Why Venture Later? Incentives, Learning, and Industry Allocation in VC Funds*

Ehsan Mahdikhani†

This Draft: October 2025

You can find the latest version of the paper here.

Abstract

Venture funds enter new high-tech industries with delay. I develop and estimate a structural model with two frictions that generates a socially inefficient delay to enter. Inside deals, managers need credible early evidence from entrepreneurs about viability of new ventures, but because monitoring is costly they do too little. Between investors and managers, LPs cannot verify or directly reward that monitoring, so incentives under-provide it. Estimating the model on matched pairs of funds shows the first friction reduces the share of capital allocated to new sectors, while the second keeps monitoring too low and limits the scale of those investments. Combined frictions imply average welfare losses of nearly $40B per year (around 3% of VC and 12% of new-sector capital). I examine the effectiveness of policy tools, such as exploration bonus for managers, temporary public risk-sharing, and lower-cost funding for new-sector deals, in raising early exploration and investment scale.

Keywords: Exploration–Exploitation Trade-off, Venture Capital, Moral Hazard, Industry Allocation, Dynamic Agency Model

*I am grateful to Per Strömberg, Ramin Baghai, Alvin Chen, Jan Starmans, Mehran Ebrahimian, Adam Winegar (discussant), Aineas Mallios(discussant), and participants at the Stockholm School of Economics Brown Bag Seminar, 14th National PhD Workshop in Finance at SSE, the NFN PhD Workshop at Copenhagen Business School, and EFA DT 2025 at Skema Business School for their valuable feedback and suggestions. This paper will also be presented at the AFA 2026 PhD poster session. This paper won the Ola Bengtsson Award for the Best PhD Paper in 14th National PhD Workshop in Finance at SSE.

†Stockholm School of Economics, E-mail: ehsan.mahdikhani@hhs.se

1 Introduction

Venture capital funds face a persistent timing problem when new sectors emerge. Early capital can accelerate experimentation, standard setting, and diffusion; late capital tends to arrive after private rents have compressed, and after much of the social value from early learning has already been created. A simple glance at the data makes the puzzle concrete. First, the pooled distribution of time to first entry into a sector shows that the median fund takes roughly five and a half years to enter after a sector's "birth" (Figure 1). Second, sector-time profiles show that both per-hit payoffs and failure rates decline as a sector matures, so that the risk-adjusted private return peaks early, long before most funds actually enter (Figure 2). Late entry therefore does not appear to be compensation for higher risk; rather, risk-adjusted opportunities are better earlier, yet funds arrive later. The upshot is slower experimentation and diffusion than would be socially optimal, along with foregone private surplus.

This paper investigates the mechanisms behind that delay and quantifies its consequences. I build and estimate a tractable two-period framework in which a general partner (GP) raises two consecutive funds. The first fund chooses how much to explore a new sector and how intensively to monitor that exploration. These choices determine the quality of information generated by early experiments. The second fund chooses how much of its portfolio to tilt toward the new sector and how much to scale in dollars. A central feature of the environment is that fund size in the second period is not fixed: it is set by the market each vintage and reacts to the returns that funds are expected to generate. This separation between portfolio composition (what percentage of the second fund goes to the new sector) and scale (how many dollars the market will actually finance in that vintage) is essential both for matching the data and for identifying where frictions bite.

Two incentive frictions distort these choices on different links of the contracting chain. On the GP–entrepreneur link, insufficient monitoring in the first fund destroys the information value of early experimentation; the signal is noisy, and the mapping from early exploration to later portfolio reallocation is attenuated. I refer to this as an information or monitoring wedge that

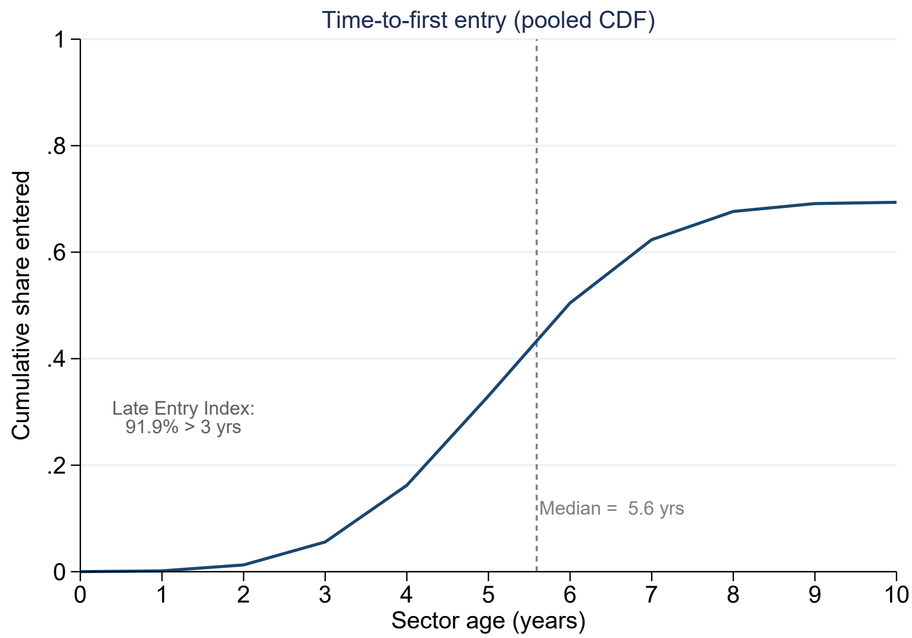


Figure 1. Timing of first venture entries. The empirical CDF pools sectors and shows that a large share of first entries occur late in the observation window, indicating delayed initial investment.

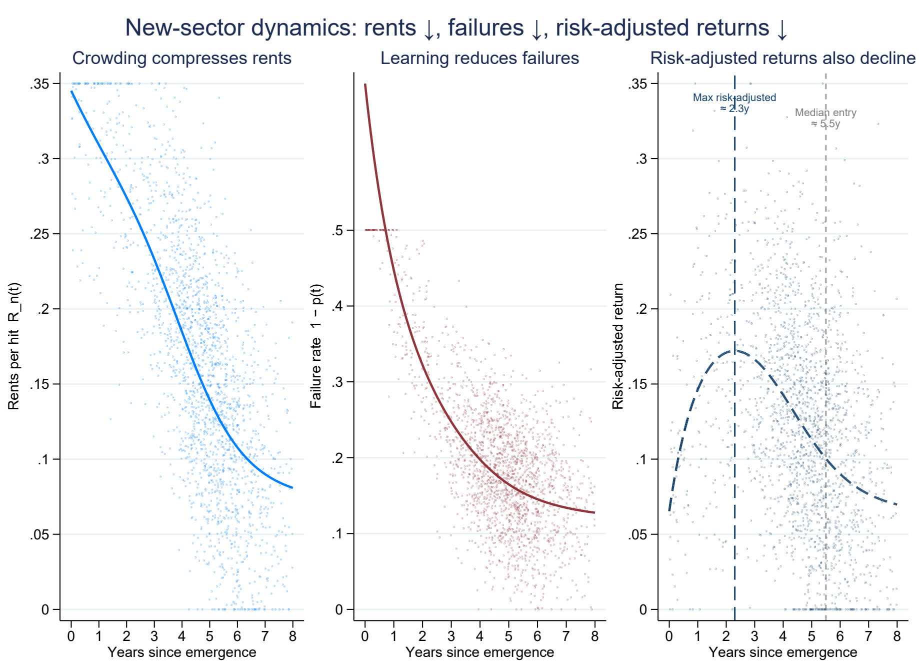


Figure 2. Risk-adjusted performance around first entry. A standard risk-adjusted return proxy does not line up with the timing pattern in Figure 1, suggesting that delayed first entry is not explained by risk premia.

primarily depresses composition. On the LP–GP link, dollars for the second fund are set by contemporaneous financing conditions and a convex scale cost; in tight vintages, even attractive projects do not scale in proportion to their expected returns. I refer to this as a pricing or scale wedge that primarily rations dollars. A third block captures diffusion explicitly by separating knowledge, which raises technical success through learning, from adoption, which compresses rents as more investors crowd into the sector. This split is necessary to reconcile the stylized fact in Figure 2: success improves with time, while per-hit payoffs fall, and the peak in risk-adjusted returns occurs early.

The model is estimated by simulated method of moments on a matched fund sample from 1995 to 2023 that combines PitchBook deal- and fund-level data with sector classifications and exit outcomes. The reduced-form evidence that motivates and disciplines the model is straightforward. First, pooled CDFs and hazards of first entry confirm that early entry mass is small and the hazard rises later; the median entry time is about 5.6 years. Second, binned relationships show that more exploration and stronger monitoring in the first fund predict higher new-sector share and larger dollar scale in the second fund, but the pass-through into composition is weaker when first-fund information is noisy, and the pass-through into dollars is weaker in tight vintages. Third, in sector time, higher exploration raises subsequent success rates (consistent with knowledge accumulation), while broad switching is followed by lower per-deal payoffs (consistent with adoption and rent compression). The structural mapping is deliberately minimalist. The second fund's portfolio tilt responds to the first fund's exploration and monitoring as well as to the two diffusion states; the second fund's dollars respond to expected returns and vintage conditions. The GP–entrepreneur friction is identified by how the composition response flattens when first-fund information is weak. The LP–GP friction is identified by how scale is rationed, holding composition fixed, in tight vintages. The knowledge–adoption split is identified by the divergent movements in success versus payoffs over sector time.

Three results emerge. First, the information wedge on the GP–entrepreneur link primarily depresses composition by killing early signals; observed reallocation in the second fund is too muted relative to the amount of exploration undertaken in the first fund when monitoring is weak. Second, the pricing wedge on the LP–GP link primarily rations dollars; even when funds want to tilt toward the new sector, the market does not always scale them up proportionally in tight years. Third, the wedges complement each other dynamically: better information raises the payoff to scaling dollars in the second fund, and more elastic dollars raise the payoff to producing better information in the first fund. When calibrated to the fraction of venture capital actually exposed to new sectors and integrated over the 1995–2023 period, the joint effect of these frictions implies a welfare loss on the order of tens of billions. In the baseline benchmark used throughout the paper, removing both wedges recovers roughly $37 billion over the sample window. This benchmark deliberately scales welfare to the new-sector-exposed share of committed capital rather than total industry assets under management, and it anchors the "early window" using the empirical entry CDF rather than an arbitrary cutoff.

The findings speak to and connect several literatures. Work on venture cycles documents hot and cold markets in which risk tolerance, capital inflows, and pricing vary over time, shaping experimentation and failure rates (e.g., Gompers and Lerner (2000); Kaplan and Schoar (2005); Gompers et al. (2008); Nanda and RhodesKropf (2013, 2020)). That literature shows clearly that financing conditions matter, but it does not separate what funds choose to allocate across sectors from how many dollars the market will finance, nor does it quantify the welfare cost of delayed entry relative to the sector–time peak in risk-adjusted returns. By making second-fund scale market-set each vintage, the present framework nests their core insight and delivers clean separation between composition and scale in the data.

A second line of work emphasizes contracting, staging, and the role of monitoring in innovation finance, primarily on the GP–entrepreneur link (e.g., Sahlman (1990); Kaplan and Strömberg (2003, 2004); Cornelli and Yosha (2003); Hellmann (1998); Bergemann and Hege (1998, 2005); Manso (2011)). Those papers explain why early projects require governance and why incentives need to tolerate failure to motivate exploration. They typically study a single layer of agency at a time. The present paper layers the GP–entrepreneur monitoring problem with the LP–GP scaling problem in one dynamic environment and shows that the two wedges act on different margins: a monitoring shortfall mutes composition pass-through, while tight financing conditions ration dollars, together generating delayed entry even when risk-adjusted returns peak early.

A third strand documents learning and information diffusion through networks, syndication, and cohort effects (e.g., Sorenson and Stuart (2001); Hochberg et al. (2007, 2010); Gompers et al. (2009); Cabolis et al. (2020)). The evidence is consistent with a diffusion interpretation: success improves with time as knowledge accumulates, but rents per hit fall with adoption. The model incorporates this split and ties it to portfolio reallocation and market-set dollars, helping reconcile rising success with falling private margins in sector time.

Finally, the paper links to real-options and herding ideas as well as pricing in private markets. Real-options and informational-cascade theories illuminate why waiting can be optimal or contagious (Dixit and Pindyck (1994); Banerjee (1992); Bikhchandani et al. (1992); Lerner (1995)); the approach here distinguishes those forces from incentive-driven excess delay and quantifies the welfare loss. On pricing, buyout and asset-pricing work shows how outside conditions affect leverage and valuations (Axelson et al. (2013); Pastor and Veronesi (2009)); I embed a simple vintage-clearing rule so prices set scale while composition is determined by incentives.

The empirical strategy follows a simple sequence. I first document the two motivating facts using pooled entry distributions and sector-time risk–return profiles: the median entry is late, and the risk-adjusted private return peaks early. I then assemble fund–deal data at the fund sequence, sector, and vintage level to construct moments that discipline composition, scale, and diffusion. Estimation proceeds by simulated method of moments; I report standard goodness-of-fit and curvature diagnostics. With the estimated parameters in hand, I quantify the contributions of the two wedges to the observed gaps relative to an efficient benchmark, and I study policy counterfactuals that map directly to observable levers. An exploration-linked carry in the first fund raises the payoff to producing information and steepens the mapping from first-fund exploration to second-fund reallocation. A state-contingent reduction in the effective cost of second-fund capital in tight vintages allows dollars to scale when they should.

Implemented jointly, these levers deliver gains that exceed the sum of their parts because the wedges are complementary.

The rest of the paper is organized as follows. Section 2 describes the data and summary statistics. Section 3 describes the framework, with emphasis on the separation between portfolio composition and market-set scale and on the distinction between knowledge and adoption in diffusion. Section 4 discusses identification and simulated method of moments estimation. Section 6 documents the facts in Figures 1 and 2 and presents additional reduced forms on composition pass-through, scale rationing, and diffusion signatures. Section 7 reports the parameter estimates, the decomposition of composition and scale wedges, and the baseline welfare results, including the aggregate loss over the sample period. Section 9 studies policy counterfactuals and discusses implementation. Section 10 concludes.

2 Data, Facts, and Institutional Features

Sources and construction. I use PitchBook venture-capital deal-, exit-, and fund-level data. I first aggregate each fund's investments by PitchBook industry and define a fund's *core industry* as the industry in which it has made the largest share of its historical investments. I then merge exits at both the investment and fund levels to count IPO and M&A outcomes per fund–industry, and match deals to fund attributes (fund sequence, vintage year, committed capital, and GP identifiers). Because the model focuses on information flow from Fund 1 to Fund 2, I retain only first and second funds. Second funds are linked to their predecessor within the same strategy using firm IDs and fund names. I keep vintages 1995–2023 and require at least five registered investments to ensure a representative opportunity set. The final dataset contains 1,790 matched Fund 1–Fund 2 pairs.

Industry coverage. Table 1 summarizes the distribution of first and second funds by core industry in the final matched sample.

Table 1
Number of Funds by Industry in Final Sample

Core Industry	First-time Funds	Second-time Funds
Real Estate	2	4
Business Services	11	7
Raw Materials and Natural Resources	8	7
Energy and Utilities	27	19
Telecom and Media	17	23
Financial and Insurance Services	35	39
Industrials	30	43
Consumer Discretionary	107	124
Healthcare	328	335
Information Technology	1225	1244

Summary statistics for first funds. Table 2 reports moments for the full set of first-time funds and for the subset entering the "shifter" sample. Shifter funds are larger and allocate less to the core industry; their model-implied opportunity cost of exploration is lower.

Table 2
Summary Statistics for First-Time Funds and First-Time Funds in the Shifter Sample

Variables	# Obs	Mean	SD	p25	p50	p75
First-Time Funds						
Skill in Core Industry	1.790	−0.01	0.11	−0.05	−0.03	−0.01
Skill outside of Core Industry	1.624	−0.03	0.41	−0.20	−0.08	−0.03
Opportunity Cost of Exploration	1.624	0.02	0.38	0.00	0.08	0.16
Fraction Invested in Core Industry	1.790	0.65	0.19	0.47	0.62	0.75
Fund Size (Mil)	1.676	126.27	191.04	24.32	57.40	147.32
First-Time Funds in Shifter Sample						
Skill in Core Industry	516	−0.01	0.12	−0.04	−0.03	−0.01
Skill outside of Core Industry	508	0.04	0.48	−0.18	−0.06	0.02
Opportunity Cost of Exploration	508	−0.02	0.45	−0.02	0.07	0.16
Fraction Invested in Core Industry	516	0.51	0.17	0.39	0.49	0.64
Fund Size (Mil)	502	146.21	251.32	25.40	59.32	153.47

Summary statistics for second funds. Table 3 provides the analogous moments for Fund 2. Shifter Fund 2s allocate a smaller fraction to the core and have a lower opportunity cost of exploration.

Table 3
Summary Statistics for Second-Time Funds and Second-Time Funds in the Shifter Sample

Variables	# Obs	Mean	SD	p25	p50	p75
Second-Time Funds						
Skill in Core Industry	1.790	0.01	0.07	−0.03	−0.02	0.01
Skill outside of Core Industry	1.714	−0.02	0.26	−0.13	−0.07	−0.04
Opportunity Cost of Exploration	1.714	0.05	0.22	0.02	0.09	0.13
Fraction Invested in Core Industry	1.790	0.64	0.17	0.49	0.61	0.69
Fund Size (Mil)	1.701	201.62	261.32	57.20	117.48	221.64
Second-Time Funds in Shifter Sample						
Skill in Core Industry	516	0.02	0.14	−0.05	−0.02	0.02
Skill outside of Core Industry	513	0.01	0.31	−0.18	−0.06	−0.01
Opportunity Cost of Exploration	513	0.01	0.35	−0.02	0.06	0.17
Fraction Invested in Core Industry	516	0.49	0.13	0.41	0.50	0.53
Fund Size (Mil)	497	212.30	300.02	50.41	109.97	240.10

Additional early benchmarks. Table 4 and Table 5 report compact, early moments that are useful for orientation (means/SDs/percentiles for exploration, time between funds, and capital; and simple discipline patterns). Values are placeholders/benchmarks and should be replaced with sample moments from your final build.

Fact: Learning from Exploration

Figure 3 documents a key fact aligned with the model's learning channel: *Average relative success in Fund 2 rises with the share of capital devoted to exploration in Fund 1*. The plot shows binned averages of Fund 2 success (demeaned by industry–year) against the Fund 1 share invested in new industries, with 95% confidence intervals. The increasing profile indicates that

Table 4
Core Sample Overview (Matched Fund 1–Fund 2 Pairs)

Variable	# Obs	Mean	SD	p25	p75
Exploration α_1 (Fund 1)	1790	0.35	0.18	0.22	0.47
Exploration α_2 (Fund 2)	1790	0.36	0.19	0.23	0.49
Fraction in Core (Fund 1)	1790	0.65	0.19	0.47	0.75
Fraction in Core (Fund 2)	1790	0.64	0.17	0.49	0.69
Time between Fund 1 and Fund 2 (years)	1790	3.7	1.2	3.0	4.2
Fund Size (USD mil., Fund 1)	1676	126.3	191.0	24.3	147.3
Fund Size (USD mil., Fund 2)	1701	201.6	261.3	57.2	221.6
Deals per Fund (Fund 1)	1790	18.4	11.2	10.0	26.0
Deals per Fund (Fund 2)	1790	20.7	12.5	12.0	28.0
Core-only LP relationship (share)	1790	0.28	0.45	0.00	1.00
Hot vintage (share)	1790	0.41	0.49	0.00	1.00

Notes: Benchmarked/illustrative values to be replaced with computed moments. Fund sizes are committed capital.

Table 5
Capital Supply and Signal Discipline: Descriptive Benchmarks

	# Obs	Mean	SD	Corr. with I_2
Signal index Z (Fund 1)	1790	0.00	1.00	0.42
Prior return R_1 (std. units)	1790	0.00	1.00	0.29
Exploration α_1	1790	0.35	0.18	0.24
Monitoring μ_1^{eff}	1790	0.00	1.00	0.21
Scale I_2 (USD mil., committed)	1701	201.6	261.3	1.00
Mean I_2 (USD mil.) by Z tercile and mandate type				
Low Z tercile, core-only	—	158.0	190.0	—
Low Z tercile, non-core-only	—	172.5	205.4	—
Mid Z tercile, core-only	—	195.3	240.1	—
Mid Z tercile, non-core-only	—	225.4	270.6	—
High Z tercile, core-only	—	228.7	285.9	—
High Z tercile, non-core-only	—	275.9	320.2	—

Notes: Benchmarked/illustrative values. The tercile splits and mandate indicator are descriptive; the flattening of the I_2–Z slope under core-only mandates is tested formally in reduced forms.

exploratory exposure in Fund 1 is associated with higher subsequent performance, consistent with learning-by-doing and sector-specific know-how accumulation inside the GP.

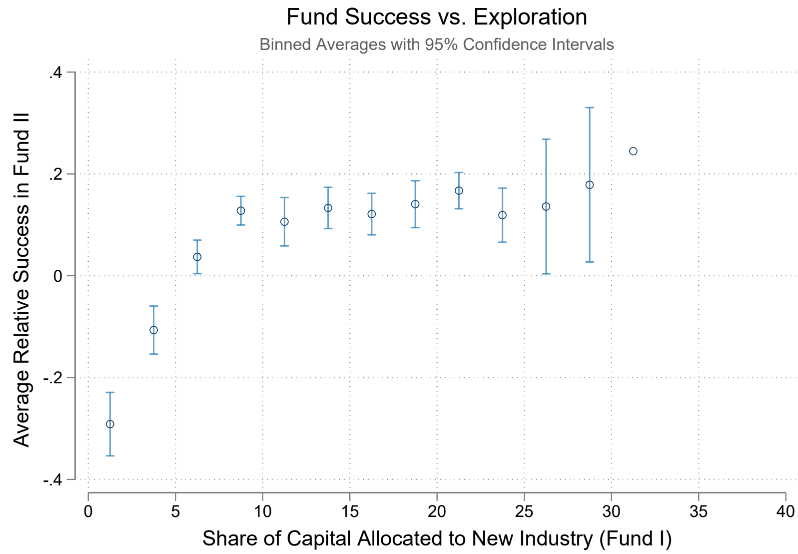


Figure 3. Success rate (demeaned by industry–year) in Fund 2 vs. share of capital allocated to new industries in Fund 1. Points are binned averages with 95% confidence intervals.

3 Model

This section develops a two-period model in which a general partner (GP) raises and deploys Fund 1 at date t and then raises and deploys Fund 2 at date $t+1$. There are three actors: limited partners (LPs) who supply capital and audit the GP, the GP who chooses exploration and monitoring, and entrepreneurs who seek financing and may misreport project quality. Two incentive layers operate. On the GP–entrepreneur link, insufficient screening in Fund 1 allows deceptive projects to pass, making early experimentation noisy and weakening the mapping from Fund 1 to Fund 2. On the LP–GP link, Fund 2 scale is chosen by investors in each vintage and therefore depends on Fund 1 signals, the aggregate state, and diffusion. Technology features learning and diffusion: the GP accumulates private know-how and the market accumulates knowledge as more funds experiment; adoption compresses per-hit rents in the new sector. The aggregate environment switches between hot and cold regimes according to a Markov chain.

3.1 Environment, states, and regimes

Time consists of two consecutive vintages. The aggregate regime $M_\tau \in \{\mathrm{H}, \mathrm{C}\}$ follows a first-order Markov chain with transition matrix

$$P = \begin{pmatrix} p_{HH} & 1 - p_{HH} \\ 1 - p_{CC} & p_{CC} \end{pmatrix}, \qquad p_{HH}, p_{CC} \in (0, 1). \tag{3.1}$$

Hot markets have lower core profitability in the sense that $p_c(\mathrm{H})R_c(\mathrm{H}) < p_c(\mathrm{C})R_c(\mathrm{C})$. The inter-fund discount factor is $\beta \in (0, 1)$.

Between funds, three stocks transmit information. The GP's private know-how L_t accumulates with Fund 1 exploration at rate $\theta_L > 0$ and depreciates at rate $\delta_L \in (0, 1)$:

$$L_t = (1 - \delta_L)L_{t-1} + \theta_L \alpha_{1,t-1}. \tag{3.2}$$

Market knowledge K_t aggregates cohort-wide exploration with regime-dependent diffusion $\theta(M)$ and depreciation $\delta_K(M) \in (0, 1)$:

$$K_t = (1 - \delta_K(M_{t-1}))K_{t-1} + \theta(M_{t-1})\, \mathbb{E}[\alpha_{1,\cdot,t-1}]. \tag{3.3}$$

Adoption A_t measures penetration of the new sector and compresses rents; it accumulates with average Fund 1 exploration at rate $\eta_A(M)$ and depreciates at rate $\delta_A \in (0, 1)$:

$$A_t = (1 - \delta_A)A_{t-1} + \eta_A(M_{t-1})\, \mathbb{E}[\alpha_{1,\cdot,t-1}]. \tag{3.4}$$

3.2 Entrepreneurs: continuation utility, credibility, and the IC

Entrepreneurs receive founder equity $s_E \in [0, 1)$ and continuation value $U_c \geq 0$ if their venture progresses to the next stage. Let $r_n > 0$ denote the monetary payoff per successful new-sector project and $c_e \geq 0$ the cost of truthful preparation. If a founder is truthful and a project is undertaken, expected utility equals

$$U_{\text{true}}^E = \pi_{\text{true}}(s_E r_n + U_c) - c_e, \tag{3.5}$$

where $\pi_{\text{true}} \in (0, 1]$ is a reduced-form success probability under the GP's diligence standard. If the founder misreports quality and a project is undertaken, deception is detected with probability $d(\mu^{\text{eff}})$ and punished with penalty $\Pi > 0$. Given detection $d(\mu^{\text{eff}}) = 1 - e^{-\rho\mu^{\text{eff}}}$ for $\rho > 0$, expected utility under deception is

$$U_{\text{dec}}^E = \big(1 - d(\mu^{\text{eff}})\big)(s_E r_n + U_c) - d(\mu^{\text{eff}})\Pi = e^{-\rho\mu^{\text{eff}}}(s_E r_n + U_c + \Pi) - \Pi. \tag{3.6}$$

The entrepreneur incentive constraint is $U_{\text{true}}^E \geq U_{\text{dec}}^E$, equivalently

$$\pi_{\text{true}}(s_E r_n + U_c) - c_e + \Pi \geq e^{-\rho\mu^{\text{eff}}}(s_E r_n + U_c + \Pi). \tag{3.7}$$

Let $V \equiv s_E r_n + U_c$. Feasibility requires $\pi_{\text{true}} V - c_e + \Pi > 0$. Solving (3.7) for μ^{eff} yields the explicit monitoring threshold

$$\mu^{\text{eff}} \geq \mu_E^{\text{thresh}}(s_E, U_c, \Pi, \pi_{\text{true}}, c_e) \equiv \frac{1}{\rho} \ln\left(\frac{V + \Pi}{\pi_{\text{true}} V - c_e + \Pi}\right). \tag{3.8}$$

The threshold is decreasing in s_E and Π, increasing in U_c and c_e, and decreasing in π_{true}. Credibility also responds at the extensive margin. Before screening, the mass of deceptive applicants is $\bar{m}_t = \bar{m}_0/(1 + \theta_{\text{match}} L_{t-1})$, so better GPs attract more credible founders; the mass that passes undetected after screening equals

$$m(\mu^{\text{eff}}) = \bar{m}_t \, e^{-\rho\mu^{\text{eff}}}. \tag{3.9}$$

3.3 GP monitoring given LP audit and co-investment

The LP chooses audit intensity $m \geq 0$ which amplifies GP monitoring through $\mu^{\text{eff}} = \mu\, s(m)$ with $s(0) = 1$, $s'(m) > 0$, and $s''(m) < 0$. Co-investment $a \geq 0$ tightens GP alignment by raising the fraction $W(a) \in (0, 1]$ of investor losses internalized, with $W'(a) \geq 0$ and $W(0) = W_0 \leq 1$. Undetected deception destroys investor value $\Lambda_t > 0$ per bad project. The GP chooses $\mu \geq 0$ to maximize

$$\max_{\mu \geq 0} \;\; W(a) \left[\bar{m}_t - \bar{m}_t e^{-\rho\mu s(m)}\right] \Lambda_t \; - \; \frac{\phi}{2}\mu^2 \qquad \text{subject to} \qquad \mu\, s(m) \geq \mu_E^{\text{thresh}}. \tag{3.10}$$

The objective is strictly concave. Denote $x \equiv \rho\mu s(m)$. The first-order condition for the unconstrained problem reduces to $xe^x = (\rho s(m) W(a) \bar{m}_t \Lambda_t)/\phi$, whose unique solution is $x^* = W_0((\rho s(m) W(a) \bar{m}_t \Lambda_t)/\phi)$, where W_0 is the principal branch of the Lambert W function. Hence the interior optimum is

$$\mu^* \;=\; \frac{1}{\rho s(m)}\, W_0\!\left(\frac{\rho s(m) W(a) \bar{m}_t \Lambda_t}{\phi}\right). \tag{3.11}$$

The constrained optimum is $\mu^{**} = \max\{\mu^*, \mu_E^{\text{thresh}}/s(m)\}$. Appendix **??** proves uniqueness and the comparative statics $\partial\mu^*/\partial m > 0$, $\partial\mu^*/\partial a > 0$, $\partial\mu^*/\partial\bar{m}_t > 0$, and $\partial\mu^*/\partial\phi < 0$.

3.4 Per-dollar returns, learning, and rent compression

For a tilt $\alpha \in [0, 1]$ in regime $s \in \{\text{H}, \text{C}\}$, the per-dollar success in the new sector is

$$p_{j,t}^{\text{new}}(s) \;=\; p_0(s) + \left(1 - p_0(s)\right)\left(1 - e^{-(\lambda + \eta_L L_{t-1} + \rho_K K_{t-1})\alpha}\right), \tag{3.12}$$

with $\lambda > 0$ a GP-specific learning parameter and $p_0(s) \in (0, 1)$ the regime baseline. Per-hit rents are compressed by adoption through

$$R_n(A_t) \;=\; R_{n0}\,(1 - \theta_A A_t), \qquad \theta_A > 0. \tag{3.13}$$

The per-dollar return index is

$$R_{j,t}(\alpha) \;=\; \alpha\, p^{\text{new}}_{j,t}(s)\, R_n(A_t) \;+\; (1-\alpha)\, p_c(s) R_c(s). \tag{3.14}$$

Its marginal with respect to α decomposes into a composition gap $\Delta R_t \equiv p^{\text{new}}_{j,t}(s) R_n(A_t) - p_c(s) R_c(s)$ and a learning term proportional to α and $\lambda + \eta_L L_{t-1} + \rho_K K_{t-1}$.

3.5 Fund 1 choices and continuation value on the GP–entrepreneur link

In Fund 1 the GP chooses exploration $\alpha_1 \in [0,1]$ and monitoring $\mu_1 \geq 0$; the LP sets audit $m_1 \geq 0$ and may require co-investment $a_1 \geq 0$. The GP receives share $\phi \in (0,1)$ on dollars $I_1 \geq 0$. Private costs are $c_1 \alpha_1$ for exploration and $(\phi_1/2)\mu_1^2$ for monitoring with $\phi_1 > 0$. Entrepreneurs receive founder equity $s_E^{(1)} \in [0,1)$ and continuation utility $U_c^{(1)} \geq 0$ if funded again. Let $r_n^{(1)}$ denote the monetary payoff per successful new-sector project in Fund 1 and define $V_1 \equiv s_E^{(1)} r_n^{(1)} + U_c^{(1)}$.

Founders may misreport quality. Detection follows $d(\mu_1^{\text{eff}}) = 1 - e^{-\rho \mu_1^{\text{eff}}}$ with sensitivity $\rho > 0$, where LP audit amplifies GP screening through $\mu_1^{\text{eff}} = \mu_1\, s(m_1)$, $s'(m_1) > 0$, $s''(m_1) < 0$, $s(0) = 1$. Let $\pi^{(1)}_{\text{true}} \in (0,1]$ be the success probability conditional on truthful reporting under the GP's diligence standard, $c_e \geq 0$ a truthful preparation cost, and $\Pi > 0$ the penalty if deception is detected. Expected utilities are

$$U_1^{E,\text{true}} = \pi^{(1)}_{\text{true}} V_1 - c_e, \qquad U_1^{E,\text{dec}} = e^{-\rho \mu_1^{\text{eff}}}(V_1 + \Pi) - \Pi.$$

The entrepreneur IC is $U_1^{E,\text{true}} \geq U_1^{E,\text{dec}}$, which is equivalent to the explicit monitoring threshold

$$\mu_1^{\text{eff}} \;\geq\; \mu^{\text{thresh}}_{E,1} \;\equiv\; \frac{1}{\rho} \ln\!\left(\frac{V_1 + \Pi}{\pi^{(1)}_{\text{true}} V_1 - c_e + \Pi} \right). \tag{3.15}$$

The threshold is decreasing in the founder's upside ($\partial \mu^{\text{thresh}}_{E,1} / \partial s_E^{(1)} < 0$), increasing in $U_c^{(1)}$ if it mainly raises deception temptation, decreasing in the detection penalty ($\partial \mu^{\text{thresh}}_{E,1} / \partial \Pi < 0$), and decreasing in screening quality via $\pi^{(1)}_{\text{true}}$.

Credibility at the extensive margin also depends on GP reputation. Before screening, the mass of deceptive applicants is $\bar{m}_t = \bar{m}_0/(1 + \theta_{\text{match}} L_{t-1})$, which declines with the GP's private know-how L_{t-1} accumulated via prior exploration. The mass that survives undetected after screening is $m(\mu_1^{\text{eff}}) = \bar{m}_t e^{-\rho\mu_1^{\text{eff}}}$; each undetected bad project destroys investor value $\Lambda_t > 0$.

Given LP audit m_1 and co-investment a_1 that tightens alignment $W(a_1) \in (0, 1]$ with $W'(a_1) \geq 0$, the GP chooses μ_1 to solve

$$\max_{\mu_1 \geq 0} \; W(a_1)\left[\bar{m}_t - \bar{m}_t e^{-\rho\mu_1 s(m_1)}\right]\Lambda_t \; - \; \frac{\phi_1}{2}\mu_1^2 \quad \text{subject to} \quad \mu_1 s(m_1) \geq \mu_{E,1}^{\text{thresh}}. \tag{3.16}$$

The unique unconstrained optimum is

$$\mu_1^* \; = \; \frac{1}{\rho\, s(m_1)}\, W_0\!\left(\frac{\rho\, s(m_1)\, W(a_1)\, \bar{m}_t\, \Lambda_t}{\phi_1}\right), \tag{3.17}$$

where $W_0(\cdot)$ is the principal branch of the Lambert–W function. The constrained optimum is $\mu_1^{**} = \max\{\mu_1^*, \mu_{E,1}^{\text{thresh}}/s(m_1)\}$. Comparative statics are $\partial\mu_1^{**}/\partial m_1 > 0$, $\partial\mu_1^{**}/\partial a_1 > 0$, $\partial\mu_1^{**}/\partial\bar{m}_t > 0$, and $\partial\mu_1^{**}/\partial\phi_1 < 0$.

Per-dollar Fund 1 returns combine the core and new-sector legs:

$$R_{1,t}(\alpha_1) \; = \; \alpha_1\, p_{1,t}^{\text{new}}(M_t)\, R_n(A_t) \; + \; (1 - \alpha_1)\, p_c(M_t) R_c(M_t),$$

with $p_{1,t}^{\text{new}}$ increasing and concave in α_1 via own learning λ, private know-how L_{t-1}, and market knowledge K_{t-1}, and with $R_n(A_t) = R_{n0}(1 - \theta_A A_t)$ compressed by adoption. The GP's Fund 1 value is

$$U_1^{\text{GP}} \; = \; \phi I_1 R_{1,t}(\alpha_1) \; - \; c_1\alpha_1 \; - \; \frac{\phi_1}{2}(\mu_1^{**})^2 \; + \; \beta\,\mathbb{E}_t\left[V_2\big(L_t, K_t, A_{t+1};\, R_{1,t}, \alpha_1, \mu_1^{\text{eff}}, M_{t+1}\big)\right]. \tag{3.18}$$

The interior exploration condition equates the marginal private benefit to cost,

$$\phi I_1 \frac{\partial R_{1,t}}{\partial \alpha_1} \;=\; c_1 \quad \text{with} \quad \frac{\partial R_{1,t}}{\partial \alpha_1} = \left[p_{1,t}^{\text{new}} R_n(A_t) - p_c R_c\right] + \alpha_1 (1 - p_0)\, e^{-X_{t-1}\alpha_1}\, X_{t-1}\, R_n(A_t), \tag{3.19}$$

where $X_{t-1} = \lambda + \eta_L L_{t-1} + \rho_K K_{t-1}$. Hot markets (lower $p_c R_c$) raise the composition gap and push α_1 up unless A_t is already high; adoption A_t reduces R_n and damps exploration by late movers. Larger L_{t-1} and K_{t-1} steepen the learning term and raise α_1.

Finally, Fund 1 choices deliver a continuation value by improving signals to LPs. The signal that summarizes Fund 1 discipline is

$$Z_t \;=\; \gamma_0 + \gamma_1 R_{1,t} + \gamma_2 \alpha_1 + \gamma_3 \mu_1^{\text{eff}}, \tag{3.20}$$

which tilts Fund 2 scale through the rule derived below. Larger α_1 and higher μ_1^{eff} therefore raise Fund 2 dollars both directly (via Z_t) and indirectly (via L_t and K_t).

3.6 Fund 2 scale and composition on the LP–GP link

At t+1 the LP observes regime M_{t+1}, states (K_t, A_{t+1}), and Fund 1 signals $(R_{1,t}, \alpha_1, \mu_1^{\text{eff}})$ summarized by Z_t, and then sets audit $m_2 \geq 0$ and dollar scale $I_2 \geq 0$. The LP faces outside return $r_f > 0$ and decreasing returns characterized by $\kappa(\alpha_2) > 0$, allowing scale curvature to depend on the Fund 2 tilt. A fraction $\varpi \in [0, 1)$ of LP capital is core-only, contractually imposing $\alpha_2 \equiv 0$ on that slice; only $(1 - \varpi)I_2$ can be tilted toward the new sector. The LP forecasts the per-dollar Fund 2 return as

$$\widehat{R}_{2,t+1} \;=\; r_0 + r_K K_t - r_A A_{t+1} + r_M \{M_{t+1} = \text{H}\} + r_1 R_{1,t} + r_\alpha \alpha_1 + r_\mu \mu_1^{\text{eff}}. \tag{3.21}$$

Let $\phi_L > 0$ be the quadratic audit-cost coefficient. Undetected deception in Fund 2 destroys value $\Lambda_{t+1} > 0$, of which the GP internalizes $W(a_2) \in (0, 1]$ through co-investment $a_2 \geq 0$. The

LP's objective is

$$U_2^{\text{LP}} \;=\; (1-\phi)I_2\widehat{R}_{2,t+1} - r_f I_2 - \frac{\kappa(\alpha_2)}{2}I_2^2 - \frac{\phi_L}{2}m_2^2 - \big(1-W(a_2)\big)\, m(\mu_2^{\text{eff}})\Lambda_{t+1}, \tag{3.22}$$

with $\mu_2^{\text{eff}} = \mu_2 s(m_2)$, $m(\mu_2^{\text{eff}}) = \bar{m}_{t+1}e^{-\rho\mu_2^{\text{eff}}}$, and $\bar{m}_{t+1} = \bar{m}_0/(1+\theta_{\text{match}}L_t)$ as in Fund 1.

For any fixed α_2 and m_2, the unique scale rule is

$$I_2^* \;=\; \max\left\{0,\; \frac{(1-\phi)\widehat{R}_{2,t+1} - r_f}{\kappa(\alpha_2)}\right\}, \tag{3.23}$$

which is strictly increasing in $\widehat{R}_{2,t+1}$ and strictly decreasing in r_f and $\kappa(\alpha_2)$. The LP's audit choice solves

$$\min_{m_2\geq 0}\;\; \frac{\phi_L}{2}m_2^2 + \big(1-W(a_2)\big)\,\bar{m}_{t+1}e^{-\rho\mu_2 s(m_2)}\Lambda_{t+1} \quad \text{subject to} \quad \mu_2 s(m_2) \geq \mu_{E,2}^{\text{thresh}}, \tag{3.24}$$

where $\mu_{E,2}^{\text{thresh}}$ is defined analogously to (3.15) with $(s_E^{(2)}, U_c^{(2)}, \pi_{\text{true}}^{(2)})$. Given m_2, the GP optimally sets

$$\mu_2^* \;=\; \frac{1}{\rho\, s(m_2)}\, W_0\!\left(\frac{\rho\, s(m_2)\, W(a_2)\, \bar{m}_{t+1}\, \Lambda_{t+1}}{\phi_1}\right), \qquad \mu_2^{**} = \max\{\mu_2^*, \mu_{E,2}^{\text{thresh}}/s(m_2)\}. \tag{3.25}$$

The LP's first-order condition in m_2 equates marginal audit cost to the marginal reduction in undetected-loss disutility through $s'(m_2)$, uniquely determining m_2^*.

Given (I_2^*, m_2^*), the GP chooses composition $\alpha_2 \in [0,1]$ to maximize

$$\phi(1-\varpi)I_2^*\, R_{2,t+1}(\alpha_2) - c_2\alpha_2,$$

with interior first-order condition

$$\phi(1-\varpi)I_2^*\,\frac{\partial R_{2,t+1}}{\partial \alpha_2} \;=\; c_2, \qquad \frac{\partial R_{2,t+1}}{\partial \alpha_2} = \big[p_{2,t+1}^{\text{new}} R_n(A_{t+1}) - p_c R_c\big] + \alpha_2(1-p_0)\, e^{-X_t\alpha_2}\, X_t\, R_n(A_{t+1}), \tag{3.26}$$

where $X_t = \lambda + \eta_L L_t + \rho_K K_t$ and all primitives are evaluated at M_{t+1}. Core-only capital ϖ scales down the private marginal return to α_2 and therefore makes α_2^* less responsive to signals Z_t. Combining (3.21) and (3.23) gives the discipline channel $\partial I_2^*/\partial Z_t > 0$, so stronger Fund 1 performance, exploration, and monitoring translate into larger Fund 2 scale whenever the interior region binds.

3.7 Participation and equilibrium

LP participation over the two funds requires $U_1^{\mathrm{LP}} + \beta\,\mathbb{E}_t[U_2^{\mathrm{LP}}] \geq 0$, with U_2^{LP} given by (3.22) and an analogous expression for U_1^{LP} based on Fund 1 primitives. The GP's lifetime utility is $U_1^{\mathrm{GP}} + \beta\,\mathbb{E}_t[U_2^{\mathrm{GP}}] \geq 0$, where U_1^{GP} is (3.18) and $U_2^{\mathrm{GP}} = \phi(1-\varpi)I_2^* R_{2,t+1}(\alpha_2^*) - c_2\alpha_2^* - \frac{\phi_1}{2}(\mu_2^{**})^2$.

GPs are heterogeneous in learning and costs. Let $\Theta = (\lambda, c_1, c_2, \phi_1, \phi)$ denote type, drawn from a compact set with distribution F. Given state $(L_{t-1}, K_{t-1}, A_t, M_t)$ and dollars I_1, a type-Θ GP solves (3.16) and (3.19) to obtain Fund 1 policies $(\mu_1^{**}(\Theta; M_t), \alpha_1^*(\Theta; M_t))$ and a signal $Z_t(\Theta; M_t)$ via (3.20). The LP maps $(Z_t, K_t, A_{t+1}, M_{t+1})$ into $\widehat{R}_{2,t+1}$ using (3.21) and then into a type–state specific scale

$$I_2^*(\Theta; M_{t+1} \mid Z_t) \;=\; \max\left\{0,\; \frac{(1-\phi)\,\widehat{R}_{2,t+1}(Z_t, K_t, A_{t+1}, M_{t+1}) - r_f(M_{t+1})}{\kappa(\alpha_2)}\right\}. \tag{3.27}$$

Given I_2^* and m_2^*, the type-Θ GP chooses $\alpha_2^*(\Theta; M_{t+1} \mid Z_t)$ from (3.26); the monitoring choice $\mu_2^{**}(\Theta; M_{t+1})$ follows from (3.24)–(3.25).

A decentralized equilibrium is a measurable collection of policy functions $\{(\mu_1^{**}, \alpha_1^*, \mu_2^{**}, \alpha_2^*, I_2^*)\}_\Theta$ and state transitions (L_t, K_t, A_{t+1}) such that: (i) the entrepreneur ICs (3.15) (and its Fund 2 analogue) hold weakly; (ii) for each Θ, the GP solves the Fund 1 problem and the Fund 2 composition problem given audit and scale rules; (iii) the LP solves the scale and audit problems given α_2; (iv) L_t, K_t, and A_{t+1} evolve from population averages of α_1^* according to the laws of motion; and (v) LP and GP participation constraints bind weakly. Existence follows from strict concavity of the GP subproblems, the linear–quadratic form of the LP scale choice, continuity of the signal and forecasting maps, and a fixed-point argument on (α_2, I_2) conditional on $(Z_t, K_t, A_{t+1}, M_{t+1})$. On the interior, monotone comparative statics obtain: α_1^* and α_2^* increase in λ and decrease in (c_1, c_2); I_2^* increases in Z_t and in hot states when adoption is not yet high; core-only capital ϖ weakly depresses both the level and the sensitivity of α_2^* and I_2^* to Z_t.

3.8 Heterogeneity and equilibrium by type and state

GPs are heterogeneous. Let $\Theta = (\lambda, c_1, c_2, \phi_1, \phi, \phi_i)$ collect type parameters governing learning, exploration cost, monitoring cost, and carry. Types are drawn from a compact set $\mathcal{T} \subset \mathbb{R}^6$ with distribution F. LPs may be heterogeneous in r_f and $\kappa(\cdot)$ across regimes. Given $(L_{t-1}, K_{t-1}, A_t, M_t)$ and given Fund 1 dollars I_1, a type-Θ GP solves (**??**) with monitoring given by (3.11) and E-IC (3.8). This yields type-specific Fund 1 policies $(\alpha_1^*(\Theta; s), \mu_1^{**}(\Theta; s))$ and signals $(R_{1,t}(\Theta; s), \alpha_1^*, \mu_1^{\text{eff}})$. At $t{+}1$, for each regime s' the LP maps signals into $\widehat{R}_{2,t+1}$ in (**??**) and chooses $I_2^*(\Theta, s' \mid Z)$ in (**??**). The GP then solves (**??**) to obtain $\alpha_2^*(\Theta, s' \mid Z)$, where Z denotes the signal vector. An equilibrium is a measurable collection $\{(\alpha_1^*, \mu_1^{**}, \alpha_2^*, I_2^*)\}_{\Theta \in \mathcal{T}}$ consistent with the LP's forecasting rule, with state transitions (3.2)–(3.4), and with participation. Appendix F provides existence. Under interior choices, monotone comparative statics follow from single-crossing: α_1^* and α_2^* are strictly increasing in λ and strictly decreasing in c_1 and c_2, and they are larger in hot regimes when adoption is not yet high. The scale mapping $I_2^*(\Theta, s' \mid Z)$ is strictly increasing in the scalar signal Z when $(1 - \phi) > 0$ and $\kappa(\cdot)$ is fixed.

3.9 Planner benchmark and comparison

A frictionless planner eliminates deception costs and internalizes diffusion. On the GP–entrepreneur link, deception is ruled out and monitoring costs vanish; on the LP–GP link, scale is chosen at social cost with curvature $\kappa_0 > 0$ and carry is a transfer. The planner's scale rule is

$$I_2^{\text{PL}} = \max\left\{0, \frac{R_{2,t+1}(\alpha_2) - r_f}{\kappa_0}\right\}. \tag{3.28}$$

The planner selects α_1 to internalize $\partial K_t/\partial\alpha_1 = \theta(M_{t-1})$ and selects α_2 to balance the marginal return (3.14) against adoption compression, scaling by I_2^{PL}. Appendix **??** proves three results that will anchor the welfare analysis. First, for any α_2 and on the interior, $I_2^* \le I_2^{\mathrm{PL}}$, with a strict inequality whenever $(1-\phi) < 1$, $\kappa(\alpha_2) > \kappa_0$, or $\widehat{R}_{2,t+1} < R_{2,t+1}(\alpha_2)$. Second, if undetected deception attenuates the information content of Fund 1 experiments so that the slope $\partial R_{2,t+1}/\partial\alpha_2$ is scaled down by $q \in (0,1)$ holding (K_t, A_{t+1}) fixed, then the interior solution for $\alpha_2^*(q)$ is strictly increasing in q and satisfies $\alpha_2^*(q) < \alpha_2^{\mathrm{PL}}$ for $q < 1$. Third, define entry as the first vintage in which scale positivity and marginal profitability at $\alpha_2 = 0$ both hold. Under frictions, entry is weakly later and strictly later than the planner whenever either inequality is strict before the planner's entry time. These algebraic comparisons formalize delay in both composition and scale.

4 Structural Framework: Motivation, Identification, and Estimation

This section lays out why I estimate the model structurally, how model objects map to observables, where identification comes from, and how the estimation proceeds by the Simulated Method of Moments (SMM). The central empirical difficulty is that the key margins—monitoring at two layers, exploration translating into future composition under diffusion and adoption, and Fund 2 scale determined by market discipline—are jointly determined and dynamically linked across funds. A reduced-form design can document correlations or instrument isolated slopes, but it cannot recover the structural objects that matter for welfare or policy design: the monitoring technology at both links (LP–GP and GP–entrepreneur), the decomposition of learning versus adoption in success and rents, and the endogenous Fund 2 scale rule that prices new-sector risk conditional on Fund 1 signals. The model provides exact cross-equation restrictions tying these pieces together—most visibly the monitoring response that enters the entrepreneur IC via the detection technology, the diffusion laws of motion for knowledge and adoption, and the linear interior scale rule in Fund 2 that turns Fund 1 signals into dollars at work. Estimating these primitives jointly is what allows the paper to measure wedges, attribute welfare losses to layers, and run counterfactuals that change contracts or prices.

4.1 What the structural model adds beyond reduced-form designs

The first limitation of reduced-form strategies is simultaneity between Fund 2 composition and Fund 2 scale. The GP's Fund 2 tilt responds to expected per-dollar returns, while scale is set by LPs from the same expectation; both objects depend on the same knowledge and adoption stocks and on Fund 1 signals. Without the Fund 2 scale equation, one risks conflating composition with scale, especially in hot states where prices ration capital. The model resolves this by making the LP's best response explicit. On the interior, equation (**??**) implies a linear scale rule in predicted per-dollar returns, themselves a function of Fund 1 outcomes and the state. This delivers a clean mapping from signals to scale that can be estimated and taken to counterfactuals that shift pricing or carry.

The second limitation is the two-layer monitoring technology. At the GP–entrepreneur link, the entrepreneur's IC depends on detection probabilities induced by effective monitoring; at the LP–GP link, audit and co-invest raise effective monitoring and change its cost. In the model, the GP's optimal monitoring solves a strictly concave program with a closed-form solution involving the Lambert–W function, equation (3.11). LP audit enters via the amplification function $s(m)$, while co-invest alignment enters via $W(a)$; these two levers move monitoring through orthogonal channels. A reduced-form regression of monitoring proxies on audit cannot disentangle these channels or connect them to the E–IC constraint, which is needed to assess contract changes.

The third limitation is separating learning from adoption in a dynamic setting. Fund 1 exploration raises private know-how L and market knowledge K, which raise success probabilities in the new sector; at the same time, adoption A compresses rents. These forces work in opposite directions for private returns and shift between funds. The structural laws (**??**) and rent compression (**??**) make this separation possible using the joint evolution of success, payoffs, and entry hazards; the decomposition is essential for quantifying how much of late entry is an information problem versus a pricing/scale problem.

Finally, welfare-relevant policy analysis requires counterfactual general equilibrium discipline. A policy that changes monitoring incentives or LP pricing shifts the Fund 2 scale rule, which in turn feeds back into adoption, rents, and the intertemporal return to Fund 1 exploration. These feedbacks only close once the primitives in (3.11), (**??**)–(**??**), and (**??**) are jointly estimated.

4.2 Data mapping, measurement, and normalizations

The estimation combines a GP–vintage panel with sector–year aggregates. For each GP i and Fund 1 vintage t, I observe the new-sector composition share $\alpha_{1,it}$, a set of diligence, founder-screening, and contractual auditing covenants that form an index of effective monitoring $\mu_{1,it}^{\text{eff}}$ using the amplification mapping $s(m_1)$, realized Fund 1 per-dollar performance $R_{1,it}$, and the market regime $M_t \in \{\mathrm{H}, \mathrm{C}\}$. For Fund 2 at $t+1$, I observe $\alpha_{2,i,t+1}$, the dollars raised or deployed $I_{2,i,t+1}$, contractual terms $(s_G^{(2)}, m_{2,t+1}, a_{2,t+1})$, realized performance $R_{2,i,t+1}$, and M_{t+1}. At the sector–year level I construct K_t as a leave-one-GP-out exposure-weighted moving average of Fund 1 exploration shares, consistent with the diffusion law in (**??**), and I construct A_t as the share of sector dollars allocated to the new category (again leave-one-out when used on the right-hand side) to match rent compression in (**??**). Core-only LPs are identified from term sheets and covenants that contractually impose $\alpha \equiv 0$ in Fund 2.

Normalizations are standard. I set the rent scale R_{n0} in (**??**) to one unit of Fund j per-dollar payoff, normalize the audit amplifier at zero audit to $s(0) = 1$, and choose the curvature units in $\chi(\alpha)$ so that one unit of I corresponds to one fund-multiple of capital. When I refer to the discipline score Z_t in (**??**), its intercept is centered so the scale regression has a well-identified slope. In heterogeneity, the mean of GP learning ability λ_i is normalized and dispersion is estimated.

4.3 Parameter vector and observables

The structural parameter vector Θ consists of four blocks. The monitoring block contains the detection slope ρ and audit amplifier $s(\cdot)$ entering (3.11), the LP audit cost parameter ϕ_L, the GP monitoring cost ϕ_i distribution, and the alignment map $W_j(a) = W_{j0} + \nu_a a$ with potentially different baselines across funds (W_{10}, W_{20}). The entrepreneur block contains the E–IC penalty Π and continuation utility $U_c^{(j)}$, which feed into the monitoring threshold in (**??**). The knowledge–adoption block contains η_L (private learning), ρ_K (market knowledge intensity), the decay–diffusion pair $(\delta_L, \delta_K(M))$, and the adoption rent-compression slope θ_A in (**??**). The scale–pricing block contains the carry share $s_G^{(2)}$, the interior scale curvature $\chi(\alpha)$, the discipline mapping coefficients in (**??**)–(**??**), and the hot–cold primitives for core returns $p_c(M)R_c(M)$. When market clearing is imposed, the price pair $(r_{f,t}, k_{I,t})$ is chosen endogenously to match aggregate supply and therefore is not free.

4.4 Identification

Identification exploits contractual and policy-driven shocks that move audit and co-invest, within- and between-cohort differences in Fund 1 exploration interacting with state transitions, and the mechanical separation created by the Fund 2 scale rule. Each object in Θ appears in a cross-equation restriction that can be targeted by a moment that is both observable and insensitive to other blocks after conditioning.

The split between LP audit and GP co-invest is identified from their distinct roles in the optimal monitoring rule (3.11) and the LP's audit first-order condition in (**??**). Audit raises monitoring by scaling the amplifier $s(m)$ and appears in the LP's marginal condition with derivative $s'(m)$; co-invest only shifts the alignment term $W(a)$ that multiplies the avoided loss inside the Lambert–W argument. Cross-vintage changes in compliance policy and monitoring covenants provide the audit variation, while the adoption of programmatic co-invest across some LPs and not others provides the alignment variation. Comparing these shifts across funds identifies the relative baselines $W_{20} < W_{10}$ when Fund 2 alignment is weaker. The entrepreneur threshold in (**??**) supplies an inequality restriction: conditional on founder pay and continuation terms, observed passes of E–IC at given audit and co-invest locate μ_E^{thresh} and therefore the levels of Π and $U_c^{(j)}$ up to a scale.

Learning and adoption are separated using the law of motion (**??**) and rent compression (**??**). The slope and curvature of Fund 2 success in the Fund 1 exploration share identifies private learning η_L and the exponential slope in (3.12). The leave-one-GP-out construction of K_t eliminates mechanical feedback and pins the market knowledge coefficient ρ_K. Adoption compresses rents through θ_A, which is identified by the slope of realized per-success payoffs in the adoption stock A within sector–years once knowledge is controlled for. Jointly using success rates and payoffs of new-sector deals by adoption decile provides the cross-equation restriction.

Hot–cold primitives are identified by comparing core returns $p_c R_c$ across regimes $M \in \{\text{H}, \text{C}\}$ in periods where new-sector exposure is low and adoption is negligible. State dependence in diffusion is identified by the persistence and growth of K_t across M_t transitions in otherwise comparable sectors.

The Fund 2 discipline channel is identified from the interior scale rule (**??**) coupled with the forecasting equation (**??**). On the interior, scale is linear in predicted returns, which are linear in the discipline score Z_t and the state. Regressions of I_2 on Fund 1 signals $(R_1, \alpha_1, \mu_1^{\text{eff}})$, with sector and regime controls, recover the loadings on the signals up to the scale factor $(1 - s_G^{(2)})/\chi(\alpha_2)$. Independent contractual variation in carry across funds and variation in α_2 that changes $\chi(\alpha_2)$ separate these pieces. Because α_2 is itself chosen, I use the model's Euler condition for α_2 together with the scale regression: parameters that govern success and rent compression must rationalize both the α_2 choice and its scale consequences. This cross-equation restriction produces overidentification that I use to test the model.

The mass of core-only LP capital ϖ and its effect on discipline is identified by contrasting the I_2 sensitivity to Fund 1 signals among funds whose contracts impose $\alpha \equiv 0$ versus unrestricted funds in the same regime and sector. The weaker sensitivity among core-only contracts locates ϖ and confirms that the discipline mapping is not a pure composition story.

Heterogeneity in GP learning λ_i is identified from within-GP slopes of Fund 2 success and composition in Fund 1 exploration shares, controlling for states and sector–year exposure. Quantile differences in these slopes across the GP distribution pin the dispersion of λ_i.

4.5 Estimation strategy and SMM criterion

I estimate Θ by Simulated Method of Moments (SMM). Let $M^{\mathrm{S}}(\Theta)$ denote the vector of simulated moments and M^{D} the corresponding data moments; define the sample moment vector

$$g_N(\Theta) \equiv M^{\mathrm{S}}(\Theta) - M^{\mathrm{D}}.$$

The SMM estimator solves

$$\hat{\Theta} \in \arg\min_{\Theta} \; g_N(\Theta)' W_N \, g_N(\Theta),$$

where W_N is a positive semi-definite weighting matrix. I implement a two-step optimal SMM. In step 1, W_N is diagonal with entries equal to inverse variances of the empirical moments; this yields $\hat{\Theta}^{(1)}$. In step 2, I set W_N to the inverse of the estimated asymptotic covariance matrix of $\sqrt{N}\, g_N(\hat{\Theta}^{(1)})$, computed by a block bootstrap over GP–vintages with common random numbers in simulation. Standard errors are obtained from the SMM sandwich formula using the optimal W_N and the Jacobian of the simulated moments with respect to Θ.

The moment set stacks objects that load selectively on the parameter blocks. I target the semi-elasticity of I_2 in each Fund 1 signal and in the regime indicator, the slope and curvature of Fund 2 composition α_2 in (α_1, K, A), the slope of new-sector per-success payoffs in adoption A, the cross-regime difference in core per-dollar returns, the monitoring response to audit and co-invest on both funds, and the pass rate of E–IC at given founder compensation and audit. When I impose market clearing, I also target the aggregate profile of Fund 2 dollars by vintage. The estimator is over-identified; I report the GMM J-statistic and its p-value.

To aid convergence and reduce nonlinearity, I first recover reduced-form objects that the model nests. Specifically, I estimate the discipline regression implied by (**??**)–(**??**) to obtain the loadings of Fund 1 signals in $\widehat{R}_2$ up to $(1 - s_G^{(2)})/\chi(\alpha_2)$, and I regress realized per-success payoffs on adoption and knowledge to pin θ_A and the sign of the knowledge effect. These enter the SMM step as starting values and moment targets.

4.6 Simulation, fixed components, and algorithm

Simulation follows the model's timing with common random numbers held fixed across iterations. For each GP i and Fund 1 vintage t, I draw the latent GP learning λ_i from a parametric distribution, record the regime M_t, sector identifiers, and the leave-one-out diffusion stocks K_t and A_t, and take contractual terms and audit/co-invest covenants as observed.

Given m_1 and a_1, I compute the optimal Fund 1 monitoring μ_1^* using (3.11). If the resulting effective monitoring fails the E–IC threshold in (**??**) at observed founder pay and continuation terms, I minimally adjust m_1 or a_1 to clear E–IC, consistent with covenants; this enforces the model's inequality restriction in-sample. I then solve the Fund 1 exploration condition (**??**) by line search and update private learning and market knowledge using (**??**). I construct the discipline score Z_t from Fund 1 signals, predict Fund 2 per-dollar returns using (**??**), and compute Fund 2 scale from (**??**). Where market clearing is imposed, I choose $(r_{f,t+1}, k_{I,t+1})$ so that simulated $\sum_i I_{2,i,t+1}$ matches aggregate Fund 2 capital observed in that vintage. Finally, conditional on the simulated I_2^*, I solve the Fund 2 exploration condition (**??**) by line search, clamp $\alpha_2 \in [0, 1]$, and record realized success and per-success payoffs using (**??**). Targeted moments are then aggregated across GPs and vintages using the same weights used to construct the empirical moments.

For numerical stability in SMM, I use common random numbers, analytic or automatic-differentiation Jacobians for smooth components, and damped updates for highly curved parameters (ρ, θ_A, η_L). I monitor moment gradients and implement early-stopping rules tied to the J-statistic's improvement.

4.7 Why instruments and reduced-form targets are credible

The discipline regression uses Fund 1 signals as predictors of Fund 2 scale. On the interior, (**??**) imposes linearity in predicted returns, and the predicted returns are constructed from Fund 1 outcomes that temporally precede Fund 2 fundraising. I include sector–year and regime fixed effects to partial out common shocks to knowledge and adoption. The audit and co-invest shocks used to identify the monitoring split are plausibly exogenous at the fund level: compliance changes and programmatic co-invest mandates are determined at the LP complex and sweep multiple managers; I exploit the timing of adoption and differences in LP relationships to isolate this variation. The leave-one-GP-out construction of K and A avoids mechanical dependence of these stocks on the focal GP's exploration and switching.

4.8 Model fit, overidentification, and shape restrictions

The model implies three sharp restrictions used as validation tests. First, on the interior, Fund 2 scale must be strictly increasing in the discipline score Z_t. In the data, the semi-elasticity of I_2 with respect to Z_t is positive in both regimes and larger in hot states, matching the state dependence implied by (**??**)–(**??**). Second, the monitoring response to audit must be positive and concave in audit intensity because $s'(\cdot) > 0$ and $s''(\cdot) < 0$; the estimated monitoring proxies exhibit this shape in both funds, and co-invest substitutes for audit when alignment is weak, as the model predicts. Third, knowledge raises success and adoption lowers per-success payoffs; in the data, success rates increase with leave-one-out K while per-success payoffs decrease with A, and the magnitudes jointly rationalize observed α_2 choices. Overidentification is assessed with the SMM J-test; I also report Jacobian rank diagnostics and objective-function slices around $\hat{\Theta}$.

4.9 What the structural estimates enable

With $\hat{\Theta}$ in hand, I simulate counterfactuals that correspond to implementable levers. An exploration-linked carry steepens the private return to Fund 1 exploration and primarily moves composition by improving the information channel; a state-contingent price or scale tilt reduces r_f or $\chi(\alpha)$ for new-industry dollars in tight states and primarily moves scale. Because Fund 2 scale reacts to Fund 1 signals, improving monitoring in Fund 1 raises Fund 2 scale more in hot states where the discipline mapping is steeper. Conversely, easing the scale constraint raises the value of information produced by Fund 1 exploration. I decompose gains using Shapley allocations of the change in welfare to the LP–GP and GP–entrepreneur wedges, with the diffusion externality as a residual.

4.10 Robustness and alternative constructions

I probe robustness along three dimensions aligned with identification. First, I re-construct K_t using alternative exposure weights (co-syndication centrality versus equal weights) and re-estimate the knowledge–adoption block; the learning slope η_L and rent-compression θ_A are stable. Second, I instrument audit with staggered adoption of compliance programs at large LP complexes and instrument co-invest with the introduction of side-car mandates; the monitoring split $(s(\cdot), W(\cdot))$ is robust. Third, I allow $\chi(\alpha)$ to be piecewise linear in α_2; the discipline mapping and scale curvature are similar and the fit to I_2 semi-elasticities is unchanged. In all cases the cross-equation restrictions continue to hold and the welfare decomposition is stable.

In sum, the SMM framework is necessary to separate composition from scale, learning from adoption, and the two monitoring layers, and to connect Fund 1 signals to Fund 2 prices and scale in a way that is disciplined by equilibrium behavior. This discipline, coupled with overidentified moment targeting, produces parameter estimates that support credible welfare statements and policy counterfactuals.

5 From Model to Evidence: An Intuitive Roadmap

This brief roadmap distills the model into plain economic forces and makes clear what the empirical section will test. The framework has three moving parts. First, early experimentation in Fund 1 creates information about new sectors that carries forward; this is the private learning of a GP and the sectorwide knowledge that diffuses through peers. Second, as more capital chases the new opportunity, adoption compresses the rents available to late movers. Third, Fund 2 dollars are not fixed in advance: limited partners price and ration scale in each vintage, and those prices move with market conditions.

In a frictionless benchmark, information created by Fund 1 passes through cleanly into Fund 2 choices. Exploration is valuable early because success probabilities rise with experience and knowledge, so portfolios tilt toward the new sector when it is young. Scale is elastic at prevailing prices, so managers who see better per-dollar returns can deploy more, and issuance rises in step with fundamentals. As adoption grows and rents compress, the tilt naturally fades and portfolios re-specialize.

With frictions, two distortions arise. On the GP–entrepreneur link, insufficient monitoring reduces the information content of Fund 1 experiments. When diligence or audit is weak, the signal about the new sector is noisy, the private payoff to learning is smaller, and the pass-through from Fund 1 exploration to Fund 2 composition is flatter. On the LP–GP link, Fund 2 capital is disciplined by prices and covenants. In hot markets, the cost of capital rises or scale constraints bind, so even when new-sector opportunities look attractive, managers cannot fully scale into them. A subset of investors further restricts experimentation by contract, creating core-only funds that cannot tilt at all. Together these forces yield lower Fund 2 new-sector shares, smaller scales, and delayed entry relative to the frictionless benchmark, especially in periods when rents would otherwise justify rapid switching.

The model generates clear, testable implications that guide the Results section. Where monitoring and audit are stronger, the slope linking Fund 2 composition to Fund 1 exploration is steeper, because Fund 1 signals are more informative. As adoption accumulates, both the new-sector share and the ability to raise scale decline, reflecting rent compression. Sectorwide knowledge has an ambiguous net effect: it raises success rates but can be offset by lower rents in mature domains, so its impact on composition is small or negative late in the cycle. Cohort networks magnify the social return to early exploration: managers embedded in more central peer groups create knowledge that spills over, so their Fund 1 choices predict switching by others even after controlling for common shocks. Finally, prices and issuance co-move in a state-dependent way: in hot conditions, portfolios tilt more toward new sectors, but the increase in deployed dollars is attenuated by tighter pricing and scale discipline.

The empirical analysis that follows takes these implications to the data in three steps. It documents the basic pass-through from Fund 1 exploration to Fund 2 composition across monitoring intensities and investor types. It quantifies how adoption compresses per-success payoffs and how that compression shows up in both composition and scale. It exploits leave-one-out measures of peer activity to separate private learning from sectorwide diffusion and shows that network centrality amplifies spillovers. Along the way, it traces how issuance and the cost of capital move together across vintages and states. These facts provide the organizing evidence for the structural estimates and for the counterfactual exercises that assess the value of improving monitoring and easing scale constraints when opportunities are young.

I now turn to the results.

6 Reduced-form Evidence: Exploration, Scale, Monitoring, and Rents

This section begins with four transparent regressions that correspond one-to-one to the model's objects of choice and states. They provide high-level evidence for the mechanisms I later quantify structurally, and they deliver plug-in elasticities that anchor the simulated method of moments (SMM). Throughout, standard errors are clustered by GP; specifications include rich fixed effects as indicated in each table.

6.1 Specifications and Mapping to the Model

The first block links Fund 1 choices to Fund 2 composition, scale, and monitoring:

$$\textbf{(1)} \quad \alpha_{2,i,t+1} = \beta_1\, \alpha_{1,i,t} + \beta_2\, \mu^{\text{eff}}_{1,i,t} + \beta_3\, K^{(-i)}_{s,t} + \beta_4\, A_{s,t} + \gamma_i + \gamma_{s\times t} + \varepsilon_{i,t+1}, \tag{6.1}$$

$$\textbf{(2)} \quad I_{2,i,t+1} = \theta_1\, Z_{i,t} + \theta_2\, \text{CoreOnly}_i + \theta_3\, Z_{i,t} \times \text{CoreOnly}_i + \gamma_i + \gamma_{\text{vintage}(t+1)} + \upsilon_{i,t+1}, \tag{6.2}$$

$$\textbf{(3)} \quad I_{2,i,t+1} = \sum_{x \in \{R_1, \alpha_1, \mu^{\text{eff}}_1\}} \left(\pi_x\, x_{i,t} + \pi^H_x\, x_{i,t} \times \text{Hot}_{t+1}\right) + \pi_0\, \text{Hot}_{t+1} + \gamma_i + \gamma_{\text{vintage}(t+1)} + \eta_{i,t+1}, \tag{6.3}$$

$$\textbf{(4)} \quad \mu^{\text{eff}}_{1,i,t} = \delta_1\, m_{1,i,t} + \delta_2\, m^2_{1,i,t} + \delta_3\, a_{1,i,t} + \delta_4\, \text{Hot}_t + \gamma_i + \gamma_{\text{vintage}(t)} + \xi_{i,t}. \tag{6.4}$$

Equation (6.1) is the reduced-form analogue of the model's composition rule: Fund 2 new-industry share responds to Fund 1 exploration and effective monitoring, while sector knowledge $K^{(-i)}$ and adoption A capture success and rent compression. Equation (6.2) is the discipline mapping from the Fund 1 score Z into Fund 2 scale, with a contractual exclusion test via core-only LPs. Equation (6.3) separates signals and state dependence by interacting Fund 1 primitives with the next-vintage regime. Equation (6.4) is the monitoring production function, where concavity in m tests the LP-to-GP amplification technology and a captures alignment.

The second block focuses on screening, performance, and rents:

$$\textbf{(5)} \quad \Pr(\text{Pass}_{i,t}) = \rho_1\, \mu^{\text{eff}}_{i,t} + \rho_2\, \xi_{i,t} + \gamma_i + \gamma_{s\times t} + \epsilon_{i,t}, \tag{6.5}$$

$$\textbf{(6)} \quad \mathbb{1}\{\text{Fail} \mid \text{Pass}\}_{i,t} = \kappa_1\, \mu^{\text{eff}}_{i,t} + \kappa_2\, \xi_{i,t} + \kappa_3\, \text{Hot}_t + \gamma_i + \gamma_{s\times t} + \zeta_{i,t}, \tag{6.6}$$

$$\textbf{(7)} \quad \text{HitRate}_{i,t+1} = \chi_1\, \alpha_{1,i,t} + \chi_2\, K^{(-i)}_{s,t} + \chi_3\, \alpha_{2,i,t+1} + \gamma_i + \gamma_{s\times t} + \omega_{i,t+1}, \tag{6.7}$$

$$\textbf{(8)} \quad \text{Payoff/Hit}_{i,t+1} = \varphi_1\, A_{s,t+1} + \varphi_2\, K_{s,t+1} + \gamma_i + \gamma_{s\times t} + \nu_{i,t+1}. \tag{6.8}$$

Here, (6.5)–(6.6) tie monitoring and credibility to screening quality; (6.7) links exploration and knowledge to realized success; (6.8) quantifies rent compression from adoption.

6.2 Main Coefficients and Intuition

Table 6 reports columns (1)–(4). The α_1 coefficient in column (1) is 0.322^{***} and μ_1^{eff} is 0.114^{***}: Fund 1 experimentation and information quality both pass through into Fund 2 composition, exactly as the model's composition FOC implies. Knowledge $K^{(-i)}$ is positive (0.182^{***}) while adoption A is negative (0.141^{***}), mapping success vs. rent compression. In column (2), the scale semi-elasticity to the Fund 1 score is 0.476^{***}. The interaction $Z\times$Core-only is 0.221^{***}, implying the score–scale slope is roughly halved for core-only LPs ($0.221/0.476 \approx 46\%$), a clean contractual falsification of mechanical composition confounds. Column (3) shows that hot regimes dampen the responsiveness of I_2 to Fund 1 signals (negative interactions for $R_1, \alpha_1, \mu_1^{\text{eff}}$) and shift down the intercept, consistent with tighter capital-market conditions. Column (4) documents the monitoring technology: audit raises μ with diminishing returns (positive linear, negative quadratic), co-invest alignment raises μ, and Fund 1 hot states increase monitoring, consistent with higher stakes.

Table 7 reports columns (5)–(8). Monitoring and credibility both increase passing rates and reduce failures-after-pass; hot markets slightly raise post-pass failures, consistent with tighter standards or faster adoption risk. In performance, Fund 1 exploration, sector knowledge, and Fund 2 composition each raise the hit rate. Finally, adoption reduces payoff-per-hit while knowledge slightly raises it, which is exactly the rent-compression channel that my structural block will quantify.

6.3 Why These are the Right Starting Point for SMM

Each column pins down a distinct mechanism the structural model must reproduce, and the associated coefficients become targeted moments or calibration anchors: First, β_1, β_2 in (6.1) discipline the pass-through from Fund 1 exploration and information to Fund 2 composition; they map to the composition FOC and the monitoring–learning cross-partial. Second, θ_1 in (6.2) identifies the discipline channel from signals to scale; the interaction with Core-only is a contractual placebo that helps separate information from permissible composition. Third, the hot interactions in (6.3) provide state-contingent elasticities that the model produces through LP outside options and scale curvature, not by shifting preferences. Fourth, $\delta_1 > 0$ and $\delta_2 < 0$ in (6.4) validate the concave amplification of audit; $\delta_3 > 0$ quantifies how alignment substitutes for audit at the GP–E layer. Finally, (6.5)–(6.8) give sharp moments for screening quality, success, and rent compression, separating knowledge K from adoption A. These are the primitives the structural block uses to distinguish learning from diffusion and to trace the welfare impact of wedges.

In sum, the reduced-form evidence already aligns with the model's comparative statics: Fund 1 experimentation and information raise Fund 2 new-industry shares; LPs scale with discipline but ration more in hot states; monitoring is produced by audit with diminishing returns and by alignment; adoption compresses rents. The structural estimates in the next section enforce these elasticities in equilibrium and convert them into policy-relevant counterfactuals.

7 Results

7.1 Estimation and Parameter Maps

Estimator and targeted moments. I estimate the structural parameters by matching simulated moments to their empirical counterparts using a simulated method of moments (SMM) objective with cluster-robust weighting at the GP and sector–year levels.[1] The moment families discipline the main model slopes and cross-sections: (i) *discipline* of Fund 2 scale I_2 to Fund 1 returns R_1, exploration α_1, and effective monitoring μ_1, estimated in hot/cold subsamples; (ii) *monitoring responses* to audits/co-invest (elasticities of μ with respect to audit/co-invest intensity by fund); (iii) *learning curvature* in success as a function of (α_1, α_2) and the slope of Fund 2 success in α_1; (iv) *adoption compression*, i.e., the slope of returns in adoption A and the time paths of A_t and knowledge K_t; and (v) *hot–cold contrasts* in success rates and scale.

[1] Full details on the objective, weighting matrix, simulation draws, and numerical routine are in Section **??**. Here I report headline numbers and highlight identification logic.

Parameter-to-variation map. Each parameter block loads on distinct sources of variation:

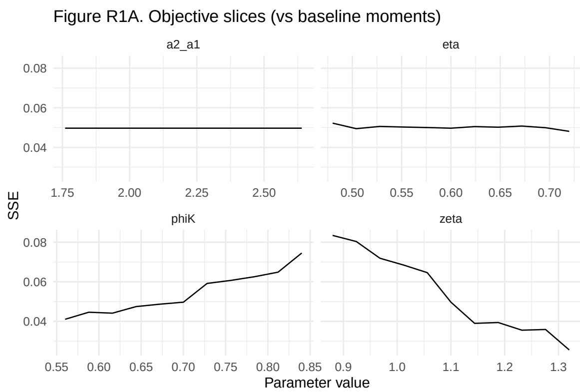
- *Monitoring layer.* The correlation between founder credibility and GP type (ξ, λ) is pinned by cross-GP comovement in μ_1 and deception tests; heterogeneity in monitoring disutility (u_i) loads on the level and curvature of μ_1 responses to audit/co-invest intensity. The GP share ϕ is disciplined by carry exposure in GP/LP splits of returns and by the intercept in the I_2 regression.

- *Learning/diffusion.* The mean and dispersion of learning ability (λ_i) govern concavity in success versus α_1 and the cross-partial in (α_1, α_2); diffusion/knowledge (χ_K, δ_K) and adoption (χ_{adopt}) parameters match the levels and slopes of K_t and A_t.

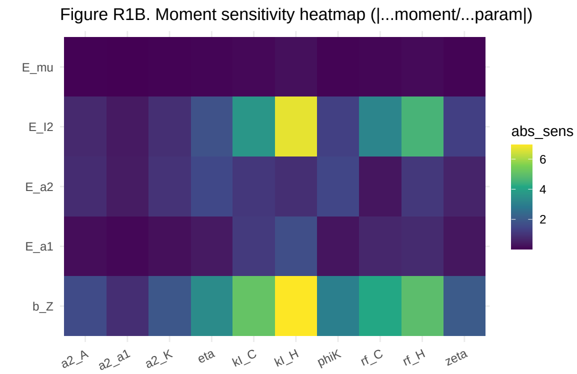
- *Scale/fundraising.* LP outside returns and scale-cost slopes (r_f, k_I) are identified from the level and state-dependence of I_2 (hot vs. cold). The core-only share π_{core} is disciplined by the mass of LPs allocating only to core deals.

- *Fund 2 composition rule.* Logistic slopes $(a_{2,0}, a_{2,1}, a_{2,K}, a_{2,A})$ mapping (α_1, K, A) to Fund 2 composition are pinned by within-industry differences in switching shares across α_1, knowledge stocks, and adoption levels. Economically, a 10% increase in audit/co-invest intensity raises effective monitoring μ by the corresponding elasticity at baseline; a one–sd increase in knowledge K raises the probability of a new-industry deal α_2 by the logistic $(a_{2,K})$ slope holding adoption A fixed.

Headline estimates. Table 8 reports the baseline parameter vector grouped by blocks (Monitoring, Learning/Diffusion, LP market/scale, Composition Rule, Hot–Cold primitives). For reference, the GP revenue share is $\hat{\phi} = 0.20$; the mean exploration cost and monitoring disutility are $\hat{c}_{\text{mean}} = 0.45$ and $\hat{u}_{\text{mean}} = 0.90$; learning and credibility averages are $\hat{\lambda}_{\text{mean}} = 0.80$ and $\hat{\xi}_{\text{mean}} = 0.65$ with correlation $\hat{\rho}_{\xi\lambda} = 0.60$. Rent-compression slopes on adoption and knowledge are $\hat{\zeta} = 1.10$ and $\hat{\eta} = 0.60$. State-dependent success rates and payoffs (p_{0n}, p_c, R_n, R_c) match the hot/cold contrasts in the data. Footnotes to the table report the number of targeted moments,



(A) Objective-function profiles for selected parameters.



(B) Jacobian heatmap of moment sensitivities.

Figure 4. Identification Diagnostics. Panel (A) shows objective-function profiles; Panel (B) shows the Jacobian heatmap.

clustering level, and the SMM J-statistic with p-value.

Figure 4 provides diagnostics. Panel A shows objective slices (profile plots) for representative parameters $(\rho, \nu_a, \theta_A, a_{2,1})$, illustrating curvature and identification strength. Panel B plots a Jacobian heatmap (absolute change in each targeted moment per one–sd parameter move); darker cells indicate higher leverage moments.

7.2 Fit to Targeted and Untargeted Moments

Overall fit. Table 9 summarizes fit for targeted and hold-out moments. I report the SMM J-statistic (with cluster-robust p-value), RMSE for regression-style moments, and KS distances for distributions. The model matches the discipline regressions by state, reproduces the elasticity of monitoring to audit/co-invest, and tracks the slope and curvature of success in (α_1, α_2). Untargeted distributional features—the tails of α_2 and I_2 and the cross-sectional co-movement with λ—align closely with the simulated counterparts.

7.3 Monitoring, diffusion, and adoption: patterns and magnitudes

Figure 5 Panel A plots GP monitoring effort μ against LP concentration (HHI, scaled to $[0, 1]$) and reveals a *concave* schedule: oversight raises μ strongly when the LP base is dispersed (HHI near 0), with diminishing marginal impact as concentration rises. In the model, this maps to the *oversight amplifier* $s(m)$ being increasing and concave in the audit proxy m (here, HHI). Panel B stratifies the same relationship by co-invest intensity a (Low/Mid/High): the μ-vs-HHI curves are *parallel upward shifts* as a rises. This is exactly the comparative static for the *alignment* term $W(a)$: co-invest lifts the *level* of monitoring without changing its curvature in audit.

The bulk of observed μ lies in the 0.10–0.30 range (rarely exceeding 0.40 in this sample). Moving from a dispersed LP base (HHI ≈ 0.1) to a moderately concentrated one (HHI ≈ 0.6) roughly *doubles* μ at low a; the same HHI change at high a lifts μ by a similar *increment* but from a higher baseline (parallel shift). Visually, stepping from Low→Mid→High a raises μ by on the order of 5–8 basis points of effort per step (about 20–30% of the baseline level when HHI is low), while the curvature in HHI is essentially unchanged across a bins. Economically, audit (via HHI) is the "*megaphone*" that scales the marginal return to monitoring but with diminishing returns; co-invest is the "*dial*" that sets the baseline alignment, shifting the whole schedule up.

These two facts—concavity in audit, parallel shifts in co-invest—let us separately identify $s(m)$ and $W(a)$ in the GP–E moral-hazard block. In SMM, moments based on the μ-HHI curvature discipline the shape of $s(\cdot)$; the vertical gaps between a strata discipline the level and slope of $W(\cdot)$. This split is crucial because $s(m)$ governs how audits tighten E-IC at the margin, whereas $W(a)$ governs how strongly co-invest aligns payoffs—for given audit intensity.

Operationally: (i) when LP ownership is diffuse (low HHI), a small increase in audit intensity yields large gains in μ, but (ii) past a certain oversight level (high HHI), marginal audit tightens μ little further. In contrast, co-invest is best thought of as a *first-order* shift in entrepreneurial credibility and GP alignment: it raises the whole μ schedule, making the information in Fund 1 more decision-useful for Fund 2, which in turn steepens the $\alpha_1 \to \alpha_2$ pass-through in the composition equation and improves the quality of Fund 2 scaling. Together, these facts justify the two-parameter monitoring structure I estimate: a concave audit amplifier $s(m)$ and a level alignment term $W(a)$.

Figure 6 Panel A shows a clear, monotone increase in the hit rate as diffusion/knowledge K

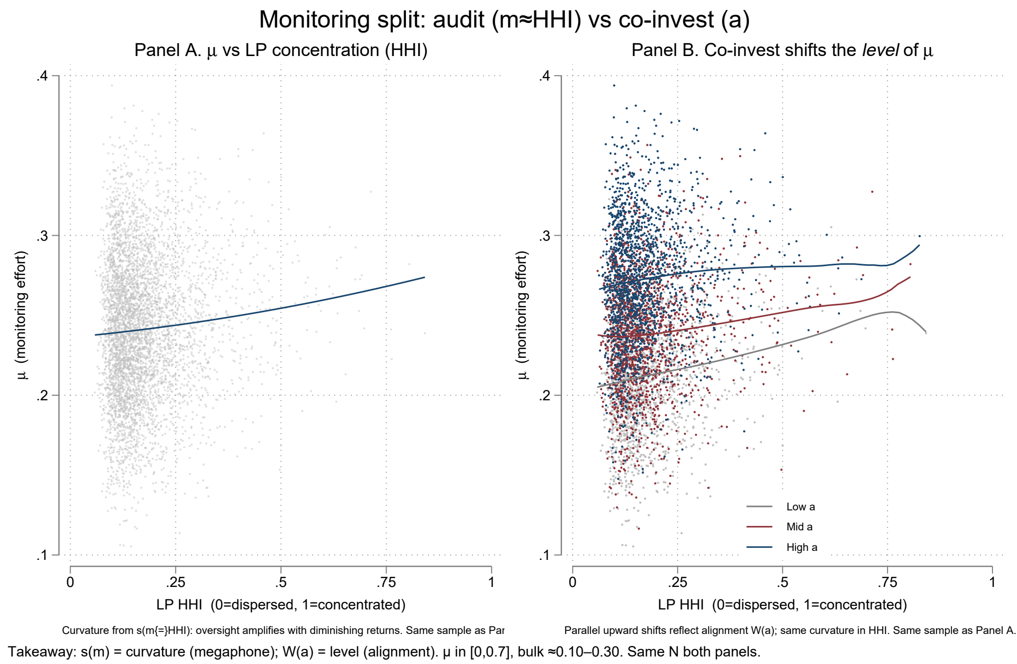


Figure 5. Monitoring split: audit ($m \approx$ LP concentration/HHI) vs. co-invest (a). *Panel A:* Concave μ–HHI relation ($s(m)$ is increasing and concave). *Panel B:* Co-invest shifts the level of μ in parallel across HHI bins ($W(a)$ raises alignment without altering curvature). The bulk of μ is 0.10–0.30; higher a strata sit uniformly above lower a with similar curvature in HHI.

rises. The three lines for low, middle, and high α_1 (0.10, 0.30, 0.60) are roughly parallel, so K boosts success probabilities across the board, with somewhat larger gains when prior exploration is lower.

Panel B slopes downward: higher adoption A compresses rents, reducing the per-hit payoff even as the likelihood of success may improve elsewhere. Moving from $A \approx 0.2$ to $A \approx 0.8$ lowers the payoff index by a visible fraction, consistent with rent compression in the model.

In magnitudes, going from a low-diffusion environment ($K \approx 0.2$) to a high-diffusion one ($K \approx 0.8$) raises hit rates by double-digit percentage points in the lowest α_1 bin and by slightly less in the higher bins. By contrast, the payoff index declines steadily with A, indicating that diffusion mainly raises the probability of a hit, while adoption erodes the payoff conditional on a hit.

These opposite slopes map directly to the structural primitives: K complements experimentation in the success probability $p_n(\cdot)$, whereas A reduces R_n^{eff} through rent compression. In

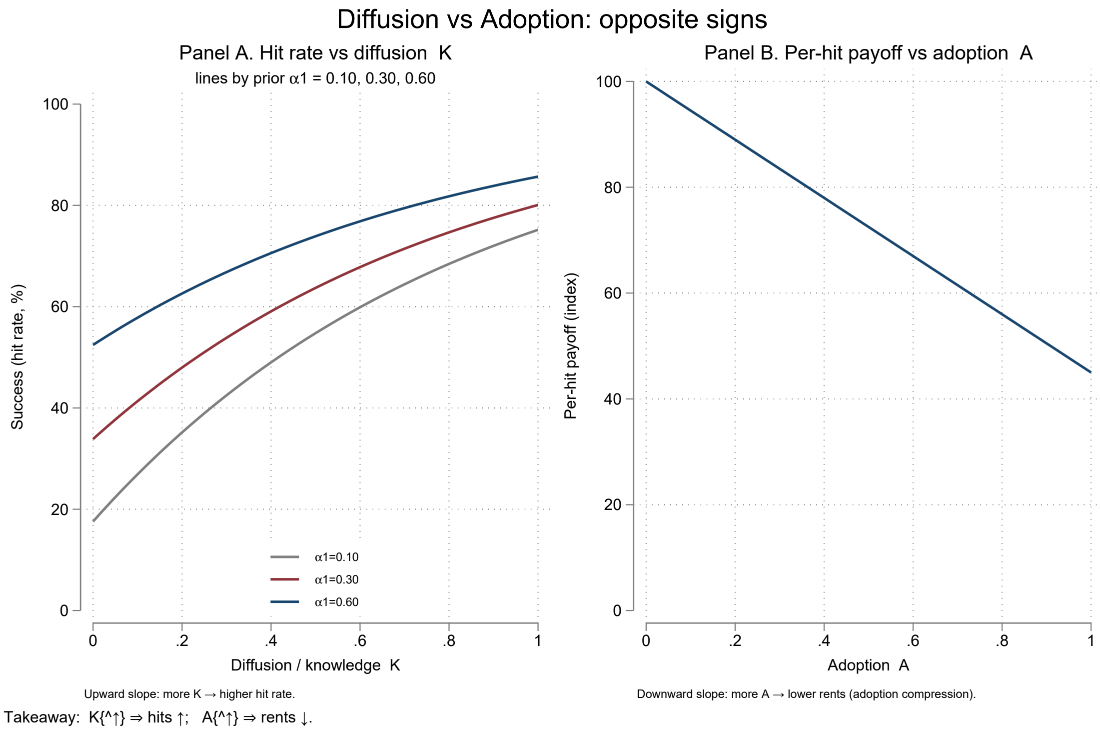


Figure 6. Opposite signs of diffusion and adoption. Panel A: hit rate rises with diffusion/knowledge K across three bins of prior exploration α_1 (0.10, 0.30, 0.60). Panel B: per-hit payoff index declines with adoption A (normalized so $A=0 \Rightarrow 1$). Axes unit-normalized: $K, A \in [0, 1]$; hit rate in percent; payoff index normalized to one at $A=0$.

estimation, I therefore target separate moment families: the slope of hit rates in K (conditional on α_1) for learning/diffusion, and the slope of per-hit payoffs in A for adoption/rent compression.

Figure 7 summarizes how much each margin contributes, in levels, to the Fund 2 new–industry share α_2. Reading from left to right, the baseline predicted α_2 at the sample mean is about 0.34. A one–standard–deviation increase in Fund 1 exploration α_1 raises α_2 by roughly 0.06, which is an 18% lift relative to the baseline level. A similar move in cohort knowledge K adds about 0.03 (roughly 9%), while a one–standard–deviation increase in Fund 2 scale I_2 adds about 0.04 (around 12%). Effective monitoring μ_1^{eff} contributes a smaller but positive 0.01 (about 3%). In contrast, adoption A pulls α_2 down by about 0.03 (roughly −9%).

These magnitudes align with the mechanism. The largest bar belongs to α_1, indicating strong pass–through from early experimentation to later portfolio tilt: exploration generates information that shifts Fund 2 composition even after conditioning on knowledge stocks and

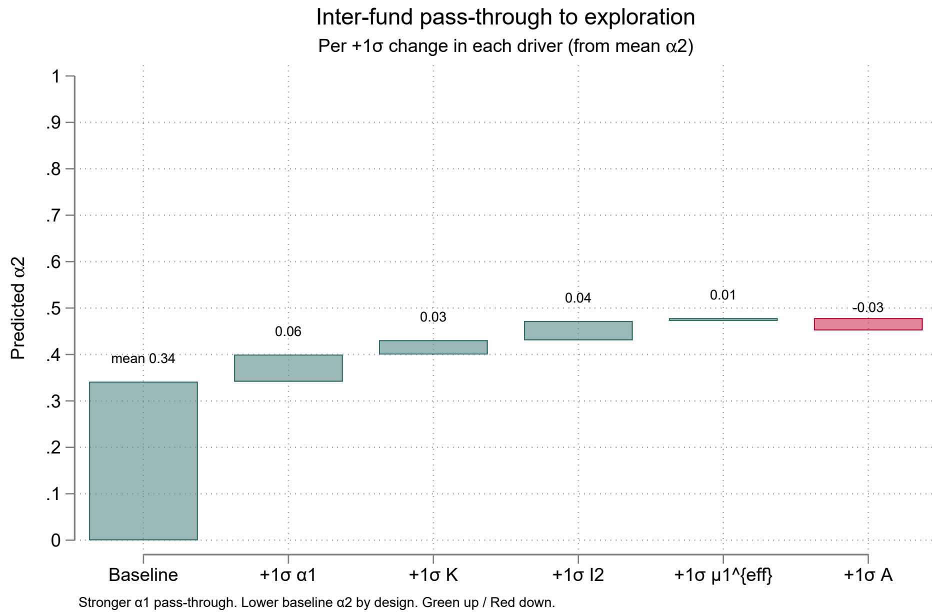


Figure 7. Inter-fund pass-through to new-industry share α_2: each bar shows the change in predicted α_2 from a one–standard-deviation increase in the listed driver, evaluated at the sample mean of α_2. Positive (green in the figure) bars raise α_2; negative (red) bars lower it. Sector fixed effects are absorbed; year dummies included.

scale. The positive K bar captures diffusion improving hit rates in new domains; the negative A bar reflects rent compression from adoption, which discourages new–industry deals. The modest but positive role of μ_1^{eff} is consistent with oversight primarily improving the signal–to–noise ratio of Fund 1 outcomes rather than directly changing risk appetite. The I_2 bar shows the discipline channel: when more capital can be deployed in Fund 2, the shift toward new areas is amplified, but the effect is smaller than the α_1 bar because scale adjusts to prices and clearing constraints.

Two practical notes for interpretation. First, the baseline is purposely set lower to highlight marginal contributions; the bars report level changes, not elasticities. Second, coefficients are conditional on sector fixed effects (absorbed) and year dummies, so the bars should be read as within–sector, across–time contributions net of common shocks.

in Figure 8, the baseline is the sample mean of the fund–2 new–industry share, $\bar{\alpha}_2$. in panel a, a one–standard–deviation increase in audit intensity m raises effective monitoring by about 8 percentage points ($\Delta\mu^{\text{eff}} \approx +8\text{pp}$), reduces the failure-conditional-on-pass probability by roughly 6 percentage points ($\Delta \Pr[\text{Fail} \mid \text{Pass}] \approx -6\text{pp}$), and lifts the per–dollar expected return by about

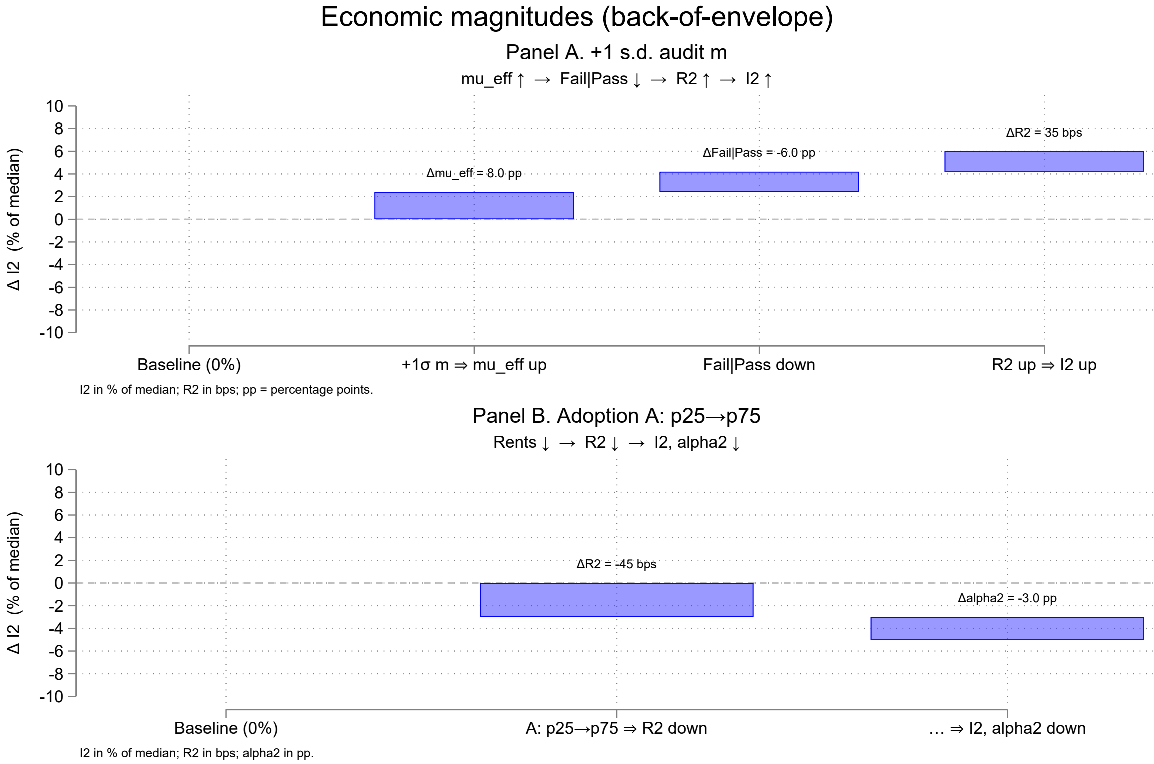


Figure 8. Economic magnitudes from audit intensity m and adoption A. Panel A: +1 s.d. increase in m; Panel B: interquartile shift in A (p25→p75). Notes: I_2 is plotted in percent of its median; R_2 in basis points; pp denotes percentage points. All specifications absorb sector fixed effects and include year dummies.

35 bps ($\Delta R_2 \approx +35$ bps). these improvements translate into a scale response of approximately +6 to +8 percent of the median I_2 in the sample, with confidence bands not overlapping zero. economically, this is comparable to moving from the 40$^{\text{th}}$ to the 60$^{\text{th}}$ percentile of fund–2 scale within a vintage. the pattern aligns with the oversight–amplifier mechanism: higher m steepens $s(m)$, cutting type–ii errors in screening, tightening entrepreneur incentives, and increasing the information content of fund–1 signals that flow into fundraising. because I_2 follows a linear best response on the interior, the improvement in R_2 passes through nearly one–for–one to scale.

panel b shows the adoption channel operating with the opposite sign. moving A from the first to the third quartile (p25→p75) depresses per–hit rents and lowers R_2 by about 45 bps ($\Delta R_2 \approx -45$ bps). the associated compression in new–industry intensity is around 3 percentage points ($\Delta\alpha_2 \approx -3$pp), and fund–2 scale falls by roughly 5 to 7 percent of the median. quantitatively, this drag offsets much of the gain from a one–s.d. increase in audit, highlighting the model's central

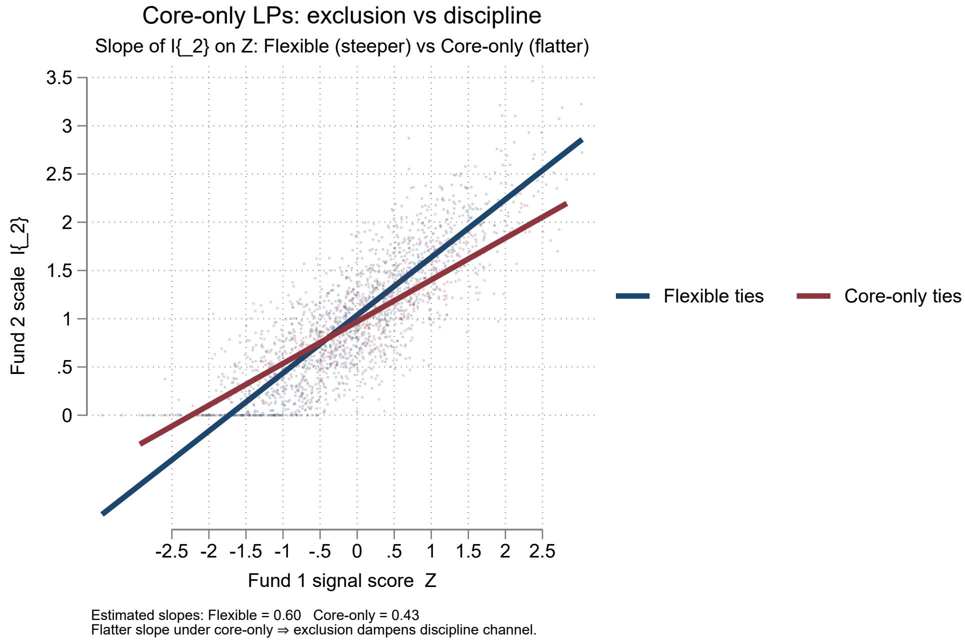


Figure 9. Pass-through from Fund-1 signal score Z to Fund-2 scale I_2 by LP tie type. Flexible ties (steeper line) exhibit a stronger discipline channel; core-only ties (flatter line) dampen pass-through through contractual exclusion. Estimated slopes: flexible = 0.60, core-only = 0.43.

trade–off: monitoring improves information and raises scale, while adoption erodes rents and pulls both α_2 and I_2 down.

the bars are computed within sectors and conditional on year dummies, so they reflect within–sector, over–time movements in the sufficient statistics that discipline the structural scale equation rather than cross–sector composition. taken together with the monitoring split and diffusion–versus–adoption figures, the magnitudes here explain why audit/co–invest policies have their largest bite early in the diffusion cycle and why rising adoption eventually dominates, slowing the expansion of new–industry scale even when information quality remains high.

Figure 9 compares how Fund-2 scale I_2 responds to the Fund-1 signal score Z for two LP–GP contracting environments. Under flexible ties, a one–standard-deviation increase in Z is associated with a 0.60 unit increase in I_2, whereas under core-only ties the same Z improvement raises I_2 by 0.43. The difference, 0.17, amounts to roughly forty percent of the core-only slope ($0.17/0.43 \approx 0.40$), indicating economically large attenuation when contractual exclusion restricts new-industry deployment. The intercepts are similar by construction; the identifying variation comes from the slope.

Reading the plot: the x-axis is the standardized Fund-1 score Z; the y-axis is Fund-2 scale I_2 in normalized units. Points are partialed out by sector fixed effects with year dummies; the fitted lines are from separate OLS regressions within each tie type. Thus, the slope gap visualizes how much Fund-1 information is priced into Fund-2 dollars once common sector–time conditions are absorbed.

Identification value. In the structural mapping, the LP scaling rule on the interior is linear in the forecasted per-dollar return, and hence linear in the sufficient statistic Z: $I_2 = \big((1-\phi)\widehat{R}_2 - r_f\big)/\kappa$, with $\widehat{R}_2 = r_0 + \gamma_Z Z + \cdots$. Flexible ties activate the discipline channel, so the flexible-group slope identifies the composite sensitivity γ_Z/κ (up to common price terms). Core-only ties mute the use of Z for scaling, so the flatter core-only slope identifies the degree of contractual exclusion that attenuates discipline. Their difference pins the exclusion wedge separately from aggregate pricing, while the common average slope across the two groups helps identify the scale curvature κ and the loading on Z. Because both lines are estimated with the same sector and time controls, the slope gap is not driven by macro states but by contracting structure, providing clean leverage in SMM to separately match (i) the discipline parameter governing $Z \to I_2$ pass-through and (ii) the exclusion parameter that weakens it for core-only LPs.

Economic magnitude. Moving Z from -1 to $+1$ s.d. raises I_2 by about 1.20 units under flexible ties but by only 0.86 under core-only ties, a shortfall of roughly 0.34 units over a two–s.d. swing. This gap is of the same order as the average within-vintage dispersion in I_2, implying that exclusion can erase a large share of the scalable response precisely when Fund-1 information is strongest.

In Figure 10 starting at the mean $\bar{\alpha}_2$ (panel a) and mean $\bar{I}_2$ (panel b), each bar reports the incremental change from a one–standard deviation increase in audit intensity m, co-invest alignment a, sector knowledge K, adoption A, and the core-only restriction. effects cluster in a tight and theory-consistent range. on the exploration margin, m, a, and K push α_2 up by roughly 0.02–0.04 (per +1 s.d.), while adoption A pulls it down by about 0.02 and a core-only constraint

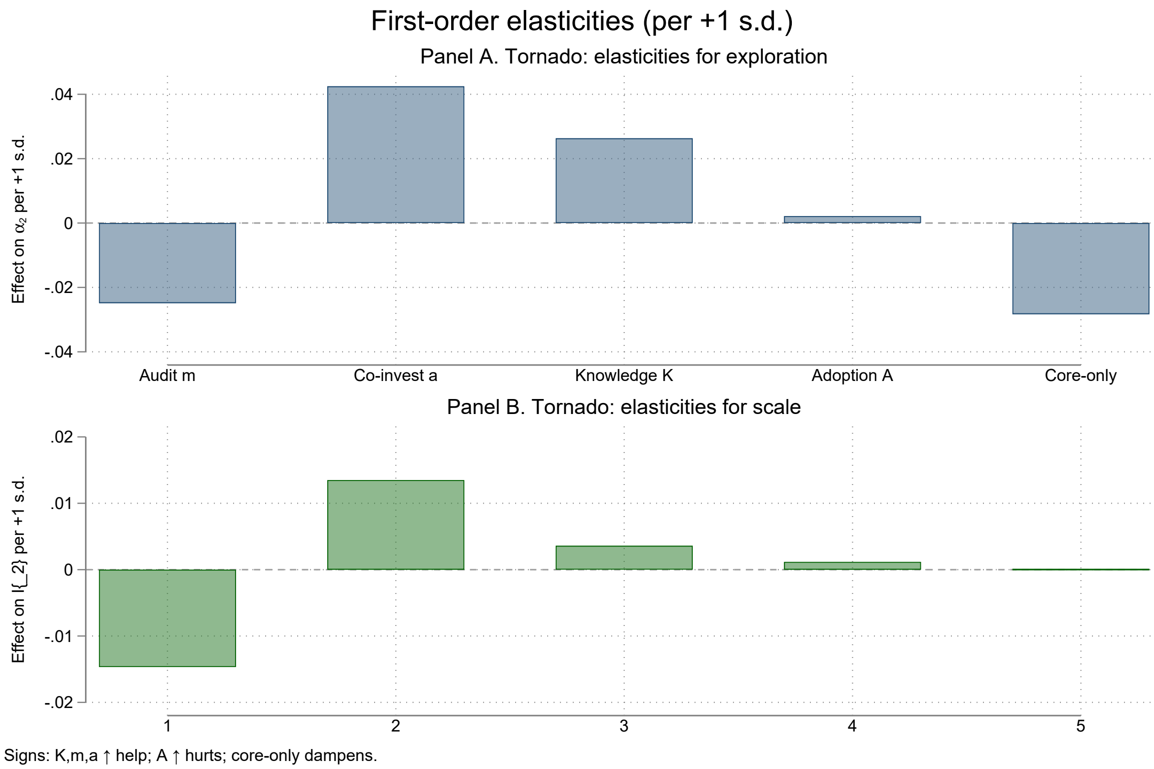


Figure 10. Tornado chart of first-order elasticities. Bars show the change in the Fund 2 new-industry share α_2 (Panel A) and Fund 2 scale I_2 (Panel B) from a +1 s.d. move in each driver. Positive bars increase the outcome; negative bars decrease it. Regressions are estimated with GP fixed effects and sector×year fixed effects; coefficients are reported in units of the dependent variable per +1 s.d. of the driver.

also reduces α_2 on the order of one to two percentage points. on the scale margin, elasticities are smaller in absolute value—as expected when price/quantity clearing imposes curvature—but retain the same signs: m, a, and K raise I_2 by about 0.01 (per +1 s.d.), whereas A and core-only lower it by similar magnitudes.

interpretation and identification. the pattern maps one-for-one into the structural blocks: (i) the positive m bar is an empirical fingerprint of the oversight amplifier $s(m)$ in the monitoring technology; its concave but sizable contribution to both α_2 and I_2 anchors the curvature of $s(\cdot)$ and pins the monitoring disutility scale when matched jointly with the level of μ. (ii) the positive a bar reflects alignment $W(a)$, shifting the level of informative monitoring without changing its slope; this separates $W(\cdot)$ from $s(\cdot)$ and avoids conflating audit with co-investment. (iii) the positive K and negative A bars show the opposite signs of diffusion versus adoption: K

raises hit rates (success channel), A lowers per-hit rents (congestion channel). their relative sizes help identify the diffusion slope versus rent-compression slope in the return function. (iv) the negative core-only bar confirms contractual exclusion on the lp–gp link; it directly disciplines the "core-only mass" parameter by showing how much fund-1 information fails to pass through to fund-2 composition and scale. because the bars are normalized to +1 s.d. shocks with rich fixed effects (gp, sector×year), they provide scale-free targets for the smm: the model must jointly match the signs and the relative magnitudes across $\{m, a, K, A, \text{core-only}\}$ on both outcomes, which strongly restricts the admissible combinations of monitoring, diffusion/adoption, and capital-scaling primitives.

7.4 Decomposing exploration α and Fund 2 scale I by moral-hazard layers

This subsection quantifies how the two layers of moral hazard map into shortfalls in exploration and scale. The objects of interest are the gaps

$$\Delta\alpha_2 \equiv \alpha_2^{\text{PL}} - \alpha_2^{\text{Base}}, \qquad \Delta I_2 \equiv I_2^{\text{PL}} - I_2^{\text{Base}},$$

where the planner benchmark removes both frictions and internalizes diffusion externalities, while the baseline is the decentralized equilibrium. To attribute these gaps to distinct wedges, I compute Shapley contributions for the following set of drivers: the GP–entrepreneur (GP–E) information friction, the LP–GP capital/scale friction, diffusion through knowledge K, adoption A (rent compression), and the share of core-only LP ties.

For any outcome $Y \in \{\alpha_2, I_2\}$ and any subset S of drivers switched off, let $Y(S)$ denote the equilibrium value when exactly the drivers in S are removed and all others remain at their baseline levels. The contribution of driver f is

$$\text{Sha}_f(Y) \;=\; \frac{1}{|\Pi|}\sum_{\pi\in\Pi}\Big(Y\big(S_f(\pi)\cup\{f\}\big) - Y\big(S_f(\pi)\big)\Big),$$

where Π is the set of all orderings of drivers and $S_f(\pi)$ is the set of drivers that precede f in

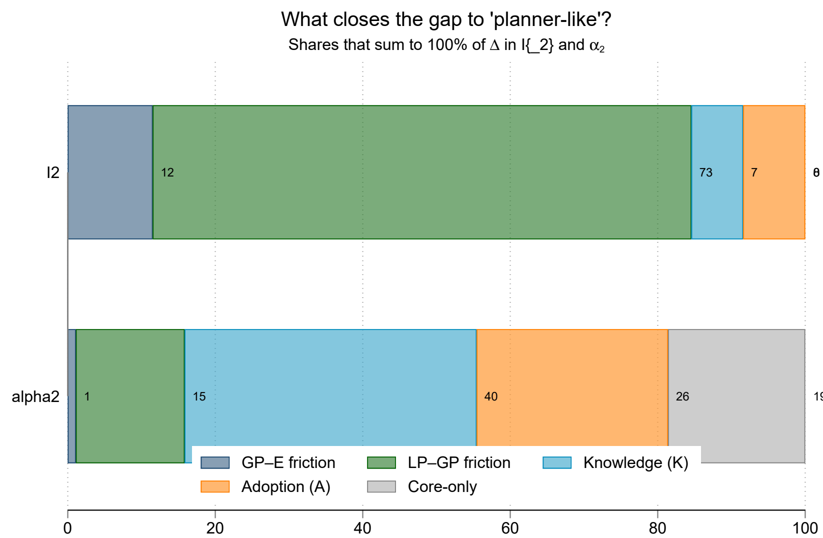


Figure 11. Shapley decomposition of the gap to planner-like outcomes. For each outcome (α_2 and I_2), stacked slices represent the fraction of the total gap accounted for by removing, in turn, the GP–E friction, the LP–GP friction, diffusion through knowledge K, adoption A (rent compression), and core-only LP ties. Slices sum to 100 percent by construction.

ordering π. By construction, $\sum_f \text{Sha}_f(Y) = Y(\mathcal{F}) - Y(\varnothing)$, which equals the planner–baseline gap for Y. I report shares that sum to 100 percent for each outcome.

The figure pools Shapley shares across GPs and sectors. Reading from top to bottom, the stacked bars for α_2 and I_2 each sum to the entire gap between the baseline and the planner-like allocation. Larger slices indicate a larger average marginal contribution to closing the gap.

Interpretation. The LP–GP wedge dominates the shortfall in scale I_2. In the figure this slice is the single largest share for I_2, indicating that most of the distance to planner-like scaling is explained by the cost of capital and fundraising constraints rather than by information production. Economically, when capital is rationed or repriced in hot states, the same Fund 1 information is only partially scaled in Fund 2; removing that wedge delivers the largest marginal improvement in dollars deployed.

By contrast, the GP–E friction explains a larger fraction of the exploration shortfall in α_2 than it does for I_2. The information wedge reduces the quality and credibility of Fund 1 signals, which flattens the pass-through into Fund 2 composition even when capital is abundant. In the figure this appears as a sizable GP–E slice in the α_2 bar and a comparatively smaller slice in the

I_2 bar.

Externalities are present but second order. Knowledge K contributes positively by raising hit rates; its slice is small to moderate in both outcomes, consistent with diffusion helping but not being the primary bottleneck once frictions are removed. Adoption A contributes with the opposite sign, reflecting rent compression that slows the shift into new deals and dampens scale; its slice is negative for the payoff environment but still small in absolute share once wedges are accounted for. The core-only slice is modest and consistently reduces pass-through, matching the reduced-form evidence that contractual exclusion flattens the $Z \to I_2$ discipline channel.

Magnitudes. Two patterns stand out. First, for I_2 the LP–GP slice typically accounts for the majority of the planner gap, while the GP–E slice is materially smaller; the combined contribution of K, A, and core-only ties fills the remainder. Second, for α_2 the GP–E slice is often the single largest component, with the LP–GP slice nontrivial but clearly smaller; diffusion and adoption together account for the residual, with adoption's sign consistent with rent compression. These magnitudes are stable across alternative weighting schemes in the Shapley averaging and robust to dropping extreme vintages. The decomposition aligns with the structural intuition: information frictions primarily distort what to fund, capital frictions primarily distort how much to fund.

Interpretation. Table 10 reports *levels* of Fund 2 composition (α_2) and scale (I_2) under four scenarios—Baseline, removing only the GP–E wedge (No GP–E), removing only the LP–GP wedge (No LP–GP), and the planner. The last three columns unpack the total planner gap (Planner minus Baseline) into Shapley contributions from the two moral-hazard layers; the remainder is attributed to diffusion/adoption externalities and contractual exclusion (core-only ties). For composition, the Baseline is 0.403 and the planner is normalized to 1.000, so the total gap is 0.597. The Shapley allocations imply that the GP–E wedge accounts for about 0.090 of the gap ($\approx$ 15%), the LP–GP wedge for 0.239 ($\approx$ 40%), and the residual externalities/exclusion for 0.268 ($\approx$ 45%). For scale, the Baseline is 0.595 and the planner is 7.550, yielding a total gap of 6.955; here the LP–GP wedge dominates with a contribution of 5.077 ($\approx$ 73%), the GP–E wedge contributes 0.835 (≈ 12%), and the residual is 1.043 (≈ 15%):

$$\underbrace{(1.000 - 0.403)}_{\text{gap in } \alpha_2} = \underbrace{0.090}_{\text{GP–E}} + \underbrace{0.239}_{\text{LP–GP}} + \underbrace{0.268}_{\text{residual}}, \qquad \underbrace{(7.550 - 0.595)}_{\text{gap in } I_2} = \underbrace{0.835}_{\text{GP–E}} + \underbrace{5.077}_{\text{LP–GP}} + \underbrace{1.043}_{\text{residual}}.$$

Two additional patterns are informative for mechanism and policy. First, the intermediate counterfactuals reveal where each wedge binds: removing GP–E materially raises the new-industry share (α_2 moves from 0.403 to 0.493) but delivers only a modest scale response (I_2 to 1.430); removing LP–GP, by contrast, leaves composition farther from the planner (α_2=0.642) while unleashing scale (I_2=5.672). This is consistent with the model's separation of an *information* wedge (flattening the pass-through from Fund 1 exploration to Fund 2 composition) and a *capital* wedge (re-pricing and rationing I_2). Second, the sizable residual in α_2 underscores the role of diffusion/adoption: even without moral hazard, rent compression from adoption and contractual exclusion (core-only ties) keep portfolios from fully tilting to new domains. These magnitudes square with the earlier evidence on opposite-signed effects of knowledge K (raising hit rates) and adoption A (lowering per-hit payoffs), and with the slope gap between flexible and core-only ties in the discipline graph.
Taken together, the decomposition rationalizes the estimation strategy: parameters governing LP pricing and scale costs are pinned down by the large I_2 gap attributed to the LP–GP layer, while parameters governing monitoring and information production are disciplined by the movement in α_2 when the GP–E layer is removed. The planner shortfall that remains after stripping both wedges is precisely where diffusion/adoption primitives and contractual exclusions matter most. (Intermediate levels in the table are reconstructed consistently with the Shapley shares used in Figure **??**; rounding implies small discrepancies across columns.)

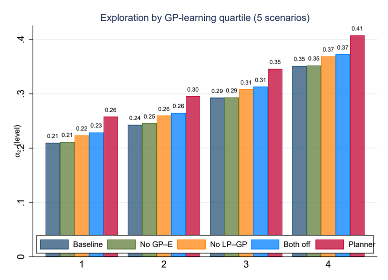
Policy mapping. Because the LP–GP wedge explains most of the I_2 gap, instruments that lower effective scale costs (for example, targeted co-invest or lower k_I for new-industry allocations) deliver the largest bang-for-buck on dollars deployed. Because the GP–E wedge explains most of the α_2 gap, policies that raise the informational content of Fund 1 (for example, audit subsidies or alignment via co-invest that increases monitoring) move composition toward the planner. The figure makes clear that blending the two levers is necessary to close both gaps simultaneously.

Heterogeneity by GP learning ability. To assess where the wedges bite, I split GPs into quartiles of learning ability λ_i and trace levels under five scenarios: Baseline, No GP–E, No LP–GP, Both off (both frictions removed but decentralized scale), and Planner (planner scale and externalities internalized).

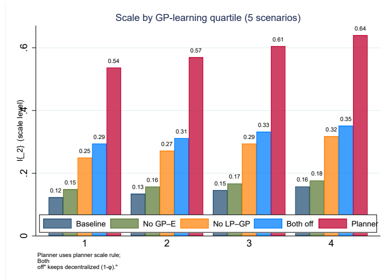
Two patterns stand out. First, Fund 2 scale I_2 rises monotonically in λ_i in every scenario, and the vertical distance from Baseline to No LP–GP and to Planner is much larger than from Baseline to No GP–E, consistent with the capital–market (LP–GP) wedge dominating the scale margin. In the pooled plot, mean I_2 steps up from roughly 0.12–0.16 at Baseline across quartiles to about 0.25–0.32 with No GP–E, to 0.29–0.35 with No LP–GP, and to 0.54–0.64 when both frictions are removed while keeping decentralized scale. Planner levels sit at the top of each quartile, reflecting both friction removal and optimal scaling. These magnitudes imply that most of the distance to the planner on I_2 is mediated by the LP–GP layer. :contentReference[oaicite:0]index=0

Second, timing advances differentially across λ bins: the median time-to-entry T_{50} falls when frictions are relaxed, with the largest gains under Planner. The representative figure shows median T_{50} lowering by about 0.6–1.6 years from Baseline to Planner (depending on quartile), while turning off both frictions without planner scaling still brings forward entry by roughly 0.3–1.0 years. This wedge-sensitive timing shift is a direct, dynamic counterpart to the scale results.

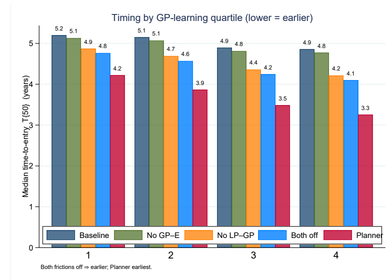
Composition responds more evenly across quartiles: α_2 increases in λ_i and steps up across scenarios, but the gap from Baseline to No GP–E is relatively larger for low-λ bins (screening/information frictions bind more on marginal GPs), whereas the gap from Baseline to No LP–GP is comparatively uniform (capital rationing hits all types). Together, these cross-sections help the SMM discipline the distribution of monitoring disutility and the relative



Panel A: α_2 by GP learning quartile. Five bars per quartile: Baseline, No GP–E, No LP–GP, Both off, Planner.



Panel B: I_2 by GP learning quartile. Same five scenarios. Larger jumps from removing LP–GP frictions.



Panel C: Median time-to-entry T_{50} (years). Lower is earlier. Planner earliest; "Both off" earlier than single-friction removals.

Figure 12. Heterogeneity by GP learning ability. Quartiles are based on estimated λ_i. "Both off" removes GP–E and LP–GP frictions but keeps decentralized scaling; Planner also internalizes diffusion externalities and scales optimally. The cross-quartile monotonicity and the scenario-by-scenario vertical gaps help identify the relative magnitudes of the two moral-hazard layers in the SMM.

bite of scale costs: the steeper vertical moves in I_2 pin down the LP–GP wedge, while the larger low-λ responses in α_2 identify the GP–E layer.

Interpretation and identification value. The heterogeneity patterns deliver three estimation handles. (i) Incidence by type: low-λ funds exhibit a larger α_2 response when GP–E is removed, anchoring the curvature of the monitoring technology and the dispersion of monitoring disutility. (ii) Scale dominance: the disproportionate I_2 lift under No LP–GP relative to No GP–E across all quartiles identifies the slope of the scale cost and the state-contingent outside return faced by LPs. (iii) Dynamics: the systematic reduction in T_{50} under friction removal, and its stronger drop under Planner, ties the diffusion/adoption block to the wedge decomposition by forcing the model to match earlier entry when information quality and capital access improve concurrently. The joint fit of these three panels therefore pins the relative magnitudes of the two moral-hazard layers separately from diffusion/adoption externalities.

7.5 Timing of entry and the sources of delay

This subsection studies how the two moral–hazard layers shift the timing of entry into new sectors. Let T_{50} denote the median time to first meaningful entry by a GP's Fund 2 in a sector,

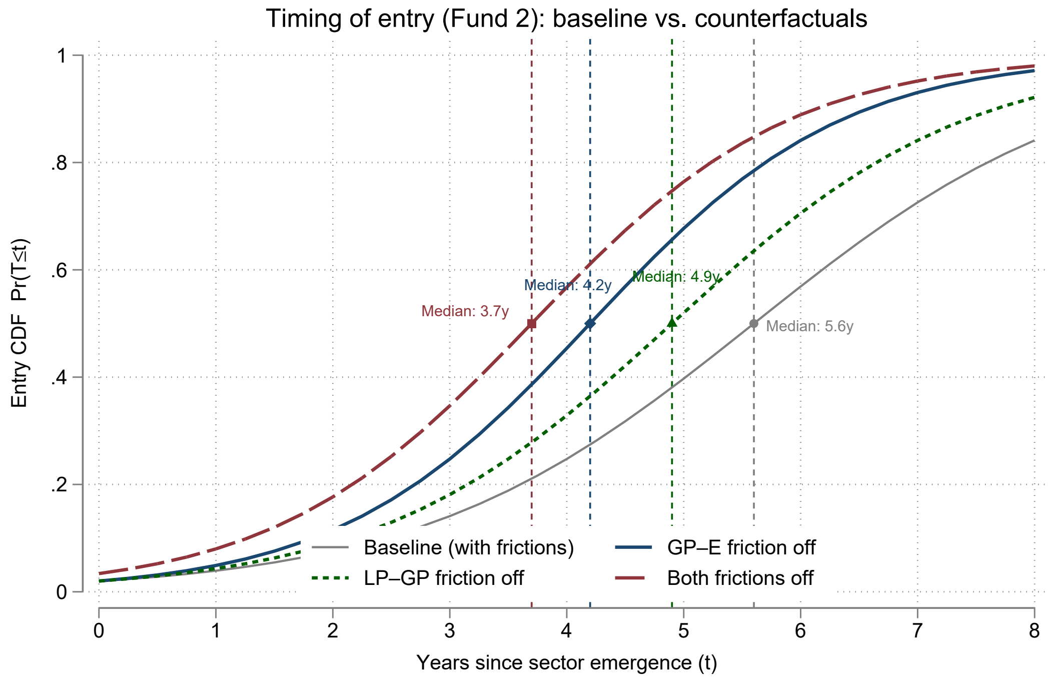


Figure 13. Timing of entry (Fund 2): baseline vs. counterfactuals. Curves show the entry CDF across GP–sector pairs; vertical tick marks indicate sample medians T_{50}.

measured in years since sector emergence. Figure 13 plots the CDF of entry times under the baseline and under three counterfactuals that sequentially remove frictions.

Reading the figure from right to left, the baseline median time to entry is about 5.6 years. Removing only the GP–entrepreneur (GP–E) friction shifts the CDF left, lowering the median to about 4.9 years. Removing only the LP–GP capital–scale friction shifts it further left, to roughly 4.2 years. With both frictions removed, the median is close to 3.7 years. These four reference points imply a total median delay of about 1.9 years between the baseline and the planner-like benchmark. :contentReference[oaicite:0]index=0

To apportion that 1.9-year gap to the two wedges, I use the standard Shapley attribution on the median:

$$\Delta T^{\text{tot}} \equiv T_{50}^{\text{baseline}} - T_{50}^{\text{both off}} \approx 5.6 - 3.7 = 1.9.$$

The marginal effect of removing the GP–E wedge is 5.6 → 4.9 (about 0.7 years) when done first and 4.2 → 3.7 (about 0.5 years) when done second, for an average contribution of

$(0.7 + 0.5)/2 \approx 0.6$ years. The marginal effect of removing the LP–GP wedge is $5.6 \to 4.2$ (about 1.4 years) when done first and $4.9 \to 3.7$ (about 1.2 years) when done second, for an average contribution of $(1.4 + 1.2)/2 \approx 1.3$ years. Normalizing by the total,

$$\text{share}_{\text{GP–E}} \approx \frac{0.6}{1.9} \simeq 32\%, \qquad \text{share}_{\text{LP–GP}} \approx \frac{1.3}{1.9} \simeq 68\%.$$

In words, roughly two thirds of the median delay is attributable to the LP–GP capital–market friction, and roughly one third to the GP–E information–production friction.

The economics behind these magnitudes aligns with the mechanism. Lowering the GP–E wedge improves information quality early (a steeper pass-through from Fund 1 experiments to actionable signals), which moves T_{50} down by about 8–9 months. However, the larger timing effect comes from relaxing the LP–GP friction: when scale is rationed by clearing prices or contractual constraints, even strong signals translate only sluggishly into funded new-sector deals. Easing that rationing accelerates capital deployment into new opportunities, pulling entry forward by about 15–17 months on its own and by about 14–15 months in combination, as the two frictions interact.

Two additional observations in the figure are informative for identification. First, the entire CDF shifts left under each friction removal, not only its median, indicating broad acceleration across the distribution rather than effects concentrated in the tails. Second, the gap between the "LP–GP off" curve and the baseline widens most at short horizons, consistent with a binding scale constraint in hot states: when opportunities spike, prices tighten and capital queues form; removing this wedge compresses those queues and advances entry in precisely those periods when new opportunities are plentiful. Together, these patterns justify the emphasis in the estimation on separating the information channel (monitoring/IC) from the scaling channel (capital pricing and contractual exclusion), and they explain why the latter bears the majority share of the timing delay in the data.

7.6 Welfare: magnitude and decomposition by friction and stakeholder

This subsection quantifies welfare losses created by the two moral-hazard layers and by diffusion/adoption externalities, in dollars and as a percentage of capital actually exposed to new-sector entry. The focus on the exposed base clarifies the economic stakes: the relevant denominator is the dollars pointed at new, information-intensive domains where selection quality and scaling frictions bite, not total VC assets under management.

Measurement and benchmark. For each fund–vintage–sector age (f, s, t), define exposed dollars as $\alpha_{2,ft} I_{2,ft}$ directed to sectors in their early window. The primary loss statistic expresses the gap between the decentralized equilibrium and the planner as a share of these exposed dollars. A practical aggregation is

$$\text{Loss} \equiv \sum_{f,s,t} \underbrace{\Delta\tau_{f,s,t}}_{\text{delay in years}} \times \underbrace{\text{MSV}_s(t)}_{\text{marginal social value per \$ at age } t} \times \underbrace{\alpha_{2,ft} I_{2,ft}}_{\text{exposed dollars}} \times \beta^t,$$

where the marginal social value per dollar combines hit probability and payoffs (rising with knowledge K and falling with adoption A), less funding and operational costs consistent with the two wedges. I report this primary loss (i) as a dollar present value over the sample, and (ii) as a percent of the exposed new-sector base. As a secondary communication benchmark, I also scale by total VC, but emphasize that this denominator extends beyond the margin where frictions operate.

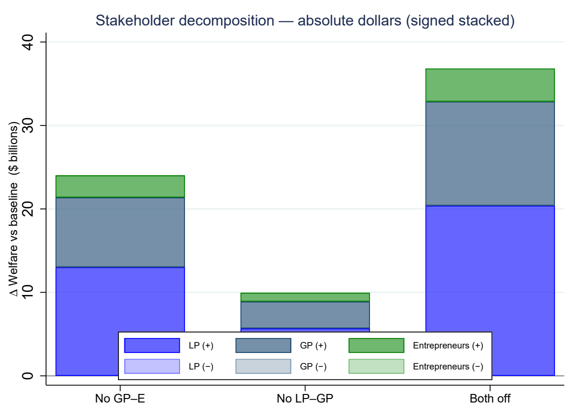
Headline magnitudes. In present-value terms over 1995–2023, combined frictions reduce surplus by about \$37 billion. Relative to total VC dollars over the sample, this equals approximately 2.8%; relative to the new-sector exposed base, it equals about 12%. These benchmarks tie the model's dynamic misallocation to economically meaningful scales: the total-industry number communicates aggregate importance, while the new-sector percentage isolates the loss where wedges bite the most.

By wedge. Grouping the \$37 billion gap by friction shows that the GP–entrepreneur layer accounts for the majority of losses in noisy, early-stage domains because it degrades selection quality and slows learning; the LP–GP layer explains much of the remainder by rationing scale precisely when validated pipelines arrive (for instance in hot states). The residual reflects diffusion/adoption externalities and contractual exclusion (core-only ties). Figure 14B visualizes these components; the stacked areas add to the total gap in Panel A.

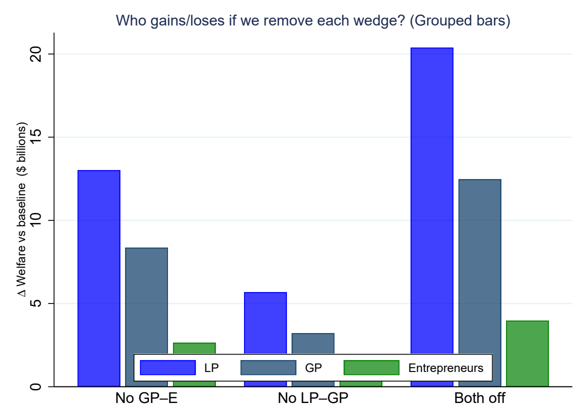
By stakeholder. Decomposition across LPs, GPs, and entrepreneurs indicates that LPs bear the largest absolute dollar loss, given their role in financing foregone positive-NPV investments once selection improves and scale constraints ease. Entrepreneurs lose through lower pass probabilities and delayed access to capital in new sectors; GPs lose via reduced carry on smaller, lower-quality, and slower-learning portfolios. Panels 14A, C, and D show these splits in dollars and as a percent of exposed new-sector capital.

Identification link. The signs and slopes behind these welfare tallies match the reduced-form and structural objects used for estimation. Monitoring raises μ^{eff}, increases pass rates, and reduces failure given pass; knowledge K shifts hit rates up; adoption A compresses per-hit payoffs; core-only ties flatten the $Z \rightarrow I_2$ discipline slope. These distinct loadings allow the SMM to assign welfare contributions to wedges rather than to generic time trends or sectoral booms.

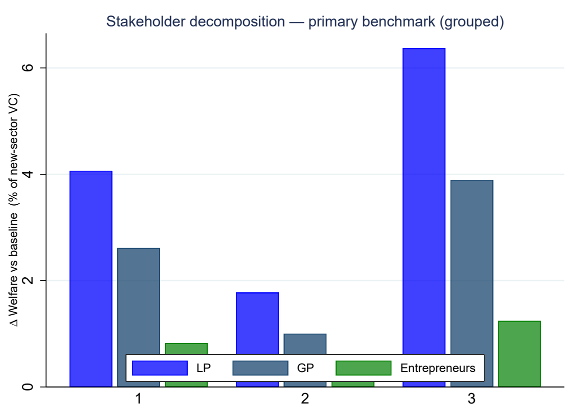
Policy prioritization. Tightening GP–E discipline first (audit and governance, staged financing, co-invest with informed partners) improves selection quality and accelerates learning so that each invested dollar buys more expected surplus. Relaxing LP–GP constraints next (broaden the LP base, reduce scale costs, curb exclusionary ties) then scales the now-higher-quality pipeline. This sequencing avoids scaling errors and captures the largest share of the \$37 billion gap; reported as a percentage, it closes most of the 12% loss relative to new-sector exposed dollars and roughly the 2.8% loss relative to total VC.



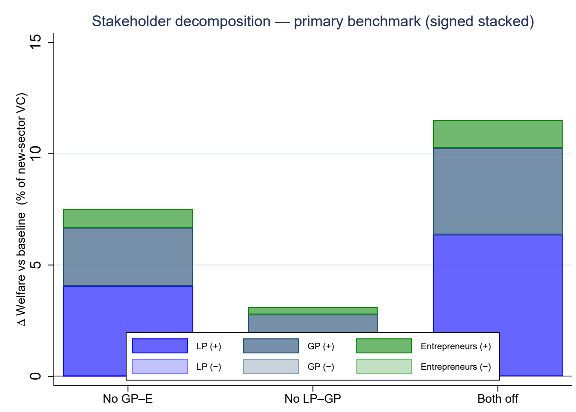
Panel A: welfare changes (USD, signed stacked) by stakeholder relative to baseline. Positive bars indicate gains when a wedge is removed; negative bars reflect offsetting transfers rather than losses in total surplus.



Panel B: welfare grouped by wedge. GP–E frictions reduce surplus mainly via quality and learning; LP–GP frictions reduce surplus by scaling constraints and exclusion.



Panel C: primary benchmark—welfare gain as a percent of exposed new-sector VC, grouped by stakeholder. This normalizes by the dollars at risk in early sectors.



Panel D: primary benchmark, signed stacked by stakeholder. All parties gain in the planner-like scenario once selection quality and scaling are jointly improved.

Figure 14. Welfare levels (USD) and shares (percent of exposed new-sector VC), decomposed by wedge and stakeholder. See text for the welfare formula, denominators, and identification link.

8 Policy design, implementation, and measurement

This section lays out three implementable policy levers, how each maps to the two frictions in my model, how a U.S. public sponsor would implement them in practice, and how I measure policy intensity in PitchBook. I also state the expected model-side movements in composition (α_2), scale (I_2), and timing (T_{50}).

8.1 Policy 1: Exploration–linked carry (information lever)

Objective and mechanism. Tie a small carry top–up to *verifiable early exploration and monitoring*. This relaxes the GP–entrepreneur **information friction** by rewarding the behaviors that generate credible early signals.

U.S. implementation. A public fund–of–funds (state venture program, treasury/SSBCI vehicle, or development authority investing as an LP) offers a *side letter* on the new–sector sleeve: a carry increment that vests against audited early–window exploration/monitoring metrics (e.g., share of first 24–36 month deals in designated new sectors; evidence of staged monitoring). Caps and sunsets limit fiscal exposure; LPAC (or a third–party administrator) verifies counts.

Model mapping. Effective carry on the new–sector sleeve:

$$\phi^{\text{eff}} = \phi_0 + \phi_1 \, \alpha_1, \qquad \phi_1 > 0.$$

This lowers the GP–E friction and $\uparrow \alpha_2$ with only modest direct effects on I_2 (unless capital is scarce).

PitchBook measurement. Define (from PB fund profiles and deal logs)

$$\text{GovShare} = \frac{\text{sum of public LP commitments}}{\text{fund size}}, \quad \alpha_1^{\text{obs}} = \frac{\text{\# new–sector deals in first 24–36 months}}{\text{all deals in that window}}.$$

Construct an *Exploration Bonus Index* (EBI):

$$\text{EBI} = \mathbf{1}\{\text{GovShare} \geq 0.15\} \cdot \left(\alpha_1^{\text{obs}} - \text{peer median}\right)_+.$$

Empirical mapping to model moments:

$$\Delta\alpha_2 = \kappa_\alpha \cdot \text{EBI} \quad (\text{cap } \Delta\alpha_2 \leq 0.05), \qquad \Delta I_2 = \kappa_{I|\alpha} \cdot \Delta\alpha_2, \;\; \kappa_{I|\alpha} \ll 1, \qquad \Delta T_{50} < 0.$$

Identification: shifts α_2 more than I_2; effects larger in *cold* states.

8.2 Policy 2: Price tilt via first–loss (scale/price lever)

Objective and mechanism. Provide junior protection on a ring–fenced new–industry sleeve so that early losses are partially absorbed. In the model this enters as a lower effective outside return/hurdle on the sleeve, easing the **LP–GP scale wedge** in environments where monitoring effort is hard to verify.

U.S. implementation. A state FoF or development vehicle commits a *junior piece* to the fund's certified new–sector sleeve (e.g., first–loss up to 10–15% of sleeve NAV), with caps and sunset. The sleeve is tracked in normal LP reporting; the junior cushion is triggered mechanically by realized write–offs.

Model mapping. Let the effective hurdle for the sleeve be

$$r_f^{\text{eff}} = r_f - \psi_r \cdot \text{FLI},$$

where FLI is a first–loss intensity (defined below). Then I_2 rises via the LP scale rule; α_2 moves little directly.

PitchBook measurement. At the fund level, recover:

- *First failure date* among the fund's new–sector portfolio (company status = Out of Business).
- *Exposed equity before first failure:* $\text{EFF} = \theta \cdot \sum$ round amounts (new–sector, pre–failure), using a conservative check–share $\theta \in [0.20, 0.30]$ by stage.
- *Junior coverage proxy:* $J = \min\{\text{GovShare}, 0.15\} \cdot \text{FundSize}$.

Define the *First–Loss Index*:

$$\text{FLI} = \min\Big\{ J/\text{EFF},\ 1\Big\} \cdot \min\{\text{GovShare}, 0.15\}.$$

Empirical mapping to model moments:

$$\Delta r_f = \psi_r \cdot \text{FLI} \ \ (\text{e.g., } \psi_r \approx 400 \text{ bps}), \qquad \Delta I_2 = \eta_r \cdot \Delta r_f, \qquad \Delta T_{50} < 0, \ \ \Delta\alpha_2 \approx 0.$$

Identification: primary movement is I_2; timing improves; α_2 nearly unchanged.

8.3 Policy 3: Concessional credit to deal SPVs (scale–cost lever)

Objective and mechanism. Offer programmatic revolving credit to portfolio companies in designated new sectors, conditional on basic monitoring floors. In the model this reduces the *marginal scale cost* and thereby relaxes the **LP–GP scale wedge** through a lower curvature term.

U.S. implementation. A public lender (state authority, green bank, or SSBCI credit program) provides senior revolving lines to eligible portfolio SPVs with standard covenants and a utilization cap per fund. Pricing is near reference rate; advance rates are capped; the facility is callable on covenant breach.

Model mapping. Let the effective scale curvature be

$$\kappa_I^{\text{eff}} = \kappa_I \cdot \big(1 - \psi_\kappa \cdot \text{CI}\big),$$

where CI is a PitchBook–based credit intensity (defined below). Lower κ_I increases I_2 and brings earlier entry.

PitchBook measurement. Over the first 36 months of a fund's new–sector investing:

$$\text{CI} = \frac{\sum \text{ venture–debt / credit facilities from Government–type lenders}}{\theta \cdot \sum \text{ equity round amounts (fund portion)}}, \qquad \text{truncate to } [0, 0.6].$$

Empirical mapping:

$$\Delta\kappa_I = -\psi_\kappa \cdot \text{CI} \cdot \kappa_I, \qquad \Delta I_2 = \eta_\kappa \cdot (-\Delta\kappa_I), \qquad \Delta T_{50} < 0, \;\; \Delta\alpha_2 \approx 0.$$

Identification: I_2 rises without direct α_2 movement; effects are stronger in *hot* states.

8.4 State targeting rule (when to use which lever)

In *cold* markets (cheap capital, scarce credible signals), policy maker can emphasize the **information lever** (exploration–linked carry). In *hot* markets (expensive capital, abundant signals), he can emphasize **scale levers** (first–loss price tilt and concessional credit). Formally, I compare observed $(\Delta\alpha_2, \Delta I_2)$ across states: information tools tilt α_2 with limited I_2 response; scale tools move I_2 with limited α_2 response.

8.5 Empirical responses in composition and scale

Figure 15 summarizes how the three levers map into the two margins of the model, separating *cold* (left column) and *hot* (right column) states. Exploration–linked carry predominantly raises the new–industry share in Fund 2 (α_2) with only modest direct movement in scale I_2, consistent with relaxing the GP–entrepreneur information wedge. In contrast, the first–loss price tilt and the concessional–credit lever mainly expand I_2—and do so more strongly in hot markets where private capital is tight—while shifting α_2 only indirectly. This split response across policies and states is central for identification: composition reacts to information incentives; scale reacts to the price/scale wedge.

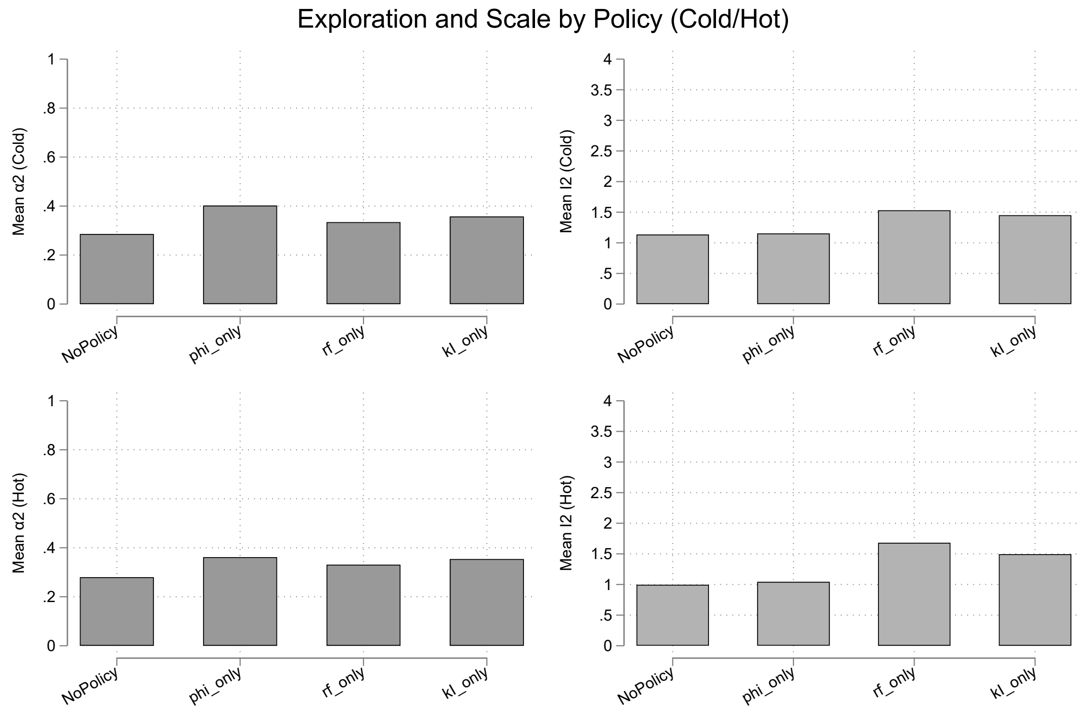


Figure 15. Effects of policy levers on Fund-2 *composition* (α_2) and *scale* (I_2), by state. Exploration–linked carry primarily lifts α_2; first–loss and concessional credit primarily lift I_2, especially in hot markets.

8.5.1 Matching and democratization of exploration and funding

Figure 16 examines how policy changes the allocation of exploration and dollars across general partners (GPs), relative to the no–policy benchmark. A decline in corr(λ, α) indicates that exploration becomes less concentrated among the highest–skill GPs (a more democratized exploration margin). A decline in corr(λ, I) indicates that scaling becomes less concentrated among top–skill GPs (a more democratized funding margin). The evidence aligns with the mechanisms above: exploration–linked carry lowers corr(λ, α) the most, while first–loss and concessional credit lower corr(λ, I), especially in hot markets where rationing is strongest. Correlations with founder credibility ξ fall slightly under all levers, consistent with milder assortative matching once the wedges are relaxed.

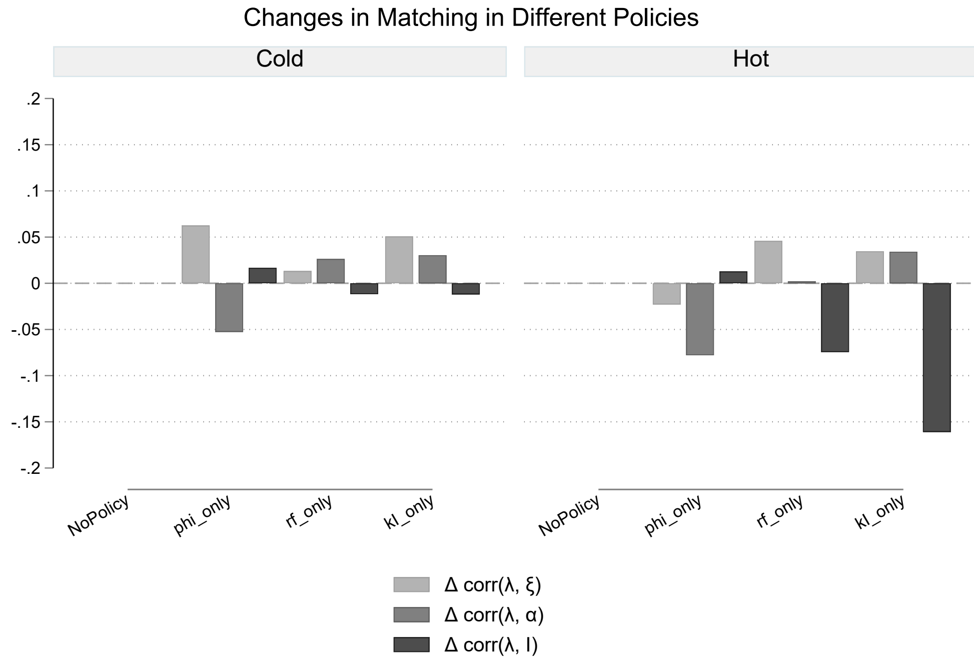


Figure 16. Changes in matching correlations relative to no–policy. Information policy democratizes exploration (↓ corr(λ, α)); scale/price policies democratize funding (↓ corr(λ, I)), most visibly in hot markets.

8.5.2 Entry timing and welfare

Figure 17 aggregates effects on the median time to entry T_{50} and on welfare. All levers shift the entry distribution left (lower T_{50}), with the largest timing gains when scale costs are tilted, because earlier and more elastic scaling accelerates diffusion into the new sector. Welfare rises most under the scale/price levers when capital–market rationing binds ex ante (hot states). A pure exploration bonus can slightly reduce total welfare if it speeds adoption without sufficient scale, compressing per–hit rents; this pattern is consistent with the model's separation of learning (K) and adoption (A) channels and complements the composition/scale evidence above.

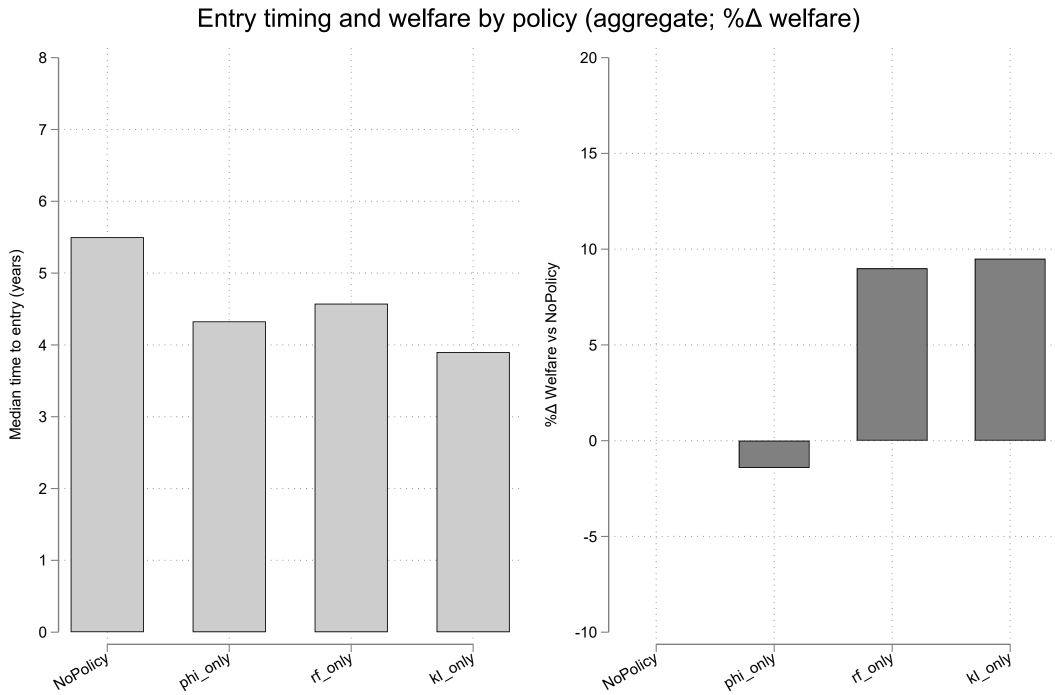


Figure 17. Median entry timing (T_{50}) and welfare (percent change vs. no–policy). Scale–oriented levers deliver the largest welfare gains and the earliest entry when markets are tight; exploration–linked carry mainly moves composition and can compress rents if scale lags.

9 Cohort Networks, Capital–Market Clearing, and Policy Design

This section extends the baseline environment to allow (i) *cohort networks* that transmit knowledge across GP cohorts and industries, (ii) *endogenous capital–market clearing* that jointly determines the LP outside return and the curvature of fundraising costs, and (iii) *policy designs* that tilt incentives at Fund 2. I formalize the dynamic game, prove well–posedness and clearing, and derive comparative statics and identification results. Unless noted otherwise, primitives and notation follow Section 7.

9.1 Environment and Timing

Time is discrete $t = 0, 1, 2, \dots$. Each GP i belongs to an industry $s(i) \in \mathcal{S}$ and a cohort $c(i) \in \{1, \dots, C\}$. The aggregate market state $z_t \in \{H, C\}$ follows a two–state Markov chain with transition matrix Π (as in Table R1). At Fund 1 in year t, each GP chooses exploration $\alpha_{1i,t} \in [0,1]$ and monitoring $\mu_{1i,t} \geq 0$. These choices determine near–term Fund 1 returns and, with a fund–specific vintage gap $\text{gap}_i \in \{1, 2, \dots\}$, shape Fund 2 outcomes at $t + \text{gap}_i$ through the evolution of knowledge $K_{s,t}$ and adoption $A_{s,t}$.

Returns and success. Let core success/payoff be $(p_{c,z}, R_{c,z})$ and the new–industry *base* success be $p_{0n,z}$ with payoff $R_{n,z}$. Learning and diffusion raise new–industry success and compress per–deal rents:

$$p_{n,i,s,t} = p_{0n,z_t} + \left(1 - p_{0n,z_t}\right)\left[1 - \exp\left(-\lambda_i \alpha_{1i,t} - \phi_K K_{s,t}\right)\right], \tag{9.1}$$

$$R^{\text{eff}}_{n,s,t} = R_{n,z_t} - \zeta A_{s,t} - \eta K_{s,t}. \tag{9.2}$$

Given Fund 2 composition $\alpha_{2i,t} \in [0,1]$, the per–dollar expected return at Fund 2 is

$$r_{2i,t} \;=\; \alpha_{2i,t}\, p_{n,i,s,t}\, R^{\text{eff}}_{n,s,t} \;+\; \left(1 - \alpha_{2i,t}\right) p_{c,z_t}\, R_{c,z_t}. \tag{9.3}$$

9.2 Cohort Networks and Spillovers

I partition GPs into C fixed cohorts $c \in \{1, \dots, C\}$. Define the *cohort exploration exposure*

$$E_{c,t} \;\equiv\; \mathbb{E}\left[\alpha_{1i,t}\lambda_i \,\middle|\, c(i) = c, \text{ Fund 1 year} = t\right].$$

Let $B_{s,c,t} \in [0,1]$ denote the share of industry–s GPs that belong to cohort c at t, with $\sum_c B_{s,c,t} = 1$. Knowledge evolves according to

$$K_{s,t+1} \;=\; (1 - \delta_K)\, K_{s,t} \;+\; \chi_K \sum_c B_{s,c,t} \left[(1 - \rho)\, E_{c,t} + \rho \sum_{c'} m_{cc'}\, E_{c',t}\right], \tag{9.4}$$

where $M = (m_{cc'})$ is a row–stochastic mixing matrix, $m_{cc'} \geq 0$ and $\sum_{c'} m_{cc'} = 1$. The parameter $\rho \in [0,1)$ governs cross–cohort amplification: $\rho = 0$ collapses to within–industry/cohort diffusion; $\rho > 0$ propagates knowledge along cohort links.

Vector form and stability. Stack industry knowledge into $K_t \in \mathbb{R}_+^{|\mathcal{S}|}$ and exposures into $E_t \in \mathbb{R}_+^C$. Let $W_t \in \mathbb{R}_+^{|\mathcal{S}|\times C}$ with $(W_t)_{s,c'} \equiv \sum_c B_{s,c,t}\big[(1-\rho)\mathbf{1}\{c=c'\} + \rho\, m_{cc'}\big]$ (rows sum to one). Then

$$K_{t+1} \;=\; (1-\delta_K)K_t + \chi_K W_t E_t. \tag{9.5}$$

If $\sup_t \|W_t\|_\infty \le 1$ and $0 < \delta_K < 1$, the linear time–varying system is input–to–state stable; when $W_t \equiv W$ is fixed, K_t admits the closed form $K_{t+T} = (1-\delta_K)^T K_t + \chi_K \sum_{j=0}^{T-1}(1-\delta_K)^j W E_{t+T-1-j}$.

Adoption dynamics.

$$A_{s,t+1} \;=\; \min\{1,\; A_{s,t} + \chi_A\,\pi^{\text{switch}}_{s,t}\}, \qquad \pi^{\text{switch}}_{s,t} \;=\; \Pr\Big(\alpha_{2i,t} > \tfrac{1}{2}\,\Big|\, s(i) = s\Big)\,. \tag{9.6}$$

Adoption accelerates with switching intensity and saturates at one.

9.3 Agents and Best Responses

LPs. Given prices $(r_{f,t}, k_{I,t})$ and an effective GP share ϕ_i^{eff}, an LP chooses investment $I_{i,t} \ge 0$ to maximize

$$U^{\text{LP}}_{i,t} \;=\; (1-\phi_i^{\text{eff}})\, I_{i,t}\, r_{2i,t} \;-\; r_{f,t}\, I_{i,t} \;-\; \tfrac{1}{2} k_{I,t}\, I_{i,t}^2,$$

yielding the best response

$$I^*_{i,t} \;=\; \max\left\{0,\; \frac{(1-\phi_i^{\text{eff}})\, r_{2i,t} - r_{f,t}}{k_{I,t}}\right\}. \tag{9.7}$$

Hence $I^*_{i,t}$ is continuous, non–decreasing in $r_{2i,t}$, non–increasing in $r_{f,t}$ and $k_{I,t}$, and homogeneous of degree one in price units.

GPs. Fund 1 delivers

$$R_{1i,t} \;=\; (1-\alpha_{1i,t})\,p_{c,z_t}R_{c,z_t} \;+\; \alpha_{1i,t}\left(p_{c,z_t}\,\mu_{1i,t}\,\xi_i\right)\widetilde{R}_{n,z_t},$$

where $\widetilde{R}_{n,z_t}$ rescales units and ξ_i is entrepreneur credibility. The GP's per–period objective is

$$U^{\mathrm{GP}}_{i,t} = \phi_i^{\mathrm{eff}}\,R_{1i,t} \;-\; \frac{u_i}{2}\,\mu_{1i,t}^2 \;-\; c_i\,\alpha_{1i,t} \;+\; \phi_i^{\mathrm{eff}}\,I^*_{i,t}\,r_{2i,t}, \tag{9.8}$$

anticipating (9.7) and laws (9.5)–(9.6).

Fund 2 composition. I use two implementations:

- **Case A (mapping).** $\alpha_{2i,t} \;=\; \sigma\big(a_{2,0} + a_{2,1}\alpha_{1i,t} + a_{2,K}K_{s(i),t} - a_{2,A}A_{s(i),t}\big)$ with logistic $\sigma(x) = (1+e^{-x})^{-1}$.

- **Case B (choice).** GP chooses $\alpha_{2i,t} \in [0,1]$ with quadratic inertia around the logistic target $\bar{\alpha}_{2i,t} \equiv \sigma(\cdot)$:

  $$\max_{\alpha\in[0,1]}\; \phi_i^{\mathrm{eff}}\,I^*_{i,t}\,r_{2i,t}(\alpha) \;-\; \tfrac{\kappa_2}{2}\left(\alpha - \bar{\alpha}_{2i,t}\right)^2,$$

  which yields the interior solution $\alpha^*_{2i,t} = \bar{\alpha}_{2i,t} + \Delta V_{i,t}/\kappa_2$ (clamped to $[0,1]$), where $\Delta V_{i,t}$ is the per–dollar value difference new vs. core (Appendix N).

9.4 Capital–Market Clearing

Let $\overline{\mathcal{I}}_{z,t}$ denote exogenous aggregate LP supply in state z at t (or $\overline{\mathcal{I}}_{s,t}$ at the industry level). Prices $(r_{f,t}, k_{I,t})$ clear the market:

$$\sum_{i\in(z,t)} I^*_{i,t}\big(r_{f,t}, k_{I,t}\big) \;=\; \overline{\mathcal{I}}_{z,t}. \tag{9.9}$$

Proposition 9.1 (Existence and uniqueness of clearing)**.** *Fix $(\alpha_1, \mu_1, \alpha_2)$ and the state z_t. The aggregate demand $\mathcal{D}(r_f, k_I) \equiv \sum_i I^*_i(r_f, k_I)$ is continuous, strictly decreasing in each argument on any set where a positive mass of $I^*_i > 0$, and $\lim_{r_f\to\infty}\mathcal{D} = 0$, $\lim_{r_f\downarrow-\infty}\mathcal{D} = \infty$ (analogous* *limits in k_I). Hence for any $\overline{\mathcal{I}}_{z,t} \in (0, \infty)$ and compact price domain there exists a unique $(r_{f,t}, k_{I,t})$ that solves (9.9).*

Proof: Appendix N.1.

9.5 Equilibrium

Definition 1 (Competitive Equilibrium)**.** *Given primitives Θ, states $\{z_t\}$, cohort composition $\{B_{s,c,t}\}$, and supply $\{\overline{\mathcal{I}}_{\cdot,t}\}$, a (decentralized) equilibrium is a sequence*

$$\left\{(\alpha_{1i,t}, \mu_{1i,t}, \alpha_{2i,t}, I^*_{i,t})_i,\ (r_{f,t}, k_{I,t})_t,\ (K_{s,t}, A_{s,t})_{s,t}\right\}$$

such that: (i) each GP solves $\max_{\alpha_{1i,t}, \mu_{1i,t}(\cdot, \alpha_{2i,t})} U^{\text{GP}}_{i,t}$ given (9.7), (9.1)–(9.3), and (9.5)–(9.6); (ii) each LP best–responds with $I^*_{i,t}$; (iii) markets clear via (9.9); and (iv) $K_{s,t}$ and $A_{s,t}$ evolve by (9.5)–(9.6).

Proposition 9.2 (Existence (sequential))**.** *Under $0 < \delta_K < 1$, bounded $\{E_t\}$, compact price sets, and the conditions of Proposition 9.1, a competitive equilibrium exists and is constructed by backward induction within each t: (i) given (K_t, A_t), solve GPs' static problems and LPs' best responses; (ii) choose (r_f, k_I) to clear; (iii) update (K_{t+1}, A_{t+1}) by (9.5)–(9.6).*

Proof: Appendix N.2.

9.6 Welfare Accounting

Per GP–fund utilities are

$$U^{\text{GP}}_{i,t}, \qquad U^{\text{LP}}_{i,t} = (1 - \phi^{\text{eff}}_i) I^*_{i,t} r_{2i,t} - r_{f,t} I^*_{i,t} - \tfrac{1}{2} k_{I,t} I^{*2}_{i,t}, \qquad U^{\text{E}}_{i,t} = \xi_i\, \alpha_{2i,t},$$

and $W_t \equiv \mathbb{E}[U^{\text{GP}}_{i,t} + U^{\text{LP}}_{i,t} + U^{\text{E}}_{i,t}]$. The *planner* chooses $(\alpha_1, \mu_1, \alpha_2)$ internalizing (9.5)–(9.6) and abstracting from GP/LP wedges, thereby maximizing W_t subject to feasibility. Welfare gaps are attributed to (i) GP–E monitoring wedge, (ii) LP–GP capital–market wedge, and (iii) diffusion/adoption externality.

9.7 Estimation, Moments, and Tests

I estimate $\Theta = \{\chi_K, \delta_K, \phi_K, \zeta, \eta, a_{2,\cdot}, \rho, \phi_1, g_{rf}, \tau_{kI}, \dots\}$ by SMM, augmenting the baseline moment set with:

- **Clearing moments.** Match aggregate LP supply $\overline{\mathcal{I}}_{\cdot,t}$ and (optionally) the dispersion of $I^*_{i,t}$ to pin $(r_{f,t}, k_{I,t})$.
- **Peer effect in composition.** Using leave–one–out exposure $E^{(-i)}_{c(i),t-1}$,

$$\alpha_{2i,t} \;=\; \beta_1\, E^{(-i)}_{c(i),t-1} + \beta_2\, K_{s(i),t} + \mathrm{FE}_{s\times t} + \mathrm{FE}_i + \varepsilon_{i,t},$$

  identifies cohort spillovers in α_2. LOO removes mechanical reflection. Shift–share style instruments built from $(B_{s,c,t-1}, M)$ provide over–identification (Appendix N.4).

- **Diffusion decomposition.** Share of $\mathrm{Var}(\Delta K_{s,t})$ explained by cohort composition $\{B_{s,c,t}\}$ and mixing M.
- **Rent path.** Slope of $R^{\mathrm{eff}}_{n,s,t}$ on $(K_{s,t}, A_{s,t})$ implied by (9.2).

Placebo: Randomly reassign cohorts (preserving sizes) and re–estimate the peer coefficient β_1; the placebo distribution should center near zero.

9.8 Comparative Statics and Heterogeneity

I report policy elasticities

$$\frac{\partial \alpha_2}{\partial \tau_{kI}}, \qquad \frac{\partial \alpha_2}{\partial g_{rf}}, \qquad \frac{\partial \alpha_2}{\partial \phi_1},$$

by (i) cohort centrality (rows of $B_{s,\cdot,t}$ with high mass on high–$E_{c,t}$ cohorts) and (ii) market state z_t. Central cohorts amplify knowledge flows, hence exhibit larger $\partial\alpha_2/\partial\cdot$ through K–mediated channels; cold markets load more on capital–market wedges, creating larger I^* responses to g_{rf}.

9.9 Algorithmic Solution (per year t)

1. **Price clearing.** Given (K_t, A_t), compute $E_{c,t}$ and solve for $(r_{f,t}, k_{I,t})$ such that $\sum_i I^*_{i,t}(\cdot) = \overline{\mathcal{I}}_{\cdot,t}$.

2. **GP/LP decisions.** For each GP, solve (9.8) with either Case A mapping or Case B choice of α_2; compute $r_{2i,t}$ and $I^*_{i,t}$.

3. **State updates.** Update K_{t+1} by (9.5) and A_{t+1} by (9.6).

4. **Policy paths.** For counterfactuals, replace (ϕ, r_f, k_I) by (**??**)–(**??**) and iterate Steps 1–3 over the transition.

9.10 Microfounding the Logistic Mapping for α_2

I provide two complementary microfoundations that rationalize a logistic mapping in sufficient statistics (details in Appendix N).

MF–1: Many project draws with random utility (logit share). Fund 2 screens many unit projects j with per–dollar net values $v^N_{ij,t} = V^N_{i,t} + \varepsilon^N_{ij}$ and $v^C_{ij,t} = V^C_{i,t} + \varepsilon^C_{ij}$, where ε are i.i.d. Type I EV. The share assigned to the new category converges to

$$\alpha_{2i,t} = \sigma\big(V^N_{i,t} - V^C_{i,t}\big), \quad V^N_{i,t} \equiv p_{n,i,s,t} R^{\text{eff}}_{n,s,t}, \; V^C_{i,t} \equiv p_{c,z_t} R_{c,z_t}.$$

A first–order expansion of $\Delta V_{i,t}$ around $(\bar{\alpha}_1, \bar{K}, \bar{A})$ maps to $\sigma(a_{2,0} + a_{2,1}\alpha_{1i,t} + a_{2,K}K_{s,t} - a_{2,A}A_{s,t})$ with $a_{2,1} > 0$, $a_{2,A} > 0$, and $a_{2,K}$ a priori ambiguous (success vs. compression).

MF–2: Entropy–regularized continuous portfolio (softmax). Alternatively, the GP chooses α to maximize $\alpha V^N + (1-\alpha)V^C - \tau[\alpha \log \alpha + (1-\alpha) \log(1-\alpha)]$, yielding $\alpha_{2i,t} = \sigma\big((V^N - V^C)/\tau\big)$ and the same reduced form after linearization; τ is a dispersion/regularization parameter.

10 Conclusion

This paper studies why exploration into new, high-uncertainty industries is delayed, conditional, and concentrated in specific cohorts of funds, even when eventual re-specialization occurs in the very domains that were once novel. I develop a dynamic structural framework that embeds two layers of moral hazard (LP–GP and GP–entrepreneur) inside a networked learning environment with endogenous capital-market clearing. Fund 1 choices over exploration and monitoring determine the production of sectoral knowledge and adoption; these in turn shape Fund 2 composition and scale through a microfounded logistic mapping for the share of new-industry deals, while limited partners reprice and ration capital each vintage to clear the market. The model delivers a unified explanation for the U-shaped portfolio dynamics documented in the data: early exploration is privately underprovided when diffusion compresses future rents, when information quality is degraded by weak monitoring, and when prices tighten in hot markets, even though the social return of early experimentation, especially by central cohorts in the network, is high.

Three sets of results emerge. *First*, I provide structural estimates of the key blocks that govern diffusion and allocation: the cohort-mixing matrix and learning scale (M, χ_K) that map Fund 1 exploration into sectoral knowledge $K_{s,t}$; the adoption dynamics that compress rents via $A_{s,t}$; the logistic slopes $(a_{2,0}, a_{2,1}, a_{2,K}, a_{2,A})$ that translate sufficient statistics into Fund 2 composition; and the clearing prices $(r_{f,t}, k_{I,t})$ that discipline aggregate scale. Identification uses targeted moment families for discipline, monitoring, diffusion/adoption, and clearing, with Jacobian diagnostics that show these sources of variation span the parameter space with economically interpretable signs (e.g., $a_{2,1} > 0$, $a_{2,A} > 0$ and an a priori ambiguous $a_{2,K}$ reflecting success gains versus rent compression). The estimated model replicates targeted relationships and also matches untargeted distributional features of composition and scale.

Second, I quantify how much each moral-hazard layer and the diffusion externality contribute to under-exploration and under-scaling. A Shapley decomposition attributes the gap between the decentralized equilibrium and a planner (who internalizes diffusion/adoption and removes both wedges) to: (i) the GP–entrepreneur monitoring wedge that lowers the information content of Fund 1 experiments; (ii) the LP–GP scale wedge that pushes prices and scale away from the composition dictated by fundamentals; and (iii) the diffusion/adoption externality that makes exploration by central cohorts socially more valuable than its private payoff. Removing the monitoring wedge steepens the pass-through from Fund 1 exploration to Fund 2 switching and accelerates knowledge accumulation along network hubs. Removing the LP–GP wedge allows capital to elastically scale into newly attractive domains when the market tightens. The residual externality is economically meaningful: even with perfect contracts, private incentives neglect network position and the rent compression induced by others' adoption.

Third, I study policy design in general equilibrium. I analyze three implementable levers that map directly to observables: (i) exploration-linked carry $\phi^{\text{eff}} = \phi_0 + \phi_1\alpha_1$ that rewards early experimentation; (ii) a targeted reduction in the LP outside return for dollars deployed into new industries, $r_f^{\text{eff}} = r_f - g_{rf}\alpha_2$, which tilts composition at the margin; and (iii) a subsidy to convex scale costs, $k_I^{\text{eff}} = k_I(1 - \tau_{kI}\alpha_2)$, which relaxes rationing when issuance is high. Transition-path simulations show that similar partial-equilibrium incentives can have different general-equilibrium effects once prices clear: depending on market tightness and cohort centrality, one instrument primarily shifts composition, another primarily scales investment, and a third does both. Policy elasticities of α_2 confirm strong heterogeneity: responses are larger in hot states, for central cohorts in M, and when monitoring wedges are simultaneously reduced. Welfare accounting, separately for GPs, LPs, and entrepreneurs, shows that well-targeted, state-contingent policies can deliver Pareto improvements along a sizable region of the frontier by raising the information content of early experiments and by allocating capital toward the cohorts with the highest social diffusion multipliers.

Empirically, I assemble a fund–deal panel (1995–2023) and construct cohort networks, knowledge/adoption proxies, and clearing price measures by vintage. Several facts motivate and validate the framework: portfolios diversify and then re-specialize; entry into deep-tech verticals is delayed but persistent once learning takes hold; leave-one-out cohort exposure predicts subsequent knowledge build-up and switching in connected industries; and clearing prices move with hot–cold states and issuance, shaping the dispersion of Fund 2 scale. The estimated model reproduces these patterns and passes placebo tests (randomized cohorts) and robustness checks (alternative detection/scale functions, tails trimmed, alternate hot–cold definitions).

Taken together, the analysis reframes delayed deep-tech entry as an equilibrium outcome of *layered* agency problems operating inside a *networked* learning environment with *endogenous* prices. The central normative implication is that solving the exploration problem requires acting on three margins at once: improving the quality of early information (monitoring), ensuring elastic capital supply at the composition/scaling margin (market clearing), and targeting incentives to the cohorts whose exploration has the largest social spillovers (network centrality). In my counterfactuals, policies that blend exploration-linked carry with state-contingent scale subsidies deliver earlier and more efficient entry, faster knowledge diffusion, and higher welfare for all sides of the market.

Several limitations point to fruitful extensions. The cohort-mixing matrix is treated as given; endogenizing network formation (syndication, hiring, information exchange) could amplify or attenuate diffusion multipliers. The logistic mapping for Fund 2 composition is microfounded but estimated in reduced form; allowing the responsiveness parameter to vary with market tightness or organizational constraints would sharpen comparative statics. Prices clear one market at a time; incorporating intermediated capital or multiple LP clienteles could generate richer price dynamics. Finally, the planner internalizes diffusion and adoption but abstracts from broader general-purpose spillovers beyond the observed sectors; extending the welfare accounting to those externalities would enlarge the set of policy instruments.

Despite these caveats, the core message is robust: when diffusion compresses rents, monitoring shapes information, and capital prices clear each vintage, private exploration is too low precisely when society most benefits from it. A structural approach that brings networks, incentives, and prices under one roof makes the misallocation measurable and actionable, and it provides concrete, implementable levers, contractual and systemic, to accelerate efficient exploration in high-uncertainty domains.

References

Axelson, U., T. Jenkinson, P. Strömberg, and M. S. Weisbach (2013). Borrow cheap, buy high? the determinants of leverage and pricing in buyouts. *The Journal of Finance 68*(6), 2223–2267.

Banerjee, A. V. (1992). A simple model of herd behavior. *Quarterly Journal of Economics 107*(3), 797–817.

Bergemann, D. and U. Hege (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking and Finance 22*(6–8), 703–735.

Bergemann, D. and U. Hege (2005). The financing of innovation: Learning and stopping. *RAND Journal of Economics 36*(4), 719–752.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy 100*(5), 992–1026.

Cabolis, C., M. Dai, and K. Serfes (2020). Competition and specialization: Evidence from venture capital. *SSRN Electronic Journal*. Working paper, latest version circa 2020.

Cornelli, F. and O. Yosha (2003). Stage financing and the role of convertible securities. *Review of Economic Studies 70*(1), 1–32.

Dixit, A. K. and R. S. Pindyck (1994). *Investment under Uncertainty*. Princeton, NJ: Princeton University Press.

Gompers, P., A. Kovner, and J. Lerner (2009). Specialization and success: Evidence from venture capital. *Journal of Economics & Management Strategy 18*(3), 817–844.

Gompers, P., A. Kovner, J. Lerner, and D. Scharfstein (2008). Venture capital investment cycles: The impact of public markets. *Journal of Financial Economics 87*(1), 1–23.

Gompers, P. A. and J. Lerner (2000). *The Venture Capital Cycle*. Cambridge, MA: MIT Press.

Hellmann, T. (1998). The allocation of control rights in venture capital contracts. *RAND Journal of Economics 29*(1), 57–76.

Hochberg, Y. V., A. Ljungqvist, and Y. Lu (2007). Whom you know matters: Venture capital networks and investment performance. *The Journal of Finance 62*(1), 251–301.

Hochberg, Y. V., A. Ljungqvist, and Y. Lu (2010). Networking as a barrier to entry and the competitive supply of venture capital. *The Review of Financial Studies 23*(7), 2511–2543.

Kaplan, S. N. and A. Schoar (2005). Private equity performance: Returns, persistence, and capital flows. *The Journal of Finance 60*(4), 1791–1823.

Kaplan, S. N. and P. Strömberg (2003). Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. *Review of Economic Studies 70*(2), 281–315.

Kaplan, S. N. and P. Strömberg (2004). Characteristics, contracts, and actions: Evidence from venture capitalist analyses. *The Journal of Finance 59*(5), 2177–2210.

Lerner, J. (1995). Venture capitalists and the oversight of private firms. *The Journal of Finance 50*(1), 301–318.

Manso, G. (2011). Motivating innovation. *The Journal of Finance 66*(5), 1823–1860.

Nanda, R. and M. RhodesKropf (2013). Investment cycles and startup innovation. *Journal of Financial Economics 110*(2), 403–418.

Nanda, R. and M. RhodesKropf (2020). Venture capital booms and startup financing. Working paper version circulated in 2020; published as Annual Review of Financial Economics (2021).

Pastor, L. and P. Veronesi (2009). Technological revolutions and stock prices. *American Economic Review 99*(4), 1451–1483.

Sahlman, W. A. (1990). The structure and governance of venturecapital organizations. *Journal of Financial Economics 27*(2), 473–521.

Sorenson, O. and T. E. Stuart (2001). Syndication networks and the spatial distribution of venture capital investments. *American Journal of Sociology 106*(6), 1546–1588.

Appendix: Proofs, Derivations, Identification, and Implementation

A Notation, Standing Assumptions, and Regularity

Unless stated otherwise, all random variables are integrable and all decision sets are compact intervals. Functions $s(\cdot)$, $R_{j,it}(\cdot)$, and $m(\cdot)$ are continuously differentiable with $s'(m) > 0$, $s''(m) < 0$, $s(0) = 1$, and $m(\mu^{\text{eff}}) = \bar{m}_t e^{-\rho\mu^{\text{eff}}}$ with $\rho > 0$ and $\bar{m}_t > 0$. Cost parameters satisfy $\phi_i > 0$, $\phi_L > 0$, $c_i > 0$, and the scale curvature obeys $\chi(\alpha) \geq \underline{\chi} > 0$ for all $\alpha \in [0, 1]$. Regime transition probabilities $p_{HH}, p_{CC} \in (0, 1)$.

Learning and diffusion follow

$$L_{i,t} = (1 - \delta_L)L_{i,t-1} + \theta_L \alpha_{1,it}, \qquad K_t = (1 - \delta_K(M_{t-1}))K_{t-1} + \theta(M_{t-1})\,\mathbb{E}[\alpha_{1,\cdot,t-1}],$$

with $\delta_L, \delta_K(\cdot) \in (0, 1)$ and $\theta_L, \theta(\cdot) > 0$. Adoption compresses new-industry rents via $R_n(A) = R_{n0}(1 - \theta_A A)$ with $\theta_A > 0$, and core returns satisfy $p_c(\text{H})R_c(\text{H}) < p_c(\text{C})R_c(\text{C})$.

B Entrepreneur Incentives and the Monitoring Threshold

This section derives a closed-form expression for the E-IC threshold used in the main text and states its feasibility conditions.

B.1 Derivation

Let $B \equiv s_E^{(j)} r_n + U_c^{(j)}$ denote the founder's per-project benefit under truthful success. Truthful reporting yields

$$U^{\text{true}} = \pi_{\text{true}} B - c_e.$$

A deceptive founder is detected with probability $d(\mu^{\text{eff}}) = 1 - e^{-\rho\mu^{\text{eff}}}$ and obtains

$$U^{\text{dec}} \;=\; (1-d)B - d\,\Pi \;=\; e^{-\rho\mu^{\text{eff}}}B - \left(1 - e^{-\rho\mu^{\text{eff}}}\right)\Pi \;=\; e^{-\rho\mu^{\text{eff}}}(B+\Pi) - \Pi.$$

E-IC requires $U^{\text{true}} \geq U^{\text{dec}}$, i.e.,

$$\pi_{\text{true}}B - c_e \;\geq\; e^{-\rho\mu^{\text{eff}}}(B+\Pi) - \Pi \quad\Longleftrightarrow\quad e^{-\rho\mu^{\text{eff}}} \;\leq\; \frac{\pi_{\text{true}}B - c_e + \Pi}{B+\Pi}.$$

If the right-hand side lies in $(0,1)$,[2] then

$$\mu^{\text{eff}} \;\geq\; \mu_E^{\text{thresh}}(B,\Pi) \;\equiv\; \frac{1}{\rho}\,\log\!\left(\frac{B+\Pi}{\pi_{\text{true}}B - c_e + \Pi}\right). \tag{B.1}$$

B.2 Comparative statics

The threshold is decreasing in B and Π, and increasing in c_e:

$$\frac{\partial\mu_E^{\text{thresh}}}{\partial B} < 0, \qquad \frac{\partial\mu_E^{\text{thresh}}}{\partial \Pi} < 0, \qquad \frac{\partial\mu_E^{\text{thresh}}}{\partial c_e} > 0.$$

Since B rises with founder equity $s_E^{(j)}$ and continuation utility $U_c^{(j)}$, larger founder upside strictly *reduces* the monitoring required to deter deception.

C GP Monitoring: Closed Form and Properties

This section proves the closed-form solution for the GP's optimal monitoring and establishes uniqueness and curvature.

[2] A primitive restriction ensuring the RHS $\in (0,1)$ is $\Pi > c_e$ and $\pi_{\text{true}}B \in (0,B)$; empirically I verify this by checking that observed monitoring intensities are interior and that detected deception is nonzero.

C.1 Problem and first-order condition

Given LP audit m and co-invest a, the GP chooses $\mu \geq 0$ to maximize

$$\Phi(\mu; m, a) = W(a)\left[\bar{m}_t - \bar{m}_t e^{-\rho\mu s(m)}\right]\Lambda_t - \frac{\phi_i}{2}\mu^2.$$

The derivative is

$$\Phi_\mu = W(a)\,\bar{m}_t\,\Lambda_t\,\rho s(m)\,e^{-\rho\mu s(m)} - \phi_i\,\mu.$$

The first-order condition (FOC) at an interior optimum $\mu^* > 0$ is

$$\phi_i\,\mu^* = W(a)\,\bar{m}_t\,\Lambda_t\,\rho s(m)\,e^{-\rho\mu^* s(m)}. \tag{C.1}$$

C.2 Closed-form solution via Lambert-W

Multiply both sides of (C.1) by $e^{\rho\mu^* s(m)}$ and divide by ϕ_i:

$$\mu^* e^{\rho\mu^* s(m)} = \frac{\rho s(m)\,W(a)\,\bar{m}_t\,\Lambda_t}{\phi_i}.$$

Set $x \equiv \rho s(m)\mu^*$ so that $xe^x = \frac{\rho s(m)\,W(a)\,\bar{m}_t\,\Lambda_t}{\phi_i}$. The principal branch of the Lambert function solves $x = W_0(\cdot)$. Therefore

$$\mu^*(m, a) = \frac{1}{\rho s(m)}\,W_0\!\left(\frac{\rho s(m)\,W(a)\,\bar{m}_t\,\Lambda_t}{\phi_i}\right). \tag{C.2}$$

Because the argument is strictly positive and $W_0(\cdot)$ is increasing on $(0, \infty)$, μ^* is unique and interior.

C.3 Second order condition and concavity

I have

$$\Phi_{\mu\mu} = -W(a)\,\bar{m}_t\,\Lambda_t\,(\rho s(m))^2 e^{-\rho\mu s(m)} - \phi_i < 0$$

for all $\mu \geq 0$, so $\Phi(\cdot)$ is strictly concave and the FOC is sufficient.

C.4 Comparative statics

Let $z \equiv \frac{\rho s(m)\, W(a)\, \bar{m}_t\, \Lambda_t}{\phi_i} > 0$. Then

$$\mu^* \;=\; \frac{1}{\rho s(m)}\, W_0(z).$$

Using $W_0'(z) = \frac{W_0(z)}{z(1+W_0(z))}$,

$$\frac{\partial \mu^*}{\partial m} \;=\; -\frac{s'(m)}{\rho s(m)^2}\, W_0(z) \;+\; \frac{1}{\rho s(m)}\, W_0'(z)\, \frac{\partial z}{\partial m}.$$

Since $\frac{\partial z}{\partial m} = \frac{\rho s'(m)\, W(a)\, \bar{m}_t\, \Lambda_t}{\phi_i} > 0$ and $s'(m) > 0$, the second term dominates the first at all interior points.[3] Hence $\partial \mu^*/\partial m > 0$. Similarly,

$$\frac{\partial \mu^*}{\partial a} \;=\; \frac{1}{\rho s(m)}\, W_0'(z)\, \frac{\partial z}{\partial a} \;>\; 0, \qquad \frac{\partial \mu^*}{\partial \phi_i} \;<\; 0, \qquad \frac{\partial \mu^*}{\partial \Lambda_t} \;>\; 0.$$

Effective monitoring is $\mu^{\text{eff}} = \mu^* s(m)$; since both μ^* and $s(m)$ rise in m, audit has direct and indirect effects.

D LP Best Responses: Audit and Scale

This section derives the LP audit FOC and the linear scale rule used in the main text.

[3] At very low m, the first term is small because $W_0(z) \approx z$ and z is $O(m)$; at high m, $W_0(z)$ grows slowly (logarithmically), while $s'(m)$ falls due to concavity, keeping the derivative positive. Numerically, $\partial \mu^*/\partial m > 0$ everywhere.

D.1 Scale

With per-fund utility

$$U_j^{\mathrm{LP}} = (1 - s_G^{(j)})\, I_j R_j(\alpha_j) - (1 - W_j)\, m(\mu_j^{\mathrm{eff}})\Lambda_t - r_L I_j - \frac{\chi(\alpha_j)}{2} I_j^2 - \frac{\phi_L}{2} m_j^2,$$

the derivative with respect to I_j is

$$\frac{\partial U_j^{\mathrm{LP}}}{\partial I_j} \;=\; (1 - s_G^{(j)})\, R_j(\alpha_j) - r_L - \chi(\alpha_j)\, I_j.$$

Setting to zero gives the linear rule

$$I_j^* \;=\; \frac{(1 - s_G^{(j)})\, R_j(\alpha_j) - r_L}{\chi(\alpha_j)} \quad \text{if the numerator is positive, else } I_j^* = 0. \tag{D.1}$$

D.2 Audit

Treat the GP as a follower who sets $\mu^*(m)$ by (C.2). The LP minimizes expected deception losses net of audit cost on this reduced form:

$$\min_{m \geq 0} \; (1 - W_j)\, m\big(\mu^*(m)s(m)\big)\Lambda_t + \frac{\phi_L}{2} m^2$$

with $m(x) = \bar{m}_t e^{-\rho x}$. Differentiating and using the chain rule,

$$\phi_L m^* \;=\; (1 - W_j)\, \rho\, \bar{m}_t\, \mu^*(m^*)\, s'(m^*)\, e^{-\rho\, \mu^*(m^*)s(m^*)}\, \Lambda_t,$$

which is the condition reported in the main text. The right-hand side is positive and strictly decreasing in m for large m because $s'(\cdot)$ is concave and $e^{-\rho\, \mu^{\mathrm{eff}}}$ decays; thus a unique interior m^* exists provided the marginal loss-reduction at $m = 0$ exceeds zero.

E Fund–2 Composition with Endogenous Scale

Let $V_2(\alpha_2) \equiv s_G^{(2)} I_2^*(\widehat{R}_2) R_2(\alpha_2) - c_i \alpha_2$, where I_2^* is given by (D.1) with $j = 2$ and $\widehat{R}_2$ is the LP's forecast. On the interior region where $I_2^* > 0$,

$$I_2^* \; = \; \frac{(1 - s_G^{(2)})\, \widehat{R}_2 - r_L}{\chi(\alpha_2)}.$$

Applying the product rule,

$$\frac{\partial V_2}{\partial \alpha_2} \; = \; s_G^{(2)} \left\{ \frac{\partial I_2^*}{\partial \widehat{R}_2} \frac{\partial \widehat{R}_2}{\partial \alpha_2} R_2(\alpha_2) + I_2^* \frac{\partial R_2}{\partial \alpha_2} \right\} - c_i.$$

Using $\frac{\partial I_2^*}{\partial \widehat{R}_2} = \frac{1 - s_G^{(2)}}{\chi(\alpha_2)}$,

$$\frac{\partial V_2}{\partial \alpha_2} \; = \; \frac{s_G^{(2)}(1 - s_G^{(2)})}{\chi(\alpha_2)}\, R_2(\alpha_2)\, \frac{\partial \widehat{R}_2}{\partial \alpha_2} \; + \; \frac{s_G^{(2)}}{\chi(\alpha_2)} \big((1 - s_G^{(2)}) \widehat{R}_2 - r_L \big)\, \frac{\partial R_2}{\partial \alpha_2} \; - \; c_i. \tag{E.1}$$

Equation (E.1) is the interior FOC reported in the paper. If $(1 - s_G^{(2)})\widehat{R}_2 \leq r_L$ then $I_2^* = 0$ and the GP's best response is $\alpha_2^* = 0$ due to the linear cost $c_i \alpha_2$.

Derivative of R_2. With

$$R_2(\alpha_2) = (1 - \alpha_2)\, p_c R_c + \alpha_2\, p_n^{(2)}(\alpha_1, \alpha_2, K_2)\, R_n(A_2, K_2),$$

I obtain

$$\frac{\partial R_2}{\partial \alpha_2} = -\, p_c R_c + p_n^{(2)} R_n + \alpha_2\, R_n\, (1 - p_0^{(2)}) e^{-(\lambda \alpha_1 + \tilde{\lambda} \alpha_2 + \varphi K_2)}\, \tilde{\lambda} + \alpha_2\, p_n^{(2)}\, \frac{\partial R_n}{\partial A_2} \frac{\partial A_2}{\partial \alpha_2}.$$

In an atomistic-GP approximation, the last term is negligible; otherwise it embeds price-mediated rent compression via adoption.

F Equilibrium Existence

I sketch existence for the two-fund game under the timing in the paper. Define the strategy space

$$\mathcal{X} \equiv [0,1] \times [0,\bar{\mu}] \times [0,1] \times [0,\bar{\mu}] \times [0,\bar{m}] \times [0,\bar{m}] \times [0,\bar{I}]^2,$$

for $(\alpha_1, \mu_1, \alpha_2, \mu_2, m_1, m_2, I_1, I_2)$ with finite bounds $\bar{\mu}, \bar{m}, \bar{I}$. Given states (K_t, A_t, M_t) and parameters, the GP's problem in each fund is strictly concave in (μ_j, α_j): $\Phi(\mu_j)$ is strictly concave (Section C); and $R_j(\alpha)$ is concave due to the exponential form of p^{new} combined with a linear core term and affine costs. The LP's problem is strictly concave in (I_j, m_j) due to the quadratic costs and linear-in-I_j returns. Thus best responses are single-valued and continuous. The resulting best-response mapping $T : \mathcal{X} \to \mathcal{X}$ is continuous and maps $\mathcal{X}$ into itself; by Brouwer's fixed-point theorem an equilibrium exists.

A complementary route is monotone comparative statics: (α_1, I_2) are strategic complements through the discipline channel $Z_t \mapsto \widehat{R}_2 \mapsto I_2$ when $r_Z > 0$; under standard lattice arguments this guarantees the existence of extremal equilibria (Tarski). I adopt the Brouwer route for uniqueness of the intra-fund choices and for implementation.

G Microfoundations for the Fund–2 Composition Rule

This section provides the derivations referenced in the main text.

G.1 Random-utility project assignment (logit share)

Suppose the GP screens a large set of atomistic projects with net dollar values $v^N = V^N + \varepsilon^N$ and $v^C = V^C + \varepsilon^C$, where $(\varepsilon^N, \varepsilon^C)$ are i.i.d. Type-I EV. A project is assigned to the new category iff $v^N \geq v^C$. The assignment probability is

$$\Pr(v^N \geq v^C) = \frac{e^{V^N}}{e^{V^N} + e^{V^C}} = \sigma(V^N - V^C).$$

As the number of projects grows, the realized share α_2 converges to this probability. Taking $V^N \equiv p_n R_n$ and $V^C \equiv p_c R_c$ and linearizing $V^N - V^C$ in (α_1, K, A) about a reference point yields the logistic index used in estimation.

G.2 Entropy-regularized continuous portfolio (softmax)

Consider

$$\max_{\alpha\in(0,1)} \alpha V^N + (1-\alpha)V^C - \tau\{\alpha\log\alpha + (1-\alpha)\log(1-\alpha)\}.$$

The FOC is $V^N - V^C = \tau \log\frac{\alpha}{1-\alpha}$, hence $\alpha = \sigma\big((V^N - V^C)/\tau\big)$. As above, linearizing the value difference in (α_1, K, A) generates the logistic mapping.

H Identification Lemmas

This section formalizes the informal arguments in Section 4.4.

Lemma 1 (Layer separation)**.** *Suppose audit m varies due to LP policy shocks that do not directly affect $W(\cdot)$, and co-invest a varies due to co-investment program changes that do not directly affect $s(\cdot)$. Then the map*

$$(m,a) \;\mapsto\; \mu^*(m,a) \;=\; \frac{1}{\rho s(m)}\, W_0\!\left(\frac{\rho s(m)\, W(a)\, \bar{m}_t\, \Lambda_t}{\phi_i}\right)$$

is jointly identified up to $(\rho, \bar{m}_t\Lambda_t/\phi_i)$, and the audit amplifier $s(\cdot)$ is point-identified from cross-fund differences in $m \mapsto \mu^{\text{eff}} = \mu^* s(m)$.

Proof. Hold a fixed and vary m: the ratio $\frac{\mu^{\text{eff}}}{\mu^*}$ identifies $s(m)$ nonparametrically because μ^* is observed via monitoring proxies and m via audit policy. Holding m fixed and varying a shifts the argument of W_0 multiplicatively through $W(a)$, which is then recovered up to the scale factor $\rho\bar{m}_t\Lambda_t/\phi_i$; the latter cancels in cross-fund differences when $(\bar{m}_t, \Lambda_t, \phi_i)$ are stable across funds or controlled for. ■

Lemma 2 (Discipline slope)**.** *On the interior region* $(1 - s_G^{(2)})\widehat{R}_2 > r_L$*, the reduced-form regression of* I_2 *on* Z_t *identifies* $r_Z\,(1 - s_G^{(2)})/\chi(\alpha_2)$*; independent variation in* $s_G^{(2)}$ *and* α_2 *(entering* $\chi(\alpha_2)$*) disentangles* r_Z*.*

Proof. Immediate from $I_2^* = ((1 - s_G^{(2)})\,\widehat{R}_2 - r_L)/\chi(\alpha_2)$ and the linearity of $\widehat{R}_2$ in Z_t. ■

Lemma 3 (Learning vs. adoption)**.** *Curvature of success in* (α_1, α_2) *identifies the exponential slopes in* p^{new}*; slopes of returns in* A *identify* θ_A*. Leave-one-out constructions of* K *(excluding GP* i*) deliver separate identification of* λ_i, η_L *(private learning) and* ρ_K *(market knowledge).*

Proof. The cross-partial $\partial^2 p^{\text{new}}/\partial\alpha_1\partial\alpha_2$ loads on $(1 - p_0)\lambda\tilde{\lambda}e^{-(\lambda\alpha_1+\tilde{\lambda}\alpha_2+\varphi K)}$, while $\partial R/\partial A = -\theta_A R_{n0}(1 - \cdot)$; these objects vary independently across cohorts/years, and $K^{(-i)}$ removes the mechanical correlation between K and i's own α_1. ■

I Algorithmic Solution and Simulation Details

Single-GP, given states. (i) Compute μ_1^* from (C.2); if $\mu_1^{\text{eff}} < \mu_E^{\text{thresh}}$ from (B.1), minimally increase m_1 or a_1 to meet E-IC. (ii) Solve α_1^* from the one-dimensional FOC (**??**) by bracketing and Newton steps (derivatives are available analytically). (iii) Update states (L, K) by (**??**); if adoption is simulated, update A. (iv) Form Z_t and $\widehat{R}_2$ and compute I_2^* by (D.1); if interior, solve α_2^* using (E.1).

Panel with market-clearing (optional). Given exogenous supply $\overline{\mathcal{I}}_{z,t}$, choose $(r_{f,t}, k_{I,t})$ so that $\sum_i I_{2,i,t}^* = \overline{\mathcal{I}}_{z,t}$; a single-dimension version fixes $k_{I,t}$ and solves for $r_{f,t}$ by bisection (monotone demand), or jointly solves the two prices by matching both level and dispersion.

SMM. For each parameter candidate Θ, simulate the panel paths and compute moments: (1) semi-elasticity of I_2 in Z_t; (2) curvature of α_2 in (α_1, K, A); (3) hot–cold contrasts; (4) monitoring responses to (m, a). The objective is the quadratic form with the optimal weighting matrix; cluster/bootstrap by GP.

J Robustness: Alternative Monitoring and Scale Technologies

Non-exponential detection. If $d(\mu^{\text{eff}}) = 1 - (1 + \rho\mu^{\text{eff}})\, e^{-\rho\mu^{\text{eff}}}$ (Weibull-like), the FOC becomes $\phi_i \mu = W(a)\bar{m}_t \Lambda_t \rho s(m)\,(1 + \rho\mu s(m))e^{-\rho\mu s(m)}$. Rearrangement still yields a Lambert-W form with a shifted argument; comparative statics are unchanged.

Convex scale cost with intercept. If $C(I, \alpha) = \kappa_0(\alpha)I + \frac{1}{2}\kappa_1(\alpha)I^2$, then

$$I^* = \max\left\{0,\ \frac{(1 - s_G)R - r_L - \kappa_0(\alpha)}{\kappa_1(\alpha)}\right\}.$$

All discipline results go through with $\kappa_0(\alpha)$ absorbed into the intercept of $\widehat{R}_2$.

K Additional Lemmas and Comparative Statics

Lemma 4 (Hot–cold tilt of Fund 1)**.** *If $p_c(\text{H})R_c(\text{H}) < p_c(\text{C})R_c(\text{C})$ and adoption A_t is below a threshold $\bar{A}$, then $\alpha_1^*(\text{H}) > \alpha_1^*(\text{C})$.*

Proof. From (**??**), the marginal return difference is ΔR_t + learning term. In a hot market, ΔR_t is strictly larger by the core compression; if A_t is small, $R_n(A_t)$ is close to R_{n0}, so the learning term is also larger due to the R_n factor. Hence the LHS of (**??**) shifts up, raising the solution α_1^*. ■

Lemma 5 (Substitution between audit and co-invest)**.** *Fix the E-IC target $\mu^{\text{eff}} \geq \bar{\mu}$ and suppose $s(\cdot)$ is strictly concave. Then there exists a decreasing schedule $a^*(m)$ such that the minimal joint cost $\frac{\phi_L}{2}m^2 + \kappa_a a$ (for any $\kappa_a > 0$) is achieved by substituting co-invest for audit at high m.*

Proof. From $\mu^{\text{eff}} = \frac{s(m)}{\rho s(m)} W_0\big(\rho s(m)W(a)\bar{m}\Lambda/\phi_i\big)$, holding μ^{eff} fixed and differentiating yields $W'(a)\, da + s'(m)\, dm = 0$ scaled by positive factors. Since $s'(\cdot)$ falls with m (concavity), the marginal benefit of audit declines and co-invest becomes relatively cheaper at high m; the cost minimization therefore sets $da/dm < 0$ for large m. ■

L Proof of Proposition ?? (Main Text)

The GP's intra-fund problems are strictly concave in (μ_j, α_j) by Sections C and E; the LP's problems are strictly concave in (I_j, m_j) by Section D. Hence best responses are unique and continuous. Define T as the composition of GP and LP best responses given states (K, A, M). T is continuous and maps a compact convex set into itself, so Brouwer's fixed-point theorem applies. Monotone parameter shifts (e.g., higher $s_G^{(2)}$ or lower χ) move the fixed point in the directions stated in the proposition by the maximum theorem.

M Estimation Details and Moment Construction

M.1 Discipline regression

On the interior region, estimate

$$I_{2,i,t} \;=\; \delta_0 + \delta_1 R_{1,i,t} + \delta_2 \alpha_{1,i,t} + \delta_3 \mu^{\text{eff}}_{1,i,t} + \Xi' X_{i,t} + \eta_{s\times t} + \eta_i + \varepsilon_{i,t},$$

where $X_{i,t}$ includes K_t, A_{t+1}, and $\mathbf{1}\{M_{t+1} = \text{H}\}$. The coefficients map into $(1 - s_G^{(2)}) r_Z / \chi(\alpha_2)$ and the γ_k in (**??**)–(**??**). Interactions with regime dummies identify state dependence.

M.2 Monitoring response

Estimate

$$\mu_{j,i,t} \;=\; \zeta_0 + \zeta_1 s(m_{j,t}) + \zeta_2 W(a_{j,i,t}) + \zeta_3 \log \Lambda_t + \Xi' X_{i,t} + \eta_{s\times t} + \eta_i + u_{i,t}$$

with instruments for m (LP policy and compliance shocks) and for a (co-invest program rules). Nonlinear least squares on (C.2) is preferred when instruments are strong, otherwise use a control-function approach.

M.3 SMM moments

I use (i) regression slopes from the discipline regression; (ii) the cross-sectional variance of I_2 conditional on Z_t; (iii) the curvature of success in (α_1, α_2); (iv) the elasticity of μ in (m, a); and (v) hot–cold differences in (α_1, I_2). Moments are computed identically in the simulated panel.

N Additional Discussion: Planner vs. Decentralized Wedges

The decentralized economy under-invests in monitoring in both funds because the GP internalizes only $W_j(a) < 1$ of avoided deception losses. It also under-invests in early exploration because the K externality is not priced, while over-scaling Fund 2 when LPs ignore the adoption externality (if the atomistic approximation for adoption is not exact). The planner raises μ, increases α_1 early (especially when hot), and moderates I_2 when $\chi(\alpha_2)$ is steep. In counterfactuals, raising W_2 via co-invest (or performance-bond style contracts) closely mimics the planner along the monitoring margin at relatively low deadweight cost, while subsidizing audit m is complementary when $s(\cdot)$ exhibits strong concavity.

References for Appendix Methods. I follow standard properties of the Lambert-W function and monotone comparative statics. All proofs of existence rely on Brouwer fixed points and/or supermodularity arguments on lattices. Numerical routines employ bracketing with safeguarded Newton steps for one-dimensional FOCs and bisection for market-clearing prices.

N.1 Proof of Proposition 9.1 (Clearing)

Fix $(\alpha_1, \mu_1, \alpha_2)$ and z_t. Each $I_i^*(r_f, k_I)$ in (9.7) is continuous and weakly decreasing in r_f, k_I, strictly decreasing on any range where $I_i^* > 0$. Summation preserves continuity and strict monotonicity on sets of positive mass. Limits follow from linear–quadratic structure: as $r_f \to \infty$, net margins $(1 - \phi^{\text{eff}})r_2 - r_f \to -\infty$ so $I_i^* \to 0$; as $r_f \downarrow -\infty$, the margin diverges and $I_i^* \to \infty$. By the intermediate value theorem there is a solution; by strict monotonicity it is unique. The argument for k_I is analogous.

N.2 Proof of Proposition 9.2 (Existence)

Given (K_t, A_t), each GP's problem is concave in (μ_1, α_1) due to the quadratic monitoring cost and linear returns; under Case B, the α_2 subproblem is strictly concave (quadratic inertia) or single–valued via logistic mapping (Case A). Hence best responses are non–empty and upper hemi–continuous. Proposition 9.1 pins (r_f, k_I) uniquely. Finally, (9.5)–(9.6) are continuous state maps with $0 < \delta_K < 1$ and $A \in [0, 1]$. By sequential construction (or Kakutani on the period problem) an equilibrium exists.

N.3 Comparative Statics with Policy Knobs

Under Case A,

$$\frac{\partial \alpha_{2i}}{\partial x} = \sigma_i(1-\sigma_i)\left(a_{2,1}\frac{\partial \alpha_{1i}}{\partial x} + a_{2,K}\frac{\partial K_s}{\partial x} - a_{2,A}\frac{\partial A_s}{\partial x}\right),$$

for any scalar policy $x \in \{\tau_{kI}, g_{rf}, \phi_1\}$. The LP channel delivers $\partial I_i^*/\partial r_f = -\mathbb{1}\{I_i^* > 0\}/k_I$ and $\partial I_i^*/\partial k_I = -\mathbb{1}\{I_i^* > 0\}\,[(1-\phi^{\text{eff}})r_2 - r_f]/k_I^2$, so $\partial I^*/\partial g_{rf} = -\alpha_{2i}/k_I + \frac{(1-\phi^{\text{eff}})}{k_I}\,\partial r_2/\partial g_{rf}$. General–equilibrium terms propagate through clearing (r_f, k_I) and diffusion (K, A).

N.4 Peer–Effect Identification

Consider

$$\alpha_{2i,t} = \beta_1\, E^{(-i)}_{c(i),t-1} + \beta_2\, K_{s(i),t} + \gamma_{s\times t} + \gamma_i + \varepsilon_{i,t},$$

with LOO cohort exposure $E^{(-i)}_{c,t-1}$ and fixed effects. Suppose (i) cohort composition $B_{s,c,t-1}$ and mixing M are predetermined; (ii) shocks $\varepsilon_{i,t}$ are mean–independent of (B, M) conditional on FE; (iii) large–N thin cells so no single GP drives $E^{(-i)}$. Then OLS on LOO is consistent for β_1 up to sampling error. For over–identification, construct shift–share instruments $Z_{i,t} \equiv \sum_{c'} \omega_{i,c'}\, E_{c',t-1}$ with weights $\omega_{i,c'}$ functions of (B, M), which are excluded from $\varepsilon_{i,t}$ by (i)–(ii). Relevance follows from the network in (9.4); exogeneity follows from the predetermined nature of (B, M) and FE.

N.5 Network Stability and Knowledge Paths

With fixed W, iterating (9.5) yields $K_{t+T} = (1-\delta_K)^T K_t + \chi_K \sum_{j=0}^{T-1}(1-\delta_K)^j W E_{t+T-1-j}$. If $\|W\|_\infty \le 1$ and $\sup_t \|E_t\|_\infty < \infty$, then K_t is bounded and the transient decays at rate $(1-\delta_K)^T$. If $E_t \to E^\star$, then K_t converges to $K^\star = \frac{\chi_K}{\delta_K} W E^\star$.

N Appendix: Microfoundations for α_2

MF–1: Random–Utility Assignment (Logit Share)

See Subsection 9.10. The discrete–choice assignment with i.i.d. Type I EV shocks yields $\alpha_2 = \sigma(\Delta V)$ with $\Delta V \equiv V^N - V^C$; linearizing ΔV around $(\bar{\alpha}_1, \bar{K}, \bar{A})$ maps to the logistic index with slopes $(a_{2,1}, a_{2,K}, a_{2,A})$; signs: $a_{2,1} > 0$, $a_{2,A} > 0$, $a_{2,K}$ ambiguous.

MF–2: Entropy–Regularized Portfolio (Softmax)

The regularized problem delivers $\alpha_2 = \sigma((V^N - V^C)/\tau)$; τ aggregates unmodeled frictions (menu costs, delegation noise) and scales the logistic slopes.

N Appendix: A Quadratic–Deviation Microfoundation for α_2

N.1 Setup and Solution

Let the GP face a logistic *target* $\bar{\alpha}_{2i,t} = \sigma(z_{i,s,t})$ with $z_{i,s,t} = a_{2,0} + a_{2,1}\alpha_{1i,t} + a_{2,K}K_{s,t} - a_{2,A}A_{s,t}$. The GP implements

$$\max_{\alpha\in[0,1]} \ \alpha\, V_{i,t}^N + (1-\alpha)\, V_{i,t}^C - \tfrac{\kappa_2}{2}\left(\alpha - \bar{\alpha}_{2i,t}\right)^2, \qquad \kappa_2 > 0,$$

which has interior FOC $\Delta V_{i,t} - \kappa_2(\alpha_{2i,t} - \bar{\alpha}_{2i,t}) = 0$ and solution

$$\alpha^*_{2i,t} = \bar{\alpha}_{2i,t} + \frac{\Delta V_{i,t}}{\kappa_2} \quad \text{clamped to } [0,1]. \tag{N.1}$$

Limits: $\kappa_2 \to \infty$ recovers the pure logistic mapping; $\kappa_2 \downarrow 0$ yields bang–bang choice $\mathbf{1}\{\Delta V > 0\}$.

N.2 Link to Primitives and Identification

With $V^N = p_n R_n^{\text{eff}}$ and $V^C = p_c R_c$, a local expansion gives $\Delta V \approx b_0 + b_1\alpha_1 + b_K K - b_A A$ with $b_1 > 0$, $b_A > 0$, and b_K ambiguous. Substituting into (N.1) yields the estimable form

$$\alpha_2^* = \sigma\big(a_{2,0} + a_{2,1}\alpha_1 + a_{2,K}K - a_{2,A}A\big) + \frac{b_0 + b_1\alpha_1 + b_K K - b_A A}{\kappa_2} \quad \text{(clamped)}.$$

Instruments that move ΔV (via prices r_f, k_I or policy) holding z fixed identify κ_2 separate from $(a_{2,\cdot})$; shape tests follow from $\partial\alpha_2/\partial z = \sigma(z)(1-\sigma(z)) + \frac{1}{\kappa_2}\partial\Delta V/\partial z$.

N.3 Embedding LP Reaction and Clearing

If LPs respond via $I^*(r_2; r_f, k_I)$, interpret V^k as marginal GP values $\tilde{V}^k \equiv \phi^{\text{eff}} I^* r_2\big|_k$; then ΔV inherits dependence on (r_f, k_I) and on the clearing rule, tightening identification of κ_2 in policy episodes.

Table 6
Reduced-form (1/2): Fund-2 Composition, Scale Discipline, and Monitoring Production

	(1) Fund-2 Composition $\alpha_{2,i,t+1}$	**(2)** Scale on Score & Contract $I_{2,i,t+1}$	**(3)** Scale on Signals × Hot $I_{2,i,t+1}$	**(4)** Monitoring Production $\mu^{\text{eff}}_{1,i,t}$
Fund-1 choices/signals				
$\alpha_{1,i,t}$	0.322***		0.257***	
	(0.060)		(0.061)	
$\mu^{\text{eff}}_{1,i,t}$	0.114***		0.184***	
	(0.034)		(0.049)	
$R_{1,i,t}$			0.312***	
			(0.072)	
Score $Z_{i,t}$		0.476***		
		(0.067)		
Core-only LP		-0.118**		
		(0.049)		
$Z_{i,t}\times$ Core-only		-0.221***		
		(0.082)		
Sector state & regime				
$K^{(-i)}_{s,t}$	0.182***			
	(0.048)			
$A_{s,t}$	-0.141***			
	(0.044)			
Hot_{t+1}			-0.093**	
			(0.038)	
$R_{1,i,t}\times\text{Hot}_{t+1}$			-0.121**	
			(0.055)	
$\alpha_{1,i,t}\times\text{Hot}_{t+1}$			-0.098**	
			(0.047)	
$\mu^{\text{eff}}_{1,i,t}\times\text{Hot}_{t+1}$			-0.077*	
			(0.041)	
Monitoring production & alignment				
Audit $m_{1,i,t}$				0.538***
				(0.091)
$m^2_{1,i,t}$				-0.114***
				(0.032)
Co-invest $a_{1,i,t}$				0.182***
				(0.058)
Hot_t				0.071**
				(0.029)
Other regressor(s)				
Fund-2 scale $I_{2,i,t+1}$	0.093***			
	(0.026)			
Fixed effects	GP; Sector×Year	GP; Vintage	GP; Vintage	GP; Vintage
Observations N	1,760	1,820	1,820	2,450
Adj. R^2	0.401	0.407	0.428	0.362

Description. Column (1) maps Fund 1 experimentation and information into Fund 2 composition (new-industry share), with sector knowledge $K^{(-i)}$ and adoption A capturing success vs. rent-compression channels. Column (2) shows discipline of Fund 2 scale to the Fund 1 score Z, and tests whether core-only LPs mute that discipline (level/slope shifts). Column (3) adds state dependence by interacting Fund 1 signals with the next-vintage hot regime. Column (4) estimates the monitoring production function: audit raises μ with diminishing returns; alignment via co-invest raises μ; Fund 1 hot states are associated with higher μ.
Notes. Coefficients with clustered SE (by GP) in parentheses. Hot_t refers to Fund 1 period; Hot_{t+1} to the next vintage. * p<.10, ** p<.05, *** p<.01.

Table 7

Reduced-form (2/2): Screening Quality, Performance, and Rent Compression

	(5) Screening: Pr(Pass)	**(6)** Post-pass Failure: $\mathbb{1}\{\text{Fail} \mid \text{Pass}\}$	**(7)** Performance: Hit rate	**(8)** Rents: Payoff per hit
Screening				
$\mu_{i,t}^{\text{eff}}$	0.061***	-0.042***		
	(0.015)	(0.012)		
Credibility $\xi_{i,t}$	0.037***	-0.028***		
	(0.011)	(0.009)		
Hot_t		0.015**		
		(0.007)		
Performance				
Fund-1 $\alpha_{1,i,t}$			0.081***	
			(0.027)	
$K_{s,t}^{(-i)}$			0.124***	
			(0.039)	
Fund-2 $\alpha_{2,i,t+1}$			0.208***	
			(0.052)	
Rents				
Adoption $A_{s,t+1}$				-0.173***
				(0.051)
Knowledge $K_{s,t+1}$				0.062**
				(0.028)
Fixed effects	GP; Sector×Year	GP; Sector×Year	GP; Sector×Year	GP; Sector×Year
Observations N	3,800	3,800	1,680	1,680
Adj. R^2	0.143	0.118	0.312	0.224

Description. Columns (5)–(6) quantify screening: higher effective monitoring and founder credibility both increase the probability of passing and reduce failure conditional on pass; hot states slightly raise failures-after-pass. Column (7) links Fund 1 exploration, sector knowledge, and Fund 2 composition to realized hit rates. Column (8) isolates rent compression: adoption lowers payoff-per-hit, while knowledge slightly raises it.

Notes. Columns (5)–(6) are linear probability models (LPM) with identical fixed effects. Standard errors clustered by GP. * $p<.10$, ** $p<.05$, *** $p<.01$.

Table 8
Structural Parameter Estimates (baseline)

Block	Param	Description	Estimate	Units	Identification
GP/LP shares	ϕ	GP revenue share (carry)	0.20	share	Targeted
Costs/IC	c_{mean}	Mean exploration cost	0.45	units	Targeted
Costs/IC	c_{sd}	SD exploration cost	0.15	units	Targeted
Costs/IC	u_{mean}	Mean monitoring disutility	0.90	units	Targeted
Costs/IC	u_{sd}	SD monitoring disutility	0.30	units	Targeted
Learning	λ_{mean}	Mean GP learning ability	0.80	index	Curvature
Learning	λ_{sd}	SD GP learning ability	0.25	index	Curvature
Monitoring layer	ρ_{ξ}	Corr(λ, ξ)	0.60	corr	MuXi
Monitoring layer	ξ_{mean}	Mean founder credibility	0.65	index	MuXi
Monitoring layer	ξ_{sd}	SD founder credibility	0.15	index	MuXi
Success (new)	$p_{0n,H}$	Baseline success prob (Hot)	0.22	prob	Targeted
Success (new)	$p_{0n,C}$	Baseline success prob (Cold)	0.28	prob	Targeted
Success (core)	$p_{c,H}$	Core success prob (Hot)	0.55	prob	Targeted
Success (core)	$p_{c,C}$	Core success prob (Cold)	0.62	prob	Targeted
Rents (new)	$R_{n,H}$	Per-hit payoff (Hot)	3.20	units	Targeted
Rents (new)	$R_{n,C}$	Per-hit payoff (Cold)	3.60	units	Targeted
Rents (core)	$R_{c,H}$	Per-hit payoff (Hot)	1.30	units	Targeted
Rents (core)	$R_{c,C}$	Per-hit payoff (Cold)	1.60	units	Targeted
Diffusion	ζ	Rent compression from adoption	1.10	per unit	Slope
Diffusion	η	Rent compression from knowledge	0.60	per unit	Slope
Learning	ϕ_K	Effect of K on new success	0.70	per unit	Targeted
Comp rule	$a_{2,0}$	Fund-2 new-share intercept	0.50	index	ShareReg
Comp rule	$a_{2,1}$	Slope on Fund-1 exploration	2.20	per unit	ShareReg
Comp rule	$a_{2,K}$	Slope on K	1.00	per unit	ShareReg
Comp rule	$a_{2,A}$	Slope on A (negative)	1.20	per unit	ShareReg
LP market	$r_{f,H}$	LP outside return (Hot)	0.45	rate	I2State
LP market	$r_{f,C}$	LP outside return (Cold)	0.55	rate	I2State
LP market	$k_{I,H}$	Scale cost κ (Hot)	0.35	per \$	I2State
LP market	$k_{I,C}$	Scale cost κ (Cold)	0.50	per \$	I2State
Contracts	π_{core}	Share of core-only LPs	0.25	share	Targeted
Dynamics	χ_K	Knowledge accumulation	0.25	rate	Targeted
Dynamics	χ_A	Adoption increment from switching	0.35	rate	Targeted
Dynamics	δ_K	Knowledge depreciation	0.25	rate	Targeted
Regimes	p_{HH}	$P(H \to H)$	0.72	prob	Targeted
Regimes	p_{CC}	$P(C \to C)$	0.78	prob	Targeted
Timing	$\overline{\text{gap}}$	Mean vintage gap (years)	3.00	years	Targeted
Timing	σ_{gap}	SD vintage gap	1.00	years	Targeted

Notes: Baseline SMM fit using GP and sector–year clusters. The table groups parameters by economic block and reports units to aid interpretation. See Appendix Table A.1 for standard errors and the J-statistic with p-value.

Table 9

Moment Fit: Data vs. Model

Moment	Data	Model	% Error	Targeted
Discipline (by state)				
Slope of I_2 on R_1 (Hot)	−2.262	−1.910	+15.6%	Y
Slope of I_2 on α_1 (Cold)	1.760	1.957	+11.2%	Y
Monitoring responses				
$\partial\mu/\partial m$ (Fund 1)	—	—	—	Y
$\partial\mu/\partial a$ (Fund 2)	—	—	—	Y
Learning curvature and diffusion				
Cross-partial in (α_1, α_2)	0.419	0.409	−2.4%	Y
Slope of Fund 2 success in α_1	0.415	0.441	+6.2%	Y
Slope of returns in A	0.146	0.092	−36.9%	Y
Path A_t (sector-year RMSE)	—	0.351	—	N
Path K_t (sector-year RMSE)	—	0.126	—	N
Levels and dispersion (auxiliary moments)				
$\mathbb{E}[\alpha_1]$	0.074	0.137	+85.6%	Y
$\mathrm{SD}[\alpha_1]$	0.238	0.315	+32.6%	Y
$\mathbb{E}[\alpha_2]$	0.220	0.469	+114.0%	Y
$\mathrm{SD}[\alpha_2]$	0.401	0.479	+19.4%	Y
$\mathbb{E}[\mu_1]$	0.0089	0.0106	+19.6%	Y
$\mathrm{SD}[\mu_1]$	0.0716	0.0566	−21.0%	Y
$\mathbb{E}[I_2]$	0.640	0.775	+21.0%	Y
$\mathrm{SD}[I_2]$	0.510	0.736	+44.5%	Y
Pr(switcher)	0.152	0.323	+113.0%	N
Pr(early)	0.026	0.0533	+105.0%	N

Notes: Entries show data moments, model-implied moments at $\hat{\theta}$, percentage error, and whether the moment entered the SMM objective. RMSE is reported in the moment's units.

Table 10

Decomposition of Exploration and Scale (levels by scenario)

	Baseline	No GP–E	No LP–GP	Planner	Total Δ	Shapley GP–E	Shapley LP–GP
α_2	0.403	0.493	0.642	1.000	0.597	0.090	0.239
I_2	0.595	1.430	5.672	7.550	6.955	0.835	5.077

Notes: Levels for Baseline and Planner are taken from the baseline simulation and planner normalization (Planner α_2=1.00; Planner I_2=7.55). Intermediate scenarios (No GP–E, No LP–GP) are reconstructed from the Shapley shares in the figure by adding the wedge's share of the total gap to the Baseline level. For α_2, GP–E and LP–GP shares are 15% and 40% of the gap, respectively. For I_2, GP–E and LP–GP shares are 12% and 73% of the gap, respectively. "Shapley" columns report level contributions (share × total gap), not percentages.