue measure ν, for a substantive economic reason rather than mathematical convenience: no single participant is pivotal, so the object that drives the platform's value is the *fraction* who stay, not any individual's choice. Each participant makes one decision, either to *stay* on the platform ($a_i = S$) or to *exit* to an outside option ($a_i = E$), and the aggregate of these decisions is the participation rate

$$n \;=\; \nu\big(\{i : a_i = S\}\big) \;=\; \int_0^1 \mathbf{1}\{a_i = S\}\, d\nu(i) \;\in\; [0, 1]. \tag{1}$$

The platform fundamental is a scalar $\theta \in \mathbb{R}$. It collects everything independent of current participation: stand-alone model quality, reliability, infrastructure robustness, legal and regulatory security, and accumulated capability. In Section 5, θ is common knowledge; in Section 6, each participant observes a private signal of θ.

Assumption 1 (Exact law of large numbers)**.** The collection $\{a_i\}_{i\in[0,1]}$ is essentially pairwise independent conditional on θ, and the index space is a Fubini extension in the sense of Sun (2006). Consequently the realized participation rate equals its conditional probability almost surely: $n = \mathbb{P}(a_i = S \mid \theta)$ for ν-a.e. i.

Assumption 1 is the standard device (Judd, 1985; Sun, 2006) that lets us treat the aggregate n as deterministic given θ and any symmetric strategy, eliminating aggregate noise from idiosyncratic signal draws. It is invoked only where we aggregate.

3.2 Quality, payoffs, and the participation hurdle

The value a participant gets from the platform has two parts: what the platform is worth on its own, and what it is worth because others are using it. We write delivered quality as the sum of the two,

$$q(\theta, n) \;=\; \theta \;+\; \Delta\, h(n), \tag{2}$$

where θ is the stand-alone fundamental of Section 3.1 and $\Delta\, h(n)$ is the participation-produced part, with $\Delta > 0$ and $h : [0, 1] \to [0, 1]$ describing how quality builds up as the user base fills in. Section 4 derives this object from a learning technology; here we record the properties the equilibrium analysis uses.

Assumption 2 (Network function)**.** h is continuous and strictly increasing on $[0, 1]$, with $h(0) = 0$ and $h(1) = 1$. Define $\bar{H} \equiv \int_0^1 h(n)\, dn$; then $\bar{H} \in (0, 1)$.

The normalization $h(0) = 0, h(1) = 1$ makes $\Delta = q(\theta, 1) - q(\theta, 0)$ the *total* quality gain from going from no participation to full participation; we call Δ the *complementarity strength*. Additive

separability (2) means Δ does not depend on θ; this is the cleanest carrier of the comparative statics and is relaxed in Remark 2. That $\bar{H} \in (0,1)$ for any strictly increasing h obeying the normalization is immediate and is the source of the panic region below.

A participant who stays obtains $u_S = q(\theta, n) - p$, where $p \geq 0$ is the platform's price (per-period fee, API charge, or subscription). A participant who exits obtains the outside option $R \geq 0$ (a rival platform, an open-source model, or an in-house system) net of a switching cost $c \geq 0$ capturing the depth of lock-in (migration effort, retraining, re-integration, lost workflow capital): $u_E = R - c$. The *payoff differential* from staying is

$$\mathcal{U}(\theta, n) \;\equiv\; u_S - u_E \;=\; q(\theta, n) - \kappa, \qquad \kappa \equiv p + R - c, \tag{3}$$

and a participant stays iff $\mathcal{U}(\theta, n) \geq 0$, i.e. iff

$$q(\theta, n) \;=\; \theta + \Delta h(n) \;\geq\; \kappa. \tag{4}$$

We call κ the *participation hurdle*: it rises with price p and the outside option R and falls with lock-in c. Because h is increasing (Assumption 2), $\mathcal{U}$ is strictly increasing in n: the game is one of strategic complementarities. It is also strictly increasing in θ (state monotonicity).

Remark 1 (Microfoundation of Δ). The complementarity decomposes into three channels, $\Delta\, h(n) = \phi_d\, h_d(n) + \phi_e\, h_e(n) + \phi_c\, k(n)$, each component continuous and increasing with the normalization absorbed into h, so that $\Delta = \phi_d + \phi_e + \phi_c$. *Data* ($\phi_d$): active participants generate interaction data and human feedback that improve fine-tuning and alignment, raising quality. *Ecosystem* (ϕ_e): participants attract developers, tools, and integrations, raising the value of being on the platform (an indirect network effect, Katz and Shapiro, 1985). *Compute* (ϕ_c): participation generates revenue ρn that finances compute, data acquisition, and talent, $k(n) = \mathcal{K}(\rho n)/\bar{\mathcal{K}}$ with $\mathcal{K}' > 0$, raising quality at rate ϕ_c. Every comparative static in Δ below therefore holds channel-by-channel in (ϕ_d, ϕ_e, ϕ_c). This decomposition is itself derived from a learning technology in Section 4, where the data channel and the curvature of h are obtained from the statistics of estimation and where the joint distribution of data informativeness and switching costs generates a mechanism, data-composition collapse, that is specific to AI platforms; Section 9 then closes the loop among the three channels.

Remark 2 (General quality functions). All qualitative results survive for any $q \in C^1$ with $\partial_\theta q > 0$ and $\partial_n q \geq 0$ (with $\partial_n q > 0$ on a positive-measure set), upon replacing $\Delta\bar{H}$ by $\int_0^1 [q(\theta^*, n) - q(\theta^*, 0)]\, dn$ and Δ by the participation sensitivity $q(\theta, 1) - q(\theta, 0)$. Additive separability buys closed forms, not the economics. We use (2) throughout and flag where separability is used.

3.3 Dominance regions

Assumption 3 (Dominance regions)**.** $\Delta > 0$, so the fundamental admits an upper-dominance and a lower-dominance region. With (2): staying is strictly dominant whenever $\theta > \overline{\theta} \equiv \kappa$ (then $q(\theta, 0) = \theta > \kappa$, so (4) holds even at $n = 0$), and exiting is strictly dominant whenever $\theta < \underline{\theta} \equiv \kappa - \Delta$

(then $q(\theta, 1) = \theta + \Delta < \kappa$, so (4) fails even at $n = 1$).

The two dominance regions and the intermediate *coordination region* $[\underline{\theta}, \overline{\theta}]$ are the only structure the equilibrium analysis needs. Their width is exactly the complementarity strength, $\overline{\theta} - \underline{\theta} = \Delta$.

3.4 Notation summary

The model's primitives and the derived quantities that recur are collected below. *Primitive* objects are taken as model inputs; *derived* objects are equilibrium outcomes of the analysis.

Symbol	Meaning	Status
θ	Stand-alone platform fundamental	primitive (random)
$n \in [0, 1]$	Participation rate (fraction staying)	equilibrium outcome
$q(\theta, n) = \theta + \Delta h(n)$	Delivered quality	derived
$h(n)$	Adoption-value curve, $h(0) = 0$, $h(1) = 1$	primitive
$\Delta = q(\theta, 1) - q(\theta, 0)$	Complementarity strength	primitive (§4), endogenous (§9)
$\bar{H} = \int_0^1 h(n)\, dn$	Average-quality curvature	derived from h
$\kappa = p + R - c$	Participation hurdle	primitive (price, outside, switching)
$\underline{\theta} = \kappa - \Delta$	Lower-dominance boundary	derived
$\overline{\theta} = \kappa$	Upper-dominance boundary	derived
$\theta^* = \kappa - \Delta\bar{H}$	Run threshold (Theorem 2)	derived
$[\underline{\theta}, \theta^*)$	Panic region; width $\Delta(1 - \bar{H})$	derived
$\pi_{\text{run}} = F(\theta^*)$	Run probability	derived
$\pi_{\text{panic}} = F(\theta^*) - F(\underline{\theta})$	Panic component	derived (Def. 3)
$\iota(u)$, $c(u)$	Informativeness and switching cost by quantile u	primitive (§4)
$\Xi(\mathcal{I})$	Learning technology mapping information to quality	primitive
$\bar{\delta}\bar{A}$	Aggregate downstream dependence weight	primitive (§10.1)
$D_t = \theta_t - \theta^*$	Distance to run	derived
$\zeta(t) = -d \ln S(t)/dt$	Run hazard rate	derived (§IA.1)
λ^{panic}	Panic premium	derived (§11.3, §IA.2.2)
$\mathbf{M} = [m_{ij}]$	Multi-homing intensity matrix	primitive (§12)

The participation hurdle $\kappa = p + R - c$ aggregates price, outside option, and switching cost; the complementarity strength Δ is microfounded in Section 4 and endogenized as the equilibrium of an interacting channel system in Section 9; the average-quality curvature $\bar{H}$ enters Theorem 2's exact threshold; the run probability $F(\theta^*)$ decomposes into fundamental and panic components in Definition 3.

4 Microfoundation: Quality as a Learning Technology, and the Data-Composition Mechanism

The model so far *posits* the participation–quality map h. A natural objection is that the central result (runs) then follows from the assumption that produces it, and that the complementarity is the same one any network-effects model writes down. This section answers both points. It derives h, Δ, and the curvature $\bar{H}$ from a primitive learning technology operating on user-generated data, and in doing so isolates a mechanism that is specific to AI platforms and absent from deposit runs and from headcount network effects: the quality of the produced good depends on the *composition*, not merely the count, of who participates, and the most valuable contributors are the most mobile. We call this the *data-composition* channel. The platform asset has two of the defining properties Crouzet et al. (2022) identify as central to intangible capital, non-rivalry of the data input (the same dataset trains many users simultaneously) and limited excludability of participation (contributors can withdraw the input on demand), and the model is, in effect, what those two properties imply once participation becomes a strategic complement.

4.1 Data generation and the estimation technology

Quality above the stand-alone level θ is not a network externality in headcount; it is the output of estimating a model from data that participants generate. Index participants by their switching-cost quantile $u \in [0,1]$, with switching cost $c(u)$ increasing in u (so low-u participants are the most footloose). Participant u generates data of *informativeness* $\iota(u) \geq 0$ (the Fisher information, or effective sample weight, her interactions and feedback contribute to training) with population mean $\bar{\iota} = \int_0^1 \iota(u)\,du > 0$. When a mass n participates, the stayers are those with the highest switching costs, in a monotone equilibrium with switching-cost-heterogeneous types, a higher switching cost means a lower hurdle κ and hence a lower signal cutoff, so high-c types stay at fundamentals where low-c types have already exited; the nested-cutoff equilibrium of Section 12.2 is the worked-out instance, i.e. the quantiles $[1-n, 1]$, and the platform's *effective information* is the order-statistic integral

$$\mathcal{I}(n) \;=\; \int_{1-n}^{1} \iota(u)\,du, \qquad \mathcal{I}(0) = 0, \;\; \mathcal{I}(1) = \bar{\iota}. \tag{5}$$

A learning technology Ξ maps information into delivered quality:

$$q(\theta, n) \;=\; \theta + \Xi\big(\mathcal{I}(n)\big), \qquad \Xi \text{ increasing, concave, } \Xi(0) = 0, \tag{6}$$

where concavity reflects *diminishing returns to data*: redundant interactions add little, and test loss falls as a sublinear power of training information (a scaling law). Define the complementarity and the network function as derived objects,

$$\Delta \;\equiv\; \Xi(\bar{\iota}), \qquad h(n) \;\equiv\; \frac{\Xi(\mathcal{I}(n))}{\Xi(\bar{\iota})}, \tag{7}$$

so that (6) is exactly $q(\theta, n) = \theta + \Delta h(n)$ with $h(0) = 0, h(1) = 1$ as in Assumption 2, but h, Δ, and (below) $\bar{H}$ are now consequences of the data process $\iota(\cdot)$ and the estimation technology Ξ, not primitives.

4.2 The homogeneous benchmark: curvature from scaling

Suppose first that informativeness is independent of switching cost, $\iota(u) \equiv \bar{\iota}$ (no selection). Then $\mathcal{I}(n) = \bar{\iota} n$ is linear, and $h(n) = \Xi(\bar{\iota} n)/\Xi(\bar{\iota})$ inherits the *concavity* of the learning technology: a homogeneous data process makes the platform *robust* (h concave, $\bar{H}$ large). In the scaling family $\Xi(\mathcal{I}) = A\mathcal{I}^{\alpha}$ with data-scaling exponent $\alpha \in (0, 1)$,

$$h(n) = n^{\alpha}, \qquad \bar{H} = \frac{1}{1+\alpha} \in \left(\tfrac{1}{2}, 1\right), \qquad \theta^* = \kappa - \frac{\Delta}{1+\alpha}. \tag{8}$$

The curvature parameter that Proposition 6 treats as a free shape parameter is thus pinned, in the homogeneous case, to the *data-scaling exponent*: with $\alpha < 1$ under diminishing returns, absent selection the adoption curve is concave and the platform sits toward the safe end of the coordination band.

4.3 The data-composition mechanism

Fragility (a convex h, small $\bar{H}$) does not come from the learning technology, which is concave. It comes from *who leaves*. The participants with the lowest switching costs are typically the most informative: sophisticated power users, multi-homing developers, and enterprises with the engineering capacity both to extract value (and so generate rich interaction data and feedback) and to migrate. When a run begins, these contributors leave first, so effective information falls faster than headcount.

Formally, let $\iota(u)$ be *decreasing* in u: informativeness falls with lock-in (the footloose are the informative). Then the marginal stayer at participation n is quantile $1 - n$, and

$$\mathcal{I}'(n) = \iota(1 - n) \ \text{ is increasing in } n,$$

so $\mathcal{I}(n)$ is *convex*: it lies below the no-selection chord $\bar{\iota} n$ for every interior n, with the gap governed by how strongly informativeness and footlooseness covary. Through (7) this makes the effective network function more convex and lowers $\bar{H}$. To keep the comparison explicit, write $h_{\text{eff}}(n) \equiv \Xi(\mathcal{I}(n))/\Xi(\bar{\iota})$ for the selection-inclusive adoption curve generated by a given $\iota(\cdot)$ (the general object of (7)) and reserve $h(n) \equiv \Xi(\bar{\iota} n)/\Xi(\bar{\iota})$ for the homogeneous ($\iota \equiv \bar{\iota}$) benchmark; $\bar{H}_{\text{eff}} \equiv \int_0^1 h_{\text{eff}}$ and $\bar{H} \equiv \int_0^1 h$ are their averages. The two coincide absent selection and satisfy $h_{\text{eff}} \leq h$ under destabilizing composition.

Proposition 1 (Microfounded quality and the data-composition mechanism)**.** *Let quality be* (6)–(7) *with* Ξ *increasing, concave,* $\Xi(0) = 0$*, and informativeness* $\iota(\cdot)$ *with mean* $\bar{\iota}$*.*

(i) Derivation. *h, $\Delta = \Xi(\bar{\iota})$, and $\bar{H} = \int_0^1 \Xi(\mathcal{I}(n))\,dn/\Xi(\bar{\iota})$ are determined by (ι, Ξ). Absent selection ($\iota \equiv \bar{\iota}$), h is concave; in the scaling family it is n^α with $\bar{H} = 1/(1+\alpha)$.*

(ii) Composition fragility. *If ι is decreasing in the switching-cost quantile (informative users are footloose), then $\mathcal{I}$ is convex, $h_{\text{eff}}(n) \le h(n)$ pointwise, $\bar{H}_{\text{eff}} < \bar{H}$, and hence $\theta^* = \kappa - \Delta\bar{H}_{\text{eff}}$ and the run probability $F(\theta^*)$ are strictly higher than in the homogeneous benchmark. In the parametric family $\iota(u) = (1+\varsigma)\bar{\iota}(1-u)^\varsigma$ and $\Xi(\mathcal{I}) = A\mathcal{I}^\alpha$,*

$$\mathcal{I}(n) = \bar{\iota}\, n^{1+\varsigma}, \qquad h_{\text{eff}}(n) = n^{\alpha(1+\varsigma)}, \qquad \bar{H}_{\text{eff}} = \frac{1}{1+\alpha(1+\varsigma)},$$

so the effective curvature is $\gamma_{\text{eff}} = \alpha(1+\varsigma)$, the product of the data-scaling exponent and one plus the selection intensity, and h_{eff} is convex (the fragile regime) iff $\varsigma > (1-\alpha)/\alpha$.

(iii) Run acceleration. *$q'(n) = \Xi'(\mathcal{I}(n))\,\iota(1-n)$: under selection the marginal stayer's informativeness exceeds the average for $n < 1$, so each exit degrades quality by more than its population share, a positive feedback through data composition with no analogue in deposit runs.*

$\bar{H}_{\text{eff}}$ is decreasing in the selection intensity ς and in the dispersion of ι; the sign of the effect is the sign of the informativeness–footlooseness covariance.

Remark 3 (Unconditional skeleton of the headline results). Proposition 1 characterizes the data-composition channel *conditional* on the sign of $\text{Cov}(\iota, c)$. The paper's headline structural results do not depend on this sign: the unique-equilibrium claim of Theorem 2, the closed form $\theta^* = \kappa - \Delta\bar{H}$ for any $\bar{H} \in (0,1)$, the hump shape of fragility in Δ (Proposition 10 and Corollary 4), and the priced panic component (Proposition 13) hold for every value of $\bar{H}$. What the data-composition channel determines is which $\bar{H}$ obtains for a given platform (lower under destabilizing composition, higher under stabilizing composition) and therefore which point on the unconditional fragility curve $P_{\text{panic}}(\Delta)$ each platform occupies. The sign-conditional content is the *level* of fragility and the *direction* of how quality declines relative to headcount during a run; the existence of fragility, its threshold structure, and its non-monotone cross-section are unconditional. The model's empirical content is in two layers: an unconditional skeleton of structural predictions, and a sign-conditional set of refinements that distinguishes empirical platform classes from each other.

4.4 Why this is a new mechanism

The composition channel is not a relabeling of existing models, and the contrast is sharp on primitives. In a deposit run (Diamond and Dybvig, 1983) the bank's assets are fixed and depositors are homogeneous claimants; the run sequences identical claims against a given asset. Here the asset (model quality) is *produced* from the participation that can be withdrawn, and its quality depends on the *composition* of who remains. In a headcount network effect (Katz and Shapiro, 1985) a participant's benefit rises with the *number* of others; here it rises with the *informativeness-weighted* composition, mediated by a statistical estimation technology, so the curvature of the adoption curve

is a scaling exponent and its convexity is manufactured by selection. In the global-games benchmark (Morris and Shin, 2003) agents are homogeneous; the engine of the new comparative statics here is the *covariance* between data informativeness and exit propensity.

The mechanism also delivers a sign that is *not* assumed: if instead lock-in correlates *positively* with informativeness (the platform's best data comes from its most captive users) then ι is increasing in u, $\mathcal{I}$ is concave, and selection is *stabilizing*, lowering run probability. Whether the data-composition channel amplifies or dampens runs therefore depends on the joint distribution of informativeness and switching costs, and is determined by primitives rather than imposed. In the leading case for frontier platforms, where the most sophisticated, multi-homing users are both the most informative and the most mobile, the channel amplifies runs, so platforms whose model quality leans on footloose power users are structurally the most run-prone; the collapse of the dominant programmer question-and-answer platform (Section 2.6), in which the routine contributions departed first and the surviving content grew harder and thinner, is the paradigm case.

5 Common Knowledge: Self-Fulfilling Runs

Take θ to be common knowledge. Building on the primitives of Section 3, this benchmark fixes ideas and locates the fragility; Section 6 then perturbs information to select among the equilibria found here. We define equilibrium, characterize the full equilibrium set (including the interior equilibrium usually left implicit), and identify which equilibria are stable.

Definition 1 (Equilibrium)**.** A (Nash) equilibrium is an action profile $a^\star$ and induced rate $n^\star$ via (1) such that for ν-a.e. i, $a_i^\star = S$ if $\mathcal{U}(\theta, n^\star) > 0$ and $a_i^\star = E$ if $\mathcal{U}(\theta, n^\star) < 0$. Equivalently, $n^\star$ is a fixed point of the aggregate best-response correspondence

$$\mathcal{B}(n) \;=\; \begin{cases} \{1\}, & q(\theta, n) > \kappa, \\ [0,1], & q(\theta, n) = \kappa, \\ \{0\}, & q(\theta, n) < \kappa. \end{cases}$$

Definition 2 (Best-response stability)**.** An equilibrium $n^\star$ is *stable* if the best-response dynamic $\dot{n} = b(n) - n$, where $b(n) = \mathbf{1}\{q(\theta, n) \geq \kappa\}$ is the (a.e.-defined) selection from $\mathcal{B}$, is locally asymptotically stable at $n^\star$; equivalently, if there is a neighborhood of $n^\star$ on which $b(n) - n$ has the sign of $n^\star - n$.

Theorem 1 (Equilibrium set and stability under common knowledge)**.** *Let Assumptions 2–3 hold and let θ be common knowledge.*

(i) If $\theta > \overline{\theta} = \kappa$: the unique equilibrium is $n^\star = 1$; the platform is run-proof*.*

(ii) If $\theta < \underline{\theta} = \kappa - \Delta$: the unique equilibrium is $n^\star = 0$; the platform fails on fundamentals.

(iii) If $\theta \in (\underline{\theta}, \overline{\theta})$: there are exactly three equilibria,

$$n^\star \in \{0,\ n_{\text{int}}(\theta),\ 1\}, \qquad n_{\text{int}}(\theta) = h^{-1}\Big(\tfrac{\kappa-\theta}{\Delta}\Big) \in (0,1).$$

The corner equilibria 0 and 1 are stable; the interior equilibrium $n_{\text{int}}(\theta)$ is unstable. At the boundaries $\theta \in \{\underline{\theta}, \overline{\theta}\}$ the interior equilibrium merges with a corner.

The width of the coordination region is $\overline{\theta} - \underline{\theta} = \Delta$.

Part (iii) makes precise the "two equilibria" intuition and disciplines it: there is also an interior equilibrium, but it is a knife-edge, a marginal increase in participation makes everyone strictly prefer to stay, pushing the system to $n = 1$, while a marginal decrease pushes it to $n = 0$. Only the corners survive any adjustment friction, so the economically relevant multiplicity is between *thriving* and *collapse*. The self-fulfilling run is the selection of $n^\star = 0$ on the coordination region even though $n^\star = 1$ would clear the hurdle: a participant expecting exodus expects $q(\theta, 0) = \theta < \kappa$ and exits; one expecting retention expects $q(\theta, 1) = \theta + \Delta \geq \kappa$ and stays. A trigger (model failure, breach, lawsuit, price hike, rival's rise) need not change θ; on the coordination region it can act purely as a device that selects the collapse equilibrium. The width of the region in which this is possible is governed one-for-one by Δ.

The benchmark leaves the run probability indeterminate, because equilibrium selection on $(\underline{\theta}, \overline{\theta})$ is unmodeled. The next section removes the indeterminacy.

6 A Unique Equilibrium via Global Games

We perturb the information structure so that participants do not share common knowledge of the platform's reliability, legal exposure, or competitive position (the realistic case) and prove that the perturbed game has a unique equilibrium. The threshold θ^* derived here is the object on which the rest of the paper rests: it organizes the main comparative statics (Section 8), is endogenized through the channel system (Section 9), determines run severity (Section 7), prices platform-run risk (Section 11), and becomes the first-passage barrier of the continuous-time model (Section IA.1).

6.1 Information structure

Assumption 4 (Diffuse prior and private signals)**.** The fundamental θ is drawn from the (improper) uniform prior on $\mathbb{R}$. Each participant i observes a private signal

$$x_i \ = \ \theta + \varepsilon\, \eta_i, \qquad \eta_i \overset{\text{i.i.d.}}{\sim} G, \tag{9}$$

with $\varepsilon > 0$. The noise distribution G has a continuous, strictly positive density g on $\mathbb{R}$, is symmetric about zero with mean zero, and the η_i are independent across i and of θ.

The improper uniform prior is the standard global-games benchmark; it can be obtained as the limit of any smooth proper prior as prior precision vanishes relative to signal precision, and it makes

the equilibrium threshold exact rather than merely a small-noise limit. A *strategy* is a measurable map $\sigma : \mathbb{R} \to [0,1]$ giving the probability of staying as a function of the signal; it is a *threshold* (monotone) strategy with cutoff $\hat{x}$ if $\sigma(x) = \mathbf{1}\{x \geq \hat{x}\}$. A (Bayes–Nash) equilibrium is a strategy profile, symmetric across participants, such that each plays a best response to the others given (9) and Assumption 1.

6.2 The aggregate map and the pivotal belief

Two lemmas do the work. The first turns a threshold strategy into a deterministic participation rate; the second establishes the Laplacian belief.

Lemma 1 (Participation map)**.** *If all participants use the threshold strategy with cutoff $\hat{x}$, then by Assumption 1 the participation rate at fundamental θ is deterministic and equal to*

$$n(\theta; \hat{x}) \;=\; \mathbb{P}(x_i \geq \hat{x} \mid \theta) \;=\; 1 - G\Big(\tfrac{\hat{x}-\theta}{\varepsilon}\Big). \tag{10}$$

It is continuous, strictly increasing in θ (from 0 to 1), and strictly decreasing in $\hat{x}$.

Lemma 2 (Laplacian belief: exact for every ε)**.** *Suppose all others use cutoff $\hat{x}$. Condition on being the* pivotal *participant, i.e. observing $x_i = \hat{x}$. Under Assumption 4, the posterior distribution of the participation rate $n(\theta; \hat{x})$ faced by the pivotal participant is uniform on $[0,1]$, for every $\varepsilon > 0$. Moreover the posterior of θ is $\theta = \hat{x} - \varepsilon z$ with $z \sim G$.*

The proof (the Internet Appendix) is the exact-belief computation: with an improper uniform prior the posterior density of θ given $x_i = \hat{x}$ is $\varepsilon^{-1} g((\hat{x} - \theta)/\varepsilon)$, so $z \equiv (\hat{x} - \theta)/\varepsilon \sim G$; since $n(\theta; \hat{x}) = 1 - G(z)$ and $G(z) \sim U[0,1]$ whenever $z \sim G$, we get $n \sim U[0,1]$ *exactly*, not just as $\varepsilon \to 0$. The marginal participant behaves as though the fraction of others who stay is equally likely to be anything.

Remark 4 (Noise distribution robustness). Lemma 2 requires only that G have a continuous strictly positive density and that the η_i be i.i.d.; symmetry of G is not used in the lemma. Symmetry enters Theorem 2's closed form only via $\mathbb{E}_z[\theta^* - \varepsilon z] = \theta^*$, which requires $\mathbb{E}_G[z] = 0$, a strictly weaker condition than symmetry. For any continuous G with mean zero, the closed form $\theta^* = \kappa - \Delta\bar{H}$ holds exactly for every $\varepsilon > 0$. For a noise distribution with non-zero mean μ_G, the closed form is $\theta^* = \kappa - \Delta\bar{H} + \varepsilon\mu_G$: the threshold acquires a noise-mean correction that vanishes as $\varepsilon \to 0$. Asymmetric or skewed noise with mean zero does not enter the threshold at all. Uniqueness is preserved in every case by the standard monotonicity-of-Ψ argument; what depends on the noise mean is only the level of θ^* at non-trivial noise, not its structure or its comparative statics.

6.3 Uniqueness and the exact threshold

Theorem 2 (Unique equilibrium; exact average-quality threshold)**.** *Under Assumptions 1–3 and 4, the global game has a unique Bayes–Nash equilibrium. It is in threshold strategies, with cutoff*

$\hat{x} = \theta^*$ *given by the indifference condition*

$$\Psi(\theta^*) \;\equiv\; \mathbb{E}_{z\sim G}\big[q(\theta^* - \varepsilon z,\, 1 - G(z))\big] \;=\; \kappa. \tag{11}$$

With the additively separable quality (2) *and zero-mean noise,* (11) *solves in closed form,* exactly for every $\varepsilon > 0$*:*

$$\theta^* \;=\; \kappa - \Delta\bar{H} \;=\; p + R - c - \Delta\bar{H}. \tag{12}$$

The realized participation rate is $n(\theta;\theta^*) = 1 - G((\theta^* - \theta)/\varepsilon)$*. As* $\varepsilon \to 0$*,* $n(\theta;\theta^*) \to \mathbf{1}\{\theta > \theta^*\}$*: the platform thrives if* $\theta > \theta^*$ *and suffers a run (collapses) if* $\theta < \theta^*$*.*

The proof (the Internet Appendix) has three main steps. (1) *Dominance and iterated deletion.* The upper- and lower-dominance regions (Assumption 3) seed iterated deletion of dominated strategies; because the game has strategic complementarities (Lemma 1, $\mathcal{U}$ increasing in n) and state-monotone payoffs, the surviving strategies are sandwiched between a largest and a smallest threshold strategy (Morris and Shin, 2003). (2) *Indifference pins any threshold.* At any threshold equilibrium the pivotal type is indifferent; by Lemma 2 her expected stay payoff is $\Psi(\hat{x})$, so $\hat{x}$ solves (11). (3) *Uniqueness.* Ψ is strictly increasing ($\Psi'(\hat{x}) = \mathbb{E}_z[\partial_\theta q] = 1 > 0$ under (2), and $\Psi' > 0$ for any q with $\partial_\theta q > 0$), so (11) has a unique root; the largest and smallest thresholds coincide, eliminating all multiplicity, including the interior equilibrium of Theorem 1. The closed form follows from $\mathbb{E}_z[\theta^* - \varepsilon z] = \theta^*$ and $\int_0^1 h(1-u)\,du = \bar{H}$.

Equation (12) is the workhorse. The marginal participant clears the hurdle iff *average* quality across participation rates does, and average quality is $\theta^* + \Delta\bar{H}$. The run probability is now a determinate object: for any (proper) distribution F of θ,

$$\mathbb{P}(\text{run}) \;=\; F(\theta^*) \;=\; F(p + R - c - \Delta\bar{H}). \tag{13}$$

Corollary 1 (Uniqueness under a proper prior)**.** *Replace Assumption 4's diffuse improper prior with any proper prior on* θ *admitting a continuous, strictly positive density* f_0 *over the coordination region* $[\underline{\theta}, \overline{\theta}]$*, and retain all other assumptions. Provided private signals are sufficiently precise relative to the prior, equivalently,* ε *is small enough relative to the dispersion of* f_0*, the standard global-games precision condition (cf. Proposition 2), the global game has a unique Bayes–Nash equilibrium, in threshold strategies with cutoff* $\hat{x}^* = \hat{x}^*(\varepsilon, f_0)$ *given implicitly by the indifference condition*

$$\Psi^{f_0}(\hat{x}^*) \;\equiv\; \mathbb{E}_{z\sim G^{f_0}_{\hat{x}^*}}\big[q(\hat{x}^* - \varepsilon z,\, 1 - G^{f_0}_{\hat{x}^*}(z))\big] \;=\; \kappa, \tag{14}$$

where $G^{f_0}_{\hat{x}}$ *is the marginal participant's posterior CDF over* $z = (\hat{x} - \theta)/\varepsilon$ *given prior* f_0*. As* $\varepsilon \to 0$*,* $\hat{x}^*(\varepsilon, f_0) \to \theta^* = \kappa - \Delta\bar{H}$ *uniformly in* f_0*. The comparative statics of Corollary 2 hold for* $\hat{x}^*(\varepsilon, f_0)$ *at every fixed* f_0*.*

The corollary separates two contents of Theorem 2. The *uniqueness* claim does not depend on the diffuse prior specifically: under any proper prior it survives as the standard global-games

consequence of dominance regions and strategic complementarities, *provided private signals are precise enough relative to the prior* (the precision condition of Proposition 2); an informative prior acts like a public signal, so sufficiently imprecise private information can restore multiplicity exactly as in Proposition 2. Lemma 2's exactness is needed only for the closed form. The *closed form* $\hat{x}^* = \kappa - \Delta\bar{H}$ does depend on the diffuse prior together with the additive separability (2): the Laplacian uniform belief on the participation rate at the marginal signal collapses the indifference condition to an unweighted average, giving the $\bar{H}$ aggregation. Under a proper prior the indifference still has a unique root, but the integrand in Ψ^{f_0} uses a posterior density derived from f_0 rather than the Lebesgue weight du; the threshold is the same in the small-noise limit and differs by $O(\varepsilon)$ terms at non-trivial noise. Empirically, the comparative-statics predictions of this paper rest on $\hat{x}^*(\varepsilon, f_0)$'s monotone dependence on $(p, R, c, \Delta, \bar{H})$, which holds at every f_0; only quantitative predictions about threshold *levels* require explicit prior assumptions.

6.4 Anatomy of a run: fundamental versus panic

The threshold lies strictly inside the coordination region, since $\bar{H} \in (0, 1)$:

$$\underline{\theta} = \kappa - \Delta \;<\; \theta^* = \kappa - \Delta\bar{H} \;<\; \kappa = \overline{\theta}.$$

Definition 3 (Fundamental and panic runs)**.** A run at fundamental $\theta < \theta^*$ is a *fundamental run* if $\theta < \underline{\theta}$ (the platform would fail even under full participation) and a *panic run* if $\theta \in [\underline{\theta}, \theta^*)$ (fundamentals would support the platform under full participation, yet it fails on a coordination failure). The width of the panic region is

$$\theta^* - \underline{\theta} \;=\; \Delta(1 - \bar{H}) \;>\; 0. \tag{15}$$

Equation (15) is the model's fragility statistic: the range of *viable* platforms that nonetheless die of a run, increasing one-for-one in the complementarity strength Δ.

Corollary 2 (Comparative statics of the threshold)**.** *From* (12)*,* $\theta^* = p + R - c - \Delta\bar{H}$*, so*

$$\frac{\partial\theta^*}{\partial p} = 1 > 0, \quad \frac{\partial\theta^*}{\partial R} = 1 > 0, \quad \frac{\partial\theta^*}{\partial c} = -1 < 0, \quad \frac{\partial\theta^*}{\partial\Delta} = -\bar{H} < 0,$$

the run probability $F(\theta^*)$ *moves with* θ^**, and the panic width* $\Delta(1 - \bar{H})$ *is strictly increasing in* Δ*.*

6.5 Triggers: the belief, not the shock

Definition 4 (Trigger taxonomy)**.** A *trigger* is an observed event that raises the run probability. Triggers act through exactly three channels: (a) *fundamental shocks* lower θ (model failure, reliability or hallucination problems, security breach, loss of key talent, loss of compute access, an adverse copyright or regulatory *ruling*); (b) *hurdle shocks* raise $\kappa = p + R - c$ and hence raise θ^* (a price increase, $p \uparrow$; a superior rival or open-source option, $R \uparrow$; or reduced lock-in, $c \downarrow$); (c) *coordinating*

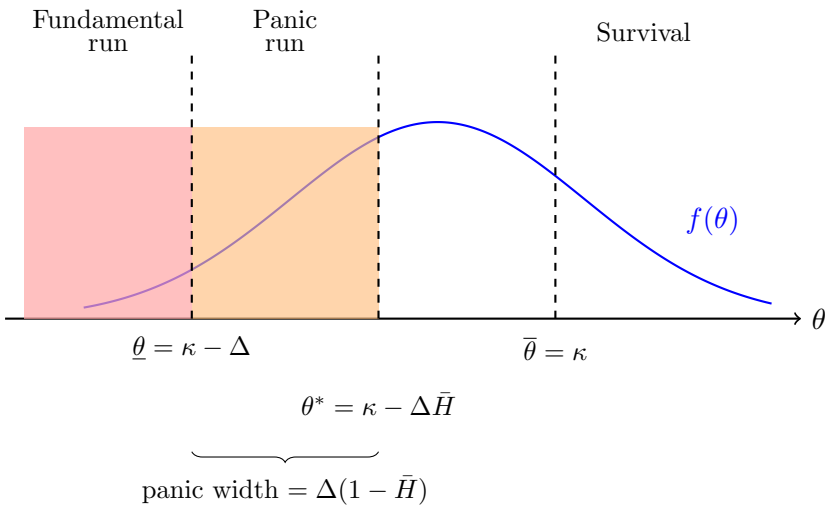


Figure 1: Equilibrium structure. The threshold $\theta^* = \kappa - \Delta\bar{H}$ separates survival from collapse. The shaded region $[\underline{\theta}, \theta^*)$ is the *panic region*: fundamentals that would support the platform under full participation but at which the global game selects collapse. Its width is $\Delta(1 - \bar{H})$, increasing one-for-one in the complementarity Δ. Run probability is $F(\theta^*)$; the panic component is $F(\theta^*) - F(\underline{\theta})$, the panic-region mass.

signals carry little fundamental content but are commonly observed (a regulatory *investigation*, a privacy *scandal*, a rumor of a rival's dominance, fear of product *discontinuation*).

Corollary 3 (Required trigger magnitude vanishes near the threshold)**.** *From a survival state (θ, κ) with $D \equiv \theta - \theta^* > 0$, a fundamental shock of size $\Delta\theta$ triggers a run iff $\Delta\theta > D$, and a hurdle shock of size $\Delta\kappa$ triggers a run iff $\Delta\kappa > D$. As $\theta \downarrow \theta^*$ (i.e. $D \downarrow 0$), arbitrarily small shocks suffice.*

This formalizes the claim that the run is not caused by the shock itself: conditional on the platform being near θ^*, the magnitude and even the fundamental content of a trigger are nearly irrelevant; what matters is that the event is commonly observed and coordinates beliefs. Two precise statements support the role of commonality.

(i) Sunspots in the coordination region. Under common knowledge (Theorem 1), for any $\theta \in (\underline{\theta}, \overline{\theta})$ a payoff-irrelevant but publicly observed event can select the run equilibrium with no change in fundamentals.

(ii) The publicity multiplier. The global-games refinement removes sunspots but assigns commonly observed signals weight out of proportion to their fundamental content. Augment (9) with a public signal $y = \theta + \nu$, $\nu \sim N(0, \tau_y^{-1})$, alongside private signals of precision $\tau_x = \varepsilon^{-2}$.

Proposition 2 (Public information and the survival of uniqueness)**.** *With the additional public signal y, the threshold equilibrium is unique provided private precision is high enough relative to public precision,*

$$\frac{\tau_x}{\tau_y^2} \;\geq\; \frac{1}{2\pi}\left(\Delta h'_{\max}\right)^2, \qquad h'_{\max} \equiv \sup_n h'(n), \tag{16}$$

the platform analogue of the Morris and Shin (2003) condition. When (16) holds, the equilibrium cutoff $\theta^*(y)$ is strictly decreasing in y, and a marginal change in the public signal moves the cutoff by more than its purely fundamental informativeness would warrant, because public information shifts beliefs about others in a correlated way (Morris and Shin, 2002). If private precision is sufficiently low relative to public precision (so that (16) is violated with room to spare), multiplicity reappears: very public, salient events can themselves restore self-fulfilling runs.

It is the commonality of the belief, not the fundamental content of the shock, that triggers the run; sufficiently public signals can even overturn uniqueness.

7 Endogenous Integration and Run Severity

The comparative statics use a notion of run *severity*, the loss a participant suffers when the platform collapses. We derive it from an integration choice rather than assume it, which also endogenizes the link between lock-in and severity that drives the stability illusion.

Add a date $t = -1$ at which each participant chooses an *integration depth* $b \geq 0$: how deeply to build workflows, retrain staff, and re-architect around the platform. Integration has three consequences. First, it raises the value of staying, modeled as reducing the effective price by $\beta_I b$ (deeper integration extracts more value from the platform), so $u_S = q(\theta, n) - p + \beta_I b$ with $\beta_I > 0$. Second, it raises the switching cost, $c = c(b)$ with $c' > 0$, $c'' \geq 0$, $c(0) = c_0 \geq 0$ (a deeply integrated participant forfeits more by migrating). Third, the integration capital is platform-specific and sunk: in a run, a participant loses it, incurring severity

$$L(b) \;=\; \ell_0 + \ell\, b, \qquad \ell > 0,\ \ell_0 \geq 0. \tag{17}$$

Integration itself costs $\chi(b)$, $\chi' > 0$, $\chi'' > 0$. At $t = -1$ participants hold a common prior belief $\pi_R \in [0,1]$ about the run probability (rational expectations require $\pi_R = F(\theta^*(c(b^*)))$ in equilibrium). The integration choice solves

$$\max_{b \geq 0}\ (1 - \pi_R)\, \beta_I b \;-\; \pi_R\, L(b) \;-\; \chi(b), \tag{18}$$

trading the participation-conditional benefit of integration against its expected sunk loss and cost.

Lemma 3 (Integration depth and the lock-in–severity link)**.** *Problem (18) has a unique solution $b^*(\pi_R)$, weakly decreasing in the perceived run probability π_R: $b^{*\prime}(\pi_R) \leq 0$, strictly where interior. Hence both the switching cost $c(b^*)$ and the severity $L(b^*)$ are weakly decreasing in π_R and co-move: platforms perceived as safe induce deep integration, high lock-in, and high severity.*

Define the equilibrium severity $\mathsf{L} \equiv L(b^*)$ and equilibrium lock-in $\mathsf{c} \equiv c(b^*)$. Lemma 3 delivers, as a derived object, the co-movement of lock-in and severity that the prose of the stability illusion needs: they rise together because both are increasing transforms of the same integration depth.

Treating c as the policy-relevant lock-in and $\mathsf{L}(\mathsf{c})$ as the induced severity (eliminating b), we have $\mathsf{L}'(\mathsf{c}) = \ell/c'(b^*) > 0$.

8 Main Results

We can now read the economics off the threshold. What follows is not a catalog of separate pathologies but four faces of a single fact: because a platform's quality is produced by the very participation that can be withdrawn, its value is confidence-sensitive, and the threshold $\theta^* = \kappa - \Delta\bar{H}$ is the hinge on which that confidence turns. Each result takes a lever a financial economist already cares about, the strength of the network effect, customer lock-in, the openness of the technology, and (in Section 10) the case for a public backstop, and asks how it moves that hinge. In each case the answer runs against the lever's reputation. The network effects prized as moats widen the region of self-fulfilling collapse. The switching costs that look like stability convert frequent small exits into rare large ones. Openness trades catastrophe for churn. And the backstop that would arrest a panic erodes the incentive to avoid one. The objects each result invokes are developed where they are first needed and linked at each use (run severity in Section 7, the endogenous complementarity in Section 9, resilience and the downstream sector in Section 10, the pricing kernel in Section 11) and all proofs are in the Internet Appendix. We state each result formally for precision, then say in words what it means.

Start with the network effect itself. The first result holds the platform's full-participation quality fixed and varies only how *sensitive* that quality is to losing users.

Proposition 3 (Fragility rises with the retention–quality feedback)**.** *Fix the full-participation quality $q(\theta, 1) = \theta + \Delta$ at any level (so the platform's stand-alone-plus-network quality is held constant). Then a stronger retention-to-quality feedback (larger Δ, equivalently larger ϕ_d, ϕ_e, or ϕ_c (Remark 1)) strictly widens the panic region: $\partial[\Delta(1-\bar{H})]/\partial\Delta = 1 - \bar{H} > 0$. The probability of a panic run, $F(\theta^*) - F(\underline{\theta})$, then rises in Δ at low-to-moderate complementarity, formally whenever $f(\underline{\theta}) \geq \bar{H}\, f(\theta^*)$, which holds in particular for Δ small or for f non-increasing across the panic band; for large Δ this probability eventually falls (the global hump shape is Proposition 10).*

The relevant fragility statistic is not the *level* of quality but its *sensitivity* to participation, $q(\theta, 1) - q(\theta, 0) = \Delta$. A platform whose quality collapses sharply when users leave, because its edge comes from data scale, ecosystem depth, and revenue-financed compute, occupies a wide band of fundamentals over which a coordinated loss of confidence is self-validating. The complementarities prized as moats in good times are the transmission mechanism of collapse in bad times.

Remark 5 (The two faces of Δ). Higher Δ *lowers* θ^* (Corollary 2), so a more network-intensive platform clears the hurdle more easily on average and suffers *fewer* fundamental failures; yet it *widens* the panic band. Fragility is a statement about panics, not about average quality.

Turn next to customer lock-in, the lever a financial economist would most naturally read as a source of stability, since high switching costs mean customers do not leave. The model agrees that they leave less often, and shows that this is the problem.

Proposition 4 (Switching costs: the stability illusion)**.** *Using the endogenous lock-in* c *and severity* $\mathsf{L}(\mathsf{c})$ *of Section 7: higher lock-in lowers the threshold,* $\partial\theta^*/\partial\mathsf{c} = -1 < 0$*, hence reduces run* frequency $F(\theta^*)$*, while raising severity,* $\mathsf{L}'(\mathsf{c}) > 0$*. Expected disruption is*

$$\mathcal{E}(\mathsf{c}) \;=\; F\big(\theta^*(\mathsf{c})\big)\,\mathsf{L}(\mathsf{c}), \qquad \frac{d\log\mathcal{E}}{d\log\mathsf{c}} \;=\; \epsilon_{\mathsf{L},\mathsf{c}} - \big|\epsilon_{F,\mathsf{c}}\big|,$$

with severity elasticity $\epsilon_{\mathsf{L},\mathsf{c}} > 0$ *and frequency elasticity* $\epsilon_{F,\mathsf{c}} < 0$*. Under the standard tail regularity condition* $\mathsf{c}f(\theta^*(\mathsf{c}))/F(\theta^*(\mathsf{c})) \to 0$ *as* $\mathsf{c} \to \infty$*, there exists* $\bar{c}$ *such that for* $\mathsf{c} > \bar{c}$ *expected disruption is* increasing *in lock-in. Thus high-lock-in platforms exhibit low routine churn (rare runs) yet larger losses conditional on a run.*

Lock-in trades a stream of small, frequent exits for a rare, catastrophic run. Observers infer stability from the absence of churn, but the absence of churn is what deepens integration (Lemma 3) and raises the social cost of the eventual coordination failure. High measured stability can be a symptom of latent fragility.

The third lever is the openness of the underlying technology (open weights, open standards, interoperability) which policy debate treats as either unambiguously stabilizing or unambiguously corrosive. The model says it is both, through channels that pull in opposite directions.

Proposition 5 (Open source: catastrophe, frequency, and rents)**.** *Open-source and interoperable alternatives act through four channels with opposing signs (Corollary 2, Section 7). Two of them raise run* frequency*: a better outside option (*$R \uparrow$*) and reduced lock-in (*$\mathsf{c} \downarrow$*) both lift the threshold* θ^**, so the platform is abandoned at a wider range of fundamentals. A third channel works the other way and cuts the* severity *of what remains: interoperability lets a shared, open commons absorb a coordinated exodus, so a departure destroys less platform-specific value; this lowers the effective complementarity* Δ*, which both shrinks the panic region* $\Delta(1-\bar{H})$ *and reduces per-participant loss (*$\partial\mathsf{L}/\partial\Delta > 0$*, via* ℓ *in* (17)*), muting the catastrophic, self-fulfilling dimension of a run. The fourth channel is a cost borne elsewhere: proprietary rents* Π *fall, and because frontier investment is financed out of rents (*$I^{*\prime}(\Pi) > 0$*, Section 10.3), openness lowers the private return to pushing the frontier (*$\partial I^*/\partial R < 0$*). Netting the first three, expected disruption* $\mathcal{E} = F(\theta^*)\mathsf{L}$ *falls if and only if the severity-and-panic reduction outweighs the frequency increase,* $|\epsilon_{\mathsf{L},\Delta}\,\epsilon_{\Delta}| > \epsilon_{F,R} + \epsilon_{F,\mathsf{c}}$*.*

This reconciles the two readings of openness. Open source *does* lower the probability of catastrophic runs on the panic-and-severity margin (channel iii): an interoperable commons mutes the self-fulfilling collapse and lets leavers land softly. The offsetting force is that easier migration and a better outside option make ordinary exit easier to coordinate, raising the *frequency* of milder runs (channels i–ii), at the cost of the rents that finance frontier capability (channel iv). The deep tension is between resilience (favored by openness) and frontier innovation (favored by proprietary rents).

8.1 The shape of the adoption curve: a result the assumptions do not contain

The results so far might be read as mechanical consequences of the setup: once the panic band is *defined* as $\Delta(1-\bar{H})$, "fragility rises with Δ" is almost an accounting statement. The following result is different in kind. It shows that the equilibrium analysis makes run risk depend on an object that is *invisible* in the assumptions and indeterminate under common knowledge: the curvature of the network-benefit function, i.e. *where along the adoption curve* the complementarity is delivered.

Hold the total complementarity Δ and the fundamental distribution F fixed, and vary only the shape of h. The relevant statistic is $\bar{H} = \int_0^1 h(n)\,dn$, the area under the adoption-value curve, which the average-quality condition of Theorem 2 (not any assumption) singles out.

Proposition 6 (Timing of network benefits and fragility)**.** *Fix Δ and F. Run probability $F(\theta^*) = F(\kappa - \Delta\bar{H})$ is strictly decreasing in $\bar{H}$. Hence, holding the* magnitude *of the network effect fixed:*

(i) platforms whose benefits accrue early *in adoption (h concave, $\bar{H}$ large, value is delivered even to small cohorts) are* more *run-proof;*

(ii) platforms that require critical mass *(h convex, $\bar{H}$ small, value materializes only near full adoption) are* more *run-prone.*

In the parametric family $h(n) = n^\gamma$ ($\gamma > 0$), $\bar{H} = 1/(1+\gamma)$, so $\theta^* = \kappa - \Delta/(1+\gamma)$ is strictly increasing in the convexity γ: $\partial\theta^*/\partial\gamma = \Delta/(1+\gamma)^2 > 0$, and run probability rises with γ. As $\gamma \to 0$ (h a step up, value from the first adopter) $\theta^* \to \underline{\theta}$; as $\gamma \to \infty$ (h flat until full adoption) $\theta^* \to \overline{\theta}$.

Two platforms with *identical* total network effect Δ, *identical* fundamentals, and *identical* prices can have very different run risk purely because of the curvature of h. A coding assistant useful to a solo developer on day one (concave h, high $\bar{H}$) is structurally safer than a collaboration or social feature whose value appears only once most of an organization has joined (convex h, low $\bar{H}$), even if the two ultimately deliver the same full-participation quality. The result does not follow from the assumptions: under common knowledge (Theorem 1) the shape of h does not pin down any run probability at all, the entire coordination band is consistent with both outcomes. The dependence on $\bar{H}$ is *created* by the equilibrium selection: the marginal participant evaluates *average* quality across participation rates, and the average is a functional of the whole curve, not of its endpoint Δ. The shape of the adoption curve is thus a primitive source of fragility that the magnitude Δ conceals; Section 4 derives that shape from the data-scaling exponent and the exit-selection intensity, so $\bar{H}$ is itself determined by the underlying data primitives.

This last point has a sharp corollary once the curvature is read as the technology's data-scaling exponent. AI labs compete to make models extract more from each unit of data, to push the scaling exponent up, weakening the diminishing returns that bound a learning technology. The model says this very improvement, holding peak quality fixed, makes the platform *more* fragile.

Proposition 7 (Capability and fragility: diminishing returns to data are stabilizing)**.** *Take the homogeneous learning technology $\Xi(\mathcal{I}) = A\mathcal{I}^\alpha$, so $h(n) = n^\alpha$ and $\bar{H} = 1/(1+\alpha)$, and hold full-*

participation quality $\Delta = \Xi(\bar{\iota})$ *fixed (adjusting the level* A*). Then the run threshold*

$$\theta^* = \kappa - \frac{\Delta}{1+\alpha}, \qquad \frac{\partial \theta^*}{\partial \alpha} = \frac{\Delta}{(1+\alpha)^2} > 0,$$

is strictly increasing in the data-scaling exponent α*: a technology with* weaker *diminishing returns to data is strictly more run-prone, even though peak quality is unchanged. Under selection the statement holds verbatim with* α *replaced by* $\gamma_{\text{eff}} = \alpha(1+\varsigma)$*, so fragility rises in both the scaling exponent and the exit-selection intensity.*

The result is the opposite of the usual reading. Diminishing returns to data are ordinarily treated as a limit on a platform's value; here they are privately and socially *stabilizing*, because a technology that keeps returns high as data accumulates back-loads quality to high participation, lowers the average quality the marginal participant internalizes, and widens the coordination region. The force is an interaction between a data-and-AI primitive (the scaling exponent) and the coordination threshold; it has no counterpart in a deposit run, where there is no production technology whose curvature can be improved. A frontier lab that succeeds in making intelligence cheaper per token thereby concentrates more of its delivered value at scale, and buys fragility with capability.

8.2 What is mechanical and what is earned

Because a recurring worry about coordination models is that the conclusions are baked into the definitions, we state plainly which results are organizing identities and which are substantive, the latter being claims that require the equilibrium analysis and could have come out otherwise.

Organizing identities (true by construction, useful for accounting but not discoveries): the panic width equals $\Delta(1-\bar{H})$ (Eq. 15); the threshold comparative statics in (p, R, c) (Corollary 2); and the decomposition of expected disruption into frequency times severity (Proposition 4). We label these as such and do not rest the paper's contribution on them.

Substantive results (require the equilibrium machinery; the sign was not assumed and could have been the opposite):

(a) *Exactness of the threshold.* That the cutoff is $\theta^* = \kappa - \Delta\bar{H}$ *for every noise level*, not merely in the small-noise limit, is special to the diffuse-prior, mean-zero-noise structure; with a proper prior or non-zero noise mean the threshold is noise-dependent (Theorem 2; Corollary 1 and Remark 4).

(b) *Curvature dependence, derived from data primitives.* Run risk depends on $\bar{H}$, a functional of the *shape* of h that is indeterminate under common knowledge and manufactured by the Laplacian average-quality condition (Proposition 6); the shape is in turn derived from the data-scaling exponent and the exit-selection intensity (Proposition 1), not assumed.

(c) *The data-composition mechanism.* The complementarity is the output of a learning technology on heterogeneous user data; fragility is generated by adverse selection in exit (informative users

leave first), a force specific to AI platforms whose sign turns on a primitive covariance and can in principle stabilize as well as destabilize (Proposition 1).

(d) *Non-monotone fragility.* Panic-run *probability* is hump-shaped in Δ even though the panic *width* is monotone, a sign reversal, not a restatement (Proposition 10).

(e) *Severity endogenized, not assumed.* $L'(\mathsf{c}) > 0$ is derived from the integration problem (Lemma 3); the model also delivers the conditions under which expected disruption rises with lock-in rather than assuming it (Proposition 4).

(f) *Priced coordination risk.* The panic component of run risk (pure coordination failure on fundamentals that would otherwise support the platform) is nonetheless priced and is undiversifiable within an ecosystem (Proposition 13 below), contrary to the intuition that belief-driven risk is idiosyncratic.

The paper's claims rest on the second group.

9 Endogenous Complementarity and the Limits to Centralization

Until now Δ has been a primitive: Remark 1 decomposed it as $\phi_d + \phi_e + \phi_c$ but treated the data, ecosystem, and compute channels as additive and exogenous. The channels are not exogenous: they feed one another, and that feedback is the structural engine of platform dominance. This section endogenizes Δ as the equilibrium of an interacting channel system, derives the sign structure of the cross-channel *reinforcement matrix* from primitives (Section 9.1), and then shows that once channels saturate, the centralization tendency *stalls* and platform fragility becomes *non-monotone* in complementarity strength, overturning the uniformly pessimistic reading on the frequency margin (Sections 9.2–9.3).

9.1 The reinforcement system

Evaluate the channels at the high-participation ($n = 1$) regime. Let $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3_+$ collect the data stock, ecosystem breadth, and compute/capability. Each channel is produced from participation and from the *other* channels:

$$x_i \;=\; f_i\big(n;\, \mathbf{x}_{-i}\big), \qquad \frac{\partial f_i}{\partial n} \geq 0, \quad \frac{\partial f_i}{\partial x_j} \geq 0, \quad f_i(0;\mathbf{0}) = 0. \tag{19}$$

Data improve with users and with compute (curation, training); the ecosystem grows with users and with model quality (itself driven by data and compute); compute is financed by revenue, which scales with users and with ecosystem monetization. At $n = 0$ all channels vanish, consistent with $q(\theta, 0) = \theta$; at $n = 1$ they solve the simultaneous fixed point $\mathbf{x}^* = f(1;\mathbf{x}^*)$. The network contribution to quality is $\Delta = \mathcal{Q}(\mathbf{x}^*)$, with $\mathcal{Q}$ increasing and $\mathcal{Q}(\mathbf{0}) = 0$, so $q(\theta, 1) = \theta + \Delta$ as before.

Definition 5 (Reinforcement matrix)**.** The *reinforcement matrix* is the Jacobian of the channel system (19) at the fixed point, $\mathbf{A}(\mathbf{x}) \equiv [\partial f_i/\partial x_j]_{ij} \geq \mathbf{0}$ (entrywise, by (19)). Its spectral radius $\rho(\mathbf{A})$ is the *loop gain* of the feedback. We use the linear benchmark $f_i(1;\mathbf{x}_{-i}) = s_i + \sum_j a_{ij}x_j$ with $\mathbf{A} = [a_{ij}] \geq \mathbf{0}$ and direct (participation-driven) effects $\mathbf{s} \geq \mathbf{0}$, for which $\Delta = \mathbf{w}'\mathbf{x}^*$ with quality weights $\mathbf{w} > \mathbf{0}$.

Proposition 8 (Endogenous complementarity and the reinforcement multiplier)**.** *Let $\mathbf{A} \geq \mathbf{0}$ with loop gain $\rho(\mathbf{A}) < 1$. Then the channel system has a unique nonnegative fixed point*

$$\mathbf{x}^* = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{s} = \Big(\textstyle\sum_{k\geq 0} \mathbf{A}^k\Big)\mathbf{s} \geq \mathbf{0},$$

and the equilibrium complementarity $\Delta = \mathbf{w}'(\mathbf{I} - \mathbf{A})^{-1}\mathbf{s}$ satisfies:

(i) it is positive and finite;

(ii) it is strictly increasing in every direct effect s_i and every reinforcement coefficient a_{ij}, signs inherited from the primitive monotonicities $\partial f_i/\partial x_j \geq 0$ in (19);

(iii) it is increasing in the loop gain, with $\Delta \to \infty$ as $\rho(\mathbf{A}) \to 1^-$.

Consequently the panic width $\Delta(1 - \bar{H})$ and threshold $\theta^* = \kappa - \Delta\bar{H}$ inherit these signs: stronger reinforcement endogenously widens the fragile region and (Corollary 2) lowers the run threshold.

The reduced form $\Delta = \phi_d + \phi_e + \phi_c$ of Remark 1 is the no-interaction case $\mathbf{A} = \mathbf{0}$, where $\Delta = \mathbf{w}'\mathbf{s}$ is just the sum of direct effects. Interactions ($\mathbf{A} > \mathbf{0}$) amplify it through the multiplier $(\mathbf{I}-\mathbf{A})^{-1} = \mathbf{I}+\mathbf{A}+\mathbf{A}^2+\cdots$, whose terms are the feedback loops: $\mathbf{A}$ is the one-step reinforcement (data → model → developers), $\mathbf{A}^2$ the two-step loops, and so on. This multiplier *is* the structural centralization force, and the off-diagonal entries of $\mathbf{A}$, derived from the primitive cross-channel elasticities, are the arrows of the data–compute–ecosystem flywheel. Winner-take-all is the regime $\rho(\mathbf{A}) \to 1$.

Remark 6 (Loop gain as a stability condition)*.* $\rho(\mathbf{A}) < 1$ is simultaneously the existence condition for an interior channel equilibrium and the local stability condition of the tâtonnement $\mathbf{x}_{t+1} = f(1;\mathbf{x}_t)$. When $\rho(\mathbf{A}) \geq 1$ the feedback is explosive: no interior fixed point exists and the model predicts unbounded scale, the knife-edge dominance-or-collapse limit of Remark 9. The bimodal market structure is thus the high-loop-gain regime of a single primitive object.

9.2 Saturation: why centralization stalls

The linear system predicts runaway centralization as $\rho(\mathbf{A}) \to 1$. But the channels saturate: the millionth near-duplicate conversation teaches a model little, ecosystems congest, and compute meets scaling limits. We model this by concavity and bounded returns.

Assumption 5 (Saturation)**.** Each $f_i(1;\cdot)$ maps the compact order interval $[\mathbf{0}, \bar{\mathbf{x}}]$ into itself, is C^2, increasing and strictly concave in $\mathbf{x}_{-i}$, and has globally bounded loop gain on that interval: there is

$\bar{a} < 1$ with $\sum_j \partial f_i/\partial x_j \le \bar{a}$ for all i and all $\mathbf{x} \in [\mathbf{0}, \bar{\mathbf{x}}]$ (decreasing returns to the aggregate channel bundle).

Proposition 9 (Centralization stalls)**.** *Under Assumption 5:*

(i) the channel fixed point $\mathbf{x}^*$ exists, is unique, and is globally stable under the tâtonnement;

(ii) the loop gain at the fixed point satisfies $\rho(\mathbf{A}(\mathbf{x}^*)) \le \bar{a} < 1$ for all primitive levels, so the winner-take-all blow-up of Proposition 8(iii) never occurs;

(iii) the equilibrium complementarity is bounded, $\Delta \le \Delta_{\max} = \mathbf{w}'\bar{\mathbf{x}}$, and the marginal effect of scale vanishes, $\partial\Delta/\partial s_i \to 0$ as $s_i \to \infty$.

Centralization therefore stalls at finite scale: beyond a point, additional participation, data, or compute raises complementarity negligibly.

The proof (the Internet Appendix) uses the nonnegative-matrix bound $\rho(\mathbf{A}) \le \max_i \sum_j a_{ij} \le \bar{a} < 1$ (Perron–Frobenius), which makes $f(1;\cdot)$ a contraction beyond scale $\bar{X}$; Banach's theorem gives existence, uniqueness, and global stability, and boundedness gives $\Delta_{\max}$. The unbounded flywheel is an artifact of linearity; with diminishing returns, the loop gain is pinned below one and the multiplier stays finite.

9.3 The fragility reversal

The next consequence concerns fragility itself. Proposition 3 showed the panic *width* $\Delta(1-\bar{H})$ rises monotonically in Δ. But the economically relevant object is the *probability* that the platform actually lands in the fragile band,

$$P_{\text{panic}}(\Delta) \;\equiv\; F(\theta^*) - F(\underline{\theta}) \;=\; F(\kappa - \Delta\bar{H}) - F(\kappa - \Delta). \tag{20}$$

Proposition 10 (Non-monotone fragility; reversal of the centralization pessimism)**.** *For any continuous fundamental distribution F with $F(-\infty) = 0$ and positive density on the relevant intervals:*

(i) $P_{\text{panic}}(0) = 0$ and $\lim_{\Delta\to\infty} P_{\text{panic}}(\Delta) = 0$; hence P_{panic} is non-monotone and attains an interior maximum at some $\widehat{\Delta} > 0$;

(ii) $P'_{\text{panic}}(\Delta) = f(\underline{\theta}) - \bar{H}\, f(\theta^*)$, which is positive for small Δ and, if F is eventually unimodal (density decreasing in the left tail), negative for all $\Delta > \widehat{\Delta}$;

(iii) therefore, among sufficiently network-intensive platforms ($\Delta > \widehat{\Delta}$), greater complementarity lowers the probability of a panic run*, reversing Proposition 3 on the frequency margin.*

Two forces collide. Higher Δ widens the panic band (Proposition 3), but it also lowers the threshold $\theta^* = \kappa - \Delta\bar{H}$, sliding the entire coordination region $[\underline{\theta}, \theta^*]$ leftward into the low-density left tail of fundamentals. At low Δ the widening dominates and fragility rises; at high Δ the

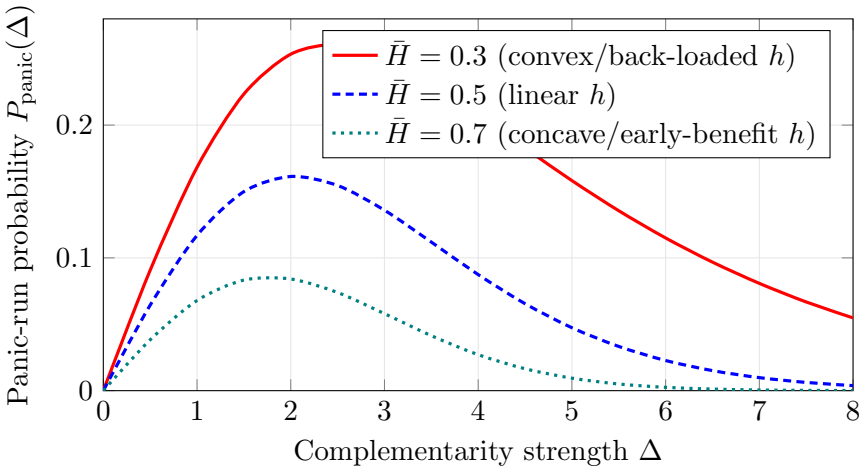


Figure 2: Hump-shaped fragility (Proposition 10). Panic-run probability $P_{\text{panic}}(\Delta) = F(\kappa - \Delta\bar{H}) - F(\kappa - \Delta)$ for $\theta \sim N(\mu_\theta = \kappa, \sigma_\theta^2)$, $\sigma_\theta = 1.5$. As Δ rises, two forces compete: the panic-width $\Delta(1-\bar{H})$ grows linearly (Proposition 3), but the band slides leftward into the low-density tail. The reversal point $\widehat{\Delta}$ is the value of Δ at which density-weighted width starts to fall; for the three curves, $\widehat{\Delta} \approx \{2.4, 2.0, 1.8\}$ as $\bar{H}$ rises. More back-loaded, critical-mass adoption curves (lower $\bar{H}$) shift the peak rightward and raise its height.

leftward migration dominates, and the platform so rarely draws a fundamental low enough to enter the fragile band that panic runs become rare. The largest, most network-intensive platforms sit on the *declining* branch: their enormous complementarity makes them clear the hurdle so easily that they seldom enter the fragile region at all, even though, conditional on entering it, their panic band is the widest. This squares the run mechanism with the observed durability of the largest platforms without abandoning it. The non-monotonicity is qualitatively akin to the integration–diversification result of Elliott et al. (2014) for financial-network cascades, where mid-range exposure produces the largest contagion and both extremes are safer; here the mechanism is different (width-versus-tail of the fundamental distribution, not cross-holdings of claims), but the cross-sectional shape is the same.

Remark 7 (Width versus probability: reconciling the two readings). Proposition 3 (width monotone in Δ) and Proposition 10 (probability hump-shaped in Δ) are both correct and not in tension. Width is Lebesgue measure on the θ-line; probability weights that width by the density f. *Conditional* severity rises in Δ; *unconditional* frequency is hump-shaped. The pessimistic reading (network effects breed fragility) holds at low-to-moderate Δ and always on the severity margin, but reverses on the frequency margin once the platform is network-intensive enough.

Calibrated illustration. Figure 2 plots $P_{\text{panic}}(\Delta)$ for $\theta \sim N(\kappa, 1.5^2)$ at three values of $\bar{H} \in \{0.3, 0.5, 0.7\}$. Three readings stand out. First, fragility's maximum is interior at $\widehat{\Delta}$ of order one to three depending on curvature, not at the upper extreme; the cross-sectional ranking of platforms by their data-and-compute intensity therefore does *not* translate into a ranking by run probability. Second, the peak height is itself a function of curvature: more back-loaded or critical-mass adoption

curves (lower $\bar{H}$) imply a wider panic band for a given Δ, raising peak fragility; early-benefit curves (higher $\bar{H}$), where value is delivered even at low participation, have lower peak fragility. Third, the right tail of all three curves is thin: very-network-intensive platforms ($\Delta > 5$) have panic probabilities below one percent in this calibration, consistent with the empirical durability of the largest platforms despite their large panic widths.

Corollary 4 (Hump shape under fat-tailed and multimodal F)**.** *Proposition 10's non-monotonicity admits two refinements that do not require eventual unimodality of F.*

(i) Polynomial decay under fat tails. *If the left tail of F is regularly varying with index $\alpha_F > 0$, formally, $f(x) \sim c|x|^{-1-\alpha_F}$ as $x \to -\infty$ for some $c > 0$, equivalently $F(x) \sim (c/\alpha_F)|x|^{-\alpha_F}$, then*

$$P_{\text{panic}}(\Delta) \;\sim\; \frac{c\,(\bar{H}^{-\alpha_F} - 1)}{\alpha_F} \cdot \Delta^{-\alpha_F} \qquad \text{as } \Delta \to \infty.$$

The coefficient is strictly positive since $\bar{H} \in (0,1) \Rightarrow \bar{H}^{-\alpha_F} > 1$; $P_{\text{panic}}(\Delta)$ still approaches zero, but at polynomial rate rather than the Gaussian's exponential rate. The reversal point $\widehat{\Delta}$ exists, and the right-tail decay of fragility is slow: very-large-Δ platforms drawn from a fat-tailed fundamental distribution carry materially higher panic-run probability than the Gaussian benchmark suggests, even as the qualitative non-monotonicity is preserved.

(ii) Multiple maxima under multimodal F. *If f has several economically relevant modes, the stationarity condition*

$$f\big(\kappa - \Delta\big) \;=\; \bar{H} \cdot f\big(\kappa - \Delta\bar{H}\big)$$

can have several solutions, so $P_{\text{panic}}(\Delta)$ need not be unimodal. Under the additional regularity that the modes are separated by monotone density segments, each local maximum is associated with the panic band moving across one of the modes of f; the dominant maximum is associated with the mode carrying the largest density mass over the band.

In both cases the headline implication of Proposition 10, that mid-intensity platforms carry higher panic-run probability than either very-small-Δ or sufficiently large-Δ platforms, survives. What is specific to eventual unimodality is the uniqueness of the maximum.

Corollary 5 (The cross-section of crash risk is hump-shaped)**.** *The equity disaster component's* probability *scales with $P_{\text{panic}}(\Delta)$ and its* severity *with the width $\Delta(1-\bar{H})$, so expected crash losses $\mathcal{C}(\Delta) = P_{\text{panic}}(\Delta) \cdot \Delta(1-\bar{H}) \cdot \bar{\ell}$ (for severity scale $\bar{\ell} > 0$) are hump-shaped in network intensity under the Gaussian benchmark and, more generally, under light-tailed unimodal F satisfying $\Delta P_{\text{panic}}(\Delta) \to 0$: mid-intensity platforms carry the largest run-risk premium and crash exposure, while both small-Δ and mega-Δ platforms carry less. The cross-sectional pattern is non-monotone, departing sharply from the uniformly pessimistic baseline.*

Remark 8 (A second reversal: channel congestion). If a channel congests at scale, low-quality or adversarial data degrading the model, moderation and safety load rising super-linearly, its own-effect turns negative ($a_{ii}(\mathbf{x}) < 0$ beyond a scale). Then $\partial\Delta/\partial s_i$ can change sign: strengthening that

channel's direct input *lowers* equilibrium complementarity once congestion dominates reinforcement. Expanding the data pipeline can reduce Δ, and with congestion strong enough the system supports several coexisting platforms (multi-homing) rather than a single winner, an endogenous force for *de*-centralization absent from the linear flywheel.

10 Resilience, the Downstream Sector, and Systemic Risk

We now specify the platform's own problem, the downstream sector that makes a platform *systemic*, and the frontier-investment technology used in Proposition 5. This delivers Proposition 11 from primitives.

10.1 The downstream sector

A unit continuum of downstream firms $j \in [0,1]$ each operate a workflow built on the platform. Firm j's output is $y_j = A_j\, Q$, where $Q \in \{Q_{\text{hi}}, Q_{\text{lo}}\}$ is realized platform value (Q_{hi} if the platform survives, $Q_{\text{lo}} < Q_{\text{hi}}$ in a run) and $A_j > 0$ is firm productivity. Firm j has *platform dependence* $\delta_j \geq 0$: the fraction of its operations running on this platform, so that a run reduces its output by $\delta_j\,(Q_{\text{hi}} - Q_{\text{lo}})\,A_j$. Aggregate downstream disruption from a run is

$$X \;\equiv\; \int_0^1 \delta_j\,(Q_{\text{hi}} - Q_{\text{lo}})\,A_j\,dj \;=\; \bar{\delta}\,(Q_{\text{hi}} - Q_{\text{lo}})\,\bar{A} \;>\; 0, \tag{21}$$

with $\bar{\delta}\bar{A} \equiv \int \delta_j A_j\,dj$. A platform is *systemic* when $\bar{\delta}$ (the economy's average reliance on it) is large. X is an externality: it falls on downstream firms, not on the platform, and the platform does not internalize it. The aggregation here is the simplest possible (linear shares δ_j on a binary platform outcome) but the structural feature it captures is the same one that drives the production-network propagation of Acemoglu et al. (2012): an upstream disturbance reaches the aggregate through a weighted sum of downstream dependences, and a sector's systemic weight is governed by its dependence shares. The difference is that the upstream disturbance here is not a sectoral productivity shock but a coordination-driven quality collapse, so a single platform with high $\bar{\delta}$ generates aggregate disruption from what is, fundamentally, a belief-driven event.

10.2 Resilience and the platform's problem

At $t = -1$ the platform chooses resilience effort $e \geq 0$ at convex cost $\psi(e)$ ($\psi' > 0, \psi'' > 0, \psi'(0) = 0$). Resilience improves fundamentals by a first-order-stochastic-dominance right-shift of the distribution of θ: the survival probability is

$$\Sigma(e) \;\equiv\; \mathbb{P}\big(\theta \geq \theta^* \mid e\big) \;=\; 1 - F\big(\theta^* - \varrho(e)\big), \qquad \varrho'(e) > 0,\ \varrho(0) = 0, \tag{22}$$

so $\Sigma'(e) = f(\theta^* - \varrho(e))\,\varrho'(e) > 0$ and Σ is increasing and (for f single-peaked and ρ concave) concave on the relevant range. Conditional on survival the platform earns rents $\Pi > 0$ (present

value of fees net of operating cost); in a run it earns $\Pi_{\text{run}} \in [0, \Pi)$. The platform solves

$$\max_{e \geq 0} \; V^{\text{priv}}(e) \;=\; \Sigma(e)\,\Pi + (1 - \Sigma(e))\,\Pi_{\text{run}} - \psi(e) \;=\; \Pi_{\text{run}} + \Sigma(e)\,(\Pi - \Pi_{\text{run}}) - \psi(e). \tag{23}$$

10.3 Frontier investment

Separately (and used in Proposition 5), the platform can raise stand-alone quality by frontier investment $I \geq 0$ at cost $\Gamma(I)$ ($\Gamma' > 0, \Gamma'' > 0$), shifting θ up by $\upsilon(I)$, $\upsilon' > 0$. Because rents accrue only on survival, the marginal return to I scales with rents: $I^* = \arg\max_I \Sigma(e)\,[\Pi + \partial\Pi/\partial\upsilon \cdot \upsilon(I)] - \Gamma(I)$ has $\partial I^*/\partial\Pi > 0$ by the standard supermodularity of $\Sigma\,\Pi$ in (I, Π). This is the channel through which competition from open source ($\Pi \downarrow$) lowers I^*.

10.4 The too-connected-to-fail wedge

Proposition 11 (Under-provision of resilience; moral hazard from rescue)**.** *Let e^{priv} solve (23) and let e^{soc} solve the planner's problem that adds downstream surplus,* $\max_e \Sigma(e)\,(\Pi - \Pi_{run} + X) - \psi(e) + const$*, with X from (21). Then:*

(i) Under-provision. *$e^{soc} > e^{priv}$ whenever $X > 0$; the gap is increasing in the platform's systemic footprint $\bar{\delta}$.*

(ii) Moral hazard. *If the platform anticipates a rescue paying $B \geq 0$ in a run (a public compute/data backstop, bailout, or forced acquisition), its private survival benefit shrinks to $\Pi - \Pi_{run} - B$, and $\partial e^{priv}/\partial B < 0$: a more generous backstop lowers private resilience.*

When many firms run pricing, marketing, service, coding, and analytics on one platform, its collapse propagates as an externality the platform does not bear, the too-big-to-fail logic of systemically important institutions transplanted to digital infrastructure. Private resilience is too low, and the standard remedy, a credible backstop, reintroduces moral hazard.

11 Asset Pricing: The Price of Platform-Run Risk

We now embed the platform in an asset-pricing environment, *derive* that runs strike in high-marginal-utility states, and price the resulting risk. The run event is the $\{\theta < \theta^*\}$ outcome of Theorem 2; the premium isolated here is decomposed into fundamental and panic parts in Section 11.3 and recast as a term structure and a CDS-style survival spread in Section IA.1.

11.1 The stochastic discount factor and the run state

Let ω index the aggregate state, with a strictly positive stochastic discount factor (SDF) $M(\omega)$, $\mathbb{E}[M] = 1/(1 + r_f)$, high in "bad" (high marginal-utility) states. The platform's hurdle and fundamental depend on the aggregate state: in downturns, enterprise budgets and financing tighten and

outside options shift, so the hurdle rises and the fundamental falls,

$$\kappa(\omega) = \kappa_0 + \kappa_1\,\mathsf{m}(\omega), \qquad \theta(\omega) = \theta_0 - \theta_1\,\mathsf{m}(\omega), \qquad \kappa_1, \theta_1 \geq 0,\ (\kappa_1, \theta_1) \neq (0,0), \tag{24}$$

where $\mathsf{m}(\omega)$ is increasing in marginal utility (comonotone with M). Realized platform value is $\tilde{Q} = Q_{\text{hi}}\mathbf{1}\{\theta(\omega) \geq \theta^*(\omega)\} + Q_{\text{lo}}\mathbf{1}\{\theta(\omega) < \theta^*(\omega)\}$, where $\theta^*(\omega) = \kappa(\omega) - \Delta\bar{H}$.

Lemma 4 (Runs are bad-state events)**.** *Under (24), the distance to run $\theta(\omega) - \theta^*(\omega) = \theta_0 - \kappa_0 + \Delta\bar{H} - (\theta_1 + \kappa_1)\mathsf{m}(\omega)$ is strictly decreasing in $\mathsf{m}(\omega)$. Hence runs occur in the high-m, high-marginal-utility states, and platform value $\tilde{Q}$ is low exactly when M is high:* $\text{Cov}(M, \tilde{Q}) < 0$.

The covariance sign is here derived rather than assumed: it follows from the single economic premise that downturns both raise the cost of staying and depress the fundamental.

11.2 The cross-section of dependent firms

Firm j's run-date cash flow inherits the downstream structure of Section 10.1: $\tilde{\pi}_j = \bar{\pi}_j + \delta_j(\tilde{Q} - \bar{Q}) + \tilde{u}_j$, with $\bar{Q} = \mathbb{E}[\tilde{Q}]$, dependence $\delta_j \geq 0$, and idiosyncratic $\tilde{u}_j$ ($\mathbb{E}[\tilde{u}_j \mid M] = 0$). With price $P_j = \mathbb{E}[M\tilde{\pi}_j]$ and return $\tilde{R}_j = \tilde{\pi}_j / P_j$:

Proposition 12 (Platform-run risk is priced)**.** *Under Lemma 4 the platform-run factor carries a positive price λ_{run} (common across firms), and firm j's expected excess return is its factor loading $\beta_j = \delta_j / P_j$ times that price:*

$$\mathbb{E}[\tilde{R}_j^e] \;=\; \beta_j \cdot \lambda_{run}, \qquad \beta_j \equiv \frac{\delta_j}{P_j}, \qquad \lambda_{run} \;\equiv\; -(1 + r_f)\ \text{Cov}(M, \tilde{Q}) \;>\; 0. \tag{25}$$

Consequently:

(i) firms with greater single-platform dependence δ_j earn higher expected returns and have valuations that fall more when the perceived fundamental deteriorates toward θ^*;

(ii) multi-homing diversifies*: a firm spread across platforms k with cash flow $\bar{\pi}_j + \sum_k \delta_{jk}(\tilde{Q}_k - \bar{Q}_k) + \tilde{u}_j$ loads on platform k's factor only through $\delta_{jk} \leq \delta_j$, lowering exposure to any single collapse;*

(iii) firms with proprietary data and in-house capability (low δ_j) command a lower run premium.

11.3 Is coordination risk priced? The panic premium

A natural objection is that the panic component of run risk (collapse on fundamentals $\theta \in [\underline{\theta}, \theta^*)$ that would have supported the platform) is a pure coordination artifact and should therefore behave like idiosyncratic noise that the market does not price. The model says the opposite, and sharply.

Decompose the run event by type (Definition 3): $\mathbf{1}\{\text{run}\} = \mathbf{1}\{\text{fundamental run}\} + \mathbf{1}\{\text{panic run}\}$, and write the platform-value shortfall as $\tilde{Q} = Q_{\text{hi}} - (Q_{\text{hi}} - Q_{\text{lo}})\mathbf{1}\{\text{run}\}$. The run-risk price λ_{run} of Proposition 12 then splits additively.

Proposition 13 (The panic premium)**.** *Under the cyclical structure* (24) *and Lemma 4, the run-risk price decomposes as*

$$\lambda_{run} \;=\; \underbrace{\lambda_{fund}}_{\geq 0} \;+\; \underbrace{\lambda_{panic}}_{>0}, \qquad \lambda_{type} \;=\; (1+r_f)\,(Q_{hi}-Q_{lo})\;\mathrm{Cov}\big(M, \mathbf{1}\{type\ run\}\big),$$

and $\lambda_{panic} > 0$ *strictly whenever the panic region has positive probability in high-marginal-utility states. Moreover the panic premium is* undiversifiable within a platform's ecosystem*: averaging firm returns over any set of firms dependent on the same platform leaves* λ_{panic} *intact, because the coordination shock is common to all of them. It is diversified only by multi-homing across platforms whose coordination shocks are independent.*

The economic content is subtle and does not follow from the setup. The coordination failure, *conditional* on the platform being in the fragile band, is not about fundamentals; it is a sunspot-like selection. Yet whether the platform *is* in the fragile band is bad-state-loaded: by (24), downturns raise the hurdle and depress the fundamental, sliding the platform into $[\underline{\theta}, \theta^*)$ exactly when marginal utility is high. The panic thus strikes hardest in bad states, so its risk is *systematic* for everyone who depends on the platform, and it is priced. The lesson reverses the naive intuition: belief-driven, fundamentals-orthogonal risk is not idiosyncratic noise to be diversified away within an ecosystem; it is a common factor, and only multi-homing across platforms (not diversification across firms on one platform) reduces exposure to it. This is the asset-pricing counterpart of Proposition 12(ii) and gives the multi-homing result a deeper rationale.

11.4 Platform equity and the disaster component

Platform equity in the stark case is $V(\theta) = \Pi\,\mathbf{1}\{\theta \geq \theta^*\} + \Pi_{\text{run}}\,\mathbf{1}\{\theta < \theta^*\}$. Its sensitivity to fundamentals is concentrated at the threshold; with the smoothing induced by finite signal noise (so survival probability is the continuous $1 - G((\theta^* - \theta)/\varepsilon)$), equity exhibits high local return volatility near the threshold, pronounced negative skewness, and crash risk. The size of this left-tail disaster component is governed by the panic width $\Delta(1-\bar{H})$: a confidence-sensitive platform is not merely high-beta but carries a built-in disaster component in the rare-disaster sense of Barro (2006), Gabaix (2012), and Wachter (2013), with the distinguishing feature that the disaster is a *coordination* event rather than a fundamental one and its severity is set by the equilibrium primitives Δ and $\bar{H}$.

Remark 9 (Market structure and mispricing)*.* Because outcomes are governed by the sign of $\theta - \theta^*$, realized scale is bimodal, near-monopoly dominance ($\theta > \theta^*$) or collapse ($\theta < \theta^*$), interior scale only on the knife-edge as $\varepsilon \to 0$. Standard platform theory (stable dominance) is the small-Δ limit; *fragile* network effects (large Δ) replace dominance with a knife-edge. If investors price platforms on expected ecosystem rents while neglecting the panic component, dominant platforms are over-funded and resilient or multi-homed firms under-funded relative to run-adjusted value.

12 Cross-Platform Contagion and the Multi-Homing Tension

The single-platform analysis credits multi-homing with diversifying platform-run risk: a firm spread across independent platforms loads less on any one platform's coordination factor (Propositions 12 and 13). That conclusion is incomplete once we recognize *who* multi-homes. The most informative contributors (the sophisticated, footloose power users whose data is most valuable (Section 4)) are precisely the ones who participate on several platforms at once. Their shared participation links platforms, but not through balance-sheet exposures as in models of interbank contagion (Allen and Gale, 2000; Acemoglu et al., 2015); it links them through the *composition of the data that produces quality*. This section shows that the shared-contributor channel makes runs contagious, and that it works against the very diversification multi-homing is credited with.

12.1 Two platforms with shared contributors

Consider two platforms $k \in \{1, 2\}$, each as in the baseline, symmetric for clarity, with run threshold $\theta^* = \kappa - \Delta\bar{H}$ from Theorem 2 (treated as a commonly known constant in this reduced form; the threshold itself is microfounded in Section 12.2). The informative contributors split into a footloose, multi-homing group of mass m with informativeness ι_H, present on both platforms, and platform-loyal groups with informativeness $\iota_L \leq \iota_H$. A common confidence shock $S \in \{0, 1\}$, $\mathbb{P}(S = 1) = \pi$, a coordinating signal of the kind catalogued in Definition 4, e.g. a sector-wide governance, safety, or regulatory event, induces the footloose group, being the least locked-in, to withdraw from *both* platforms at once. Their departure strips informative data from each platform, raising each effective threshold from θ^* to $\theta^* + \Phi_m$, where

$$\Phi_m \;\equiv\; \Xi(\mathcal{I}) - \Xi(\mathcal{I} - m\,\iota_H) \;>\; 0, \qquad \frac{\partial \Phi_m}{\partial m} > 0, \quad \frac{\partial \Phi_m}{\partial \iota_H} > 0, \tag{26}$$

by monotonicity and concavity of the learning technology Ξ. Conditional on S, the platforms' fundamentals θ_1, θ_2 are independent draws from F; we set their exogenous correlation to zero to isolate the new channel. Platform k runs iff $\theta_k < \theta^* + \Phi_m S$.

Proposition 14 (Contagion through shared contributors)**.** *Under this structure, with $F_1 \equiv F(\theta^* + \Phi_m)$ and $F_0 \equiv F(\theta^*)$:*

(i) each platform's marginal run probability is $\bar{p} = \pi F_1 + (1 - \pi)F_0$;

(ii) the joint run probability is $\pi F_1^2 + (1 - \pi)F_0^2$, and the run covariance is

$$\mathrm{Cov}\big(\mathbf{1}_{\mathrm{run}_1}, \mathbf{1}_{\mathrm{run}_2}\big) = \pi(1 - \pi)\,(F_1 - F_0)^2 \;\geq\; 0,$$

strictly positive whenever $\pi \in (0, 1)$ and $m > 0$, and strictly increasing in the multi-homing share m and in the informativeness gap $\iota_H - \iota_L$;

(iii) runs are therefore positively correlated even though fundamentals are independent*: the correlation is manufactured entirely by the shared footloose contributors and vanishes as $m \to 0$.*

Corollary 6 (The multi-homing tension)**.** *A downstream firm that splits its dependence equally across the two platforms suffers disruption variance proportional to* $\frac{1}{2}\big[\bar{p}(1-\bar{p}) + \mathrm{Cov}(\mathbf{1}_{\mathrm{run}_1}, \mathbf{1}_{\mathrm{run}_2})\big]$*, whose covariance term is increasing in contributor multi-homing* m*. The diversification that the pricing analysis attributes to spreading dependence (Proposition 13) is thus fully available only when contributors are platform-specific, and is progressively eroded as contributor multi-homing rises. Interoperability and portability, the very instruments that reduce lock-in and idiosyncratic fragility (Proposition 5), raise the multi-homing share and so raise the systematic co-movement of runs. The same instrument that diversifies idiosyncratic platform risk synchronizes the system: in the rare-runs regime* ($\bar{p} < \frac{1}{2}$) *the equal-split firm's disruption variance is* strictly increasing *in contributor multi-homing* m*, rising from the fully diversified* $\frac{1}{2}\bar{p}(1-\bar{p})$ *at* $m = 0$ *toward the comovement-dominated level as* $m \to 1$*.*

12.2 Microfoundation: a two-platform global game

The common shock S above captured the joint exit of footloose contributors with a reduced form. The same conclusion is derivable from a two-platform global game with private signals and correlated fundamentals, in which the joint exit of multi-homers and the threshold shift Φ_m emerge endogenously rather than by assumption.

Let the platforms' fundamentals (θ_1, θ_2) have a jointly positively-quadrant-dependent distribution F_{12} on $\mathbb{R}^2$ with identical marginals F and dependence parameter $\rho_\theta \in [0, 1)$ (e.g., bivariate normal with correlation ρ_θ). The diffuse-prior limit is taken pointwise as in Section 6. Each contributor of type $\tau \in \{M, L\}$ on platform k observes a private signal $x_{i,k} = \theta_k + \varepsilon\sigma\xi_{i,k}$ with $\xi_{i,k}$ iid standard normal; multi-homers observe two signals (one per platform), loyalists one.

The platform populations are as in Section 12: a mass m of multi-homers (footloose, informativeness ι_H, participation hurdle κ_M), and a mass $(1-m)/2$ of loyalists on each platform (sticky, informativeness $\iota_L \le \iota_H$, hurdle $\kappa_L < \kappa_M$). Quality on platform k is $q_k = \theta_k + \Xi(I_k)$ with $I_k = \iota_H m\, n_{M,k} + \iota_L \frac{1-m}{2}\, n_{L,k}$.

Restrict attention to symmetric monotone equilibria in which type τ on platform k stays iff her signal exceeds a cutoff $\hat{x}_{\tau,k}$. With $\kappa_L < \kappa_M$, the equilibrium has nested cutoffs $\hat{x}_{L,k} < \hat{x}_{M,k}$. In the small-noise limit $\varepsilon \to 0$ the marginal contributor of each type forms Laplacian beliefs (Lemma 2) about her own type's participation rate, uniform on $[0, 1]$. Two regimes pin down the indifference conditions. At signal $\hat{x}_{M,k}$ the loyalists, sitting above their own cutoff, fully participate, so the marginal multi-homer's indifference is

$$\hat{x}_{M,k} = \kappa_M - \bar{\Xi}_M(m), \qquad \bar{\Xi}_M(m) \equiv \int_0^1 \Xi\big(\iota_H m\, n + \iota_L \tfrac{1-m}{2}\big)\, dn. \tag{27}$$

At signal $\hat{x}_{L,k}$ the multi-homers have already exited, so the marginal loyalist's indifference is

$$\hat{x}_{L,k} = \kappa_L - \bar{\Xi}_L(m), \qquad \bar{\Xi}_L(m) \equiv \int_0^1 \Xi\big(\iota_L \tfrac{1-m}{2}\, n\big)\, dn. \tag{28}$$

A *multi-homer run* on platform k is $\theta_k < \hat{x}_{M,k}$ (the informative contributors leave); a *full run* is $\theta_k < \hat{x}_{L,k}$ (loyalists exit too and the platform collapses). The reduced-form threshold shift of Section 12 is recovered as $\Phi_m = \bar{\Xi}_M(m) - \bar{\Xi}_L(m) > 0$, the loss in average quality the marginal loyalist faces once multi-homers are gone.

Remark 10 (Joint Laplacian belief in the nested-cutoff equilibrium). The type-specific Laplacian arguments leading to (27)–(28) are exact in the small-noise limit because, given the strict ordering $\kappa_L < \kappa_M$ and the resulting nested cutoffs $\hat{x}_{L,k} < \hat{x}_{M,k}$, the off-marginal type's participation rate becomes deterministic at each margin. Formally: at a multi-homer's marginal signal $\hat{x}_{M,k}$, the conditional posterior places mass arbitrarily close to $\hat{x}_{M,k}$ as $\varepsilon \to 0$, and since $\hat{x}_{M,k} > \hat{x}_{L,k}$, $\mathbb{P}(\theta_k > \hat{x}_{L,k} \mid x_{i,k} = \hat{x}_{M,k}) \to 1$, so loyalist participation $n_{L,k} \to 1$ in probability. By the Laplacian identity (Lemma 2), the marginal multi-homer's own-type participation rate $n_{M,k} = 1 - G\big((\hat{x}_{M,k} - \theta_k)/(\varepsilon\sigma)\big)$ is uniformly distributed on $[0,1]$ for every $\varepsilon > 0$ at her cutoff signal, regardless of the cross-platform structure. The joint belief at the multi-homer's margin is therefore a Dirac at $n_{L,k} = 1$ tensored with $U[0,1]$ on $n_{M,k}$, and the indifference condition (27) is exact. The argument at the loyalist's margin is symmetric, with $n_{M,k} \to 0$ deterministically and own-type $n_{L,k} \sim U[0,1]$, yielding (28). The same logic extends to any finite number of types with strictly ordered hurdles, with each marginal type's belief over off-marginal types collapsing to a deterministic in/out vector.

Lemma 5 (Multi-homing raises the loyalist threshold). $\dfrac{\partial \bar{\Xi}_L}{\partial m} = -\dfrac{\iota_L}{2} \displaystyle\int_0^1 n\, \Xi'\big(\iota_L \tfrac{1-m}{2} n\big)\, dn < 0$, *so* $\partial \hat{x}_{L,k}/\partial m > 0$*: a larger multi-homing share lowers the platform-loyal informativeness and raises the full-run threshold.*

Theorem 3 (Microfounded cross-platform contagion). *Assume* (θ_1, θ_2) *is positively quadrant dependent with* $\rho_\theta \in (0,1)$ *and* F_{12} *admits a continuously differentiable density. Then:*

(i) Positive contagion (global). *Full-run events* $\{\theta_k < \hat{x}_{L,k}\}$ *have strictly positive covariance for all admissible thresholds, and the covariance is monotone increasing in the dependence parameter* ρ_θ*. This part holds without any tail restriction.*

(ii) Amplification by multi-homing (rare-runs regime). *Assume in addition an elliptical dependence structure for* (θ_1, θ_2) *(in particular the Gaussian copula of Corollary 7), under which the conditional CDF* $\Pr(\theta_2 < a \mid \theta_1 = a)$ *exceeds the marginal* $F(a)$ *for* $F(a) < \frac{1}{2}$ *and falls below it for* $F(a) > \frac{1}{2}$*. Then whenever the full-run threshold satisfies* $F(\hat{x}_{L,k}) < \frac{1}{2}$ *(runs are rare events, the economically relevant case)* $\mathrm{Cov}\big(\mathbf{1}_{\text{full run}_1}, \mathbf{1}_{\text{full run}_2}\big)$ *is strictly increasing in the multi-homing share* m*; for a general PQD copula the sign of* $\partial\,\mathrm{Cov}\,/\partial m$ *equals that of* $\Pr(\theta_2 < \hat{x}_{L,k} \mid \theta_1 = \hat{x}_{L,k}) - F(\hat{x}_{L,k})$.

(iii) Reduced-form recovery. *The covariance formula* $\mathrm{Cov} = \pi(1-\pi)(F_1 - F_0)^2$ *of Proposition 14 obtains as the special case in which* F_{12} *is a two-point Bernoulli mixture over a common factor and* $F_1 - F_0$ *corresponds to the threshold shift* $\hat{x}_L|_{\text{multi-homers out}} - \hat{x}_L|_{\text{multi-homers in}}$.

Corollary 7 (Bivariate normal: the exact regime)**.** *With bivariate normal fundamentals and standardized marginals,* $\mathrm{Cov}(\hat{x}_L) \equiv F_{12}(\hat{x}_L, \hat{x}_L) - F(\hat{x}_L)^2$ *is strictly hump-shaped in the common threshold* $\hat{x}_L$ *and peaks at the marginal median* $F(\hat{x}_L) = \frac{1}{2}$*. Consequently* $d\,\mathrm{Cov}/dm > 0$ *if and only if* $F(\hat{x}_L) < \frac{1}{2}$*: multi-homing amplifies contagion exactly in the rare-runs regime, is neutral at the median, and dampens contagion only in the implausible regime where runs are more likely than not. Within the rare-runs regime the amplification is monotone in* m *on the entire interval* $[0, 1)$*.*

Remark 11 (Positive dependence is doing the work for contagion, not for fragility)*.* The *contagion* result of Theorem 3 depends on positive quadrant dependence of (θ_1, θ_2). The *per-platform* fragility results (unique threshold (Theorem 2), hump shape (Proposition 10), priced panic component (Proposition 13), term structure (Proposition IA.1)) require no cross-platform structure at all and hold whatever the joint distribution of fundamentals across platforms. PQD is implied by, but strictly weaker than, positive correlation; it holds for bivariate normal with positive correlation, for bivariate copulas with positive dependence parameter, and for any joint distribution generated by shared exogenous drivers (compute supply, the labor market for AI talent, the regulatory environment, capital-market sentiment about AI). It does not hold under *negative* quadrant dependence (NQD) of fundamentals, formally, $F_{12}(x, y) < F(x)F(y)$ on the relevant region. NQD is the formal expression of winner-take-all dynamics: one platform's gain is another's loss in fundamentals, as when a small number of close substitutes compete for the same enterprise contracts or compute capacity. Under NQD, Theorem 3 reverses: positive multi-homing exposure *reduces* cross-platform run covariance, because a run on one platform is now an inflow signal to the other. Per-platform fragility, by contrast, is unchanged in either regime. The model's contagion content therefore applies to the regime where common-factor exposure dominates platform-level substitution (the empirically natural regime for frontier providers facing the same upstream constraints) and reverses to a negative-contagion content in the close-substitutes regime. The sign of cross-platform fundamental dependence is the key empirical question that determines which regime obtains, while neither regime affects the per-platform fragility analysis.

The honest content of the theorem is this. Positive cross-platform contagion holds globally under any positive dependence in fundamentals; that part is uncontroversial. The interesting claim, that contributor multi-homing *amplifies* the contagion, requires the rare-runs regime, but that regime is exactly the one a model of platform fragility should care about. For bivariate normal the regime has a sharp characterization: $\hat{x}_L$ below the marginal median, equivalently $\bar{p} < \frac{1}{2}$. The reduced-form result of Section 12 is then a clean special case.

The derivation shows that the contagion result is not an artifact of the reduced-form shock: it holds whenever fundamentals are positively dependent and contributors are heterogeneous in informativeness and mobility, with m amplifying co-movement through the endogenous threshold shift in (28). A further channel, absent from the per-platform Laplacian analysis, is that a multi-homer's signals on the two platforms are themselves correlated through the common-factor component of the fundamentals, so her posterior on one platform's state shifts with her signal on the other. In the small-noise limit signals reveal fundamentals and this collapses to ρ_θ; at non-trivial noise it

strengthens contagion further, as a separate, signal-correlation channel reinforcing the threshold-shift channel of Theorem 3.

The interaction is the point. Two of the model's own results, that the informative users are the footloose ones (data composition, Section 4) and that multi-homing diversifies priced run risk (Section 11), combine to produce a conclusion neither contains alone and that contradicts the natural reading of each: making it easier to multi-home, which any single-platform analysis would call stabilizing, raises systemic fragility. The channel is specific to platforms whose asset is produced by shared, mobile contributors; it has no analogue in bank-run contagion, where banks are linked by claims on one another rather than by depositors who are simultaneously the productive input on several banks at once.

13 Welfare and Policy

The panic runs of Definition 3 are coordination failures on fundamentals that would support the platform, and the downstream sector of Section 10.1 means their cost spills beyond the platform's participants. Three policy questions follow. Are panic runs inefficient, and by how much? Which externalities drive the inefficiency, and what instruments target each? And which tensions between instruments must any second-best policy resolve? We answer the first formally, classify the externalities and instruments, and structure the second-best problem.

13.1 Constrained inefficiency and the two externalities

Definition 6 (Welfare)**.** Utilitarian welfare is the sum of participant surplus, platform profit, and downstream surplus: $\mathcal{W} = \int_0^1 \max\{q(\theta, n) - p,\, R - c\}\, di + \Pi(\text{outcome}) + \int_0^1 \delta_j A_j Q\, dj$, evaluated at the equilibrium participation and outcome.

Proposition 15 (Panic runs are constrained-inefficient)**.** *For $\theta \in [\underline{\theta}, \theta^*)$ (the panic region) the equilibrium selects the run outcome, yet the full-participation outcome Pareto-dominates it: every participant obtains $q(\theta, 1) - p \geq R - c$ and downstream firms obtain $Q_{\text{hi}} > Q_{\text{lo}}$. The welfare loss from a panic run is $\mathcal{W}_{\text{stay}} - \mathcal{W}_{\text{run}} = \int_0^1 [\, q(\theta, 1) - \kappa \,]\, di + (\Pi - \Pi_{\text{run}}) + X > 0$. The loss is increasing in the systemic footprint $\bar{\delta}$ (through X) and in lock-in (through severity, Section 7).*

Two distinct externalities drive this loss and any policy must target one or both. The first is the *coordination externality*: each participant's exit lowers everyone's quality through $\Delta h(n)$, and individuals do not internalize the contribution of their stay decision to others' payoffs. This externality is the same as in classical bank runs. The second is the *downstream externality*: a run costs the downstream sector $X = \bar{\delta}(Q_{\text{hi}} - Q_{\text{lo}})\bar{A}$, but the platform's private incentive to choose resilience effort and integration depth (Section 10) does not weight X at all. This externality is distinct from the coordination externality and is specific to platforms that serve as digital infrastructure for real firms. It scales with systemic footprint $\bar{\delta}$ and with the loss-given-run, but not directly with the panic width $\Delta(1 - \bar{H})$, which the coordination externality alone determines. Effective policy

must therefore deploy different instruments against the two externalities, or accept that any single instrument addresses one at the cost of worsening the other.

13.2 Lowering severity: portability and interoperability

The first class of intervention reduces the severity of runs without reducing their frequency. Portability standards (data export, weight escrow, API stability), interoperability mandates (cross-platform model migration, common protocols), and open-source requirements all reduce the platform-specific value of integration b that enters the severity $L(b) = \ell_0 + \ell b$ of Section 7 (Eq. 17), and hence the severity of a run. By Proposition 4 (the stability illusion), this trade is exact: cutting severity raises frequency, because participants are quicker to exit at the first sign of trouble when they can take their data and complements with them. Formally, the proposition shows that lock-in trades frequency for severity along a curve whose slope is governed by $\bar{H}$ and the curvature of h; portability moves the platform up the curve.

The further consequence, also formal, is the open-source result of Proposition 5: portability and open access reduce catastrophic-run severity but raise run frequency and erode the rents that finance frontier investment. The cross-sectional implication is that the platforms whose downstream sectors would benefit most from portability (the most systemic ones, with the largest $\bar{\delta}$) are also the ones whose frontier investment is most rents-financed, so the policy trade-off is steepest exactly where it appears most attractive. The model does not select among portability regimes; it identifies the trade-off that any portability regime must resolve.

13.3 Eliminating the bad equilibrium: backstops

A second class of intervention works on the coordination externality directly: a credible commitment to support the platform if a run begins makes participants confident that a coordination failure cannot deliver the run outcome, so the run equilibrium is no longer self-enforcing. The classical analogue is the lender of last resort against bank runs (Diamond and Dybvig, 1983; Rochet and Vives, 2004); the platform analogue is a public commitment to compute and data, model-weight escrow that guarantees service continuity, or a contractual right of independent operation under a regulated successor. In the model, such a commitment raises the lower-dominance fundamental $\underline{\theta}$ (the threshold below which exit is dominant even at full participation), narrowing the panic region $[\underline{\theta}, \theta^*)$ or eliminating it entirely if the commitment is sufficient.

The cost is the moral-hazard wedge of Proposition 11: when the platform anticipates rescue, it underinvests in private resilience, raising the equilibrium hurdle κ and partially offsetting the gain from removing the bad equilibrium. The proposition gives this as a strict comparative static, not as a qualitative warning. The implication is that any backstop scheme must be conditioned on a verifiable resilience standard, with the same incentive-compatibility constraints familiar from banking, paid-in deposit insurance, capital ratios, and resolution authority. The natural mapping is: model-weight escrow is the analogue of deposit insurance; minimum compute capacity and a

documented disaster-recovery plan are the analogue of capital ratios; a regulated successor or open licensing on failure is the analogue of an orderly-resolution regime.

13.4 Targeting the downstream externality: systemic designation

The third class of intervention addresses the downstream externality directly. Designating platforms with $\bar{\delta}$ above a threshold as *systemically important digital infrastructure* (SIDI) and imposing additional resilience and transparency requirements internalizes part of X in the platform's choice of e and b. This is the analogue of systemically important financial institution (SIFI) designation in banking, but with a tighter mapping: $\bar{\delta}$ is directly the dependence share of downstream output on the platform, the very quantity that drives the welfare loss in Proposition 15.

The cost is entrenchment: SIDI status confers an implicit reputational and regulatory advantage, raising barriers to entry by smaller competitors and limiting the channel-saturation force of Section 9 that pushes very large platforms toward the safer right tail of the fragility hump. The cross-sectional implication is delicate: SIDI status raises a platform's resilience but also raises its run threshold $\theta^* = \kappa - \Delta\bar{H}$ (through higher κ from compliance costs) and concentrates the multi-homing intensity matrix $\mathbf{M}$ of Theorem IA.1 further on the designated platforms, raising eigenvector centrality. Whether the net effect of SIDI designation on systemic risk is positive or negative is a quantitative question the model frames but does not answer; the qualitative content is that the policy targets the right externality and bears predictable entrenchment costs.

13.5 Disclosure and information policy

Information policy is a fourth instrument, distinct from the three above because it works on beliefs rather than on payoffs. By Proposition 2 and Subsection IA.2.2, increased disclosure precision raises the survival probability *conditional on uniqueness* but threatens uniqueness when it crosses the Morris–Shin contraction threshold. The implication for mandated transparency (real-time customer counts, benchmark performance, governance events) is that it should be set below the uniqueness boundary, not above it. Disclosure regimes that publicize information faster than markets can process it create coordination risk rather than alleviating it; the cross-platform empirical signature is that runs cluster around earnings or benchmark releases above and beyond their fundamental information content. The widely observed reticence of platform operators about real-time customer counts and benchmark performance is, in the model, second-best optimal under coordination risk, not opportunism.

13.6 The second-best problem

The four instruments (portability, backstops, SIDI designation, and disclosure) each address one externality at a different cost, and any unified policy must combine them. The structure of the second-best problem can be stated compactly: portability and open-access regulation reduces severity at the cost of frequency and rents; backstops eliminate the bad equilibrium at the cost of resilience

effort; SIDI designation internalizes the downstream externality at the cost of entrenchment; disclosure raises survival probability at the cost of uniqueness above a threshold. No single instrument is dominated by another, and the optimal mix depends on (i) the relative weight of severity versus frequency in social welfare, (ii) the strength of moral hazard in resilience choice, (iii) the elasticity of entry to designation-induced barriers, and (iv) the precision of public disclosure already in the market.

The model identifies what each instrument targets and what it sacrifices, frames the joint optimization, and rules out two common arguments. First, that the bank-runs analogy implies a deposit-insurance-style response is direct and uncomplicated: in fact, the closest analogue is a complex of all four instruments, not deposit insurance alone, because the downstream and information externalities have no exact banking analogues. Second, that AI-platform fragility is a non-issue because the largest platforms have been durable: Proposition 10 explains the durability but also shows that mid-sized platforms can be more run-prone than either small or very large ones, so the policy-relevant cross-section is non-monotone in platform scale.

14 Central Predictions

The analysis yields many comparative statics, but five theoretical predictions are central: they are the claims that separate the confidence-sensitive view of AI platforms from the network-effects-as-moats benchmark, and the remaining results are refinements of them.

1. **Dominance is punctuated by collapse, not eroded by decline.** Because the outcome is governed by the sign of $\theta - \theta^*$ with $\theta^* = \kappa - \Delta\bar{H}$ (Theorem 2), a platform operates at near-full participation until the fundamental crosses the threshold and then collapses discontinuously. Realized scale is bimodal, and a dominant platform's history is long spells of dominance interrupted by sudden runs rather than a glide down a demand curve (Remark 9).

2. **Fragility is set by the sensitivity and timing of quality, not its level.** What matters is the participation-sensitivity $\Delta = q(\theta, 1) - q(\theta, 0)$ and the curvature $\bar{H}$ of the adoption curve, not the level of quality (Propositions 3 and 6). Holding fundamentals and full-participation quality fixed, a platform whose value requires critical mass (a convex adoption curve, low $\bar{H}$) is more run-prone than one useful to small cohorts, a concave curve, high $\bar{H}$.

3. **Quality collapses faster than headcount when informative users are mobile.** When the most informative contributors are also the most footloose, exit is adversely selected on data quality, so a run degrades the model by more than the loss of users alone implies and the platform is more fragile (Proposition 1). The direction of this data-composition channel is governed by the covariance between informativeness and switching costs, and can stabilize as well as destabilize.

4. **Fragility is hump-shaped in network intensity.** The panic *band* widens with the complementarity Δ, but the *probability* of a panic run rises and then falls (Proposition 10): beyond a

point, a stronger network effect slides the entire fragile region into the low-density tail of fundamentals, so the largest, most network-intensive platforms are less panic-prone than mid-sized ones, and crash risk in the cross-section is non-monotone (Corollary 5).

5. **Platform-run risk is priced, and its panic component is systematic when runs fall in bad aggregate states.** Firms whose operations rest on a single platform bear a positive run-risk premium that rises with dependence and falls with multi-homing (Proposition 12); when runs strike in high-marginal-utility states (Lemma 4), the panic component (a pure coordination failure) is priced rather than idiosyncratic and does not diversify away across firms tied to the same platform, only across independent platforms (Proposition 13); and platform equity carries a left-tail disaster component whose size is the panic width (Section IA.1).

The stability illusion, the too-connected-to-fail externality, and the open-source trade-offs (Sections 8 and 10) are corollaries that follow once these five are in place.

15 Discussion and Conclusion

The paper places AI platforms within finance's theory of fragile institutions and identifies what is structurally distinct: the asset run on is produced by the run-able flow, with its quality depending on the composition of who stays. The most informative contributors are also the most mobile, so a confidence shock strips out the best data first and quality falls faster than headcount. A global-games refinement converts the loose "two equilibria" intuition into a unique equilibrium with an exact average-quality threshold $\theta^* = \kappa - \Delta\bar{H}$, an endogenous run probability decomposable into fundamental and panic components of width $\Delta(1-\bar{H})$, sharp comparative statics, and priced risk with a left-tail disaster component scaled by the panic width. The forces are the forces of a bank run; only the object of withdrawal (trust, data, developer effort, adoption, complements) is new.

What is distinctive relative to existing finance theory comes to three things. First, the asset and the claim are not separable, so the standard maturity-mismatch framing does not apply; fragility is set instead by the sensitivity of quality to participation, Δ, and the curvature of the adoption curve, $\bar{H}$. Second, when runs concentrate in high-marginal-utility states the panic component of run risk is priced and undiversifiable within an ecosystem (Proposition 13); the diversification a single firm engineers across platforms is precisely what makes the system more correlated (Theorem 3). Third, the cross-section of run risk is non-monotone in network intensity (Proposition 10, Corollary 4): the largest, most network-intensive platforms can be safer than mid-sized ones, reconciling the run mechanism with the observed durability of frontier incumbents without abandoning it.

Three operational consequences follow for policy and asset pricing. The policy menu (Section 13) is a complex of four instruments, portability, backstops, systemic-designation, and disclosure, each targeting a distinct externality, with the second-best mix determined by relative welfare weights on severity versus frequency, the moral-hazard wedge, and the entrenchment cost of designation. For asset pricing, the panic premium scales as $\gamma\bar{\delta}\pi_{\text{panic}}$ (Proposition IA.4): cross-sectionally, two

platforms with identical fragility primitives but different systemic weights carry different panic premia. For dynamics, the survival curve has a closed-form term structure (Proposition IA.1) and the CDS-style price of platform survival sits within the same first-passage frame as structural credit risk.

15.1 Design choices and the scope of the contribution

Several design choices make the central mechanism operative. Each is active (it makes a specific structural feature of AI platforms transparent) not a residual simplification.

A user-side mechanism. The model studies runs on the participation flow that produces platform quality. Investor-side runs (rollover of equity or debt financing) and platform-side strategic responses are left for adjacent analyses, the focus is on what is structurally distinct to AI platforms, where the live-participation production of quality has no banking analogue. The decomposition of run risk into fundamental and panic components and the priced panic premium both ride on this scoping: they are the run-side counterpart of credit risk in a Merton model, with the same first-passage frame but a different underlying.

Global-games refinement. Coordination is resolved by private signals of fundamentals (Theorem 2), not by sunspots or ad-hoc selection. The refinement delivers a determinate run probability and the clean fundamental–panic decomposition that drives the pricing of Section 11; without it the run probability would be at best a basin-of-attraction argument and the panic premium would lack a probability scale.

Diffuse prior and additive separability. These give the closed form $\theta^* = \kappa - \Delta\bar{H}$ that organizes the paper's comparative statics. The diffuse prior buys the closed form, not the uniqueness of equilibrium; the additive separability buys the average-quality interpretation, not the existence of an equilibrium. The choice is what makes the threshold's economic content visible, average quality across participation rates is the object that the marginal participant compares to the hurdle κ.

Continuum of participants. The continuum $i \in [0, 1]$ is essentially exact for consumer-facing platforms with billions of users and a controlled approximation for enterprise platforms with hundreds of customers. The choice keeps the noise-aggregation Laplacian argument exact (Lemma 2); without it the equilibrium would be a function of population size, and the cross-platform comparisons of Section 12 would carry a finite-population correction term that obscures the dependence-of-fundamentals channel the paper studies.

Static channels, dynamic first-passage. The complementarity Δ is the equilibrium of an interacting channel system (Section 9) at the static level; the threshold becomes the absorbing barrier of an arithmetic-Brownian first-passage problem at the dynamic level (Section IA.1). The combination is what lets us price platform survival like a defaultable claim while keeping the channel structure microfounded, a fully dynamic channel system would couple the time-evolution of Δ to the run window itself and lose the closed-form term structure.

15.2 What survives relaxation

The headline results survive each natural weakening of the design choices; the boundary of survival separates what is generic in the framework from what is specific to the choice.

Uniqueness without the diffuse prior. Theorem 2's unique-equilibrium claim is the standard global-games consequence of dominance regions plus strategic complementarities. It holds for any proper prior with continuous strictly-positive density over the coordination region (Corollary 1). The closed form $\hat{x}^* = \kappa - \Delta\bar{H}$ requires the diffuse prior; the comparative-statics signs of Corollary 2 do not. Empirically, predictions about *relative* fragility across platforms are robust to the prior; predictions about *levels* require explicit prior assumptions.

Non-monotone fragility without eventual unimodality. Proposition 10's hump shape extends to fat-tailed and multimodal fundamental distributions (Corollary 4). Under a left tail regularly varying with index $\alpha_F > 0$, $P_{\text{panic}}(\Delta) \sim c\Delta^{-\alpha_F}$, the right-tail decay is polynomial rather than exponential, but the hump exists and the cross-sectional non-monotonicity is preserved. Under multimodal F, multiple local maxima can appear but the qualitative non-monotonicity holds. Eventual unimodality buys uniqueness of the peak, not the existence of a hump.

Conditional predictions across primitive signs, on an unconditional skeleton. Two primitives carry the model's substantive directional content rather than buying it, and they do so without affecting the headline structural results. The covariance $\text{Cov}(\iota, c)$ between informativeness and switching cost determines whether data composition is destabilizing (Proposition 1) or stabilizing, but the existence of a unique threshold, the hump shape, and the priced panic component hold for any sign (Remark 3). Cross-platform fundamental dependence determines whether contagion is positive (PQD; Theorem 3) or negative (NQD), but per-platform fragility is unchanged in either regime (Remark 11). The cross-section of AI platforms maps to both sides of each split: frontier providers and participation-produced knowledge platforms to destabilizing data composition, enterprise SaaS with reputation-locked users to stabilizing; frontier providers sharing common upstream drivers to PQD-contagion, close-substitute markets to NQD. What looks like an assumption in each case is the structural variable whose sign distinguishes empirically distinct classes of platform, an empirical question the model identifies, not a theoretical question it leaves open.

Functional forms and their precision. The continuum-of-agents limit is exact as the population size $K \to \infty$; the qualitative content holds for finite-K corrections. The noise distribution G is required to have a continuous density and mean zero; symmetry is not required, and asymmetric or skewed mean-zero noise leaves the threshold formula $\theta^* = \kappa - \Delta\bar{H}$ unchanged (Remark 4). The fundamental dynamics are arithmetic Brownian for closed forms; the qualitative term structure of Proposition IA.1 extends to mean-reverting and jump-diffusion fundamentals at the cost of closed-form expressions (Remark IA.2). The Frankel and Pauzner (2000) static-threshold approximation in Section IA.1 is correct to terms vanishing as μ, σ become small relative to the revision rate (Remark IA.1); for high-volatility platforms in fast-news environments the moving-barrier correction is second-order but *positive*, so CDS-style spreads from (IA.2) are conservative. The bounded-returns assumption on the reinforcement matrix $\mathbf{A}$ holds tightly in the saturation regime where

the largest platforms operate; in the rapid-scaling regime the multiplier is larger but the non-monotonicity of Proposition 10 survives with attenuated magnitude. The CRRA scaling of the panic premium (Proposition IA.4) is informative precisely for platforms with non-trivial $X_{\text{run}}/\bar{C}$ (systemic platforms) and shrinks to third-order for small platforms, a feature rather than a bug: it predicts that panic premia are large where systemic-importance designation is needed and small elsewhere.

15.3 The research agenda the paper creates

The mechanism developed here is the foundation for several natural extensions, each itself a paper-sized question rather than a missing piece of this paper.

Coupled user–investor runs. Frontier AI platforms are heavily venture-financed. Embedding two populations with a non-additive payoff structure (failure of one population imposing a discrete cost on the other's stay value) generates a coupled global game that connects this paper's user-side mechanism to the dynamic-debt-runs apparatus of He and Xiong (2012). The natural cross-fragility prediction is that AI-platform run probabilities load on financing signals (funding-round news, layoff announcements) above and beyond their fundamental information content.

Endogenous forking. Contributors collectively coordinating on a fork during a run create a three-action global game with two productive attractors. Whether the fork attractor exists depends on portability of data, code, and weights, and on the cost of coordinated migration. Equilibrium-selection theory in three-action global games with two productive attractors is open; resolving it would give Proposition 5's open-source result a structural mechanism (the fork's basin of attraction) rather than a parametric one.

Dynamic data composition. The first-passage analysis treats the threshold as static during the run window, with the architectural cost bounded in Remark IA.1. The fully dynamic version requires a Hamilton–Jacobi–Bellman system with a moving barrier $\theta^*(t) = \kappa - \Delta\bar{H}(\bar{\iota}(t))$ and type-dependent exit hazards governing $\bar{\iota}(t)$. The static-threshold CDS spreads of Proposition IA.2 are conservative; the dynamic spreads are an open object whose magnitude is the natural empirical quantification of the data-composition mechanism.

Optimal information design. Platform disclosure is a Bayesian-persuasion problem under a coordination-uniqueness constraint, with the optimum bounded above by the Morris–Shin contraction threshold of Proposition 2. The framing is in place; the formal solution, including the optimal disclosure schedule across fundamental states and the empirical signature of optimally-opaque platforms around governance events, is open. The natural test is whether platforms exhibit predictable opacity around exactly the events Proposition 2 identifies as uniqueness-threatening.

Endogenous downstream dependence. Firms' dependence shares δ_j are exogenous in the model; endogenizing them generates a stabilizing fixed-point feedback between platform run probability and aggregate systemic weight. The natural setting is a downstream-industry-dynamics paper that takes the platform's run probability as given and solves for the cross-section of δ_j; the answer feeds back into the panic-premium scaling of Proposition IA.4 and sharpens the systemic-importance

designation analysis of Section 13.

Empirical companion. Table IA.1 maps the five central predictions of Section 14 to specific empirical proxies. Implementing these on a panel of frontier foundation-model providers, participation-produced knowledge platforms, and AI-application ecosystems would estimate the cross-sectional shape of the fragility curve, the magnitude of the panic premium under realistic systemic weights, and the sign of $\mathrm{Cov}(\iota, c)$ in each platform class. The measurement is a substantial separate enterprise; the model identifies the objects to be measured and the variation that identifies them.

15.4 Closing

The substantive overlap with classical financial fragility is precise rather than analogical: every formal claim derives from primitives, and every coupling with the existing run-theory literature is made explicit at the point where it bites. The mechanism is robust in a precise sense, fragility is set by Δ and $\bar{H}$ and is hump-shaped in network intensity, and relaxing any design choice locates platforms differently along these curves but does not alter their shape. The research agenda the paper creates inherits the framework's primitives and is the natural sequence in which the analysis would be extended.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80, 1977–2016.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015): "Systemic Risk and Stability in Financial Networks," *American Economic Review*, 105, 564–608.

Allen, F., and D. Gale (2000): "Financial Contagion," *Journal of Political Economy*, 108, 1–33.

Barro, R. J. (2006): "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 121, 823–866.

Black, F., and J. C. Cox (1976): "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions," *Journal of Finance*, 31, 351–367.

Carlsson, H., and E. van Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018.

Chen, Q., I. Goldstein, and W. Jiang (2010): "Payoff Complementarities and Financial Fragility: Evidence from Mutual Fund Outflows," *Journal of Financial Economics*, 97, 239–262.

Crouzet, N., J. C. Eberly, A. L. Eisfeldt, and D. Papanikolaou (2022): "The Economics of Intangible Capital," *Journal of Economic Perspectives*, 36(3), 29–52.

Diamond, D. W., and P. H. Dybvig (1983): "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 91, 401–419.

Elliott, M., B. Golub, and M. O. Jackson (2014): "Financial Networks and Contagion," *American Economic Review*, 104, 3115–3153.

Falato, A., I. Goldstein, and A. Hortaçsu (2021): "Financial Fragility in the COVID-19 Crisis: The Case of Investment Funds in Corporate Bond Markets," *Journal of Monetary Economics*, 123, 35–52.

Farboodi, M., and L. Veldkamp (2021): "A Growth Model of the Data Economy," NBER Working Paper 28427.

Frankel, D. M., and A. Pauzner (2000): "Resolving Indeterminacy in Dynamic Settings: The Role of Shocks," *Quarterly Journal of Economics*, 115, 285–304.

Gabaix, X. (2012): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," *Quarterly Journal of Economics*, 127, 645–700.

Gill, B. S. (2025): "Algorithmic Stablecoins as Managed Bubbles: Network Robustness and Optimal Termination," working paper, Stevens Institute of Technology.

Gill, B. S., and K. John (2026a): "Transparency and Stablecoin Runs," working paper, Stevens Institute of Technology and New York University.

Gill, B. S., and K. John (2026b): "Run Dynamics in Restaking Protocols: A Microfounded Spectral Theory of Operator-Capacity Contagion," working paper, Stevens Institute of Technology and New York University.

Gill, B. S., and K. John (2026c): "Cross-Issuer Substitution and Stablecoin Bailouts," working paper, Stevens Institute of Technology and New York University.

Goldstein, I., H. Jiang, and D. T. Ng (2017): "Investor Flows and Fragility in Corporate Bond Funds," *Journal of Financial Economics*, 126, 592–613.

Goldstein, I., and A. Pauzner (2005): "Demand-Deposit Contracts and the Probability of Bank Runs," *Journal of Finance*, 60, 1293–1327.

Harrison, J. M. (1985): *Brownian Motion and Stochastic Flow Systems*, Wiley, New York.

He, Z., and W. Xiong (2012): "Dynamic Debt Runs," *Review of Financial Studies*, 25, 1799–1843.

Jones, C. I., and C. Tonetti (2020): "Nonrivalry and the Economics of Data," *American Economic Review*, 110, 2819–2858.

Judd, K. L. (1985): "The Law of Large Numbers with a Continuum of IID Random Variables," *Journal of Economic Theory*, 35, 19–25.

Kacperczyk, M., and P. Schnabl (2013): "How Safe Are Money Market Funds?" *Quarterly Journal of Economics*, 128, 1073–1122.

Katz, M. L., and C. Shapiro (1985): "Network Externalities, Competition, and Compatibility," *American Economic Review*, 75, 424–440.

Merton, R. C. (1974): "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29, 449–470.

Morris, S., and H. S. Shin (1998): "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *American Economic Review*, 88, 587–597.

Morris, S., and H. S. Shin (2002): "Social Value of Public Information," *American Economic Review*, 92, 1521–1534.

Morris, S., and H. S. Shin (2003): "Global Games: Theory and Applications," in *Advances in Economics and Econometrics*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, Cambridge University Press, 56–114.

Rochet, J.-C., and X. Vives (2004): "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?" *Journal of the European Economic Association*, 2, 1116–1147.

Schroth, E., G. A. Suarez, and L. A. Taylor (2014): "Dynamic Debt Runs and Financial Fragility: Evidence from the 2007 ABCP Crisis," *Journal of Financial Economics*, 112, 164–189.

Sun, Y. (2006): "The Exact Law of Large Numbers via Fubini Extension and Characterization of Insurable Risks," *Journal of Economic Theory*, 126, 31–69.

Wachter, J. A. (2013): "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *Journal of Finance*, 68, 987–1035.

Internet Appendix

Runs on Artificial Intelligence Platforms

This appendix collects supplementary material. Section IA.1 develops the continuous-time version of the model; Section IA.2 develops three extensions and delimits the model's scope; Section IA.3 maps the central predictions to empirical proxies; Section IA.4 contains all proofs. Equations, results, and floats in this appendix are numbered with an "IA" prefix.

IA.1 Dynamics: The Run as First Passage and the Price of Survival

We now let the fundamental move over time. A run becomes the *first passage* of the fundamental across the threshold, and platform survival is priced as a defaultable claim, turning the static disaster component into closed-form objects. The construction parallels the dynamic-debt-runs analysis of He and Xiong (2012), with two substantive differences: the coordinating barrier is the average-quality threshold $\theta^* = \kappa - \Delta\bar{H}$ rather than a debt face value, and what crosses it is the participation-produced fundamental of a non-financial platform.

IA.1.1 Setup

Fix a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ carrying a standard $\mathbb{P}$-Brownian motion W. The fundamental evolves as the arithmetic diffusion

$$d\theta_t \;=\; \mu\, dt + \sigma\, dW_t, \qquad \theta_0 \text{ given}, \tag{IA.1}$$

with drift μ (trend in stand-alone quality, reliability, and competitive standing) and volatility $\sigma > 0$ (the flow of news). The run threshold $\theta^* = \kappa - \Delta\bar{H}$ is the constant barrier from Theorem 2. Define the *distance to run* $D_t \equiv \theta_t - \theta^*$, with $D_0 > 0$, and the *run time* $\tau \equiv \inf\{t \geq 0 : D_t \leq 0\}$.

Remark IA.1 (Dynamic foundation for the barrier and its limits). Using the *static* threshold θ^* as the absorbing barrier is justified by Frankel and Pauzner (2000): when fundamentals follow a random walk and participants revise actions at dispersed Poisson times, the dynamic coordination game has a unique equilibrium in which the regime switches the first time the fundamental crosses a threshold, and that threshold coincides with the static global-games threshold up to terms vanishing as μ, σ become small relative to the revision rate λ. The first-passage analysis below treats θ^* as a constant barrier and the corrections as second order. Two distinct sources of correction merit acknowledgment. First, the Frankel–Pauzner correction itself scales as $(\mu/\lambda, \sigma^2/\lambda)$: for very high-volatility platforms in fast-news environments, the second-order correction need not be small, and the dynamic threshold drifts toward the running fundamental. Second, the data-composition mechanism of Section 4 introduces a force the Frankel–Pauzner argument does not

contemplate: during a run window, the most informative users exit first, so the average informativeness of remaining participants declines and the effective threshold $\theta^*(t) = \kappa - \Delta\bar{H}(\bar{\iota}(t))$ drifts upward inside the run window itself, raising the local hazard. This is the dynamic content of the data-composition channel, and is the natural focus of the moving-barrier extension discussed in Section 15's research agenda. The closed-form term structure below should be read as the static-threshold benchmark; CDS-style spreads computed from (IA.2) understate run risk for platforms with steep informativeness-vs-mobility gradients.

IA.1.2 Term structure of run risk

Proposition IA.1 (Survival curve, run density, and lifetime run probability)**.** *Under* (IA.1) *with* $D_0 > 0$*, the probability of no run by horizon* t *is*

$$S(t) \;=\; \Phi\Big(\tfrac{D_0+\mu t}{\sigma\sqrt{t}}\Big) - e^{-2\mu D_0/\sigma^2}\,\Phi\Big(\tfrac{-D_0+\mu t}{\sigma\sqrt{t}}\Big), \tag{IA.2}$$

the run-time density is inverse-Gaussian,

$$f_\tau(t) \;=\; \frac{D_0}{\sigma\sqrt{2\pi t^3}}\,\exp\Big(-\tfrac{(D_0+\mu t)^2}{2\sigma^2 t}\Big), \tag{IA.3}$$

the run intensity (hazard) is $\zeta(t) = f_\tau(t)/S(t) = -d\ln S(t)/dt$*, and the lifetime run probability is*

$$\mathbb{P}(\tau < \infty) = \begin{cases} 1, & \mu \le 0, \\ e^{-2\mu D_0/\sigma^2}, & \mu > 0. \end{cases}$$

$\zeta(t)$ *is decreasing in* D_0*; with* $\mu < 0$ *the platform eventually runs with probability one, while with* $\mu > 0$ *it survives forever with probability* $1 - e^{-2\mu D_0/\sigma^2}$*.*

The survival curve (IA.2) is the term structure of run risk: hump-shaped in the near term when the platform sits close to the barrier with little drift, downward-sloping when fundamentals improve. Every static comparative static enters through $D_0 = \theta_0 - (p + R - c) + \Delta\bar{H}$: a price increase or stronger rival lowers D_0 and steepens the term structure; deeper lock-in raises D_0 and flattens it.

Remark IA.2 (Alternative fundamental dynamics). The first-passage analysis uses arithmetic Brownian motion (IA.1) for closed forms, but the qualitative term-structure content extends to a broader class of fundamental processes. Under *mean-reverting* fundamentals $d\theta_t = \mu(\bar{\theta} - \theta_t)\,dt + \sigma\,dW_t$ (Ornstein–Uhlenbeck), if $\bar{\theta} > \theta^*$ the process is recurrent above the barrier with positive probability and the lifetime run probability is the absorbing-probability of a constrained OU process, a Sturm–Liouville object with no closed form but a standard series representation; the hazard is decreasing in $\bar{\theta} - \theta^*$, the steady-state distance to run, replacing D_0. Under *jump-diffusion* fundamentals $d\theta_t = \mu\,dt + \sigma\,dW_t - dJ_t$ with downward-jump process J of intensity λ_J and mean jump size $\bar{\jmath}$, the run hazard inherits a jump component: $\zeta(t) = \zeta_{\text{diff}}(t) + \lambda_J\mathbb{P}(J_t > D_t)$, and CDS-style spreads acquire a state-independent floor proportional to λ_J that an arithmetic-Brownian benchmark misses.

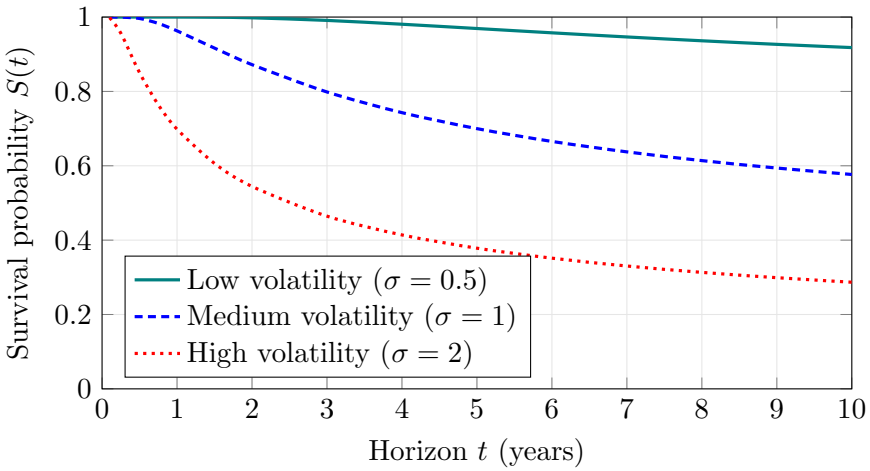


Figure IA.1: Term structure of run risk (Proposition IA.1). Survival curve $S(t)$ for arithmetic-Brownian fundamental with initial distance to run $D_0 = 2$, drift $\mu = 0.1$, and three volatility levels $\sigma \in \{0.5, 1, 2\}$. Low-volatility platforms have a near-flat term structure: the platform sits comfortably above the threshold and survives indefinitely with high probability. Medium volatility produces a downward-sloping curve, the closest analogue to standard credit-survival term structures. High volatility makes the platform run-prone within months, with survival probability roughly halving by years two to three. The hazard $\zeta(t) = -d \ln S(t)/dt$ rises sharply with σ and falls sharply with D_0.

Both extensions preserve the qualitative content of Proposition IA.1, a non-trivial term structure pinned by the platform's distance to the average-quality threshold, while changing the closed-form expressions; the arithmetic-Brownian benchmark is the tractable case, and the empirical companion would estimate which process best fits each platform's fundamental dynamics.

Calibrated illustration. Figure IA.1 plots $S(t)$ for $D_0 = 2$, $\mu = 0.1$, and three volatility levels. The economic content is that platform survival has a meaningful term structure: even with positive fundamental drift, sufficient news-flow volatility produces near-term run risk that compounds with horizon. CDS-style protection spreads (Proposition IA.2) inherit this shape under the risk-neutral measure, with the wedge between physical and risk-neutral spreads pricing the panic component of run risk. The empirical proxy for D_0 is the platform's headroom relative to its average-quality threshold (compute, data, governance health, revenue cushion); the empirical proxy for σ is the volatility of news flow about benchmark performance, regulatory action, and key personnel.

IA.1.3 The price of platform survival

Let the pricing kernel follow $d\Lambda_t/\Lambda_t = -r_f\, dt - \lambda\, dW_t$ with risk-free rate r_f and market price of fundamental risk λ. By Girsanov, $W_t^{\mathbb{Q}} = W_t + \lambda t$ is a $\mathbb{Q}$-Brownian motion and the fundamental drifts at $\mu^{\mathbb{Q}} = \mu - \lambda\sigma$ under $\mathbb{Q}$; (IA.2)–(IA.3) hold under $\mathbb{Q}$ with $\mu \mapsto \mu^{\mathbb{Q}}$, giving $S^{\mathbb{Q}}, f_\tau^{\mathbb{Q}}, \zeta^{\mathbb{Q}}$. Because runs strike in high-marginal-utility states ($\lambda > 0$, Lemma 4), $\mu^{\mathbb{Q}} < \mu$: the risk-neutral run intensity exceeds the physical one, and the wedge is the dynamic run-risk premium.

Proposition IA.2 (CDS-style survival spread and platform equity)**.** *Consider protection paying loss-given-run* L *at* τ *if* $\tau \le T$ *against a premium paid continuously at rate* c *while the platform is alive. The fair spread is*

$$c^*(T) \;=\; \frac{\mathsf{L}\displaystyle\int_0^T e^{-r_f s} f_\tau^{\mathbb{Q}}(s)\,ds}{\displaystyle\int_0^T e^{-r_f s} S^{\mathbb{Q}}(s)\,ds} \;\xrightarrow[T\to 0]{}\; \mathsf{L}\cdot\zeta^{\mathbb{Q}}(0^+), \tag{IA.4}$$

the credit-triangle relation. Platform equity, paying flow $\pi > 0$ *while alive and zero in a run, is worth*

$$V_0 \;=\; \mathbb{E}^{\mathbb{Q}}\Big[\int_0^\tau e^{-r_f s}\pi\,ds\Big] \;=\; \frac{\pi}{r_f}\big(1-e^{-\beta D_0}\big), \qquad \beta \;=\; \frac{\mu^{\mathbb{Q}}+\sqrt{(\mu^{\mathbb{Q}})^2+2r_f\sigma^2}}{\sigma^2} \;>\; 0. \tag{IA.5}$$

Hence $\partial V_0/\partial D_0 = (\pi\beta/r_f)e^{-\beta D_0} > 0$*, largest as* $D_0 \to 0$*: value is most responsive to fundamentals near the run. As* $D_0 \to \infty$*,* $V_0 \to \pi/r_f$ *(the run-free perpetuity); as* $D_0 \to 0$*,* $V_0 \to 0$*.*

Equation (IA.5) is the disaster component in closed form: the convexity of V_0 in the distance to run generates amplified return volatility, negative skewness, and crash risk as θ deteriorates toward θ^*. The loss-given-run L inherits the cross-sectional structure of the static model, rising with the panic width $\Delta(1-\bar{H})$ (Proposition 3) and with lock-in (Proposition 4), so the spread (IA.4) is high precisely for platforms that are both close to the barrier and catastrophic to leave.

IA.1.4 Distance-to-run

Define the standardized distance-to-run over horizon T, $\mathrm{DTR}(T) \equiv (D_0 + \mu^{\mathbb{Q}}T)/(\sigma\sqrt{T})$; to leading order the risk-neutral run probability is $\Phi(-\mathrm{DTR}(T))$, the platform analogue of the Merton–KMV distance-to-default (Merton, 1974; Black and Cox, 1976).

IA.2 Extensions and the Model's Scope

The baseline model isolates a specific mechanism, participation-produced quality, refined by a global game and disciplined by data-composition heterogeneity, and prices the resulting risk. Three extensions strengthen the main results: strategic platform response during a run (Subsection IA.2.1, which microfounds the two-faces logic of Proposition 4), an endogenous stochastic discount factor (Subsection IA.2.2, which sharpens the panic premium of Proposition 13), and N-platform contagion (Subsection IA.2.3, which generalizes Theorem 3). Two further subsections probe robustness to the information and composition assumptions (Subsection IA.2.4) and give a calibrated illustration of the run prices (Subsection IA.2.5).

IA.2.1 Strategic platform response during a run

The baseline platform commits resilience effort e and integration depth b ex ante, then observes the run passively. A real platform observing a run forming can lower price, offer commitments, or publicly disclose information. Add an intermediate stage in which the platform observes a public signal y of θ with posterior $(\theta|y) \sim N(\mu(y), \sigma_y^2)$ and chooses a subsidy $\tau(y) \geq 0$ that lowers the participation hurdle to $\kappa - \tau(y)$ at quadratic cost $\psi_\tau(\tau) = (\kappa_\tau/2)\tau^2$. The threshold shifts to $\theta^*(y) = \kappa - \tau(y) - \Delta\bar{H}$, and the platform solves

$$\max_{\tau(\cdot)\geq 0} \mathbb{E}_y\big[\Pi \cdot S(\theta^*(y); y)\big] - \mathbb{E}_y\big[\tfrac{\kappa_\tau}{2}\tau(y)^2\big], \qquad S(\theta^*; y) = 1 - \Phi\big((\theta^* - \mu(y))/\sigma_y\big). \tag{IA.6}$$

Proposition IA.3 (Optimal intra-game subsidy)**.** *The interior first-order condition for $\tau^*(y)$ is $\tau^*(y)\kappa_\tau\sigma_y = \Pi\,\phi\big((\kappa - \tau^*(y) - \Delta\bar{H} - \mu(y))/\sigma_y\big)$. The optimal subsidy is strictly single-peaked in the posterior mean $\mu(y)$, attains its maximum where $\mu(y) = \kappa - \tau^*(y) - \Delta\bar{H}$ (the posterior mean coincides with the post-subsidy threshold), and falls to zero as $|\mu(y)| \to \infty$. Expected platform value is strictly increasing in the public-signal precision τ_y at $\tau_y = 0$.*

Subsidies concentrate exactly where fragility bites: at posteriors near the threshold, where the marginal user is pivotal. For clearly-safe platforms and clearly-doomed ones, intervention is wasted. The asymmetry sharpens the two-faces logic of Proposition 4: ex-ante commitment that lowers θ^* at small cost is strictly cheaper than ex-post intervention that buys back already-departed users, because the marginal cost of moving the threshold by intervention is governed by the cost density ϕ, which is highest at the threshold and lowest at posteriors far from it. A platform can substitute ex-ante information acquisition (higher τ_y) for ex-post intervention; the substitution is strict, not merely weak.

IA.2.2 Endogenous stochastic discount factor

Section 11 posits an SDF $M(\omega)$ with the property (Lemma 4) that M is higher in run states than in survival states. Combining with the downstream sector of Section 10.1 closes the loop. Aggregate consumption is $C(\omega) = \bar{C} - X(\omega)$, where $X(\omega)$ is realized downstream disruption from a platform run (21), taking value $X_{\text{run}} = \bar{\delta}(Q_{\text{hi}} - Q_{\text{lo}})\bar{A}$ in run states and 0 in survival states.

Proposition IA.4 (Endogenous SDF and panic-premium scaling)**.** *Under a representative agent with CRRA utility $U(C) = C^{1-\gamma}/(1-\gamma)$ and consumption $C(\omega) = \bar{C} - X(\omega)$,*

$$M(\omega) = \beta\Big(1 - \tfrac{X(\omega)}{\bar{C}}\Big)^{-\gamma}, \qquad \frac{M_{\text{run}}}{M_{\text{survive}}} = \Big(1 - \tfrac{\bar{\delta}(Q_{\text{hi}} - Q_{\text{lo}})\bar{A}}{\bar{C}}\Big)^{-\gamma} > 1. \tag{IA.7}$$

Lemma 4 is therefore a consequence, not an assumption. The panic component of a downstream firm's run-risk premium of Proposition 13 scales linearly with risk aversion, with the platform's systemic weight $\bar{\delta}$, and with the panic-component probability π_{panic}:

$$\lambda_{\text{run}}^{\text{panic}} \;\propto\; \gamma \cdot \bar{\delta} \cdot \pi_{\text{panic}}\,(1 - \pi_{\text{run}}), \tag{IA.8}$$

with proportionality factor depending on the firm's own loss share and the run-state cash-flow magnitude.

Two new comparative statics follow. First, panic premia scale linearly with relative risk aversion, a testable cross-period prediction. Second, panic premia scale linearly with the platform's systemic weight $\bar{\delta}$ holding fragility primitives $(\Delta, \bar{H})$ fixed. Cross-sectionally, two platforms with identical $(\Delta, \bar{H})$ but different systemic weights carry different panic premia, with the ratio approximately equal to the ratio of systemic weights. The pricing claims of Section 11 should be read with M understood as (IA.7), not as a free parameter.

The magnitude of the panic premium is conditional on $X_{\text{run}}/\bar{C}$ being non-trivial. For platforms whose downstream disruption is a meaningful fraction of aggregate consumption, the largest frontier providers and the most-depended-upon infrastructure platforms, the wedge $M_{\text{run}}/M_{\text{survive}} = (1 - X_{\text{run}}/\bar{C})^{-\gamma}$ is materially above one and the linearization $\lambda^{\text{panic}} \propto \gamma\bar{\delta}\pi_{\text{panic}}$ is informative. For platforms that are small in aggregate-consumption terms, early-stage frontier providers, specialized vertical platforms with modest enterprise footprints, $X_{\text{run}}/\bar{C}$ is small and the panic premium is third-order in the platform's dependence share. The model thus predicts a sharp cross-section: panic premia are large and economically relevant precisely for systemic platforms, vanishing as platform size relative to the aggregate economy shrinks. This is the formal version of the qualitative statement that systemic-importance designation is needed for the same platforms whose panic premia are large; the scaling and the designation rationale are the same object.

IA.2.3 N-platform contagion

Section 12 extends to two platforms. With N platforms, let multi-homing matrix $\mathbf{M} = [m_{ij}]_{i,j=1}^{N}$ with $m_{ii} = 1$ and $m_{ij} \in [0, 1]$ the share of platform j's informative contributors who also participate on platform i. Let fundamentals $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$ be jointly positively quadrant dependent with covariance matrix Σ_θ and continuously differentiable joint density.

Theorem IA.1 (N-platform contagion)**.** *Under positive quadrant dependence of $\boldsymbol{\theta}$ and the rare-runs regime $F(\hat{x}_{L,k}) < 1/2$ for all k:*

(i) the run-event covariance matrix $\boldsymbol{\Sigma}_{\text{run}} = [\text{Cov}(\mathbf{1}_{\text{run}_k}, \mathbf{1}_{\text{run}_l})]_{k,l}$ has strictly positive off-diagonal entries;

(ii) under the elliptical-dependence condition of Theorem 3(ii), applied pairwise, each entry $(\boldsymbol{\Sigma}_{\text{run}})_{kl}$ is strictly increasing in the multi-homing intensity $m_{kl} + m_{lk}$ and in the fundamental correlation $(\Sigma_\theta)_{kl}$;

(iii) write $\boldsymbol{\Sigma}_{\text{run}} = \boldsymbol{\Sigma}^{(F)} + \boldsymbol{\Sigma}^{(M)}$ as the decomposition into contributions from fundamental correlation and multi-homing intensity respectively. When the multi-homing contribution dominates in the spectral norm, $\|\boldsymbol{\Sigma}^{(M)}\| > \|\boldsymbol{\Sigma}^{(F)}\|$, Weyl's perturbation inequality gives that the leading eigenvector of $\boldsymbol{\Sigma}_{\text{run}}$ lies within angular distance $\|\boldsymbol{\Sigma}^{(F)}\|/\text{gap}(\boldsymbol{\Sigma}^{(M)})$ of the leading eigenvector of $\boldsymbol{\Sigma}^{(M)}$, which in the rare-runs limit aligns with the leading eigenvector of $(\mathbf{M} + \mathbf{M}^\top)/2$.

The two-platform case of Theorem 3 is the $N = 2$, $m_{12} = m_{21} = m$ specialization. Part (i) establishes positive cross-platform contagion globally and part (ii) gives monotonicity in both linkage channels; both are tight under PQD. Part (iii) is the operationally useful but conditional claim: when contributor multi-homing is the dominant channel of cross-platform linkage, the case relevant for AI platforms, which share contributors more than they share owners or balance-sheet exposures, systemic risk concentrates on the platforms with highest eigenvector centrality in the multi-homing network, and targeted regulation of these central nodes is more efficient than uniform regulation. The dominance condition is the empirical content of (iii). Proof, via element-wise PQD arguments parallel to Theorem 3 combined with Weyl's bound, is in Section IA.4.

IA.2.4 Robustness: correlated signals and non-monotone informativeness

Two maintained features deserve scrutiny: that private signals are conditionally independent, and that informativeness $\iota(u)$ is monotone in the switching-cost rank u. Neither drives the two central objects, uniqueness and the threshold $\theta^* = \kappa - \Delta\bar{H}_{\text{eff}}$.

Proposition IA.5 (Robustness of uniqueness and the threshold)**.** (a) Correlated private signals. *Let* $x_i = \theta + \varepsilon(\eta_i + \rho\zeta)$ *with* $\eta_i \sim N(0,1)$ *i.i.d. across agents,* $\zeta \sim N(0,1)$ *a common component, and* $\rho \geq 0$ *fixed. Signal* differences $x_i - x_j = \varepsilon(\eta_i - \eta_j)$ *are independent of* ζ*, so the strategic uncertainty that selects the equilibrium is governed by the idiosyncratic component alone: as* $\varepsilon \to 0$ *the threshold equilibrium is unique and the threshold converges to* $\theta^* = \kappa - \Delta\bar{H}$*, for every fixed* ρ*. A common component that does* not *vanish (a genuinely observed public signal of* θ*) is instead the configuration of Proposition 2, where uniqueness requires the precision condition* (16)*.* (b) Non-monotone informativeness. *For* any *integrable* $\iota : [0,1] \to \mathbb{R}_+$ *(monotone or not) the effective adoption map* h_{eff} *is increasing and bounded, the equilibrium is unique, and* $\theta^* = \kappa - \Delta\bar{H}_{\text{eff}}$ *with* $\bar{H}_{\text{eff}} = \int_0^1 h_{\text{eff}}(n)\,dn$*. Only the* sign *of the composition effect on fragility (whether selection in exit widens or narrows the panic region) depends on the local behavior of* ι *over the marginal quantiles* $[1-n, 1]$*, through the convexity of* $I(n) = \int_{1-n}^1 \iota(u)\,du$*.*

Uniqueness and the threshold are properties of the *average* adoption curvature and the idiosyncratic noise. Correlation in signals matters only through the public-information channel already analyzed in Proposition 2; the shape of ι matters only for the second-order question of whether composition amplifies or dampens fragility (Proposition 6), not for whether a unique run threshold exists or where it sits. The mechanism is thus robust where it must be (in the existence and location of the run boundary) and delivers its sharper comparative statics only under the stated structure.

IA.2.5 A calibrated illustration

Figure IA.2 calibrates the two prices the run mechanism generates (the probability of a panic run and the priced panic spread it commands) under the Gaussian benchmark $\theta \sim N(\kappa, \sigma_\theta^2)$, $\sigma_\theta = 1.5$, of the body's Figure 2. The panic-run probability $P_{\text{panic}}(\Delta) = F(\kappa - \Delta\bar{H}) - F(\kappa - \Delta)$ is hump-shaped in the complementarity Δ (Proposition 10). The priced panic spread $\mathcal{C}(\Delta) = P_{\text{panic}}(\Delta) \cdot \Delta(1 - \bar{H})\,\bar{\ell}$

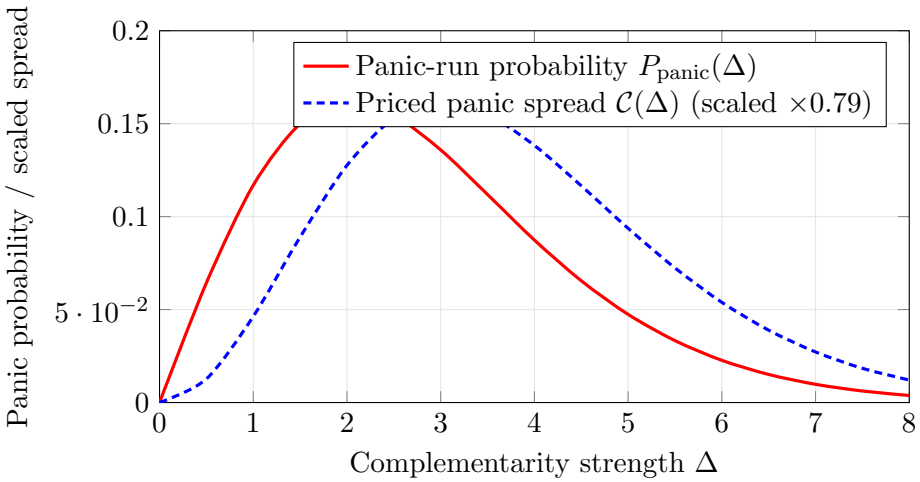


Figure IA.2: Fragility and the priced panic spread under $\theta \sim N(\kappa, 1.5^2)$, $\bar{H} = 0.5$. The panic-run probability $P_{\text{panic}}(\Delta)$ (solid) peaks at $\Delta \approx 2.0$. The priced panic spread $\mathcal{C}(\Delta) = P_{\text{panic}}(\Delta) \cdot \Delta(1 - \bar{H})\,\bar{\ell}$ ($\bar{\ell} = 1$) of Corollary 5 (dashed, rescaled to a common axis by 0.79) peaks at $\Delta \approx 3.0$. The most fragile platform (highest run probability) is therefore not the one carrying the highest priced run risk: because the priced spread also loads on the panic-region width $\Delta(1-\bar{H})$, which grows linearly in Δ, the cross-sectional maximum of priced exposure sits at a strictly higher complementarity than the maximum of fragility. The underlying stochastic-discount-factor covariance is positive in the rare-runs regime, where the panic band falls in high-marginal-utility states (Proposition 13).

of Corollary 5 (severity scale $\bar{\ell}$ normalized to one), panic probability times the panic-region width that the marginal downstream claim loads on, is also hump-shaped but peaks materially to the right: priced run risk is maximized at a higher complementarity than fragility itself.

IA.3 Empirical Proxies for the Central Predictions

The predictions involve model objects (Δ, $\bar{H}$, θ^*, π_{panic}, λ^{panic}) that are not directly observed but each maps to a measurable proxy. Table IA.1 states each mapping, the variation in the data that identifies the prediction, and the principal alternative explanation that would have to be ruled out. The aim is not to test the model here, but to make the testable content unambiguous.

Table IA.1: Mapping the five central predictions to empirical proxies.

Prediction	**Model object**	**Empirical proxy**	**Identifying variation / alternative to rule out**
Punctuated dominance, not gradual decline	$\mathbf{1}\{\theta < \theta^*\}$; distribution of Δ(share) over event windows	Bimodality of cross-sectional growth rates around governance/policy events; jumps in usage at common triggers	Step-function shape of share losses around dated triggers, not smooth decay; alternative is gradual quality decline, which predicts smooth share trajectories
Fragility set by $(\Delta, \bar{H})$, not by quality level	$\Delta = q(\theta, 1) - q(\theta, 0)$; curvature $\bar{H}$	Steepness of usage-vs-active-user curves; benchmark improvement per added user/datapoint at platform level	Cross-platform run frequency conditional on quality level; alternative is uniform pessimism (all large platforms equally fragile)
Quality collapses faster than headcount	$\mathrm{Cov}(\iota, c)$ in user-level data	Cross-sectional change in benchmark scores per percentage-point drop in active-user count; quality measures restricted to retained users	Quality decline exceeds linear prediction in headcount; alternative is proportional degradation, distinguishable by joint dynamics of headcount and quality
Hump-shaped fragility	$P_{\text{panic}}(\Delta)$ (Fig. 2)	Run frequency across quartiles of platform-level network intensity (active users, ecosystem partners, compute spend)	Inverted-U pattern in run frequency vs. network-intensity quartiles; alternative is monotone relationship, which baseline finance theory would predict
Panic premium is priced	$\lambda^{\text{panic}} \propto \gamma \bar{\delta} \pi_{\text{panic}}$ (eq. (IA.8))	Excess returns on platform-dependent firm portfolios over independent-platform portfolios; loadings of equity returns on coordination shocks (governance, benchmark, regulatory events)	Cross-sectional spread in expected returns on dependence quartiles; alternative is that risk loads only on fundamental cash-flow factors, not on coordination events

Three identification challenges run across all five predictions. First, the panic component is by construction not separately observable from the fundamental component on any single realization; identification rests on the joint behavior of survival, news flow, and price reactions across many platforms or many episodes. Second, Δ is endogenous through the channel system of Section 9; cross-sectional comparisons must use proxies that capture the equilibrium Δ, not its underlying components, since the model's predictions are about Δ itself. Third, the threshold $\theta^* = \kappa - \Delta\bar{H}$ shifts with policy and platform choices, so structural break tests should center on dated policy or governance events. The companion to this paper, which we do not undertake here, would implement these tests on a panel of frontier and major established AI platforms.

IA.4 Proofs

Proof of Proposition 1. (i) By construction $h(n) = \Xi(\mathcal{I}(n))/\Xi(\bar{\iota})$ with $\mathcal{I}(1) = \bar{\iota}$, so $h(1) = 1$, $h(0) = 0$, and h is increasing because Ξ and $\mathcal{I}$ are; $\Delta = \Xi(\bar{\iota})$ and $\bar{H} = \int_0^1 h$. With $\iota \equiv \bar{\iota}$, $\mathcal{I}(n) = \bar{\iota}n$ and $h(n) = \Xi(\bar{\iota}n)/\Xi(\bar{\iota})$, concave as the composition of concave Ξ with a linear map; for $\Xi(\mathcal{I}) = A\mathcal{I}^\alpha$, $h(n) = n^\alpha$ and $\bar{H} = \int_0^1 n^\alpha\, dn = 1/(1+\alpha)$. (ii) If ι is decreasing, $\mathcal{I}'(n) = \iota(1-n)$ is increasing in n, so $\mathcal{I}$ is convex on $[0,1]$ with $\mathcal{I}(0) = 0$, $\mathcal{I}(1) = \bar{\iota}$; a convex function with these endpoints lies weakly below the chord $\bar{\iota}n$, strictly on the interior, so $\mathcal{I}(n) \le \bar{\iota}n$ and, Ξ increasing, $h_{\text{eff}}(n) = \Xi(\mathcal{I}(n))/\Xi(\bar{\iota}) \le \Xi(\bar{\iota}n)/\Xi(\bar{\iota}) = h(n)$ pointwise; integrating, $\bar{H}_{\text{eff}} \le \bar{H}$, strict since the inequality is strict on a positive-measure set. Then $\theta^* = \kappa - \Delta\bar{H}_{\text{eff}} > \kappa - \Delta\bar{H}$ and $F(\theta^*)$ rises. For the parametric family, $\int_0^1 (1+\varsigma)\bar{\iota}(1-u)^\varsigma du = \bar{\iota}$ confirms the mean, and $\mathcal{I}(n) = \int_{1-n}^1 (1+\varsigma)\bar{\iota}(1-u)^\varsigma du = (1+\varsigma)\bar{\iota}\int_0^n t^\varsigma dt = \bar{\iota}n^{1+\varsigma}$; hence $h_{\text{eff}}(n) = (\bar{\iota}n^{1+\varsigma})^\alpha/\bar{\iota}^\alpha = n^{\alpha(1+\varsigma)}$ and $\bar{H}_{\text{eff}} = 1/(1+\alpha(1+\varsigma))$, convex iff $\alpha(1+\varsigma) > 1$, i.e. $\varsigma > (1-\alpha)/\alpha$. (iii) Differentiating (6), $q'(n) = \Xi'(\mathcal{I}(n))\,\mathcal{I}'(n) = \Xi'(\mathcal{I}(n))\,\iota(1-n)$; under (ii) $\iota(1-n) > \bar{\iota}$ for $n < 1$, so the marginal stayer is above-average informative and quality falls faster than headcount. Monotonicity of $\bar{H}_{\text{eff}}$ in ς and in dispersion follows since larger ς (or a mean-preserving spread that increases ι on low-u, the early leavers) increases the convexity of $\mathcal{I}$ and lowers each $h_{\text{eff}}(n)$. □

Proof of Theorem 1. The equilibrium concept (Definition 1) requires a fixed point of the best-response correspondence $\mathcal{B} : [0,1] \rightrightarrows [0,1]$, where $\mathcal{B}(n) = \{1\}$ if $\Theta(n) > 0$, $\{0\}$ if $\Theta(n) < 0$, and $[0,1]$ if $\Theta(n) = 0$, with $\Theta(n) \equiv q(\theta, n) - \kappa = \theta + \Delta h(n) - \kappa$. By Assumption 2, Θ is continuous and strictly increasing in n, with $\Theta(0) = \theta - \kappa$ and $\Theta(1) = \theta + \Delta - \kappa$.

Existence. The graph of $\mathcal{B}$ is closed (continuity of Θ), nonempty- and convex-valued (the image is $\{0\}$, $\{1\}$, or all of $[0,1]$), so Kakutani's fixed-point theorem applies on the compact convex set $[0,1]$ and a fixed point exists. We now characterize all fixed points by region.

Case (i): $\theta > \kappa$. Then $\Theta(0) = \theta - \kappa > 0$ and Θ is increasing, so $\Theta(n) > 0$ for all $n \in [0,1]$. Hence $\mathcal{B}(n) = \{1\}$ for all n; the only fixed point is $n^\star = 1$. *Existence and uniqueness in this region.*

Case (ii): $\theta < \kappa - \Delta$. Then $\Theta(1) = \theta + \Delta - \kappa < 0$ and Θ is increasing, so $\Theta(n) < 0$ for all n. $\mathcal{B}(n) = \{0\}$ for all n; the only fixed point is $n^\star = 0$. *Existence and uniqueness in this region.*

Case (iii): $\theta \in (\kappa - \Delta, \kappa)$. Then $\Theta(0) < 0 < \Theta(1)$. Three fixed points:

- At $n = 0$: $\Theta(0) < 0$, so $\mathcal{B}(0) = \{0\} \ni 0$.
- At $n = 1$: $\Theta(1) > 0$, so $\mathcal{B}(1) = \{1\} \ni 1$.
- Interior: n is a fixed point iff $\Theta(n) = 0$, i.e. $h(n) = (\kappa - \theta)/\Delta \in (0, 1)$. Since h is a continuous strictly increasing bijection $[0, 1] \to [0, 1]$ (Assumption 2), there is a unique $n_{\text{int}}(\theta) = h^{-1}((\kappa - \theta)/\Delta) \in (0, 1)$.

No other n is a fixed point: for $n \notin \{0, n_{\text{int}}, 1\}$, $\Theta(n) \neq 0$, so $\mathcal{B}(n) \subset \{0, 1\}$ and $n \notin \mathcal{B}(n)$.

Stability. Use the tâtonnement Definition 2 with $b(n) = \mathbf{1}\{\Theta(n) \geq 0\}$. For $n = 0 + \delta$ with $\delta > 0$ small, $\Theta(n) < 0$ (continuity), so $b(n) = 0 < n$, $\dot{n} < 0$: $n^\star = 0$ is locally stable. For $n = 1 - \delta$, $\Theta(n) > 0$, so $b(n) = 1 > n$, $\dot{n} > 0$: $n^\star = 1$ is locally stable. For $n = n_{\text{int}} \pm \delta$, Θ has opposite signs above and below, so trajectories diverge: n_{int} is unstable. At $\theta = \kappa$, $n_{\text{int}} = h^{-1}(0) = 0$ (the interior root collides with the lower corner); at $\theta = \kappa - \Delta$, $n_{\text{int}} = h^{-1}(1) = 1$ (collides with upper corner). □

Proof of Lemma 1. Under the threshold strategy with cutoff $\hat{x}$, participant i stays iff $x_i = \theta + \varepsilon\eta_i \geq \hat{x}$, i.e. iff $\eta_i \geq (\hat{x} - \theta)/\varepsilon$. Conditional on θ, $\mathbb{P}(\text{stay}) = \mathbb{P}(\eta_i \geq (\hat{x} - \theta)/\varepsilon) = 1 - G((\hat{x} - \theta)/\varepsilon)$. By Assumption 1 the realized fraction equals this probability a.s., giving (10). Continuity and monotonicity follow because G is a continuous, strictly increasing CDF (its density $g > 0$): $\partial n/\partial\theta = \varepsilon^{-1} g((\hat{x} - \theta)/\varepsilon) > 0$ and $\partial n/\partial\hat{x} = -\varepsilon^{-1} g(\cdot) < 0$; $n \to 0$ as $\theta \to -\infty$ and $n \to 1$ as $\theta \to +\infty$. □

Proof of Lemma 2. With the improper uniform prior, the posterior density of θ given $x_i = \hat{x}$ is proportional to the likelihood $\varepsilon^{-1} g((\hat{x} - \theta)/\varepsilon)$; the normalizing constant is $\int_{\mathbb{R}} \varepsilon^{-1} g((\hat{x} - \theta)/\varepsilon)\, d\theta = \int_{\mathbb{R}} g(u)\, du = 1$ (substituting $u = (\hat{x} - \theta)/\varepsilon$). Hence the posterior density of θ is exactly $\varepsilon^{-1} g((\hat{x} - \theta)/\varepsilon)$. Define $z \equiv (\hat{x} - \theta)/\varepsilon$. Then $\theta = \hat{x} - \varepsilon z$ and, by the change of variables, z has posterior density $g(z)$, i.e. $z \sim G$. By Lemma 1, the participation rate faced by the pivotal participant is $n = 1 - G((\hat{x} - \theta)/\varepsilon) = 1 - G(z)$. For any random variable z with continuous CDF G, the probability-integral transform gives $G(z) \sim U[0, 1]$; therefore $n = 1 - G(z) \sim U[0, 1]$, exactly and for every $\varepsilon > 0$. □

Proof of Theorem 2. The proof has six steps: (1) iterated deletion of dominated strategies yields extremal threshold strategies; (2) any threshold equilibrium cutoff $\hat{x}$ satisfies the indifference condition $\Psi(\hat{x}) = \kappa$; (3) closed-form computation of Ψ under additive separability gives $\Psi(\hat{x}) = \hat{x} + \Delta\bar{H}$; (4) *existence* of a cutoff θ^* solving $\Psi(\theta^*) = \kappa$ by the intermediate value theorem applied to Ψ on the coordination region $[\underline{\theta}, \overline{\theta}]$; (5) *uniqueness* from strict monotonicity of Ψ; (6) convergence of the realized participation rate as $\varepsilon \to 0$.

Step 1 (Iterated deletion yields extremal threshold strategies). The game has three structural properties:

e complementarities. Let σ, σ' be symmetric strategies with $\sigma \leq \sigma'$ pointwise (i.e., σ exits more aggressively). By Lemma 1 the realized participation rate under σ is weakly below that under σ' at each θ, so a deviator's payoff from staying, $q(\theta, n(\theta;\sigma)) - \kappa$, is weakly higher under σ' than under σ by the monotonicity of q in n (Assumption 2). The best response to σ' stays in more situations than the best response to σ: the best-response correspondence is monotone in the strong-set order.

State monotonicity. $q(\theta, n)$ is strictly increasing in θ by Assumption 2, so the stay payoff is increasing in the signal: best responses are monotone (threshold) strategies whenever beliefs are monotone, which the diffuse prior plus translation-invariant noise guarantees.

Dominance regions. Assumption 3 ensures that signals with $x_i > \overline{\theta} + M(\varepsilon)$ stay in any rationalizable profile and signals with $x_i < \underline{\theta} - M(\varepsilon)$ exit, where $M(\varepsilon)$ is a noise-implied buffer that vanishes as $\varepsilon \to 0$.

Under (a)–(c), the iterated-deletion correspondence is monotone on the lattice of symmetric monotone (threshold) strategies, ordered by cutoff. Tarski's fixed-point theorem (Morris and Shin, 2003) delivers a largest fixed point with cutoff $\bar{x}$ and a smallest fixed point with cutoff $\underline{x}$ in the set of strategies surviving iterated deletion, with $\underline{x} \leq \bar{x}$; both are pure-strategy Bayes–Nash equilibria of the global game. Every rationalizable strategy is sandwiched between these two. *The existence of at least one equilibrium is established at this step*: the threshold strategy with cutoff $\bar{x}$ (or $\underline{x}$) is a Bayes–Nash equilibrium.

Step 2 (Indifference at the threshold cutoff). Suppose all participants use cutoff $\hat{x}$. The participant with signal exactly $\hat{x}$ is pivotal. By Lemma 2 her posterior is $\theta = \hat{x} - \varepsilon z$ with $z \sim G$, and the participation rate she faces is $n = 1 - G(z)$. Her expected stay payoff (net of the price) is

$$\mathbb{E}_z\big[q(\hat{x} - \varepsilon z, 1 - G(z))\big] - p \;=\; \Psi(\hat{x}) - p,$$

and her exit payoff is $R - c$. For $\hat{x}$ to be an equilibrium threshold cutoff, the marginal type must be indifferent (slightly higher signals strictly prefer staying and slightly lower strictly prefer exiting, by continuity of Ψ and monotonicity in the signal). Indifference is $\Psi(\hat{x}) - p = R - c$, i.e.

$$\Psi(\hat{x}) \;=\; p + R - c \;=\; \kappa, \tag{11'}$$

which is (11). Any threshold equilibrium cutoff solves (11).

Step 3 (Closed form for Ψ under additive separability). Under (2) and the zero-mean-noise condition $\mathbb{E}_G[z] = 0$ (which holds under Assumption 4; see Remark 4 for the weakening from

symmetric to mean-zero),

$$\begin{aligned}\Psi(\hat{x}) \;=\; \mathbb{E}_z[\hat{x} - \varepsilon z + \Delta\, h(1 - G(z))] \;&=\; \hat{x} - \varepsilon \mathbb{E}_G[z] + \Delta \int_{\mathbb{R}} h(1 - G(z))\, g(z)\, dz \\ &=\; \hat{x} + \Delta \int_0^1 h(1-u)\, du \qquad (\text{sub. } u = G(z)) \\ &=\; \hat{x} + \Delta \int_0^1 h(v)\, dv \;=\; \hat{x} + \Delta \bar{H}.\end{aligned}$$

Thus $\Psi(\hat{x}) = \hat{x} + \Delta\bar{H}$ is a linear continuous function of $\hat{x}$ with $\Psi'(\hat{x}) \equiv 1$.

Step 4 (Existence). The function $\Psi - \kappa$ is continuous and strictly increasing. Evaluating at the dominance boundaries:

$$\Psi(\underline{\theta}) - \kappa = \underline{\theta} + \Delta\bar{H} - \kappa = (\kappa - \Delta) + \Delta\bar{H} - \kappa = -\Delta(1 - \bar{H}) < 0,$$

$$\Psi(\overline{\theta}) - \kappa = \overline{\theta} + \Delta\bar{H} - \kappa = \kappa + \Delta\bar{H} - \kappa = \Delta\bar{H} > 0,$$

since $\Delta > 0$ and $\bar{H} \in (0, 1)$. By the intermediate value theorem, there exists $\theta^* \in (\underline{\theta}, \overline{\theta})$ with $\Psi(\theta^*) = \kappa$. Explicitly, solving $\Psi(\theta^*) = \theta^* + \Delta\bar{H} = \kappa$ gives

$$\theta^* \;=\; \kappa - \Delta\bar{H} \;=\; p + R - c - \Delta\bar{H},$$

which lies strictly inside the coordination region $(\underline{\theta}, \overline{\theta})$. Hence at least one threshold equilibrium with cutoff θ^* exists.

Step 5 (Uniqueness). Strict monotonicity $\Psi' \equiv 1 > 0$ implies Ψ is injective; so $\Psi^{-1}(\{\kappa\})$ is a singleton, namely $\theta^* = \kappa - \Delta\bar{H}$. By Step 1, the extremal cutoffs $\underline{x} \le \bar{x}$ both satisfy (11) (since they are equilibria, and Step 2 shows every equilibrium cutoff solves (11)). Hence $\underline{x} = \bar{x} = \theta^*$. By the sandwich property of iterated deletion, every rationalizable strategy is the threshold strategy with cutoff θ^*: the equilibrium is unique and equal to this threshold strategy. The interior common-knowledge equilibrium of Theorem 1 is eliminated by the global-games refinement; the corner equilibria $n^\star \in \{0, 1\}$ collapse into the threshold rule "stay iff $\theta > \theta^*$" in the small-noise limit.

Step 6 (Realized participation rate). By Lemma 1, the realized rate at the threshold equilibrium is $n(\theta; \theta^*) = 1 - G((\theta^* - \theta)/\varepsilon)$ for any $\varepsilon > 0$. As $\varepsilon \to 0$, the argument $(\theta^* - \theta)/\varepsilon \to +\infty$ if $\theta < \theta^*$ and $\to -\infty$ if $\theta > \theta^*$, so $G((\theta^* - \theta)/\varepsilon) \to \mathbf{1}\{\theta < \theta^*\}$ and $n(\theta; \theta^*) \to \mathbf{1}\{\theta > \theta^*\}$ pointwise. The platform survives ($n = 1$) for $\theta > \theta^*$ and collapses ($n = 0$) for $\theta < \theta^*$. □

Proof of Corollary 1. *Step 1 (extremal threshold strategies).* The dominance regions of Assumption 3 do not involve the prior. The monotone-methods argument of Theorem 2 Step 1 (strategic complementarities and state-monotone payoffs) goes through unchanged: iterated deletion of dominated strategies leaves a largest and smallest threshold strategy with cutoffs $\underline{x} \le \bar{x}$, both seeded by the dominance regions.

Step 2 (posterior at the marginal type). With prior density f_0 on θ and signal $x_i = \theta + \varepsilon\eta_i$ with

$\eta_i \sim G$, the posterior density of θ given $x_i = \hat{x}$ is, by Bayes' rule, $f_0(\theta) \cdot \varepsilon^{-1} g((\hat{x}-\theta)/\varepsilon)/Z_{\hat{x}}$ where $Z_{\hat{x}} = \int f_0(\theta)\varepsilon^{-1} g((\hat{x}-\theta)/\varepsilon)\, d\theta > 0$ by the positivity of f_0 and g. Equivalently, $z \equiv (\hat{x}-\theta)/\varepsilon$ has posterior density $g^{f_0}_{\hat{x}}(z) = f_0(\hat{x}-\varepsilon z)g(z)/Z_{\hat{x}}$. Let $G^{f_0}_{\hat{x}}$ denote the corresponding CDF.

Step 3 (indifference). If all others use cutoff $\hat{x}$, the participation rate at fundamental θ is $n(\theta) = 1 - G((\hat{x}-\theta)/\varepsilon)$ (Lemma 1), so at the marginal type's posterior the participation rate is $1 - G(z)$ with $z \sim g^{f_0}_{\hat{x}}$. Her expected stay payoff is

$$\Psi^{f_0}(\hat{x}) \;=\; \mathbb{E}_{z \sim g^{f_0}_{\hat{x}}}\big[q(\hat{x}-\varepsilon z, 1-G(z))\big],$$

so any threshold equilibrium satisfies $\Psi^{f_0}(\hat{x}) = \kappa$, which is (14).

Step 4 (strict monotonicity of Ψ^{f_0}, existence, and uniqueness). Differentiating under the integral (justified by uniform continuity of f_0 on compact sets and integrability of g),

$$\frac{\partial \Psi^{f_0}(\hat{x})}{\partial \hat{x}} = \mathbb{E}_z[\partial_\theta q] \;+\; \mathbb{E}_z[q \cdot \partial_{\hat{x}} \ln g^{f_0}_{\hat{x}}(z)].$$

The first term equals 1 under (2) (with $\partial_\theta q \equiv 1$) and is bounded away from zero under any q with $\partial_\theta q > 0$. The second term is $O(\varepsilon)$ since $g^{f_0}_{\hat{x}}$'s dependence on $\hat{x}$ enters only through $f_0(\hat{x}-\varepsilon z)$, whose log-derivative in $\hat{x}$ is $\varepsilon \cdot f_0'/f_0$ at first order. For ε sufficiently small (and in particular at every ε small enough that the second term does not dominate the first), Ψ^{f_0} is strictly increasing in $\hat{x}$ on $[\underline{\theta}, \overline{\theta}]$.

Existence. By the dominance regions and continuity of Ψ^{f_0}, $\Psi^{f_0}(\underline{\theta}) < \kappa < \Psi^{f_0}(\overline{\theta})$ for ε small (since $\Psi^{f_0}(\hat{x}) \to \hat{x} + \Delta\bar{H}$ uniformly, and $\underline{\theta} + \Delta\bar{H} < \kappa < \overline{\theta} + \Delta\bar{H}$). By the intermediate value theorem, there exists at least one $\hat{x}^*(\varepsilon, f_0) \in (\underline{\theta}, \overline{\theta})$ with $\Psi^{f_0}(\hat{x}^*) = \kappa$. *Uniqueness.* Strict monotonicity of Ψ^{f_0} implies $\hat{x}^*$ is the unique solution of (14). Both extremal cutoffs solve (14) (by Step 3 applied to $\underline{x}$ and $\bar{x}$), hence $\underline{x} = \bar{x} = \hat{x}^*$: the equilibrium of the global game with proper prior f_0 is unique and equal to the threshold strategy with cutoff $\hat{x}^*$.

Step 5 (small-noise convergence). As $\varepsilon \to 0$, $f_0(\hat{x}-\varepsilon z) \to f_0(\hat{x})$ uniformly on any compact set of z values, so $g^{f_0}_{\hat{x}}(z) \to g(z)$ pointwise. The dominated-convergence theorem gives $\Psi^{f_0}(\hat{x}) \to \hat{x} + \Delta\bar{H}$ (using the closed-form calculation of Step 4 of Theorem 2's proof under the now-uniform belief over n). Therefore $\hat{x}^*(\varepsilon, f_0) \to \theta^* = \kappa - \Delta\bar{H}$, uniformly in any prior f_0 bounded away from zero on the coordination region.

Step 6 (comparative statics). Apply the implicit function theorem to $\Psi^{f_0}(\hat{x}^*; p, R, c, \Delta, \bar{H}) = p + R - c$. The derivatives $\partial_p \Psi^{f_0} = \partial_R \Psi^{f_0} = 0$ and $\partial_c \Psi^{f_0} = 0$ (none of p, R, c enter q), so $\partial\hat{x}^*/\partial p = \partial\hat{x}^*/\partial R = 1/(\partial_{\hat{x}}\Psi^{f_0}) > 0$ and $\partial\hat{x}^*/\partial c = -1/(\partial_{\hat{x}}\Psi^{f_0}) < 0$. For Δ: $\partial_\Delta \Psi^{f_0} = \mathbb{E}_z[h(1-G(z))] > 0$, so $\partial\hat{x}^*/\partial\Delta = -\mathbb{E}_z[h(1-G(z))]/\partial_{\hat{x}}\Psi^{f_0} < 0$. The signs match Corollary 2. □

Proof of Corollary 2. Differentiate $\theta^* = p + R - c - \Delta\bar{H}$. Monotonicity of F gives the direction of $F(\theta^*)$. The panic width $\Delta(1-\bar{H})$ has derivative $1 - \bar{H} > 0$ in Δ (with $\bar{H} \in (0,1)$ fixed by h). □

Proof of Corollary 3. The platform survives at (θ, κ) iff $\theta \geq \theta^* = \kappa - \Delta\bar{H}$, i.e. iff $D = \theta - \theta^* \geq 0$. A fundamental shock $\Delta\theta$ sends $\theta \mapsto \theta - \Delta\theta$, triggering a run iff $\theta - \Delta\theta < \theta^* \Leftrightarrow \Delta\theta > D$. A hurdle shock $\Delta\kappa$ sends $\theta^* \mapsto \theta^* + \Delta\kappa$, triggering a run iff $\theta < \theta^* + \Delta\kappa \Leftrightarrow \Delta\kappa > D$. As $\theta \downarrow \theta^*$, $D \downarrow 0$. $\square$

Proof of Proposition 2. Conjecture a monotone equilibrium in which an agent stays iff her private signal exceeds a cutoff $\hat{x}$. By the aggregation of Lemma 1 with Gaussian noise, the participation rate at state θ is $n(\theta) = \mathbb{P}(x_i \geq \hat{x} \mid \theta) = \Phi\big(\sqrt{\tau_x}\,(\theta - \hat{x})\big)$. The marginal agent (signal $\hat{x}$) has, given the public signal $y \sim N(\theta, \tau_y^{-1})$, the Gaussian posterior $\theta \mid (\hat{x}, y) \sim N\big(\mu, (\tau_x + \tau_y)^{-1}\big)$ with $\mu = (\tau_x \hat{x} + \tau_y y)/(\tau_x + \tau_y)$. Her indifference condition is

$$V(\hat{x}; y) \;\equiv\; \mu + \Delta \int_{\mathbb{R}} h\big(\Phi(\sqrt{\tau_x}(\theta - \hat{x}))\big)\, d\Phi_{\mu,\, \tau_x + \tau_y}(\theta) - \kappa \;=\; 0,$$

which defines the equilibrium cutoff(s) $\hat{x}(y)$; a monotone equilibrium is unique iff $V(\cdot; y)$ has a single zero, i.e. is strictly monotone in $\hat{x}$. Differentiating,

$$\frac{\partial V}{\partial \hat{x}} \;=\; \underbrace{\frac{\tau_x}{\tau_x + \tau_y}}_{\text{belief anchoring } >0} \;+\; \Delta\, \frac{\partial}{\partial \hat{x}}\, \mathbb{E}\big[h(n(\theta)) \mid \hat{x}, y\big].$$

The first term is stabilizing: a higher conjectured cutoff raises the marginal type's posterior mean in proportion to the private weight. The second is the destabilizing coordination feedback: a higher $\hat{x}$ lowers participation ($\partial n/\partial \hat{x} = -\sqrt{\tau_x}\,\phi(\cdot) < 0$), which through the complementarity $\Delta h'$ lowers the stay incentive and rationalizes a still-higher cutoff. To evaluate the feedback explicitly, take h at its steepest, $h' \equiv h'_{\max}$, so locally $\mathbb{E}[h(n(\theta)) \mid \hat{x}, y] = \text{const} + h'_{\max}\, \mathbb{E}[n(\theta) \mid \hat{x}, y]$. With $n(\theta) = \Phi(\sqrt{\tau_x}(\theta - \hat{x}))$ and $\theta \mid (\hat{x}, y) \sim N(\mu, (\tau_x + \tau_y)^{-1})$, the Gaussian identity $\int \Phi(a(\theta - \hat{x}))\, dN(\theta; \mu, s^2) = \Phi\big(a(\mu - \hat{x})/\sqrt{1 + a^2 s^2}\big)$ with $a = \sqrt{\tau_x}$, $s^2 = (\tau_x + \tau_y)^{-1}$, and $\mu - \hat{x} = \tau_y(y - \hat{x})/(\tau_x + \tau_y)$ gives

$$\mathbb{E}[n(\theta) \mid \hat{x}, y] = \Phi\left(\frac{\sqrt{\tau_x}\,\tau_y\,(y - \hat{x})}{\sqrt{(\tau_x + \tau_y)(2\tau_x + \tau_y)}}\right).$$

Differentiating in $\hat{x}$ and maximizing the magnitude over y (worst case $\mu = \hat{x}$, where the standard-normal density attains $1/\sqrt{2\pi}$),

$$\max_y \left| \frac{\partial}{\partial \hat{x}} \mathbb{E}[n(\theta) \mid \hat{x}, y] \right| = \frac{1}{\sqrt{2\pi}} \cdot \frac{\sqrt{\tau_x}\,\tau_y}{\sqrt{(\tau_x + \tau_y)(2\tau_x + \tau_y)}}.$$

Thus $\partial V/\partial \hat{x} > 0$ for every y (so the equilibrium is unique) if and only if the anchoring term dominates this maximum, i.e.

$$\frac{\tau_x}{\tau_x + \tau_y} \;\geq\; \frac{\Delta h'_{\max}}{\sqrt{2\pi}} \cdot \frac{\sqrt{\tau_x}\,\tau_y}{\sqrt{(\tau_x + \tau_y)(2\tau_x + \tau_y)}} \quad \Longleftrightarrow \quad (\Delta h'_{\max})^2 \;\leq\; 2\pi\, \frac{\tau_x(2\tau_x + \tau_y)}{\tau_y^2(\tau_x + \tau_y)},$$

the sharp single-crossing threshold for the linear benchmark. Since $(2\tau_x+\tau_y)/(\tau_x+\tau_y) \in (1,2)$, this is implied by the simpler sufficient condition $\tau_x/\tau_y^2 \geq (\Delta h'_{\max})^2/(2\pi)$, which is (16), the platform analogue of the Morris and Shin (2003) condition (conservative relative to the exact threshold by at most a factor of two, and reported in that canonical form; cf. the public-information analysis of Morris and Shin, 2002). For general h, replacing h' by its supremum $h'_{\max}$ keeps (16) sufficient. When it holds, $V(\cdot; y)$ has the unique zero $\hat{x}(y)$ (equivalently a unique state threshold $\theta^*(y)$); the implicit-function theorem applied to $V = 0$ gives $\partial\hat{x}/\partial y = -(\partial V/\partial y)/(\partial V/\partial\hat{x}) < 0$ (since $\partial V/\partial y = \tau_y/(\tau_x+\tau_y) > 0$ and $\partial V/\partial\hat{x} > 0$), so $\theta^*(y)$ is strictly decreasing in y, and its magnitude exceeds the pure Bayesian weight $\tau_y/(\tau_x+\tau_y)$ because the public signal additionally shifts beliefs about others' actions in a correlated way. When (16) fails, $\partial V/\partial\hat{x} < 0$ over a range, $V(\cdot; y)$ has multiple zeros, and self-fulfilling multiplicity reappears. □

Proof of Lemma 3. The objective in (18) is $J(b) = (1-\pi_R)\beta_I b - \pi_R(\ell_0 + \ell b) - \chi(b)$, with $J'(b) = (1-\pi_R)\beta_I - \pi_R\ell - \chi'(b)$ and $J'' = -\chi'' < 0$ (strict concavity), so the maximizer is unique: $b^* = 0$ if $(1-\pi_R)\beta_I - \pi_R\ell \leq \chi'(0)$, else the interior root of $\chi'(b^*) = (1-\pi_R)\beta_I - \pi_R\ell$. The right side is strictly decreasing in π_R (derivative $-\beta_I - \ell < 0$), and χ' is increasing, so by the implicit-function theorem $b^{*\prime}(\pi_R) = (-\beta_I - \ell)/\chi''(b^*) < 0$ where interior, and b^* is weakly decreasing throughout. Since $c(\cdot)$ and $L(\cdot)$ are increasing, $c(b^*)$ and $L(b^*)$ are weakly decreasing in π_R and co-move (both increasing transforms of b^*); eliminating b, $\mathsf{L}'(\mathsf{c}) = L'(b^*)/c'(b^*) = \ell/c'(b^*) > 0$. □

Proof of Proposition 3. The panic region is $\theta^* - \underline{\theta} = \Delta(1-\bar{H})$ (Eq. 15), with derivative $1-\bar{H} > 0$ in Δ. Holding $q(\theta, 1) = \theta + \Delta$ fixed isolates the slope Δ from the level. The panic-run probability is $F(\theta^*) - F(\underline{\theta}) = F(\kappa - \Delta\bar{H}) - F(\kappa - \Delta)$; its derivative in Δ is $-\bar{H}f(\theta^*) + f(\underline{\theta})$, which is positive whenever the density does not fall too fast across the panic band (in particular for f non-increasing on $[\underline{\theta}, \theta^*]$, or for $\bar{H}$ near 1), and the region width rises unconditionally. By Remark 1, $\Delta = \phi_d + \phi_e + \phi_c$, so the result holds channel-by-channel. □

Proof of Proposition 4. By Corollary 2, $\partial\theta^*/\partial\mathsf{c} = -1$, so $F(\theta^*(\mathsf{c}))$ is decreasing in c: $\epsilon_{F,\mathsf{c}} < 0$. By Lemma 3, $\mathsf{L}'(\mathsf{c}) > 0$: $\epsilon_{\mathsf{L},\mathsf{c}} > 0$. For $\mathcal{E}(\mathsf{c}) = F(\theta^*(\mathsf{c}))\mathsf{L}(\mathsf{c})$, $\frac{d\log\mathcal{E}}{d\log\mathsf{c}} = \epsilon_{F,\mathsf{c}} + \epsilon_{\mathsf{L},\mathsf{c}} = \epsilon_{\mathsf{L},\mathsf{c}} - |\epsilon_{F,\mathsf{c}}|$. Under the stated tail regularity condition, $|\epsilon_{F,\mathsf{c}}| = \mathsf{c}\, f(\theta^*)/F(\theta^*) \to 0$ as $\mathsf{c} \to \infty$, while the severity elasticity remains positive for any integration technology with $\mathsf{L}'(\mathsf{c}) > 0$ and non-vanishing marginal severity. Hence there is $\bar{c}$ with $\mathcal{E}'(\mathsf{c}) > 0$ for $\mathsf{c} > \bar{c}$. □

Proof of Proposition 11. *(i)* The private FOC from (23) is $\Sigma'(e^{\text{priv}})(\Pi - \Pi_{\text{run}}) = \psi'(e^{\text{priv}})$; the social FOC is $\Sigma'(e^{\text{soc}})(\Pi - \Pi_{\text{run}} + X) = \psi'(e^{\text{soc}})$. Define $\Lambda(e; G) \equiv \Sigma'(e)G - \psi'(e)$; the optimum satisfies $\Lambda = 0$. Since $\Sigma' > 0$, Λ is increasing in the gain G, so the root $e(G)$ is increasing in G (as $\partial\Lambda/\partial e = \Sigma''G - \psi'' < 0$ at an interior max). With $X > 0$ the social gain exceeds the private, so $e^{\text{soc}} = e(\Pi - \Pi_{\text{run}} + X) > e(\Pi - \Pi_{\text{run}}) = e^{\text{priv}}$. The gap is increasing in $X = \bar{\delta}(Q_{\text{hi}} - Q_{\text{lo}})\bar{A}$, hence in $\bar{\delta}$. *(ii)* With a rescue B, the private survival benefit becomes $\Pi - \Pi_{\text{run}} - B$, and $e^{\text{priv}} = e(\Pi - \Pi_{\text{run}} - B)$ is decreasing in B since $e(\cdot)$ is increasing in its argument; $\partial e^{\text{priv}}/\partial B < 0$. □

Proof of Proposition 5. Channels (i),(ii): $\partial\theta^*/\partial R = 1 > 0$ and $\partial\theta^*/\partial \mathsf{c} = -1 < 0$ (Cor. 2), both raising $F(\theta^*)$. Channel (iii): the panic region $\Delta(1-\bar{H})$ is increasing in Δ (Eq. 15), and severity inherits $\partial \mathsf{L}/\partial\Delta > 0$ because the destroyed platform-specific capital scales with the complementarity (the ℓ-term in (17) is proportional to platform-specific value, itself increasing in Δ); so $\Delta \downarrow$ shrinks both. Channel (iv): rents Π fall in R by competition; with $\Sigma\Pi$ supermodular in (I,Π) (Section 10.3), $I^{*\prime}(\Pi) > 0$, so $\partial I^*/\partial R = I^{*\prime}(\Pi)\cdot\partial\Pi/\partial R < 0$. For $\mathcal{E} = F(\theta^*)\mathsf{L}$, $d\log\mathcal{E} = (\epsilon_{F,R} + \epsilon_{F,\mathsf{c}}) - |\epsilon_{\mathsf{L},\Delta}\epsilon_{\Delta}|$, so $\mathcal{E}$ falls iff the severity-and-panic reduction dominates. □

Proof of Proposition 7. With $\Xi(\mathcal{I}) = A\mathcal{I}^{\alpha}$ and homogeneous informativeness, $h(n) = \Xi(\bar{\iota}n)/\Xi(\bar{\iota}) = n^{\alpha}$, so $\bar{H} = \int_0^1 n^{\alpha}\,dn = 1/(1+\alpha)$ and $\theta^* = \kappa - \Delta\bar{H} = \kappa - \Delta/(1+\alpha)$. Differentiating at fixed Δ, $\partial\theta^*/\partial\alpha = \Delta/(1+\alpha)^2 > 0$, and F increasing gives $\partial F(\theta^*)/\partial\alpha > 0$. Under selection, Proposition 1(ii) gives $h_{\text{eff}}(n) = n^{\alpha(1+\varsigma)}$ and $\bar{H}_{\text{eff}} = 1/(1+\alpha(1+\varsigma))$; substituting $\gamma_{\text{eff}} = \alpha(1+\varsigma)$ for α reproduces the displayed derivative with $\partial\gamma_{\text{eff}}/\partial\alpha = 1+\varsigma > 0$ and $\partial\gamma_{\text{eff}}/\partial\varsigma = \alpha > 0$. □

Proof of Proposition 14. $\Phi_m > 0$ and its monotonicity in m and ι_H follow from Ξ increasing and strictly concave with $\Xi(0) = 0$: $\Phi_m = \Xi(\mathcal{I}) - \Xi(\mathcal{I} - m\iota_H)$, so $\partial\Phi_m/\partial m = \iota_H\,\Xi'(\mathcal{I} - m\iota_H) > 0$ and $\partial\Phi_m/\partial\iota_H = m\,\Xi'(\mathcal{I} - m\iota_H) > 0$; concentrating informativeness in the footloose group (a larger $\iota_H - \iota_L$ at fixed mean) raises $m\iota_H$ and hence Φ_m. (i) Conditioning on S, $\bar{p} = \mathbb{P}(\theta_k < \theta^* + \Phi_m\,S) = \pi F_1 + (1-\pi)F_0$. (ii) Conditional on S the fundamentals are independent, so $\mathbb{P}(\text{run}_1, \text{run}_2) = \pi F_1^2 + (1-\pi)F_0^2$. Then

$$\text{Cov} = \pi F_1^2 + (1-\pi)F_0^2 - \big(\pi F_1 + (1-\pi)F_0\big)^2 = \pi(1-\pi)(F_1 - F_0)^2,$$

the variance-of-a-two-point-mixture identity. As F is increasing and $\Phi_m \geq 0$, $F_1 - F_0 = F(\theta^* + \Phi_m) - F(\theta^*) \geq 0$ is increasing in Φ_m, hence in m and $\iota_H - \iota_L$; the covariance, being $\pi(1-\pi)(F_1-F_0)^2$, inherits the monotonicity and is strictly positive for $\pi \in (0,1)$, $m > 0$. (iii) Immediate, since $\Phi_0 = 0 \Rightarrow F_1 = F_0 \Rightarrow \text{Cov} = 0$. □

Proof of Corollary 6. Let $\delta_1 = \delta_2 = \frac{1}{2}$ be the firm's dependence shares, so total disruption is

$$X \;=\; \delta_1 \mathbf{1}_{\text{run}_1} + \delta_2 \mathbf{1}_{\text{run}_2} \;=\; \tfrac{1}{2}\big(\mathbf{1}_{\text{run}_1} + \mathbf{1}_{\text{run}_2}\big).$$

Computing the variance,

$$\text{Var}(X) \;=\; \tfrac{1}{4}\big[\,\text{Var}(\mathbf{1}_{\text{run}_1}) + \text{Var}(\mathbf{1}_{\text{run}_2}) + 2\,\text{Cov}(\mathbf{1}_{\text{run}_1}, \mathbf{1}_{\text{run}_2})\big].$$

With symmetric platforms each having marginal run probability $\bar{p} = \mathbb{P}(\text{run}_k)$ and run variance $\text{Var}(\mathbf{1}_{\text{run}_k}) = \bar{p}(1-\bar{p})$,

$$\text{Var}(X) \;=\; \tfrac{1}{4}\big[2\bar{p}(1-\bar{p}) + 2\,\text{Cov}(\mathbf{1}_{\text{run}_1}, \mathbf{1}_{\text{run}_2})\big] \;=\; \tfrac{1}{2}\big[\bar{p}(1-\bar{p}) + \text{Cov}(\mathbf{1}_{\text{run}_1}, \mathbf{1}_{\text{run}_2})\big],$$

which is the form in the corollary statement. Decompose $\partial \operatorname{Var}(X)/\partial m$ into two terms:

$$\frac{\partial \operatorname{Var}(X)}{\partial m} \;=\; \tfrac{1}{2}\frac{\partial[\bar{p}(1-\bar{p})]}{\partial m} \;+\; \tfrac{1}{2}\frac{\partial \operatorname{Cov}(\mathbf{1}_{\text{run}_1}, \mathbf{1}_{\text{run}_2})}{\partial m}.$$

By Proposition 14, the per-platform marginal run probability and the run covariance are, as functions of m,

$$\bar{p}(m) = \pi F_1(m) + (1-\pi)F_0, \qquad \operatorname{Cov}(m) = \pi(1-\pi)\big(F_1(m) - F_0\big)^2,$$

with $F_1(m) = F(\theta^* + \Phi_m)$, $F_0 = F(\theta^*)$, and Φ_m increasing in m with $\Phi_0 = 0$ (Eq. 26). Hence $\bar{p}'(m) = \pi f(\theta^* + \Phi_m)\,\Phi_m' > 0$ and $\operatorname{Cov}'(m) = 2\pi(1-\pi)(F_1 - F_0)\,F_1' \geq 0$, with $\operatorname{Cov}'(0) = 0$. Therefore

$$\frac{d \operatorname{Var}(X)}{dm} = \tfrac{1}{2}\big[(1-2\bar{p})\,\bar{p}'(m) + \operatorname{Cov}'(m)\big].$$

In the rare-runs regime $\bar{p}(m) < \frac{1}{2}$ for all $m \in [0,1]$, so $(1-2\bar{p}) > 0$; both terms are then nonnegative and the first is strictly positive, giving $d\operatorname{Var}(X)/dm > 0$. Thus $\operatorname{Var}(X)$ is *strictly increasing* in m: it rises from the fully diversified level $\operatorname{Var}(X)|_{m=0} = \frac{1}{2}\bar{p}(1-\bar{p})$ (at $m = 0$, Cov $= 0$ and the platforms are independent) toward the comovement-dominated level as $m \to 1$. Contributor multi-homing therefore monotonically erodes the firm's cross-platform diversification; the diversification credited to spreading dependence (Proposition 13) is fully available only at $m = 0$ and is progressively lost as m rises. Outside the rare-runs regime, where $\bar{p}$ rises above $\frac{1}{2}$, the idiosyncratic term $(1-2\bar{p})\bar{p}'$ turns negative and $\operatorname{Var}(X)$ can become non-monotone in m; the economically relevant case is $\bar{p} < \frac{1}{2}$. □

Proof of Lemma 5. Differentiating (28) under the integral sign,

$$\frac{\partial \bar{\Xi}_L}{\partial m} = \int_0^1 \Xi'\big(\iota_L \tfrac{1-m}{2} n\big) \cdot \Big(-\frac{\iota_L n}{2}\Big)\, dn = -\frac{\iota_L}{2}\int_0^1 n\,\Xi'\big(\iota_L \tfrac{1-m}{2} n\big)\, dn,$$

which is strictly negative because $\Xi' > 0$ and $n > 0$ on a positive-measure set. By (28), $\partial \hat{x}_{L,k}/\partial m = -\partial \bar{\Xi}_L/\partial m > 0$. □

Proof of Theorem 3. *Preliminary: Existence and uniqueness of equilibrium in the two-platform global game.* The two-platform global game of Section 12.2 is a Bayesian game with type space $\mathbb{R}^2$ (signals on each platform), action space $\{0,1\}^2$ per agent (stay or exit on each platform), and strategic complementarities both within and across platforms (the within-platform complementarity is inherited from the single-platform game; the cross-platform complementarity arises through multi-homers whose threshold on platform 1 depends on the perceived run probability on platform 2). The payoff structure is supermodular in others' actions and monotone in each platform's signal. Applying the monotone-methods argument of Theorem 2 Step 1 platform-by-platform (with the multi-homer's joint threshold rule treated as a vector cutoff in the lattice of monotone strategies), iterated deletion converges to a largest and smallest nested-cutoff equilibrium with thresholds

$(\underline{x}_L, \underline{x}_H) \leq (\bar{x}_L, \bar{x}_H)$, where the subscripts denote loyalist (L) and multi-homer (H) cutoffs on each platform (suppressing the platform index in the symmetric case). Each cutoff satisfies an indifference condition analogous to (11): the loyalist's indifference depends only on the participation rate of contributors on her platform, while the multi-homer's joint indifference incorporates both platforms' participation. The Laplacian belief of Remark 10 ensures the indifference conditions are linear in the cutoffs, and the resulting linear system has a unique solution by strict diagonal dominance (each platform's own cutoff has unit coefficient in its own indifference condition, while the cross-platform coupling is bounded by the multi-homing share $m < 1$). Hence $\underline{x}_L = \bar{x}_L = \hat{x}_L$ and $\underline{x}_H = \bar{x}_H = \hat{x}_H$, and the two-platform equilibrium is unique. The threshold shift $\hat{x}_L(m) - \hat{x}_L(0) = \Phi_m \geq 0$ is computed in Lemma 5.

With the equilibrium thresholds in hand, the rest of the proof characterizes the cross-platform run covariance.

(i) Positive contagion, global. Positive quadrant dependence gives $F_{12}(a, b) \geq F(a)F(b)$ for all (a, b), with strict inequality on the interior of the support whenever $\rho_\theta > 0$, so

$$\mathrm{Cov}\big(\mathbf{1}_{\theta_1 < \hat{x}_L}, \mathbf{1}_{\theta_2 < \hat{x}_L}\big) = F_{12}(\hat{x}_L, \hat{x}_L) - F(\hat{x}_L)^2 > 0,$$

without restriction on the threshold. Monotonicity in ρ_θ follows from the standard property that, in standard one-parameter families of bivariate distributions (such as the bivariate normal), F_{12}^ρ is non-decreasing in ρ in the PQD/concordance order, so $F_{12}^\rho(a, a)$ is non-decreasing in ρ while $F(a)^2$ is independent of ρ.

(ii) Amplification by multi-homing, in the rare-runs regime. Write $C(m) \equiv F_{12}(\hat{x}_L(m), \hat{x}_L(m)) - F(\hat{x}_L(m))^2$. Using the diagonal symmetry $\partial_1 F_{12}(a, a) = \partial_2 F_{12}(a, a)$ and the chain rule,

$$\frac{dC}{dm} = 2\Big[\partial_1 F_{12}(\hat{x}_L, \hat{x}_L) - F(\hat{x}_L)\, f(\hat{x}_L)\Big] \frac{\partial \hat{x}_L}{\partial m}.$$

Lemma 5 gives $\partial \hat{x}_L / \partial m > 0$, so $\mathrm{sign}(dC/dm) = \mathrm{sign}(B)$ for $B \equiv \partial_1 F_{12}(\hat{x}_L, \hat{x}_L) - F(\hat{x}_L) f(\hat{x}_L)$. Writing the partial as $\partial_1 F_{12}(a, a) = f(a)\, \mathbb{P}(\theta_2 < a \mid \theta_1 = a)$, $B = f(\hat{x}_L)[\mathbb{P}(\theta_2 < \hat{x}_L \mid \theta_1 = \hat{x}_L) - F(\hat{x}_L)]$. For any PQD copula this gives the local characterization $\mathrm{sign}(dC/dm) = \mathrm{sign}(B)$. Under the elliptical-dependence assumption of part (ii), the conditional CDF $a \mapsto \mathbb{P}(\theta_2 < a \mid \theta_1 = a)$ crosses the unconditional $a \mapsto F(a)$ exactly at the marginal median, exceeding it on the lower half ($F(a) < \frac{1}{2}$) and lying below it on the upper half, verified for the Gaussian copula in Corollary 7 and inherited by every elliptical copula through its linear conditional-location property. Hence $B > 0$ iff $F(\hat{x}_L) < \frac{1}{2}$, which is the stated regime, and $dC/dm > 0$ in that regime.

(iii) Reduced-form recovery. Take F_{12} to be the law of $(\theta_1, \theta_2) = (\theta(S), \theta(S))$ with $S \in \{0, 1\}$, $\mathbb{P}(S = 1) = \pi$, and $\theta(s)$ having CDF G_s on each platform conditional on $S = s$. Then $F_1 \equiv G_1(\hat{x}_L|_{\text{out}})$, $F_0 \equiv G_0(\hat{x}_L|_{\text{in}})$, and the bivariate identity reduces to $\pi(1 - \pi)(F_1 - F_0)^2$, the variance of a two-point mixture, matching Proposition 14. □

Proof of Corollary 7. For bivariate normal with correlation $\rho_\theta \in (0,1)$ and standardized marginals, $\mathbb{P}(\theta_2 < a \mid \theta_1 = a) = \Phi\big(a\sqrt{(1-\rho_\theta)/(1+\rho_\theta)}\big)$ by the conditional-normal formula. Define $\beta_{\rho_\theta} \equiv \sqrt{(1-\rho_\theta)/(1+\rho_\theta)} \in (0,1)$. Then

$$B(a) = \phi(a)\big[\Phi(\beta_{\rho_\theta} a) - \Phi(a)\big]$$

in standard-normal notation, with ϕ the marginal density. Since $\beta_{\rho_\theta} < 1$, $\beta_{\rho_\theta} a$ is closer to zero than a; comparing $\Phi(\beta_{\rho_\theta} a)$ to $\Phi(a)$, $\Phi(\beta_{\rho_\theta} a) > \Phi(a)$ iff $a < 0$, with equality at $a = 0$. Hence $B(a) > 0$ iff $a < 0$ iff $F(a) = \Phi(a) < \frac{1}{2}$, and $B(0) = 0$. Since $\partial \hat{x}_L / \partial m > 0$ throughout (Lemma 5), dC/dm inherits the same sign pattern, so Cov is hump-shaped in $\hat{x}_L$ with a unique maximum at the marginal median; and amplification by m is monotone on the entire interval $[0,1)$ within the rare-runs regime. □

Proof of Proposition 6. From (12), $\theta^* = \kappa - \Delta \bar{H}$, so $\partial \theta^* / \partial \bar{H} = -\Delta < 0$; with F increasing, $F(\theta^*)$ is strictly decreasing in $\bar{H}$. Concavity of h (benefits early) raises $\bar{H} = \int_0^1 h$ toward 1, convexity lowers it toward 0, holding $h(0) = 0, h(1) = 1$ fixed (so Δ unchanged). For the family $h(n) = n^\gamma$, $\bar{H} = \int_0^1 n^\gamma \, dn = 1/(1+\gamma)$, whence $\theta^* = \kappa - \Delta/(1+\gamma)$ and $\partial \theta^* / \partial \gamma = \Delta/(1+\gamma)^2 > 0$; run probability $F(\theta^*)$ rises with γ. The limits follow from $\bar{H} \to 1$ as $\gamma \to 0$ and $\bar{H} \to 0$ as $\gamma \to \infty$. □

Proof of Proposition 13. By linearity of covariance and $\tilde{Q} = Q_{\text{hi}} - (Q_{\text{hi}} - Q_{\text{lo}})\mathbf{1}\{\text{run}\}$, $\text{Cov}(M, \tilde{Q}) = -(Q_{\text{hi}} - Q_{\text{lo}}) \, \text{Cov}(M, \mathbf{1}\{\text{run}\}) = -(Q_{\text{hi}} - Q_{\text{lo}})[\text{Cov}(M, \mathbf{1}\{\text{fund}\}) + \text{Cov}(M, \mathbf{1}\{\text{panic}\})]$. Hence $\lambda_{\text{run}} = -(1 + r_f) \, \text{Cov}(M, \tilde{Q}) = \lambda_{\text{fund}} + \lambda_{\text{panic}}$ with the stated form $\lambda_{\text{type}} = (1 + r_f)(Q_{\text{hi}} - Q_{\text{lo}}) \, \text{Cov}(M, \mathbf{1}\{\text{type run}\})$ (the two minus signs cancel, so each λ_{type} loads positively on its run indicator). By Lemma 4 the distance to run is decreasing in m, so each run type, including the panic type, occurs on a high-m (high-M) set; hence $\text{Cov}(M, \mathbf{1}\{\text{panic}\}) > 0$ whenever that set has positive probability, giving $\lambda_{\text{panic}} > 0$. Undiversifiability: for firms j in the ecosystem set $\mathcal{J}$ on the same platform, $\tilde{\pi}_j = \bar{\pi}_j + \delta_j(\tilde{Q} - \bar{Q}) + \tilde{u}_j$ share the *common* factor $\tilde{Q}$; the equal-weighted ecosystem portfolio $\frac{1}{|\mathcal{J}|} \sum_{j \in \mathcal{J}} \tilde{R}_j$ has loading $\frac{1}{|\mathcal{J}|} \sum_{j \in \mathcal{J}} \delta_j$ on $\tilde{Q}$, which does not vanish as $|\mathcal{J}| \to \infty$ (the $\tilde{u}_j$ average out, $\tilde{Q}$ does not). Only loadings on *independent* platform factors $\tilde{Q}_k$ (multi-homing) diversify, by the same averaging applied across k. □

Proof of Proposition 8. *Existence and uniqueness of the fixed point.* The map $T(\mathbf{x}) \equiv \mathbf{A}\mathbf{x} + \mathbf{s}$ is affine on $\mathbb{R}^N_+$. Since $\mathbf{A} \geq \mathbf{0}$ entrywise and $\rho(\mathbf{A}) < 1$, the Neumann series

$$\sum_{k=0}^{\infty} \mathbf{A}^k \;=\; (\mathbf{I} - \mathbf{A})^{-1}$$

converges absolutely, and each $\mathbf{A}^k \geq \mathbf{0}$ entrywise, so $(\mathbf{I} - \mathbf{A})^{-1} \geq \mathbf{0}$ entrywise. Setting $\mathbf{x}^* = T(\mathbf{x}^*)$ gives the unique solution

$$\mathbf{x}^* \;=\; (\mathbf{I} - \mathbf{A})^{-1}\mathbf{s} \;\in\; \mathbb{R}^N_+,$$

since $\mathbf{s} \geq \mathbf{0}$. Existence is proven (the explicit Neumann-series expression constructs the fixed point); uniqueness follows because $\mathbf{I} - \mathbf{A}$ is invertible (its eigenvalues are $1 - \lambda_i(\mathbf{A})$, all nonzero since $\rho(\mathbf{A}) < 1$). Alternatively, T is a contraction on $(\mathbb{R}^N, \|\cdot\|_{\rho(\mathbf{A})})$, the norm induced by the Perron eigenvalue, with Lipschitz constant $\rho(\mathbf{A}) < 1$, and Banach's fixed-point theorem yields existence, uniqueness, and global convergence of iterates $\mathbf{x}^{(k+1)} = T(\mathbf{x}^{(k)})$ from any starting point.

(i) Strict positivity of Δ. $\Delta = \mathbf{w}'\mathbf{x}^* \geq 0$, finite, and > 0 whenever $\mathbf{s} \neq \mathbf{0}$ and $\mathbf{w} > \mathbf{0}$ (since $(\mathbf{I} - \mathbf{A})^{-1}\mathbf{s}$ has all entries strictly positive when $\mathbf{s}$ has a positive component and $\mathbf{A}$ has at least one positive off-diagonal entry connecting that component to the others).

(ii) Comparative statics. Each entry of $(\mathbf{I} - \mathbf{A})^{-1} = \sum_k \mathbf{A}^k$ is a polynomial in the $a_{ij} \geq 0$ with nonnegative coefficients (every monomial in $\prod a_{i_k j_k}$ is a path-product), so $\partial[(\mathbf{I} - \mathbf{A})^{-1}]_{ij}/\partial a_{kl} \geq 0$ for every (i, j, k, l); strict inequality holds where the corresponding loop $i \to \cdots \to k \to l \to \cdots \to j$ exists with positive weight. Likewise $\partial\Delta/\partial s_i = \mathbf{w}'(\mathbf{I} - \mathbf{A})^{-1}\mathbf{e}_i > 0$. The signs match the primitive monotonicities $a_{ij} = \partial f_i/\partial x_j \geq 0$ (Section 9.1).

(iii) Limit as $\rho(\mathbf{A}) \to 1$. Let $\mathbf{A}(t) \to \mathbf{A}$ with $\rho(\mathbf{A}(t)) \uparrow 1$. By Perron–Frobenius, $\mathbf{A}$ has a simple largest eigenvalue $\rho(\mathbf{A})$ with nonnegative right eigenvector $\mathbf{v}$ and left eigenvector $\mathbf{u}$, normalized $\mathbf{u}'\mathbf{v} = 1$. The resolvent admits the Laurent expansion

$$(\mathbf{I} - \mathbf{A}(t))^{-1} \;=\; \frac{\mathbf{v}\mathbf{u}'}{1 - \rho(\mathbf{A}(t))} \;+\; O(1)$$

as $\rho(\mathbf{A}(t)) \to 1$. For $\mathbf{s} > \mathbf{0}$, the inner product $\mathbf{u}'\mathbf{s} > 0$, so $\mathbf{x}^* = (\mathbf{I} - \mathbf{A}(t))^{-1}\mathbf{s}$ has Perron component $\mathbf{v} \cdot (\mathbf{u}'\mathbf{s})/(1 - \rho(\mathbf{A}(t))) \to \infty$ in norm; hence $\Delta(\mathbf{A}(t)) = \mathbf{w}'\mathbf{x}^* \to \infty$. The inheritance by $\theta^* = \kappa - \Delta\bar{H} \to -\infty$ and the panic width $\Delta(1 - \bar{H}) \to \infty$ is immediate. □

Proof of Proposition 9. *Existence and uniqueness of the saturated fixed point.* Let $T(\mathbf{x}) \equiv f(1; \mathbf{x})$. By Assumption 5, T is a continuous self-map of the compact convex order interval $[\mathbf{0}, \bar{\mathbf{x}}]$. Brouwer's theorem gives at least one fixed point. For uniqueness and global stability, take any $\mathbf{x}, \mathbf{y} \in [\mathbf{0}, \bar{\mathbf{x}}]$. The mean-value theorem and the row-sum bound in Assumption 5 imply

$$\|T(\mathbf{x}) - T(\mathbf{y})\|_\infty \leq \left(\max_i \sum_j \sup_{\mathbf{z} \in [\mathbf{0}, \bar{\mathbf{x}}]} \left| \frac{\partial f_i}{\partial x_j}(1; \mathbf{z}) \right| \right) \|\mathbf{x} - \mathbf{y}\|_\infty \leq \bar{a} \|\mathbf{x} - \mathbf{y}\|_\infty,$$

with $\bar{a} < 1$. Hence T is a contraction on $[\mathbf{0}, \bar{\mathbf{x}}]$, and Banach's fixed-point theorem gives a unique fixed point $\mathbf{x}^*$ together with global convergence of the tâtonnement $\mathbf{x}^{(k+1)} = T(\mathbf{x}^{(k)})$ from every initial condition in that interval.

(ii) Spectral-radius bound. At $\mathbf{x}^*$, the Jacobian $\mathbf{A}(\mathbf{x}^*)$ has nonnegative entries and row sums $\sum_j a_{ij}(\mathbf{x}^*) \leq \bar{a} < 1$ by Assumption 5. Perron–Frobenius for nonnegative matrices gives $\rho(\mathbf{A}(\mathbf{x}^*)) \leq \max_i \sum_j a_{ij}(\mathbf{x}^*) \leq \bar{a} < 1$.

(iii) Saturation limits. $\mathbf{x}^* \leq \bar{\mathbf{x}}$ (since $\bar{\mathbf{x}}$ is an upper bound of the order interval) gives $\Delta = \mathcal{Q}(\mathbf{x}^*) \leq \mathbf{w}'\bar{\mathbf{x}} = \Delta_{\max}$. Since each f_i is bounded above by $\bar{x}_i < \infty$, $\partial x_i^*/\partial s_i \to 0$ as $s_i \to \infty$ (the constraint $x_i^* \leq \bar{x}_i$ binds in the limit, and by the envelope theorem applied to the fixed-point

equation, the marginal effect of s_i on x_i^* vanishes), so $\partial\Delta/\partial s_i \to 0$. □

Proof of Proposition 10. (i) At $\Delta = 0$, $\theta^* = \underline{\theta} = \kappa$, so $P_{\text{panic}}(0) = F(\kappa) - F(\kappa) = 0$. As $\Delta \to \infty$, both $\theta^* = \kappa - \Delta\bar{H} \to -\infty$ and $\underline{\theta} = \kappa - \Delta \to -\infty$, so $F(\theta^*) \to 0$ and $F(\underline{\theta}) \to 0$, giving $P_{\text{panic}}(\Delta) \to 0$. For $\Delta > 0$, $\theta^* > \underline{\theta}$ (as $\bar{H} < 1$) and F is increasing, so $P_{\text{panic}}(\Delta) > 0$. A continuous function that is zero at the endpoints and strictly positive inside attains an interior maximum at some $\widehat{\Delta} > 0$. (ii) Differentiating (20), $P'_{\text{panic}}(\Delta) = -\bar{H}f(\kappa - \Delta\bar{H}) + f(\kappa - \Delta) = f(\underline{\theta}) - \bar{H}f(\theta^*)$. For small Δ, both θ^* and $\underline{\theta}$ approach κ, giving $P'_{\text{panic}}(0^+) = f(\kappa)(1 - \bar{H}) > 0$. If f is decreasing in the left tail (eventual unimodality), then for large Δ the point $\underline{\theta} = \kappa - \Delta$ lies deeper in the tail than $\theta^* = \kappa - \Delta\bar{H}$, so $f(\underline{\theta}) < f(\theta^*)$; once $f(\underline{\theta})/f(\theta^*) < \bar{H}$ we have $P'_{\text{panic}}(\Delta) = f(\underline{\theta}) - \bar{H}f(\theta^*) < 0$, which holds for all Δ beyond a finite $\widehat{\Delta}$. (iii) is the economic reading of (ii). □

Proof of Corollary 4. *Part (i): Polynomial decay under fat tails.* Assume $f(x) \sim c|x|^{-1-\alpha_F}$ as $x \to -\infty$ for some $c > 0$ and $\alpha_F > 0$. Integrating, for $x < 0$ with $|x|$ large,

$$F(x) \;=\; \int_{-\infty}^{x} f(t)\,dt \;=\; \int_{|x|}^{\infty} c\,u^{-1-\alpha_F}\,du\,(1 + o(1)) \;=\; \frac{c}{\alpha_F}|x|^{-\alpha_F}(1 + o(1)),$$

where we substituted $t = -u$ for $t < 0$. For $\Delta \to \infty$, $\kappa - \Delta\bar{H} \to -\infty$ and $|\kappa - \Delta\bar{H}| = \Delta\bar{H} - \kappa \sim \Delta\bar{H}$; similarly $|\kappa - \Delta| \sim \Delta$. Therefore

$$F(\kappa - \Delta\bar{H}) \sim \frac{c}{\alpha_F}(\Delta\bar{H})^{-\alpha_F}, \qquad F(\kappa - \Delta) \sim \frac{c}{\alpha_F}\Delta^{-\alpha_F}.$$

Substituting into (20),

$$P_{\text{panic}}(\Delta) \;=\; F(\kappa - \Delta\bar{H}) - F(\kappa - \Delta) \;\sim\; \frac{c}{\alpha_F}\left[(\Delta\bar{H})^{-\alpha_F} - \Delta^{-\alpha_F}\right] \;=\; \frac{c(\bar{H}^{-\alpha_F} - 1)}{\alpha_F}\,\Delta^{-\alpha_F}.$$

Since $\bar{H} \in (0, 1)$ gives $\bar{H}^{-\alpha_F} > 1$, the coefficient is strictly positive. The hump-shape conclusion of part (i) of Proposition 10 follows because $P_{\text{panic}}(0) = 0$, $P_{\text{panic}}(\Delta) > 0$ for $\Delta > 0$, and $\lim_{\Delta\to\infty} P_{\text{panic}}(\Delta) = 0$; the interior maximum exists by continuity. The decay rate is polynomial $\Delta^{-\alpha_F}$ rather than exponential.

Part (ii): Multiple maxima under multimodal F. The first-order condition for critical points of P_{panic} is, from part (ii) of Proposition 10's proof, $P'_{\text{panic}}(\Delta) = 0 \Leftrightarrow f(\kappa - \Delta) = \bar{H}\,f(\kappa - \Delta\bar{H})$. Let $\xi(\Delta) \equiv f(\kappa - \Delta) - \bar{H}\,f(\kappa - \Delta\bar{H})$. When f is multimodal, both terms in ξ inherit the modes of f, one evaluated at the left endpoint of the panic band and one at the right endpoint. As Δ changes, the band can pass over different modes, and the sign of ξ can therefore switch more than once. Hence P_{panic} is non-monotone but need not be unimodal. If the modes are separated by monotone density segments, each switch corresponds to the band entering or leaving one of these modal regions; the largest local maximum is the one for which the band captures the largest density mass. This is the regular multimodal case stated in the corollary. □

Proof of Corollary 5. $\mathcal{C}(\Delta) = P_{\text{panic}}(\Delta) \cdot \Delta(1-\bar{H})\bar{\ell}$ is the product of $P_{\text{panic}}(\Delta)$, Δ, and constants. By Proposition 10, $P_{\text{panic}}(0) = 0$, hence $\mathcal{C}(0) = 0$. For $\Delta > 0$, $P_{\text{panic}} > 0$ and $\Delta > 0$, hence $\mathcal{C}(\Delta) > 0$. For $\Delta \to \infty$, the right endpoint $\kappa - \Delta\bar{H}$ tends to $-\infty$, and $P_{\text{panic}}(\Delta) \leq F(\kappa - \Delta\bar{H})$. Under the Gaussian benchmark, Mills's ratio implies $F(\kappa - \Delta\bar{H}) = O(\Delta^{-1}\exp[-c\Delta^2])$ for some $c > 0$, so $\mathcal{C}(\Delta) \to 0$. More generally, the maintained light-tail condition $\Delta P_{\text{panic}}(\Delta) \to 0$ gives the same conclusion. Thus $\mathcal{C}$ is continuous, vanishes at 0 and at ∞, and is strictly positive in between, so it attains an interior maximum. Non-monotonicity follows. For the asymmetry of the peak relative to P_{panic}'s peak: $\mathcal{C}'(\Delta) = (1-\bar{H})\bar{\ell}[P_{\text{panic}}(\Delta) + \Delta \cdot P'_{\text{panic}}(\Delta)]$; setting this to zero gives $P'_{\text{panic}}(\Delta)/P_{\text{panic}}(\Delta) = -1/\Delta$. In the Gaussian calibration of Figure IA.2, this condition occurs to the right of the zero of $P'_{\text{panic}}(\Delta)$. □

Proof of Lemma 4. From (24), $\theta(\omega) - \theta^*(\omega) = [\theta_0 - \theta_1 \mathsf{m}] - [\kappa_0 + \kappa_1 \mathsf{m} - \Delta\bar{H}] = (\theta_0 - \kappa_0 + \Delta\bar{H}) - (\theta_1 + \kappa_1)\mathsf{m}(\omega)$, strictly decreasing in m since $\theta_1 + \kappa_1 > 0$. Thus there is a cutoff m^* with a run iff $\mathsf{m}(\omega) > \mathsf{m}^*$, i.e. runs occur in high-m states. Since m is comonotone with M and $\tilde{Q} = Q_{\text{hi}}$ off the run set, Q_{lo} on it, $\tilde{Q}$ is a decreasing function of m and hence of M; for any decreasing function and any random variable, $\text{Cov}(M, \tilde{Q}) = \text{Cov}(M, \text{decr}(M)) \leq 0$, strict when the run set has interior probability. □

Proof of Proposition 12. From $P_j = \mathbb{E}[M\tilde{\pi}_j]$, $\tilde{R}_j = \tilde{\pi}_j / P_j$, the Euler equation $\mathbb{E}[M\tilde{R}_j] = 1$ and $\mathbb{E}[M] = 1/(1+r_f)$ give $\mathbb{E}[\tilde{R}^e_j] = -(1+r_f)\,\text{Cov}(M, \tilde{R}_j)$. Substituting $\tilde{\pi}_j = \bar{\pi}_j + \delta_j(\tilde{Q} - \bar{Q}) + \tilde{u}_j$ with $\mathbb{E}[\tilde{u}_j \mid M] = 0$ (so $\text{Cov}(M, \tilde{u}_j) = 0$) and $\bar{\pi}_j$ constant yields $\text{Cov}(M, \tilde{R}_j) = \delta_j \,\text{Cov}(M, \tilde{Q})/P_j$, hence $\mathbb{E}[\tilde{R}^e_j] = \beta_j \lambda_{\text{run}}$ with common factor price $\lambda_{\text{run}} = -(1+r_f)\,\text{Cov}(M, \tilde{Q})$ and firm-specific loading $\beta_j = \delta_j / P_j$. By Lemma 4, $\text{Cov}(M, \tilde{Q}) < 0$, so $\lambda_{\text{run}} > 0$ (claim i). For multi-homing, $\tilde{\pi}_j = \bar{\pi}_j + \sum_k \delta_{jk}(\tilde{Q}_k - \bar{Q}_k) + \tilde{u}_j$ gives loading $\delta_{jk} \leq \delta_j$ on each factor (claim ii). Low δ_j implies low premium and low sensitivity (claim iii). □

Proof of Proposition IA.1. Equations (IA.2)–(IA.3) are the first-passage law of arithmetic Brownian motion (IA.1) to the lower barrier at distance D_0, obtained by the reflection principle after a Girsanov change of measure removing the drift (Harrison, 1985): for driftless $\tilde{W}$, $\mathbb{P}(\min_{s \leq t} \tilde{W}_s \leq -D_0) = 2\Phi(-D_0/(\sigma\sqrt{t}))$, and the Cameron–Martin density reweights by $\exp(-\frac{\mu}{\sigma^2}(\theta_t - \theta_0) - \frac{\mu^2}{2\sigma^2}t)$, yielding (IA.2) after collecting terms; differentiating $-dS/dt$ gives (IA.3). The lifetime probability is the $t \to \infty$ limit of $1 - S(t)$: $S(\infty) = 0$ for $\mu \leq 0$ (the process is recurrent and hits the barrier almost surely) and $S(\infty) = 1 - e^{-2\mu D_0/\sigma^2}$ for $\mu > 0$ (transience), the standard dichotomy. Monotonicity of ζ in D_0 is immediate from (IA.3). □

Proof of Proposition IA.2. *Spread.* Equating discounted legs gives

$$\mathsf{L} \int_0^T e^{-r_f s} f^{\mathbb{Q}}_\tau(s)\, ds = c \int_0^T e^{-r_f s} S^{\mathbb{Q}}(s)\, ds.$$

Solving for c gives (IA.4); the $T \to 0$ limit uses $\lim_{s\to 0} f_\tau^{\mathbb{Q}}(s)/S^{\mathbb{Q}}(s) = \zeta^{\mathbb{Q}}(0^+)$ and l'Hôpital on the discounted integrals. *Equity.* Let $x = D_0 > 0$ and $V(x) = \mathbb{E}_x^{\mathbb{Q}}[\int_0^\tau e^{-r_f s}\pi\, ds] = \frac{\pi}{r_f}(1-\mathbb{E}_x^{\mathbb{Q}}[e^{-r_f\tau}])$. Let $g(x) = \mathbb{E}_x^{\mathbb{Q}}[e^{-r_f\tau}]$. By the Feynman–Kac/optional-sampling argument, g solves the ODE $\frac{1}{2}\sigma^2 g'' + \mu^{\mathbb{Q}} g' - r_f g = 0$ on $x > 0$, with boundary conditions $g(0) = 1$ (run is immediate when $D_0 = 0$) and g bounded as $x \to \infty$ (the run recedes, $\tau \to \infty$, $e^{-r_f\tau} \to 0$, so $g(\infty) = 0$). The characteristic equation $\frac{1}{2}\sigma^2 z^2 + \mu^{\mathbb{Q}} z - r_f = 0$ has roots $z_\pm = (-\mu^{\mathbb{Q}} \pm \sqrt{(\mu^{\mathbb{Q}})^2 + 2r_f\sigma^2})/\sigma^2$ with $z_- < 0 < z_+$. Boundedness at ∞ rules out $e^{z_+ x}$, so $g(x) = Ae^{z_- x}$; $g(0) = 1$ gives $A = 1$, hence $g(x) = e^{z_- x} = e^{-\beta x}$ with $\beta \equiv -z_- = (\mu^{\mathbb{Q}} + \sqrt{(\mu^{\mathbb{Q}})^2 + 2r_f\sigma^2})/\sigma^2 > 0$. Therefore $V(x) = \frac{\pi}{r_f}(1 - e^{-\beta x})$, which is (IA.5). Differentiation gives $\partial V/\partial x = (\pi\beta/r_f)e^{-\beta x} > 0$, maximal at $x = 0$; limits follow from $\beta > 0$. □

Proof of Proposition 15. For $\theta \in [\underline{\theta}, \theta^*)$ the equilibrium of Theorem 2 selects collapse ($\theta < \theta^*$). But $\theta \geq \underline{\theta} = \kappa - \Delta$ means $q(\theta, 1) = \theta + \Delta \geq \kappa$, so at full participation each participant's stay payoff $q(\theta, 1) - p \geq R - c$ weakly exceeds the exit payoff, and downstream firms obtain $Q_{\text{hi}} > Q_{\text{lo}}$ and the platform $\Pi > \Pi_{\text{run}}$. Hence full participation Pareto-dominates the run. The welfare difference is the sum of the per-participant gains $\int_0^1 [q(\theta, 1) - \kappa]\, di \geq 0$, the platform's $\Pi - \Pi_{\text{run}} > 0$, and downstream $X > 0$ (Eq. 21); it is increasing in $\bar{\delta}$ through X and in lock-in through severity (Section 7). □

Proof of Proposition IA.3. Write $g \equiv (\theta^*(y) - \mu)/\sigma_y = (\kappa - \tau - \Delta\bar{H} - \mu)/\sigma_y$, so $S(\theta^*; y) = 1 - \Phi(g)$. Since $\partial\theta^*/\partial\tau = -1$, $\partial S/\partial\tau = \phi(g)/\sigma_y$, and differentiating (IA.6) in τ gives the first-order condition

$$\Pi\, \phi(g)/\sigma_y = \kappa_\tau \tau. \tag{FOC}$$

Existence. The marginal benefit $\Pi\phi(g)/\sigma_y$ is bounded above by $\Pi/(\sigma_y\sqrt{2\pi})$ and strictly positive at $\tau = 0$, while the marginal cost $\kappa_\tau\tau$ is linear and unbounded; the two cross, so an interior $\tau^*(y) > 0$ exists. It is the unique maximizer when the cost is convex enough that marginal cost cuts marginal benefit from below, i.e. $\mathcal{F}_\tau > 0$ below, which holds whenever $\kappa_\tau\sigma_y^2 > \Pi \sup_g |g|\phi(g) = \Pi/\sqrt{2\pi e}$ (the supremum is attained at $|g| = 1$); assume this throughout. *Single-peakedness in μ.* Write the FOC as $\mathcal{F}(\tau, \mu) \equiv \kappa_\tau\sigma_y\tau - \Pi\phi(g) = 0$. Using $\phi'(g) = -g\phi(g)$ and $\partial g/\partial\tau = \partial g/\partial\mu = -1/\sigma_y$,

$$\mathcal{F}_\tau = \kappa_\tau\sigma_y - \tfrac{\Pi}{\sigma_y} g\, \phi(g) > 0 \;\; \text{(the second-order condition)}, \qquad \mathcal{F}_\mu = -\tfrac{\Pi}{\sigma_y} g\, \phi(g),$$

so by the implicit function theorem

$$\frac{d\tau^*}{d\mu} = -\frac{\mathcal{F}_\mu}{\mathcal{F}_\tau} = \frac{(\Pi/\sigma_y)\, g\, \phi(g)}{\mathcal{F}_\tau}, \qquad \operatorname{sign}\frac{d\tau^*}{d\mu} = \operatorname{sign}(g).$$

Hence τ^* strictly increases in μ while $g > 0$ ($\mu < \kappa - \tau^* - \Delta\bar{H}$) and strictly decreases while $g < 0$ ($\mu > \kappa - \tau^* - \Delta\bar{H}$): $\tau^*(\mu)$ is strictly single-peaked, with its unique maximum at $g = 0$, i.e. where the posterior mean coincides with the post-subsidy threshold $\mu = \kappa - \tau^* - \Delta\bar{H}$. As $|\mu| \to \infty$, $|g| \to \infty$ and $\phi(g) \to 0$, so the FOC forces $\tau^*(y) \to 0$. *Precision.* At $\tau_y = 0$ the signal is uninformative and the optimal subsidy is state-independent. A marginal increase in precision lets

the platform condition τ on y; since expected value is the expectation of a pointwise maximum, $W(\tau_y) = \mathbb{E}_y[\max_\tau\{\Pi S(\theta^*; y) - \frac{\kappa_\tau}{2}\tau^2\}]$ is convex in the informativeness of y and weakly exceeds the constant-policy value, strictly so whenever θ^* lies in the interior of the support of $\mu(y)$. Hence $dW/d\tau_y > 0$ at $\tau_y = 0$. □

Proof of Proposition IA.4. The CRRA Euler condition $M(\omega) = \beta U'(C(\omega))/U'(\bar{C})$ with $U'(C) = C^{-\gamma}$ gives $M(\omega) = \beta(1 - X(\omega)/\bar{C})^{-\gamma}$, which is (IA.7). In run states $X = X_{\text{run}} = \bar{\delta}(Q_{\text{hi}} - Q_{\text{lo}})\bar{A} > 0$ and in survival states $X = 0$, yielding the run/survive wedge. For the panic-premium scaling, a downstream firm j has cash flow A_jQ with $Q = Q_{\text{hi}} - (Q_{\text{hi}} - Q_{\text{lo}})\mathbf{1}_{\text{run}}$. The risk premium is $\lambda_j = -\operatorname{Cov}(M, A_jQ)/\mathbb{E}[M] = A_j(Q_{\text{hi}} - Q_{\text{lo}})\operatorname{Cov}(M, \mathbf{1}_{\text{run}})/\mathbb{E}[M]$. Linearizing the SDF for small $X/\bar{C}$, $M \approx \beta(1 + \gamma X/\bar{C})$, gives $\operatorname{Cov}(M, \mathbf{1}_{\text{run}}) = \beta\gamma X_{\text{run}}\pi_{\text{run}}(1 - \pi_{\text{run}})/\bar{C}$ and $\mathbb{E}[M] \approx \beta$. Hence $\lambda_j \approx A_j(Q_{\text{hi}} - Q_{\text{lo}})\gamma X_{\text{run}}\pi_{\text{run}}(1 - \pi_{\text{run}})/\bar{C}$. Decomposing $\pi_{\text{run}} = \pi_{\text{fund}} + \pi_{\text{panic}}$, the panic component is the contribution of π_{panic}: $\lambda_j^{\text{panic}} \approx A_j(Q_{\text{hi}} - Q_{\text{lo}})\gamma X_{\text{run}}\pi_{\text{panic}}(1 - \pi_{\text{run}})/\bar{C}$. With $X_{\text{run}} = \bar{\delta}(Q_{\text{hi}} - Q_{\text{lo}})\bar{A}$ this yields (IA.8), proportional to $\gamma\bar{\delta}\pi_{\text{panic}}(1 - \pi_{\text{run}})$ with proportionality factor $A_j(Q_{\text{hi}} - Q_{\text{lo}})^2\bar{A}/\bar{C}$. □

Proof of Theorem IA.1. *Preliminary: Existence and uniqueness of the N-platform equilibrium.* The N-platform global game has type space $\mathbb{R}^N$, action space $\{0, 1\}^N$ per agent (stay or exit on each platform), and a multi-homing matrix $\mathbf{M} = [m_{ij}]$ encoding cross-platform participation. The payoff structure is supermodular within each platform and pairwise across platforms via the multi-homing channel. The monotone-methods argument of Theorem 3 extends to N platforms: iterated deletion yields a largest and smallest vector of cutoffs $(\bar{x}_{L,k}, \bar{x}_{H,k})_{k=1}^N$ and $(\underline{x}_{L,k}, \underline{x}_{H,k})_{k=1}^N$, with each cutoff satisfying a Laplacian indifference condition. The resulting $2N$-dimensional linear system has a strictly diagonally dominant coefficient matrix (each cutoff has unit coefficient in its own indifference condition; the cross-platform couplings are bounded by $\sum_j m_{ij} < 1$ for the multi-homers, ensured by $\mathbf{M}$'s row sums), so the system has a unique solution and the equilibrium is unique with cutoff vector $\hat{\mathbf{x}}_L = (\hat{x}_{L,1}, \ldots, \hat{x}_{L,N})$ for loyalists and analogously for multi-homers. The threshold-shift Lemma 5 extends pair-wise, giving $\partial\hat{x}_{L,k}/\partial m_{kl} > 0$ for each $k \neq l$.

With equilibrium thresholds in hand:

(i) For each pair (k, l), the marginal full-run event has $\operatorname{Cov}(\mathbf{1}_k, \mathbf{1}_l) = F_{kl}(\hat{x}_{L,k}, \hat{x}_{L,l}) - F_k(\hat{x}_{L,k})F_l(\hat{x}_{L,l})$, where F_{kl} is the bivariate marginal of (θ_k, θ_l). PQD of $\boldsymbol{\theta}$ implies PQD of each pair, hence each off-diagonal entry is non-negative; strict positivity follows from $(\Sigma_\theta)_{kl} > 0$ in the interior of the support. (ii) The bivariate-derivative argument of Theorem 3 applies pair-wise: $\partial \operatorname{Cov}_{kl}/\partial m_{kl} = [\partial_1 F_{kl}(\hat{x}_{L,k}, \hat{x}_{L,l}) - F_k(\hat{x}_{L,k})f_k(\hat{x}_{L,k})] \cdot (\partial\hat{x}_{L,k}/\partial m_{kl}) > 0$ in the rare-runs regime by the same conditional-versus-unconditional CDF comparison as in the proof of Theorem 3 (invoking that theorem's elliptical-dependence condition); symmetric in m_{lk}. Likewise, $\partial \operatorname{Cov}_{kl}/\partial(\Sigma_\theta)_{kl} > 0$ follows from the standard monotonicity of bivariate quadrant probability in the correlation parameter. (iii) Decompose $\boldsymbol{\Sigma}_{\text{run}} = \boldsymbol{\Sigma}^{(F)} + \boldsymbol{\Sigma}^{(M)}$, where $\boldsymbol{\Sigma}^{(F)}$ collects the contribution from fundamental correlations (the part with $m_{ij} = 0$ for $i \neq j$) and $\boldsymbol{\Sigma}^{(M)}$ the remainder. By a standard linearization in m, the off-diagonal entries of $\boldsymbol{\Sigma}^{(M)}$ are proportional to $m_{kl} + m_{lk}$ in the rare-runs

limit, so $\mathbf{\Sigma}^{(M)}$ has the same leading eigenvector as $(\mathbf{M}+\mathbf{M}^\top)/2$. Suppose $\|\mathbf{\Sigma}^{(M)}\| > \|\mathbf{\Sigma}^{(F)}\|$, and let $\lambda_1(\mathbf{\Sigma}^{(M)}) > \lambda_2(\mathbf{\Sigma}^{(M)})$ be the top two eigenvalues with gap $g \equiv \lambda_1 - \lambda_2$. By Weyl's inequality, the eigenvalues of $\mathbf{\Sigma}_{\text{run}} = \mathbf{\Sigma}^{(M)} + \mathbf{\Sigma}^{(F)}$ differ from those of $\mathbf{\Sigma}^{(M)}$ by at most $\|\mathbf{\Sigma}^{(F)}\|$. By the Davis-Kahan $\sin\Theta$ theorem, the angular distance between the leading eigenvectors of $\mathbf{\Sigma}_{\text{run}}$ and $\mathbf{\Sigma}^{(M)}$ is at most $\|\mathbf{\Sigma}^{(F)}\|/g$. Hence under multi-homing dominance ($\|\mathbf{\Sigma}^{(M)}\| > \|\mathbf{\Sigma}^{(F)}\|$, g bounded away from zero) the leading eigenvector of $\mathbf{\Sigma}_{\text{run}}$ approximates that of $(\mathbf{M}+\mathbf{M}^\top)/2$ with controllable error. □

Proof of Proposition IA.5. *(a) Correlated signals.* Conditional on the common draw ζ, signals are $x_i = (\theta + \varepsilon\rho\zeta) + \varepsilon\eta_i$ with i.i.d. idiosyncratic noise $\varepsilon\eta_i$: the baseline global game with state displaced by $\varepsilon\rho\zeta$ and idiosyncratic precision $\tau_x = \varepsilon^{-2}$. The displacement is common, so it cancels in signal differences $x_i - x_j = \varepsilon(\eta_i - \eta_j)$, and the marginal agent's belief about others' signals (hence about the participation rate) is governed entirely by the idiosyncratic noise. Theorem 2 with the exact Laplacian belief of Lemma 2 then gives a unique conditional threshold $\theta^*(\zeta) = \kappa - \Delta\bar{H} + O(\varepsilon)$, the remainder absorbing the displacement $\varepsilon\rho\zeta$. As $\varepsilon \to 0$ the displacement vanishes and the unconditional threshold concentrates at $\kappa - \Delta\bar{H}$ for every fixed ρ. If instead agents observe a public signal of θ whose precision does not vanish, the marginal agent's posterior loads on it and the destabilizing coordination feedback can dominate; that is the configuration of Proposition 2, where uniqueness requires (16).

(b) Non-monotone ι. The learning map is $h_{\text{eff}}(n) = g(I(n))$ with g increasing (the bounded-returns learning technology) and $I(n) = \int_{1-n}^{1} \iota(u)\,du$. For any integrable $\iota \geq 0$, I is absolutely continuous with $I'(n) = \iota(1-n) \geq 0$, so I is nondecreasing and $h_{\text{eff}} = g \circ I$ is nondecreasing and bounded above by $g\big(\int_0^1 \iota\big)$; Assumption 2 holds irrespective of the monotonicity of ι. Theorem 2 then gives uniqueness and Lemma 1 gives $\theta^* = \kappa - \Delta \int_0^1 h_{\text{eff}}(n)\,dn = \kappa - \Delta\bar{H}_{\text{eff}}$. The composition effect on fragility operates through the curvature $h''_{\text{eff}} = g''(I')^2 + g'\,I''$ with $I''(n) = -\iota'(1-n)$, so its sign at the marginal quantile is governed by $\iota'(1-n)$ over $[1-n, 1]$ (the local, not global, monotonicity of ι) which is precisely the content of Proposition 6. □