

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY


Beyond the Hype: Venture Capital Funding Outcomes of AI Startups

Citation

Jamal, Hawraz. 2026. Beyond the Hype: Venture Capital Funding Outcomes of AI Startups. Bachelors Thesis, Harvard University Engineering and Applied Sciences.

Link

https://dash.harvard.edu/handle/1/42742225

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility

https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

Beyond the Hype:
Venture Capital Funding Outcomes of AI Startups

A thesis by
Hawraz Jamal

Presented to the Department of Applied Mathematics
in partial fulfillment of the requirements
for a Bachelor of Arts degree
March 27, 2026

Abstract

Artificial intelligence startups have attracted extraordinary levels of venture capital in recent years, raising questions about whether these firms are genuinely more attractive investments or simply beneficiaries of heightened investor enthusiasm. This paper examines whether startups classified as AI firms raise more cumulative venture capital than comparable non-AI startups, using a large dataset of 494,333 companies sourced from PitchBook and narrowed to a final sample of 7,918 U.S. early-stage venture-backed firms founded between 2010 and 2022. AI classification is based on keyword matching across company descriptions and tags, and the sample is deliberately restricted to a period predating the generative AI boom to minimize the influence of recent hype on funding outcomes. Coarsened Exact Matching is applied on founding year, state, and initial financing size to ensure that AI and non-AI startups are compared within structurally similar groups. Three nested OLS specifications with heteroskedasticity-robust standard errors are then estimated on the matched sample. Our results consistently show that AI-classified startups raise between 33 and 42 percent more cumulative venture capital than matched non-AI counterparts. While the analysis does not establish causality, the findings suggest that the AI funding premium reflects structural differences in how capital has been allocated to AI-focused ventures over time, rather than purely speculative or hype-driven dynamics.

Keywords: Artificial intelligence, venture capital, startup funding, coarsened exact matching, regression analysis

1. Introduction

The recent surge in artificial intelligence–focused startups has become one of the defining features of today's venture capital landscape. News headlines frequently highlight newly founded AI companies raising exceptionally large funding rounds at unprecedented valuations. In this environment, artificial intelligence is increasingly framed not just as a technological tool, but as its own business category.

This shift is clearly reflected in capital flows. In 2024 alone, approximately $100 billion of venture capital was invested in AI startups, the highest annual total on record (Crunchbase). More broadly, global venture funding reached roughly $184 billion in 2024, representing a 21 percent increase relative to the previous year, with AI accounting for a disproportionately large share of this growth (Crunchbase). These trends show little evidence of slowing, indicating sustained investor enthusiasm for AI-driven ventures.

Such extraordinary investment levels raise important questions. Are these large capital allocations being deployed efficiently, ultimately leading AI startups to achieve greater long-term success? Or do they reflect a speculative dynamic in which investors systematically overvalue AI-labeled firms? We believe that venture capital funding is not simply a financial outcome. It also reflects beliefs about future technological impact and economic value. Understanding whether these beliefs are justified is therefore of both academic and practical importance.

To address this question, this paper focuses on U.S. venture-backed startups, examining whether firms classified as AI companies raise more cumulative funding than comparable non-AI

startups. The analysis draws on a large dataset of approximately 494,333 companies spanning 2000 to 2024, sourced from PitchBook. To ensure a well-populated and temporally consistent sample, the analysis is restricted to startups founded between 2010 and 2022. The early years before 2010 are excluded because AI-classified startups were relatively uncommon, limiting a meaningful comparison. The most recent years are also excluded for two reasons: funding outcomes for recently founded startups remain incomplete, as many are still in early stages of raising capital, and the period from roughly 2022 onward coincides with the rapid mainstream adoption of large language models and generative AI, which introduced a wave of investor enthusiasm that could distort funding comparisons. By focusing on the 2010 to 2022 window, our analysis aims to assess whether an AI funding premium reflects structural differences in capital allocation rather than the outsized hype that has come to define the most recent era.

Moreover, this paper focuses on U.S. venture-backed startups, narrowed the dataset to 98,147 U.S. venture-backed firms after filtering by country and funding source. The sample is further restricted to startups founded between 2010 and 2022, yielding a pre-match sample of 44,349 companies. After applying Coarsened Exact Matching on founding year, state, and initial financing size, the final analytical sample consists of 7,918 companies (3,959 AI and 3,959 non-AI) on which three nested OLS specifications are estimated with heteroskedasticity-robust standard errors.

The main finding is consistent across all three models: AI-classified startups raise between 33 and 42 percent more cumulative venture capital than matched non-AI counterparts, with the AI coefficient stable and statistically significant at the 1 percent level throughout. This result holds

under alternative matching specifications and across pre- and post-2017 subperiods, indicating a structural advantage for AI-focused ventures.

The remainder of the paper is as follows. Section 2 reviews the relevant literature. Section 3 describes the data and the AI classification method. Section 4 outlines the matching strategy and regression framework. Section 5 presents the main results and robustness checks. Section 6 discusses the findings in relation to prior work and their broader implications. Section 7 concludes and notes limitations.

2. Literature Review

2.1 Prior Evidence on AI Funding Outcomes

Existing research offers mixed evidence on AI investment outcomes, though they do provide a consistent picture regarding uncertainty, hype factor, and real economic impact.

Montanaro et al. (2024) provide the most direct examination of VC funding in AI ventures, analyzing over 5,000 investments across AI and comparable non-AI firms. They find that AI ventures receive less funding per investment round than non-AI counterparts, citing this gap primarily to higher uncertainty and information asymmetry that makes it harder for investors to evaluate AI-specific claims. However, this penalty does not always hold. The gap narrows substantially as firms mature, and becomes statistically insignificant for later-stage ventures, suggesting that investors remain engaged with AI companies over time once uncertainty is partially resolved. Their study also shows that investor experience and sector-specific knowledge affect the relationship, with more experienced VCs better equipped to assess AI ventures and thus more willing to commit larger amounts. A key limitation, however, is that Montanaro et al.

focus on per-round deal size rather than cumulative capital raised over a firm's lifetime. A startup that raises smaller individual rounds but continues to attract investment across many rounds could still accumulate substantially more total funding than a peer that closes fewer, larger rounds. This distinction is critical to the question this paper addresses.

Rady et al. (2025) approach the problem from a different angle, focusing on how the rhetoric AI startups use to attract investors shapes funding outcomes. Analyzing 302 AI startups across 880 financing rounds, they document a non-linear relationship between startup hype and valuation. Moderate enthusiasm tends to raise investor interest and improve valuations, but excessive hype triggers skepticism, as investors begin to view that rhetoric as implausible. This finding carries an important implication for studying AI funding. If the framing surrounding AI startups systematically distort how investors allocate capital at individual rounds, then per-round funding outcomes may be a noisy signal for underlying firm quality. Investors may overreact to narrative signals in ways that inflate or deflate deal sizes regardless of a startup's actual prospects. By focusing on total funding raised across a firm's full history, this paper removes some of this noise, capturing sustained investor conviction rather than any single round's valuation.

Babina et al. (2024) offer a different but complementary perspective by examining AI adoption within established firms rather than AI-labeled startups. Using a measure of firm-level AI investment derived from hiring data across 535 million individual job histories, they find that firms investing more heavily in AI experience higher growth through increased product innovation, as reflected in trademarks, product patents, and portfolio updates. These results, derived from data spanning 2010 to 2018, suggest that AI investment carries real productive value rather than functioning purely as a speculative label. However, one limitation to the

analysis is that Babina et al. focus on large, established firms rather than early-stage startups, where information asymmetry is more severe and the link between AI investment and observable outcomes is less direct. Still, their findings provide important context. If AI generates genuine economic value at the firm level, a funding premium for AI startups, if it exists, need not reflect pure speculation.

Taken together, these studies suggest that the relationship between AI classification and capital allocation is shaped by a combination of real productivity differences, investor uncertainty, and the rhetoric startups use to attract capital. What remains less clear is whether, after controlling for firm and environmental characteristics, AI startups accumulate more total venture capital over their lifetimes than comparable non-AI firms and consequently more successful. That is the gap this paper aims to fill.

Unlike prior work, which either examines per-round deal sizes, focuses on established firms, or deals with a small dataset, this paper asks whether the AI label is associated with greater cumulative capital raised across a startup's full funding history. We believe this question speaks more directly to long-run investor conviction and startup success.

2.2 Determinants of VC Funding

Geographic location plays a significant role in startup funding outcomes. Chen, Gompers, Kovner, and Lerner (2010) document that more than half of all U.S. venture capital offices are concentrated in just three metropolitan areas — San Francisco/San Jose, Boston, and New York. Nearly 49% of all VC-backed companies are in those same cities, and the reason is structural. Since VCs actively monitor portfolio companies, their ability to add value depends on proximity.

Monitoring costs rise with distance, so VCs prefer to invest locally, and when they do invest remotely they apply a stricter return threshold. This creates a self-reinforcing cycle. Capital, talent, and successful exits keep clustering in the same places, and Chen et al. find that a one standard deviation increase in local VC office presence is associated with a nearly 50% increase in local investment activity.

For this paper, that dynamic matters because a startup based in California or Massachusetts has structural access advantages that would independently drive higher cumulative funding, regardless of AI status. Failing to account for location could therefore cause a geographic funding premium to be mistakenly attributed to AI classification. To address this, one of the regression models includes state bucket dummies for California, New York, Massachusetts, Texas, and Washington, isolating the AI funding premium from these well-documented geographic advantages.

Moreover, early-stage startups can enter the venture ecosystem through several channels, which is also documented in this dataset. Examples include angel rounds, seed rounds, accelerator programs, and early VC rounds — each with meaningfully different starting capital. Angel investments typically range from $25,000 to $100,000, seed rounds from $500,000 to $2 million, while accelerators provide modest initial checks in exchange for equity and network access. Early VC rounds tend to be the largest, coming with increased due diligence and a credentialed investor on the cap table from the start.

We believe that the type and size of a startup's first financing round can also affect total funding raised, as a larger or more prestigious entry round signals quality to future investors and shapes the trajectory of subsequent fundraising. Since entry channel differences could influence the

estimated AI premium, the full regression model includes first financing deal type dummies to ensure the AI coefficient is estimated within comparable entry channels.

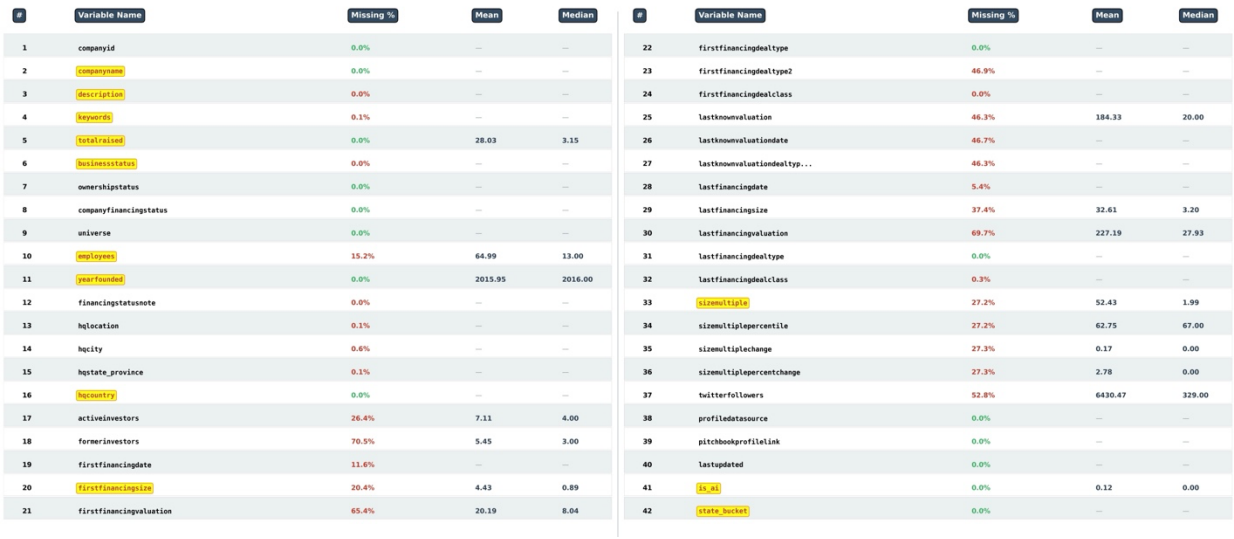
3. Data

3.1 Data Source and Variables

The primary data source for this study is Pitchbook, a popular financial data platform that tracks private and public companies, investors, and deals across the global venture and private equity landscape. The raw dataset contained information on 494,333 companies worldwide founded from 2000 to 2024, including variables such as total funding raised, first and last financing size, deal types, investor details, headquarters location, founding year, and company descriptions, among others.

The sample was first restricted to U.S.-based companies, reducing it to 201,976 firms. Since this paper focuses specifically on venture-backed startups, companies backed by other capital sources were excluded (including private equity-backed firms, publicly listed companies, angel-only backed firms, and corporate-backed entities) leaving 98,147 U.S.-based venture-backed companies as the base sample from which the final analytical dataset was constructed.

Figure 1 - Variable Descriptions Table



#	Variable Name	Missing %	Mean	Median
1	companyid	0.0%	–	–
2	companyname	0.0%	–	–
3	description	0.0%	–	–
4	keywords	0.1%	–	–
5	totalraised	0.0%	28.03	3.15
6	businessstatus	0.0%	–	–
7	ownershipstatus	0.0%	–	–
8	companyfinancingstatus	0.0%	–	–
9	universe	0.0%	–	–
10	employees	15.2%	64.99	13.00
11	yearfounded	0.0%	2015.95	2016.00
12	financingstatusnote	0.0%	–	–
13	hqlocation	0.1%	–	–
14	hqcity	0.6%	–	–
15	hqstate_province	0.1%	–	–
16	hqcountry	0.0%	–	–
17	activeinvestors	26.4%	7.11	4.00
18	formerinvestors	70.5%	5.45	3.00
19	firstfinancingdate	11.6%	–	–
20	firstfinancingsize	20.4%	4.43	0.89
21	firstfinancingvaluation	65.4%	20.19	8.04

#	Variable Name	Missing %	Mean	Median
22	firstfinancingdealtype	0.0%	–	–
23	firstfinancingdealtype2	46.9%	–	–
24	firstfinancingdealclass	0.0%	–	–
25	lastknownvaluation	46.3%	184.33	20.00
26	lastknownvaluationdate	46.7%	–	–
27	lastknownvaluationdealtyp...	46.3%	–	–
28	lastfinancingdate	5.4%	–	–
29	lastfinancingsize	37.4%	32.61	3.20
30	lastfinancingvaluation	69.7%	227.19	27.93
31	lastfinancingdealtype	0.0%	–	–
32	lastfinancingdealclass	0.3%	–	–
33	sizemultiple	27.2%	52.43	1.99
34	sizemultiplepercentile	27.2%	62.75	67.00
35	sizemultiplechange	27.3%	0.17	0.00
36	sizemultiplepercentchange	27.3%	2.78	0.00
37	twitterfollowers	52.8%	6430.47	329.00
38	profiledatasource	0.0%	–	–
39	pitchbookprofilelink	0.0%	–	–
40	lastupdated	0.0%	–	–
41	is_ai	0.0%	0.12	0.00
42	state_bucket	0.0%	–	–

The key variables used in this study are highlighted in the table above. The outcome variable is **totalraised**, measuring total venture capital raised across all financing rounds in millions of USD. The primary independent variable is **is_ai**, a binary indicator equal to 1 if the company is classified as an AI firm and 0 otherwise. Classification was based on each company's **description** and **keywords** fields, which provide a short text summary of the firm's business focus and associated technology tags. **yearfounded** is self explanatory, and **firstfinancingsize** captures the dollar size of the company's first financing round, and **state_bucket** groups companies by state into California, New York, Massachusetts, Texas, Washington, or Other. **businessstatus** indicates the company's current operational status (active, acquired, or defunct) and **employees** records headcount as a measure of firm size. **hqcountry** identifies the company's headquarters country and was used to restrict the sample to U.S.-based firms.

3.2 AI Classification and Filtering

To classify whether a company in the dataset was AI-focused, we constructed a binary indicator variable using each firm's keywords and description fields. We compiled a list of AI-related terms, including "artificial intelligence," "machine learning," "deep learning," "natural language processing," "computer vision," and "large language model," among others, and flagged a company as AI if any of these terms appeared in either field. A company was assigned is_ai = 1 if at least one AI-related term was matched in its keywords or description, and is_ai = 0 otherwise.

For example, a company with the description "We build machine learning infrastructure for enterprise data pipelines" and keywords tagged as "machine learning, AI, data infrastructure" would be classified as AI based on matches in both fields. A company with a description focused e-commerce logistics and no AI-related keywords would be assigned a zero.

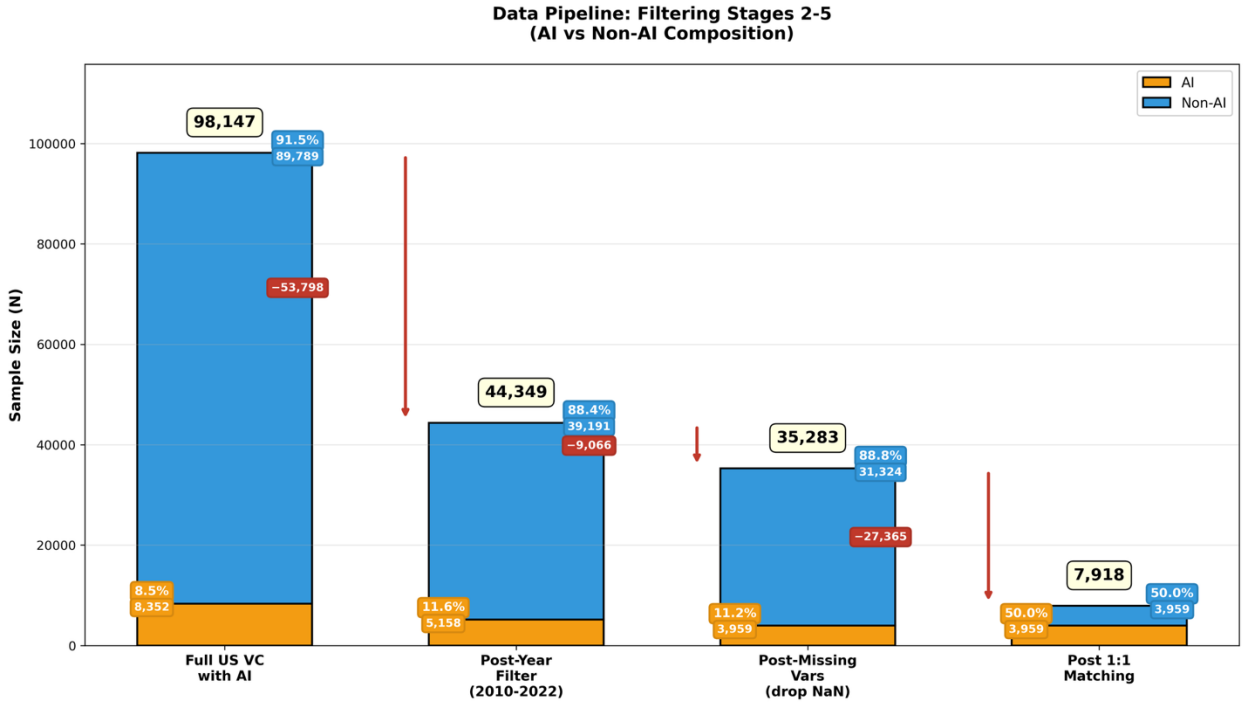
This approach is intentionally straightforward. Instead of relying on third-party classifications or manual labeling, this straightforward approach uses the company's own language to determine AI status, which is consistent with how investors and founders themselves frame the business. It does however rely on the quality and completeness of Pitchbook’s description and keyword data. Companies with sparse or generic descriptions may be misclassified in either direction.

Furthermore, the sample was restricted to companies founded between 2010 and 2022. As discussed before, this window was chosen for two reasons. The early cutoff in 2010 ensures a sufficient number of AI-classified startups for meaningful comparison. AI companies were relatively uncommon before this period, making pre-2010 cohorts too sparse for reliable inference. Also, the late cutoff at 2022 is for two reasons: companies founded after 2022 have had limited time to accumulate funding, making their total capital figures incomplete and

potentially misleading, and the period from 2023 onward coincides with the rapid mainstream adoption of large language models and generative AI, which introduced a wave of investor enthusiasm that could distort the estimated funding premium. Restricting to 2010–2022 therefore aims to capture a window of substantial venture activity while keeping the analysis removed from the most recent hype cycle. This reduced the sample to 44,349 companies. Companies with missing first financing size were then also dropped, as this variable is critical for the matching procedure described in the next section, bringing the final pre-match sample to 35,283 firms.

One concern some may have with all the filtering is that dropped companies may disproportionately be AI or non-AI, introducing selection bias into the estimated premium. To assess this, we tracked the AI share of the sample at every filtering stage. In the full U.S. VC sample, AI companies account for 8.5% of firms (8,352 out of 98,147). After applying the 2010–2022 year filter, the AI share rises modestly to 11.6% (5,158 out of 44,349), and remains essentially stable at 11.2% (3,959 out of 35,283) after dropping observations with missing first financing size. The consistency of the AI share across filtering stages suggests that no stage disproportionately removed AI or non-AI companies, and that the composition of the sample remained broadly stable throughout the data construction process.

Figure 2- Data Pipeline: AI Share Across Filtering Stages



3.3 EDA

To ensure an accurate comparison between AI and non-AI startups, we matched firms on three key characteristics: founding year, state bucket, and log of first financing size.

Founding year ensures that AI and non-AI companies are compared within the same economic and market environment. A startup founded in 2012 operated under very different conditions than one founded in 2019. Matching on this dimension removes cohort-level differences that could otherwise confound the estimated AI premium.

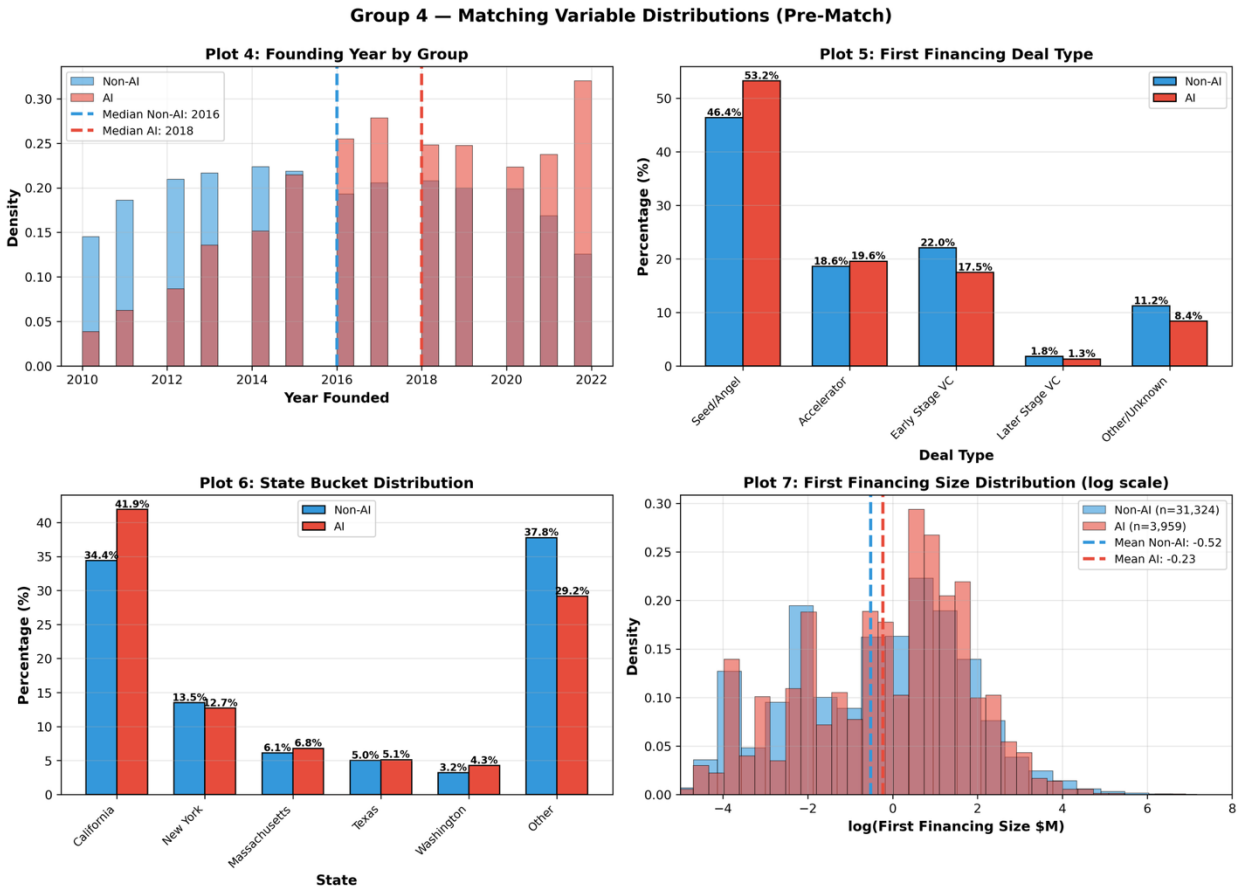
State bucket controls for the geographic concentration of venture capital. As Chen et al. (2010) show, startups located in major VC hubs like California or Massachusetts have structural access

advantages. Matching on state ensures that AI and non-AI companies face broadly similar funding environments.

Finally, first financing size serves as a proxy for initial firm quality and investor conviction at entry. A larger first round signals that early investors already viewed the company favorably, which tends to attract more capital in subsequent rounds. Matching on this variable ensures that observed differences in total funding are not simply a reflection of AI startups starting from a higher capital base.

These three variables capture the most important structural differences between startups that are unrelated to AI status.

Figure 3, Pre-Match Distributions of Matching Variables



These pre-match distributions reveal differences between AI and non-AI startups across all three matching variables. AI companies tend to be founded later (median 2018 vs 2016), are more concentrated in California (41.9% vs 34.4%), and enter with slightly more initial capital, a mean log first financing size of -0.23 (approximately $0.79M) for AI startups compared to -0.52 (approximately $0.59M) for non-AI firms.

4. Methodology

4.1 Identification and CEM

A raw comparison of funding outcomes between AI and non-AI startups would conflate the effect of AI classification. As we see in the pre-match distributions, the two groups differ in founding year, state concentration, and initial financing size, all of which independently affect total funding. To isolate the effect of AI status, we apply Coarsened Exact Matching (CEM) to align the two groups on these characteristics before running any regression.

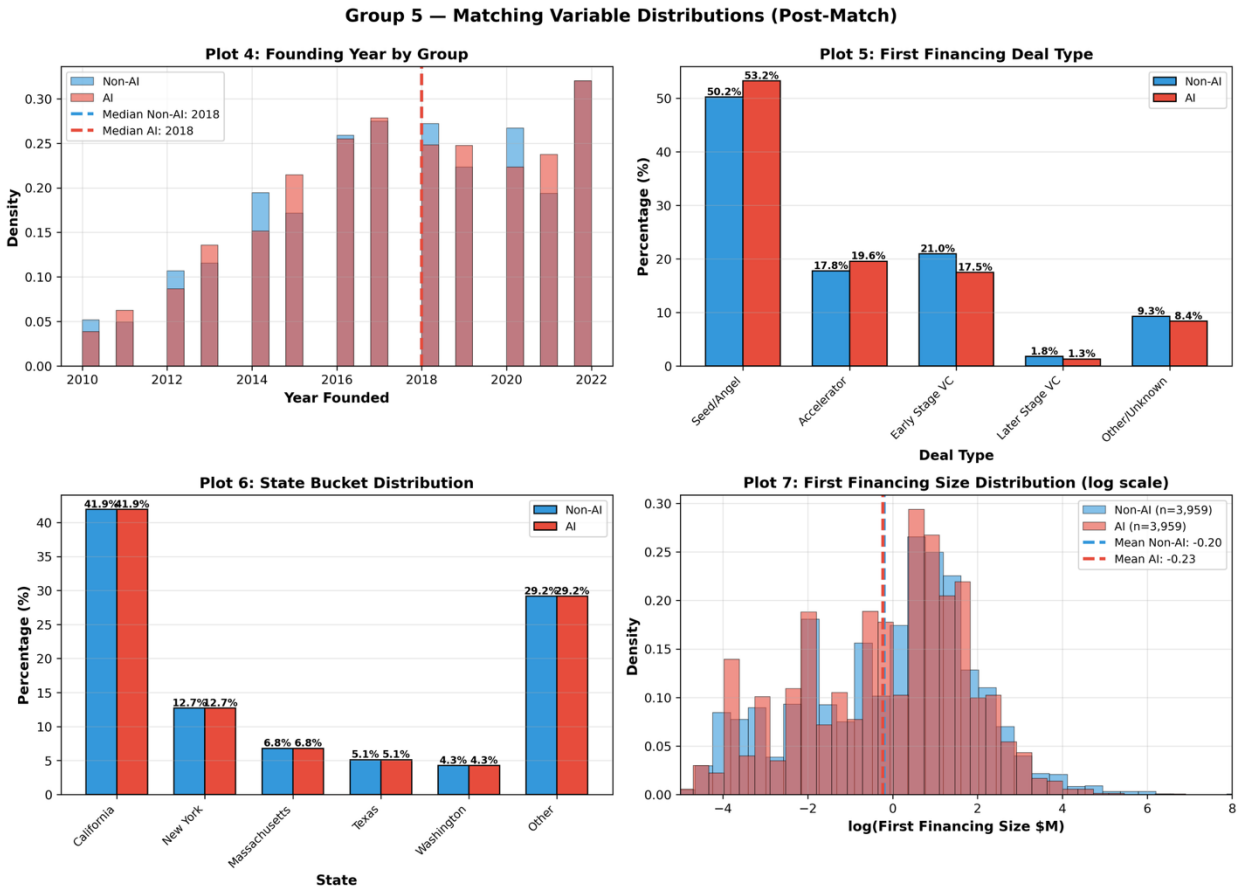
CEM works by temporarily coarsening each matching variable into bins, then retaining only observations that fall into the same bin across all variables simultaneously (Iacus, King, and Porro, 2012). Any observation that cannot find a match in every dimension is removed from the sample. Unlike propensity score matching, which collapses all covariates into a single estimated score, CEM guarantees overlap on each matching variable by construction.

In our dataset, we coarsen founding year into two-year bins, state into six buckets (California, New York, Massachusetts, Texas, Washington, and Other), and first financing size into quartiles of its log-transformed distribution. Founding year is coarsened into two-year bins to allow for minor cohort differences while still grouping companies that operated under broadly similar market conditions. State is collapsed into six buckets reflecting the five states with the highest concentration of VC activity (California, New York, Massachusetts, Texas, and Washington) with all remaining states grouped together. First financing size is log-transformed to reduce skewness and then divided into quartiles, grouping companies by their relative starting capital rather than exact dollar amounts.

After matching, AI companies are treated as the treatment group and retained in full; non-AI companies are matched 1:1 within each cell. The result is a balanced sample of 7,918 companies (3,959 AI and 3,959 non-AI).

4.2 Post Match EDA

Figure 4, Post-Match Distributions of Matching Variables



After matching, the distributions of all three matching variables are closely aligned between AI and non-AI startups, and CEM successfully balanced the two groups.

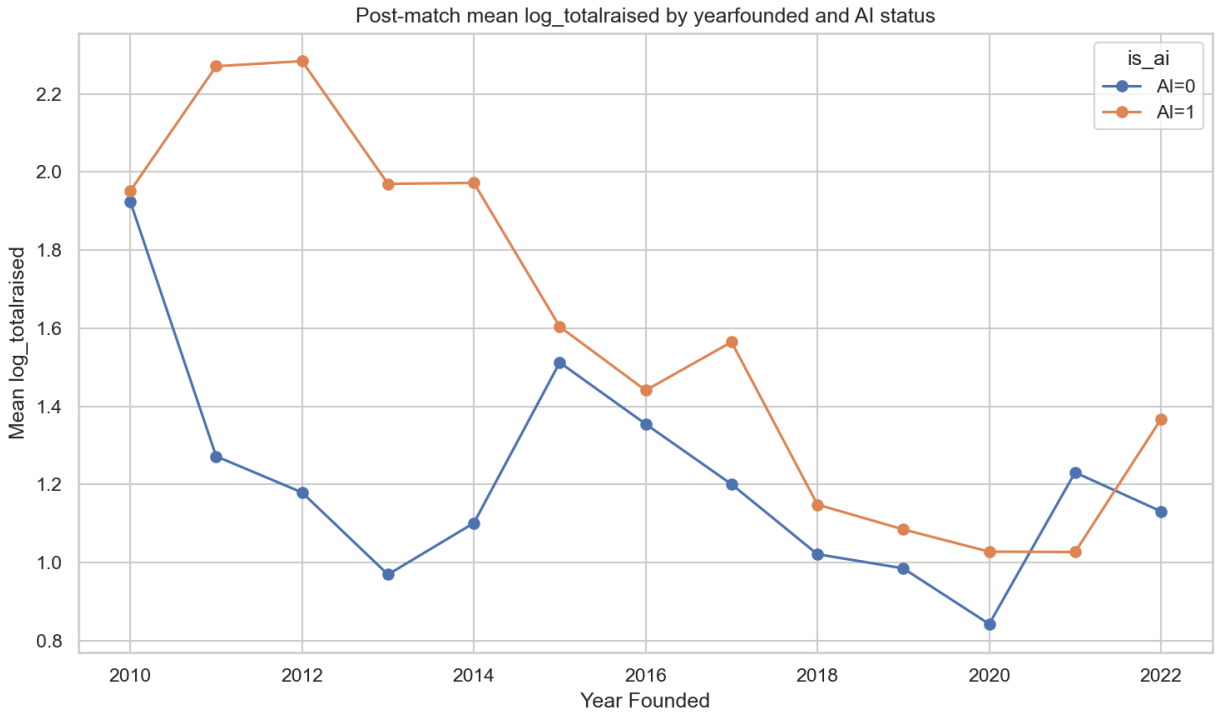
The founding year distribution is now nearly identical, with both groups sharing a median founding year of 2018, down from a two-year gap in the pre-match sample. State distributions are perfectly aligned across all six buckets, with California at 41.9%, New York at 12.7%, Massachusetts at 6.8%, Texas at 5.1%, Washington at 4.3%, and Other at 29.2% for both groups.

First financing size distributions are similarly close, with mean log financing sizes of -0.20 for non-AI and -0.23 for AI companies, compared to a gap of -0.52 versus -0.23 before matching.

More importantly, CEM achieved this balance while retaining 100% of AI companies. All 3,959 AI startups found a match, with an equal number of non-AI companies selected to mirror them, resulting in a final matched sample of 7,918 firms evenly split between the two groups. This is key. Since AI companies are the treatment group of interest, we prioritized retaining every AI observation and matched non-AI firm to them, rather than removing from both sides.

The matched dataset is now ready for regression analysis, but we let us check if we can already see any insight.

Figure 5, Post-Match Mean Log Total Funding by Founding Year and AI Status



The figure above plots mean log total funding raised by founding year, separately for AI and non-AI companies in the matched sample, visually capturing the AI funding premium before any regression

Two patterns stand out. First, AI companies consistently sit above non-AI companies across nearly every founding year cohort, suggesting a persistent funding premium that is visible even in the raw matched data. The gap is particularly large for earlier cohorts. Companies founded in 2011 and 2012 show AI means of around 2.27 log points compared to roughly 1.27 and 1.19 for non-AI, which is a substantial difference. Second, both lines trend downward over time, which is expected. Companies founded more recently have had less time to accumulate funding, so their total raised figures are naturally lower and likely still incomplete.

Importantly, the gap between the two lines narrows for more recent cohorts. By 2018, AI and non-AI companies are much closer together, with the premium nearly disappearing in some years.

4.3 Regression and Hypothesis

This study uses OLS regression to estimate whether AI-classified startups raise more cumulative venture capital than comparable non-AI startups. It is easy to interpret and widely used in the entrepreneurial finance literature. The outcome variable is log-transformed total funding raised.

Model 1 — Baseline:

$$\log(totalraised_i) = \beta_0 + \beta_1(is_{ai_i}) + \epsilon_i$$

The baseline model regresses log total funding on AI status alone. On the matched sample, this gives the raw within-match funding difference between AI and non-AI companies. Our matching has already aligned the two groups on founding year, state, and initial financing size. This simple specification, while baseline, is still pretty credible, and the coefficient β_1 is the primary quantity of interest.

Model 2 – Controlling Year and First Financing:

$$\log(totalraised_i) =$$

$$\beta_0 + \beta_1(is_{ai_i}) + \beta_2(yearfounded_i) + \beta_3(\log firstfinancingsize_i) + \epsilon_i$$

Model 2 adds founding year and log first financing size as continuous controls. Both variables can affect total funding. Companies founded earlier have had more time to raise follow-on rounds, and a larger first financing round signals investor quality and tends to attract more capital downstream. Although both variables were matched on, companies within the same bin can still differ on their exact values. For instance, a startup founded in 2016 and one founded in 2017 fall in the same two-year bin but are not identical. Including these variables as controls improves precision and reduces the risk that any remaining imbalance on these dimensions affects the estimated AI premium.

Model 3 — Full Specification:

$$\log(totalraised_i) =$$

$$\beta_0 + \beta_1(is_{ai_i}) + \beta_2(yearfounded_i) + \beta_3(\log firstfinancingsize_i)$$

$$+\beta_4(state_k) + \beta_5(dealtype_j) + \epsilon_i$$

Model 3 adds state bucket dummies and first financing deal type dummies. State dummies absorb any residual variation in funding outcomes associated with geographic location that is not already captured by matching, improving the precision of the AI coefficient although not by much. The more important addition is the deal type dummies, which were not matched on at all. The entry channel through which a startup first receives funding can independently affect its total capital trajectory. For example, accelerator-backed companies typically start with smaller checks and follow different fundraising paths than those backed by institutional early-stage VCs from beginning. Including deal type dummies ensures these differences do not confound the estimated AI premium.

The central hypothesis of this study is that AI startups raise more cumulative venture capital than comparable non-AI startups, evaluated by the coefficient on the AI indicator variable, β_1, across all three model specifications. The null and alternative hypotheses are defined as:

$$H_0: \beta_1 = 0$$

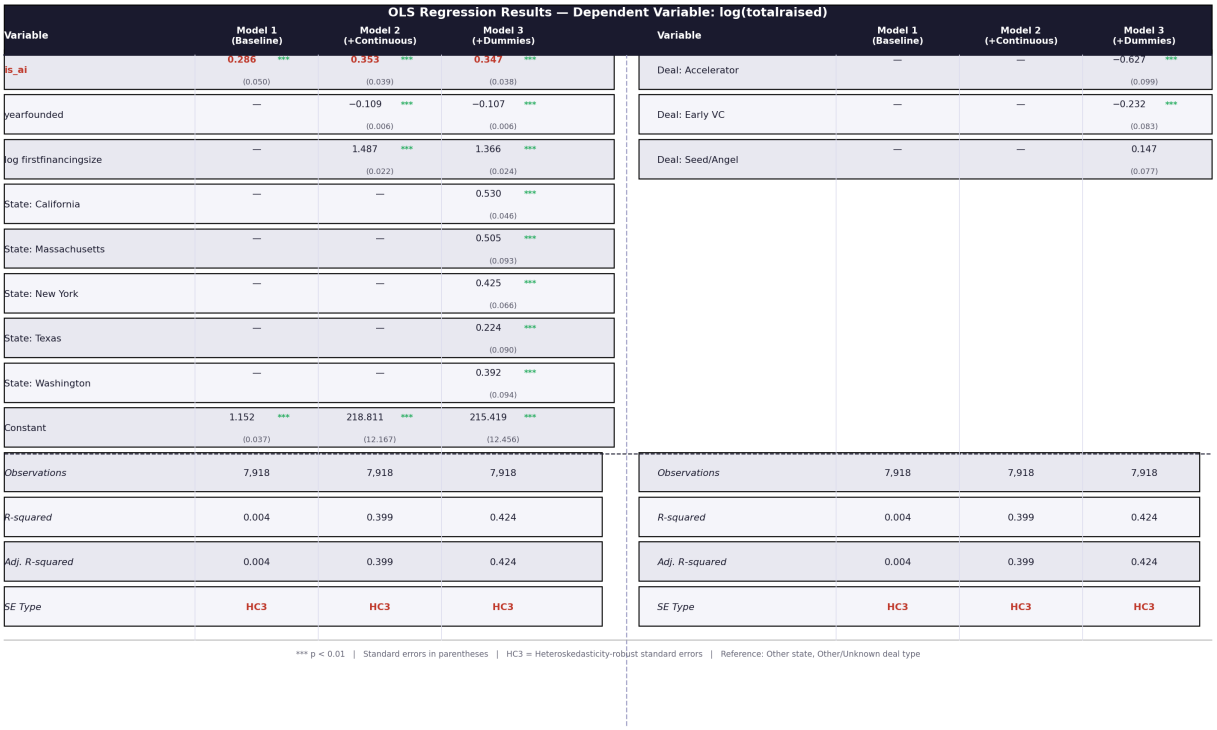
$$H_a: \beta_1 > 0$$

The null hypothesis states that there is no difference in cumulative funding between AI and non-AI startups after controlling for observable characteristics. The alternative hypothesis states that AI startups raise more venture capital on average than comparable non-AI startups. If the estimated β_1 is positive and statistically significant across all three models, this provides evidence in favor of the alternative hypothesis. Crucially, the stability of β_1 as controls are added from Model 1 through Model 3 is important to note. A coefficient that remains consistent across specifications is less likely to be driven by omitted variable bias and more likely to indicate a genuine relationship between AI classification and funding outcomes. Hypothesis testing is conducted using standard t-tests on the estimated coefficient β_1 at the 1% significance level.

5. Results

5.1 Main Reg Results and Coefficient Stability

Figure 6, OLS Regression Results Across All Three Models



OLS Regression Results — Dependent Variable: log(totalraised)

Variable	Model 1 (Baseline)	Model 2 (+Continuous)	Model 3 (+Dummies)
is_ai	0.286 *** (0.050)	0.353 *** (0.039)	0.347 *** (0.038)
yearfounded	—	−0.109 *** (0.006)	−0.107 *** (0.006)
log firstfinancingsize	—	1.487 *** (0.022)	1.366 *** (0.024)
State: California	—	—	0.530 *** (0.046)
State: Massachusetts	—	—	0.505 *** (0.093)
State: New York	—	—	0.425 *** (0.066)
State: Texas	—	—	0.224 *** (0.090)
State: Washington	—	—	0.392 *** (0.094)
Constant	1.152 *** (0.037)	218.811 *** (12.167)	215.419 *** (12.456)
Observations	7,918	7,918	7,918
R-squared	0.004	0.399	0.424
Adj. R-squared	0.004	0.399	0.424
SE Type	HC3	HC3	HC3

Variable	Model 1 (Baseline)	Model 2 (+Continuous)	Model 3 (+Dummies)
Deal: Accelerator	—	—	−0.627 *** (0.099)
Deal: Early VC	—	—	−0.232 *** (0.083)
Deal: Seed/Angel	—	—	0.147 (0.077)
Observations	7,918	7,918	7,918
R-squared	0.004	0.399	0.424
Adj. R-squared	0.004	0.399	0.424
SE Type	HC3	HC3	HC3

*** p < 0.01 | Standard errors in parentheses | HC3 = Heteroskedasticity-robust standard errors | Reference: Other state, Other/Unknown deal type

Figure 6 reports the OLS regression results across all three model specifications. In the baseline model, the coefficient on the AI indicator is 0.286 and is statistically significant at the 1 percent level, implying that within the matched sample, AI-classified startups raise approximately 33.1% more cumulative venture capital than comparable non-AI startups.

In model 2, the coefficient on the AI indicator increases to 0.353 and remains statistically significant at the 1 percent level, implying a funding premium of approximately 42.3%. The increase in the AI coefficient relative to Model 1 suggests that once residual within-cell variation in founding year and initial financing size is accounted for, the estimated AI premium becomes more pronounced. Also, we note that the coefficient on founding year is −0.109 and is significant at the 1 percent level, consistent with the expectation that more recently founded

companies have had less time to accumulate follow-on funding. The coefficient on log first financing size is 1.487 and is also significant at the 1 percent level, confirming that a larger initial round is strongly associated with higher cumulative funding. The addition of these two controls substantially improves model fit, with R-squared rising from 0.004 to 0.399.

In model 3, the coefficient on the AI indicator is 0.347, statistically significant at the 1 percent level, implying a funding premium of approximately 41.5%. The near-identical coefficient relative to Model 2 confirms that the AI premium is not driven by geographic concentration or entry channel differences. Among the state dummies, California carries the largest premium (0.530, $p < 0.01$), followed by Massachusetts (0.505, $p < 0.01$), Washington (0.392, $p < 0.01$), New York (0.425, $p < 0.01$), and Texas (0.224, $p < 0.01$), all relative to the Other state reference category. This is consistent with the geographic concentration of venture capital documented by Chen et al. (2010), where startups in major VC hubs systematically attract more capital. Model fit improves further with R-squared rising to 0.424.

Taken together, the stability of the AI coefficient across all three models provides strong evidence that the estimated funding premium is robust and real. AI-classified startups raise between 33 and 42 percent more cumulative venture capital than matched non-AI startups, after accounting for when and where they were founded and how they first entered the venture ecosystem.

5.2 OLS Assumptions

Figure 7, OLS Diagnostic Tests Across Models

Table: OLS Diagnostic Tests Across Models

Assumption	Statistic	Model 1 (is_ai only)	Model 2 (+ year, first fin.)	Model 3 (+ state & deal dummies)
Normality (Jarque-Bera)	Jarque-Bera stat	234.22	246.42	240.19
	Jarque-Bera p-value	<0.001	<0.001	<0.001
Homoskedasticity (Breusch-Pagan)	Breusch-Pagan stat	19.20	754.72	800.80
	Breusch-Pagan p-value	<0.001	<0.001	<0.001
Independence (Durbin-Watson)	Durbin-Watson	1.164	1.800	1.824
Multicollinearity (Max VIF)	Max VIF	—	3.15	4.87

ge ■ Red = threshold exceeded — = not applicable (single predictor)

Figure 7 reports OLS diagnostic tests across all three model specifications. Normality is rejected in all three models based on the Jarque-Bera test ($p < 0.001$), though this is unsurprising given the sample size of nearly 8,000 observations. With this many observations, the Central Limit Theorem ensures the sampling distribution of the coefficients is approximately normal.

While heteroskedasticity is present across all models, as confirmed by the Breusch-Pagan test ($p < 0.001$ in all cases), this is addressed throughout by using HC3 heteroskedasticity-robust standard errors.

The Durbin-Watson statistic improves notably from Model 1 (1.164) to Models 2 and 3 (1.800 and 1.824), both of which fall within the acceptable range around 2.0, suggesting that the additional controls in Models 2 and 3.

Multicollinearity is also not a concern. Max VIF reaches 4.87 in Model 3, well below the conventional threshold of 10, Overall, the diagnostic picture is stable across all three models, and any violations are either inconsequential or directly addressed using robust standard errors.

5.3 Descriptive Supporting Evidence

While total funding is the primary measure of a startup's success, there are also other interesting variables of interest. Briefly describe what business status and size multiple are.

Figure 8, Business Status Distribution: AI vs Non-AI Companies

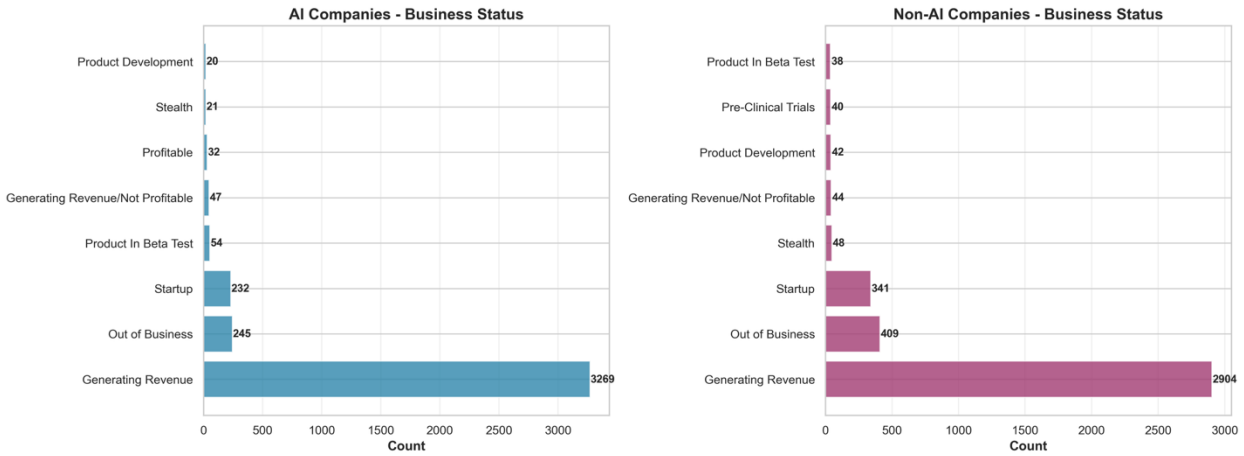
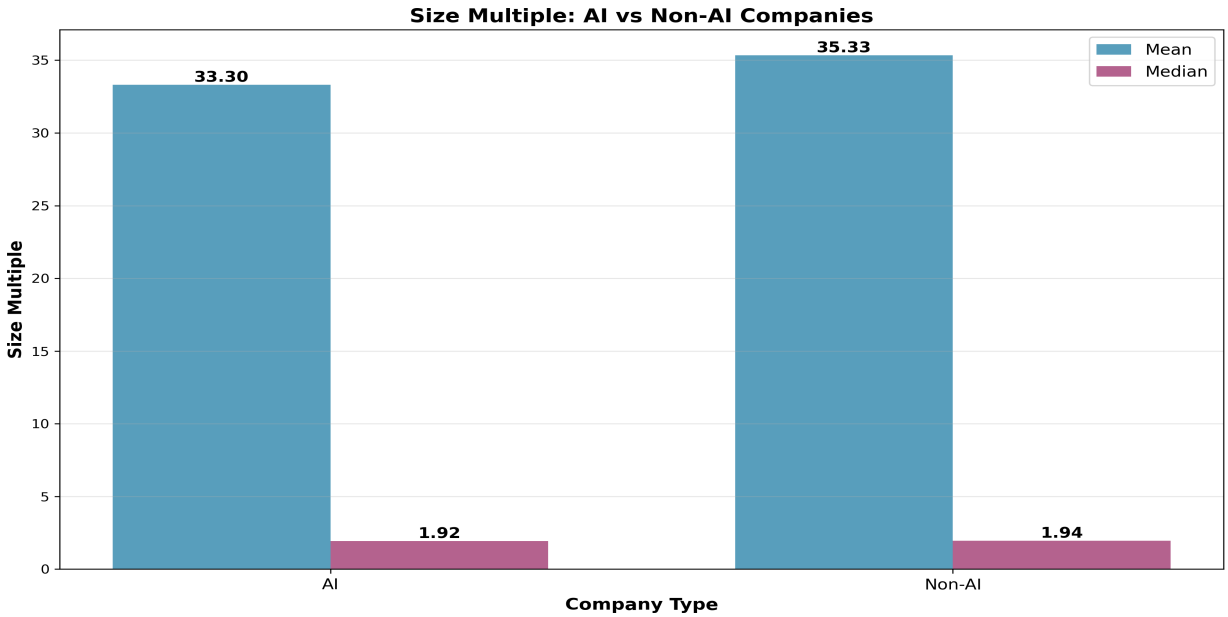


Figure 9, Size Multiple Comparison: AI vs Non-AI Companies



While total funding raised is the primary outcome of interest, two additional variables offer more helpful insight. Business status captures a company's current operational state, and size multiple measures a company's current size relative to peers.

The business status distributions are broadly similar across the two groups. The large majority of both AI (3,269) and non-AI (2,904) companies are classified as Generating Revenue. AI companies show slightly fewer Out of Business classifications (245 vs 409), which could suggest modestly better survival rates.

The size multiple comparison shows virtually identical medians, suggesting that the typical AI and non-AI company in the matched sample are comparable in relative size despite the funding premium documented in the regression.

6. Robustness Checks

6.1 Alternative Years

To check the robustness of the main finding, two additional checks are done. The first reruns the matching procedure using 3-year founding year bins instead of 2-year bins, then re-estimates the full Model 3 specification on the resulting matched sample. This tests whether the results are sensitive to how finely founding year groups are divided. If the AI premium holds under this alternative matching, it means the main result is not due to a year specific binning choice.

The second check introduces an interaction term between AI status and a post-2017 indicator, motivated by how earlier finding of the AI funding gap narrowing for more recent cohorts. We ultimately want to check whether and how much stronger the AI the premium was before 2017, when AI was a less common and potentially stronger signal to investors.

Figure 10, Robustness Checks: 3-Year Bin and Interaction Model Results

REGRESSION COMPARISON: COEFFICIENTS & MODEL FIT

MODEL 1: 3-YEAR BINS

(post_miss dataset)

Variable	Coefficient	Sig.
is_ai	0.3182	***
yearfounded	-0.1084	***
log_firstfinancingsize	1.3919	***
state_California	0.5753	***
state_Massachusetts	0.6860	***
state_New York	0.4045	***
state_Texas	0.3017	**
state_Washington	0.4538	***
deal_Accelerator	-0.5800	***
deal_Early_VC	-0.2945	***
deal_Seed_Angel	0.0766	

N Observations	R-squared	Adj. R-squared
7,918	0.4193	0.4185

MODEL 2: INTERACTION (is_ai × post2017)

(cem_matched dataset)

Variable	Coefficient	Sig.
is_ai	0.5530	***
post2017	-0.0538	
is_ai × post2017	-0.3338	***
yearfound	-0.0793	***
log_firstfinancingsize	1.3645	***
state_California	0.5261	***
state_Massachusetts	0.5031	***
state_New York	0.4202	***
state_Texas	0.2187	**
state_Washington	0.3841	***
deal_Accelerator	-0.6305	***
deal_Early_VC	-0.2277	**
deal_Seed_Angel	0.1498	*

N Observations	R-squared	Adj. R-squared
7,918	0.4265	0.4255

Both are reassuring. The 3-year bin model results in an AI coefficient of 0.318, significant at the 1 percent level, implying that AI startups raise approximately 37.4% more than matched non-AI startups. This is close with the 41.5% premium estimated in the 2-year binning, so the result is not sensitive to the binning choice.

The interaction model tells an interesting story. The baseline AI coefficient for pre-2017 companies is 0.553, a funding premium of approximately 73.8% for earlier cohorts. The interaction term is −0.334, significant at the 1 percent level, bringing the implied post-2017 premium down to approximately 0.219 log points, or about 24.5%.

7. Discussion

7.1 Interpretation of Main Findings

The central finding of this paper is that AI-classified startups raise between 33 and 42 percent more cumulative venture capital than comparable non-AI startups, after matching on founding year, state, and initial financing size, and controlling for additional observable characteristics. This is a substantial premium by any reasonable standard. It represents a structural advantage in a startup's ability to grow and survive long enough to reach profitability or exit.

What makes this finding noteworthy is its stability. The AI coefficient moves very little as controls are added across the three model specifications — from 0.286 in the baseline to 0.353 with continuous controls to 0.347 in the fully specified model. A coefficient that remains essentially unchanged as the model becomes more saturated is strong evidence that the estimated premium is not being driven by observable confounders.

This stability points toward a structural rather than purely speculative interpretation. If the premium were driven primarily by investor enthusiasm or narrative framing, we would expect it to be more sensitive to the period studied and more concentrated in specific geographic markets or entry channels. Instead, it appears broadly consistent. AI-classified startups, it seems, do not simply raise more because they happen to be in California or because they entered the ecosystem through a prestigious early VC round. While they help for sure, they ultimately raise more even after these advantages are accounted for.

7.2 Comparison to Prior Literature

Babina et al. (2024) noted how AI investment is associated with real productivity gains and product innovation in established firms. This is consistent with the premium documented here. If

AI generates genuine economic value, it is reasonable that forward-looking investors would direct more capital toward AI-classified startups in anticipation of those benefits.

While Montanaro et al. (2024) find that AI ventures receive less funding per individual round than comparable non-AI firms, this paper reaches the opposite conclusion when cumulative funding is the outcome. Still, the two findings are not contradictory. A startup that raises smaller individual rounds but sustains investor engagement over many rounds can still accumulate substantially more total capital. The per-round penalty Montanaro et al. document may understate the long-run cumulative advantage that AI startups have, and this paper extends their analysis by shifting focus from the deal to the firm's full funding history.

7.3 Implications

The findings carry implications for founders, investors, and policymakers thinking about the role of AI in the venture capital ecosystem.

For founders, the results suggest that genuine AI integration into a company's core business model, rather than superficial adoption of AI terminology, is associated with sustained investor support over time. The funding premium documented here spans more than a decade and persists across geographic markets, entry channels, and founding cohorts. Founders who build businesses where AI is central to the product may be better positioned to attract not just initial capital but continued investor engagement as the company matures. In the current environment, where the AI label has become too standard, this distinction between genuine and superficial AI integration may matter more than ever.

For investors, the historical record provides a useful calibration point. The AI funding premium predates the current hype cycle by more than a decade, suggesting that the preference for AI-classified startups is not simply a product of recent enthusiasm around large language models and generative AI. Investors appear to have consistently perceived AI-focused ventures as more attractive. At the same time, the declining premium in post-2017 cohorts is a reminder that signaling value diminishes as categories become crowded. In today's environment, where virtually every startup claims some form of AI capability, investors face a harder identification problem than they did in 2012 or 2015, and historical premiums may not translate directly into the current landscape without more careful diligence on the depth and proprietary nature of a startup's AI capabilities.

For policymakers, the premium could be interpreted as evidence of rational capital allocation. Investors directing resources toward a category that has historically been associated with real economic value creation rather than pure speculation or herding behavior. However, this does not mean that all AI investment is always well-allocated, particularly given the concerns about GPT wrappers that have emerged in the post-2022 period. But it does suggest that the historical pattern of elevated funding for AI-classified startups has a structural basis, and that policy interventions aimed at redistributing capital away from AI toward other sectors should account for the possibility that some of this concentration reflects genuine differences in economic potential.

8. Conclusion

Several limitations are worth noting. AI classification relies on keyword matching across Pitchbook descriptions and tags, which may misclassify firms that use AI terminology loosely or

fail to capture those that are substantively AI-driven but describe themselves differently. A more precise classification based on patent data, hiring records, or product-level analysis would strengthen the treatment indicator. Additionally, cumulative funding may be an imperfect measure of success. It captures investor conviction but not revenue, profitability, or exit outcomes. Unobserved confounders such as founder quality and investor network connections also remain a limitation and are not available in the data.

Several options are open for future research. Access to more detailed data would allow matching on additional characteristics such as industry, further tightening the comparison between AI and non-AI firms. We can also extend the analysis beyond 2022, which would help assess whether the historical premium persists in the middle of the increased noise surrounding generative. Finally, studying alternative outcome measures like exit valuations or revenue generated, would further explain if the funding premium documented here translates into sustained economic outperformance.

References

Babina, Tania, et al. "Artificial Intelligence, Firm Growth, and Product Innovation." *Journal of Financial Economics*, vol. 151, 2024, article 103745, https://doi.org/10.1016/j.jfineco.2023.103745.

Chen, Henry, et al. "Buy Local? The Geography of Venture Capital." *Journal of Urban Economics*, vol. 67, no. 1, 2010, pp. 90–102, https://doi.org/10.1016/j.jue.2009.09.013.

Iacus, Stefano M., et al. "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis*, vol. 20, no. 1, 2012, pp. 1–24, https://doi.org/10.1093/pan/mpr013.

Montanaro, Benedetta, et al. "Venture Capital Investments in Artificial Intelligence." *Journal of Evolutionary Economics*, vol. 34, 2024, pp. 1–28, https://doi.org/10.1007/s00191-024-00857-7.

Rady, Judy, et al. "The Expectations Game: The Contingent Value of Hype as a Rhetorical Strategy in Resource Mobilization Processes among AI Startups." *Journal of Business Venturing*, vol. 40, 2025, article 106499, https://doi.org/10.1016/j.jbusvent.2025.106499.

Van Romburgh, Marlize. "The State of Startups in 12 Charts: AI Soars, Asia Tanks, Seed Stalls and More." *Crunchbase News*, 3 Feb. 2025, https://news.crunchbase.com/venture/startups-ai-seed-investors-data-charts-ye-2024/.