



Institute of Economic Studies  
Charles University

Bachelor's Thesis in Economics and Finance

**Prevailing trends in venture capital funding for startups in the  
artificial intelligence sector**

Author: Tomáš (Thomas) Turek

Study program: Economics and Finance

Supervisor: Doc. PhDr. Martin Gregor, Ph.D.

Year of defense: 2025

Length of thesis: 79,626 characters incl. spaces

## **Dedication**

This thesis is dedicated to my great-grandfather, Ulrich (Oldřich) Kindl, whose perseverance, integrity, and determination through struggle faced with unwavering resolve have left an indelible mark on my life.

*“A man can be himself only so long as he is alone; and if he does not love solitude, he will not love freedom; for it is only when he is alone that he is really free”*

- Arthur Schopenhauer, *Essays of Schopenhauer*

## **Acknowledgements**

I would like to acknowledge that this thesis would not have been possible without the guidance and help of several individuals and organizations. First and foremost, I extend my sincere gratitude to doc. PhDr. Martin Gregor, Ph.D. for his invaluable support and supervision throughout this project. His expertise, insights, and guidance have been instrumental in shaping this thesis. I would also like to express my appreciation to Crunchbase for providing the data that formed a crucial foundation for my research. Their comprehensive database has been an essential resource in conducting this study.

## **Declaration**

I hereby declare that I have compiled this thesis using the listed literature and resources only. I declare that my thesis has not been used to gain any other academic title. I fully agree to my work being used for study and scientific purposes. During the preparation of this thesis, the author used Open AI Chat GPT and Grammarly in order to improve language. After using these services, the author reviewed and edited the content as necessary and takes full responsibility for the content of the publication.

A handwritten signature in black ink, appearing to be 'A. Kindl', written over a horizontal line.

# 1. Table of Contents

- 2. Abstract 5
- 3. Literature Review 6
- 4. Data Collection 9
  - 4.1. Data Source 9
  - 4.2. Dataset Description 9
  - 4.3. Defining AI Startups 10
  - 4.4. Data Collection Process 10
  - 4.5. Data Preprocessing 10
    - 4.5.1. Metric Calculation 10
    - 4.5.2. Data Merging and Enhancement 11
    - 4.5.3. Data Quality Assurance 11
  - 4.6. Derived Variables from Analysis 12
  - 4.7. Methodological Limitations and Considerations 12
  - 4.8. Final Dataset Characteristics 12
- 5. Data Analysis 13
  - 5.1. Exploratory Analysis 14
    - 5.1.1. Follow-on Funding Rates 14
    - 5.1.2. Funding Amounts and Distributions 14
    - 5.1.3. Temporal Funding Patterns 15
    - 5.1.4. Summary of Exploratory Findings 15
  - 5.2. Methodology 16
    - 5.2.1. Predicting Follow-on Funding Success 16
    - 5.2.2. Analyzing Follow-on Funding Amounts 17
    - 5.2.3. Examining Timing Between Funding Rounds 18
    - 5.2.4. Comparative Analysis Framework 19
  - 5.3. Results and Comparative Analysis 19
    - 5.3.1. Factors Predicting Follow-on Funding Success 19
    - 5.3.2. Evolutionary Patterns in Follow-on Funding Success Factors 25
    - 5.3.3. Enhanced Logistic Regression Analysis with Progressive Controls 35
    - 5.3.4. Characteristics Associated with Follow-on Funding Amounts 36

5.3.5.	Evolution in Follow-on Funding Amount Determinants	39
5.3.6.	First-to-Second Round Dynamics Analysis	39
5.3.7.	Time Between Funding Rounds	41
5.3.8.	Temporal Acceleration and Structural Consistency	43
5.3.9.	Synthesized Evolutionary Patterns	44
5.3.10.	Alternative Explanations for Observed Patterns	46
6.	Conclusions	46
7.	Works Cited	48

## 2. Abstract

This thesis investigates how venture capital (VC) funding patterns for artificial intelligence (AI) startups evolved between the pre-2014 and post-2014 periods, using a dataset of 15,608 AI startups. The study finds three key transformations: (1) increased investor selectivity, with follow-on funding rates declining from 81.6% to 71.5%, (2) accelerated funding cycles, with median time between rounds decreasing by 24% (from 526 to 398 days), and (3) standardization of the relationship between time to first funding and follow-on success, where this relationship became more consistently negative and significant across all funding sizes in the post-2014 period. These findings suggest a maturing AI sector characterized by more competitive funding environments and increased emphasis on early validation, with significant implications for entrepreneurs, investors, and policymakers.

### Abstrakt

Tato práce zkoumá, jak se vyvíjely vzorce financování rizikového kapitálu (VC) pro startupy v oblasti umělé inteligence (AI) mezi obdobími před rokem 2014 a po roce 2014, s využitím datasetu 15 608 AI startupů. Studie identifikuje tři klíčové transformace: (1) zvýšenou selektivitu investorů, kdy míra následného financování klesla z 81,6 % na 71,5 %, (2) zrychlené cykly financování, kdy se medián času mezi koly snížil o 24 % (z 526 na 398 dní), a (3) standardizaci vztahu mezi dobou do prvního financování a úspěšností následného financování, kdy se tento vztah stal konzistentněji negativním a významným napříč všemi velikostmi financování v období po roce 2014. Tyto nálezy naznačují dozrávající sektor AI charakterizovaný konkurenčnějším prostředím financování a zvýšeným důrazem na včasnou validaci, s významnými implikacemi pro podnikatele, investory a tvůrce politik.

### Key Words

Venture capital, artificial intelligence startups, funding dynamics, investment patterns, startup funding success, AI sector, temporal evolution, funding predictors

**Klíčová slova**

Rizikový kapitál, startupy umělé inteligence, dynamika financování, investiční vzorce, úspěšnost financování startupů, sektor umělé inteligence, časový vývoj, prediktory financování

**Title**

Prevailing trends in venture capital funding for startups in the artificial intelligence sector

**Název práce**

Převládající trendy ve financování rizikového kapitálu pro startupy v sektoru umělé inteligence

**3. Literature Review**

The evolution of AI as a transformative technology has attracted substantial VC investment over the past decade. Despite the growing body of literature on the characteristics of AI startups and VC decision-making, critical gaps remain in understanding how these funding patterns have evolved over time. This literature review synthesizes current research across key dimensions, including theoretical foundations, the unique characteristics of AI startups, the evolution of VC decision-making criteria, and institutional influences, to justify a temporal analysis of AI startup funding patterns.

Existing literature provides valuable insights into the theoretical underpinnings of technology adoption and investment cycles. Rogers' diffusion of innovation theory (2) offers a foundational framework for understanding how technologies traverse the stages of adoption, from early innovation to mainstream integration. The year 2014 aligns with a critical juncture in this diffusion process, a finding empirically supported by Shao et al., who identified 2014 as a significant inflection point in AI research and application trends (8). Their analysis of publication patterns and emerging subfields like reinforcement learning validates our temporal division. Similarly, Jacobides et al. document how specialized AI ecosystems emerged and became interlinked precisely during this post-2014 period (420). AI technology, following early

breakthroughs in deep learning in 2012, began crossing the chasm from niche adoption to widespread commercialization during this pivotal transition.

Supporting this transition, Gompers et al. emphasize the influence of public market signals on VC behavior (22), noting that market conditions play a pivotal role in guiding investment strategies during periods of technological transformation. The introduction of Generative Adversarial Networks (GANs) and Google's acquisition of DeepMind in 2014 are emblematic of the shift from experimental research to impactful applications, further establishing the year as a significant inflection point in AI development and investment.

Signaling theory (Spence 434) provides an additional lens through which to analyze VC decision-making. As markets mature, the nature and reliability of signals change, affecting how investors evaluate startups. Early-stage AI startups may rely on technical achievements or founder credentials as primary signals, while later-stage ventures emphasize market traction and scalability. This dynamic process is further illuminated by Gompers' foundational work on staged financing, which demonstrates how sequential investments help mitigate information asymmetries and agency costs over multiple rounds ("Optimal Investment", 1462). The model provides a theoretical foundation for understanding the evolving relationship between initial rounds and follow-on funding that is observed in the AI sector, where staged financing serves both monitoring and signaling functions. Perrone highlights the role of intellectual property and talent concentration as key signals in AI startup funding, reflecting broader shifts in the types of indicators valued by VCs pre- and post-2014 (44). This framework offers a theoretical basis for understanding how signaling dynamics evolved as AI transitioned into a mainstream investment category.

The characteristics of AI startups further underscore the unique dynamics of this investment landscape. Warzyńska et al. conclude that AI startups attract a larger number of investors overall, with government entities and business angels showing stronger interest in AI ventures, though venture capitalists demonstrated greater preference for general software companies (23). Similarly, Weber et al. provide a comprehensive taxonomy of AI startup business models, identifying archetypes such as "AI Development Facilitators" and "Deep Tech Researchers" (100) that illustrate the diversity of approaches within the AI ecosystem. These

characteristics underscore the complexity and heterogeneity of the AI startup landscape, further justifying a nuanced temporal analysis to capture shifts in investor priorities and startup strategies over time.

The evolution of VC decision-making criteria reflects broader trends in the technology investment landscape. Gottfried et al. provide a comparative analysis of VC priorities across the United States, China, and Germany, highlighting the interplay of cultural, economic, and technological factors. Their findings reveal significant regional differences in the weighting of criteria such as team experience, product innovation, and personal relationships, suggesting that temporal shifts in AI funding patterns may also reflect institutional variations (19). Torssell complements this perspective by identifying team composition (25) as a critical determinant of follow-on funding success, emphasizing the importance of balanced expertise in technical and business domains. These insights highlight the evolving priorities of VCs as the AI market matures, underscoring the need for a temporal framework that accounts for changing evaluation criteria.

Despite these contributions, significant gaps persist in the existing literature. Few studies explicitly examine the temporal dynamics of AI funding, leaving a critical void in understanding how patterns of investment and follow-on funding have evolved over time. Moreover, while existing research provides valuable insights into the drivers of VC decisions, it often lacks a robust theoretical integration that connects these drivers to broader patterns of technological and market evolution. The absence of studies explicitly addressing 2014 as a turning point represents a missed opportunity to contextualize the transformative changes in AI technology and investment that occurred during this period.

This review thus establishes the necessity of a temporal analysis of AI funding patterns by integrating theoretical and empirical insights from diverse strands of literature. The findings underscore the importance of a multidisciplinary approach that incorporates perspectives from technology diffusion, signaling theory, and regional economics to illuminate the complex interplay of factors driving AI funding patterns over time.

## **4. Data Collection**

This section describes the systematic methodology used to collect and process data on venture capital funding for AI startups across two time periods: pre-2014 and 2014-2023. The approach supports comparative analysis addressing three research objectives: predicting follow-on funding success, identifying characteristics associated with funding amounts, and analyzing temporal patterns between funding rounds.

### **4.1) Data Source**

Crunchbase was selected as the primary data source due to its comprehensive global coverage of startup funding activities. The database aggregates information from multiple sources including press releases, user submissions, and automated news scanning, providing detailed startup data on funding rounds, investor profiles, and company characteristics. This source is widely recognized in academic research for tracking venture capital investments (Perrone 46; Weber et al. 95). However, limitations include potentially incomplete reporting for early-stage startups, reporting delays, and selective self-reporting of positive developments.

### **4.2) Dataset Description**

The initial combined dataset contains information on 15,608 AI-related startups (3,256 pre-2014 and 12,352 post-2014), with key variables including company identifiers (name, website), company descriptions, founding date, funding metrics (total funding in USD, number of funding rounds), latest funding date, and operational status (operating, acquired, or closed).

For each startup, detailed funding round information is available through separate datasets: funding round identifiers, announcement dates, investment types (seed, series A, series B, etc.), money raised per round, post-money valuations (where available), and number of investors per round. The combined funding rounds dataset contains 45,479 individual funding events (12,399 for pre-2014 startups and 33,080 for post-2014 startups).

Additional derived metrics specifically created for this analysis include: binary indicator for follow-on funding, time intervals between consecutive funding rounds, total follow-on funding amounts, maximum follow-on funding round size, and average time between rounds. A

particularly important metric, time to first funding, was calculated to measure the period between company founding and initial funding acquisition, providing insights into early-stage investor interest.

### **4.3) Defining AI Startups**

This study defines an AI startup as a company whose primary product or service relies fundamentally on artificial intelligence technologies or incorporates AI as a significant differentiator. This excludes companies merely using off-the-shelf AI tools without core integration. Genuine AI ventures are identified using the term "artificial intelligence" as the primary filter in Crunchbase descriptions.

### **4.4) Data Collection Process**

Data collection employed custom Python scripts utilizing the Crunchbase API with academic research access. The process followed a two-phase approach: (1) retrieving basic company information for AI startups using "artificial intelligence" as the primary search criterion, and (2) gathering detailed funding round data including announcement dates, investment types, and funding amounts. Throughout collection, rate limiting and error handling mechanisms ensured data completeness while respecting Crunchbase API usage policies.

### **4.5) Data Preprocessing**

Raw data underwent preprocessing to ensure quality and consistency across both time periods. This involved converting nested API responses to tabular format, calculating essential metrics, merging datasets, and validating data consistency while handling missing values.

#### **4.5.1) Metric Calculation**

Several essential metrics were calculated for each startup: follow-on funding status, total funding rounds, first round amount, subsequent funding totals, maximum follow-on amount, and average time between rounds. Each startup's funding events were sorted chronologically to

derive accurate metrics. A critical calculation determined time-to-first-funding, addressing challenges such as complex date formats, missing information, and implausible values.

The calculation of time-to-first-funding successfully generated initial values for 2,715 startups from the pre-2014 cohort (83.4% of the total) and 11,895 startups from the post-2014 cohort (96.3% of the total). After removing extreme outliers and implausible values (such as negative times or values exceeding 10 years), the final analysis dataset contained 2,715 pre-2014 startups (unchanged after filtering) and 11,469 post-2014 startups with valid time-to-first-funding data. This consistent dataset was used throughout all temporal analyses to ensure comparable results. For analyses involving multiple variables, such as regression models, the effective sample size was further reduced due to the requirement that all included variables have non-missing values for each observation.

#### **4.5.2) Data Merging and Enhancement**

The merging process combined multi-round and single-round startup datasets for both time periods, verified funding indicators to ensure consistent classification, and corrected discrepancies between funding indicators and actual round counts. Alternative success indicators were created: whether startups received more than two rounds, whether follow-on funding exceeded cohort medians, and whether time between rounds was below median interval. Processing logs confirm successful correction of missing values and inconsistencies in both period datasets.

#### **4.5.3) Data Quality Assurance**

Data validation procedures corrected several specific issues: 579 missing first round amounts (pre-2014) and 2,261 (post-2014), 109 missing follow-on amounts (pre-2014) and 481 (post-2014), and implausible temporal outliers. The process ensured consistency in funding classifications, addressed missing values through statistical methods, and standardized data types across financial variables. Validation confirmed appropriate distributions of follow-on funding outcomes across both periods, resulting in datasets suitable for statistical analysis.

#### **4.6) Derived Variables from Analysis**

Specialized variables were derived to address the research questions. The primary dependent variable indicates follow-on funding success (binary). Secondary variables include total follow-on amount and maximum follow-on round size. Temporal variables comprise average time between rounds and time-to-first-funding. Alternative success indicators include "substantial follow-on" (above-median funding) and "quick follow-on" (below-median time to next round), providing multidimensional measures of funding outcomes.

#### **4.7) Methodological Limitations and Considerations**

Several methodological considerations warrant discussion. Our reliance on "artificial intelligence" as a keyword filter may have excluded companies that use AI but don't explicitly market it or pivoted to AI after their founding, particularly affecting the pre-2014 cohort when AI terminology was less standardized and artificial intelligence had not yet achieved its current prominence in academic and commercial discourse. The definition of what constitutes "AI" has evolved over time, potentially creating inconsistencies between periods. While 2014 represents a theoretically justified temporal boundary, this binary split may oversimplify continuous technological evolution, and alternative approaches such as rolling window analysis, multiple time period divisions, or continuous time variables could have yielded different insights. Finally, our regression models identify associations but cannot establish causality, meaning relationships such as the connection between time to first funding and follow-on success could be bidirectional, with faster initial funding potentially predicting success or inherently successful ventures attracting funding more quickly.

#### **4.8) Final Dataset Characteristics**

The preprocessing efforts resulted in two comprehensive and clean datasets for analysis. The pre-2014 AI startups dataset contains information on 3,256 startups founded before 2014, with 2,658 (81.6%) having secured follow-on funding. This dataset includes 12,399 individual funding rounds and derived metrics for each startup. The post-2014 AI startups dataset encompasses 12,352 startups founded between 2014 and 2023, with 8,839 (71.5%) having

received follow-on funding. This dataset incorporates 33,080 individual funding rounds. Valid time-to-first-funding data was available for 2,715 pre-2014 startups (2,347 with follow-on funding) and 11,469 post-2014 startups (8,217 with follow-on funding). The time intervals range from immediate funding (0 days) to nearly 10 years (3,635 days for pre-2014 and 3,606 days for post-2014), showing substantial variation in early-stage funding acquisition across the AI sector. Both datasets contain parallel structures with identical variables, enabling comparative analysis. Having established these datasets, the next section details the analytical methodology employed to address the research questions regarding follow-on funding dynamics in the AI sector.

## **5. Data Analysis**

This section presents a comprehensive analysis of factors influencing follow-on funding dynamics for AI startups. The analysis employs multiple complementary statistical approaches to address our three primary research questions: what factors predict an AI startup's likelihood of securing follow-on funding, which characteristics are associated with higher follow-on funding amounts, and how startup attributes relate to the timing between funding rounds.

The analytical approach leverages the comparative structure of the dataset, examining patterns across two distinct time periods in the evolution of the artificial intelligence sector. This temporal division enables identification of both persistent patterns in venture capital decision-making and potential shifts in investment criteria as the AI industry has matured.

For each research question, appropriate statistical methodologies were implemented to address the specific nature of the dependent variable. To analyze follow-on funding probability, logistic regression models were employed that account for the binary nature of the outcome variable. For examining funding amounts, multiple regression analysis was utilized to identify factors associated with larger follow-on investments. Finally, to investigate the temporal aspects of funding patterns, survival analysis techniques were implemented that appropriately handle time-to-event data structures.

Throughout the analysis, a consistent approach was maintained across both time periods, enabling meaningful comparisons between the pre-2014 and post-2014 cohorts. For each model, independent variables were carefully selected based on theoretical considerations and existing

literature on venture capital decision-making. These include startup characteristics (such as founding year), initial funding attributes (including first round amount), and temporal factors (such as time to first funding). The analysis also incorporates appropriate data transformations to address the distributional characteristics of funding data. Financial variables often exhibit significant right-skew, necessitating logarithmic transformations to improve model fit and interpretability.

## **5.1) Exploratory Analysis**

Before engaging in formal inferential analysis, an exploratory examination of the dataset was conducted to understand the fundamental characteristics and patterns in AI startup funding across the two time periods.

### **5.1.1) Follow-on Funding Rates**

In the pre-2014 period, 81.6% of AI startups in the dataset secured follow-on funding beyond their initial round. By contrast, this figure declined to 71.5% for startups founded between 2014 and 2023. The distribution of funding rounds varied substantially within both cohorts. While most startups with follow-on funding secured 2-3 rounds, both periods contained notable outliers with significantly higher numbers of funding rounds. The pre-2014 dataset included startups with up to 29 distinct funding events, while the post-2014 period contained companies with up to 27 rounds.

### **5.1.2) Funding Amounts and Distributions**

As shown in Table 1, the data exhibits a counterintuitive pattern where startups that did not secure follow-on funding had substantially higher initial round amounts in both periods. This finding challenges conventional assumptions that larger initial funding predicts follow-on success. Rather, it suggests a strategic division in funding approaches: startups securing larger initial rounds may intentionally raise sufficient capital to reach significant milestones without requiring additional rounds, while those with smaller initial rounds may adopt a more staged approach to funding. This pattern persisted across both time periods but intensified post-2014, suggesting this strategic differentiation has become more pronounced as the AI sector matured.

For the pre-2014 cohort, startups that did not secure follow-on funding raised an average of \$23.03 million in their first round, compared to \$6.27 million for those that did secure additional funding. The post-2014 cohort exhibited a similar pattern, with startups not securing follow-on funding raising \$8.60 million versus \$3.23 million for those that did. Mann-Whitney U tests confirmed that these differences were statistically significant ( $p < 0.0001$ ) for both cohorts. Follow-on funding amounts displayed considerable skewness, with median values substantially lower than means in both periods. The pre-2014 cohort showed a median follow-on funding amount of \$10.33 million, while the post-2014 period demonstrated a median of \$4.79 million.

**Table 1: First Round Characteristics by Follow-on Funding Status**

Period	Follow-on Status	Mean (\$M)	Median (\$M)	Count	Mann-Whitney U	p-value
Pre-2014	No	23.03	2.90	598	500,872.5	<0.0001
Pre-2014	Yes	6.27	1.50	2,658	--	--
Post-2014	No	8.60	2.00	3,519	9,980,038.0	<0.0001
Post-2014	Yes	3.23	1.20	8,839	--	--

### 5.1.3) Temporal Funding Patterns

As noted in Section 4.8, the time from founding to first funding demonstrated substantial variation, ranging from immediate funding (0 days) to nearly 10 years (approximately 3,600 days). For the pre-2014 cohort, 2,715 startups with valid time-to-first-funding data were identified. The post-2014 cohort included 11,469 startups with valid temporal data.

### 5.1.4) Summary of Exploratory Findings

The exploratory analysis revealed key patterns that informed statistical modeling. Follow-on funding rates declined between the two time periods, suggesting potentially increased selectivity in follow-on investment decisions. Counterintuitively, larger first rounds appeared associated with lower likelihood of securing follow-on funding in both periods. Time to first funding demonstrated a complex relationship with follow-on funding success that varies across

funding size categories, with further analysis needed to determine the precise nature of this relationship. These exploratory findings indicated the need for multivariate regression approaches to disentangle the complex relationships between startup characteristics and funding outcomes.

## 5.2) Methodology

This section details the analytical approaches employed to address three research questions concerning the dynamics of follow-on funding for AI startups. For each question, a methodology specifically tailored to the nature of the dependent variable and the theoretical relationships being investigated was implemented.

### 5.2.1) Predicting Follow-on Funding Success

To identify factors that predict an AI startup's probability of securing follow-on funding, logistic regression was employed, a method suited to modeling binary outcomes. The logistic regression framework enables estimation of the odds ratios associated with various predictor variables, providing interpretable measures of their influence on follow-on funding likelihood. The primary dependent variable in this analysis is whether a startup secured any funding beyond its initial round. This binary outcome indicates success in obtaining follow-on investment. The general form of the logistic regression model is:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1\chi_1 + \beta_2\chi_2 + \dots + \beta_n\chi_n$$

Where  $p$  represents the probability of securing follow-on funding, and  $\chi_1$  through  $\chi_n$  represent the predictor variables. The coefficients  $\beta_1$  through  $\beta_n$  indicate the change in log-odds associated with a one-unit change in the corresponding predictor.

Based on theoretical considerations and exploratory analysis, several predictor variables were selected. Financial characteristics included the logarithm of first-round funding amounts, which accounts for the right-skewed distribution of funding values. Categorical indicators were also created for very small (bottom quintile) and very large (top quintile) first rounds to capture potential non-linear effects. Temporal factors encompassed the logarithm of time to first funding,

which showed significant associations with follow-on success in the exploratory analysis. The logarithmic transformation addresses the right-skewed distribution of this variable. Finally, founding period indicators were introduced to control for temporal effects within each cohort, distinguishing between startups founded in the earlier versus later years of each period.

For model estimation, L1 regularization (LASSO) was applied with cross-validation to determine the optimal regularization strength. This approach helps prevent overfitting while identifying the most relevant predictors. Stratified k-fold cross-validation was used with five folds to estimate model performance, employing the area under the receiver operating characteristic curve (AUC) as the primary performance metric.

### **5.2.2) Analyzing Follow-on Funding Amounts**

To investigate the factors associated with the magnitude of follow-on funding, multiple regression models were implemented for the subset of startups that successfully secured follow-on funding. This approach allowed for the estimation of the relationships between startup characteristics and subsequent funding amounts while controlling for potential confounding factors. Two complementary models were specified: one with total follow-on funding as the dependent variable and another with maximum follow-on round size. The form of these regression models is:

$$Y = \beta_0 + \beta_1\chi_1 + \beta_2\chi_2 + \dots + \beta_n\chi_n + \varepsilon$$

Where  $Y$  represents either the total follow-on funding or maximum follow-on round,  $\chi_1$  through  $\chi_n$  represent the predictor variables, and  $\varepsilon$  is the error term.

For these models, predictor variables were selected that were expected to influence follow-on funding amounts based on theoretical considerations and existing literature. These include first round characteristics, including amount raised and investment type, which may signal startup quality to subsequent investors. Temporal factors, including the number of funding rounds (for the total follow-on funding model) and the time interval between founding and initial funding. Additional operational characteristics, including company status and founding timing,

were included to control for potential effects of business maturity and market entry timing on funding amounts.

To address the substantial right-skew in funding amount distributions, logarithmic transformations were applied to both dependent and key independent financial variables. This transformation improves model fit and makes the relationships more interpretable, as the coefficients can be interpreted in terms of percentage changes. The models were estimated using ordinary least squares (OLS) with robust standard errors to account for potential heteroscedasticity. Model diagnostics included tests for multicollinearity, influential observations, and residual normality to ensure the validity of statistical inference.

### **5.2.3) Examining Timing Between Funding Rounds**

To analyze the temporal aspects of funding patterns, survival analysis techniques were employed, which are specifically designed for time-to-event data. This approach is appropriate for examining the factors influencing the duration between funding rounds, as it properly handles the temporal structure of the data.

Two complementary methods were employed within the survival analysis framework. First, non-parametric Kaplan-Meier estimation to visualize the survival function, which represents the probability of not receiving the next funding round as a function of time since the previous round. Second, semi-parametric Cox proportional hazards models to estimate the effects of covariates on the hazard rate (the instantaneous risk of receiving the next funding round). The general form of the Cox model is:

$$h(t) = h_0(t) \exp(\beta_1 \chi_1 + \beta_2 \chi_2 + \dots + \beta_n \chi_n)$$

Where  $h(t)$  is the hazard function,  $h_0(t)$  is the baseline hazard, and  $\chi_1$  through  $\chi_n$  are the predictor variables with corresponding coefficients  $\beta_1$  through  $\beta_n$ .

For the survival analysis, the dataset was restructured to focus on the time intervals between consecutive funding rounds for startups with multiple rounds. The dependent variable is the time (in days) from one funding round to the next. Predictor variables include round-specific

factors such as the round number and the logarithm of the amount raised in the previous round. To address potential convergence issues in the Cox model, a regular penalization term (0.1) was applied, and robust standard errors were used. Tests of the proportional hazards assumption were conducted to ensure model validity.

#### **5.2.4) Comparative Analysis Framework**

To systematically compare funding dynamics between the pre-2014 and post-2014 periods, parallel modeling approaches were implemented for each cohort. This structure enables direct comparison of coefficient magnitudes, significance levels, and overall model performance across time periods. Separate models were estimated for each period to allow for complete flexibility in coefficient estimates. This approach acknowledges the potential for structural changes in the relationships between predictor variables and funding outcomes as the AI sector evolved. For all three methodological approaches, consistent variable specifications were maintained across time periods to facilitate direct comparisons.

### **5.3) Results and Comparative Analysis**

This section presents the empirical findings addressing the three primary research questions of this study: (1) what factors predict an AI startup's likelihood of securing follow-on funding, (2) which characteristics are associated with higher follow-on funding amounts, and (3) how startup attributes relate to the timing between funding rounds. Through comparative analysis of pre-2014 and post-2014 AI startup cohorts, the results reveal both persistent patterns and evolutionary shifts in funding dynamics as the artificial intelligence sector has matured.

#### **5.3.1) Factors Predicting Follow-on Funding Success**

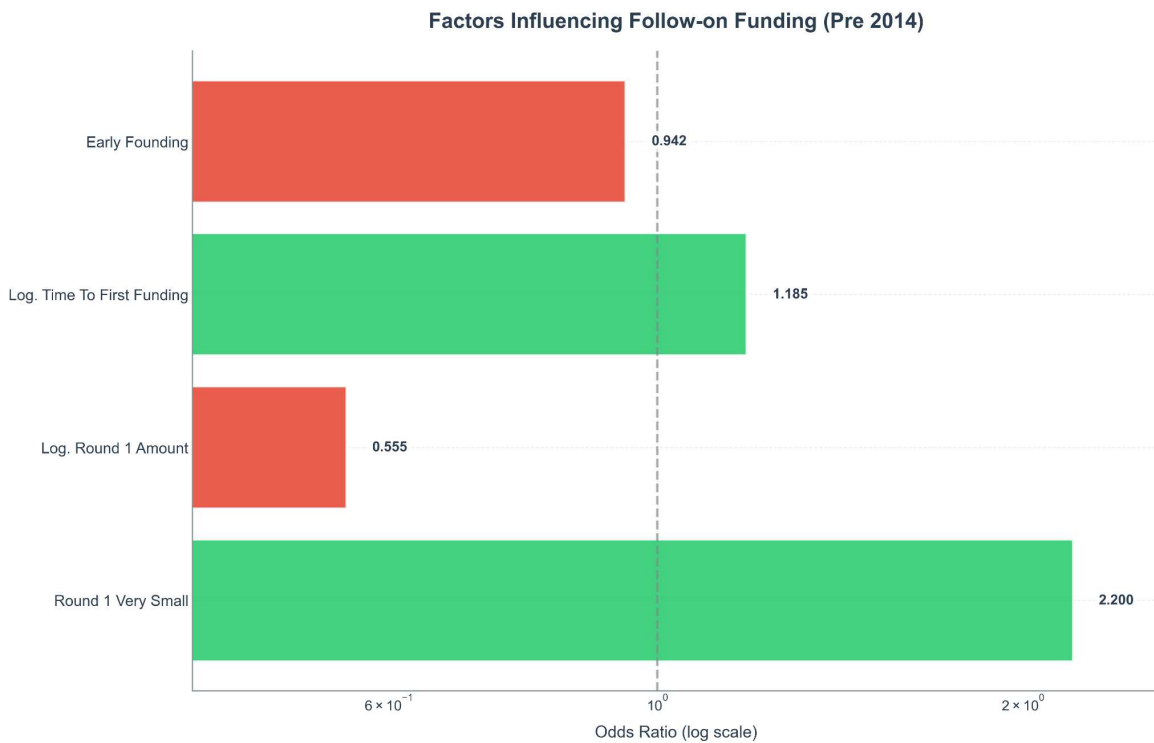
##### **Pre-2014 Cohort**

The logistic regression analysis for pre-2014 AI startups revealed several significant predictors of follow-on funding success. The model demonstrated good predictive performance with an AUC of 0.694, building on the exploratory patterns identified earlier in section 5.1.

**Table 2: Logistic Regression Results Predicting Follow-on Funding Success**

Variable	Pre-2014 Coefficient	Pre-2014 Odds Ratio	Post-2014 Coefficient	Post-2014 Odds Ratio
First Round Very Small	+0.789	2.200	+0.089	1.094
First Round Very Large	0.000	1.000	+0.274	1.315
Log First Round Amount	-0.590	0.555	-0.808	0.446
Log Time to First Funding	+0.170	1.185	-0.292	0.747
Early Founding	-0.060	0.942	+0.978	2.660

Footnote: The 'First Round Very Large' variable was included in models for both periods but eliminated by L1 regularization in the pre-2014 model.

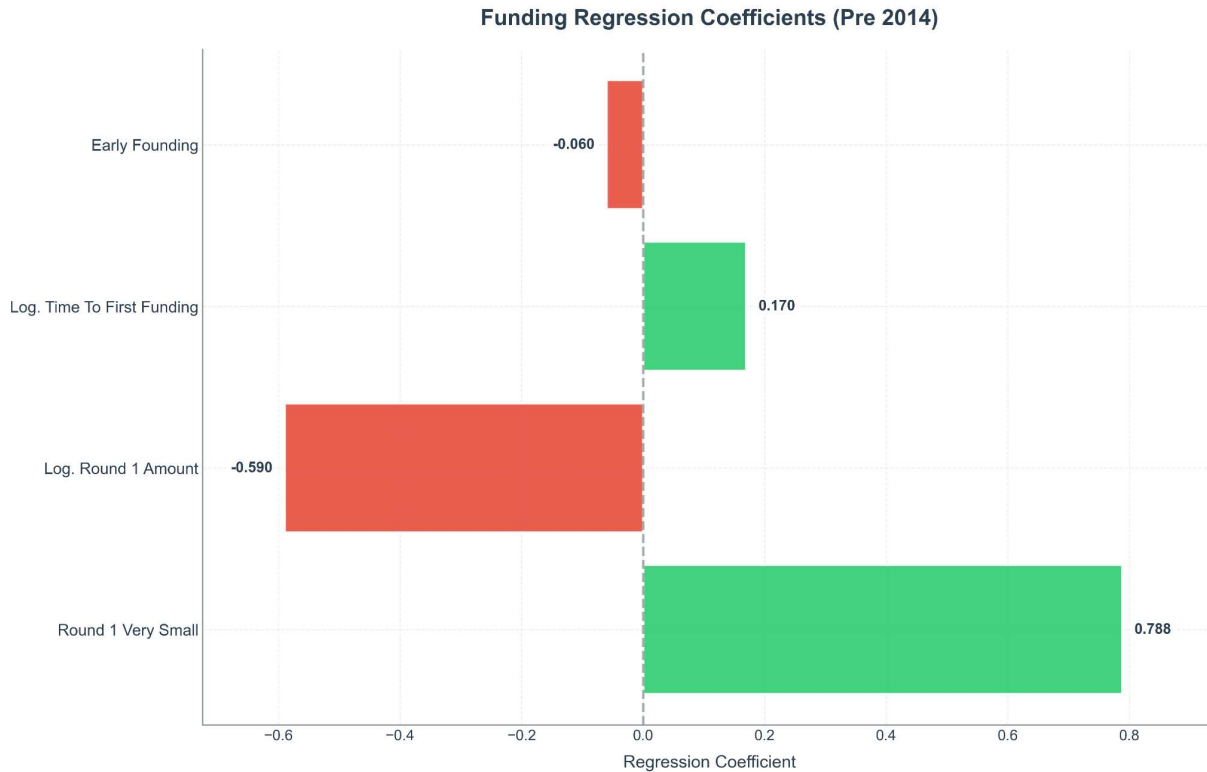


**Figure 1: Factors Influencing Follow-on Funding (Pre 2014)**

First round size indicators emerged as the most influential predictors, with the relative magnitudes for the pre-2014 period visualized in Figure 1. Having a very small first round (bottom quintile) was positively associated with follow-on funding probability (coefficient +0.789, odds ratio 2.200,  $p < 0.001$ ), indicating that startups with minimal initial funding were significantly more likely to raise subsequent rounds. Conversely, log-transformed first round amount showed a negative association (coefficient -0.590, odds ratio 0.555), suggesting that larger initial rounds generally reduced the likelihood of subsequent funding. Notably, while having a very large first round (top quintile) did not emerge as a significant predictor in the pre-2014 period, it would later become important in the post-2014 era, suggesting an evolution in funding dynamics.

Initial analysis of time to first funding showed a surprising positive association with follow-on funding probability (coefficient +0.170, odds ratio 1.185). This result appears counterintuitive, as our exploratory data analysis (see Figure 6, Section 5.3.2) clearly demonstrates declining follow-on rates as time to first funding increases, from 94.6% for "Very Fast" startups to 73.7% for "Very Slow" startups. Further investigation revealed this contradiction stems from a statistical phenomenon known as 'perfect separation' in the smallest funding quintile (Q1). In this quintile, nearly all startups received follow-on funding regardless of their time to first funding, creating an extreme positive coefficient (+4.162) that distorted the aggregate model results.

When analyzing each funding quintile separately, it was found that all other quintiles (Q2-Q5) showed significant negative coefficients (ranging from -0.217 to -0.745), confirming that faster funding generally predicted higher follow-on success. This finding was further corroborated by alternative approaches, including categorical time variables and a modified model excluding Q1, which yielded a strongly negative coefficient (-0.361, odds ratio 0.697). To account for potential temporal effects within each cohort period, a binary 'early founding' variable was created, identifying startups founded before their cohort's median year versus those founded during or after it.

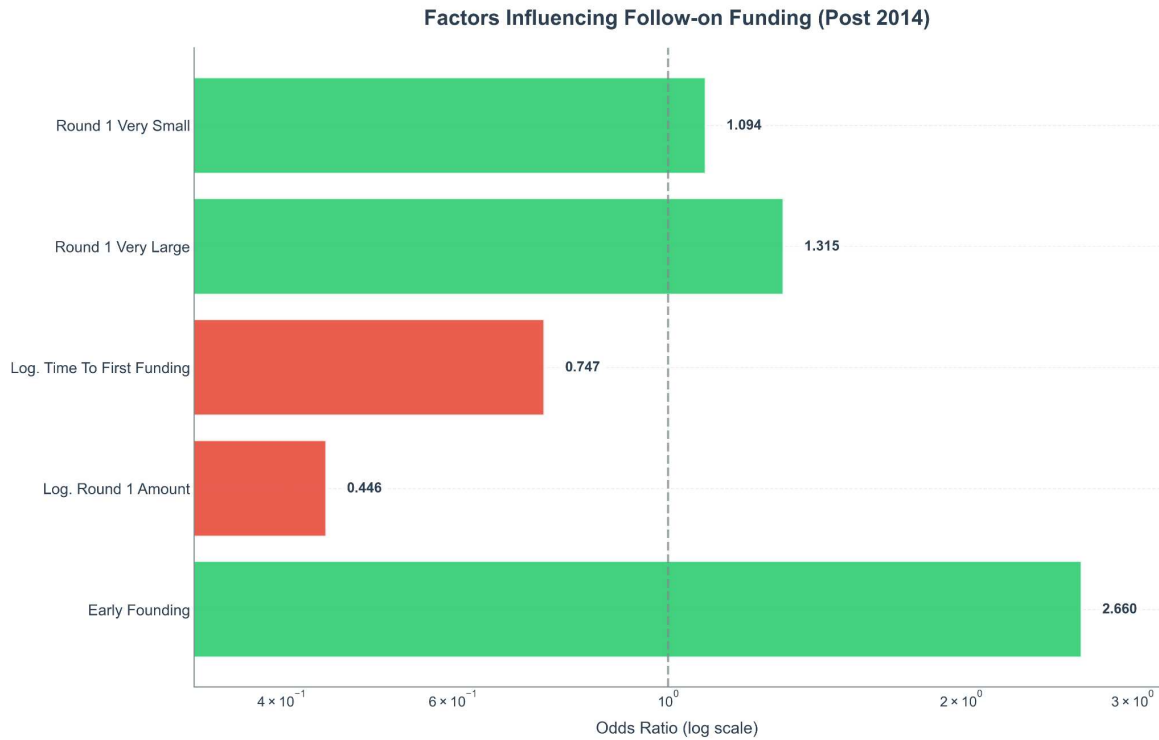


**Figure 2: Funding Regression Coefficients (Pre 2014)**

Figure 2 presents the same logistic regression results in coefficient form rather than as odds ratios. The coefficient visualization highlights the magnitude and direction of each predictor's effect, with positive coefficients (green bars) indicating factors that increase follow-on probability and negative coefficients (red bars) indicating factors that decrease it. The coefficient values directly show that having a very small first round (+0.788) and longer time to first funding (+0.170) were positively associated with follow-on success, while larger first round amounts (-0.590) and earlier founding (-0.060) negatively impacted follow-on probability.

### **Post-2014 Cohort**

The logistic regression for the post-2014 cohort demonstrated stronger predictive performance compared to the pre-2014 period, with an AUC of 0.739. The pattern of significant predictors showed both similarities and notable differences compared to the pre-2014 period. Early founding within the post-2014 period emerged as the strongest positive predictor (coefficient +0.978, odds ratio 2.660), indicating that startups established in the earlier years of this period had substantially higher odds of securing follow-on funding.



**Figure 3: Factors Influencing Follow-on Funding (Post 2014)**

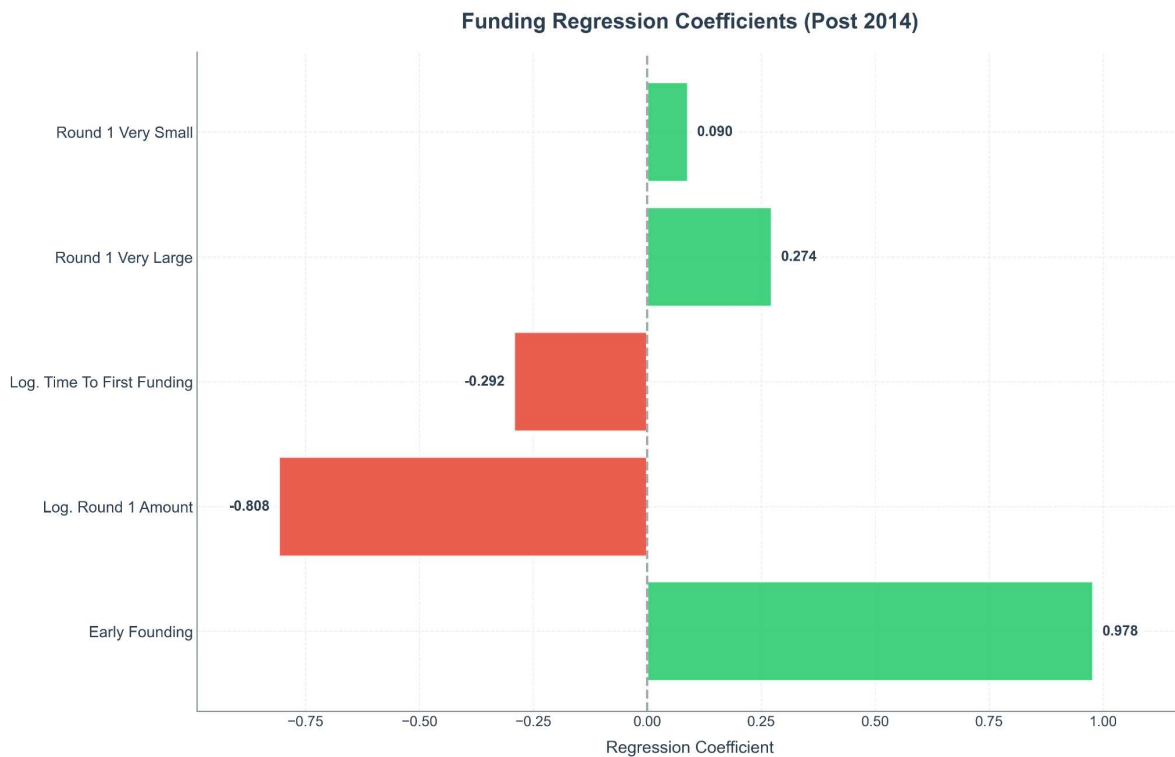
Early founding within the post-2014 period emerged as the strongest positive predictor (coefficient +0.978, odds ratio 2.660), indicating that startups established in the earlier years of this period had substantially higher odds of securing follow-on funding. This finding may reflect both the additional time available for these companies to secure additional funding and potentially more favorable market conditions in the earlier part of the period.

As with the pre-2014 period, first round amount showed a negative association with follow-on funding likelihood (coefficient -0.808, odds ratio 0.446). This relationship was even stronger than in the earlier period, suggesting an intensified pattern whereby larger initial rounds reduced the need for or likelihood of subsequent financing.

Unlike the pre-2014 period, time to first funding demonstrated a negative association with follow-on funding probability (coefficient -0.292, odds ratio 0.747), indicating that in the more recent period, faster initial funding was associated with follow-on success. This reversal represents a significant shift in the temporal dynamics of AI funding between the two periods.

First round size extremes showed somewhat different patterns compared to the pre-2014 period. While having a very small first round was still positively associated with follow-on funding (odds ratio 1.094), the effect was much weaker. Very large first rounds showed a positive association (odds ratio 1.315), in contrast to the general negative relationship between first round size and follow-on probability. This suggests a potential U-shaped relationship in the more recent period, with both very small and very large rounds associated with higher follow-on rates compared to mid-sized rounds.

Figure 4 displays the same findings as regression coefficients rather than odds ratios. The coefficient visualization more directly shows the magnitudes of effects, highlighting the strong positive influence of early founding (+0.978) and the substantial negative effect of log first round amount (-0.808). The coefficient visualization also clearly illustrates the reversal in the time to first funding variable, with the post-2014 coefficient being negative (-0.292) compared to the positive coefficient in the pre-2014 period, supporting the finding that faster initial funding became associated with higher follow-on success in the more recent period.



**Figure 4: Funding Regression Coefficients (Post 2014)**

### 5.3.2) Evolutionary Patterns in Follow-on Funding Success Factors

Comparing the results across time periods reveals several key evolutionary patterns in the determinants of follow-on funding success, clearly visible in Figures 1 through 4. First, the pattern of association between first round amount and follow-on funding probability remained consistently negative across both periods, though with a stronger effect in the post-2014 era (odds ratio decreasing from 0.555 to 0.446). This intensified negative relationship suggests an evolution in funding strategies, with larger initial rounds increasingly designed to provide sufficient runway to reach significant milestones without the need for frequent follow-on financing.

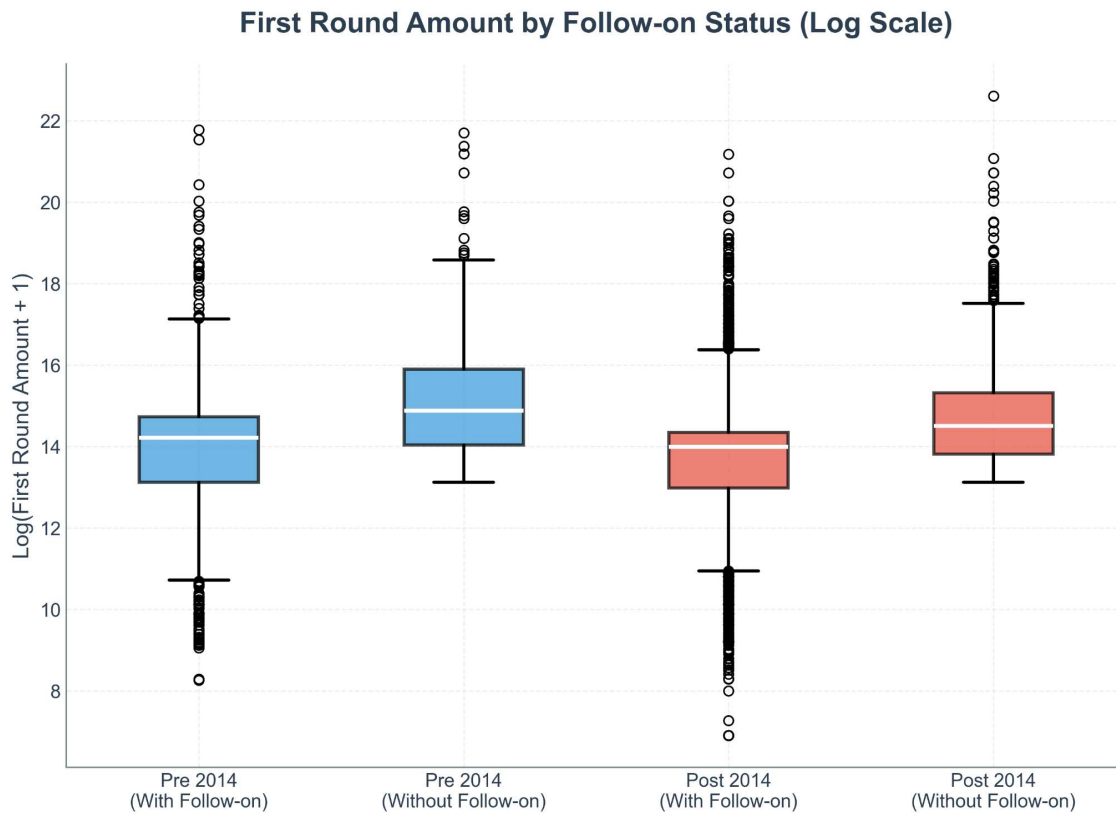
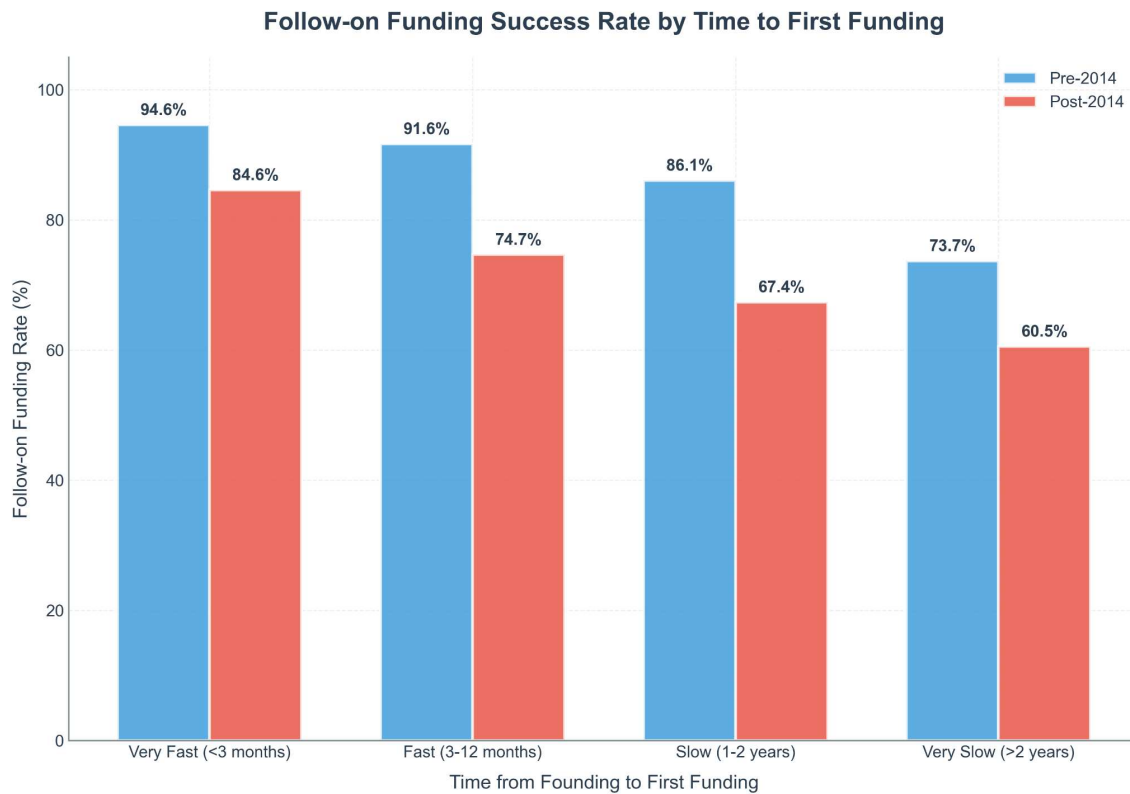


Figure 5: First Round Amount by Follow-on Status (Log Scale)

Figure 5 presents a comparative boxplot of first round funding amounts (on logarithmic scale) across both time periods, separated by follow-on funding status. This visualization confirms the counterintuitive relationship identified in the logistic regression analysis: in both

time periods, startups that did not secure follow-on funding showed higher median first round amounts compared to those that did.

The relationship between time to first funding and follow-on probability initially appears to show differences between the two periods, with the pre-2014 cohort showing a positive coefficient (+0.170, odds ratio 1.185) and the post-2014 cohort showing a negative coefficient (-0.292, odds ratio 0.747). However, Figure 6 illustrates that this apparent contradiction doesn't align with the raw data. The chart shows that startups receiving initial funding more quickly after founding demonstrated higher follow-on funding success rates in both time periods. For the pre-2014 cohort, follow-on rates declined from 94.6% for the "Very Fast" category to 73.7% for the "Very Slow" category. Similarly, in the post-2014 cohort, rates decline from 84.6% to 60.5% across these categories. Chi-squared tests confirmed these associations were statistically significant, with  $\chi^2 = 138.0$  ( $p < 0.0001$ ,  $df = 3$ ) for the pre-2014 cohort and  $\chi^2 = 400.2$  ( $p < 0.0001$ ,  $df = 3$ ) for the post-2014 cohort.



**Figure 6: Follow-on Funding Success Rate by Time to First Funding**

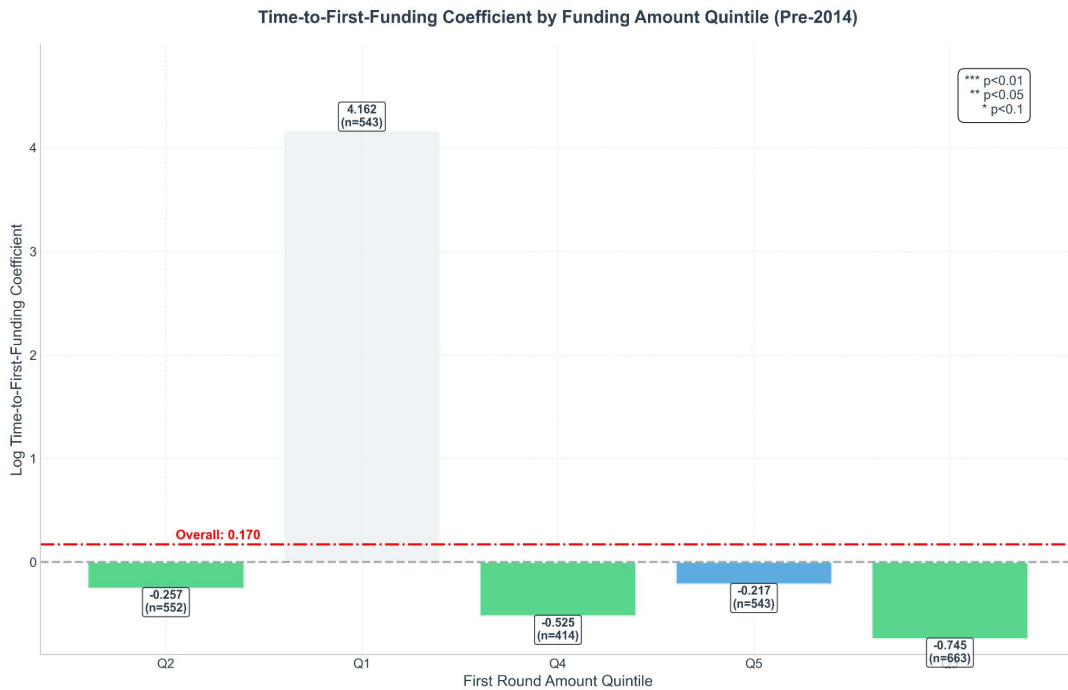
To investigate this contradiction, multiple complementary analyses were conducted. These included a modified logistic regression excluding Q1 (coefficient -0.513, odds ratio 0.598), categorical time variables (showing monotonically increasing negative effects: Fast -0.161, Slow -0.505, Very Slow -1.267), and quintile-specific models (showing negative coefficients for Q2-Q5). All these analyses demonstrated that faster funding predicted higher follow-on success in both periods. The primary evolutionary pattern was that this relationship became more consistent and uniform across all funding sizes in the post-2014 era, whereas the pre-2014 period showed greater heterogeneity and was statistically distorted by perfect separation in Q1.

To understand why our initial regression showed a positive coefficient for the pre-2014 period despite this clear negative relationship in the raw data, detailed statistical investigations were conducted. The quintile-specific analysis, shown in Figures 9 and 10, reveals the source of this apparent contradiction. The smallest funding quintile (Q1) exhibits perfect separation, a statistical condition in logistic regression where the outcome variable is perfectly predicted by a combination of predictors. In our case, nearly all startups in Q1 received follow-on funding regardless of their time to first funding, creating an extreme positive coefficient (+4.162) for Q1 that distorted the aggregate model results. In contrast, all other quintiles (Q2-Q5) show negative coefficients (ranging from -0.217 to -0.745), confirming that faster funding predicted higher follow-on success across most of the distribution.

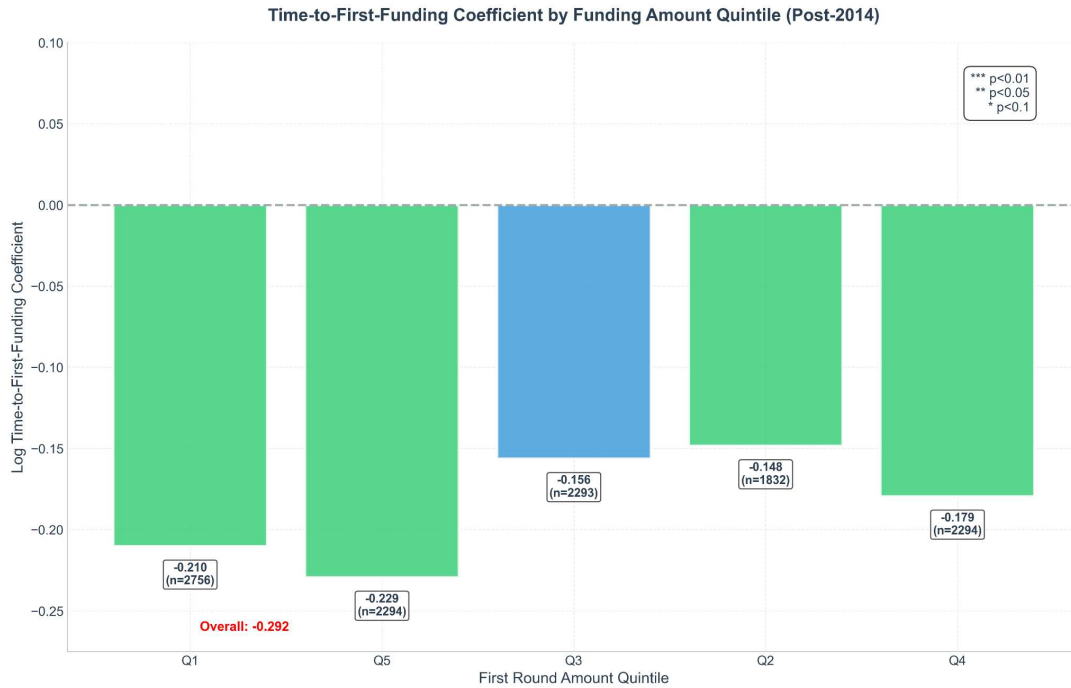
Two additional analyses confirm this explanation. First, when using categorical variables for time-to-first-funding instead of logarithmic transformation, a clear monotonic relationship emerged. Using "Very Fast" as the reference category, the coefficients became increasingly negative as funding speed decreased: Fast (-0.334,  $p = 0.609$ ), Slow (-0.821,  $p < 0.001$ ), and Very Slow (-1.455,  $p < 0.001$ ). While the Fast category coefficient wasn't statistically significant, the pronounced negative trend across categories and the strong significance for slower funding categories reinforces our findings. Additionally, when modeling the data excluding Q1, the coefficient for time-to-first-funding became strongly negative (-0.361, odds ratio 0.697), suggesting that longer time to funding is associated with decreased follow-on probability by approximately 30%. Model diagnostics confirmed improved fit, showing reduced AIC values and no evidence of specification errors in residual analysis.

These complementary analytical approaches collectively indicate that the positive coefficient in our initial aggregate model was an artifact of the perfect separation in Q1 rather than reflecting the predominant pattern across the dataset. The initial regression simultaneously controlled for first round size (-0.590 coefficient) and very small first round status (+0.789 coefficient), which, combined with the perfect separation in Q1, created the misleading positive coefficient for time-to-first-funding in the pre-2014 period.

Figures 7 and 8 provide a visual comparison of this evolutionary pattern. Figure 7 illustrates the pre-2014 heterogeneity with the extreme positive coefficient for Q1 (+4.162) contrasting sharply with negative coefficients for Q2-Q5. Figure 8 shows the post-2014 standardization, with consistently negative coefficients across all quintiles, ranging from -0.148 to -0.229. This comparison highlights how the relationship between time to first funding and follow-on success became more uniformly negative and consistent in the post-2014 period, representing an evolution in how early market validation influences subsequent funding success.

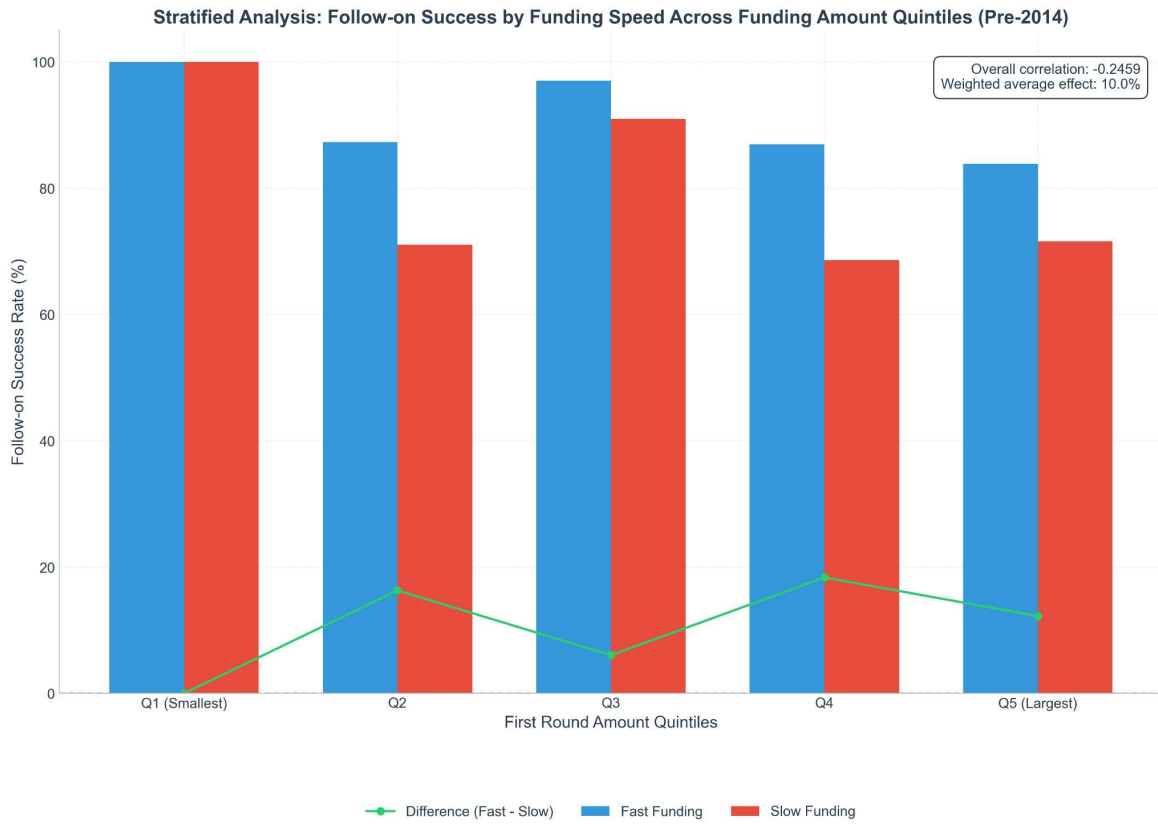


**Figure 7: Time to First Funding Coefficient by Funding Amount (Pre-2014)**

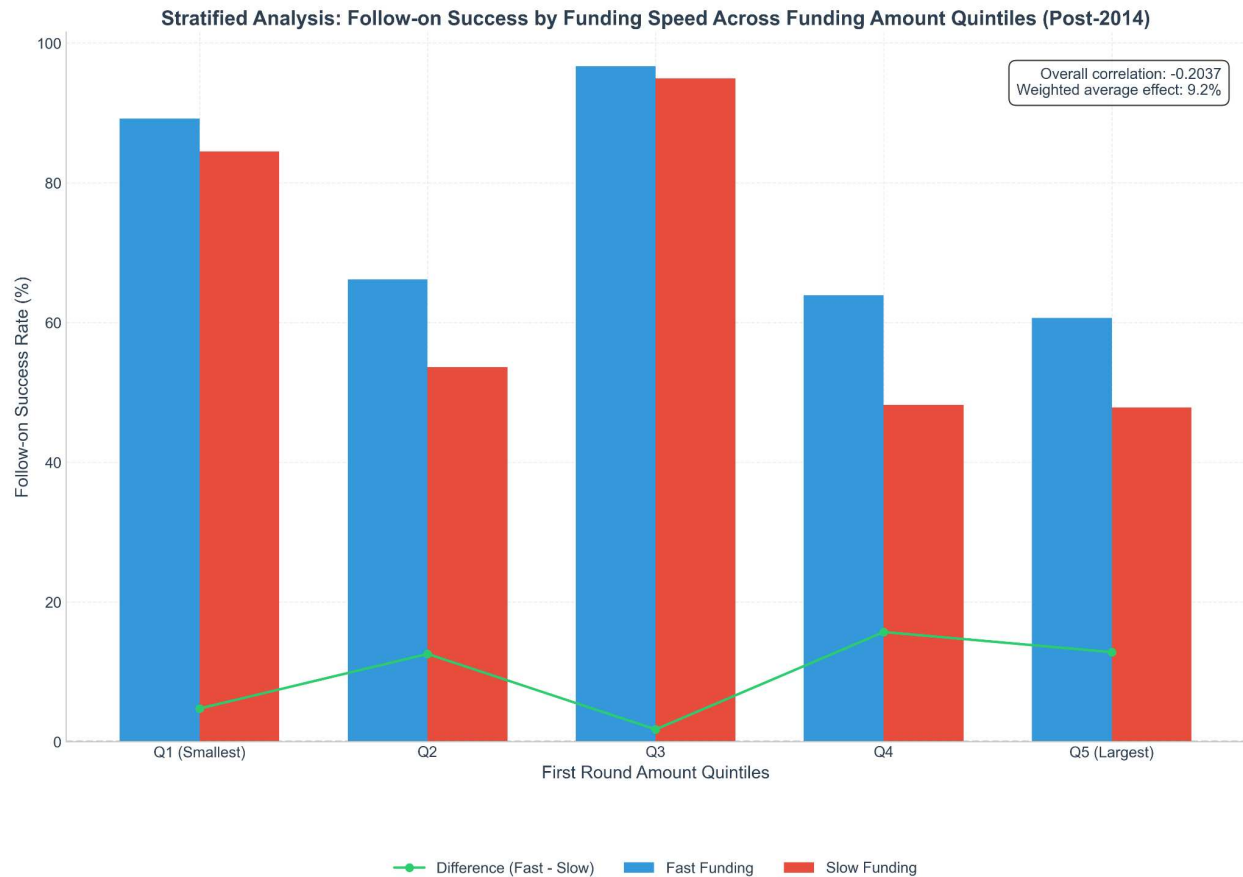


**Figure 8: Time to First Funding Coefficient by Funding Amount (Post-2014)**

For the post-2014 cohort, both the regression coefficient (-0.292, odds ratio 0.747) and the overall correlation (-0.2037) demonstrate a consistent negative relationship between time to first funding and follow-on success. This relationship remains uniform when stratifying by first round amount quintiles, with all quintiles showing negative correlations ranging from -0.0732 to -0.1863 (Figure 10). Startups with faster-than-median time to first funding consistently outperformed their slower counterparts across all funding size categories, with a weighted average effect of 9.2 percentage points higher follow-on rates.



**Figure 9: Stratified Analysis Across Funding Amount Quintiles (Pre-2014)**

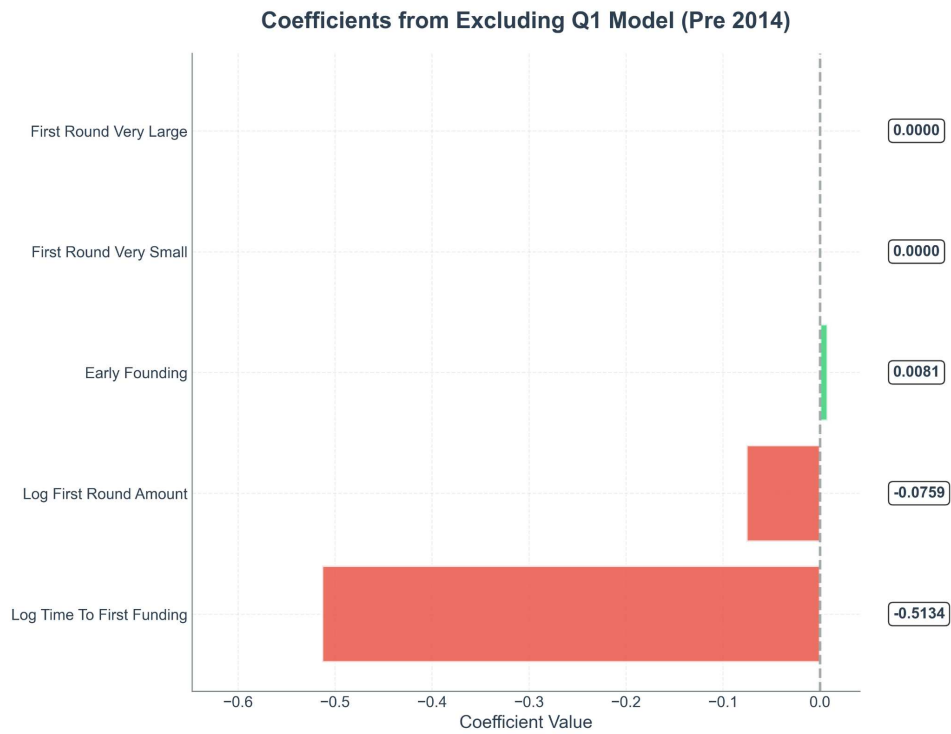


**Figure 10: Stratified Analysis Across Funding Amount Quintiles (Post-2014)**

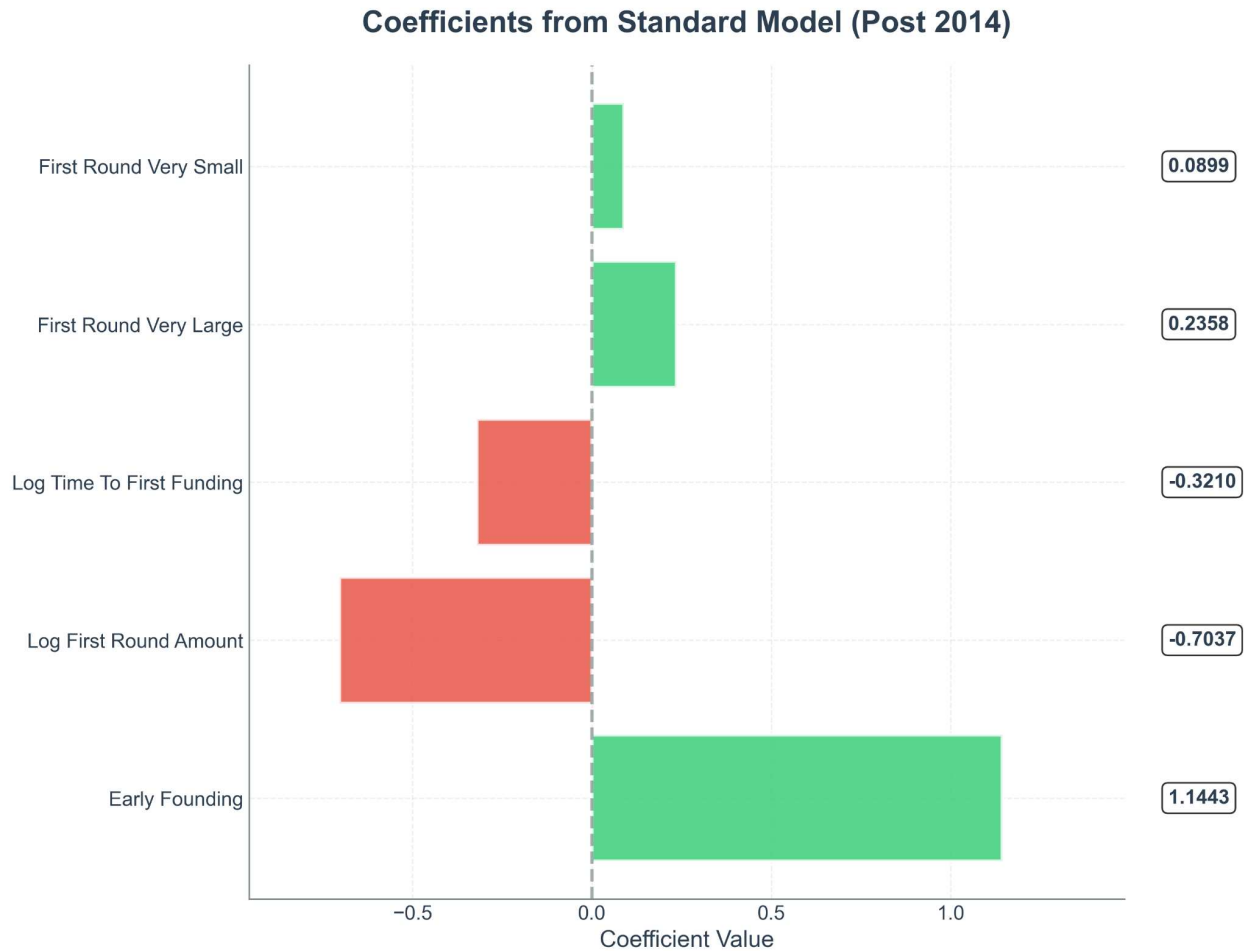
Figures 11-13 provide additional statistical evidence confirming our findings. Figure 11 illustrates the consistently negative time-to-first-funding coefficients across all funding quintiles in the post-2014 period. Figure 12 displays the coefficients from our modified logistic regression model excluding Q1 for the pre-2014 period. In this specification, the coefficient for time-to-first-funding ( $-0.513$ ,  $p < 0.001$ ) emerges as the most influential factor in the model, with an odds ratio of 0.598, indicating that longer time to funding decreases follow-on probability by approximately 40%. Figure 13 further supports this by showing the standard model coefficients for the post-2014 period. For future research addressing similar perfect separation challenges, Firth's penalized likelihood approach as implemented by Heinze and Schemper could provide a more elegant statistical solution than our exclusion approach. This method specifically corrects for the small-sample bias in maximum likelihood estimation while producing finite parameter estimates even in cases of separation, offering methodological advantages over the variable exclusion strategy employed in this study.



**Figure 11: Time to First Funding by Funding Amount Quintile (Post 2014)**



**Figure 12: Coefficients from Excluding Q1 Model (Pre 2014)**

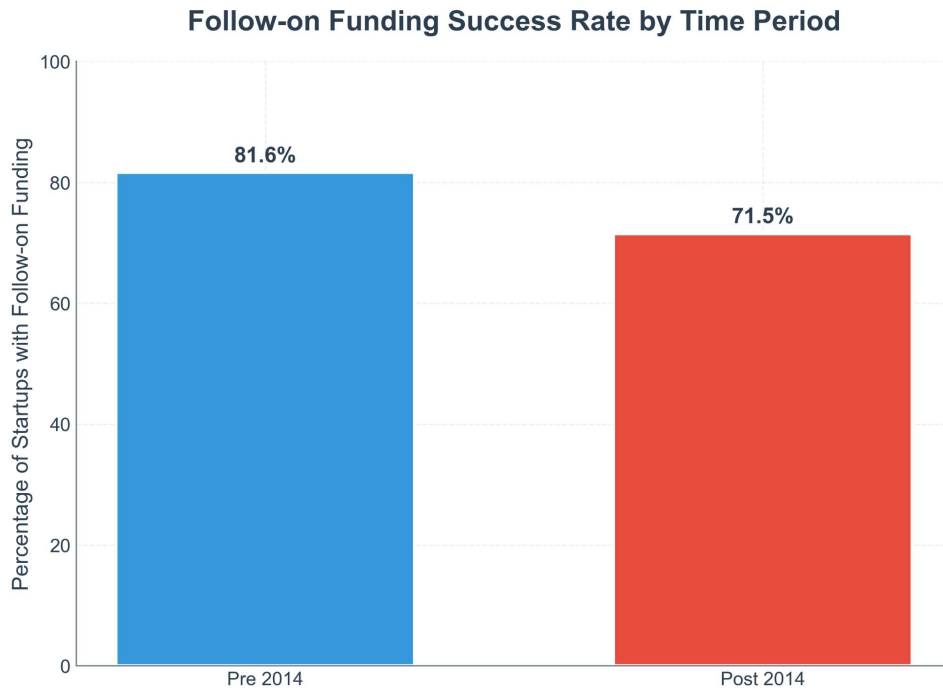


**Figure 13: Coefficients from Standard Model (Post 2014)**

Figure 13 provides the standard model coefficients for the post-2014 period, showing how the time-to-first-funding coefficient (-0.3210) becomes clearly negative in the aggregate model without requiring the exclusion of Q1. This contrasts with the pre-2014 period, where such exclusion was necessary to reveal the negative relationship. The post-2014 standardization across quintiles (with coefficients ranging from -0.148 to -0.229) represents an important evolution in funding dynamics, signaling more consistent investor evaluation criteria as the AI sector matured.

In summary, our comprehensive analysis employing multiple methodological approaches reveals that faster time-to-first-funding consistently predicted higher follow-on success across both periods, though with important evolutionary differences. The key transformation was in

consistency and standardization. In the pre-2014 period, this relationship showed heterogeneity across funding sizes and was significantly distorted by the statistical anomaly in Q1. This heterogeneity largely disappeared in the post-2014 period, where the relationship became more uniformly negative across all quintiles (ranging from -0.148 to -0.229).



**Figure 14: Follow on Funding Success Rate by Time Period**

This standardization suggests that as the AI sector matured, the importance of early market validation intensified, and investor evaluation criteria became more consistent across the funding spectrum. Simultaneously, the role of founding timing evolved substantially, with early founding becoming a much stronger positive predictor in the post-2014 cohort (odds ratio 2.660 vs. 0.942).

Figure 14 visualizes the overall decline in follow-on funding success rates identified in Section 5.1.1, decreasing from 81.6% to 71.5%. This 10 percentage point decrease represents a broader shift toward investor selectivity as the AI sector matured and expanded, providing important context for interpreting the regression results and evolutionary patterns identified above.

### **5.3.3) Enhanced Logistic Regression Analysis with Progressive Controls**

To examine predictor robustness, two models were estimated per time period: base model (fundamental predictors) and company controls (adding operational status and founding timing). Pre-2014 results showed minimal performance gains across specifications (AUC: 0.695 base to 0.702 with company controls), indicating fundamental predictors capture most variance. The models showed no significant multicollinearity issues ( $VIF < 5$  for all variables) and maintained consistent coefficient directions across specifications.

Post-2014 results differed markedly. Company controls improved performance by 5.2% (AUC: 0.726 to 0.764). Cross-validation confirmed these performance differences were statistically significant ( $p < 0.01$ ). This stark difference in model improvement patterns, minimal in pre-2014 versus substantial in post-2014, provides statistical evidence that company-specific operational characteristics have become significantly more important determinants of follow-on funding in the maturing AI sector.

#### **Alternative Success Metrics Analysis**

Beyond the binary follow-on funding outcome, three alternative success metrics were examined to provide more nuanced insights into funding dynamics. These metrics include whether startups secured more than two funding rounds, whether follow-on funding exceeded the median amount, and whether follow-on rounds occurred faster than the median interval.

Analysis of these alternative metrics revealed consistent patterns with the primary analysis while highlighting important nuances. For the multiple funding rounds outcome (pre-2014 AUC: 0.636, post-2014 AUC: 0.713), early founding emerged as a particularly strong predictor in the post-2014 period (odds ratio 2.734), suggesting temporal advantages compound for companies securing multiple rounds. The analysis of above-median follow-on funding demonstrated that first round size maintained a strong positive association with subsequent funding magnitude in both periods, with very large first rounds showing particularly strong effects in the post-2014 era (odds ratio 2.402).

Notably, the analysis of faster-than-median follow-on revealed that early founding showed opposite effects on speed of follow-on funding between periods. While early founding

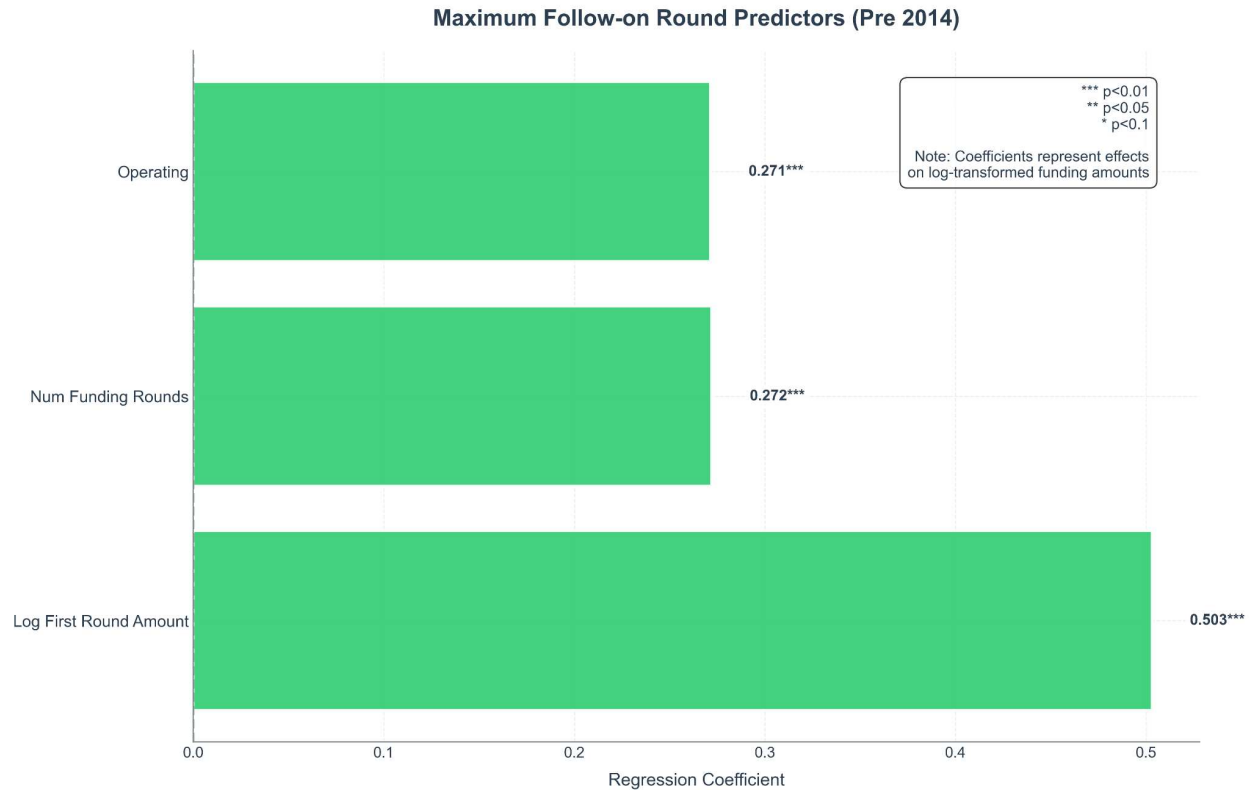
was associated with faster follow-on in the pre-2014 cohort (odds ratio 0.626), it predicted slower follow-on in the post-2014 period (odds ratio 0.517). This reversal may reflect changing market dynamics where early entrants in the pre-2014 period faced less competition and could secure rapid follow-on, while early entrants in the more mature post-2014 market prioritized achieving substantial milestones before subsequent rounds.

#### **5.3.4) Characteristics Associated with Follow-on Funding Amounts**

##### **Pre-2014 Cohort**

The analysis examining factors associated with follow-on funding amounts for the pre-2014 cohort employed both correlational and multiple regression approaches to understand the determinants of subsequent investment magnitudes. For startups that secured follow-on funding, the Pearson correlation between first round amount and total follow-on funding was  $r = 0.340$  (95% *CI* [0.307, 0.372],  $p < 0.0001$ ), while the Spearman rank correlation was  $\rho = 0.377$  (95% *CI* [0.345, 0.408],  $p < 0.0001$ ). These significant positive correlations indicate that startups raising larger initial rounds subsequently secured larger follow-on investments. The slightly stronger Spearman correlation compared to the Pearson coefficient suggests some non-linearity in the relationship, with rank order being more strongly preserved than exact proportionality between the variables.

The multiple regression analysis yielded more comprehensive insights into the determinants of follow-on funding amounts. For the pre-2014 cohort analyzing total follow-on funding, log-transformed first round amount ( $\beta = 0.518$ ,  $p < 0.001$ ) and number of funding rounds ( $\beta = 0.364$ ,  $p < 0.001$ ) emerged as the strongest predictors. The operating status of the startup also showed a significant positive effect ( $\beta = 0.272$ ,  $p < 0.001$ ), suggesting that actively operating companies secure substantially higher follow-on investments.



**Figure 15: Total Follow-on Funding Amount Predictors (Pre 2014)**

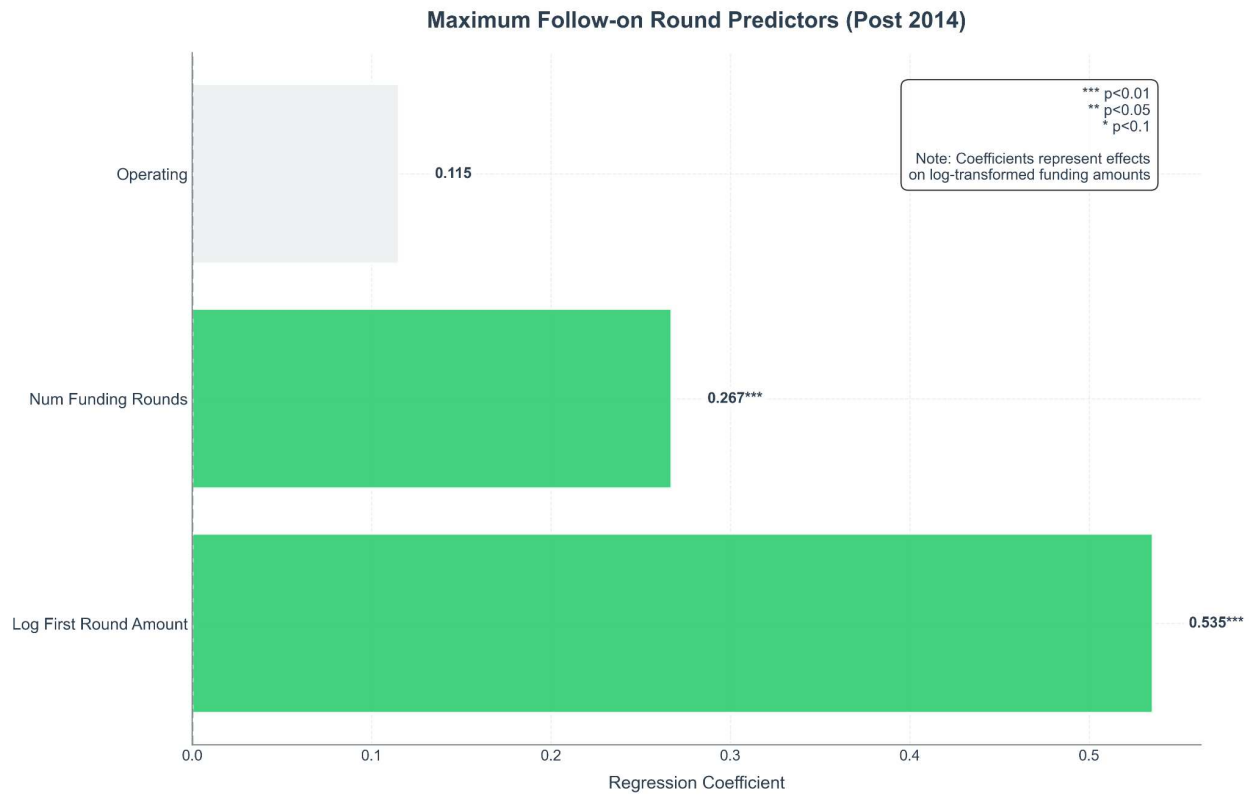
The analysis of maximum follow-on round size revealed parallel patterns with slightly lower explanatory power. Log-transformed first round amount maintained its significance as the strongest predictor ( $\beta = 0.535, p < 0.001$ ), indicating that larger initial investments consistently predicted larger peak follow-on rounds. The number of funding rounds showed a substantial effect on maximum round size ( $\beta = 0.268, p < 0.001$ ), though expectedly lower than its impact on total follow-on funding as maximum size is less mechanically tied to round count.

### Post-2014 Cohort

For the post-2014 cohort, both correlational and regression analyses revealed stronger relationships than in the earlier period. The Pearson correlation between first round amount and total follow-on funding increased to  $r = 0.398 (p < 0.0001)$ , while the Spearman correlation rose to  $\rho = 0.438 (p < 0.0001)$ . This enhancement in correlation strength

suggests an intensification of the signaling effect of initial funding in the more recent AI funding environment.

The regression analysis demonstrated even stronger explanatory power, with log-transformed first round amount showing an increased coefficient ( $\beta = 0.550, p < 0.001$ ) compared to the pre-2014 cohort. Number of funding rounds maintained a strong effect ( $\beta = 0.368, p < 0.001$ ), while operating status showed a reduced but still significant impact ( $\beta = 0.121, p < 0.05$ ).



**Figure 16: Total Follow-on Funding Amount Predictors (Post 2014)**

The analysis of maximum follow-on round size for the post-2014 cohort revealed similarly strengthened relationships. Log-transformed first round amount showed an enhanced effect ( $\beta = 0.535, p < 0.001$ ) compared to the pre-2014 cohort, reinforcing that initial funding signals have become more powerful predictors of subsequent investment success.

The number of funding rounds ( $\beta = 0.267, p < 0.001$ ) and operating status ( $\beta = 0.115, p < 0.05$ ) also maintained significant effects on maximum follow-on round size.

### **5.3.5) Evolution in Follow-on Funding Amount Determinants**

The strengthening relationships in both correlational and regression analyses indicate that first round size has become a more reliable predictor of subsequent funding success as the AI sector has developed. The approximately 17% increase in correlation strength and the enhancement in regression coefficients suggest that initial round size has become a more powerful signal of startup quality to subsequent investors. This phenomenon aligns with Svetek's systematic review of signaling in early-stage financing, which documents how prior investments serve as increasingly important quality signals to subsequent investors in technology-intensive sectors (Svetek 87). The observed strengthening correlations provide empirical support for Svetek's theoretical framework on how signal interpretation evolves as markets mature.

Analysis of maximum follow-on round size revealed parallel patterns with slightly lower explanatory power, confirming that factors predicting total follow-on funding also influence peak investment amounts. The consistent strengthening of coefficients between periods across both models indicates intensifying relationships between early funding characteristics and subsequent investment outcomes as the AI sector has matured.

This strengthening relationship indicates increasing investor coordination and information flow in the maturing AI sector. As the field has developed, initial round size has become a more reliable signal of startup quality, creating stronger path dependencies in funding trajectories. This pattern may reflect more standardized venture capital evaluation criteria or increased investor sophistication in the AI sector, leading to greater consistency in funding decisions across sequential rounds.

### **5.3.6) First-to-Second Round Dynamics Analysis**

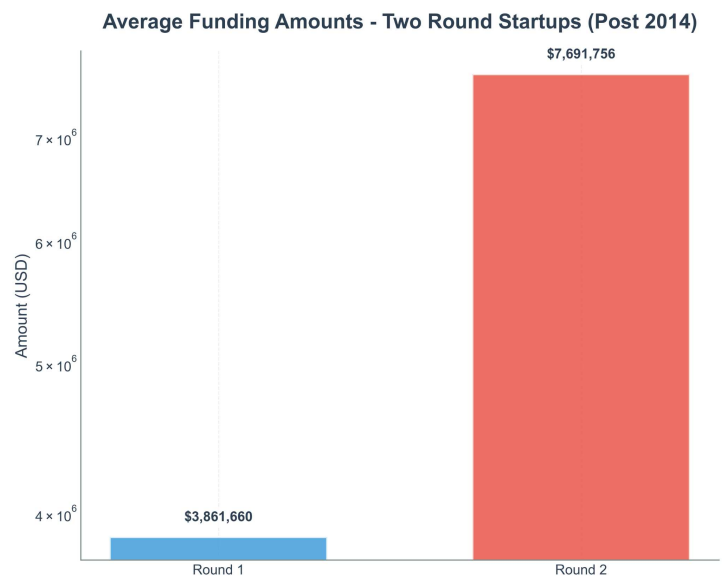
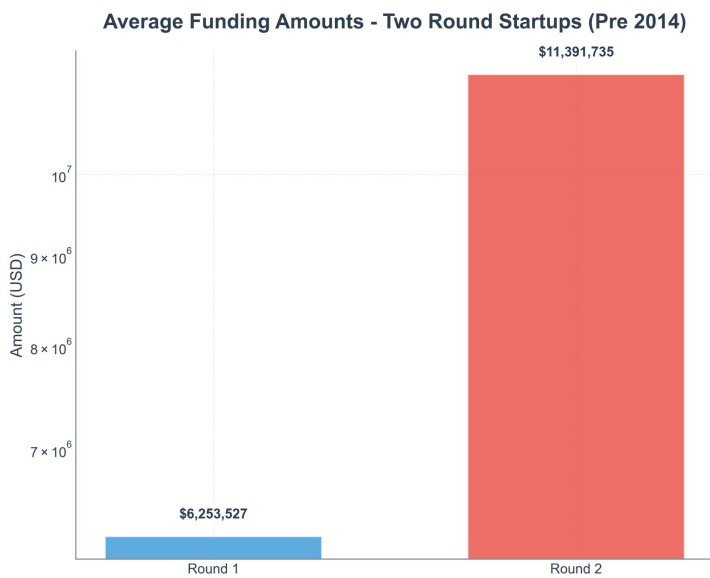
To examine early-stage funding dynamics in greater detail, a focused analysis was conducted of startups with exactly two funding rounds. This subset analysis provides insights into the factors influencing both the magnitude and timing of initial follow-on funding, offering a

cleaner examination of early funding trajectories by controlling for the complexity introduced by multiple subsequent rounds.

The dataset included 649 pre-2014 startups and 3,029 post-2014 startups with exactly two funding rounds. Regression analysis revealed that first round amount strongly predicted second round magnitude across both periods, with elasticities suggesting a 1% increase in first round funding associated with 0.337% (pre-2014) and 0.385% (post-2014) increases in second round amounts. The strengthening of this relationship in the post-2014 period further supports the intensification of signaling effects identified in the broader analysis.

**Figure 17: Average Funding - Two Round Startups (Pre 2014)**

**Figure 18: Average Funding - Two Round Startups (Post 2014)**



Time to first funding demonstrated contrasting effects on the interval to second round between periods. In the pre-2014 cohort, longer time to first funding predicted longer intervals to second round (coefficient 0.125,  $p < 0.01$ ), while in the post-2014 period, this relationship became negligible (coefficient -0.003,  $p > 0.10$ ). This shift suggests that early validation has become decoupled from subsequent funding velocity as the AI sector has matured, potentially reflecting more standardized evaluation processes for follow-on rounds.

### **5.3.7) Time Between Funding Rounds**

#### **Pre-2014 Cohort**

The survival analysis examining factors influencing the time between funding rounds for the pre-2014 cohort yielded insights into the temporal dynamics of AI startup financing. The Kaplan-Meier estimation revealed a median time of 526 days between consecutive funding rounds, with considerable variation around this central tendency.

The Cox proportional hazards model successfully converged with both single-feature and multiple-feature specifications, indicating robust temporal patterns in the data. Addressing the third research question regarding factors influencing timing between rounds, the analysis found round number was negatively associated with the hazard rate (coefficient -0.11, hazard ratio 0.90), indicating that higher-numbered rounds (i.e., later in the funding sequence) typically experienced longer intervals before the subsequent round.

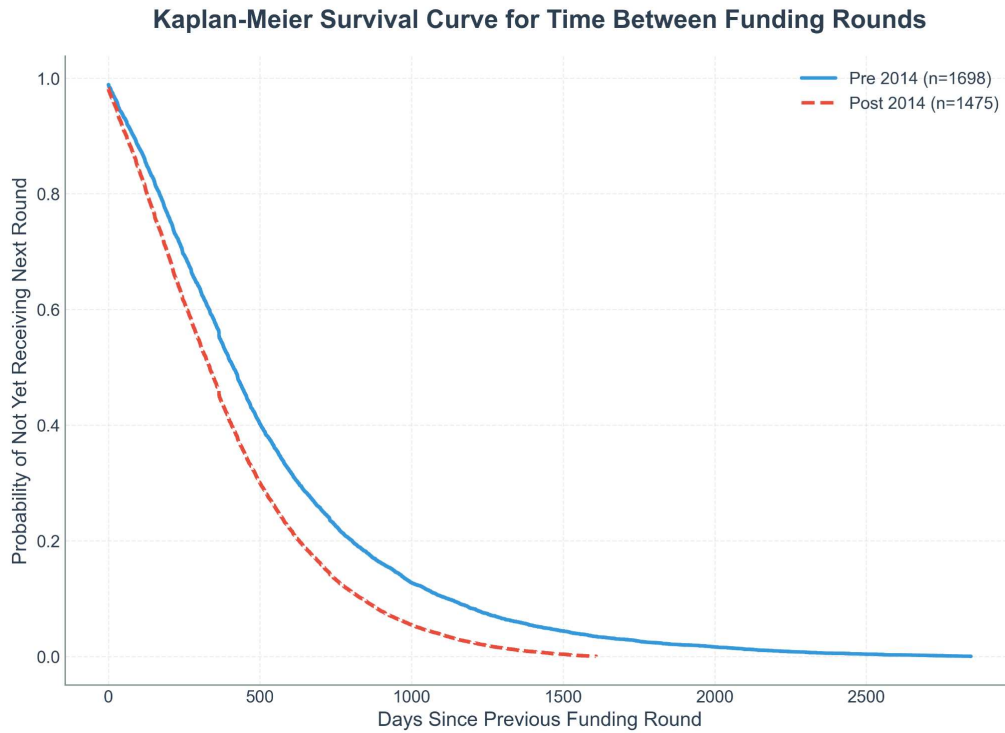
This finding reveals how the funding sequence itself shapes the temporal dynamics of AI venture capital, with earlier rounds proceeding more rapidly than later rounds, suggesting increasing due diligence and higher stakes as startups progress through their funding lifecycle. The log-transformed round amount also showed a significant negative association with the hazard rate (coefficient -0.15, hazard ratio 0.86), suggesting that larger rounds were followed by longer intervals before the next funding event. This finding aligns with the intuition that substantial funding reduces need for additional capital.

#### **Post-2014 Cohort**

For the post-2014 cohort, the Kaplan-Meier estimation revealed a shorter median time between rounds of 398 days, approximately 24% less than the pre-2014 period. This acceleration in funding cycles suggests increasing velocity in the more recent AI venture capital environment.

The Cox proportional hazards model again successfully converged for this period. As with the pre-2014 cohort, round number and round amount maintained their negative associations with the hazard rate, confirming the consistency of these temporal effects across both periods. The survival curve for the post-2014 period showed a steeper initial decline

compared to the pre-2014 curve, indicating that startups in the more recent period secured their next funding rounds more quickly on average, particularly in the early months following a round.



**Figure 19: Kaplan-Meier Survival Curve for Time Between Funding Rounds**

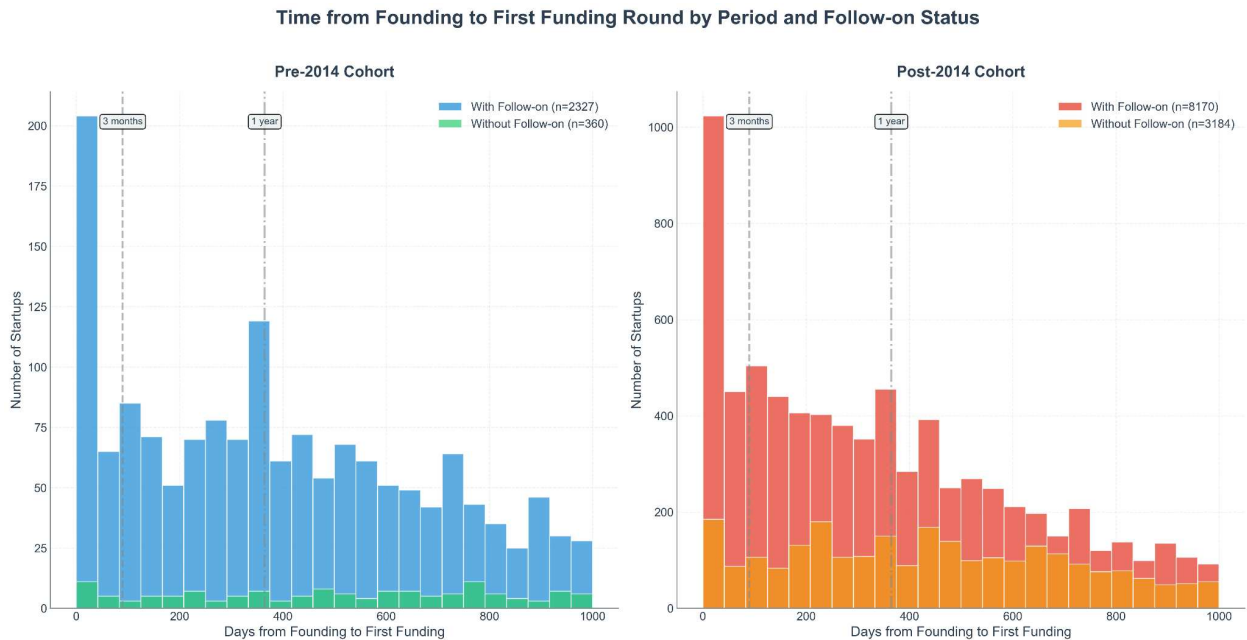
Figure 19 displays Kaplan-Meier survival curves for the time between funding rounds in both cohorts. In this visualization, the y-axis represents the probability of not yet receiving the next funding round, while the x-axis shows time since the previous round in days. The post-2014 curve (shown in orange) demonstrates a steeper initial decline compared to the pre-2014 curve (in blue), indicating that startups in the more recent period secured their next funding rounds more quickly on average.

The post-2014 curve consistently lies below the pre-2014 curve across most time points, visually confirming the finding that the time between funding rounds has decreased significantly in the more recent period.

### 5.3.8) Temporal Acceleration and Structural Consistency

The most striking evolution in AI funding dynamics revealed by the comparative survival analysis is the fundamental acceleration of the funding cycle. The 24% reduction in median time between consecutive rounds indicates significantly increased velocity in the AI venture capital ecosystem. This acceleration reflects increasing competition among investors and greater emphasis on rapid scaling in the maturing AI market. Despite this acceleration, the factors influencing the relative timing of funding rounds showed consistency across the periods. Both round number and round amount maintained their negative associations with hazard rates, indicating that fundamental aspects of venture capital staging patterns have remained stable despite the overall increase in velocity.

The survival curves visually confirmed this acceleration pattern, with the post-2014 curve showing systematically higher progression rates to subsequent rounds across most time points. This increased funding velocity represents one of the clearest indicators of the evolution in AI venture capital dynamics between the two periods.



**Figure 20: Time from Founding to First Funding by Period and Follow-on Status**

Figure 20 presents histograms showing the distribution of time from founding to first funding across both cohorts, with separate distributions for startups that did and did not secure

follow-on funding. The distributions show substantial concentration of startups in both cohorts that secured initial funding relatively quickly after founding, particularly within the first year. However, there are also long tails in all distributions, with some startups experiencing much longer intervals before securing their first funding.

The distributions for startups with follow-on funding show greater concentration at shorter time intervals compared to those without follow-on funding. This visual pattern supports the finding that faster initial funding is associated with higher follow-on funding probability, particularly in the post-2014 period.

### **5.3.9) Synthesized Evolutionary Patterns**

These evolutionary patterns align with the theoretical frameworks established in our literature review. The observed changes support Rogers' theory of technology diffusion, as the declining follow-on rates and increased selectivity reflect AI's movement from the "early adopter" to "early majority" phase, where investment criteria become more stringent and market-oriented. The strengthening correlation between initial round size and subsequent funding validates signaling theory predictions, demonstrating how initial funding amounts became more reliable quality signals as the AI market matured, creating stronger path dependencies in funding trajectories.

Integrating the findings across all three analytical approaches reveals several overarching evolutionary patterns in AI startup funding dynamics as the sector has matured. First, increased selectivity has emerged, as evidenced by the declining follow-on funding rates between periods (see Section 5.1.1). This 10 percentage point decrease suggests more discriminating investment decisions in the more competitive recent landscape.

Second, the increasing importance of early market validation is evident in both the standardization of the relationship between time to first funding and follow-on success, and the overall acceleration of funding cycles. While faster initial funding generally predicted higher follow-on success in both periods, this relationship became more uniform and consistent across all funding sizes in the post-2014 period. In the more competitive recent environment, speed to

market and rapid investor validation have become increasingly critical and standardized determinants of startup success.

Third, the intensified negative relationship between first round size and follow-on probability, alongside the strengthening positive correlation between first round and follow-on amounts, indicates evolving funding strategies. Larger initial rounds increasingly appear to function as substitutes for multiple smaller rounds, potentially reflecting both founder preferences for fewer dilutive events and investor strategies aimed at supporting startups through longer development cycles characteristic of modern AI technologies.

Fourth, the dramatic acceleration of funding cycles indicates fundamental changes in the speed of decision-making and competitive dynamics in AI venture capital. This increased velocity may reflect more standardized evaluation processes, greater investor competition for promising startups, and broader shifts in the technology development timeline for AI companies.

These evolutionary patterns collectively indicate a transformation in AI funding dynamics between the two periods, characterized by acceleration, increased signaling effects, and evolving funding strategies. These shifts reflect the maturation of artificial intelligence from an emerging technology to an established sector with more defined funding processes and expectations.

The enhanced analysis incorporating progressive controls, alternative success metrics, and focused examination of early-stage dynamics reinforces and extends the evolutionary patterns identified. The stability of fundamental predictors across model specifications indicates that core funding dynamics have remained consistent even as the importance of operational characteristics has increased. The divergent patterns observed in alternative success metrics, particularly the reversal in quick follow-on relationships, highlight the multifaceted nature of funding success in the maturing AI ecosystem. Together, these findings paint a comprehensive picture of an industry transitioning from exploratory investment patterns to more structured funding processes, with increasing emphasis on early validation, operational maturity, and strategic capital deployment.

### **5.3.10) Alternative Explanations for Observed Patterns**

While the analysis reveals clear evolutionary patterns in AI funding dynamics, alternative explanations merit consideration. The observed decline in follow-on funding rates may reflect broader macroeconomic conditions rather than AI-specific evolution, as the post-2014 period included significant monetary policy shifts, including the end of quantitative easing, which could have increased investor selectivity across all sectors. Additionally, survivorship bias may influence our findings, as failed startups from the pre-2014 era could be underrepresented in Crunchbase, potentially inflating success rates for that period, while the post-2014 cohort may include more early-stage startups that haven't yet had time to fail. The negative relationship between initial round size and follow-on probability might also reflect market saturation rather than evolving funding strategies, as the increasingly crowded AI sector post-2014 may have led investors to concentrate larger initial investments in fewer, high-potential startups. Finally, changes in Crunchbase's data collection methods over time could affect reporting consistency between periods, with the platform's increased coverage and visibility after 2014 potentially capturing more funding events and artificially creating the appearance of accelerated funding cycles.

## **6. Conclusions**

This research examined the evolution of venture capital funding dynamics for artificial intelligence startups across the pre-2014 and post-2014 periods. Through analysis of 15,608 AI startups and their 45,479 funding events, several significant shifts in funding patterns emerged as the AI sector matured. The empirical findings reveal three fundamental transformations in AI venture funding dynamics. First, increased investor selectivity emerged, reflected in declining follow-on funding rates. This suggests a more competitive landscape with more discriminating investors despite sector expansion.

Second, funding cycles accelerated substantially, with median time between rounds decreasing by 24% (from 526 to 398 days). This reflects heightened investor competition and industry emphasis on rapid scaling, while structural influences on funding intervals remained consistent.

Third, the analysis reveals an intensified importance of early market validation, demonstrated by the evolving relationship between time to first funding and follow-on success. Through multiple analytical approaches, including quintile-specific models, categorical analysis, and modified regression excluding Q1, it was demonstrated that faster initial funding predicted higher follow-on success in both periods. However, while this relationship showed significant heterogeneity across funding sizes in the pre-2014 period and was statistically distorted by perfect separation in Q1, it became more uniformly negative and significant across all funding quintiles in the post-2014 period. This standardization indicates increasing investor emphasis on early market validation and more consistent evaluation criteria as the AI sector matured.

Additional evolutionary patterns emerged from the comparative analysis. The negative association between initial round size and follow-on funding probability intensified in the post-2014 era, suggesting an evolution in funding strategies. Simultaneously, the signaling effect of initial round size strengthened considerably, with the correlation between first round amount and subsequent funding increasing by approximately 17%.

These findings have significant implications for multiple stakeholders. For entrepreneurs, the results emphasize the strategic importance of securing early validation and carefully calibrating initial funding rounds. For investors, the findings suggest increasing standardization in evaluation criteria and growing emphasis on early validation as a quality signal. For policymakers, the identified trends indicate the potential need for targeted support mechanisms for early-stage AI ventures, particularly those pursuing applications with longer development horizons.

The limitations of this research include potential reporting biases in the Crunchbase data, challenges in maintaining consistent classification of AI startups across the evolving definition of the field, and a focus on funding outcomes without addressing long-term company performance. Future research could examine whether similar patterns exist in other technology sectors, explore the mechanisms driving these shifts, and incorporate additional outcome measures beyond funding metrics.

These findings carry broader implications for the development of the AI industry. The increased selectivity and accelerated funding cycles suggest a shift from exploratory to

efficiency-driven innovation in AI, potentially leading to concentration of resources in fewer, more conservative AI applications, reduced funding for fundamental AI research requiring longer development horizons, and barriers to entry for novel AI approaches that don't fit established patterns. The evolving funding patterns indicate AI's transition from an emerging to a maturing technology sector, which may result in increased industry consolidation as larger rounds favor scale over diversity, emergence of AI-specific investment criteria distinct from general technology investing, and potential reduction in breakthrough innovations as risk appetite decreases. From a policy perspective, these trends suggest the need for alternative funding mechanisms for long-term AI research, the importance of maintaining competitive dynamics in an increasingly selective funding environment, and a potential role for public funding to support AI applications with longer development cycles but high societal value.

As artificial intelligence continues to advance as both a technology and an industry, understanding the evolving dynamics of startup funding in this sector remains critical. The patterns revealed in this analysis demonstrate how venture capital has adapted to the maturing AI landscape, with increased selectivity, accelerated cycles, and intensified emphasis on early validation emerging as defining characteristics of the contemporary AI funding environment.

## **7. Works Cited**

Gompers, Paul A. "Optimal Investment, Monitoring, and the Staging of Venture Capital." *The Journal of Finance*, vol. 50, no. 5, Dec. 1995, pp. 1461-1489. DOI: 10.2307/2329323.

Gompers, Paul, et al. "Venture Capital Investment Cycles: The Impact of Public Markets." *Journal of Financial Economics*, vol. 87, no. 1, 2008, pp. 1-23. DOI: 10.1016/j.jfineco.2006.12.002.

Gottfried, Oliver, et al. "Investigation of Venture Capital Firms' Criteria for Investing into Technology Startups: A Comparative Analysis of Chinese, US, and German Investors." SSRN, 26 June 2023. DOI: 10.2139/ssrn.4223114.

Heinze, Georg, and Michael Schemper. "A Solution to the Problem of Separation in Logistic Regression." *Statistics in Medicine*, vol. 21, no. 16, Aug. 2002, pp. 2409-2419. DOI: 10.1002/sim.1047.

Jacobides, Michael G., et al. "The Evolutionary Dynamics of the Artificial Intelligence Ecosystem." *Strategy Science*, vol. 6, no. 1, May 2021, pp. 412-435. DOI: 10.1287/stsc.2021.0148.

Perrone, C. "Drivers of Venture Capitalist's Investments in Artificial Intelligence." Master's thesis, Politecnico di Torino, 2020. Politecnico di Torino Digital Library, [webthesis.biblio.polito.it/id/eprint/16792](http://webthesis.biblio.polito.it/id/eprint/16792).

Rogers, Everett M. *Diffusion of Innovations*. 5th ed., Free Press, 2003.

Shao, Zhendong, et al. "Tracing the Evolution of AI in the Past Decade and Forecasting the Emerging Trends." *Expert Systems with Applications*, vol. 209, no. 10, July 2022, p. 118221. DOI: 10.1016/j.eswa.2022.118221.

Spence, A. Michael. "Signaling in Retrospect and the Informational Structure of Markets." *American Economic Review*, vol. 92, no. 3, June 2002, pp. 434-459. DOI: 10.1257/00028280260136200.

Svetek, Mojca. "Signaling in the Context of Early-Stage Equity Financing: Review and Directions." *Venture Capital*, vol. 24, no. 4, Apr. 2022, pp. 1-34. DOI: 10.1080/13691066.2022.2063092.

Torssell, J. "The Most Influential Team Attributes When Predicting Start-Up Success: A Quantitative Study of 25,430 European New Ventures." Master's thesis, KTH Royal Institute of Technology, 2022. DiVA Portal, [www.diva-portal.org/smash/get/diva2:1697289/FULLTEXT01.pdf](http://www.diva-portal.org/smash/get/diva2:1697289/FULLTEXT01.pdf).

Warzyńska, D. A., et al. "Dutch Artificial Intelligence Startups: A Case Study Analysis of Twenty-Four Dutch Artificial Intelligence New Ventures Characteristics and Financing." Tilburg University, 2020.

Weber, Michael, et al. "AI Startup Business Models." *Business & Information Systems Engineering*, vol. 64, 2022, pp. 91-109. DOI: 10.1007/s12599-021-00732-w.