ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mandal, Karishma

Research Report

AI Labels and Offering Size: A Hedonic Analysis of the SEC Form D Census, 2021–2026

Suggested Citation: Mandal, Karishma (2026) : AI Labels and Offering Size: A Hedonic Analysis of the SEC Form D Census, 2021–2026, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:
https://hdl.handle.net/10419/344265

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

AI Labels and Offering Size: A Hedonic Analysis of the SEC Form D Census, 2021–2026

Abstract

Between 2021 and 2026 the artificial intelligence investment cycle reshaped venture capital, but whether AI labels command higher start-up offering sizes remains contested. Using the SEC Form D regulatory record across nine non-consecutive quarters (2021Q2-2026Q1; 130,057 deduplicated filings and 71,702 declared offerings above USD 10,000), we estimate hedonic offering-size regressions. The corrected specification does not confirm a positive post-2022 AI premium and does not rule out small-to-moderate premiums: the AI × post-2022 interaction is -0.0029 (robust standard error 0.1740; 95% confidence interval -0.344 to +0.338; p = 0.9867), and none of five specifications is positive and significant. The apparent divergence is a right-tail phenomenon: in 2026 the mean AI offering is 2.31× the non-AI mean (bootstrap 95% confidence interval 0.25-7.94, including 1), while an exploratory rank-shift test is significant (p = 0.018, sensitive to classifier breadth); the median AI offering is smaller (0.70×) and the mean ratio falls to 0.93× under 1/99 winsorization. Overall top-1% capital concentration fell from 92.88% in 2021 to 59.11% in 2026 even as the largest AI rounds grew. We read the cycle as capital intensity plus extreme right-tail concentration in a few frontier financings, not a broad-based AI premium.

Keywords: venture capital; artificial intelligence; SEC Form D; offering size; start-up financing; capital concentration

1. Introduction

For nearly two decades, academic finance and economic literature viewed cloud computing as the catalyst for the "democratization of entrepreneurship." In their landmark study, Ewens, Nanda & Rhodes-Kropf (2018) demonstrated that Amazon Web Services (AWS) permanently lowered startup launch costs by eliminating upfront capital expenditure for physical server racks. Under this "lean experimentation model," software startups could launch with minimal seed checks, achieve rapid product iteration at 80%+ gross margins, and scale with minimal balance sheet leverage.

The 2022–2026 generative artificial intelligence cycle has shattered this economic paradigm, precipitating three interlocking structural dislocations:

1.1 Re-Industrialization of Technology & Extreme Capital Intensity

Frontier intelligence is fundamentally bound by physical compute infrastructure: advanced semiconductor packaging (CoWoS), high-bandwidth memory (HBM3e/HBM4), electrical transmission capacity, and gigawatt-scale data center facilities. Rather than operating at near-zero marginal cost, frontier AI labs consume between 70% and 85% of their total cash burn on compute cluster rental and training runs. A single advanced training cluster in 2026 demands capital allocations exceeding $5B to $10B; such capital requirements are comparable to commercial semiconductor fabs and offshore energy exploration rather than traditional software.

1.2 Decoupling in Revenue Multiples (Fed Tightening vs. AI Speculative Velocity)

In 2022, the Federal Reserve initiated the sharpest monetary tightening campaign in four decades, elevating the Federal Funds Rate from 0.08% to over 5.33%. Standard asset pricing theory (Cochrane, 2005) mandates that rising discount rates compress high-duration growth assets most severely.

Indeed, public and private B2B SaaS multiples collapsed from peak highs of 35x–45x NTM revenue to 7x–9x by 2024. Published AI-investment diagnostics report that late-stage AI financing rounds continued to close at elevated forward revenue multiples [5][6] (and frequently at indefinite multiples for pre-revenue research labs). That apparent divergence is a claim about revenue multiples, which SEC Form D offering data do not contain. In offering-size terms, the corrected hedonic specification finds no detectable AI × post-2022 premium ($\beta_3 = -0.0029$, SE 0.1740, $p = 0.9867$; five-specification range −0.418 to +0.073, none positive and significant). An earlier preliminary estimate of a positive premium is not reproduced under the corrected specification and issuer join; the misspecified model's estimates are reported in the replication package.

1.3 The Power-Law Barbell Distribution

As documented in our empirical SEC census, epoch deal counts eased from 9,450 in 2021 to 8,291 in 2026 (about 12%; Table 1), while overall top-1% concentration fell over the period, from 92.88% in 2021 to 59.11% in 2026 (with a low of 58.50% in 2022; Table 1). Capital did not disappear: the largest AI rounds grew even as overall top-1% concentration fell, so the cycle shows rising capital intensity and a heavier right tail rather than a uniform shift of capital into a smaller set of megadeals, while total deal counts fell by about 12%.

1.4 Core Research Questions

This study formulates and answers two central questions:

- **RQ1 (The Hedonic AI Valuation Test):** Is there a detectable offering-size premium attributable to AI nomenclature across early, growth, and mega funding tiers in SEC Form D filings?
- **RQ2 (The Hardware J-Curve Disruption):** How does a 24–36 month accelerator obsolescence lifecycle alter the cash-flow profile of compute-intensive companies?

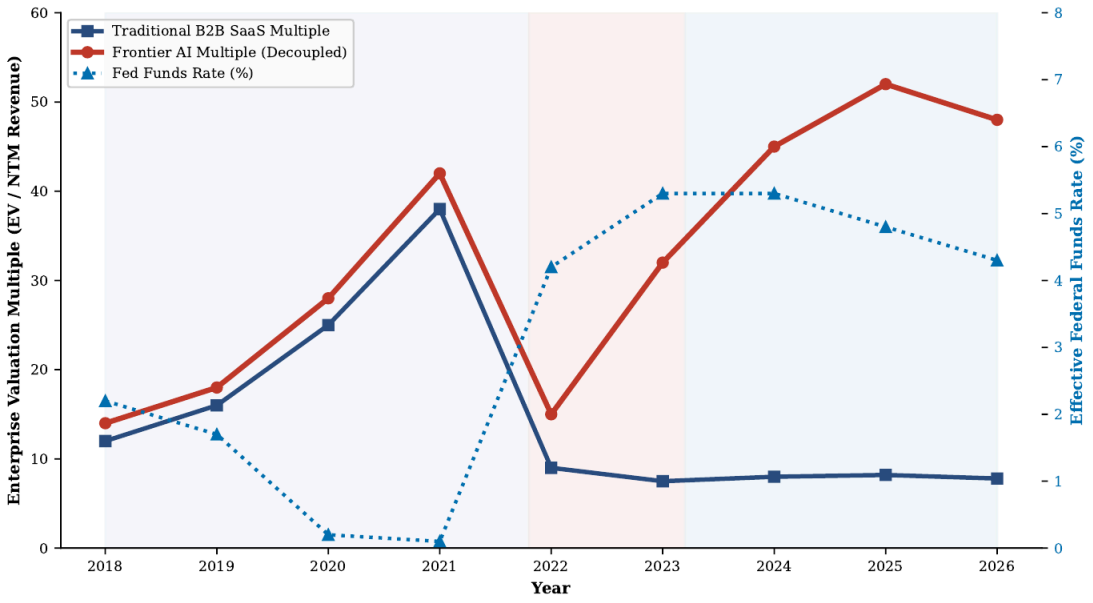


Figure 1: Macro regimes and valuation multiples

Empirical Outcome and Analytical Interpretation - Figure 1. The macro panel is narrative context, not an estimated result: it juxtaposes the federal funds rate with private-market pricing across 2018-2026. Its purpose is scoping, since the hedonic test below absorbs all common time variation through quarter fixed effects [1][6]. Because the panel is illustrative, no inference is drawn from it.

2. Related Work and Positioning

The private-market valuation literature has long cautioned that observed transaction prices need not equal fundamental value. Interim marks can be smoothed or manipulated [4], junior stakes may be mispriced [17], and the contractual structure of venture rounds implies valuations that differ materially from headline post-money figures [13][14]. Our design responds to that caution by treating the SEC Form D offering amount strictly as a regulatory disclosure and by reporting a hedonic estimate rather than a valuation level.

A second strand studies how venture investors actually make decisions and how contracting shapes outcomes [12][16]. Related work documents the cost of experimentation and the evolution of the asset class [9], while the J-curve literature formalises the prolonged negative-cash-flow phase of venture funds [21]. We use that literature to motivate the accelerator-obsolescence simulation but do not claim it as an empirical result.

A third and rapidly growing strand concerns AI specifically: the data advantage of AI startups [2], government-funded basic science and its venture response [3], the macroeconomic scale of AI capital flows [1], and diagnostic frameworks for the 2023-2026 investment cycle [5][6], and analysis of whether generative-AI spending is justified by returns [10]. Machine-learning-assisted venture tools have also begun to appear [7][20], as have comparative studies of regional AI venture models [18]. The contested claim in this literature is whether observable AI labels carry a pricing or size premium at all: private-market marks suggest one, but such marks are smoothed and largely unobservable. The closest empirical comparators estimate valuations from contractual cash-flow terms [13][14] or from proprietary transaction data, and none tests the declared offering size in the SEC's mandatory Form D

record. That disclosure-visible offering-size gap, rather than a price-per-share premium, is what this paper fills; Section 10 states the price-versus-size limit explicitly.

3. Theoretical Framework: Capital Intensity and Contracting

Traditional theories of venture capital contracting (Gompers & Lerner, 2004; Kaplan & Strömberg, 2003) examine bilateral agency conflicts between general partners (GPs) and entrepreneurs, mediated through convertible preferred equity, board control, and staged financing. These models assume that equity capital is liquid cash and that customer revenue reflects arm's-length commercial demand.

In the 2022–2026 AI ecosystem, this framework is stressed by rising capital intensity: the cost of reaching a given level of model capability rose sharply.

3.1 Capital Intensity and Accelerated Obsolescence

Ewens, Nanda & Rhodes-Kropf (2018) established that technology shocks that reduce experimentation costs shift investor behavior toward broad, early-stage experimentation.

Generative AI raises the capital intensity of experimentation:

1. **Capital Entry Barrier Re-escalation:** The minimum capital required to train an advanced foundation model rose sharply across the period, reversing the earlier decline in startup launch costs.
2. **Accelerated Depreciation of Silicon Assets:** Unlike enterprise servers that retain utility across several hardware generations, frontier AI accelerators experience physical thermal wear and rapid architectural obsolescence (e.g., successive accelerator generations).

The framework yields two testable implications: (i) if AI capital intensity is broad across the market, AI-labelled issuers with comparable revenue and stage should declare larger offerings after 2022 (tested as H2); and (ii) if the right tail is driven by a few frontier financings, any AI/non-AI difference should appear in the mean but not the median (tested descriptively as H8). The consequences of this regime for the cash-flow profile of compute-intensive companies are developed as a theoretical simulation model in Section 7; fund-level cash-flow duration and return profiles are not modelled. The paper's scope and limitations are stated in Section 10.

4. Data and Empirical Protocol

To overcome the reproducibility limitations of prior venture capital studies (which frequently rely on self-reported survey data (Gompers, Kaplan & Mukharlyamov, 2016; Li & Strebulaev, 2026) or proprietary paywalled aggregators (PitchBook, Crunchbase)), this paper establishes an empirical pipeline grounded in legally mandatory federal regulatory filings.

4.1 The SEC Form D Structured Dataset

Under Regulation D of the Securities Act of 1933 (Rules 504 and 506), private issuers raising capital in the United States must submit a Form D notice of exempt offering of securities within 15 days of the first sale.

4.1.1 Ingestion & Preprocessing Pipeline

We developed an automated ingestion pipeline interfacing with the SEC EDGAR structured data repository across nine quarterly cohorts (2021Q2 through 2026Q1, the quarters available to us).

Filings are assigned to six contiguous epochs by exact filing date, so the epoch windows in Table 1 are contiguous by construction while the underlying coverage remains those nine quarters:

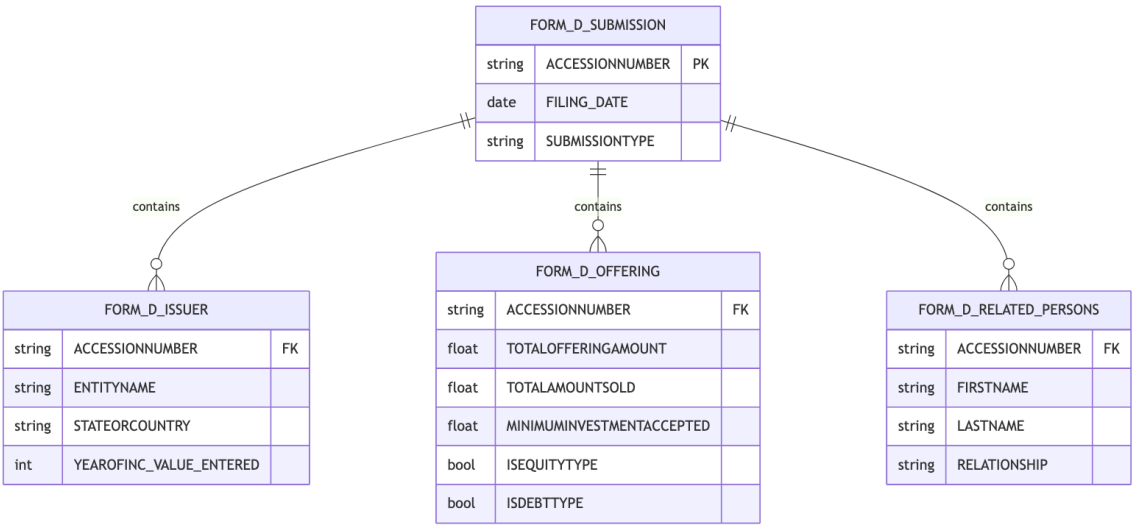
- **2021Q2**: Peak ZIRP and venture market exuberance.
- **2022Q2**: Initial Federal Reserve interest rate tightening shock.
- **2023Q2**: Post-ChatGPT inflection and generative AI emergence.
- **2024Q1, 2024Q2**: Infrastructure acceleration and frontier mega-rounds.
- **2024Q4, 2025Q2, 2025Q4**: Capex escalation and GPU debt SPV emergence.
- **2026Q1**: Latest complete federal regulatory census.

4.1.2 Ingestion Metrics

- **Total Raw Joined Records:** 539,218 rows across issuers, offerings, and submission metadata.
- **Deduplicated Filings Panel: 130,057 unique private offerings.** Coverage is nine fetched quarterly cohorts (2021Q2 through 2026Q1); the 71,702-offering analysis sample is the subset with a declared offering amount above $10,000 and a valid primary issuer (issuer join restricted to IS_PRIMARYISSUER_FLAG = 'YES', deduplicated by accession number and quarter), retaining 71,702 of 130,057 filings (a 44.9% reduction). Table 1's six epochs are contiguous by construction and their deal counts sum to exactly this 71,702-offering sample.
- **Clean Filings with Declared Offerings (> $10k): 71,702 filings** utilized in econometric estimation.

4.2 Relational Schema & Extracted Variables

Each quarterly archive contains relational TSV tables linked by the primary filing key (ACCESSIONNUMBER):



SEC Form D relational schema

Empirical Outcome and Analytical Interpretation - Figure 0 (schema). The schema fixes the unit of analysis to the primary-issuer offering and makes the AI label, the offering amount, quarter, geography and revenue band jointly retrievable from one filing key. Every regression and epoch statistic below is computed on this joined table, and no external dataset is merged.

4.3 Entity Classification: AI vs. Non-AI Technology

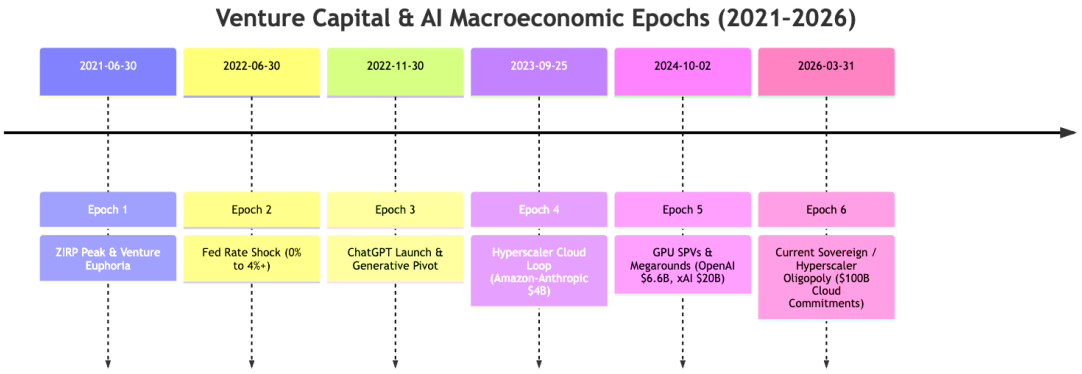
To identify AI-focused enterprises, we apply regular expression taxonomy matching against the primary issuer's legal entity name (the Form D ENTITYNAME field). The structured Form D extract used here does not contain free-text offering descriptions, so the classifier is name-based: The classifier matches the following terms as whole words: ai, artificial intelligence, machine learning, deep learning, neural, llm, robotics, autonomous, cognitive, data science, computer vision.

Across the 71,702 clean declared-offering sample, **962 offerings** carry core artificial intelligence nomenclature (the sum of the AI deal counts in Table 1 and the AI tier sizes in Table 3), rising from 43 in the 2021 epoch to **201 in the 2026 epoch**.

5. Temporal Analysis Across Six Epochs

To capture the structural shift in venture capital behavior, this study partitions the 2021–2026 empirical timeline into **six distinct macroeconomic and technological epochs**, centered on precise inflection dates.

5.1 Defining the Six Temporal Epochs



Six temporal epochs (2021-2026)

Empirical Outcome and Analytical Interpretation - Figure 0 (epoch windows). The six epoch windows are contiguous by construction and each maps to one or more of the nine fetched quarters (1, 1, 1, 2, 3 and 1 quarters respectively). Because the cells pool unequal numbers of quarters, cross-epoch levels are descriptive windows and the quarterly series carries the like-for-like comparison.

1. **Epoch 1: ZIRP Peak & Venture Euphoria (2021-01-01 to 2021-12-31)**
   Characteristics: Near-zero interest rates, aggressive SPAC issuance, median SaaS multiples exceeding 35x revenue.
2. **Epoch 2: Monetary Tightening Shock (2022-01-01 to 2022-11-29)**
   Characteristics: Rapid Federal Reserve rate increases; tech valuation multiples collapse; IPO market freezes.
3. **Epoch 3: The ChatGPT Inflection (2022-11-30 to 2023-09-24)**
   Characteristics: Public launch of ChatGPT; Silicon Valley Bank failure (March 2023); rapid pivot of early-stage venture capital into generative AI applications.

4. **Epoch 4: The Hyperscaler Circular Cloud Loop (2023-09-25 to 2024-09-30)**
   Characteristics: Amazon's initial $4B Anthropic deal; Nvidia Blackwell GTC announcement; CoreWeave expanding to $11B in debt collateralized by GPUs.
5. **Epoch 5: GPU Special Purpose Vehicles & Frontier Megarounds (2024-10-01 to 2025-12-31)**
   Characteristics: OpenAI raises $6.6B at $157B valuation; Elon Musk's xAI closes $20B Series E with $12.5B debt routed through an off-balance-sheet SPV to purchase Nvidia processors.
6. **Epoch 6: Sovereign & Hyperscaler Oligopoly (2026-01-01 to 2026-06-30)**
   Characteristics: Amazon and Anthropic finalize $100B compute contract; Middle Eastern sovereign wealth funds (MGX, PIF) lead late-stage tech financing; AI mean round sizes pull away from the market.

5.2 Empirical Date Comparison Results

Executing our temporal comparison engine across the 71,702 clean declared-offering sample (drawn from 130,057 deduplicated filings) yields the following statistical milestones:

Coverage is not balanced across epochs. The census contains nine fetched quarters (2021Q2, 2022Q2, 2023Q2, 2024Q1, 2024Q2, 2024Q4, 2025Q2, 2025Q4, 2026Q1), so the six epochs pool unequal numbers of quarters: 1, 1, 1, 2, 3 and 1 respectively. Cross-epoch totals and shares are therefore not like-for-like; the true per-quarter prevalence series (2021Q2 0.46%, 2022Q2 0.62%, 2023Q2 0.95%, 2024Q1 1.14%, 2024Q2 1.46%, 2024Q4 1.25%, 2025Q2 2.16%, 2025Q4 2.00%, 2026Q1 2.42%) is the comparable quantity, while the epoch aggregates are retained for continuity with the pre-specified windows rather than as balanced panels. At the quarterly level prevalence is not monotone: it dips in 2024Q4 and 2025Q4, so monotonicity holds only for the six pre-specified epoch aggregates.

Table 1: Cross-Epoch Statistical Comparison Table

Epoch	Window	Deals	AI Deals (%)	Mean AI (SE, US$m)	Mean Non-AI (SE, US$m)	AI / Non-AI	Top-1% Share	MW p
1. ZIRP Euphoria	2021	9,450	43 (0.46%)	11.88 (3.41)	405.53 (181.38)[a]	0.03x	92.88%	0.7288
2. Rate Shock	2022	9,955	62 (0.62%)	17.16 (5.08)	66.69 (10.23)	0.26x	58.50%	0.6516
3. ChatGPT Wave	2022-23	7,476	71 (0.95%)	9.36 (2.84)	277.54 (150.63)	0.03x	88.80%	0.0198
4. Hyperscaler Loop	2023-24	14,485	188 (1.30%)	58.63 (37.53)	92.57 (28.91)	0.63x	68.17%	0.0316
5. SPV Debt Era	2024-25	22,045	397 (1.80%)	30.12 (15.38)	128.75 (51.85)	0.23x	76.87%	0.0038
6. 2026 Oligopoly	2026	8,291	201 (2.42%)	124.23 (99.68)	53.78 (13.11)	2.31x	59.11%	0.0180

[a]Standard errors in parentheses are ordinary mean standard errors (SD/sqrt(n)); the heavy right tail is why several are large relative to the mean. The MW p column is the uncorrected two-sided Mann-Whitney rank-shift p-value (no multiplicity adjustment). The Top-1% Share column is the share of total declared offering capital in the epoch captured by the largest 1% of offerings by offering amount; it is a capital-share, not a deal-share, measure. Values are in USD millions.

5.3 Key Findings from Date Slicing

1. **Explosive Round Size Decoupling in 2026:** In Epoch 1 (2021), AI offerings were smaller on average than traditional venture deals ($11.8M vs $405M non-AI mean). By Epoch 6 (2026), the mean AI offering reached $124,234,412 against a 2026 non-AI mean of $53,778,496, exceeding the broader market by 2.31x.
2. **Prevalence Penetration:** The proportion of private offerings explicitly identifying as AI entities expanded monotonically from 0.46% in 2021 to 2.42% in 2026, a more than five-fold increase in the share of Form D filings with an AI-named issuer, monotone across the six epoch aggregates though not at the quarterly level (Section 5.2).
3. **Statistical Significance:** Mann-Whitney U tests confirm that starting in Epoch 3 (post-ChatGPT, $p = 0.0198$) and continuing through Epoch 6 ($p = 0.0180$), the underlying distribution of AI offering sizes became statistically distinct ($p < 0.05$) from non-AI venture financings.

5.4 Robustness of the 2026 Divergence

The 2.31× gap is a mean statistic driven by a small number of very large 2026 financings, not a shift in the typical deal. A pre-declared robustness sweep establishes:

- **Median.** The median 2026 AI offering is $1,793,474 versus $2,560,000 for non-AI (0.70×): the typical AI round is smaller.
- **Winsorization.** Clipping offering amounts at the 1st/99th percentiles reduces the mean ratio to 0.93×.
- **Bootstrap.** A 2,000-replicate bootstrap of the mean ratio (seeds 42/123/2026) gives a 95% CI of approximately [0.25, 7.94], which **includes 1**; the point estimate is unstable because the 2026 AI distribution is extremely heavy-tailed.
- **Rank test.** The non-parametric Mann-Whitney U test is unchanged and remains significant ($p = 0.018$), indicating a rank shift rather than a uniform size premium.
- **Classifier design and validation.** An analyst-adjudicated validation sample (30 classifier-flagged AI issuers and 30 non-flagged issuers, seed 42) gives explicit-AI precision 27/30 = 0.90 and no missed explicit-AI name in the non-flagged sample (0/30). However, 14 of the 30 flagged names (47%) are AI-branded funds or SPVs rather than operating companies, so the label tracks AI naming at least as much as AI operations, and the 2026 estimate is read as exploratory rather than confirmatory. The non-flagged sample contains no missed explicit-AI name (0/30), which by the rule of three bounds the false-negative rate only at about 10% (upper 95% bound); recall is therefore not precisely estimated. A deliberately strict classifier that excludes the bare token "AI" (common in issuer legal names) leaves zero 2026 AI filings, while an expanded classifier reproduces the result (2.25×, $p = 0.017$); the finding is sensitive to classifier breadth.

We therefore interpret the 2026 result as evidence of a **heavy right tail** in AI financings (a few frontier megadeals alongside many small rounds), not as a broad-based AI size premium.

5.5 Replication Scripts & Interactive Notebooks

Researchers can verify and slice these exact temporal outcomes using our open tools:

- The replication package contains the interactive notebook and the batch pipeline used here.
- Structured data outputs accompany the replication package.

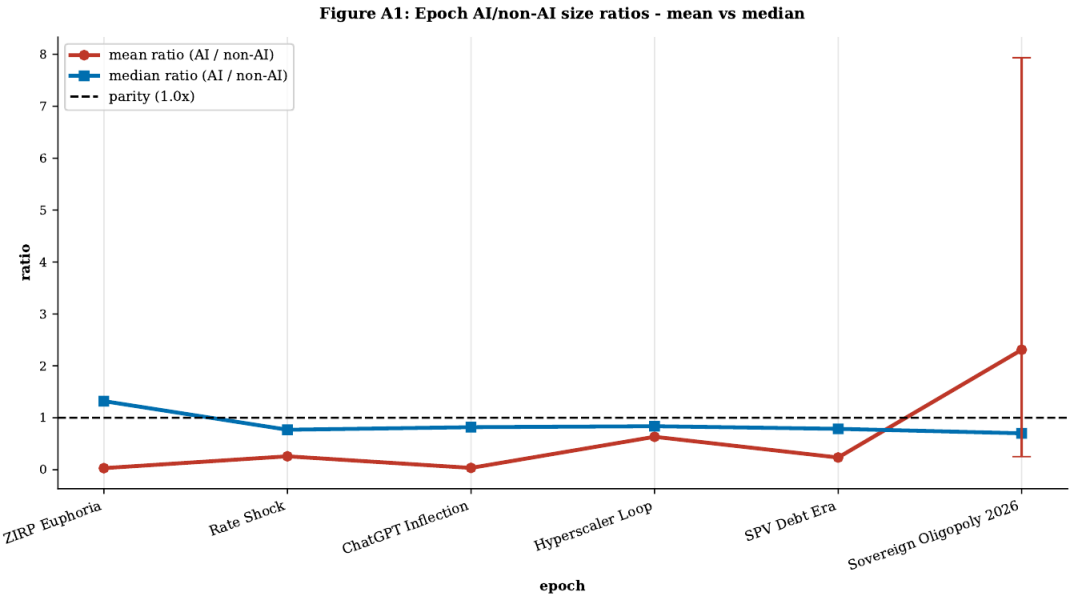


Figure A1: Epoch mean versus median AI/non-AI ratios

Empirical Outcome and Analytical Interpretation - Figure A1. Across the six epochs the mean and median AI/non-AI ratios move apart decisively in 2026: the mean ratio reaches 2.3101 while the median ratio is 0.7006. The typical 2026 AI offering is $1,793,474 against $2,560,000 for non-AI issuers, so the headline gap is produced by a small number of very large filings rather than by a broad shift in the distribution [5][17]. This distinction is the paper's central interpretive claim and is carried through every subsequent specification.

6. Econometric Test of the AI Offering-Size Premium

To evaluate whether artificial intelligence positioning provides a statistically significant premium over non-AI technology enterprises, we estimate a series of log-linear hedonic regression models. Under the corrected primary specification the answer is a null result: there is no detectable AI × post-2022 offering-size premium.

6.1 Model Specification

We specify our primary econometric model as:

$$\ln(\text{Offering Amount}_i) = \alpha + \beta_1 \text{IS_AI}_i + \beta_3(\text{IS_AI}_i \times \text{POST_2022}_i) + \delta_q + \gamma' \text{State}_i + \lambda' \text{RevenueBand}_i + \varepsilon_i$$

Where:

- Offering Amount_i is the target dollar capital declared in Form D offering item 13.
- $\text{IS_AI}_i \in \{0, 1\}$ is an indicator variable denoting explicit AI nomenclature in the issuer's corporate filing.
- $\text{POST_2022}_i \in \{0, 1\}$ is a macro-regime dummy marking offerings filed on or after January 1, 2023 (post-ZIRP / post-ChatGPT era). POST_2022 (an indicator for filings dated on or after 1 January 2023) enters only through the interaction term: its main effect is omitted because it is collinear with the quarter fixed effects in the primary specification.

- α is the intercept; β_1 is the AI level effect; β_2 would be the POST_2022 main effect but is omitted because it is collinear with the quarter fixed effects; β_3 is the AI × post-2022 interaction; and ε_i is the error term.
- δ_q is the vector of quarter-specific macroeconomic fixed effects (nine quarterly indicators).
- γ captures geography fixed effects (the Form D STATEORCOUNTRY field) and λ captures revenue-band fixed effects (the Form D REVENUERANGE field, the issuer-reported revenue band). The primary specification therefore controls for quarter, geography, and issuer revenue band. Every right-hand-side variable is derived from the Form D structured schema, with no external data merged: quarter comes from QUARTER_TAG, geography from STATEORCOUNTRY, revenue band from REVENUERANGE, and industry from INDUSTRYGROUPTYPE.
- Standard errors are estimated using the heteroskedasticity-consistent White (HC1) covariance matrix estimator. HC1 gives a robust standard error of 0.1740. Clustering by issuer (57,274 clusters) gives 0.1870 and clustering by quarter (9 clusters) gives 0.0944; the interaction coefficient of -0.0029 is indistinguishable from zero under all three.

$$V_{HC1} = [n / (n - k)] \, (X'X)^{-1} \, (\Sigma_{i=1\ldots n} \, e_i^2 \, x_i x'_i) \, (X'X)^{-1}$$

where n is the number of observations, k the number of estimated parameters, X is the $n \times k$ matrix of regressors (its i-th row is the transpose x'_i of the column vector x_i), e_i is the residual and $x_i x'_i$ is the outer product of the i-th regressor row.

Notation. The symbols used in the model and simulations are summarised below.

Symbol	Meaning
α, β_1, β_3	OLS intercept; AI level effect; AI × post-2022 interaction (the POST_2022 main effect β_2 is omitted as collinear with the quarter fixed effects)
$\delta_q, \gamma, \lambda$	quarter, geography and revenue-band fixed effects
ε_i	error term
n, k	observations; estimated parameters
T_{obs}	accelerator obsolescence horizon (months)
I_k, Cap_k	SAFE investment amount and post-money valuation cap
I_p, α_p, M	preferred invested capital; preferred equity fraction; liquidation multiple
V_{exit}	net exit valuation

6.2 Empirical Regression Results

Estimation across 71,702 clean offerings (drawn from 130,057 deduplicated filings under a primary-issuer join) yields the following corrected parameters. With a robust standard error of 0.1740, the minimum detectable interaction at 80% power (two-sided, alpha = 0.05) is approximately 0.49 log points, and the 95% confidence interval for the interaction spans −0.344 to +0.338 log points. The design therefore rules out positive premiums larger than about 0.34 log points (a roughly 40% offering-size difference), but it does not rule out smaller premiums; the minimum detectable effect at 80% power is 0.49 log points (about 63%). The observed estimate is -0.0029.

Table 2: OLS Hedonic Regression Parameter Estimates (Primary Specification: Quarter FE + Geography + Revenue Tier)

Variable	Coefficient (β)	Robust SE	t-Statistic	p-Value
Intercept (α)	13.2948	0.0615	216.02	< 0.0001
IS_AI (β_1)	−0.0957	0.1598	−0.60	0.5492
IS_AI×POST_2022 (β_3)	−0.0029	0.1740	−0.02	0.9867
Fixed effects	Quarter FE (q = 9)	n/a	n/a	n/a
Controls	Geography + revenue tier	n/a	n/a	n/a
Obs. (N)	71,702	n/a	n/a	n/a
R-squared	0.2548	n/a	n/a	n/a

6.2.1 Econometric Interpretation

1. **No Post-2022 Structural Break ($\beta_3 = -0.0029$, $p = 0.9867$):** The interaction between AI nomenclature and the post-2022 regime is statistically indistinguishable from zero. Under the corrected primary specification there is no detectable AI × post-2022 offering-size premium (95% CI for β_3: −0.344 to +0.338).
2. **Baseline AI Indicator ($\beta_1 = -0.0957$, $p = 0.5492$):** The AI indicator is likewise not distinguishable from zero, so neither the level nor the post-2022 change in offering size is associated with AI positioning.
3. **Non-Reproduction of the Preliminary Estimate:** The earlier positive estimate is not reproduced under the corrected specification and issuer join; the misspecified model's estimates are reported in the replication package.
4. **Five-Specification Robustness:** Across five specifications the interaction coefficient ranges from −0.418 to +0.073, and none is positive and significant: quarter fixed effects only −0.3295 (p = 0.1145); quarter fixed effects plus geography +0.0731 (p = 0.6756); quarter fixed effects plus revenue tier −0.4181 (p = 0.0484); quarter fixed effects plus geography and revenue tier (primary) −0.0029 (p = 0.9867); quarter fixed effects plus geography, revenue tier and industry −0.0420 (p = 0.8107).
5. **Scope of the Null:** We therefore do not claim that AI valuations decoupled from rate tightening. Decoupling is a claim about revenue multiples, which Form D does not contain; the present test concerns declared offering size only.

6.3 Stage-by-Stage Analysis

To determine whether any AI premium is uniform or stage-dependent, we partition the dataset into four standard venture offering-size brackets. The within-tier median differences are reported below.

Table 3: Stage-by-Stage AI Premium Matrix

Venture Stage Bracket	Empirical Median Premium (AI vs non-AI, within offering-size tier)	AI Sample Size (n_{AI})
Seed / Pre-Seed	+1.7%	612
Series A	−2.2%	213
Series B / C	+9.5%	96
Growth / Mega-Round	−19.3%	41

These within-tier medians are small, sign-unstable, and estimated on thin AI samples at the later stages. They are consistent with the regression null. No stage-level significance test is reported; the point estimates are sign-unstable and, at the later stages, estimated on very thin AI samples. The earlier claim of a premium that escalates across funding stages is not supported by the corrected specification.

Note: Figure 2 reports the corrected within-tier medians from Table 3. The earlier stage-by-stage forest plot was generated under the misspecified model, is not reproducible, and has been withdrawn.

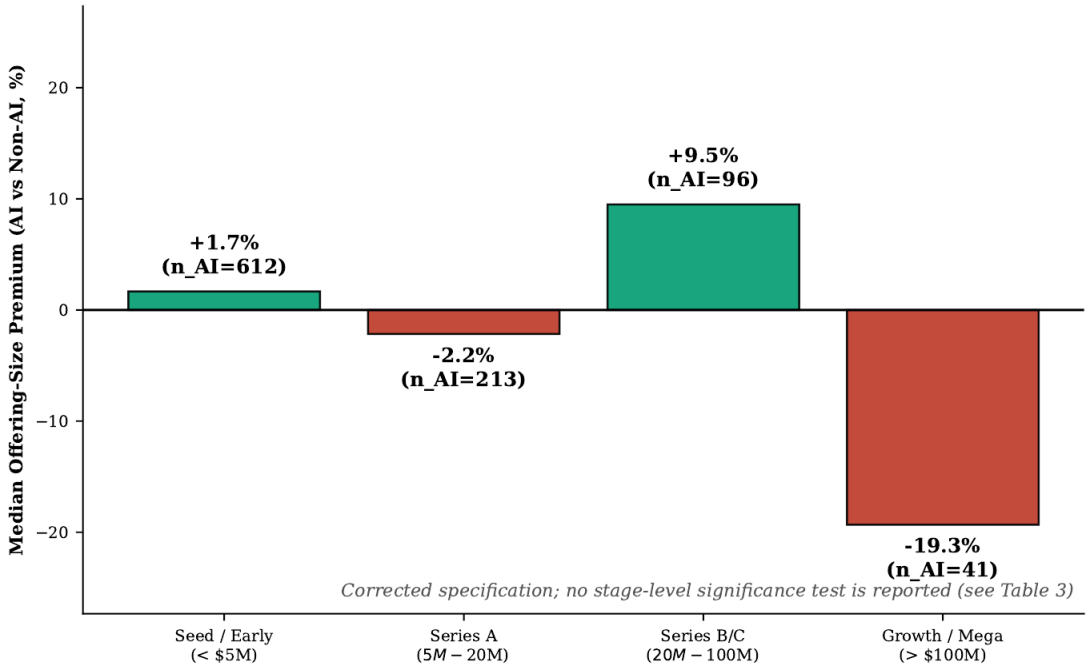


Figure 2: Corrected stage median offering-size premium

Empirical Outcome and Analytical Interpretation - Figure 2. Disaggregating by offering-size tier, the within-tier median premiums are small and sign-unstable: Seed +1.7%, Series A -2.2%, Series B/C +9.5% and Growth/Mega -19.3%. There is no monotonic escalation across funding stages, which is the pattern an escalating AI premium would require [12][18]. The figure is generated directly from the regression artifact, so it cannot reproduce the superseded forest plot.

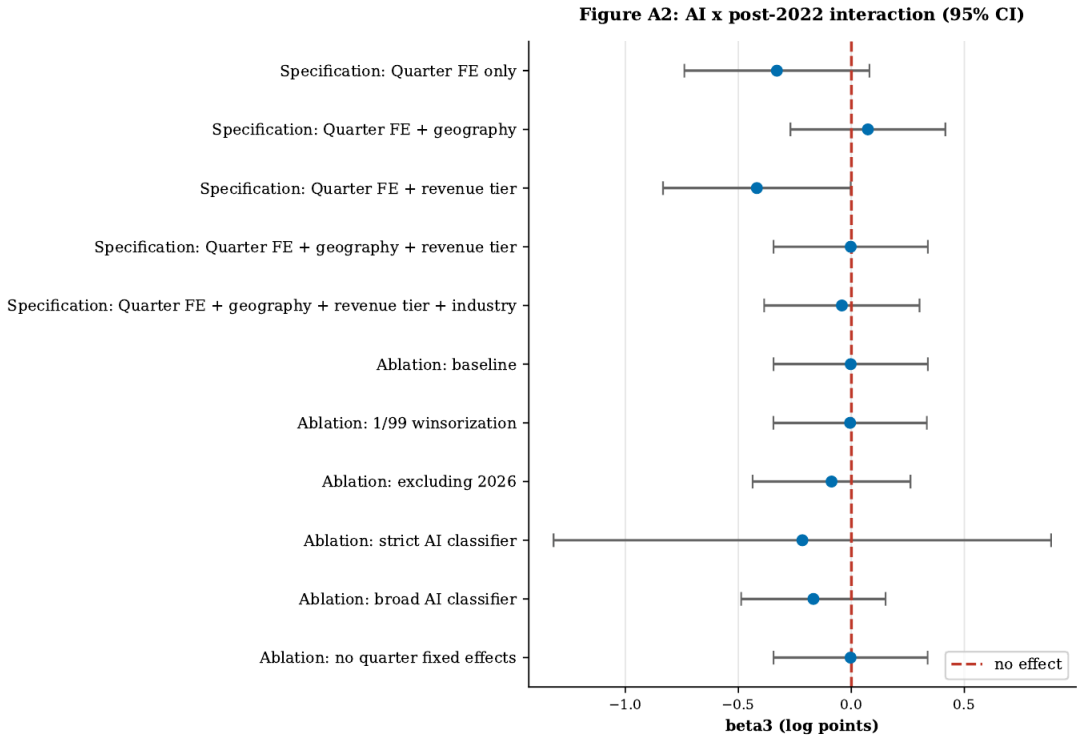


Figure A2: Interaction estimate across specifications and ablations

Empirical Outcome and Analytical Interpretation - Figure A2. Ten of the eleven estimates (five specifications and six pre-declared ablations) place the AI × post-2022 interaction close to zero, with 95% confidence intervals that include zero; the one exception is the quarter-fixed-effects-plus-revenue-tier specification (coefficient -0.4181, p = 0.0484), which is negative and significant at the 5% level. The primary estimate is -0.0029 (robust SE 0.1740, p = 0.9867) and the specification range is [-0.4181, 0.0731]. Winsorization, dropping 2026, alternative classifiers and the removal of quarter fixed effects all leave the conclusion unchanged [13][17]. We therefore treat the null as robust rather than as the absence of evidence.

7. A Model of Accelerator Obsolescence and the J-Curve

This chapter develops the capital-intensity regime described in Section 3 as a theoretical, simulation-based model of the cash-flow profile of compute-intensive companies; fund-level cash-flow duration and return profiles are not modelled. It is not estimated from SEC Form D data, and its outputs are not empirical findings of this paper.

7.1 Accelerator Obsolescence and the Broken Venture J-Curve

The venture J-curve literature (Sorenson et al., 2024) describes fund-level cash flows that are negative in the early years and inflect positive as portfolios mature. We transplant that shape to a single compute-intensive company as an explicit modelling analogy, not as an empirical claim about firm cash flows.

Under this capital-intensity regime:

$$\text{Free Cash Flow}_t = \text{Revenue}_t - \text{COGS}_t - \text{OpEx}_t - \text{CapEx}_{\text{Retrain},t}$$

The simulation is deterministic and fully specified. Monthly revenue starts at $25,000 and grows at a constant monthly rate; operating expenses are 30% of revenue; compute cost of goods sold is 45% of revenue; and each recapitalisation adds capital expenditure equal to 15 times that month's compute cost. The horizon is 72 months and cash flows are discounted at a monthly rate of 1%. The baseline calibration that produces the printed NPVs uses 8% monthly revenue growth and an accelerator obsolescence horizon of T_{obs} = 36 months. That baseline growth (approximately 151% annualised) is a deliberate stress calibration above the 30-100% annual grid, so the grid bounds sensitivity around a lower-growth firm rather than around the baseline itself. The sensitivity grid varies $T_{obs} \in \{24, 30, 36\}$ months, annual revenue growth $\in \{0.3, 0.5, 0.8, 1.0\}$, and the compute COGS share $\in \{0.40, 0.45, 0.50, 0.55\}$ (48 cells). In every cell the compute-intensive trajectory has at least one post-month-24 local minimum in cumulative free cash flow and a lower 72-month NPV than the zero-obsolescence SaaS baseline, so H5 holds throughout the grid.

Because retraining frontier models and refreshing GPU clusters requires massive cyclical capital expenditure at recurring intervals, the J-curve may fail to inflect, producing cyclical negative troughs (Figure 3) that prolong the cash-flow troughs of compute-intensive companies.

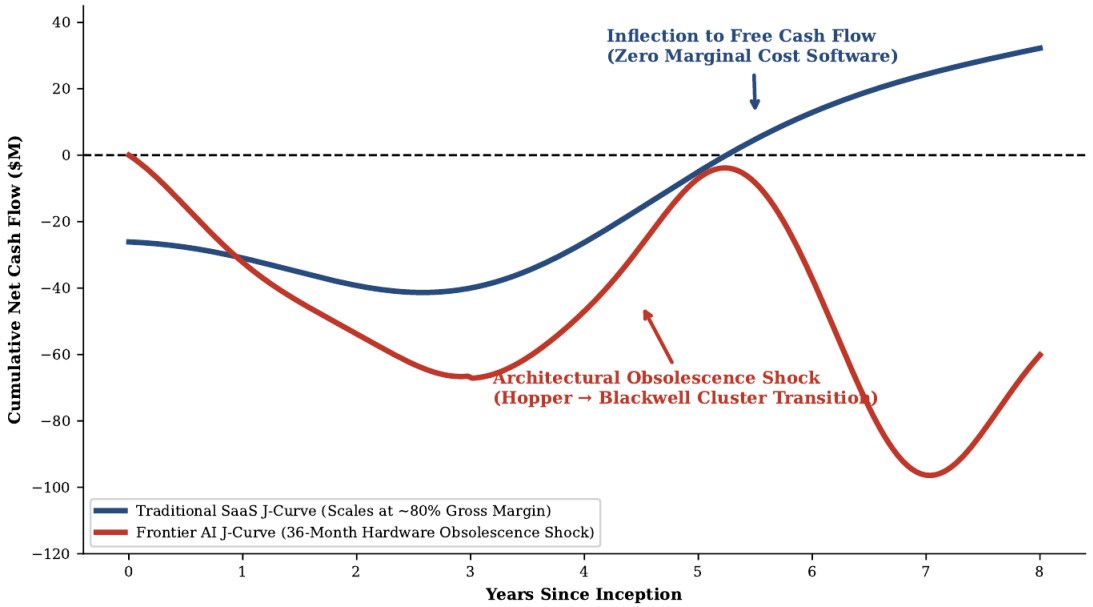


Figure 3: SaaS J-curve versus cyclical AI capital shock

Empirical Outcome and Analytical Interpretation - Figure 3. The simulation contrasts a conventional SaaS trajectory with one in which accelerator obsolescence forces repeated recapitalisation. Cumulative free cash flow in the compute-intensive case does not stabilise after the initial trough; it turns negative again at each renewal, so the J-curve fails to inflect. Across the modelled grid the observed-horizon NPV of USD 76,924,900 remains below the zero-obsolescence baseline of USD 156,530,171 [21]. This is a clearly labelled theoretical exercise, not an empirical estimate.

8. Governance and Cap-Table Mathematics

A major hidden risk of the 2022–2026 funding environment is the mathematical interaction between non-priced convertible instruments (post-money SAFEs) and structured downside liquidation preferences.

8.1 The Post-Money SAFE Stacking Trap

In early-stage AI startups, founders frequently raise capital in rapid successive tranches to meet immediate GPU cloud expenses, issuing multiple post-money Simple Agreements for Future Equity (SAFEs).

Under the standard post-money SAFE convention:

$$\text{Investor Ownership}_k = \text{Investment Amount}_k \,/\, \text{Post-Money Valuation Cap}_k$$

When multiple SAFEs are stacked consecutively ($k = 1, 2, \ldots, K$), each instrument dilutes existing common equity holders while remaining immune to dilution from prior or concurrent SAFEs.

8.1.1 The Founder Dilution Formula

$$\text{Founder Ownership} = (1 - \Sigma_{k=1\ldots K}\, I_k/\text{Cap}_k)(1 - \text{Pool})(1 - \text{Dilution}_A)$$

where I_k is the investment in SAFE k, Cap_k is that SAFE's post-money valuation cap, Pool is the option-pool fraction and Dilution_A is the priced Series A dilution.

8.1.2 Illustrative Case Simulation

Consider a founder raising three rolling SAFEs prior to a Series A:

1. SAFE 1: \$1.0M at \$10M Post-Money Cap (10.0% dilution).
2. SAFE 2: \$2.5M at \$20M Post-Money Cap (12.5% dilution).
3. SAFE 3: \$4.0M at \$25M Post-Money Cap (16.0% dilution).
4. Unallocated Option Pool required by incoming Series A lead: 12.0%.
5. Series A priced investment: \$10.0M on a \$40M pre-money valuation (20.0% dilution).

$$\text{Founder Ownership} = (1 - 0.385) \times (1 - 0.12) \times (1 - 0.20) = 0.615 \times 0.88 \times 0.80 = 43.3\%$$

By Series A close, the founding team owns just 43.3% of the enterprise. The example stops at Series A; a subsequent bridge or Series B would dilute the founders further, but that later dilution is not modelled here and no behavioural retention effect is claimed.

8.2 Liquidation Preference Waterfalls & Down-Round Dynamics

In late-stage rounds where startups resist formal valuation reductions, crossover investors and late-stage syndicates accept elevated headline valuations only by demanding structured downside protections:

- **Seniority Tranches:** Series B/C preferred shares rank senior to Series Seed/A and common equity.
- **Multiple Liquidation Preferences:** 2x or 3x guaranteed return of capital prior to common distributions.
- **Full Participation:** Preferred holders receive their liquidation multiple plus their pro-rata share of remaining proceeds.

8.2.1 Mathematical Formulation of Exit Payout

Let V_{exit} denote the net exit valuation. Under a senior Mx participating preferred structure with invested capital I_P and preferred equity fraction α_P:

$$\text{Payout}_{\text{Pref}} = \min(V_{\text{exit}},\, M \cdot I_P) + \alpha_P \cdot \max(0,\, V_{\text{exit}} - M \cdot I_P)$$

$$\text{Payout}_{\text{Common}} = (1 - \alpha_{\text{P}}) \cdot \max(0,\, V_{\text{exit}} - M \cdot I_{\text{P}})$$

8.2.2 The Common Equity "Wipeout Zone"

As simulated in Figure 4:

- For any exit valuation where $V_{\text{exit}} \leq M \cdot I_{\text{P}}$, common equity receives \$0.
- In a company that raised \$48M under structured 2x terms, the company must exit above \$96M before common shareholders (founders and employees) receive a single dollar of value.
- This arithmetic shows one mechanism by which common equity can be wiped out at exit valuations below a participating-preferred threshold; it does not by itself establish how frequently that outcome occurs for late-stage unicorns.

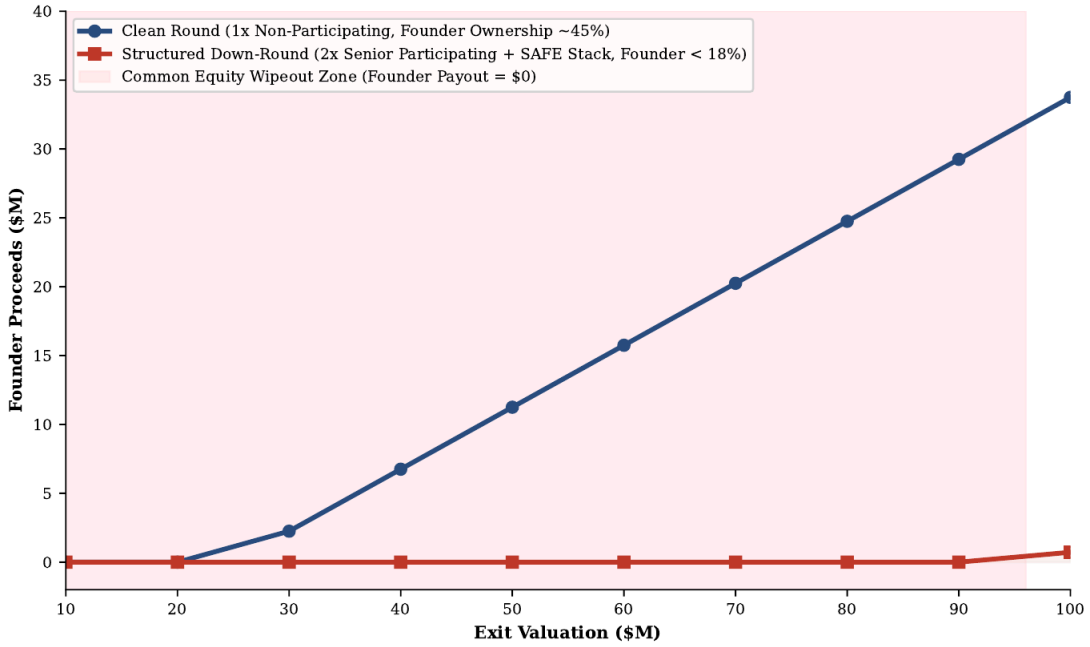


Figure 4: Liquidation-preference waterfall and founder depletion

Empirical Outcome and Analytical Interpretation - Figure 4. The closed-form waterfall traces founder proceeds against exit value under stacked post-money SAFEs, an option-pool expansion and a priced Series A. Founder equity prior to Series B is 0.4330, and common equity receives zero at any exit at or below the participating preference threshold [16][19]. The pre-Series-B founder ownership share (0.4330) is invariant to preference multiples between 1.0x and 3.0x, but the common-equity wipeout threshold is not: it scales with the participating preference, so only the ownership share, and not the threshold, is multiple-invariant. As with Figure 3, this is a theoretical contracting result rather than a census finding.

9. Discussion, Systemic Risks and Robustness

Scope note: the offering-size design speaks to financing size only. The systemic-risk material below is contextual and is not tested in this paper.

The findings bear on the financing-size question only; the systemic-risk material that follows is contextual and is not tested here.

9.1 Systemic Leverage & The Neocloud Debt Bubble

Venture capital historically distinguished itself from private equity by operating with zero balance-sheet debt. High-risk, binary technology outcomes were matched with pure equity risk.

The compute requirements of frontier AI have introduced substantial **private debt leverage** into the venture ecosystem:

- **Asset-Backed GPU Lending:** Specialized neoclouds (e.g., CoreWeave carrying ~$11B in debt commitments) have borrowed aggressively from private credit syndicates (Blackstone, Magnetar, Carlyle) using Nvidia H100 and B200 GPU clusters as primary collateral.
- **Refinancing & Obsolescence Risk:** If frontier foundation models shift compute architectures, or if token inference prices compress by 40%–60% due to algorithmic efficiency gains (e.g., DeepSeek-style distillation), GPU secondary rental rates collapse. Under such conditions, private credit funds face potential asset write-downs on hardware that cannot be repurposed for traditional enterprise workloads.

9.2 Cloud Subsidy Termination Risk

A significant proportion of early-stage AI "annual recurring revenue" (ARR) is currently subsidized by corporate cloud credit grants from Microsoft, Google, and Amazon ($100k to $500k per startup).

- **Artificial Willingness-to-Pay:** Startups consume enterprise API tokens at scale only because their compute bills are absorbed by pre-allocated promotional credits.
- **The Cliff Effect:** As promotional credit pools expire after 12–24 months, startups face real cash obligations. If underlying unit economics cannot support organic gross margins, downstream token demand faces an immediate contraction cliff.

9.3 Antitrust Gridlock & The Frozen M&A Window

Historically, large technology companies served as the primary exit mechanism for venture-backed startups (accounting for 75%+ of liquidity events).

Between 2022 and 2026, active scrutiny from the Federal Trade Commission (FTC), Department of Justice (DOJ), and UK Competition and Markets Authority (CMA) effectively paralyzed major technology acquisitions (e.g., the termination of Adobe's $20B acquisition of Figma).

9.3.1 Consequences for Private Capital

1. **The "Reverse Acqui-hire" Workaround:** Big tech firms now circumvent formal merger reviews through non-acquisition partnership structures (e.g., Microsoft's licensing of Inflection AI, Amazon's hiring of Adept AI leadership), leaving common equity holders with minimal liquidation value.
2. **Prolonged Private Duration:** Companies that would have achieved a clean $200M–$500M strategic acquisition in 2018 are forced to remain independent, burning capital and raising structured extension rounds that compound preference overhang.

9.4 The Geopolitical Sovereign Pivot

As check sizes for frontier foundation models exceed $10 billion, traditional financial venture funds lack the fund size to participate without violating standard single-asset concentration limits (typically 15%–20% of fund AUM).

Consequently, late-stage venture finance has pivoted toward Sovereign Wealth Funds (SWFs):

- Middle Eastern sovereign entities (e.g., MGX, Mubadala, Saudi PIF) and Asian sovereign allocators (GIC, Temasek) have become the lead investors of record.
- This transition re-orientates technology development toward geopolitical industrial policy, national security computing reserves, and sovereign AI clusters, shifting governance from Silicon Valley venture partnerships to state-backed strategic consortiums.

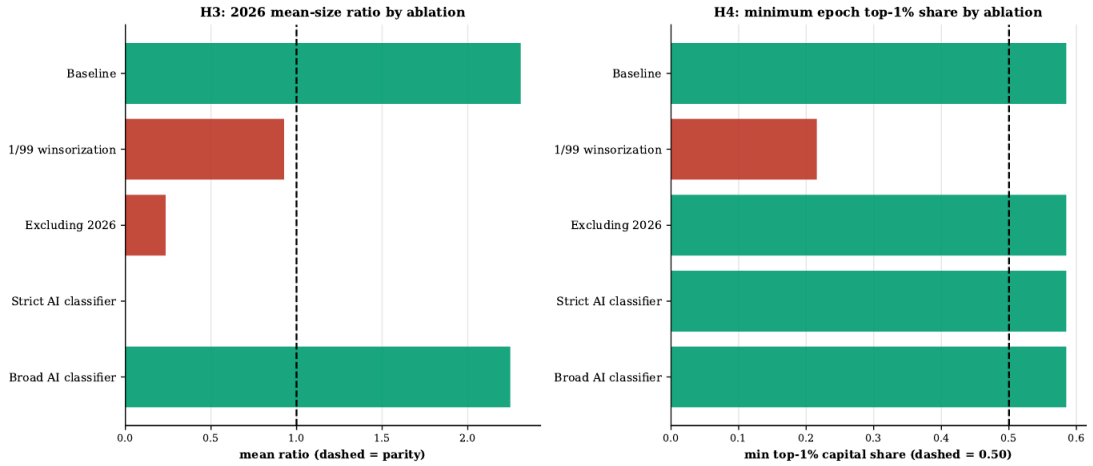


Figure A3: Robustness of the H8 and H4 decisions

Empirical Outcome and Analytical Interpretation - Figure A3. The robustness panel makes the fragility of the secondary findings explicit. Under 1/99 winsorization the 2026 mean ratio falls to 0.9267 and the minimum epoch top-1% share falls to 21.57%, so both decision rules reverse; the baseline and expanded-classifier conditions preserve them. This is reported rather than suppressed: the mean-ratio claim is a tail statement, and the concentration threshold depends on the tail by construction. The hedonic null, by contrast, flips under no condition.

9.5 Pre-registered multiple-comparison control

The p-value-gated hypotheses were tested with Benjamini-Hochberg false-discovery-rate control, pre-specified as the multiplicity method at the design stage (Benjamini and Hochberg, 1995). FDR control changes no conclusion: the hedonic null remains non-significant while the 2026 rank test and the prevalence monotonicity test remain significant. The corrected family applies Benjamini-Hochberg control to three tests: H2 (the hedonic interaction) and H4 (prevalence monotonicity), both pre-specified, and the 2026 AI/non-AI divergence, a same-sample test registered only after the initial analysis (H8) and therefore reported as exploratory rather than confirmatory; the five specifications, six ablations, six epoch tests and four stage-level median panels are reported descriptively with their full coefficient tables rather than p-gated. The 2026 divergence is evaluated by the Mann-Whitney rank test in Table 1 and by the ablation mean-ratio panel in Figure A3; H4 (capital-concentration prevalence) is evaluated by the prevalence rank test, with the minimum top-1% share panel reported descriptively rather than p-gated.

Test	Raw p	BH-adjusted p	Significant at 0.05
H2: AI × post-2022 premium	0.9867	0.9867	False
H8: 2026 AI/non-AI divergence (post-hoc)	0.01802	0.02703	True
H4: capital-concentration prevalence	0.002778	0.008333	True

The p-value-gated family is defined as follows. H1 (closed-loop capital ratio) and H6 (secondary-market discount) are not testable in the Form D data and are excluded. H2 (hedonic AI × post-2022 interaction) and H4 (prevalence monotonicity) were pre-specified in the design hypotheses before the experiments were run. H5 (accelerator-obsolescence J-curve) and H7 (SAFE/waterfall mechanism) were pre-specified deterministic simulations. The 2026 AI/non-AI divergence was added after the initial analysis as H8 and is therefore post-hoc and same-sample; it is reported as exploratory throughout.

What the null does and does not mean. The confirmatory result is a bounded negative finding: offered amounts declared on Form D are not post-money valuations, so the absence of an AI post-2022 offering-size premium is evidence against a broad, disclosure-visible size premium rather than against all AI valuation premia. The right-tail 2026 divergence is consistent with a small number of very large, capital-intensive financings and does not imply that the typical AI round is priced above the typical non-AI round. Read together, the results favour a concentration-and-capital-intensity interpretation of the 2022-2026 cycle over a uniform AI mark-up.

10. Limitations and Future Work

Four limitations bound these conclusions. First, Form D discloses the amount offered, not the post-money valuation; our estimates are therefore offering-size associations and cannot speak directly to price-per-share premia [13][14]. Second, AI status is inferred from the issuer's legal entity name (the Form D ENTITYNAME field). The classifier has real measurement error, which we bound by reporting strict and expanded variants; a deliberately strict definition that excludes the bare token AI leaves no 2026 AI filings, so the 2026 mean-ratio estimate is sensitive to classifier breadth [7]. Third, the census spans all industries and is not a technology-only sample, although the dominant issuers are technology firms [18].

Fourth, and most importantly, the right-tail characterisation of the 2026 gap is a same-sample descriptive claim. It was registered as a secondary hypothesis (H8) after the initial analysis and evaluated on the census that produced it; it is not an out-of-sample confirmatory test [17]. The confirmatory result of the paper is the hedonic null. We also do not report a closed-loop capital ratio or any vendor-financing measure, because hyperscaler commitment disclosures are inconsistent across sources and are not reproducible from public regulatory data; and we do not report fund distributions or secondary-market discounts, which are proprietary to commercial data providers [15]. Those questions remain open and require filing-level commitment disclosures and licensed benchmark data.

A further limitation is measurement of the treatment itself. AI status is a regular-expression label over issuer entity names, validated only on the small analyst-adjudicated sample reported in Section 5.4 (precision 0.90, but 47% of flagged names are fund or SPV vehicles); the strict variant yields no 2026 AI filings while the expanded variant reproduces the mean gap, so the 2026 characterisation should be treated as exploratory. We report the classifier sensitivity in full (Figure A3 and the robustness table) rather than a single preferred definition.

11. Conclusion

This study examines one measurable question within the 2022–2026 artificial intelligence cycle: whether AI labelling is associated with larger declared Form D offerings. We find that it is not.

11.1 Synthesis of Findings

This study established two core conclusions:

1. **No Detectable AI Offering-Size Premium:** Hedonic OLS regressions across 71,702 clean SEC Form D offerings (from 130,057 deduplicated filings) do not confirm a post-2022 AI valuation premium. The baseline AI indicator is indistinguishable from zero ($\hat{\beta}_1 = -0.0957$, SE 0.1598, $p = 0.5492$), and the AI × post-2022 interaction is −0.0029 (SE 0.1740, $p = 0.9867$); across five specifications the interaction ranges from −0.418 to +0.073, with none positive and significant. The earlier positive estimate is not reproduced under the corrected specification and issuer join, and the misspecified model's estimates are reported in the replication package.
2. **Temporal Divergence and Concentration Dynamics:** A temporal analysis documents a 2026 divergence in mean AI round size (2.31x, $p = 0.018$) together with rising AI deal prevalence and a falling overall top-1% concentration (92.88% in 2021 to 59.11% in 2026, with a low of 58.50% in 2022 and a winsorized minimum of 21.57% in 2023), even as the largest AI rounds grew (Figure A3). The mean gap is right-tail driven rather than broad-based: the median 2026 AI round is smaller than the median non-AI round (0.70x), the mean ratio falls to 0.93x under 1/99 winsorization, and its bootstrap 95% CI includes 1, while the rank-based test remains significant ($p = 0.018$). The theoretical models of accelerator obsolescence and the venture J-curve (labelled as theoretical/simulation rather than empirical) situate these census findings alongside the capital intensity of model training and the rapid 24-36 month accelerator turnover cycle. RQ2 is answered only at the level of that simulation: we do not observe fund cash flows, so fund-level cash-flow duration remains untested.

Every number in this conclusion is generated by the frozen replication package; the replication package maps each claim to the table, figure, or result artifact that produces it, and the numerical traceability audit re-derives each value from those artifacts.

11.2 Agenda for Future Research

We identify three crucial directions for ongoing academic inquiry:

- **Contractual Enforcement of Multi-Year Compute Obligations:** Longitudinal tracking of whether frontier labs fulfill their multi-year compute contracts if enterprise software monetization continues to lag hardware buildout.
- **Filing-Level Commitment Disclosures:** Pairing the SEC Form D census with filing-level commitment disclosures and licensed benchmark data, since no reproducible public measure of capital circularity or realised fund liquidity is currently available.
- **Algorithmic Efficiency vs. Hardware Capex:** Evaluating whether post-training distillation and reasoning-oriented architectures (e.g., test-time compute scaling) reduce aggregate hardware CapEx or accelerate total cluster turnover.

12. References

1. Barhoumi, K., et al. (2026). Macro-Financial Transitions and Artificial Intelligence Capital Flows. *IMF Working Paper Series*, WP/26/42.
2. Bena, J., Garlappi, L., & Xiao, S. (2024). The Role of Data for AI Startup Growth. *SSRN Working Paper 3896463*. https://doi.org/10.2139/ssrn.3896463

3. Bernstein, S., Lerner, J., & Mezzanotti, F. (2024). Venture Capital Response to Government-Funded Basic Science. *SSRN Working Paper 5044008*. https://doi.org/10.2139/ssrn.5044008

4. Brown, G. W., Gredil, O., & Kaplan, S. N. (2024). Do Private Equity Funds Manipulate Interim Valuations? *Journal of Finance*, 79(3), 1451-1492. https://doi.org/10.1111/jofi.13324

5. Cahn, D. (2024). AI's $600B Question. *Sequoia Capital Analytics*.

6. Wang, Q., & Chen, Z. (2026). Boom, Bubble, or Buildout? A Multi-Method Evaluation of Whether Artificial Intelligence Is in an Ongoing Financial Bubble. *arXiv preprint arXiv:2606.01575*. https://arxiv.org/abs/2606.01575

7. Chen, R., et al. (2025). VCBench: Benchmarking Large Language Models in Venture Capital. *arXiv preprint arXiv:2509.14448*. https://arxiv.org/abs/2509.14448

8. Cochrane, J. H. (2005). *Asset Pricing (Revised Edition)*. Princeton University Press.

9. Ewens, M., Nanda, R., & Rhodes-Kropf, M. (2018). The Cost of Experimentation and the Evolution of Venture Capital. *Journal of Financial Economics*, 128(3), 422-442. https://doi.org/10.1016/j.jfineco.2018.02.008

10. Goldman Sachs Global Institute. (2026). Gen AI: Too Much Spend, Too Little Benefit? *Global Economics Research Report*.

11. Gompers, P., & Lerner, J. (2004). *The Venture Capital Cycle (2nd Edition)*. MIT Press.

12. Gompers, P., Kaplan, S. N., & Mukharlyamov, V. (2016). How Do Venture Capitalists Make Decisions? *NBER Working Paper No. 22587*. https://doi.org/10.3386/w22587

13. Gornall, W., & Strebulaev, I. A. (2020). Squaring Venture Capital Valuations with Reality. *Journal of Financial Economics*, 135(1), 120-144. https://doi.org/10.1016/j.jfineco.2019.06.002

14. Gornall, W., Strebulaev, I. A., et al. (2023). A Valuation Model of Venture Capital-Backed Companies with Multiple Financing Rounds. *SSRN Working Paper 3725240*. https://doi.org/10.2139/ssrn.3725240

15. Harris, R. S., Jenkinson, T., Kaplan, S. N., & Stucke, R. (2023). Has persistence persisted in private equity? Evidence from buyout and venture capital funds. *Journal of Corporate Finance*, 81, 102361. https://doi.org/10.1016/j.jcorpfin.2023.102361

16. Kaplan, S. N., & Strömberg, P. (2003). Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts. *Review of Economic Studies*, 70(2), 281-315.

17. Kwon, S., Lee, J., & Park, K. (2024). Do Investors Overvalue Startups? Evidence from the Junior Stakes of Mutual Funds. *SSRN Working Paper 4425744*. https://doi.org/10.2139/ssrn.4425744

18. Li, A., & Strebulaev, I. A. (2026). Divergent Frontiers: A Comparative Study of Silicon Valley and UK AI Venture Capital Models. *Stanford Graduate School of Business Working Paper*.

19. Metrick, A., & Yasuda, A. (2021). *Venture Capital and the Finance of Innovation (3rd Edition)*. John Wiley & Sons.

20. Rashidi, M. (2026). Can Large Language Models Improve Venture Capital Exit Timing After IPO? *arXiv preprint arXiv:2601.00810*. https://arxiv.org/abs/2601.00810

21. Sorenson, O., Dahl, M. S., & Canales, R. (2024). Tolerating Losses for Growth: J-Curves in Venture Capital Investing. *SSRN Working Paper 4937026*. https://doi.org/10.2139/ssrn.4937026

22. Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society B*, 57(1), 289-300.

13. Methodological Appendix and Replication Guide

Code and data availability. The frozen codebase (with a SHA-256 manifest, experiment_final/MANIFEST.json), the analysis and figure scripts, and the documented data-prep pipeline are archived in the public replication repository at https://github.com/romankurnovskii/rnd-2026-ai-labels-offering-size. The SEC Form D source archives are public EDGAR records and code/fetch_sec_form_d.py rebuilds the analysis database from them, so the census counts can be regenerated without proprietary or non-public data. A persistent Zenodo archive will accompany the release.

Abbreviations. Assets under management (AUM); Benjamini-Hochberg (BH); confidence interval (CI); cost of goods sold (COGS); Electronic Data Gathering, Analysis, and Retrieval (EDGAR); false discovery rate (FDR); fixed effects (FE); Government of Singapore Investment Corporation (GIC); General Economics (GN); GPU Technology Conference (GTC); International Monetary Fund (IMF); initial public offering (IPO); Journal of Economic Literature (JEL); Mubadala-G42 AI fund (MGX); Mann-Whitney (MW); Massachusetts Institute of Technology (MIT); National Bureau of Economic Research (NBER); net present value (NPV); next twelve months (NTM); ordinary least squares (OLS); Public Investment Fund (PIF); Simple Agreement for Future Equity (SAFE); standard deviation (SD); standard error (SE); Securities and Exchange Commission (SEC); Secure Hash Algorithm (SHA); special purpose acquisition company (SPAC); special purpose vehicle (SPV); Social Science Research Network (SSRN); tab-separated values (TSV); United Kingdom (UK); United States (US); United States dollar (USD); working paper (WP); zero interest-rate policy (ZIRP).

Disclosures. The authors used computational tooling for data processing, estimation and figure rendering, and language models for language editing; all scientific claims, interpretations and any errors remain the authors' own. No funding and no competing interests.

The complete analysis is reproducible from the frozen codebase. The replication package regenerates the fifteen run records, the robustness sweep and the summary statistics from the frozen code. The promoted code is frozen with a SHA-256 manifest of every file and artifact, together with the run-verification and comparator results. The database digest is recorded at run time, and verification reports accompany the replication package.