ALMA MATER STUDIORUM — UNIVERSITY OF BOLOGNA

ECHO

Equity Capital Hype Oscillation

A Quantitative Analysis of Speculative Dynamics in the Artificial Intelligence Market

Final Dissertation presented by

Manuel Stella

Disclaimer:

This work is made available by the Author as an Open Access publication of the First Cycle Degree dissertation in Financial Analysis, submitted within the degree programme in Business and Economics at the University of Bologna "Alma Mater Studiorum". The work was supervised by Professor Lorenzo Dal Maso; it was defended by the Candidate on July the 22nd 2026, during the first graduation session of the 2025/2026 Academic Year. The present version has been slightly amended to comply with the editorial guidelines of the platform.

Contacts and Info:

Author's contact address: stella.manuel@libero.it
Author's ORCID account: https://orcid.org/0009-0005-5399-8972

Publication's DOI: https://doi.org/10.6092/unibo/amsacta/9058

Contents

1 – Introduction

"Nature uses only the longest threads to weave her patterns, so each small piece of her fabric reveals the organization of the entire tapestry."

Richard Feynman offered this image in 1964, in the *Messenger Lectures* at *Cornell University*. He was speaking about the *law of gravitation*, and of a property that delighted him: the rule that draws two stones together is the same rule that holds the planets in their orbits, identical across scales so vast that the small case must be magnified many millions of times to reach the large one. If a single law runs through every scale of nature, then a patient look at one small fragment can reveal the order of the whole.

This work borrows that habit of mind and applies it to a very different object: the market that has grown around *Artificial Intelligence*. Nothing here assumes that a market obeys laws in Feynman's sense; yet rather than value the *AI* sector only as one great mass, this work starts by examining the smallest pieces of its 'fabric', the individual financing rounds raised by *private companies* and the individual price paths of the listed firms, asking whether they carry a recognisable 'signature', repeated from one scale to the next, of a system organised by *speculation* rather than by *value*, approaching its *critical point*. If the signature is there, those fragments will reveal the whole tapestry.

1.1 Historical Overview of Tech Speculative Waves

Every few decades a new technology arrives carrying a promise large enough to rearrange the economy, and capital rushes towards it faster than anyone can work out what it is really worth. The shape of the episode is old and remarkably consistent: a genuine breakthrough supplies the premise, a story about the future grows bolder with each retelling, prices climb in a way that feeds on its own momentum, and a sharp reversal follows once the distance between price and substance becomes impossible to ignore.[1] The two most recent episodes, the *dotcom boom* of the late 1990s and the *cryptocurrency cycles* of the past decade, are worth recalling with some care, because they are the natural reference points against which the present enthusiasm for artificial intelligence will be judged.

The dotcom boom grew out of the arrival of the *World Wide Web*. As households and businesses came online through the second half of the 1990s, investors became convinced that almost any company with a foothold on the internet would inherit a slice of an enormous new market. Capital flowed into young firms with thin

[1] See, among others, Kindleberger and Aliber [35], Shiller [50], and Pástor and Veronesi [46].

revenue and no profits, and a company's share price could jump on little more than the addition of *".com"* to its name. The *Nasdaq Composite*, the index where most of these companies were listed, climbed more than fivefold between 1995 and its peak of 5,048 on 10 March 2000, reaching valuations far beyond anything the firms' earnings could support. What followed was as severe as the climb had been steep: the index lost about 78% of its value by October 2002, erasing some $5 trillion and taking down a long list of companies, from *Pets.com* to *Webvan*, that never found a way to make money. Studies of the period have since argued that the collapse was less a sudden change of heart than a mechanical unwinding, set off once the earliest investors were finally free to sell and did so in numbers.[2] Nor were the most sophisticated investors standing apart from the excitement; several rode the rising market on purpose, buying in and stepping out ahead of the fall rather than betting against it.[3]

The cryptocurrency episodes followed a similar script in a very different setting. *Bitcoin*, launched in 2009 as a form of money that no government issued, rose from a few hundred dollars in 2015 to around $20,000 by the end of 2017, then gave up roughly 80% of that value over the following year. The second wave was larger still. *Bitcoin* reached close to $69,000 in November 2021 and the combined value of all cryptocurrencies briefly touched $3 trillion, before the market shed about three quarters of that value over the next year. The fall was hastened by a chain of failures from inside *crypto* itself: the algorithmic *stablecoin TerraUSD* lost its dollar peg in May 2022 and erased tens of billions of dollars within days, and in November the large exchange *FTX* collapsed into bankruptcy. For all the wealth destroyed, the damage stayed largely within the sector, leaving the wider financial system relatively untouched.[4] A body of formal study has since confirmed what the price charts suggested, identifying clear episodes of explosive growth in cryptocurrency prices with the same family of statistical tools that the later chapters of this research apply to artificial intelligence.[5]

Seen side by side, the two episodes share a recognisable outline. In each a real innovation supplied the original premise, a compelling story pulled prices far beyond what current activity could justify, the rise drew in steadily wider groups of buyers, and the eventual fall was swift and costly. They part company on one point that matters for what follows. The dotcom bubble inflated and burst in plain sight on public exchanges, whereas the crypto cycles drew much of their force from a private and lightly regulated layer beneath the visible market. The boom in artificial intelligence carries both traits at once, with very large sums committed privately to

[2] Ofek and Richardson [45]. [3] Brunnermeier and Nagel [7]. [4] Bank for International Settlements [4]. [5] See Cheah and Fry [11], Cheung et al. [12], and Wheatley et al. [57].

young companies and a smaller set of listed firms absorbing the public enthusiasm, which is the reason the chapters that follow examine it from each side in turn.

1.2 The Expansion of the AI Startup Market

The wave that this research examines began with a single technology. In 2020 the firm *OpenAI* released *GPT-3*, a *large language model* that could write fluent text on almost any subject. Two years later, in November 2022, it launched *ChatGPT*, a free *chatbot* built on that technology, and the public response was immediate. Within three years generative AI had reached an estimated 53% of the world's population, spreading faster than either the personal computer or the internet before it.[6] A technology that ordinary people could use, and that businesses could see working, set off one of the largest investment booms in recent memory.

Most of that money went into private companies rather than onto public exchanges. By 2025 private investment in artificial intelligence had reached around $345 billion worldwide, more than double the figure for the year before, with *generative AI* alone accounting for nearly half of it.[7] Venture capital tells the same story from another angle: about half of all venture funding raised anywhere in the world in 2025 went to AI companies, up from roughly a third the year before.[8] For three years running, AI has been the single largest destination for startup capital, and by a wide margin.

The striking feature is not only the size of the flow but how few companies receive it. A small group of laboratories that build the underlying models has absorbed the greater part of the capital. By the end of 2025 *OpenAI* had become the most valuable private company ever, worth around $500 billion, while its closest rival *Anthropic* was valued at $183 billion. Together the two firms took about 14% of all venture capital raised across the world that year, and close to 60% of AI funding went into individual rounds of $500 million or more.[9] The figures have since grown larger still: in the first quarter of 2026 *OpenAI* alone raised $122 billion, the largest venture round ever recorded.[10] The sums are now so large that ordinary venture funds can no longer supply them on their own, and sovereign wealth funds from the *Gulf* and *Asia* have become the main backers of the biggest deals.

The boom reaches well beyond software. Building and running these models takes vast quantities of specialised *chips* and electricity, and the spending on that physical layer has been on a similar scale. *Nvidia*, whose processors train most leading models, became in 2025 the first listed company to be worth $4 trillion, and the largest cloud providers committed hundreds of billions of dollars between them to new capacity. The *Stargate* project, a joint venture announced in 2025, alone plans

[6] Stanford HAI [54]. [7] Stanford HAI [54]. [8] Crunchbase [19]. [9] Crunchbase [19]. [10] Crunchbase [20].

to invest between \$100 billion and \$500 billion in *data centres* in the *United States* by the end of the decade.[11]

There is an important difference between this episode and the dotcom boom that came before it. The internet companies of 2000 were valued on little more than a promise, whereas the AI products of today are used by hundreds of millions of people and already earn real revenue. Yet the central tension is plain to see. Valuations have been set in private, at extraordinary levels, against revenues that remain far smaller, and the capital has gathered into a handful of firms at a *speed* and a *concentration* with few precedents.

1.3 Research Question

Whether the rush of capital into artificial intelligence marks the early stage of a lasting transformation or the familiar shape of a speculative bubble is the question this research sets out to examine. No single number can settle it. The difficulty, and the reason the question is worth asking afresh, is that the *AI boom* lives in two markets at once: a vast private market of young companies funded by venture capital, and a public market of listed firms whose shares trade every day. Most existing work looks at only one side. The recent *econometric* studies of speculation in the boom concentrate on listed stocks, above all the *Nasdaq* index and the group of large firms known as the *"Magnificent Seven"*,[12] while the literature on private valuations examines venture deals on their own terms.[13] Few studies bring the two together. This research tries to close that gap by studying the private and the public markets within a single framework.

[11] Stanford HAI [54]. [12] Basele et al. [5]. [13] Gornall and Strebulaev [30], Gompers and Lerner [29].

2 – Methodology and Literature Review

The question raised in the previous chapter does not sit comfortably within a single field. To ask whether the capital gathering around artificial intelligence reflects lasting value or speculative excess is to ask, at the same time, how that capital is shared among firms, how the prices attached to it behave through time, and how the largest events compare with the rest. No one discipline answers all three. The analysis in this research therefore proceeds along three lines, taken up in turn and in the order that follows.

The first line is *economic* and *financial*: it measures how evenly or unevenly capital is spread across companies, with concentration measures such as the *Gini coefficient* and the *Herfindahl-Hirschman* index, and it weighs the valuations investors pay against the fundamentals those valuations are meant to represent. The second is *statistical* and *econometric*: it studies the shape of the data and the movement of prices, fitting distributions by *maximum likelihood*, judging their fit with the *bootstrap*, measuring *skewness*, *kurtosis* and the other features of a *long tail*, and applying *recursive tests* such as *GSADF* to detect and date explosive episodes. The third is *statistical mechanics*, the branch of *physics* whose methods have been carried into finance under the name of *econophysics*: it treats a market as a system of many interacting parts and looks for the traces such systems leave behind, power laws in the size of events and the *accelerating*, *oscillating* climb described by the *log-periodic power law* model.

One caution governs all three, and it is the reason the net is cast so wide. No single one of these measures, taken on its own, proves that a *bubble* is present; each is consistent with speculative excess, but also with a calmer account of the same facts. The purpose of using all three is not to find one decisive test but to see whether independent lines of evidence agree. In the image with which this work opened, each measure is one of nature's long threads, and the aim is to weave them together, on the plain principle that one clue does not make a proof, but several that point the same way might. The three sections that follow take up each approach in turn, setting out what it measures, where it comes from, and how this thesis puts it to use.

2.1 Economic and Financial Approach

The first set of tools comes from economics and answers a plain question: when capital flows into a sector, is it spread broadly or does it pool in a few hands? Two measures with a long history serve the purpose. The *Gini coefficient*, introduced by *Corrado Gini* in 1912, runs from zero, when every firm receives an equal share, to one, when a single firm receives everything.[1] The Herfindahl-Hirschman index, which appears in the independent work of *Hirschman* and of *Herfindahl* and is now standard in competition policy, adds up the squared shares of all the firms and so rises as a market grows more concentrated.[2] Applied to funding rounds rather than to market shares, both measures describe how much of the money raised in a sector is captured by how few companies.

Alongside concentration sits the question of value. In private markets the prices investors pay are not set by daily trading but negotiated round by round, and there is good reason to treat the headline figures with care. *Gornall* and *Strebulaev* have shown that the reported valuations of companies backed by venture capital systematically overstate their true worth, because the favourable terms granted to the newest investors are ignored when the whole company is valued at the price of its latest shares.[3] *Gompers* and *Lerner*, studying an earlier cycle, found that when large sums chase a limited set of deals the valuations themselves are pushed upward.[4] A note of caution belongs here as well. *Concentration* describes how capital is arranged; it is not in itself a sign that anything is *mispriced*, and recent work on public equity has shown that a highly concentrated market need not be a riskier one.[5] These measures set the scene, then, rather than deliver a verdict.

2.2 Statistical and Econometric Approach

The second set of tools is *statistical*, and it handles two separate things: the shape of a distribution and the *behaviour* of a series through time. To describe the shape, one fits a candidate distribution to the data and estimates its parameters by *maximum likelihood*, the standard way of choosing the values that make the observed data most probable. Whether the fit is any good is then judged by resampling the data many times over, the procedure known as the *bootstrap*, introduced by *Efron* in 1979.[6] Because distributions with long tails are easily mistaken for one another, it is rarely enough to fit a single curve; the analysis also measures *skewness* and other features of the *tail* and, where two candidate models are close, compares them with the test of *Vuong*, which can say whether one fits significantly better or whether the data simply cannot tell them apart.[7] For fitting and testing distributions of this kind the thesis follows the now standard procedure of *Clauset, Shalizi* and *Newman*.[8]

The second task is to find speculation in prices as it unfolds. The starting point is the test of *Dickey* and Fuller, *which* asks whether a series wanders freely like a *random walk* or is instead pulled back towards a trend.[9] A *bubble* is the opposite of

[1] Gini [28]. [2] Hirschman [33], Herfindahl [32]. [3] Gornall and Strebulaev [30]. [4] Gompers and Lerner [29]. [5] Kritzman and Turkington [36]. [6] Efron [23]. [7] Vuong [56]. [8] Clauset et al. [13]. [9] Dickey and Fuller [22].

both: a series that runs away from itself and grows explosively. *Phillips, Shi* and *Yu* turned this idea into a working instrument, the *GSADF* test, which applies the same logic again and again over a moving window of subsamples, and so can detect, and date, several separate episodes of explosive behaviour within a single history.[10] The test has since been used to identify *bubbles* in cryptocurrency prices,[11] and, most relevant here, in the recent prices of the *Nasdaq* and the largest technology stocks during the *AI boom*.[12]

2.3 Physics and Statistical Mechanics Approach

The third set of tools comes from *physics*, and more precisely from the field of *econophysics*, which applies the methods of *statistical mechanics* to financial markets.[13] It rests on an observation that recurs throughout the natural world: many quantities are distributed not in the gentle bell shape of the normal curve but according to a *power law*, in which events of every size occur and the very large ones, though rare, are far more frequent than a bell curve would permit.[14] A distribution of this kind has no typical scale, so the same statistical shape appears whether one looks at small events or large ones, the property that the opening quotation of this thesis captures in Feynman's words. Firm sizes are a familiar example, following a power law across the whole economy,[15] and the sizes of funding rounds, this thesis argues, can be read in the same way.

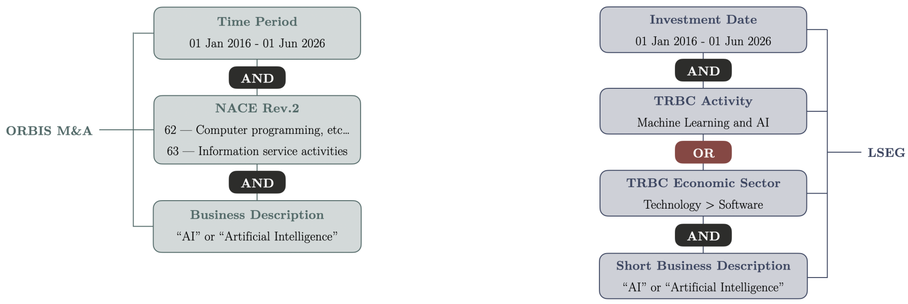
Physics also supplies a model of how a *bubble* ends. Borrowing the *mathematics* of *critical points*, the moments at which a physical system changes state abruptly, *Johansen, Ledoit* and *Sornette* proposed that a speculative *bubble* is a market approaching such a point.[16] In their *log-periodic power law* model the price does not simply rise but *accelerates*, climbing faster than any *exponential* as buyers imitate one another, while the small waves of correction crowd closer together as the *critical moment* nears. Fitted to data, the model returns an estimate of that moment, the most probable time of a crash, and it has been applied to many markets, including a careful study of *Bitcoin*.[17] It is the natural counterpart, in continuous time, to the measures of concentration and explosiveness in the two preceding sections, and together the three approaches form the toolkit that the rest of this thesis applies.

---

[10] Phillips et al. [48]. [11] Cheung et al. [12], Cheah and Fry [11]. [12] Basele et al. [5]. [13] Mantegna and Stanley [38]. [14] Newman [44], Gabaix [27]. [15] Axtell [3]. [16] Johansen et al. [34], Sornette [52]. [17] Wheatley et al. [57].

3 – Static Criticalities

The whole analysis begins by examining different types of 'criticality'. The first one is 'Static', meaning that it does not rely on continuous price observations but on the discrete price points captured at venture capital financing rounds in the private market. To obtain a concrete sample of the Artificial Intelligence private market the following filters were applied to two leading data sources: ORBIS M&A (ex Zephyr) by Bureau van Dijk (Moody's Analytics) and Workspace by LSEG (London Stock Exchange Group), as outlined in the steps below:



3.1 Data Universe

After obtaining the 'raw' data, the construction of the dataset begins with an exploration of the latter, followed by a careful assessment of the available observations. Finally, an integrative approach was adopted aiming at merging both sources letting them complement each other.

The first step towards the merge involved assigning a unique Identification Number (*ID*) to every company across the two datasets (1,853 in ORBIS and 5,595 in LSEG), with cross-source duplicates resolved through a fuzzy matching algorithm: exact matches were accepted directly (62), while the remaining 520 candidate pairs, ranked by a similarity index (from 0.6 to 1.0), were manually validated, yielding a further 168 matches. Using those *IDs*, the merge was completed according to the following set of rules:

- Records from the two sources were merged into a single financing round whenever their dates fell within ±15 days of one another, since after this interval no 'similar' financing round (not even in the *Round Total Value*) was observed; anything further apart was retained as a separate round.
- Where a financing round was matched across both providers, the LSEG *Round Total Value* was kept in place of the Orbis figure, as the more precisely recorded of the two.

- Deal stage, the pre- and post-money valuations, and the remaining round-level descriptors come from LSEG; for rounds originating in Orbis these were filled in from the matched LSEG entry, with each valuation kept against its own round date.
- The *Country* field carries the location reported by each source, comma separated where the two disagree (at most two distinct codes), rather than defaulting to a single provider.
- Near identical investor names were collapsed to one entry, the resulting syndicate listed in a single cell as a comma separated string.
- Financing rounds are ordered from most to least recent; an unmatched round inserted between two neighbours takes an index interpolated between them, so the chronological sequence is preserved.
- Company names follow the LSEG spelling.

Using a custom Python ETL (Extract Transform Load) script, these rules were applied, producing the following excerpt of the merged dataset, which contains *11,317* financing round observations across *7,218* companies in *84* countries worldwide. The dataset represents a total flow of *EUR 548,792,112,862.11* (≈ EUR 549 billion), accounting for slightly more than 50%[1] of global capital flowed into the private Artificial Intelligence market over the past decade.

The first seven columns contain company-related variables: The assigned ID, the company's known name, the company's principal geographic location(s) (two-letter ISO 3166-1 country code), and the LSEG's *TRBC Sector classification*, listed in order of increasing granularity, represented by the letters ***E***, ***B*** and ***A*** (Economic, Business and Activity). Variable ***D*** provides a short business description by LSEG (e.g. "Anthropic PBC Develops and provides AI models and safety-focused AI services").

The last seven columns are, instead, round-specific variables: ***The Round Date*** states the exact date of the relative financing round; ***No.*** the financing round number, ***Stage*** indicating the maturity of the company or the transaction type, ***Investors*** the name of the firms participating in the relative financing round, separated by a comma in the same cell, and ***Round Total Value***, ***Pre-Valuation*** and ***Post-Valuation***, respectively state the total amount collected by the financing round and the disclosed Pre and Post round valuation (where applicable).

[1] Author's own calculation. The denominator is the cumulative global private AI investment, aggregated from the country level figures in Stanford HAI, *The AI Index 2026 Annual Report*, Chapter 4: Economy, Figure 4.2.12, p. 184 (underlying data: Quid, 2025). Available at https://hai.stanford.edu/ai-index/2026-ai-index-report.

Table 3.1: Excerpt of the dataset (Round Total Values in EUR; pre and post valuations in USD, as reported by LSEG).

ID	Company	Country	E	B	A	D	Round date	No.	Stage	Investors	Round Total Value	Pre-Valuation	Post-Valuation
0109	Anthropic	US	Te...	So...	So...	De...	2026-05-28	20	Expansion	Blackstone Inc, ...	55,789,500,000.00	900,000,000,000	965,000,000,000
0109	Anthropic	US	Te...	So...	So...	De...	2026-02-12	19	Expansion	JP Morgan, ...	25,275,898,651.95	350,000,000,000	380,000,000,000
...	...	...	...	...	...	...	...	...	...	...	...	...	...
0109	Anthropic	US	Te...	So...	So...	De...	2021-05-28	1	Early Stage	Eric Schmidt, ...	107,508,000.00	n.a.	n.a.
...	...	...	...	...	...	...	...	...	...	...	...	...	...
1153	OpenAI	US	n.a.	n.a.	n.a.	n.a.	2026-03-31	n.a.	n.a.	Microsoft, ...	105,774,000,000.00	n.a.	n.a.
1153	OpenAI	US	n.a.	n.a.	n.a.	n.a.	2025-12-31	n.a.	n.a.	Thrive Capital	867,000,000.00	n.a.	n.a.
...	...	...	...	...	...	...	...	...	...	...	...	...	...
1153	OpenAI	US	n.a.	n.a.	n.a.	n.a.	2016-08-22	n.a.	n.a.	Y Combinator	104,040.00	n.a.	n.a.
...	...	...	...	...	...	...	...	...	...	...	...	...	...
5129	Zymbly Ltd	GB	Te...	So...	Ma...	Of...	2026-03-24	1	Early Stage	Y Combinator, ...	1,344,018.00	2,248,000	3,808,000

Following this methodology the relative parallel ‘control’ dataset has been created reverting the filters described at the opening of this chapter, selecting only the Software producing companies that did not contain the keywords *“AI”*, *“artificial intelligence”* and *“Crypto”*, *“NFT”*, *“Blockchain”* in the short business description (***D***), in order to preserve the control dataset from further speculative dynamics concerning another recent phenomenon that will be discussed later. It contains *52,310* different financing rounds of *27,720* companies across 114 countries. From now on the datasets concerning artificial intelligence companies will be addressed as ***“AI”*** and the control as ***“Control AI”***.

3.2 Global Statistics

The first exploratory descriptive statistics computed on the previous dataset show that the distribution of round values is extremely skewed. A handful of rounds, like those of *OpenAI* and *Anthropic*, sit far above everything else, with the largest 1% absorbing roughly 72% of all the capital raised (Figures 3.1 and 3.2), while the vast majority pile up at the bottom of the scale and are almost indistinguishable from one another. The same holds at the company level: *OpenAI* and *Anthropic* alone account for about 28% and 18% of the total respectively, close to half between the two of them. On the logarithmic plots the data fall along a smooth, slowly declining line, and the mean (about EUR 48.5 M) lies far above the median (about EUR 3.5 M). As a result, the mean is not a statistically meaningful summary of a typical round: it is pulled upward by a few very large deals and tells us very little about the bulk of the sample.

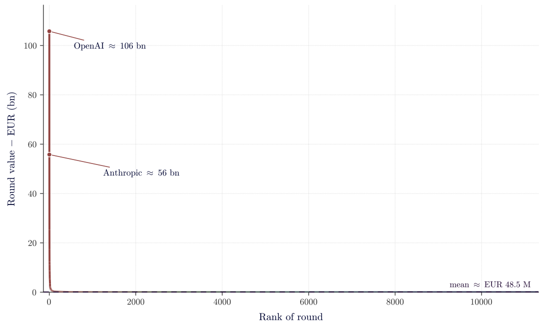


Figure 3.1: *AI* dataset *Round Total Values* across all companies, ranked from largest to smallest (linear scale)

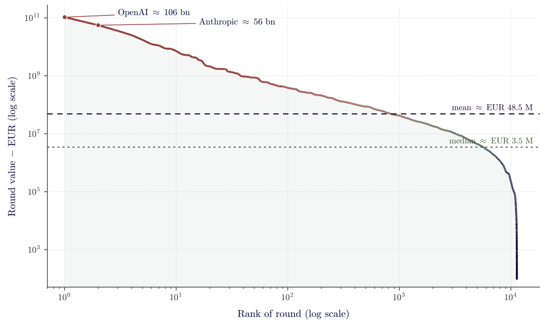


Figure 3.2: *AI* dataset *Round Total Values* across all companies, ranked from largest to smallest on logarithmic axes

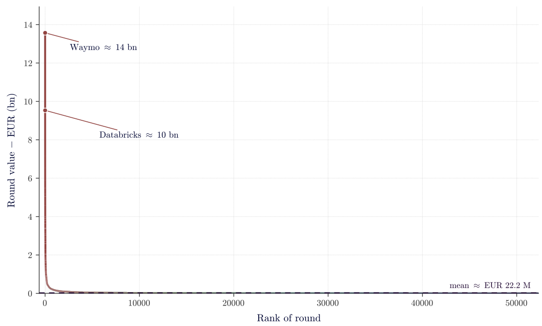


Figure 3.3: *Control AI* dataset *Round Total Values* across all companies, ranked from largest to smallest (linear scale)

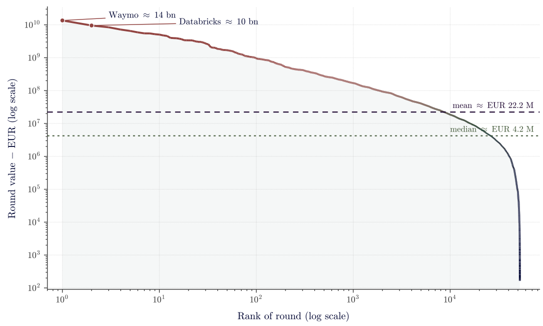


Figure 3.4: *Control AI* dataset *Round Total Values* across all companies, ranked from largest to smallest on logarithmic axes

The *Control AI* group tells a similar story, but in a milder key. The same skew is there: a few large rounds at the top and a long tail of small ones (Figure 3.3 and 3.4), yet the concentration is far less pronounced: the leading firms, *Waymo* and *Databricks*, hold only about 2% and 1.5% of the total, against the roughly 28% of *OpenAI*, and the largest rounds reach about EUR 14 bn rather than EUR 100 bn. The mean (about EUR 22 M) still sits above the median (about EUR 4 M), but the two lie much closer together, so the distribution, though its tail is heavy, is far less dominated by its largest deals than the *AI* dataset.

To investigate this concentration and imbalance, the *Gini coefficient*[2] of the company funding totals is computed for each of the two datasets.

Geometrically the coefficient is read from the *Lorenz curve*, which plots the cumulative share of total funding against the cumulative share of companies ranked from smallest to largest: the *Gini* is twice the area between this curve and the diagonal of perfect equality. Equivalently, it is the mean absolute difference in funding between all pairs of companies, scaled by twice the mean

$$G = 1 - 2\int_0^1 L(p)\,\mathrm{d}p = \frac{1}{2\,n^2\,\mu}\sum_{i,j=1}^{n}|x_i - x_j| \tag{3.1}$$

where $L(p)$ is the *Lorenz Curve*, p the cumulative share of companies, x_i the total funding raised by company i, n the number of companies and μ the mean funding. The two datasets are then analysed in two ways: on the full sample, and split into the periods before and after the release of *ChatGPT* by *OpenAI* at the end of 2022, so that the change in concentration over time can also be captured, obtaining the following values:

Table 3.2: Gini coefficient of company funding totals, full sample and split around the release of ChatGPT (end of 2022). Δ is the change between periods.

	Full Sample	≤ 2022	> 2022	Δ
AI	0.926	0.795	0.947	+0.152
Control AI	0.817	0.811	0.823	+0.012

On the full sample the *AI* dataset shows a Gini of 0.926 against 0.817 for the *Control AI* group, so capital is allocated more unequally in *AI* than in the broader benchmark (Table 3.2). The temporal split makes this clearer. In the *Control AI* group the coefficient barely moves between the two periods, rising only from 0.811 to 0.823 ($+\approx 1.5\%$). In the *AI* dataset it climbs from 0.795 to 0.947 ($+\approx 19\%$), by far the larger change of the two. Concentration in *AI* funding was therefore close to that of the *Control AI* group before 2022 and well above it afterwards.

This suggests that the increase in concentration is specific to AI rather than a general feature of private markets, since the *Control AI* group, used here as a benchmark,

[2] The Gini coefficient is a standard measure of inequality, introduced by Gini [28]: it summarises how evenly a quantity is shared among the units that hold it, here how evenly total capital is spread across companies. It ranges from 0, when every company has received the same amount, to 1, when a single company holds everything; a higher value therefore means a more concentrated dataset.

stays almost flat over the same window. The timing also matches the onset of the recent boom at the end of 2022.

The same contrast appears in the *Herfindahl–Hirschman index* (HHI)[3], a complementary measure of concentration whose inverse gives the effective number of equally sized firms in the market. For the *AI* dataset the index is 0.118, equivalent to a market of only about 8.5 such firms, while for the *Control AI* group it is 0.0018, equivalent to roughly 560. Capital in *AI* is therefore spread across far fewer effective players than in the benchmark.

3.3 Analysis of the Tail

To shift from a "global" to a more in depth perspective, the shape of the distribution is firstly assessed through its skewness and kurtosis: the third and fourth standardised central moments, which capture its asymmetry and the weight of its tails. They are defined as:

$$\text{Skewness} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^3 \qquad \text{Kurtosis} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^4 \tag{3.2}$$

And they yield the following results:

Table 3.3: *Skewness* and *kurtosis* of the distribution of round values, against the values of a *Gaussian distribution* as a benchmark.

	AI	*Control AI*	*Normal*
Skewness	67.29	40.08	0
Kurtosis	5232.25	2476.85	3

Both values lie far above their Gaussian benchmarks: the large positive skewness (67.29 against 0) reflects a strong right hand asymmetry, with a long upper tail, while the very high kurtosis (5232.25 against 3) indicates extreme *leptokurtosis*; a distribution whose spread is governed by a handful of very large rounds rather than by the bulk of the sample. The *Control AI* group shows the same pattern in a milder form: a skewness of about 40.08 and a kurtosis of about 2476.85, both well below the *AI* figures and consistent with its less concentrated distribution.

Taken together, these observations suggest that ordinary summary statistics are of little use here and that whatever is interesting about the sample lies in its upper tail. This is no surprise, as heavy tails and power laws are a well documented feature of economic and financial size distributions [27], venture capital being no exception; the

[3] The HHI is the sum of the squared funding shares of all companies, HHI $= \sum_{i=1}^{n} s_i^2$ with $s_i = x_i / \sum_{j=1}^{n} x_j$; its inverse 1/HHI gives the effective number of equally sized firms [32, 33].

descriptive picture is therefore set aside in favour of a more formal and model based examination of the tail, asking whether the distribution of round values follows a power law and, if so, with what exponent, which gives a compact measure of how unequal the allocation of capital is.

3.4 Power-law

The *Analysis of the Tail* of the previous section pointed clearly towards a heavy upper tail, but said little about its precise shape. To characterise it properly, the distribution of round values is treated as a candidate power law and estimated through the framework of Clauset et al. [13], already introduced in the literature review. The choice is deliberate: the usual shortcut of drawing a straight line through the log-log plot, though still widespread, is known to deliver biased and unreliable estimates of the exponent [13]. Their procedure avoids this by estimating the scaling exponent α by maximum likelihood above a lower cut off $x_{\min}$ (itself chosen so as to minimise the distance between the data and the fitted model) and by judging the quality of the fit through a bootstrap goodness-of-fit test rather than by eye. This lets the tail be described, and the value of α read, on firmer ground than visual inspection alone.

Following this methodology the first step is the choice of the $x_{\min}$ value; defined as the lower cut off above which the distribution is taken to follow a power law: everything below it is treated as a different regime and left out of the fit, so $x_{\min}$ simply marks where the tail begins. For all possible $x_{\min}$ values ranging from 10^6 to 10^{10} *Round Total Value* a corresponding α is estimated through *Maximum Likelihood Estimation* (MLE):

$$\hat{\alpha} = 1 + n \left(\sum_{i=1}^{n} \ln \frac{x_i}{x_{\min}} \right)^{-1} \tag{3.3}$$

Where n is the number of observations such that $x_i \geq x_{min}$ [13]

Having paired all possible combinations between x_{min} and $\hat{\alpha}$ the chosen pair is the one that minimises the *Kolmogorov-Smirnov* (KS) distance:

$$D = \max_i \left| \frac{i}{n} - \left(1 - \left(\frac{x_i}{x_{\min}} \right)^{-(\alpha-1)} \right) \right| \tag{3.4}$$

The statistic compares two cumulative curves: the empirical share of rounds at or below each value (i/n) and the same quantity under the fitted power law ($1 - (x_i/x_{\min})^{-(\alpha-1)}$). It then takes the widest vertical gap between them, and the smaller that gap, the more closely the fitted law follows the data. This yields the following values:

Table 3.4: Power-law tail fit for the *AI* and *Control AI* datasets

Sample	$\hat{\alpha}$	95% *CI*	x_{min}	n	n_{tail}	*KS*
AI	1.911	[1.874 ; 1.950]	16.3 M	11,310	2,221	0.035
Control AI	2.296	[2.252 ; 2.348]	79.2 M	52,280	2,552	0.029

* The gap $\Delta\hat{\alpha}_{AI} : \hat{\alpha}_{AI} - \hat{\alpha}_{Control\ AI} = -0.386$ (95% CI $[-0.448, -0.326]$) has been tested using a Nonparametric Bootstrap and remained < 0 in all 2000 resamples.
** Here and in Tables 3.6 and 3.8, n counts the rounds entering the fit after removing the few records identical in company, date and value; the totals reported in Section 3.1 count all recorded rows.

The exponent α measures how heavy the tail of the *continuous power-law* distribution is:

$$p(x) = \frac{\alpha - 1}{x_{\text{min}}} \left(\frac{x}{x_{\text{min}}} \right)^{-\alpha}, \qquad x \geq x_{\text{min}} \tag{3.5}$$

The lower it is, the more of the total capital sits in the very largest rounds. Both samples show heavy tails above their thresholds x_{min} (the point where the power law sets in), but the *AI* tail is clearly heavier, with $\hat{\alpha} \approx 1.91$ against ≈ 2.30 for the *Control AI* dataset. This difference ($\Delta\hat{\alpha}_{AI} \approx -0.39$) matters because it falls either side of $\alpha = 2$: below that value, the distribution has no finite mean[4], so the average is dominated by a few extreme deals and is therefore an unreliable summary of a typical round; and the *AI* dataset sits in this regime while the *Control AI* does not.

The power law only describes a small slice of each sample. In the *AI* data it starts at about EUR 16 M and covers the largest fifth of the rounds (2,221 out of 11,310), so roughly 80% of the rounds lie below this point and are left out. In the *Control AI* it starts much higher, at about EUR 79 M, and covers only the top 5% or so. The heavy tail, then, belongs to the largest rounds alone; but in the *AI* dataset it begins earlier and takes in more rounds. The *KS* distances are small in both cases (about 0.03), meaning the fitted power-law line stays close to the empirical tail. A close fit, however, is not the same as a correct one, and on its own it does not establish that the tail truly follows a power law.
To further understand whether the power-law distribution fits the *AI* dataset, it is mandatory (according to Clauset et al. [13]) to apply a *parametric bootstrap* to estimate the *goodness of fit* (GoF).

[4] This is a standard property of power-law distributions: the mean is $\langle x \rangle = \int_{x_{\text{min}}}^{\infty} x\, p(x)\, dx \propto \int x^{1-\alpha}\, dx$, which converges only for $\alpha > 2$; more generally the k-th moment is finite only when $\alpha > k + 1$, so the mean requires $\alpha > 2$ and the variance $\alpha > 3$ [13, 27, 44]

$$p = \frac{1}{B}\sum_{b=1}^{B}\mathbf{1}(D_b \geq D_{\text{obs}}) \tag{3.6}$$

For the *AI* dataset the bootstrap p falls below the 0.1 threshold, which formally rejects the pure power law. This is a routine outcome at this sample size rather than a real problem[5]: the KS test becomes stricter as the number of observations grows, so with several thousand points in the tail even a tiny difference from a perfect power law is enough to cross the threshold. The difference itself stays small, the fitted curve never moving more than a few per cent from the data, and an exact power law is in any case an idealisation that real data almost never follow exactly.

A sharper way to judge the power law is the *likelihood ratio* test of Vuong [56] for choosing between two competing distributions, which Clauset et al. [13] recommend as the proper complement to the *GoF* test, to set a fitted power law against its rivals. It does not ask whether the power law is perfect, but whether a rival fits the same tail better. For each observation x_i one takes the difference between the logarithms of the two densities, sums it over the tail and standardises it,

$$\ell_i = \ln p_{\text{PL}}(x_i) - \ln p_{\text{alt}}(x_i), \qquad R = \frac{\sum_{i=1}^{n}\ell_i}{\sqrt{n}\,s}, \qquad s^2 = \frac{1}{n}\sum_{i=1}^{n}\big(\ell_i - \bar{\ell}\big)^2, \tag{3.7}$$

where a positive R favours the power law and a negative one the alternative, with significance read from $p = \operatorname{erfc}\big(|R|/\sqrt{2}\big)$.

The first natural rival is the *lognormal*, with density:

$$p_{\text{LN}}(x) = \frac{1}{x\,\sigma\sqrt{2\pi}}\exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right). \tag{3.8}$$

Here the comparison does not point to a genuine lognormal. The closest lognormal fit places its body entirely below $x_{\min}$, so that within the observed range only its upper tail is in play, where a lognormal and a power law trace practically the same curve. The data thus show no sign of a lognormal body, and the test gives no reason to prefer one: the small lean towards the lognormal ($R = -1.17$, $p = 0.24$) is not significant, and the power law remains a fully adequate description of the tail.

The decisive comparison is with the *exponential*, the natural benchmark for a light

[5] The verdict of this test depends on sample size. Clauset et al. [13] show that at small samples the power law cannot be ruled out, since "there is simply not enough data to go on", whereas at large samples the test becomes sensitive enough to flag even the small, unavoidable differences between real data and an exact power law. Several of the datasets they examine are themselves rejected on the same grounds.

tail, with density

$$p_{\exp}(x) = \lambda\, e^{-\lambda(x - x_{\min})} \quad , \qquad x \geq x_{\min} \tag{3.9}$$

Against it the power law wins clearly ($R = +5.35$, $p < 0.001$): the *AI* tail decays far more slowly than any exponential, so it is genuinely heavy rather than merely steep. Taken together, these comparisons leave the power law as the most reliable description of the tail, which is exactly the footing α needs to serve as a comparable measure of its heaviness.

Table 3.5: Summary of the bootstrap and likelihood ratio tests on the *AI* tail. A positive R favours the power law over the alternative; the GoF p is the bootstrap value.

Test	*Statistic*	p
Bootstrap GoF (KS)	$D = 0.035$	< 0.01
Likelihood ratio vs. lognormal	$R = -1.17$	0.24
Likelihood ratio vs. exponential	$R = +5.35^{***}$	< 0.001

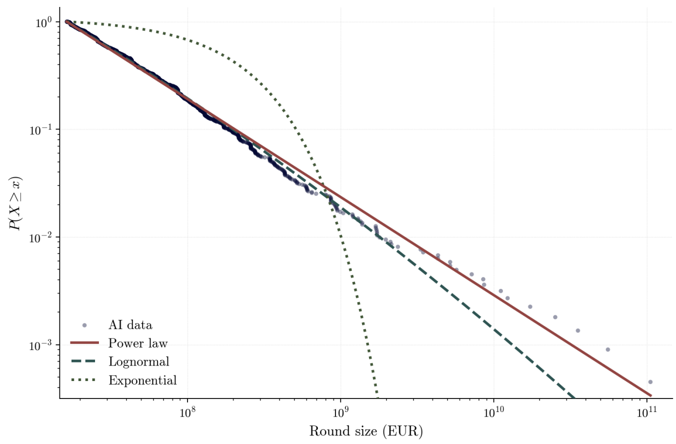


Figure 3.5: Empirical tail of the *AI* funding rounds (points) against the fitted power law, lognormal and exponential, plotted as $P(X \geq x)$ on logarithmic axes. The power law and lognormal both track the data closely, whereas the exponential falls away well before the tail, confirming that the tail is genuinely heavy.

To explore the dynamics of $\hat{\alpha}_{AI}$ and its response to the AI market boom, as in the previous analysis of the Gini coefficient, the same split before and after the introduction of ChatGPT is used to estimate the change in the exponent (≤ 2022 and > 2022). This produces the following values:

The change in the exponent $\hat{\alpha}_{AI}$ is striking. After the boom the AI tail crosses below two, falling to 1.884 from 2.052 before ($\Delta\hat{\alpha}_{AI} = 0.168$) ; a threshold the control group never reaches, since its exponent stays above two in both periods. This matters because below two the tail is at its heaviest: the largest deals dominate so strongly that, for an idealised power law, even the average size of a round ceases to be well defined. The same boom shows up just as sharply in the global measure,

Table 3.6: Power-law exponent before and after the introduction of ChatGPT, *AI* and *Control AI*

Sample	t	$\hat{\alpha}$	95% *CI*	n	n_{tail}
AI	≤ 2022	2.052	[1.988 ; 2.130]	3,533	678
AI	> 2022	1.884	[1.840 ; 1.933]	7,777	1,403
Control AI	≤ 2022	2.387	[2.293 ; 2.491]	29,855	701
Control AI	> 2022	2.219	[2.168 ; 2.273]	22,425	2,009

where the AI Gini jumps from 0.795 to 0.947, a rise of 0.152 and the largest change in the whole comparison (Table 3.2). Tail and aggregate therefore move together and in the same direction, and the agreement between a local measure and a global one is the strongest signal so far of how intensely capital concentrated in AI after the boom. The concentration runs well above a comparable benchmark and keeps deepening as the boom proceeds, signalling a market in which a few names absorb an ever larger share of the capital.

3.5 Finding the Pattern

Concentration

To assess whether these dynamics are specific to *AI* or are also present in other episodes, the same analysis is repeated on two markets widely described as speculative *bubbles* and their control datasets: the *dotcom bubble* and the *crypto bubble*. Their datasets are crafted using the same screening logic described at the opening of this chapter, considering the relative search keywords.

The *Dotcom* dataset contains 11,170 financing rounds across 6,310 distinct companies worldwide, from 1995 to 2002. Its *Control Dotcom* counterpart, built with the same keyword inversion methodology used to craft the *Control AI* dataset, collects 12,243 rounds across 5,694 distinct companies over the same span.

The *Crypto* dataset contains 1,292 financing rounds across 866 distinct companies worldwide, from 2016 to 2026. Since its period and sector are close to those of AI, no separate control is built: the Control AI dataset is duplicated and renamed *Control Crypto*.

The split into two periods is set separately for each episode, with the cut placed near the turning point of its cycle, so that the change in concentration can be read on a before and after basis within the same market. *Crypto* shares the same window and the same cut as *AI*, its own cycle having peaked in 2021 and turned down through 2022, so that the first period covers the rise and the second the decline. The *dotcom* episode covers an earlier window, from 1995 to 2002, and is split at its peak in 2000.

Table 3.7: Gini coefficient of *Round Total values*, by episode and matched control.

	Full sample	*Pre*	*Post*	Δ
Crypto	0.750	0.761	0.746	−0.015
Control Crypto	0.817	0.811	0.823	+0.012
Dotcom	0.643	0.662	0.598	−0.064
Control Dotcom	0.668	0.684	0.628	−0.056

The boundaries differ across episodes only because the episodes occur at different times.

Across the three episodes the contrast is clear (Table 3.7). *AI* is the only one whose funding is more concentrated than its control, with a *Gini* of 0.926 against 0.817, while both *Crypto* at 0.750 and *Dotcom* at 0.643 sit below their own benchmarks. The same holds over time: concentration deepens only in *AI*, where the coefficient rises by 0.152 between the two periods, while in *Crypto* and *Dotcom* there is no deepening relative to the control. Both the level and the change are therefore specific to *AI* rather than common to bubbles in general. The *Crypto* figures rest on a small number of rounds and should be read with some caution, but they point the same way. That neither *Dotcom* nor *Crypto*, both firmly regarded as bubbles, shows the same concentration makes the *AI* result more striking rather than less. Continuing the same analytical cycle, the four new datasets are interrogated to compute the power law tail exponent for each episode and its control, as for *AI*, complementing the comparison of concentration, assessing the weight of the tails.

Table 3.8: Power-law tail fit for the *Dotcom*, *Control Dotcom* and *Crypto*, *Control Crypto* datasets

Sample	$\hat{\alpha}$	95% *CI*	$x_{\min}$	n	n_{tail}	*KS*
Dotcom	3.045	[2.850 ; 3.239]	41.8 M	11,167	425	0.025
Control Dotcom	3.020	[2.862 ; 3.179]	27.4 M	12,243	626	0.022
Crypto	1.907	[1.845 ; 1.976]	4.0 M	1,288	641	0.035
Control Crypto	2.296	[2.252 ; 2.348]	79.2 M	52,280	2,552	0.029

* The gap $\Delta\hat{\alpha}_{Crypto}$: $\hat{\alpha}_{Crypto} - \hat{\alpha}_{Control\ Crypto} = -0.387$ (95% CI [−0.467, −0.305]) has been tested using a Nonparametric Bootstrap and remained < 0 in all 2000 resamples.

Dotcom sits apart from the other two episodes. Its tail exponent is close to three, $\hat{\alpha}_{\text{Dotcom}} \approx 3.0$, and indistinguishable from its control, the gap straddling zero across the bootstrap. This is the expected pattern, because the *dotcom bubble* was a public markets phenomenon. The speculation showed up in quoted stock prices on the *Nasdaq*,[6] not in the distribution of private capital, and so the funding rounds of

[6] The dotcom episode on the Nasdaq is documented in detail by Ofek and Richardson [45], who trace the rise and subsequent collapse of internet stock prices in the public equity market over 1998 to 2000.

that period carry no unusual weight in their tail.

Crypto behaves in the opposite way. Its tail exponent falls firmly below the $\alpha = 2$ line, $\hat{\alpha}_{\text{Crypto}} = 1.907$, and sits well under its control, the gap ($\Delta\hat{\alpha}_{Crypto} = -0.387$) remaining below zero in every resample. Here the speculation is private. Capital concentrated in late stage funding rounds rather than in a quoted index, and the heavy tail reflects a funding market in which a handful of rounds absorb a disproportionate share of the money. The parallel with *AI* is close to exact: the gaps, $\Delta\hat{\alpha}_{Crypto}$ (-0.387) and $\Delta\hat{\alpha}_{AI}$ (-0.386), differ only in the third decimal, and both episodes sit on the same side of the $\alpha = 2$ line, against a control that never crosses it.

The tail exponent therefore places *AI* and *Crypto* together and sets *Dotcom* apart, but it does not separate *AI* from *Crypto*. That separation comes from the Gini (Tables 3.2, 3.7). *AI* is the only one of the three episodes whose funding is more concentrated than its control, and the only one whose concentration deepens over time. *Crypto* is heavy in the tail but not across the whole distribution: its *Gini coefficient* stays below that of its control, and does not rise. *AI* is concentrated at every level and still rising, a depth the other episodes do not reach.

There is an encouraging signal in how tightly these pieces fit. The tail exponent and the *Gini*, built separately and on different parts of the data, agree in placing *AI* at the extreme. That independent agreement is reason enough to look more closely at how these markets, and the financing rounds that run alongside them, move together over time.

Dynamic Volumes

Concentration describes how capital is shared, not how much of it there is or how fast it moves. A speculative phase, though, is as much about quantity and pace as about distribution: the amount of money entering a sector, and the speed at which it arrives. This section measures that flow directly.

The quantity tracked is the total funding volume raised each year, computed for *AI* and for *Crypto* against the shared *Control* baseline of comparable venture capital. Each series is indexed to its first year of observation, so that markets of very different absolute size can be read on a common scale.

$$V_t \;=\; \sum_{i=1}^{n_t} v_i, \qquad I_t \;=\; 100\,\frac{V_t}{V_{t_0}}, \tag{3.10}$$

where n_t is the number of funding rounds in year t and v_i the value of round i, so that V_t is the total volume raised that year. The base year t_0 is the first year of observation, which fixes $I_{t_0} = 100$, and each dataset is normalised to its own V_{t_0}. The horizontal axis is the relative year $\Delta t = t - t_0$. The question is whether the capital flowing into *AI* is not only large but accelerating, and whether its path departs from the broad market and from the earlier *crypto* cycle.

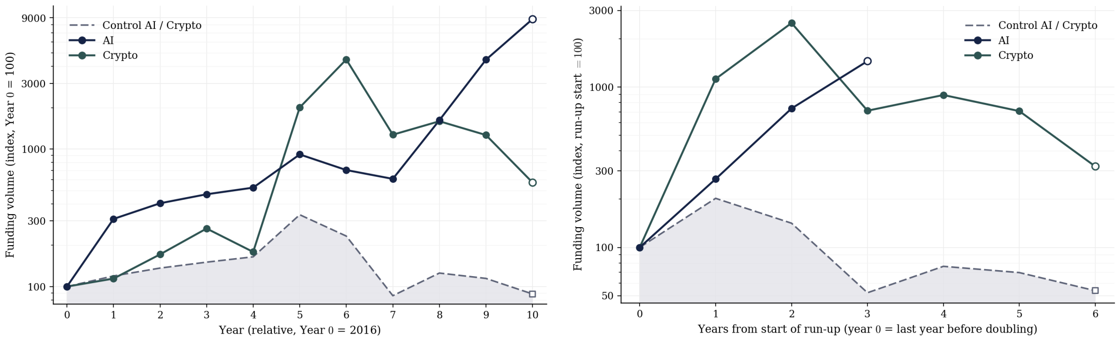


Figure 3.6: Annual funding volume for the *AI* and *Crypto* datasets and their shared *Control*, indexed to 100 as in equation (3.10) on a logarithmic scale. The left panel uses calendar time (Year 0 = 2016); the right panel aligns each series to its own run-up, with Year 0 the last year before doubling.

Three features stand out. The scale of the *AI* inflow is the first: annual funding volume rises from EUR 2.9 billion in 2016 to EUR 255 billion in 2026, about 87 times its starting level, and the series is still climbing at the end of the sample with no peak yet. The *Control* sharpens the point. The comparable venture capital market is essentially flat and ends below its own 2016 level, so the *AI* surge is specific to the sector rather than the product of a rising market; by 2025 *AI* already absorbs close to 60% of comparable funding, against roughly 10% before 2023. *Crypto*, by contrast, traces a complete arc, peaking around 2022 at some 44 times its 2016 level and then falling back, whereas *AI* accelerates rather than turns, its annual growth stepping up sharply after 2022. A sharp and accelerating inflow of capital into a single sector, while the broader market stays flat, is consistent with the boom phase of a speculative episode,[7] though here the inflow has not yet reversed.

The right panel shows the two episodes from the same point, the last year before each begins to double, so they can be compared at the same stage of their rise rather than by the calendar. They differ in *velocity*, but the telling difference is where each one now sits. *Crypto* rose quickly and was already falling within two years, after reaching about 25 times its starting level. *AI* has risen more steadily, about 7 times up after two years, and unlike *crypto* it is still going up. The year each series is pinned to is a choice, so this is a precedent rather than a forecast. What the panel shows is *AI* on the rising part of an arc that, in *crypto*, ended in a fall.

The same difference in velocity shows up round by round. Capital goes into *AI*

[7] On the surge of speculative capital in the boom phase of financial manias see Kindleberger and Aliber [35]; on inflows of capital into venture markets inflating deal activity and valuations specifically, see Gompers and Lerner [29].

faster than into the *Control*: round size grows by 1.80 times a year against 1.51, a clearly significant gap on a *Mann-Whitney* test ($p < 0.0001$), and the jump from one round to the next is larger too, 2.11 times against 1.78. The one thing *AI* shares with the market is acceleration: the share of rounds speeding up is 39.3% against 39.8%, which a test of proportions cannot separate ($p = 0.79$). So the marker is velocity, not acceleration: money moves into *AI* much faster than anywhere else, even if the rate at which that speed builds is the market's. As with the tail exponent, *AI* is set apart by its level, here the sheer speed of the inflow.

Valuation gap

A speculative episode shows not only in how capital concentrates but in how fast the value placed on a company is marked up. Where a *valuation* is observed at more than one round, the rate of revaluation can be read directly. For each company the valuation at one round is set against the valuation at the previous round, and the ratio is annualised over the time between them,

$$v_i = \left(\frac{EV_{i,t}}{EV_{i,t-1}} \right)^{1/\Delta t_i} \tag{3.11}$$

where Δt_i is the time in years between the two rounds. A value of $v = 2$ means the valuation doubles each year. The median is used throughout, being robust to a small number of extreme pairs.

Table 3.9: Median annualised valuation growth v, eq. (3.11); lower row, valuations $\geq$ EUR 50M only.

Sample	*AI*	*Control*
Full dataset	2.50	1.84
$EV \geq 50\ M$	3.37	1.83

At real, consolidated companies (those carrying a valuation of at least EUR 50 million), *AI* valuations grow at a median of about 3.4 times a year, against about 1.8 times in the *Control* (*Mann-Whitney* $p < 0.0001$). Unlike the level of valuation relative to capital, which does not separate the two groups, the pace of revaluation does, and the gap widens rather than narrows once the smallest companies are removed. The result is therefore not an artefact of a few tiny firms.

The mechanism is one of frequency rather than size. *AI* companies return to the market far more often: the typical gap between two valued rounds is about 6 months, against about 11 in the *Control*. Each single step is in fact slightly smaller in *AI*, a median markup of 1.63 against 1.76. It is the rapid succession of rounds, each lifting the valuation again before the previous one has settled, that compounds into the faster annual rate.

This measures the speed at which prices are set, not whether they are deserved. A valuation marked up every few months may reflect genuine growth as easily as exuberance. What the figure shows is that, on this pace, *AI* stands apart from the wider private market in the same direction, and to a similar degree, as it does in the concentration of its capital: serial revaluation at a rate the rest of the market does not match.

A converging picture

Taken together, the static analysis returns a set of structural features that line up, one after another, with the shape associated with a speculative phase. Capital is concentrated beyond anything the broader market reaches: the top 1% of rounds hold 72% of all funding, and the *AI Gini* of 0.926 exceeds the *Control*'s 0.817 and climbs by 0.152 after 2022 while the benchmark stays flat, alone among the three episodes in pulling away from its own control. The tail of the round size distribution is correspondingly heavy, its exponent falling below 2 after the boom, the regime in which a few deals dominate the whole. That tail places *AI* alongside *Crypto*, their gaps differing only in the third decimal, and apart from the public market *Dotcom*: any speculation here is private and structural, though *AI* alone is concentrated at every level rather than in the tail only. The flow agrees, annual volume up about 87 times since 2016 against a flat *Control*, and valuations marked up about 3.4 times a year against 1.8 as companies raise roughly twice as often as the market.

None of this, on its own, shows that prices have come loose from fundamentals: concentration and speed are equally the mark of a genuine *winner take all* market, and the reported valuations themselves run well above fair value once preferential terms are counted.[8] What the chapter establishes is firmer. On every measure built here, tail, concentration, flow and the pace of revaluation, the private *AI* market sits at the extreme, and in the same direction as *Crypto*, the episode it most resembles.

[8] Gornall and Strebulaev [30] value 135 US unicorns from their legal filings and find reported post-money valuations average about 48% above fair value, the gap driven by the preferential terms granted to the latest investors.

4 – Dynamic Criticalities

The previous chapter examined private markets, where the picture was necessarily a 'static' one. A private company is valued only when it raises capital, so its value is recorded at a few scattered moments rather than followed through time, and this limits what can be said about how that value moves between one round and the next. Publicly traded companies, by contrast, are priced continuously, every day the market is open, which makes it possible to study prices as they move rather than as a set of fixed points. The analysis here is therefore a 'dynamic' one, working with the time series of publicly traded companies to ask whether their movement through time can reveal speculative behaviours that a static view cannot capture.

4.1 Prices, Velocity and Acceleration

The arrival of *ChatGPT* at the end of 2022 turned artificial intelligence from a research subject into a market theme, and public equity responded at once. Money moved towards the companies seen as most exposed to the technology, and a small group of them pulled sharply away from everything else. The clearest case is *Nvidia* (NVDA), which designs the chips on which most large AI models are trained. Its share price rose by 571% in the 18 months to June 2024,[1] and its market value then climbed through a run of milestones no company had reached before, passing $1 trillion in 2023 and $4 trillion in July 2025, the first firm ever to do so.[2] At that point *Nvidia* alone was worth more than every listed company in the United Kingdom put together.[3]

What these figures hide is how narrow the rise was. As Table 4.1 shows, the market as a whole barely changed pace: measured by the trailing one year growth of its price, the *Nasdaq Composite* (IXIC) ran only 2.3 percentage points faster after the boom than before, against 40.9 for *Nvidia* and 36.2 for *Meta* (META). Several of the largest names in fact slowed, with *Apple* (AAPL), *Microsoft* (MSFT) and *Tesla* (TSLA) all growing more slowly after 2022 than before. The boom in listed AI equity is therefore not a broad lift in the market but a sharp rise concentrated in a handful of names, those whose exposure to the technology is most direct.

The same names also changed the shape of their growth, not just its level. Before the *boom* most were slowing, each year adding less than the one before; afterwards they were speeding up, with prices bending upwards rather than following a straight line. The turn is general, with acceleration rising from negative to positive across the board, from −0.18 to 0.18 for *Meta* and from −0.17 to 0.08 for *Nvidia*. Part

[1] Basele et al. [5]. [2] NBC News [43]. [3] Al Jazeera [2].

of it is the rebound from the market low of late 2022 rather than the boom alone, since any recovery reads as an acceleration, and the shift appears even in names that grew more slowly overall.

	Velocity (%)			Acceleration		
Ticker	Pre	Post	Δ	Pre	Post	Δ
NVDA	49.5	90.4	40.9	−0.17	0.08	0.26
META	7.0	43.2	36.2	−0.18	0.18	0.36
AIQ	14.4	24.2	9.8	−0.13	0.15	0.28
GOOGL	21.2	27.8	6.5	−0.05	0.14	0.19
IXIC	16.3	18.6	2.3	−0.08	0.10	0.17
AMZN	23.8	17.2	−6.6	−0.14	0.11	0.25
MSFT	33.6	16.6	−17.0	−0.06	0.01	0.08
AAPL	36.8	15.4	−21.5	−0.04	0.00	0.04
TSLA	68.7	10.8	−57.9	−0.04	0.06	0.10
CHAT	–	43.6	–	–	0.28	–

Table 4.1: Trailing one year velocity and acceleration of log prices, averaged before and after the boom.

Figure 4.1: Velocity and acceleration of log prices for selected tickers

The same velocity and acceleration analysis was carried out on two earlier episodes that are now widely accepted as speculative *bubbles*, the *dotcom boom* of the late 1990s and the *cryptocurrency boom* of the 2010s, using the same construction so that the three can be read on a single scale. For these two the vertical line marks the peak rather than the start, March 2000 for the dotcom market and November 2021 for *cryptocurrency*, so that the run sits before it and the collapse after.

Before each peak the picture is the one a *bubble* produces. The *dotcom* names were running at extreme speed, with *Cisco* (CSCO) growing at 84% a year, *Microsoft* at 68% and *Qualcomm* (QCOM) at 80%, and the same held in *cryptocurrency*, where *Bitcoin* (BTC) ran at 135% a year and the smaller and *meme* names higher still, *Dogecoin* (DOGE) above 170%. After the peak both turn over together and every velocity falls hard into negative ground, the mark of a whole market deflating rather than a few names cooling. The prices leave no doubt about what followed: *Cisco* reached $132 in March 2000 and fell to $16 within the year,[4] while *Bitcoin* reached a record near $69,000 in November 2021 and then lost 64% of its value in 2022.[5]

What sets the *AI boom* apart from these two is not its speed but its shape. In both earlier episodes the run was broad: the market kept pace with its leaders, the *Nasdaq* growing at 33% a year before the *dotcom* peak and the whole *cryptocurrency* complex rising as one. In the *AI boom* the run is narrow. The broad market grows at only 19% a year while Nvidia runs at 90%, far ahead of it, so the gain piles

[4] Basele et al. [5]. [5] CNBC [14].

into a few names rather than lifting the market as a whole. The same explosive velocity and the same upward turn in acceleration that preceded the *dotcom* and *cryptocurrency* crashes are present in the *AI* names now, but concentrated as those episodes never were, and placed in the ascending phase rather than after a peak that has already passed. A single stock that runs at this rate, becomes the first company worth \$4 trillion and then loses close to \$600 billion in one day on a single piece of news,[6] carries a fragility that a run spread across a whole market does not. On the measures used here the *AI boom* is consistent with the early, ascending stage of the two clearest *bubbles* of the last three decades, with a concentration that arguably leaves it more exposed rather than less.

4.2 Fundamentals and Displacement

A valuation angle completes the picture drawn so far from price dynamics. The standard way to judge how dear a market is over long stretches of time is the *cyclically adjusted price-to-earnings* ratio, which divides the real price of the index by the average of its real earnings over the previous 10 years:

$$\text{CAPE}_t = \frac{P_t^r}{1/10 \sum_{k=1}^{10} E_{t-k}^r} \tag{4.1}$$

where P_t^r and E_s^r are the price and the earnings expressed in real terms, that is adjusted for inflation through the *consumer price index* (CPI). The point of averaging earnings over a decade is that a single bad year, when profits fall close to zero, no longer throws the ratio to meaningless heights, the very breakdown that makes a plain *price-to-earnings* ratio unusable around the 2000 *crash*. The measure was built for the market as a whole, but it carries over to single sectors, an approach worked out for the main sectors of the US economy over more than a century.[7]

Table 4.2: CAPE, current readings against the dotcom peak. The Nasdaq 100 figure for 2000 is not publicly available.

Index	AI boom	Dotcom peak
S&P 500	40	44
Nasdaq 100	54	—

Set against the *dotcom* episode, the broad market today looks almost as expensive as it did then. The *S&P 500* cyclically adjusted ratio stands near 40, close to its all-time high of about 44, a level reached only once before, at the *dotcom* peak in March 2000.[8] The ratio has averaged near 17 since 1881, with a median close to 16, and even against the higher average of about 27 over the past three decades the current

[6] CNBC [15]. [7] Bunn and Shiller [8]. [8] multpl [41].

level stands well clear. On this single measure the market is priced now much as it was in the months before the *dotcom* crash. The resemblance, though, hides where the height comes from. The elevated reading is not spread evenly across the market but piles up in the names that lead the *AI* rise, which fits the concentration found earlier in the price data. The *Nasdaq 100*, the index of large technology and growth stocks where the *Magnificent Seven* alone account for close to half the total weight, trades at a cyclically adjusted ratio near 54, well above the broad market, up from 34 at the end of 2022.[9] The market looks dear, in short, because a narrow set of technology names is dear, not because every part of it has been repriced (Table 4.2).

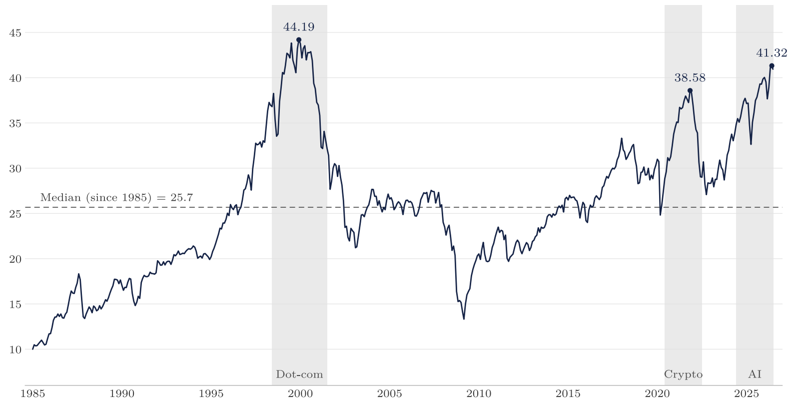


Figure 4.2: Cyclically adjusted price/earnings ratio (Shiller CAPE) of the S&P 500, January 1985 to June 2026. The shaded bands mark the peak windows of the dot-com, crypto, and AI episodes, and the dashed line is the median over the period.

4.3 GSADF – Generalized Supremum Augmented Dickey Fuller

The analysis in this section follows *Basele, Phillips* and *Shi*, whose study of the recent artificial intelligence *boom* is taken here as the reference for testing speculative behaviour in listed prices.[10] Their work, published in the *Journal of Time Series Analysis*, examines the *Nasdaq* market and the group of large technology stocks known as the *Magnificent Seven* over the period January 2017 to January 2025, and asks whether the rise in prices carries the marks of a speculative *bubble*. The study begins with the index itself, testing whether it moves in an explosive way that goes beyond steady growth. It then turns to each of the stocks on its own, holding the wider market and the industry constant, so that the exuberance of a single firm can be told apart from a general rise in the market. It finally measures how strong any explosive movement is, by placing a confidence interval around the degree of departure from ordinary price behaviour. What follows rests entirely on their framework.

The framework starts from the standard way of valuing an asset from its future payoffs. The price today is the stream of expected future payoffs, discounted back

[9] Siblis Research [51]. [10] Basele et al. [5].

to the present, together with a further term that captures anything not tied to those payoffs,

$$P_t = \sum_{i=0}^{\infty} \left(\frac{1}{1 + r_f} \right)^i \mathbb{E}_t(D_{t+i} + U_{t+i}) + B_t \tag{4.2}$$

where P_t is the price, r_f the risk free rate, D_t the payoff such as a dividend, and U_t collects factors that cannot be observed directly. The final term B_t is the *bubble*, defined by the property that it is expected to grow at the rate of discount,

$$\mathbb{E}_t(B_{t+1}) = (1 + r_f)\, B_t \tag{4.3}$$

When this term starts at zero the price is carried by payoffs alone and grows in a steady way. When it starts above zero it tends to take over the movement of the price and pushes it to grow faster and faster, and it is this behaviour that the test is built to find.

Let y_t be the *dividend adjusted index*. The case of no *bubble* is a series that drifts like a *random walk*, with a drift small enough to fade as the sample grows,

$$y_t = d\,T^{-\eta} + y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \text{iid}\,(0, \sigma^2) \tag{4.4}$$

and the case of a *bubble* is mildly explosive growth, where each value sits a little above the one before,

$$y_t = \theta_T\, y_{t-1} + \varepsilon_t, \qquad \theta_T = 1 + a\,T^{-\alpha}, \quad a > 0,\ \alpha \in (0, 1) \tag{4.5}$$

The choice between the two hypotheses is made by fitting an augmented Dickey and Fuller regression,

$$\Delta y_t = \phi + \beta\, y_{t-1} + \sum_{i=1}^{K} \psi_i\, \Delta y_{t-i} + \varepsilon_{K,t} \tag{4.6}$$

in which no *bubble* corresponds to $\beta = 0$ and a *bubble* to $\beta > 0$. The figure of interest is the t ratio on β, written *ADF*. A single regression of this kind reads the sample as a whole and can miss a *bubble* that fills only part of it. To guard against this the regression is run repeatedly, on windows that begin at different points and widen through the sample, and the largest statistic across the starting points is retained,

$$\text{PSY}_\tau(\tau_0) = \sup_{\tau_1 \in [\,0,\, \tau_2 - \tau_0\,],\ \tau_2 = \tau} \left\{ \text{ADF}_{\tau_1}^{\tau_2} \right\} \tag{4.7}$$

with the smallest window set to $\tau_0 = 0.01 + 1.8/\sqrt{T}$. The test for a *bubble* at any point in the sample then takes the largest of these values across every end point,

$$\mathrm{GSADF} = \sup_{\tau \in [\tau_0, 1]} \mathrm{PSY}_\tau(\tau_0) \tag{4.8}$$

A value above the *critical level* is read as evidence of explosive behaviour. The critical levels are not drawn from a fixed table but from a *wild bootstrap*, which keeps the test dependable when the size of the price swings changes over the sample.

Testing a single stock calls for more care, since a rise in its price may do no more than track a rise in the whole market. One approach divides the stock price by the market index and applies the same test to the ratio, so that any movement shared with the market is removed. A second approach splits the price into a part the market explains and a part that is left over, using the *Fama and French three factor model*. The return of the stock is regressed on the three factors,

$$\Delta p_{i,t} = \beta_i X_t + v_{i,t} \tag{4.9}$$

where $\Delta p_{i,t}$ is the change in the log price of stock i and X_t holds the three factors. The fitted returns are summed into the part of the price the market accounts for,

$$\widehat{M}_{i,t} = \widehat{M}_{1,i} + \sum_{t=2}^{T} \widehat{\Delta p}_{i,t}, \tag{4.10}$$

and what remains is the part specific to the stock,

$$\widehat{IC}_{i,t} = p_{i,t} - \widehat{M}_{i,t} \tag{4.11}$$

The *bubble* test is then applied to this leftover part. The price to index ratio and the stock specific component are used side by side, and agreement between them lends more weight to any *bubble* that is found. The same framework also measures the strength of an explosive episode, through the near unit root method of *Phillips*, which sets a confidence interval around how far the price process sits from ordinary behaviour.

The tests return clear evidence of speculative *bubbles*. The *Nasdaq* index and the *Magnificent Seven* all show explosive behaviour, *Nvidia* and *Tesla* most strongly in the first period to 2021 and *Meta* in the second, with *Apple* the only stock that shows no *bubble* in the second period from 2022.[11]

Tests of the same kind have fired on the two earlier episodes used here as reference points. Applied to the 1990s *Nasdaq*, the method found explosive behaviour and

[11] Basele et al. [5].

placed the start of the *dotcom* exuberance in the middle of the decade, well before the crash.[12] Applied to the major *cryptocurrencies*, it found repeated explosive episodes in every coin examined, most of them lasting only a few days.[13] The statistics from separate studies cannot be read against one another as magnitudes, since the samples, the frequencies and the windows all differ, so what carries across is the shared detection of explosive behaviour rather than a ranking of its size. On that reading the *AI* era *Nasdaq* sits in the same company as the *dotcom* and *cryptocurrency* markets, consistent with the velocity and valuation evidence set out earlier in the chapter.

4.4 LPPLS – Log Periodic Power Law Singularity

The tests so far ask whether prices are exploding. The *log periodic power law* asks the harder question of when the explosion might end, and it answers with a piece of *physics*. It treats a crash the way statistical physics treats a *critical point*, the moment a system tips from one *state* into another, and it borrows the same mathematics used for the rupture of a stressed material or the onset of a *phase transition*.[14] The picture is of a market climbing toward such a point under its own momentum, with a pair of warning signs along the way: the price grows faster than an exponential, lifted by speculative buying that feeds on itself, and it *oscillates* with swings that come faster and faster as the point draws near. The model captures both at once and reads off the *critical time* t_c, the date at which the climb is most likely to break.

The starting point is a *no arbitrage* argument. The asset carries a small but real risk of a sudden crash, and rational investors hold it only if the expected return pays them for bearing that risk. This condition ties the expected growth of the log price to the crash *hazard rate* $h(t)$, the rate at which a crash becomes imminent at time t,

$$\frac{\mathrm{d}\ln p(t)}{\mathrm{d}t} = \kappa\, h(t) \tag{4.12}$$

where κ is the size of the drop a crash would bring. Integrating, the log price is carried along by the accumulated hazard,

$$\ln p(t) = \ln p(t_0) + \kappa \int_{t_0}^{t} h(u)\,\mathrm{d}u \tag{4.13}$$

The form of the *hazard* is where the physics does its work. The market is made of traders who copy one another, and as the strength of that imitation approaches a critical value the copying locks the whole crowd into the same move, the cooperative

[12] Phillips et al. [47]. [13] Enoksen et al. [24]. [14] Johansen et al. [34].

behaviour that marks a system poised at a critical point. Near that point the *hazard* follows a power law in the time remaining to t_c, modulated by an oscillation whose *frequency* climbs without bound as t_c is approached,

$$h(t) \;\propto\; (t_c - t)^{-\alpha} \left[1 + C' \cos\big(\omega \ln(t_c - t) + \phi'\big) \right], \qquad 0 < \alpha < 1 \tag{4.14}$$

Putting this hazard back through the integral yields the law that gives the model its name, written here for the log price,

$$\ln p(t) = A + B\,(t_c - t)^m + C\,(t_c - t)^m \cos\big(\omega \ln(t_c - t) + \phi\big) \tag{4.15}$$

with exponent $m = 1 - \alpha$ lying between 0 and 1, a coefficient $B < 0$ that makes the price accelerate into the peak, and $|C| < 1$ so the oscillations ride on the trend without ever reversing it. The first term carries the faster than exponential rise; the cosine carries the oscillations, which are periodic not in time but in the logarithm of the time left to t_c, the reason they crowd together as the critical point nears.

The parameters of the model, $\{A, B, C, t_c, m, \omega, \phi\}$, are awkward to estimate, since the function is strongly non linear and its error surface is full of local minima that trap a naive search. The fit follows the scheme of Filimonov and Sornette, which expands the cosine into its slow and fast parts and regroups them,

$$\ln p(t) = A + B\,(t_c - t)^m + C_1 (t_c - t)^m \cos\big(\omega \ln(t_c - t)\big) + C_2 (t_c - t)^m \sin\big(\omega \ln(t_c - t)\big) \tag{4.16}$$

with $C_1 = C\cos\phi$ and $C_2 = -C\sin\phi$. For any fixed *triple* (t_c, m, ω) the coefficients A, B, C_1, C_2 now enter linearly and follow from least squares, so only the *triple* is searched over numerically. This shrinks the search from the full parameter set to three dimensions and keeps the local minima at bay.[15]

One feature of the model demands care: it returns a *critical time* for any series at all, *bubble* or no *bubble*, so a raw fit proves nothing on its own. The estimated parameters are therefore screened against the conditions the theory asks of a genuine *bubble*. A fit is accepted only if $B < 0$, the exponent falls in $0.1 \le m \le 0.9$ and the frequency in $4.8 \le \omega \le 13$, and the oscillations stay damped enough not to swamp the trend, a balance measured by $D = m|B|/(\omega|C|) \ge 0.8$, with a further requirement that the window hold enough oscillations to be informative.[16] A last difficulty is that the fit drifts with the window it is read over, so the calibration is repeated across many windows. The share of windows that pass the screen serves as a measure of confidence in the signal, and the spread of the *critical time* across

[15] Filimonov and Sornette [25]. [16] Sornette et al. [53].

those windows shows whether t_c is a stable feature of the data or an artefact of the window chosen.

The model was fitted to the *Nasdaq Composite* and to each of the *Magnificent Seven* in turn, over the window running from the market low at the end of 2022 through to the latest observation, on *split adjusted closing prices.* The fit follows the scheme of *Filimonov* and *Sornette* set out above, with the three *nonlinear parameters* searched over from a wide grid of starting points to keep the optimisation clear of *local minima.*

On the individual stocks the model fails in a way that is itself informative. The fit converges for every name, and the proportion of variance it accounts for is high throughout, between 0.79 and 0.96, but this fidelity is empty, since the log periodic power law bends to almost any rising series whether or not a *bubble* is present. The telling sign is elsewhere, in the parameters. Of the 7 stocks, 6 drive the exponent m to the upper edge of its admissible range at 0.9, and 5 also pin the frequency ω to a boundary, 3 of them at the lower limit of 4.8. A parameter resting against its bound is the mark of an optimiser that has found no genuine structure to settle on and has run to the edge of the space it is allowed. The 3 stocks pinned at the lowest frequency are trying to slow the oscillation below what the model permits, which is to say the data carry no real log periodicity for the model to lock onto. The *damping ratio* confirms the picture, falling below the threshold of 0.8 for most of the names, so the oscillations, where they appear at all, are too weak to count. Read together, these fits do not support a log periodic *bubble* signature in any of the 7 stocks.

The *Nasdaq index* is the one series that holds. Its fit sits inside the admissible region on every count, with an exponent away from the bounds, a frequency of about 6.8 well within range, and a damping ratio above 0.8, the only one of the 8 series to pass the full screen. The estimated *critical time* falls in the autumn of 2026.

Figure 4.3: LPPL calibration on the Nasdaq Composite, fitted over the window from the end of 2022 to the latest observation ($R^2 = 0.97$). The blue line is the log price, the red line the fitted log periodic power law extrapolated to the critical time, and the grey band marks the critical-time window across the valid fits, from September 2026 to January 2027.

A single fit, however, settles little, since the result can move with the window it is read over, so the *calibration* was repeated across windows of decreasing length, all ending at the latest observation, and each fit was tested against the same admissibility conditions. Of the 81 windows examined, 12 pass the screen, a confidence of 0.15, and they are exactly the 12 longest, the windows that span the full ascent from the end of 2022. Across these the critical time is tightly grouped, between *September 2026* and *January 2027*, with a median in the middle of *October* and a spread of 1.3 months. The agreement is not a coincidence of the chosen window but a stable feature of the long climb. As the window is shortened from the start the fits leave the admissible region one by one, the damping ratio sliding below 0.8 as the earliest part of the rise is dropped, until the signature dissolves entirely on the shortest windows.

Two readings follow, and they pull in the same direction. The log periodic signature is absent from the individual stocks and present only in the index, which places whatever speculative structure exists at the level of the market rather than in any single name. And even for the index the *critical time* is to be read as an indication rather than a forecast, since it rests on the longest windows alone and the overall confidence is low, a caution that sits naturally with the known fragility of the model when it is asked to extrapolate a crash that has not yet happened.

4.5 Nasdaq Private Market

The public side of the analysis rests on companies that are exposed to artificial intelligence rather than defined by it alone. Firms such as *Nvidia* and *Microsoft* run large and varied businesses whose prices respond to many forces beyond the AI theme alone, and the names that sit at the centre of the present wave are in any case still private, so for them the listed market offers nothing to read, because a private valuation arrives only as a discrete figure set at each funding round and not as the continuous series that the explosive root and log periodic methods require. The final interpretation of this study uses a source that removes this obstacle, the daily valuation series published by *Nasdaq Private Market* (NPM).

NPM is an institutional venue for secondary trading in private company shares, and its data arm produces a daily evaluated price, marketed as *Tape D*, for several hundred of the most actively traded names. The series is not a record of executed trades, since the shares of any single private company change hands only rarely, but an estimate that an algorithm rebuilds on each business day from six independent signals, namely reported secondary trades, the history of *bids* and offers, primary round prices, *409A* valuations, mutual fund marks, and its own earlier estimates, with the resulting price placed at Level 2 of the fair value hierarchy of ASC 820 and

IFRS 13.[17] Every daily price (for any of the 390 available companies) was extracted directly from the HTML of the NPM platform by an AI agent, producing, to the author's knowledge, the first aggregate visualisation of a private market's continuous prices.

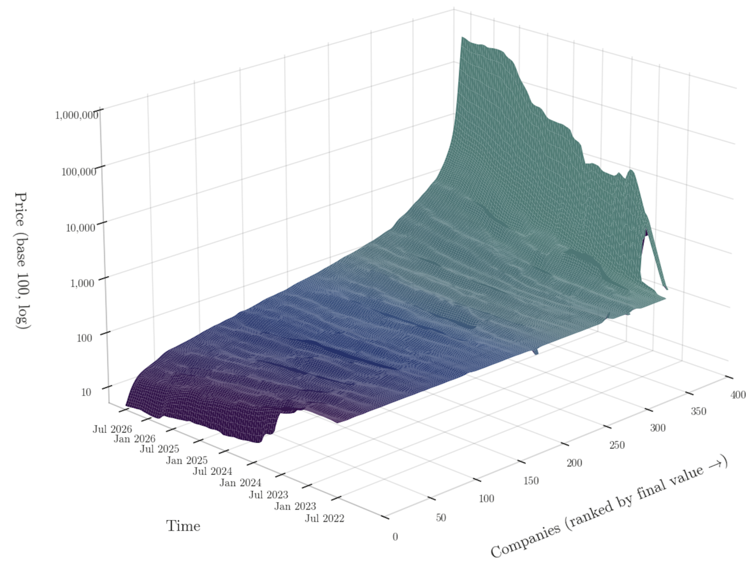


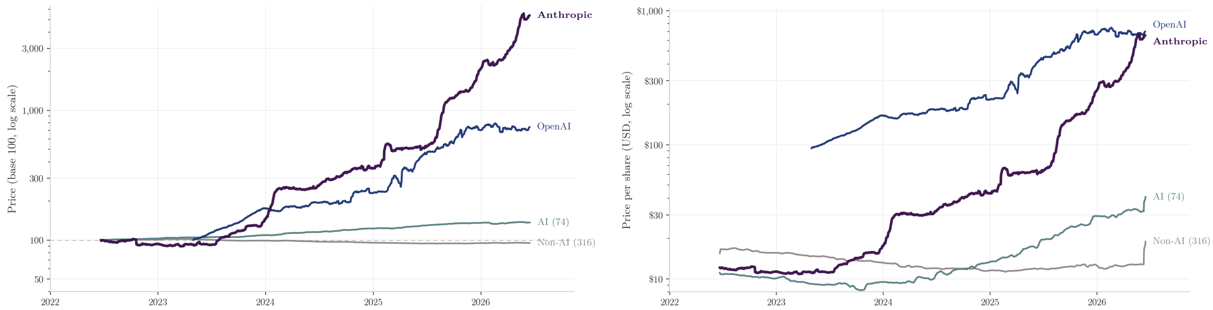
Figure 4.4: Daily evaluated price of 390 private companies on *NPM Price*, rebased to 100 at each company's first observation and shown on a logarithmic scale. Companies are ordered along the horizontal axis by final value, increasing to the right.

Within the sample, daily evaluated prices are available for 390 private companies, of which 74, close to 19%, fall under the *NPM Price AI and Machine Learning* sector. Their weight rises sharply with performance, since once the companies are ranked by final value 3 of the 4 names in the top 1% are *AI* companies, and *AI* companies make up 54% of the top decile while forming only 19% of the sample, almost 3 times their share. The same tilt is visible at the centre of the distribution, where the median *AI* company ends the period at 187 against a base of 100 while the median company outside the sector slips to 77, and 66% of *AI* names finish above their starting level against 40% of the rest.

The two groups also move at different speeds. Price velocity, the annual rate at which it grows, stays close to flat for the median company outside the sector at 0.94x a year, against 1.22x for the median *AI* company, 1.90x for *OpenAI*, and 2.72x for *Anthropic*. The frontier names are not only faster but accelerating, with *Anthropic's* annual growth rising from 1.59x in the first half of the window to 4.67x in the second, a movement absent from the rest of the market.

What the private data add is a direct view of the names the listed market cannot show. The same tilt found in the public prices returns here in a sharper form:

[17] Nasdaq Private Market [42].



(a) Rebased to 100 at each company's first observation.

(b) Evaluated price per share, in dollars.

Figure 4.5: Price paths of the *NPM Price* companies on a log scale, rebased to 100 in panel (a) and as price per share in panel (b). Group lines are cross-sectional medians.

value, growth and acceleration all gather in a thin band of *AI* companies, with the two frontier names pulling away from a market that is otherwise close to flat. The reading has its limits, since an evaluated price is an estimate rebuilt from scattered signals rather than a traded one, and the methods used earlier in the chapter were therefore set aside as ill suited to it. Yet the structure that survives this caution is the same one seen throughout, concentration rather than a broad rise, and it reaches the private market as plainly as the public one.

Read together, the public and private evidence of this chapter point the same way. Prices, valuation and the explosive root tests place the listed *AI* market in the early phase of a speculative episode, alongside the *dotcom* and *cryptocurrency* markets on the measures used here, while the log periodic fit locates whatever structure exists in the index rather than in any single stock. What separates this episode from the two that precede it is not its speed but its *narrowness*, a rise that gathers in a handful of names instead of lifting the market as a whole, and the same narrowness carries into the private valuations of the firms at its centre. A concentration of this kind leaves the market more exposed to the fortunes of a few companies than a broad advance would, and that exposure, rather than any single critical date, is the central finding of the analysis.

5 – Final Considerations and Conclusion

5.1 Towards the IPO

The two firms at the centre of this private market are now moving towards a public listing, and that listing fixes a date on the question of what their valuations rest upon. *Anthropic* filed a confidential prospectus with the *United States regulator* in June 2026, and *OpenAI* filed its own within days, so both could list as early as the autumn.[1] The accounts they bring to that listing make the distance between price and profit plain. *Anthropic* projects its first quarterly *operating profit*, of about $559 million, in the second quarter of 2026, on *revenue* of $10.9 billion. The company itself cautions that this figure depends on a temporary discount in its compute contract, and that it is unlikely to hold once the discount ends and planned spending resumes.[2] Even if that single quarter were repeated across a full year, the profit, set against the roughly $965 billion at which its shares are privately valued, would imply a *price-to-earnings* ratio near 430. That is several times the level at which established companies trade, and on the company's own guidance the true figure would be higher still, or there would be no earnings to divide and so no ratio at all.[3]

The same holds for *OpenAI*, which offers no such ratio at all, because it has no earnings to divide. Accounts obtained by the *Financial Times* show the company losing about $21 billion on *revenue* of about $13 billion,[4] and its own projections do not reach a profit before the end of the decade.[5] At a private valuation near $852 billion,[6] it would list as one of the most valuable companies in the world, while still losing money on a vast scale. The value the private market places on these two firms, and the value the coming listings will ask the public to accept, is therefore a claim on profits that do not yet exist. This is the same gap between price and present earnings that the rest of this study has traced, first in the listed market and then in the private one, and it is the same gap that marked the *dotcom boom* and the *cryptocurrency cycle* before it.

5.2 External Factors

So far the analysis has stayed inside the market, reading speculation from prices and valuations alone. Those figures rest on a base that the price series cannot show, and that base is as narrow as the financial picture around it. The value of the listed AI names, and of the private firms behind them, depends in the end on a single physical input, which is *computation*. But that computation depends in turn on a chain of suppliers so concentrated that almost every link is held by one company or

[1] Bloomberg News [6]. [2] CNBC [17]. [3] CNBC [16]. [4] Fortune [26]. [5] The Information [55].
[6] CNBC [18].

one place. The concentration this study has measured in capital has a twin in the physical world, and the second is arguably the more fragile of the two.

For more than a decade, progress in artificial intelligence has been carried by the steady growth of computing power. The computation used to train the leading models has roughly doubled every six months since the early 2010s. Across 12 years that is a rise of more than 16 million times, far faster than the older pace set by improvements in the chips themselves.[7] Demand for AI is therefore, at one step removed, demand for *chips*, and the financial boom of the earlier chapters rests on a physical input that must keep rising if the promise behind the valuations is to hold.

That input rests on a remarkably thin supply chain. *Nvidia* designs the *processors* on which most large models are trained, but it does not make them. The *chips* are made by a single firm, the *Taiwan Semiconductor Manufacturing Company* (TSMC), which produces close to 90% of the world's most advanced *logic chips* and about a third of all new computing power added worldwide each year.[8] The chain then narrows further. Only three companies in the world can make chips at the leading edge, and the machines that draw the circuits, using extreme *ultraviolet lithography*, come from one supplier alone, the Dutch firm *ASML*. Every stage that ends in an *AI data centre* passes through a point held by a single hand.

The chain is not only narrow. It is also fixed in one place, and that place is among the most contested on earth. The advanced fabrication sits in *Taiwan*, an island that China claims and whose security rests on an uneasy balance, so much so that its chip industry is often called a *silicon shield*. The main plants stand on a *fault line* that produced a large earthquake in 1999.[9] Economic theory gives a clear reading of what a shock to such a point would do. When production is organised as a network, a disturbance to one sector that supplies many others does not stay local. It spreads through the links to the wider economy, so that a shock at a single important point can move the whole.[10] This is not only a theoretical claim. The *Great East Japan Earthquake* of 2011 travelled along supply chains from the stricken region to firms across the country, and is estimated to have lowered Japan's output by about 1.2 percentage points in the following year.[11] A comparable shock to chip fabrication would reach every company in the *AI* complex at once, and would carry the market built on them down with it.

A second physical limit sits beside the first, in the *energy* that the computation consumes. *Data centres* already use roughly 1% of the world's electricity, a share held down so far only by steady gains in efficiency.[12] The spread of *AI* is now

[7] Sevilla et al. [49]. [8] Miller [40]. [9] Miller [40]. [10] Acemoglu et al. [1]. [11] Carvalho et al. [10]. [12] Masanet et al. [39].

pushing that demand up. A single conversational system is estimated to use on the order of hundreds of megawatt hours a day, and each query draws far more power than an ordinary web search.[13] The growth that the valuations assume is limited not only by *chips* but by the power and the infrastructure needed to run them, and neither of these can be expanded as fast as capital can be raised.

These external factors place the financial concentration of this study on a physical base that is concentrated in the same way. The value of the listed *AI* names rests on one design house, one fabricator, one tool maker and one island, and on a power supply that is already under strain. A fall in those prices need not begin in the market at all. It could be set off by an event with no connection to any valuation, such as a blockade in the *Taiwan Strait*, an earthquake, a tightening of export controls, or a shortage of electricity. The price methods of previous chapters cannot see this risk, because it lies outside the price, yet it defines the surface on which the structure stands.

5.3 Findings and Conclusions

This study set out to ask whether the capital gathering around artificial intelligence is the start of a lasting change or the familiar shape of a speculative *bubble*, and to ask it in two markets at once: the private one, where the firms at the centre of the wave are still funded round by round, and the public one, where their listed neighbours trade every day. No single number can settle a question of this kind. Concentration is the clearest reason why. A market can be narrow simply because a few firms genuinely lead it, and a narrow market is not, on its own, a riskier one.[14] What this study has tried to show is more specific, and harder to dismiss: that a concentration this extreme begins to read as risk once measures of a quite different kind, drawn from *economics*, from *statistics* and from *physics*, all point to the same place. The question is therefore not how high any one measure stands, but whether instruments that share no method agree on where they point.

It is worth starting with the measures that do not separate *AI* at all, since they are left exactly as they fall. The valuation a round commands, set against the capital it raises, does not mark these firms out from the wider private market, which is unsurprising, since reported valuations sit above fair value across the whole of venture capital once preferential terms are counted.[15] The pace at which *AI* funding accelerates is much the same as the market's. The pure power law is rejected on the tail, as it is on almost any sample of this size. No single listed stock carries a clean log periodic signature. These flat results are not a weakness in the argument but the thing that keeps it honest, since they show that no one measure is quietly

[13] de Vries [21]. [14] Kritzman and Turkington [36]. [15] Gornall and Strebulaev [30].

carrying all the weight.

On the measures that do separate it, the private market sets *AI* apart from everything around it. Its funding is far more concentrated: the *AI Gini coefficient* stands at 0.926, against 0.817 for the *Control*, and alone among the three episodes studied here it climbs as the *boom* goes on, reaching 0.947 after 2022 while the control holds flat. The shape of its largest rounds is more extreme still, since the tail of its distribution follows a power law exponent of about 1.9, against 2.3 for the control, and below an exponent of 2 a handful of deals can outweigh all the rest combined. The money also arrives faster and is repriced faster. Annual funding into *AI* is now roughly 87 times its level of a decade ago, while the comparable market has scarcely moved, and the value placed on these firms is marked up about 3.4 times a year, against 1.8 elsewhere, not in larger steps but in quicker ones, round following round before the last has settled.

The listed market is a different kind of data, read with different tools, prices that move every day rather than valuations struck a few times in a decade, and it points the same way. The explosive root tests place the AI era *Nasdaq* in the same company as the *dotcom* and *cryptocurrency* markets. The *cyclically adjusted price-to-earnings* ratio of the broad index stands near 40, a height reached only once before, in the months before the 2000 crash, and on the *Nasdaq 100*, where a few technology names carry close to half the weight, it reaches about 54, up from 34 at the end of 2022. The climb is not the market's but the work of a few names: *Nvidia* rises at 90% a year, almost five times as fast as the 19% of the index it sits in. The most striking thing, though, is what sits at the centre of all this money. The firms drawing in the largest share of it, and carrying the highest valuations ever placed on private companies, are precisely the ones with the least to show for it in profit. *Anthropic*, on the most generous reading of its own accounts, would list at a *price-to-earnings ratio* near 430. *OpenAI* cannot be given such a ratio at all, since it earns no profit to divide, and is in fact losing about $21 billion on revenue of about $13 billion. The more capital a firm here commands, the less, in earnings, it has to justify it. What the valuations rest on is not present profit but a claim on profit that does not yet exist.

Set against the speculative episodes this study has used as its mirrors, the resemblance is closest not to the *dotcom boom*, which swelled in plain sight on the public exchanges, but to the *cryptocurrency* cycle, whose speculation ran beneath the visible market in private rounds. The likeness is almost exact, since the gap that separates *AI* from its control and the gap that separates *crypto* from its own differ only in the third decimal. That likeness holds for now, while the wave is still a private one. It is only once the coming listings move these firms onto the public exchanges that the episode could take on the shape of the *dotcom boom* as well, the speculation passing from private rounds into prices the whole market can see. This thesis opened with Feynman's image of a single pattern that holds at every scale of nature, magnified or shrunk but never changing its shape. That image was a method as much as a metaphor. If something is real, it shows up the same way no matter which tool you measure it with, and that is what happened here.

The same concentration appears everywhere you look. In the private market, the largest 1% of funding rounds hold 72% of all the money raised. In the stock market, a single company, *Nvidia*, is rising almost five times faster than the index around it. The two firms the whole boom rests on, *OpenAI* and *Anthropic*, are seeing their valuations grow two to three times a year, while the rest of the private market barely moves. And underneath all of it lies the hardware, where every *AI* company depends on a single chip designer, which depends on a single factory in *Taiwan*, which depends on a single *Dutch* firm for the machines that make the chips. The closer you look, the more concentrated it becomes, and the same pattern runs through the funding, the prices, and the physical supply chain alike.

A market can hold a concentration like this and still reward the people who bought into it. It would do so if the few names at its heart grow into the value the world has placed on them, and that possibility deserves to be put plainly. But a structure that gathers so much onto so small a base is, by that same measure, a fragile one. That fragility is hard to dismiss as chance, because so many independent readings agree, and not because any single one of them stands high. One measure standing high could be a coincidence; a dozen of them, built in fields that do not speak to one another and all pointing the same way, are not. Its fall would not need a broken promise or a missed quarter to begin. It could begin with the weather over a strait. It could begin with a tremor along a fault line. It could begin with a line drawn on a trade minister's map. The earlier bubbles this study compared itself to deflated slowly, because their weight was spread across an entire market. Here the weight sits on only a few companies, so nothing would slow the fall.

No single measure could have proved this. It took all of them, and all of them agreed.

5.4 Further Analysis for Future Studies and Development

The framework used here could be extended in four directions, each building on a tool this study already relied on.

Accuracy

The first would add a measure of reliability to the log periodic fits. The fitting procedure always returns a critical time, whether or not a real *bubble* is present, so the estimate is only as good as the confidence placed in it. Recent work trains a classifier on millions of fitted parameter sets to score how far each fit can be trusted, and folds that score into the crash signal to cut the rate of false alarms.[16]

Sentiment

The second would build a measure of *hype* directly, rather than reading it only from prices. How widely a technology spreads depends in part on how it is talked about, on the justifications that make backing it seem reasonable,[17] and that signal can be combined with the tone of news coverage to turn attention into a variable of its own.

Hyped Log Periodic Power Law

The third would feed that variable into the log periodic model itself. A recent study does exactly this, adding a *hype index* and a news *sentiment score* to the model so that genuine speculative phases can be told apart from price moves that simply lack attention.[18]

Machine Learning Modelling

The fourth would pass the full set of signals to a *machine learning model*. The measures kept separate here, concentration, tail weight, valuation, the explosive root tests and a hype variable, could serve together as inputs to a model trained to distinguish speculative episodes from structural growth, so that the agreement this thesis read across instruments by hand could instead be learned from the data.

[16] Lee et al. [37]. [17] Green Jr. [31]. [18] Cao et al. [9].

Appendix A – Private Companies Analysis

Configuration

Everything below is driven by this block: file paths, the firm-identifier column, the sample window, the temporal cut and the simulation sizes. Edit here, run everywhere.

```r
path <- "~/Desktop/ECHO.xlsx"
ai_sheet <- "ECHO AI"
ctrl_sheet <- "ECHO CONTROL AI"
firm_col <- "Company Name"

YEARS <- c(2016, 2026)  # sample window: rows with a dated round outside are dropped
CUT   <- 2021           # last pre-ChatGPT year: <= CUT is "pre", > CUT is "post"
NSIM  <- 2000           # bootstrap replications (GoF and confidence intervals)
NDK   <- 20000          # Monte Carlo draws for the Dragon-King test
```

Core functions

The Clauset toolkit, written by hand so that billion-scale deal values do not trip the internal limits of `poweRlaw`.

```r
# Maximum-likelihood alpha for a given threshold (xmin) (Clauset et al. 2009)
alpha_mle <- function(x, xmin) {
  x <- x[x >= xmin]
  1 + length(x) / sum(log(x / xmin))
}

# Kolmogorov-Smirnov distance between the empirical tail and the fitted power law
ks_stat <- function(x, xmin, a) {
  x <- sort(x[x >= xmin])
  n <- length(x)
  max(abs((1:n) / n - (1 - (x / xmin)^(-(a - 1)))))
}

# Draw n values from a continuous power law (inverse-CDF method)
rplcon <- function(n, xmin, a) {
  xmin * (1 - runif(n))^(-1 / (a - 1))
}

# Gini index: 0 = capital spread evenly, 1 = all capital in a single deal
gini <- function(x) {
  x <- sort(x[x > 0])
  n <- length(x)
  2 * sum((1:n) * x) / (n * sum(x)) - (n + 1) / n
}

# Threshold scan: keep the xmin that minimises the KS distance.
# Log-spaced grid from 1m to 10bn; tails thinner than `min_tail` are skipped.
fit_powerlaw <- function(vals, min_tail = 50) {

  grid <- 10^seq(6, 10, length.out = 100)
  grid <- grid[grid < max(vals)]
```

```
  best <- list(ks = Inf)

  for (xm in grid) {                    # for each candidate threshold xm in the grid

    if (sum(vals >= xm) < min_tail)     # tail too thin? skip to the next candidate value
      next

    a <- alpha_mle(vals, xm)            # estimate alpha at this threshold
    d <- ks_stat(vals, xm, a)           # KS distance between empirical tail and power law

    if (is.finite(d) && d < best$ks)    # best one so far?
      best <- list(xmin = xm, alpha = a, ks = d, n_tail = sum(vals >= xm))
  }

  best
}
```

Data

Zeros are undisclosed amounts, not real values, so they are dropped along with missing entries and one impossible outlier. Raw exports repeat each round once per investor, so rounds are deduplicated, and dated rounds outside the sample window are dropped.

```
ai <- read_excel(path, sheet = ai_sheet)
ctrl <- read_excel(path, sheet = ctrl_sheet)

ai <- ai %>% transmute(id   = as.character(.data[[firm_col]]),
                       val  = suppressWarnings(as.numeric(`Round Total Value`)),
                       date = as_date(`Round Date`)) %>%
          filter(!is.na(val), val > 0, val < 2e11)  %>%
          distinct(id, val, date) %>%                        # one row per round
          mutate(year = year(date))  %>%
          filter(is.na(year) | between(year, YEARS[1], YEARS[2]))

ctrl <- ctrl %>% transmute(id   = as.character(.data[[firm_col]]),
                           val  = suppressWarnings(as.numeric(`Round Total Value`)),
                           date = as_date(`Round Date`)) %>%
            filter(!is.na(val), val > 0, val < 2e11)  %>%
            distinct(id, val, date) %>%                        # one row per round
            mutate(year = year(date))  %>%
            filter(is.na(year) | between(year, YEARS[1], YEARS[2]))


cat("AI:", nrow(ai), "rows  |  CONTROL:", nrow(ctrl), "rows\n")
```

```
## AI: 11310 rows  |  CONTROL: 52280 rows
```

1. Down-rounds and top-k share

A down-round is a round priced below the firm's previous one, the standard VC sign of cooling. Top-k share is the slice of capital captured by the largest deals. Both read alongside alpha and Gini.

```
# Share of rounds priced below the same firm's previous round
down_round <- function(d, name) {
```

```r
  down <- d  %>% filter(!is.na(id), !is.na(date)) %>%
                 arrange(date)  %>%
                 group_by(id) %>%
                 mutate(down = val < lag(val)) %>% # TRUE when a round is priced below the firm's
      previous one
                 ungroup() %>%
                 filter(!is.na(down)) %>%
                 pull(down)

  cat(paste0("[",name,"] down-rounds: ", round(100 * mean(down), 1), "% (n=", length(down), ")\n"
    ))
}

# Same split before vs after the cut
down_round_time <- function(d, name) {

  z <- d %>% filter(!is.na(id), !is.na(date)) %>%
             arrange(date) %>%
             group_by(id) %>%
             mutate(down = val < lag(val)) %>%
             ungroup() %>%
             filter(!is.na(down))

  pre  <- z %>% filter(year <= CUT) %>% pull(down) #cut = 2021
  post <- z %>% filter(year >  CUT) %>% pull(down)

  cat(paste0("[", name, "] pre: ", round(100 * mean(pre), 1), "% (n=", length(pre),
           ") -> post: ", round(100 * mean(post), 1), "% (n=", length(post), ")\n"))
}

# Share of total capital held by the top 1% / 5% / 10% of deals
top_k <- function(vals, name) {

  v <- sort(vals, decreasing = TRUE)
  tot <- sum(v)
  n <- length(v)

  cat(paste0("[", name, "] top 1%=", round(100 * sum(v[1:ceiling(.01 * n)]) / tot, 1),
             "%  top 5%=",  round(100 * sum(v[1:ceiling(.05 * n)]) / tot, 1),
             "%  top 10%=", round(100 * sum(v[1:ceiling(.10 * n)]) / tot, 1), "%\n"))
}

down_round(ai, "AI")
```

```
## [AI] down-rounds: 24.3% (n=4092)
```

```r
down_round(ctrl, "CONTROL")
```

```
## [CONTROL] down-rounds: 31.3% (n=24560)
```

```r
cat("\n")
down_round_time(ai, "AI")
```

```
## [AI] pre: 24% (n=1494) -> post: 24.4% (n=2598)
```

```r
down_round_time(ctrl, "CONTROL")
```

```
## [CONTROL] pre: 30.6% (n=13657) -> post: 32.3% (n=10903)
```

```r
cat("\n")
```

```r
top_k(ai$val, "AI")
```

```
## [AI] top 1%=72%  top 5%=83.9%  top 10%=89.2%
```

```r
top_k(ctrl$val, "CONTROL")
```

```
## [CONTROL] top 1%=37.4%  top 5%=61.2%  top 10%=72.9%
```

2. *Gini, overall and over time*

A standard, scale-invariant concentration measure; read alongside alpha, not instead of it.

```r
pre_ai  <- ai   %>% filter(year <= CUT) %>% pull(val)
post_ai <- ai   %>% filter(year >  CUT) %>% pull(val)
pre_ct  <- ctrl %>% filter(year <= CUT) %>% pull(val)
post_ct <- ctrl %>% filter(year >  CUT) %>% pull(val)

cat(paste0("Overall : AI ", round(gini(ai$val), 3), "  |  CONTROL ", round(gini(ctrl$val), 3), "\
    n"))
```

```
## Overall : AI 0.926  |  CONTROL 0.817
```

```r
cat(paste0("AI      : pre ", round(gini(pre_ai), 3), " -> post ", round(gini(post_ai), 3), "\n"))
```

```
## AI      : pre 0.795 -> post 0.947
```

```r
cat(paste0("CONTROL : pre ", round(gini(pre_ct), 3), " -> post ", round(gini(post_ct), 3), "\n"))
```

```
## CONTROL : pre 0.811 -> post 0.823
```

3. *Herfindahl-Hirschman index*

Each firm's rounds are pooled over the window, converted to shares of group capital, squared and summed. The reciprocal 1/HHI is the number of equally sized firms that would produce the same concentration; Gini reads inequality across deals, HHI reads dominance across firms.

```r
# Firm-level HHI: sum of squared shares of total capital raised
hhi <- function(d) {
  tot <- d %>% filter(!is.na(id)) %>%
             group_by(id) %>%
             summarise(tot = sum(val), .groups = "drop") %>%
             pull(tot)
  sum((tot / sum(tot))^2)
}

h_ai <- hhi(ai)
h_ct <- hhi(ctrl)

cat(paste0("Overall : AI ", signif(h_ai, 4), " (1/HHI=", round(1 / h_ai, 1),
           ", firms=", n_distinct(ai$id), ")  |  CONTROL ", signif(h_ct, 4),
           " (1/HHI=", round(1 / h_ct, 1), ", firms=", n_distinct(ctrl$id), ")\n"))
```

```
## Overall : AI 0.118 (1/HHI=8.5, firms=7218)  |  CONTROL 0.001783 (1/HHI=560.9, firms=27720)
```

```r
cat(paste0("AI      : pre ", signif(hhi(ai %>% filter(year <= CUT)), 4),
           " -> post ", signif(hhi(ai %>% filter(year > CUT)), 4), "\n"))
```

```
## AI      : pre 0.01005 -> post 0.1599
```

```r
cat(paste0("CONTROL : pre ", signif(hhi(ctrl %>% filter(year <= CUT)), 4),
           " -> post ", signif(hhi(ctrl %>% filter(year > CUT)), 4), "\n"))
```

```
## CONTROL : pre 0.001879 -> post 0.004992
```

4. Skewness and kurtosis

```r
# Population moments on log deal values; kurtosis is reported in excess form
skew <- function(x) { m <- mean(x); mean((x - m)^3) / mean((x - m)^2)^1.5 }
kurt <- function(x) { m <- mean(x); mean((x - m)^4) / mean((x - m)^2)^2 - 3 }

shape <- function(vals, name) {
  lx <- log(vals)
  cat(paste0("[", name, "] skew=", round(skew(lx), 3),
             "  excess kurt=", round(kurt(lx), 3), "  (n=", length(lx), ")\n"))
}

shape(ai$val, "AI")
```

```
## [AI] skew=-0.249  excess kurt=0.777  (n=11310)
```

```r
shape(ctrl$val, "CONTROL")
```

```
## [CONTROL] skew=-0.565  excess kurt=1.075  (n=52280)
```

```r
# Same moments before vs after the cut (vectors from section 2)
shape(pre_ai,  "AI   pre ")
shape(post_ai, "AI   post")
shape(pre_ct,  "CTRL pre ")
shape(post_ct, "CTRL post")
```

```
## [AI   pre ] skew=-0.604  excess kurt=1.184  (n=3533)
## [AI   post] skew=-0.098  excess kurt=0.751  (n=7777)
## [CTRL pre ] skew=-0.596  excess kurt=1.323  (n=29855)
## [CTRL post] skew=-0.527  excess kurt=0.788  (n=22425)
```

5. Power-law fit and the alpha gap

For each group: the KS-optimal threshold and the MLE exponent, with a nonparametric bootstrap confidence interval for alpha (resampling the tail, NSIM draws). The same resampling then bootstraps the gap alpha(AI) - alpha(CONTROL): if its 95% interval excludes zero, the two tails differ.

```r
# Power-law fit: KS-optimal xmin and MLE alpha, plus a nonparametric
# bootstrap CI for alpha (resampling the tail, NSIM draws)
fit_tail <- function(vals, name) {

  f      <- fit_powerlaw(vals)
  tail_x <- vals[vals >= f$xmin]
  n      <- length(tail_x)

  # bootstrap: resample the n tail values with replacement, refit alpha
  a_boot <- replicate(NSIM, alpha_mle(sample(tail_x, n, replace = TRUE), f$xmin))
  ci     <- quantile(a_boot, c(.025, .975))
  sigma  <- sd(log(vals)) # log-normal dispersion, used later
```

```r
  cat(paste0("[", name, "] alpha=", round(f$alpha, 3),"  CI[", round(ci[1], 3),
             ", ", round(ci[2], 3),"]  xmin=", round(f$xmin), "  n_tail=", n, "\n"))
  list(fit = f, ci = ci, sigma = sigma, tail = tail_x)
}

res_ai   <- fit_tail(ai$val,   "AI")
```

```
## [AI] alpha=1.911  CI[1.876, 1.948]  xmin=16297508  n_tail=2221
```

```r
res_ctrl <- fit_tail(ctrl$val, "CONTROL")
```

```
## [CONTROL] alpha=2.296  CI[2.248, 2.348]  xmin=79248290  n_tail=2552
```

```r
# Nonparametric bootstrap of the gap alpha(AI) - alpha(CONTROL):
# resample both tails, take the difference of the two MLE alphas, repeat
gap_boot <- replicate(NSIM, {
  alpha_mle(sample(res_ai$tail,   length(res_ai$tail),   TRUE), res_ai$fit$xmin) -
  alpha_mle(sample(res_ctrl$tail, length(res_ctrl$tail), TRUE), res_ctrl$fit$xmin)
})

ci_gap <- quantile(gap_boot, c(.025, .975))
cat("\n")
cat(paste0("alpha gap (AI - CONTROL): ", round(mean(gap_boot), 3),
           "  CI[", round(ci_gap[1], 3), ", ", round(ci_gap[2], 3),
           "]  P(gap >= 0) = ", round(mean(gap_boot >= 0), 4), "\n"))
```

```
## alpha gap (AI - CONTROL): -0.386  CI[-0.448, -0.322]  P(gap >= 0) = 0
```

6. Goodness-of-fit (parametric bootstrap)

How often would a true power law of the fitted size produce a KS distance at least as large as the observed one? Simulate NSIM synthetic tails, refit each, and read off the share that exceed the data. A high p means the power law is not rejected.

```r
# GoF: how often would a true power law produce a KS distance this large?
gof_test <- function(vals, res, name) {

  # pull the fit found earlier: f$xmin (threshold) and f$alpha (exponent)
  f      <- res$fit
  tail_x <- vals[vals >= f$xmin]    # keep only the tail: values at/above xmin
  n      <- length(tail_x)          # how many points are in the tail

  # observed misfit: KS distance between the real tail and the fitted power law
  d_obs <- ks_stat(vals, f$xmin, f$alpha)

  # --- NULL WORLD ---------------------------------------------------------------
  # We can't collect 2000 new datasets, so we manufacture them: draw NSIM fake
  # tails from a power law we KNOW is true (same xmin and alpha as the fit).
  # For each fake tail we refit alpha from scratch (alpha_mle) and measure its
  # own KS distance. This builds the distribution of KS values you'd see when
  # the power law really IS the right model -- i.e. how far KS naturally drifts
  # just from random sampling, with no real misfit involved.
  d_sim <- replicate(NSIM, { s <- rplcon(n, f$xmin, f$alpha)          # one fake tail
                             ks_stat(s, f$xmin, alpha_mle(s, f$xmin)) }) # its own KS

  # --- P-VALUE ------------------------------------------------------------------
  # Compare the real misfit (d_obs) against that null distribution: what fraction
  # of the "honest" power-law tails miss at least as badly as your data did?
```

```
  # If lots of them do (HIGH p), your data's misfit is nothing unusual for a true
  # power law -> the model is NOT rejected. If almost none do (LOW p, <0.1), your
  # data strays further than a real power law ever would -> reject the power law.
  # NOTE: this is backwards from the usual test -- here a HIGH p is GOOD.
  p_gof <- mean(d_sim >= d_obs)

  cat(paste0("[", name, "] GoF p=", round(p_gof, 3), "  (KS_obs=", round(d_obs, 4), ")\n"))
  invisible(p_gof)
}

gof_test(ai$val,   res_ai,   "AI")
```

```
## [AI] GoF p=0  (KS_obs=0.0351)
```

```
gof_test(ctrl$val, res_ctrl, "CONTROL")
```

```
## [CONTROL] GoF p=0.002  (KS_obs=0.0291)
```

7. Alternative tails, log-normal and exponential

Two non-nested likelihood-ratio comparisons on the same tail: power law against log-normal (Vuong) and against exponential. A positive R favours the power law, a negative one the alternative; $|R| < 2$ means the data cannot tell them apart. sigma is the log-normal dispersion, sd(log(x)).

```
compare_tails <- function(vals, res, name) {

  f <- res$fit                          # the earlier power-law fit (holds xmin and alpha)

  m_pl <- conpl$new(vals)               # create a continuous power-law model object
  m_pl$setXmin(f$xmin)                  # set its threshold to our fitted xmin
  m_pl$setPars(f$alpha)                 # set its exponent to our fitted alpha

  m_ln <- conlnorm$new(vals)            # create a continuous log-normal model object
  m_ln$setXmin(f$xmin)                  # same threshold, so it's judged on the same tail
  m_ln$setPars(estimate_pars(m_ln))     # estimate its best-fit parameters by MLE

  m_ex <- conexp$new(vals)              # create a continuous exponential model object
  m_ex$setXmin(f$xmin)                  # same threshold again
  m_ex$setPars(estimate_pars(m_ex))     # estimate its best-fit rate by MLE

  v_ln <- compare_distributions(m_pl, m_ln)   # Vuong test: power law vs log-normal
  v_ex <- compare_distributions(m_pl, m_ex)   # Vuong test: power law vs exponential

  cat(paste0("[", name, "] sigma=", round(res$sigma, 3), "\n"))   # print log-normal dispersion
     sigma
  cat(paste0("     vs log-normal : R=", round(v_ln$test_statistic, 2),
           " (p=", round(v_ln$p_two_sided, 3), ")",
           if (abs(v_ln$test_statistic) < 2) "  - indistinguishable" else "", "\n"))
  cat(paste0("     vs exponential: R=", round(v_ex$test_statistic, 2),
           " (p=", round(v_ex$p_two_sided, 3), ")",
           if (abs(v_ex$test_statistic) < 2) "  - indistinguishable" else "", "\n"))
}

compare_tails(ai$val,   res_ai,   "AI")
```

```
## [AI] sigma=2.101
##      vs log-normal : R=-1.17 (p=0.24)  - indistinguishable
##      vs exponential: R=3.66 (p=0)
```

```r
compare_tails(ctrl$val, res_ctrl, "CONTROL")
```

```
## [CONTROL] sigma=2.042
##      vs log-normal : R=-0.47 (p=0.636)  - indistinguishable
##      vs exponential: R=16.49 (p=0)
```

8. Temporal split, before vs after 01/01/2022

Same fit, run separately on each sub-period (cut at end of 2021). A lower alpha after the cut means the tail has grown heavier.

```r
fit_slim <- function(vals, name) {

  f      <- fit_powerlaw(vals)
  tail_x <- vals[vals >= f$xmin]

  a_boot <- replicate(NSIM, alpha_mle(sample(tail_x, length(tail_x), TRUE), f$xmin))
  ci     <- quantile(a_boot, c(.025, .975))

  cat(paste0("[", name, "] alpha=", round(f$alpha, 3),
             "  CI[", round(ci[1], 3), ", ", round(ci[2], 3),
             "]  n=", length(vals), "  n_tail=", f$n_tail, "\n"))
}

fit_slim(ai   %>% filter(year <= CUT) %>% pull(val), "AI   pre ")
```

```
## [AI   pre ] alpha=2.052  CI[1.985, 2.123]  n=3533  n_tail=678
```

```r
fit_slim(ai   %>% filter(year >  CUT) %>% pull(val), "AI   post")
```

```
## [AI   post] alpha=1.884  CI[1.837, 1.932]  n=7777  n_tail=1403
```

```r
fit_slim(ctrl %>% filter(year <= CUT) %>% pull(val), "CTRL pre ")
```

```
## [CTRL pre ] alpha=2.387  CI[2.296, 2.491]  n=29855  n_tail=701
```

```r
fit_slim(ctrl %>% filter(year >  CUT) %>% pull(val), "CTRL post")
```

```
## [CTRL post] alpha=2.219  CI[2.17, 2.274]  n=22425  n_tail=2009
```

9. Velocity and acceleration

Velocity is the annualised growth factor between a firm's consecutive rounds; the step-up is the raw multiple. Acceleration asks whether a firm's later steps run faster than its earlier ones, a crude super-exponential check, Sornette-flavoured.

```r
step_metrics <- function(d, name) {

  steps <- d %>% filter(!is.na(id), !is.na(date)) %>%
               arrange(date) %>%
               group_by(id) %>%
               mutate(yrs   = c(NA, diff(as.numeric(date)) / 365.25),
                      ratio = val / lag(val)) %>%
               filter(yrs > 0) %>%    # skip same-day rounds
               mutate(vel = ratio^(1 / yrs)) %>%
               filter(is.finite(vel), vel > 0) %>%
               ungroup()

  acc <- steps %>% group_by(id) %>%
```

```r
                    filter(n() >= 2) %>%
                    summarise(accelerating =
                                {v <- vel
                                 m <- ceiling(length(v) / 2)
                                 median(v[(m + 1):length(v)]) > median(v[1:m])}, .groups = "drop")
      %>%
                    pull(accelerating)

  cat(paste0("[", name, "] step-up=", round(median(steps$ratio), 2),
             "x | velocity=", round(median(steps$vel), 2),
             " x/yr (n=", nrow(steps),
             ") | accelerating=", round(100 * mean(acc), 1),
             "% (n=", length(acc), ")\n"))
  list(vel = steps$vel, acc = acc)
}

s_ai <- step_metrics(ai,   "AI")
```

```
## [AI] step-up=2.11x | velocity=1.8 x/yr (n=4043) | accelerating=39.3% (n=914)
```

```r
s_ct <- step_metrics(ctrl, "CONTROL")
```

```
## [CONTROL] step-up=1.78x | velocity=1.51 x/yr (n=24342) | accelerating=39.8% (n=5680)
```

```r
# Sector differences: velocity levels (Mann-Whitney) and acceleration shares
p_vel <- wilcox.test(s_ai$vel, s_ct$vel)$p.value
p_acc <- prop.test(c(sum(s_ai$acc), sum(s_ct$acc)),
                   c(length(s_ai$acc), length(s_ct$acc)))$p.value

cat(paste0("\nVelocity AI vs CONTROL: p=", format.pval(p_vel, digits = 3),
           "  |  acceleration share: p=", format.pval(p_acc, digits = 3), "\n"))
```

```
##
## Velocity AI vs CONTROL: p=<2e-16  |  acceleration share: p=0.79
```

10. *Post-money valuations*

Round value measures capital raised, not price. Where the post-money column is populated enough, the same tail fit is repeated on valuations as a cross-check. Coverage is sparse, supporting evidence only.

```r
post_fit <- function(sheet, name) {

  pv <- read_excel(path, sheet = sheet)

  pv <- pv %>% transmute(pv = suppressWarnings(as.numeric(`Post-Valuation`))) %>%
             filter(!is.na(pv), pv > 0, pv < 2e11) %>%
             pull(pv)

  if (length(pv) < 100) {
    cat(paste0("[", name, "] post-money: only ", length(pv), " values - skipped\n"))
    return(invisible(NULL))
  }

  f  <- fit_powerlaw(pv, min_tail = 30)
  tl <- pv[pv >= f$xmin]
  b  <- replicate(NSIM, alpha_mle(sample(tl, length(tl), TRUE), f$xmin))
  ci <- quantile(b, c(.025, .975))
```

```r
  cat(paste0("[", name, "] post-money: n=", length(pv),
             "  alpha=", round(f$alpha, 3),
             "  CI[", round(ci[1], 3), ", ", round(ci[2], 3),
             "]  n_tail=", f$n_tail, "\n"))
}

post_fit(ai_sheet,   "AI")
```

```
## [AI] post-money: n=845  alpha=2.071  CI[1.871, 2.362]  n_tail=59
```

```r
post_fit(ctrl_sheet, "CONTROL")
```

```
## [CONTROL] post-money: n=2475  alpha=2.324  CI[2.175, 2.509]  n_tail=213
```

Summary

```
## PRIVATE LEG - RECAP
## -------------------
## alpha : AI 1.911 [1.876, 1.948]  vs  CONTROL 2.296 [2.248, 2.348]
## gap   : -0.386  CI[-0.448, -0.322]
## sigma : AI 2.101  vs  CONTROL 2.042
## Gini  : AI 0.926  vs  CONTROL 0.817
##
## Bubble hypothesis: alpha(AI) < alpha(CONTROL) and Gini(AI) > Gini(CONTROL);
## a gap interval excluding zero confirms the difference is significant.
```

Appendix B – Public Markets Analysis

Data

Daily adjusted close prices from Yahoo Finance via `tidyquant`, split and dividend adjusted, over the last 10 years: the Mag-7, 2 AI ETFs (AIQ, Global X Artificial Intelligence and Technology; CHAT, Roundhill Generative AI) and the Nasdaq Composite as market proxy. BTQK does not resolve on Yahoo because it is an LSEG RIC, and the proprietary LSEG AI index is not on Yahoo either, so the market here is `^IXIC`. The run is dated 5 July 2026: equity series stop at the close of 2 July, crypto series run to 5 July.

```
lookup <- tribble(~name,        ~ticker,
                  "Alphabet",  "GOOGL",
                  "Amazon",    "AMZN",
                  "Apple",     "AAPL",
                  "Meta",      "META",
                  "Microsoft", "MSFT",
                  "Nvidia",    "NVDA",
                  "Tesla",     "TSLA",
                  "AIQ",       "AIQ",      # Global X AI & Technology
                  "CHAT",      "CHAT",     # Roundhill Generative AI
                  "Nasdaq",    "^IXIC")    # Nasdaq Composite (market proxy)

from <- Sys.Date() %m-% years(10) #10 years ago from today
raw  <- tq_get(lookup$ticker, from = from, to = Sys.Date(), get = "stock.prices")

prices <- raw %>% left_join(lookup, by = c("symbol" = "ticker")) %>%
                  select(name, date, adjusted) %>%
                  arrange(name, date)

# import check
prices %>% group_by(name) %>%
           summarise(obs = n(), from = min(date), to = max(date),
                     n_na = sum(is.na(adjusted)), .groups = "drop") %>%
           print(n = Inf)
```

```
## # A tibble: 10 x 5
##    name        obs from       to          n_na
##    <chr>     <int> <date>     <date>     <int>
##  1 AIQ        2043 2018-05-16 2026-07-02     0
##  2 Alphabet   2513 2016-07-05 2026-07-02     0
##  3 Amazon     2513 2016-07-05 2026-07-02     0
##  4 Apple      2513 2016-07-05 2026-07-02     0
##  5 CHAT        783 2023-05-18 2026-07-02     0
##  6 Meta       2513 2016-07-05 2026-07-02     0
##  7 Microsoft  2513 2016-07-05 2026-07-02     0
##  8 Nasdaq     2513 2016-07-05 2026-07-02     0
##  9 Nvidia     2513 2016-07-05 2026-07-02     0
## 10 Tesla      2513 2016-07-05 2026-07-02     0
```

1. Velocity and acceleration

Velocity is the trailing 1 year log growth, $v_t = \ln P_t - \ln P_{t-h}$ with $h = 252$ trading days. Acceleration is its second difference, $a_t = v_t - v_{t-h}$; a sustained positive a marks super-exponential growth.

```r
h <- 252L   # days corresponding to one year of trade

dyn <- prices %>% filter(!is.na(adjusted)) %>%
                  group_by(name) %>%
                  arrange(date, .by_group = TRUE) %>%
                  mutate(logp = log(adjusted),
                         velocity = logp - lag(logp, h),
                          acceleration = velocity - lag(velocity, h)) %>%
                  ungroup()
```

2. Pre and post boom differences

Mean velocity and acceleration on each side of the cut, 30 November 2022, the ChatGPT release. Mean velocities are converted to compound annual rates in percent. CHAT lists in May 2023, so its pre columns are empty.

```r
boom_date <- as.Date("2022-11-30")

pre_post <- dyn %>% mutate(period = if_else(date < boom_date, "pre", "post")) %>%
                    group_by(name, period) %>%
                    summarise(vel_pct = (exp(mean(velocity, na.rm = TRUE)) - 1) * 100,
                              acc     =  mean(acceleration, na.rm = TRUE),
                              .groups = "drop") %>%
                    pivot_wider(names_from = period, values_from = c(vel_pct, acc)) %>%
                    mutate(vel_change = vel_pct_post - vel_pct_pre,
                           acc_change = acc_post - acc_pre) %>%
                    mutate(across(where(is.numeric), \(x) round(x, 2))) %>%
                    arrange(desc(vel_change))

print(pre_post, width = Inf)
```

```
## # A tibble: 10 x 7
##    name      vel_pct_post vel_pct_pre acc_post acc_pre vel_change acc_change
##    <chr>            <dbl>       <dbl>    <dbl>   <dbl>      <dbl>      <dbl>
##  1 Nvidia            89.8       48.9      0.08   -0.17      40.8        0.25
##  2 Meta              42.5        6.86     0.18   -0.18      35.6        0.36
##  3 AIQ               24.4       14.4      0.15   -0.13       9.98       0.28
##  4 Alphabet          28.3       21.2      0.14   -0.05       7.15       0.19
##  5 Nasdaq            18.7       16.2      0.1    -0.08       2.45       0.17
##  6 Amazon            17.2       23.7      0.11   -0.15      -6.56       0.25
##  7 Microsoft         16.2       33.6      0.01   -0.07     -17.4        0.07
##  8 Apple             15.6       36.7      0      -0.04     -21.1        0.04
##  9 Tesla             10.9       68.7      0.06   -0.04     -57.8        0.1
## 10 CHAT              44.4       NA        0.28   NA         NA         NA
```

3. LPPL on the Nasdaq and Mag-7

Johansen-Ledoit-Sornette fit on split-adjusted close prices, not total return, over the AI run-up window from January 2023. The linear parameters are concentrated out by least squares in the Filimonov and Sornette reformulation, and the nonlinear triple (tc, m, omega) is searched by multi-start L-BFGS-B inside the Sornette bounds. Fits that pin m or omega at a search bound are degenerate and flagged, together with a positive B and a damping D below 1.

```r
tickers <- c("^IXIC","GOOGL","AMZN","AAPL","META","MSFT","NVDA","TSLA")
labels  <- c("Nasdaq","Alphabet","Amazon","Apple","Meta","Microsoft","Nvidia","Tesla")
lppl_start <- as.Date("2023-01-01")    # window start, try also 2022-11-01
m_bounds   <- c(0.1, 0.9)              # Sornette, power-law exponent
w_bounds   <- c(4.8, 13)               # Sornette, log-periodic frequency
tc_frac    <- 0.35                     # tc within this fraction of the span ahead
raw    <- tq_get(tickers, from = "2016-01-01", to = Sys.Date(), get = "stock.prices")
prices <- raw %>% mutate(asset = labels[match(symbol, tickers)]) %>%
                  select(asset, date, close) %>%
                  drop_na(close)
# SSE for a trial (tc, m, omega); the linear part is solved by least squares
lppl_sse <- function(par, tt, yy) {

  tc <- par[1]; m <- par[2]; w <- par[3]
  dt <- tc - tt
  if (any(dt <= 1e-6)) return(1e12)
  f <- dt^m; ln <- log(dt)
  X <- cbind(1, f, f*cos(w*ln), f*sin(w*ln))
  fit <- tryCatch(.lm.fit(X, yy), error = function(e) NULL)
  if (is.null(fit)) return(1e12)
  sum(fit$residuals^2)

}
# multi-start; returns parameters, fitted log price and diagnostics
fit_lppl <- function(tt, yy) {

  tmax <- max(tt); span <- tmax - min(tt)
  grid <- expand.grid(tc = seq(tmax + 0.02*span, tmax + 0.30*span, length.out = 8),
                      w  = seq(w_bounds[1] + 0.2, w_bounds[2] - 0.2, length.out = 8),
                      m  = c(0.2, 0.5, 0.8))

  best <- list(value = Inf)

  for (i in seq_len(nrow(grid))) {
    r <- tryCatch(
      optim(c(grid$tc[i], grid$m[i], grid$w[i]), lppl_sse, tt = tt, yy = yy,
            method = "L-BFGS-B",
            lower = c(tmax + 1e-3, m_bounds[1], w_bounds[1]),
            upper = c(tmax + tc_frac*span, m_bounds[2], w_bounds[2])),
      error = function(e) NULL)
    if (!is.null(r) && r$value < best$value) best <- r
  }

  tc <- best$par[1]
  m <- best$par[2]
  w <- best$par[3]

  dt <- tc - tt
  f <- dt^m
  ln <- log(dt)

  X  <- cbind(1, f, f*cos(w*ln), f*sin(w*ln))
  co <- .lm.fit(X, yy)$coefficients          # A, B, C1, C2
```

```
  A <- co[1]; B <- co[2]; C1 <- co[3]; C2 <- co[4]
  C    <- sqrt(C1^2 + C2^2)
  yhat <- as.vector(X %*% co)
  R2   <- 1 - sum((yy - yhat)^2) / sum((yy - mean(yy))^2)
  D    <- (m * abs(B)) / (w * C)             # Sornette damping, D >= 1
  list(tc = tc, m = m, w = w, A = A, B = B, C = C, R2 = R2, D = D, yhat = yhat)
}
# fit each series on the window
results <- list()
curves <- list()
for (a in labels) {
  d  <- prices %>% filter(asset == a, date >= lppl_start) %>% arrange(date)
  tt <- decimal_date(d$date)
  yy <- log(d$close)
  ft <- fit_lppl(tt, yy)
  flag <- c(if (ft$m <= m_bounds[1] + 0.02) "m@low"  else NA,
            if (ft$m >= m_bounds[2] - 0.02) "m@high" else NA,
            if (ft$w <= w_bounds[1] + 0.05 || ft$w >= w_bounds[2] - 0.05) "omega@bound" else NA,
            if (ft$B >= 0) "B>0" else NA,
            if (ft$D < 1) "D<1" else NA)
  flag <- paste(na.omit(flag), collapse = ",")
  results[[a]] <- tibble(asset = a,
                         tc    = as.Date(date_decimal(ft$tc)),
                         m     = round(ft$m, 3), omega = round(ft$w, 2),
                         B     = round(ft$B, 3), C = round(ft$C, 3),
                         R2    = round(ft$R2, 3), D = round(ft$D, 2),
                         flag  = if_else(flag == "", "clean", flag))
  curves[[a]]  <- tibble(asset = a, date = d$date, logprice = yy, fit = ft$yhat)
}
results <- bind_rows(results)
curves  <- bind_rows(curves) %>% mutate(asset = factor(asset, levels = labels))
print(results, n = Inf, width = Inf)
```

```
## # A tibble: 8 x 9
##   asset     tc             m omega      B     C    R2     D flag
##   <chr>     <date>     <dbl> <dbl>  <dbl> <dbl> <dbl> <dbl> <chr>
## 1 Nasdaq    2026-09-29 0.802  6.73 -0.288 0.031 0.963  1.11 clean
## 2 Alphabet  2026-12-17 0.1    6.06 -6.32  0.123 0.962  0.85 m@low,D<1
## 3 Amazon    2027-09-21 0.9    4.8  -0.324 0.056 0.937  1.08 m@high,omega@bound
## 4 Apple     2027-04-02 0.9   13    -0.221 0.03  0.87   0.5  m@high,omega@bound,D<1
## 5 Meta      2027-09-21 0.9    4.8  -0.383 0.107 0.935  0.67 m@high,omega@bound,D<1
## 6 Microsoft 2027-09-21 0.9   11.0  -0.137 0.045 0.77   0.25 m@high,D<1
## 7 Nvidia    2027-09-21 0.9    4.8  -0.798 0.126 0.96   1.19 m@high,omega@bound
## 8 Tesla     2027-04-17 0.9   13    -0.378 0.093 0.853  0.28 m@high,omega@bound,D<1
```

4. LPPLS confidence on the Nasdaq

DS LPPLS confidence on the Nasdaq: the end date is fixed at the last observation and the start date slides forward in steps of 14 days, so the same regime is refit over shrinking windows. Search bounds are deliberately wide and the filtering is left to the Demirer, Demos, Gupta and Sornette qualification conditions: negative B, m in [0.1, 0.9], omega in [4.8, 13], damping of at least 0.8 and at least 2.5 oscillations before tc. Confidence is the share of windows whose fit passes.

```
window_starts <- seq(as.Date("2022-11-01"), as.Date("2025-12-01"), by = "14 days")
min_obs       <- 100         # skip windows shorter than this (100)
```

```r
nasdaq <- tq_get("^IXIC", from = "2016-01-01", to = Sys.Date(), get = "stock.prices") %>%
                  select(date, close) %>%
                  drop_na(close) %>%
                  arrange(date)
t2_date <- max(nasdaq$date)
t2      <- decimal_date(t2_date)

# wide bounds, so the qualification filter is informative
lppl_sse <- function(par, tt, yy) {

  tc <- par[1]
  m <- par[2]
  w <- par[3]
  dt <- tc - tt
  if (any(dt <= 1e-6)) return(1e12)

  f <- dt^m
  ln <- log(dt)
  X <- cbind(1, f, f*cos(w*ln), f*sin(w*ln))
  fit <- tryCatch(.lm.fit(X, yy), error = function(e) NULL)
  if (is.null(fit)) return(1e12)
  sum(fit$residuals^2)

}

fit_lppl <- function(tt, yy) {

  tmax <- max(tt)
  grid <- expand.grid(tc = seq(tmax + 0.02, tmax + 0.8, length.out = 5),
                      w  = seq(3, 20, length.out = 5),
                      m  = c(0.3, 0.6))
  best <- list(value = Inf)

  for (i in seq_len(nrow(grid))) {
    r <- tryCatch(
      optim(c(grid$tc[i], grid$m[i], grid$w[i]), lppl_sse, tt = tt, yy = yy,
            method = "L-BFGS-B",
            lower = c(tmax + 1e-3, 0.01, 1),
            upper = c(tmax + 1.0,  1.2, 25)),
      error = function(e) NULL)
    if (!is.null(r) && r$value < best$value) best <- r
  }

  tc <- best$par[1]
  m <- best$par[2]
  w <- best$par[3]

  dt <- tc - tt; f <- dt^m; ln <- log(dt)
  co <- .lm.fit(cbind(1, f, f*cos(w*ln), f*sin(w*ln)), yy)$coefficients
  C  <- sqrt(co[3]^2 + co[4]^2)
  list(tc = tc, m = m, w = w, B = co[2], C = C)

}

res <- list()

for (i in seq_along(window_starts)) {
  t1d <- window_starts[i]
```

```r
  w   <- nasdaq %>% filter(date >= t1d, date <= t2_date)
  if (nrow(w) < min_obs) next
  tt <- decimal_date(w$date)
  yy <- log(w$close)
  t1 <- decimal_date(t1d)
  f  <- fit_lppl(tt, yy)
  D  <- (f$m * abs(f$B)) / (f$w * f$C)                                        # damping
  O  <- if (f$tc > t2) (f$w / (2*pi)) * log((f$tc - t1) / (f$tc - t2)) else 0  # oscillation
  qualified <- (f$B < 0) && (f$m >= 0.1 && f$m <= 0.9) &&
               (f$w >= 4.8 && f$w <= 13) && (D >= 0.8) && (O >= 2.5)
  res[[as.character(t1d)]] <- tibble(start = t1d, n = nrow(w),
                                     tc = as.Date(date_decimal(f$tc)),
                                     m = round(f$m, 2), omega = round(f$w, 2),
                                     D = round(D, 2), O = round(O, 1),
                                     qualified = qualified)
}

res <- bind_rows(res)

# confidence and tc stability
confidence <- mean(res$qualified)
qual       <- res %>% filter(qualified)
cat(sprintf("windows tested: %d\n", nrow(res)))
```

```
## windows tested: 81
```

```r
cat(sprintf("LPPLS confidence (valid fits): %.2f  (%d/%d)\n",
            confidence, sum(res$qualified), nrow(res)))
```

```
## LPPLS confidence (valid fits): 0.15  (12/81)
```

```r
if (nrow(qual) > 0) {
  cat(sprintf("valid tc: median %s, range %s to %s, sd %.1f months\n",
              median(qual$tc), min(qual$tc), max(qual$tc),
              sd(decimal_date(qual$tc)) * 12))
}
```

```
## valid tc: median 2026-10-15, range 2026-09-07 to 2026-12-24, sd 1.2 months
```

```r
print(res, n = Inf)
```

```
## # A tibble: 81 x 8
##    start          n tc             m omega     D     O qualified
##    <date>     <int> <date>     <dbl> <dbl> <dbl> <dbl> <lgl>
##  1 2022-11-01   919 2026-09-07  0.85  6.2   1.24   3   TRUE
##  2 2022-11-15   909 2026-09-12  0.84  6.31  1.21   3   TRUE
##  3 2022-11-29   900 2026-09-12  0.84  6.31  1.21   3   TRUE
##  4 2022-12-13   890 2026-09-13  0.84  6.33  1.21   3   TRUE
##  5 2022-12-27   881 2026-09-22  0.82  6.55  1.15   2.9 TRUE
##  6 2023-01-10   872 2026-10-09  0.78  6.96  1.06   2.9 TRUE
##  7 2023-01-24   863 2026-10-22  0.76  7.24  1      2.9 TRUE
##  8 2023-02-07   853 2026-10-28  0.75  7.38  0.98   2.9 TRUE
##  9 2023-02-21   844 2026-11-01  0.74  7.46  0.96   2.9 TRUE
## 10 2023-03-07   834 2026-11-15  0.72  7.78  0.91   2.8 TRUE
## 11 2023-03-21   824 2026-12-06  0.69  8.23  0.86   2.8 TRUE
## 12 2023-04-04   814 2026-12-24  0.66  8.6   0.82   2.8 TRUE
## 13 2023-04-18   805 2027-01-10  0.64  8.96  0.78   2.8 FALSE
## 14 2023-05-02   795 2027-02-04  0.61  9.48  0.74   2.8 FALSE
## 15 2023-05-16   785 2027-03-06  0.58 10.1   0.7    2.8 FALSE
## 16 2023-05-30   776 2027-03-29  0.55 10.5   0.68   2.7 FALSE
## 17 2023-06-13   766 2027-04-11  0.54 10.8   0.66   2.7 FALSE
```

```
## 18 2023-06-27   757 2027-04-14  0.53 10.8   0.66   2.7 FALSE
## 19 2023-07-11   748 2027-04-14  0.53 10.8   0.66   2.7 FALSE
## 20 2023-07-25   738 2027-03-30  0.55 10.5   0.67   2.7 FALSE
## 21 2023-08-08   728 2027-03-09  0.59 10.1   0.69   2.7 FALSE
## 22 2023-08-22   718 2027-03-03  0.6  10     0.69   2.6 FALSE
## 23 2023-09-05   709 2027-02-18  0.62  9.74  0.7    2.6 FALSE
## 24 2023-09-19   699 2027-01-31  0.66  9.37  0.72   2.6 FALSE
## 25 2023-10-03   689 2027-02-10  0.64  9.57  0.71   2.6 FALSE
## 26 2023-10-17   679 2027-02-24  0.6   9.86  0.7    2.6 FALSE
## 27 2023-10-31   669 2027-05-16  0.44 11.4   0.67   2.5 FALSE
## 28 2023-11-14   659 2027-07-02  0.31 12.5   0.64   2.6 FALSE
## 29 2023-11-28   650 2027-07-02  0.27 12.6   0.63   2.6 FALSE
## 30 2023-12-12   640 2027-07-02  0.23 12.7   0.62   2.6 FALSE
## 31 2023-12-26   631 2027-07-02  0.22 12.8   0.62   2.6 FALSE
## 32 2024-01-09   622 2027-07-02  0.2  12.8   0.62   2.5 FALSE
## 33 2024-01-23   613 2027-07-02  0.18 12.9   0.62   2.5 FALSE
## 34 2024-02-06   603 2027-07-02  0.2  12.8   0.62   2.5 FALSE
## 35 2024-02-20   594 2027-07-02  0.25 12.8   0.62   2.5 FALSE
## 36 2024-03-05   584 2027-07-02  0.32 12.7   0.61   2.4 FALSE
## 37 2024-03-19   574 2027-07-02  0.39 12.7   0.6    2.4 FALSE
## 38 2024-04-02   565 2027-07-02  0.48 12.7   0.59   2.4 FALSE
## 39 2024-04-16   555 2027-07-02  0.55 12.7   0.58   2.4 FALSE
## 40 2024-04-30   545 2027-07-02  0.54 12.7   0.59   2.3 FALSE
## 41 2024-05-14   535 2027-07-02  0.55 12.7   0.59   2.3 FALSE
## 42 2024-05-28   526 2027-07-02  0.6  12.8   0.58   2.3 FALSE
## 43 2024-06-11   516 2027-07-02  0.66 12.9   0.57   2.3 FALSE
## 44 2024-06-25   507 2027-07-02  0.79 13.1   0.55   2.3 FALSE
## 45 2024-07-09   498 2027-07-02  0.97 13.3   0.54   2.3 FALSE
## 46 2024-07-23   488 2027-07-02  1.2  13.7   0.51   2.3 FALSE
## 47 2024-08-06   478 2027-07-02  1.2  13.8   0.5    2.3 FALSE
## 48 2024-08-20   468 2027-07-02  1.2  13.8   0.51   2.3 FALSE
## 49 2024-09-03   459 2027-07-02  1.2  13.9   0.5    2.3 FALSE
## 50 2024-09-17   449 2027-07-02  1.2  13.7   0.5    2.2 FALSE
## 51 2024-10-01   439 2027-07-02  1.2  13.6   0.5    2.2 FALSE
## 52 2024-10-15   429 2027-07-02  1.2  13.4   0.48   2.1 FALSE
## 53 2024-10-29   419 2027-07-02  1.13 13.2   0.45   2.1 FALSE
## 54 2024-11-12   409 2027-07-02  1.1  13.1   0.42   2   FALSE
## 55 2024-11-26   399 2027-07-02  1.13 13.1   0.4    2   FALSE
## 56 2024-12-10   390 2027-07-02  1.15 13.1   0.39   2   FALSE
## 57 2024-12-24   380 2027-07-02  1.14 13.1   0.39   1.9 FALSE
## 58 2025-01-07   372 2027-07-02  1.13 13.1   0.4    1.9 FALSE
## 59 2025-01-21   364 2027-07-02  1.12 13.1   0.4    1.9 FALSE
## 60 2025-02-04   354 2026-10-06  0.5   6.63  0.37   1.9 FALSE
## 61 2025-02-18   345 2026-07-31  0.39  4.64  0.4    2.1 FALSE
## 62 2025-03-04   335 2026-07-22  0.52  4.12  0.46   2.1 FALSE
## 63 2025-03-18   325 2026-07-18  0.63  3.81  0.52   2.1 FALSE
## 64 2025-04-01   315 2026-07-11  0.82  3.33  0.62   2.1 FALSE
## 65 2025-04-15   305 2026-07-15  0.7   3.61  0.56   2   FALSE
## 66 2025-04-29   296 2026-07-23  0.43  4.18  0.45   2   FALSE
## 67 2025-05-13   286 2026-07-28  0.23  4.56  0.4    2   FALSE
## 68 2025-05-27   277 2026-07-29  0.18  4.65  0.39   2   FALSE
## 69 2025-06-10   267 2026-07-31  0.09  4.81  0.38   2   FALSE
## 70 2025-06-24   258 2026-08-01  0.01  4.91  0.37   2   FALSE
## 71 2025-07-08   249 2026-08-01  0.01  4.9   0.37   2   FALSE
## 72 2025-07-22   239 2026-08-01  0.01  4.9   0.37   2   FALSE
## 73 2025-08-05   229 2026-08-01  0.01  4.91  0.37   1.9 FALSE
## 74 2025-08-19   219 2026-08-02  0.01  4.96  0.37   1.9 FALSE
## 75 2025-09-02   210 2026-08-02  0.01  4.94  0.37   1.9 FALSE
## 76 2025-09-16   200 2026-08-01  0.01  4.9   0.37   1.8 FALSE
```

```
## 77 2025-09-30   190 2026-08-03  0.01  5.03  0.36   1.8 FALSE
## 78 2025-10-14   180 2026-08-05  0.01  5.23  0.36   1.8 FALSE
## 79 2025-10-28   170 2026-08-08  0.01  5.43  0.36   1.8 FALSE
## 80 2025-11-11   160 2026-08-18  0.15  6.22  0.36   1.8 FALSE
## 81 2025-11-25   150 2026-08-20  0.26  6.41  0.36   1.7 FALSE
```

Crypto companies

Major coins from Yahoo, ticker suffix -USD, with Bitcoin as the reference series.

```
lookup <- tribble(~name,       ~ticker,
                  "Bitcoin",   "BTC-USD",
                  "Ethereum",  "ETH-USD",
                  "XRP",       "XRP-USD",
                  "Litecoin",  "LTC-USD",
                  "Dogecoin",  "DOGE-USD",
                  "Cardano",   "ADA-USD",
                  "Solana",    "SOL-USD")

from <- as.Date("2015-01-01")
to   <- Sys.Date()
raw  <- tq_get(lookup$ticker, from = from, to = to, get = "stock.prices")

prices <- raw %>% left_join(lookup, by = c("symbol" = "ticker")) %>%
                  select(name, date, adjusted) %>%
                  arrange(name, date)

prices %>% group_by(name) %>%
           summarise(obs = n(), from = min(date), to = max(date),
                     n_na = sum(is.na(adjusted)), .groups = "drop") %>%
           print(n = Inf)
```

```
## # A tibble: 7 x 5
##   name       obs from       to          n_na
##   <chr>    <int> <date>     <date>     <int>
## 1 Bitcoin   4204 2015-01-01 2026-07-05     0
## 2 Cardano   3161 2017-11-09 2026-07-05     0
## 3 Dogecoin  3161 2017-11-09 2026-07-05     0
## 4 Ethereum  3161 2017-11-09 2026-07-05     0
## 5 Litecoin  4204 2015-01-01 2026-07-05     0
## 6 Solana    2278 2020-04-10 2026-07-05     0
## 7 XRP       3161 2017-11-09 2026-07-05     0
```

Same construction with $h = 365$, since coins trade 7 days a week.

```
h <- 365L   #exchanged 7/7 days a week

dyn <- prices %>% filter(!is.na(adjusted)) %>%
                  group_by(name) %>%
                  arrange(date, .by_group = TRUE) %>%
                  mutate(
                    logp = log(adjusted),
                    velocity = logp - lag(logp, h),
                    acceleration = velocity - lag(velocity, h)) %>%
                  ungroup()
```

The cut is the Bitcoin peak of 10 November 2021. Solana lists in April 2020, so no acceleration exists before the cut, hence the NaN, and its pre velocity rests on 7 months of data only.

```
boom_date <- as.Date("2021-11-10")    # Bitcoin peak (~69k)

pre_post <- dyn %>% mutate(period = if_else(date < boom_date, "pre", "post")) %>%
                    group_by(name, period) %>%
                    summarise(vel_pct = (exp(mean(velocity, na.rm = TRUE)) - 1) * 100,
                              acc     =  mean(acceleration, na.rm = TRUE),
                              .groups = "drop") %>%
                    pivot_wider(names_from = period, values_from = c(vel_pct, acc)) %>%
                    mutate(vel_change = vel_pct_post - vel_pct_pre,
                           acc_change = acc_post - acc_pre) %>%
                    mutate(across(where(is.numeric), \(x) round(x, 2))) %>%
                    arrange(desc(vel_change))

print(pre_post, width = Inf)
```

```
## # A tibble: 7 x 7
##   name     vel_pct_post vel_pct_pre acc_post acc_pre vel_change acc_change
##   <chr>           <dbl>       <dbl>    <dbl>   <dbl>      <dbl>      <dbl>
## 1 XRP             23.9         6.52    -0.18    0.86       17.4      -1.04
## 2 Ethereum         8.23       59.6     -0.44    1.58      -51.3      -2.02
## 3 Cardano        -17          77.2     -0.67    1.93      -94.2      -2.6
## 4 Bitcoin         18.2       135.      -0.31    0.16     -117.       -0.46
## 5 Litecoin       -15.6       104.      -0.29    0.16     -119.       -0.45
## 6 Dogecoin        11.9       172.      -0.81    1.97     -160.       -2.78
## 7 Solana          45.2      4617.      -0.92  NaN       -4571.       NaN
```

Dot-com companies

Dotcom darlings plus the Nasdaq Composite as market, January 1995 to January 2003, all with Yahoo history reaching back to the 1990s.

```
lookup <- tribble(~name,       ~ticker,
                  "Cisco",     "CSCO",
                  "Intel",     "INTC",
                  "Oracle",    "ORCL",
                  "Microsoft", "MSFT",
                  "Amazon",    "AMZN",
                  "Qualcomm",  "QCOM",
                  "Nasdaq",    "^IXIC")  #market

from <- as.Date("1995-01-01")
to   <- as.Date("2003-01-01")
raw  <- tq_get(lookup$ticker, from = from, to = to, get = "stock.prices")

prices <- raw %>% left_join(lookup, by = c("symbol" = "ticker")) %>%
                  select(name, date, adjusted) %>%
                  arrange(name, date)

prices %>% group_by(name) %>%
           summarise(obs = n(), from = min(date), to = max(date),
                     n_na = sum(is.na(adjusted)), .groups = "drop") %>%
                     print(n = Inf)
```

```
## # A tibble: 7 x 5
##   name        obs from       to          n_na
##   <chr>     <int> <date>     <date>     <int>
## 1 Amazon     1416 1997-05-15 2002-12-31     0
```

```
## 2 Cisco      2015 1995-01-03 2002-12-31     0
## 3 Intel      2015 1995-01-03 2002-12-31     0
## 4 Microsoft  2015 1995-01-03 2002-12-31     0
## 5 Nasdaq     2015 1995-01-03 2002-12-31     0
## 6 Oracle     2015 1995-01-03 2002-12-31     0
## 7 Qualcomm   2015 1995-01-03 2002-12-31     0
```

Back to trading days, $h = 252$.

```
h <- 252L

dyn <- prices %>% filter(!is.na(adjusted)) %>%
                  group_by(name) %>%
                  arrange(date, .by_group = TRUE) %>%
                  mutate(
                    logp = log(adjusted),
                    velocity = logp - lag(logp, h),
                    acceleration = velocity - lag(velocity, h)) %>%
                  ungroup()
```

The cut is the Nasdaq peak of 10 March 2000.

```
boom_date <- as.Date("2000-03-10")   # Nasdaq peak

pre_post <- dyn %>% mutate(period = if_else(date < boom_date, "pre", "post")) %>%
                    group_by(name, period) %>%
                    summarise(vel_pct = (exp(mean(velocity, na.rm = TRUE)) - 1) * 100,
                              acc     =  mean(acceleration, na.rm = TRUE),
                              .groups = "drop") %>%
                    pivot_wider(names_from = period, values_from = c(vel_pct, acc)) %>%
                    mutate(vel_change = vel_pct_post - vel_pct_pre,
                           acc_change = acc_post - acc_pre) %>%
                    mutate(across(where(is.numeric), \(x) round(x, 2))) %>%
                    arrange(desc(vel_change))

print(pre_post, width = Inf)
```

```
## # A tibble: 7 x 7
##   name      vel_pct_post vel_pct_pre acc_post acc_pre vel_change acc_change
##   <chr>            <dbl>       <dbl>    <dbl>   <dbl>      <dbl>      <dbl>
## 1 Oracle            -4.3        43.0    -0.47    0.14      -47.3      -0.61
## 2 Nasdaq           -21.8        32.8    -0.27    0.07      -54.6      -0.34
## 3 Intel            -16.5        50.4    -0.26    0.03      -66.9      -0.3
## 4 Microsoft        -17.0        68.4    -0.23    0.04      -85.5      -0.27
## 5 Qualcomm         -13.1        79.7    -0.92    0.56      -92.8      -1.48
## 6 Cisco            -28.8        83.9    -0.41    0.03     -113.       -0.44
## 7 Amazon           -40.5       392.     -0.34   -0.98     -433.        0.65
```

Summary

```
## PUBLIC LEG - RECAP
## ------------------
## LPPL, Nasdaq  : clean interior fit, m=0.802, omega=6.73, R2=0.963, tc=2026-09-29
## LPPL, Mag-7   : no clean fit; 6 of 7 pin m at the upper bound (Alphabet at the
##                 lower); omega pinned for 5 of 7 (3 low at 4.8, 2 high at 13)
## DS confidence : 0.15 (12/81 windows); valid tc median 2026-10-15,
##                 range 2026-09-07 to 2026-12-24, sd 1.2 months
## Velocity, AI  : Nvidia +40.8 and Meta +35.6 points of annual velocity after the
##                 2022-11-30 cut; post accelerations at or above 0 for all names
```

```
## Controls      : crypto and dotcom velocities collapse after their peaks
##
## Bubble hypothesis: velocity accelerates after the ChatGPT cut and the Nasdaq
## carries a clean log-periodic signature with tc in autumn 2026; single names
## fit only at degenerate parameter values.
```

Bibliography

[1] Daron Acemoglu, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. The network origins of aggregate fluctuations. *Econometrica*, 80(5): 1977–2016, 2012. doi:10.3982/ECTA9623.

[2] Al Jazeera. Nvidia becomes first US company to reach $4 trillion market cap. Al Jazeera, July 2025. URL https://www.aljazeera.com/economy/2025/7/9/nvidia-becomes-first-us-company-to-reach-4-trillion-market-cap.

[3] Robert L. Axtell. Zipf distribution of U.S. firm sizes. *Science*, 293(5536):1818–1820, 2001. doi:10.1126/science.1062081.

[4] Bank for International Settlements. Crypto shocks and retail losses. BIS Bulletin 69, Bank for International Settlements, 2022.

[5] Rerotlhe B. Basele, Peter C. B. Phillips, and Shuping Shi. Speculative bubbles in the recent AI boom: Nasdaq and the Magnificent Seven. *Journal of Time Series Analysis*, 46(5):814–828, 2025. doi:10.1111/jtsa.12835.

[6] Bloomberg News. Can OpenAI and Anthropic IPOs live up to expectations? Bloomberg, 2026. URL https://www.bloomberg.com/news/articles/2026-06-01/can-openai-and-anthropic-ipos-live-up-to-expectations. 1 June 2026.

[7] Markus K. Brunnermeier and Stefan Nagel. Hedge funds and the technology bubble. *The Journal of Finance*, 59(5):2013–2040, 2004. doi:10.1111/j.1540-6261.2004.00690.x.

[8] Oliver D. Bunn and Robert J. Shiller. Changing times, changing values: A historical analysis of sectors within the US stock market 1872–2013. Cowles Foundation Discussion Paper 1950, Cowles Foundation for Research in Economics, Yale University, 2014. Also NBER Working Paper No. 20370.

[9] Zheng Cao, Xingran Shao, Yuheng Yan, and Hélyette Geman. Identifying and quantifying financial bubbles with the hyped log-periodic power law model. *arXiv preprint arXiv:2510.10878*, 2025. doi:10.48550/arXiv.2510.10878.

[10] Vasco M. Carvalho, Makoto Nirei, Yukiko U. Saito, and Alireza Tahbaz-Salehi. Supply chain disruptions: Evidence from the great east japan earthquake. *The Quarterly Journal of Economics*, 136(2):1255–1321, 2021. doi:10.1093/qje/qjaa044.

[11] Eng-Tuck Cheah and John Fry. Speculative bubbles in Bitcoin markets? an empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, 130:32–36, 2015. doi:10.1016/j.econlet.2015.02.029.

[12] Adrian Cheung, Eduardo Roca, and Jen-Je Su. Crypto-currency bubbles: An application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox Bitcoin prices. *Applied Economics*, 47(23):2348–2358, 2015. doi:10.1080/00036846.2015.1005827.

[13] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009. doi:10.1137/070710111.

[14] CNBC. Bitcoin rises to record above $69,000, then quickly tumbles. CNBC, March 2024. URL https://www.cnbc.com/2024/03/05/bitcoin-all-time-high.html.

[15] CNBC. Nvidia sheds almost $600 billion in market cap, biggest drop ever. CNBC, January 2025. URL https://www.cnbc.com/2025/01/27/nvidia-sheds-almost-600-billion-in-market-cap-biggest-drop-ever.html.

[16] CNBC. Anthropic confidentially files IPO prospectus with SEC. CNBC, 2026. URL https://www.cnbc.com/2026/06/01/anthropic-ipo-s1-prospectus.html. 1 June 2026.

[17] CNBC. Anthropic set to hit $10.9 billion in revenue in Q2, source says. CNBC, 2026. URL https://www.cnbc.com/2026/05/20/anthropic-revenue-explosive-growth-ipo-profitable-quarter.html. 20 May 2026.

[18] CNBC. OpenAI closes funding round at an $852 billion valuation. CNBC, 2026. URL https://www.cnbc.com/2026/03/31/openai-funding-round-ipo.html. 31 March 2026.

[19] Crunchbase. Six charts that show the big AI funding trends of 2025. Crunchbase News, 2025. URL https://news.crunchbase.com/ai/big-funding-trends-charts-eoy-2025/.

[20] Crunchbase. Q1 2026 shatters venture funding records as AI boom pushes startup investment to $300b. Crunchbase News, 2026. URL https://news.crunchbase.com/venture/record-breaking-funding-ai-global-q1-2026/.

[21] Alex de Vries. The growing energy footprint of artificial intelligence. *Joule*, 7 (10):2191–2194, 2023. doi:10.1016/j.joule.2023.09.004.

[22] David A. Dickey and Wayne A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427–431, 1979. doi:10.1080/01621459.1979.10482531.

[23] Bradley Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979. doi:10.1214/aos/1176344552.

[24] Fredrik A. Enoksen, Ch. J. Landsnes, K. Lučivjanská, and Peter Molnár. Understanding risk of bubbles in cryptocurrencies. *Journal of Economic Behavior & Organization*, 176:129–144, 2020. doi:10.1016/j.jebo.2020.05.005.

[25] Vladimir Filimonov and Didier Sornette. A stable and robust calibration scheme of the log-periodic power law model. *Physica A: Statistical Mechanics and its Applications*, 392(17):3698–3707, 2013. doi:10.1016/j.physa.2013.04.012.

[26] Fortune. OpenAI's financials have leaked, showing $21 billion in losses against $13 billion in revenue. Fortune, 2026. URL https://fortune.com/2026/06/16/openai-financials-leaked-losses-revenue-profit/. 16 June 2026.

[27] Xavier Gabaix. Power laws in economics and finance. *Annual Review of Economics*, 1:255–294, 2009. doi:10.1146/annurev.economics.050708.142940.

[28] Corrado Gini. *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche*. Tipografia di Paolo Cuppini, Bologna, 1912.

[29] Paul Gompers and Josh Lerner. Money chasing deals? the impact of fund inflows on private equity valuations. *Journal of Financial Economics*, 55(2): 281–325, 2000. doi:10.1016/S0304-405X(99)00052-5.

[30] Will Gornall and Ilya A. Strebulaev. Squaring venture capital valuations with reality. *Journal of Financial Economics*, 135(1):120–143, 2020. doi:10.1016/j.jfineco.2018.04.015.

[31] Sandy Edward Green Jr. A rhetorical theory of diffusion. *Academy of Management Review*, 29(4):653–669, 2004. doi:10.5465/amr.2004.14497653.

[32] Orris C. Herfindahl. *Concentration in the Steel Industry*. PhD thesis, Columbia University, 1950.

[33] Albert O. Hirschman. *National Power and the Structure of Foreign Trade*. University of California Press, Berkeley, 1945.

[34] Anders Johansen, Olivier Ledoit, and Didier Sornette. Crashes as critical points. *International Journal of Theoretical and Applied Finance*, 3(2):219–255, 2000. doi:10.1142/S0219024900000115.

[35] Charles P. Kindleberger and Robert Z. Aliber. *Manias, Panics, and Crashes: A History of Financial Crises*. Palgrave Macmillan, Basingstoke, 7th edition, 2015.

[36] Mark Kritzman and David Turkington. The fallacy of concentration. *Financial Analysts Journal*, 82(2):8–18, 2026. doi:10.1080/0015198X.2026.2640988.

[37] Ganghyeok Lee, Minhyuk Jeong, Taeyoung Park, and Kwangwon Ahn. More than ex-post fitting: log-periodic power law and its AI-based classification. *Humanities and Social Sciences Communications*, 12(1):1664, 2025. doi:10.1057/s41599-025-05920-7.

[38] Rosario N. Mantegna and H. Eugene Stanley. *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, Cambridge, 2000.

[39] Eric Masanet, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan Koomey. Recalibrating global data center energy-use estimates. *Science*, 367(6481):984–986, 2020. doi:10.1126/science.aba3758.

[40] Chris Miller. *Chip War: The Fight for the World's Most Critical Technology*. Scribner, New York, 2022.

[41] multpl. Shiller PE ratio for the S&P 500. multpl.com, 2026. URL https://www.multpl.com/shiller-pe.

[42] Nasdaq Private Market. Data and intelligence: NPM price and Tape D methodology. https://www.nasdaqprivatemarket.com/data-intelligence/, 2026.

[43] NBC News. Nvidia becomes first company worth $4 trillion: what to know. NBC News, July 2025. URL https://www.nbcnews.com/business/business-news/nvidia-becomes-first-company-worth-4-trillion-what-to-know-rcna217721.

[44] M. E. J. Newman. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, 46(5):323–351, 2005. doi:10.1080/00107510500052444.

[45] Eli Ofek and Matthew Richardson. DotCom mania: The rise and fall of internet stock prices. *The Journal of Finance*, 58(3):1113–1137, 2003. doi:10.1111/1540-6261.00560.

[46] Ľuboš Pástor and Pietro Veronesi. Technological revolutions and stock prices. *American Economic Review*, 99(4):1451–1483, 2009. doi:10.1257/aer.99.4.1451.

[47] Peter C. B. Phillips, Yangru Wu, and Jun Yu. Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review*, 52(1):201–226, 2011. doi:10.1111/j.1468-2354.2010.00625.x.

[48] Peter C. B. Phillips, Shuping Shi, and Jun Yu. Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review*, 56(4):1043–1078, 2015. doi:10.1111/iere.12132.

[49] Jaime Sevilla, Lennart Heim, Anson Ho, Tamay Besiroglu, Marius Hobbhahn, and Pablo Villalobos. Compute trends across three eras of machine learning. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2022. doi:10.1109/IJCNN55064.2022.9891914.

[50] Robert J. Shiller. *Irrational Exuberance*. Princeton University Press, Princeton, NJ, 2000.

[51] Siblis Research. NASDAQ 100 index P/E, EPS & earnings growth. Siblis Research, 2026. URL https://siblisresearch.com/data/nasdaq-100-pe-ratio/.

[52] Didier Sornette. *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton University Press, Princeton, NJ, 2003.

[53] Didier Sornette, Guilherme Demos, and Qun Zhang. Real-time prediction and postmortem analysis of the Shanghai 2015 stock market bubble and crash. *Journal of Investment Strategies*, 4(4):77–95, 2015. doi:10.21314/JOIS.2015.063.

[54] Stanford HAI. The 2026 AI index report. Technical report, Stanford Institute for Human-Centered Artificial Intelligence, Stanford University, 2026. URL https://hai.stanford.edu/ai-index/2026-ai-index-report.

[55] The Information. OpenAI projections imply losses tripling to $14 billion in 2026. The Information, 2026. URL https://www.theinformation.com/articles/openai-projections-imply-losses-tripling-to-14-billion-in-2026.

[56] Quang H. Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2):307–333, 1989. doi:10.2307/1912557.

[57] Spencer Wheatley, Didier Sornette, Tobias Huber, Max Reppen, and Robert N. Gantner. Are Bitcoin bubbles predictable? combining a generalized Metcalfe's

law and the log-periodic power law singularity model. *Royal Society Open Science*, 6(6):180538, 2019. doi:10.1098/rsos.180538.