V1.0 · SEPTEMBER 2026 · PAPER 11

The survivors wrote the manual

Startup advice is assembled by reading the companies that came back. Against 1,404 companies worth a billion dollars and 410 that went to zero, here is what that method is actually worth, and what to do instead.

Nikolaos Broikos · broikos.gr September 2026

1,404 winners · 410 failures · 44 years of audited precedent

63.5%

OF STARTUPS THAT WENT TO ZERO FIT THE PRODUCT CATEGORIES SAID TO PREDICT SUCCESS

10.5

PERCENTAGE POINTS: ALL THE ROOM THE IDEA HAS LEFT TO BE TRUE IN

26.1%

OF BILLION-DOLLAR COMPANIES NO SUCH CATEGORY CAN DESCRIBE AT ALL

45

MEAN AGE OF THE FOUNDERS OF THE FASTEST-GROWING ONE IN A THOUSAND VENTURES

There is an idea underneath almost all practical startup advice: that you can tell in advance which companies will succeed by looking at what kind of product they are. It arrives as a list of categories. Marketplaces. Things that let ordinary people earn money. Things that make an expensive service cheap.

The list is always assembled the same way, by reading the companies that reached a billion dollars and writing down what recurs. This paper asks whether that works, using the half of the evidence the method never looks at.

FINDING 01

A quarter of the answer is out of reach

366 of 1,404 billion-dollar companies build drugs, rockets, batteries and semiconductors. No product-shape category describes them. Whatever share of winners the categories explain, it cannot exceed 73.9 per cent.

FINDING 02

The same categories describe most of the dead

Of 52 failed startups classified from their own descriptions, 33 sit inside at least one category. That is 63.5 per cent, against a ceiling of 73.9. The whole idea has about ten points of room to be true in.

FINDING 03

What wins does not hold still

Consumer companies fell from 41.5 per cent of new billion-dollar entrants to 8.7. Enterprise rose from 14.9 to 46.6. China fell from 46.8 to 5.7. A rule fitted to one cohort describes a population that has been replaced.

PRECEDENT

This was run twice before, at scale

Two of the best-selling business books ever written used this exact method, in 1982 and 2001. One company appears in both lists as an exemplar of durability. It went bankrupt.

> **Who this is for.** Founders choosing what to build, and investors deciding what to believe about a category. It ends in seven things to do differently, not in a verdict.

01 The armour went on the wrong part of the plane

In 1943 the American military had what it took to be an engineering problem. Bombers were returning from Europe full of holes, and the holes were not evenly spread. They clustered on the wings, the fuselage, the tail gunner's position. The obvious move was to armour where the bullets were going.

Abraham Wald, a statistician working out of Columbia, told them to armour everywhere else.[1]

The bullet holes were a map of where a plane could be hit and still fly home. The clean areas on the returning aircraft were not lucky. They were the places where a hit meant the plane did not come back to be measured. Every observation in the dataset had been filtered by the outcome the study was trying to explain, and the pattern in the data was the exact inverse of the truth.

This is the most quoted story in the literature on selection bias, and everyone in business has heard it. It is quoted so often that it has become decorative. And then people go straight back to studying the planes that came home.

> **The problem is not the content of the list. The problem is that the list is a map of the holes in the returning planes.**

The categories are not stupid, and that is what makes this worth writing about. They are drawn from real companies, by people who have seen a great many of them, and every item describes something that genuinely did work at least once.

02 The same method, run twice before, and audited by time

Before asking whether the startup version works, it is worth knowing that this exact method has been run at the level of large corporations, twice, by serious people with far more resources than any startup writer, and that we now have decades of outcome data on both attempts.

1982: forty-three excellent companies

In Search of Excellence created the modern business book. Tom Peters and Robert Waterman, both at McKinsey, selected forty-three companies that were performing superbly, studied what they had in common, and extracted eight attributes of excellence.[2] It sold in the millions and set the template every book of this kind has followed since.

Then time passed, which is the one experimental condition this method can never control for.

Of the thirty-five that were publicly traded, **twenty went on to underperform the market average.** Over the following two decades roughly one in seven of the portfolio ceased to exist independently. Amdahl, Data General, Digital Equipment and Raychem were absorbed. Kmart, Wang and Circuit City went bankrupt.

The eight attributes did not stop being true of those companies. They stopped mattering, or they had never been what was doing the work.

2001: eleven good-to-great companies

Nineteen years later Jim Collins ran a more rigorous version of the same procedure. *Good to Great* screened a large universe of public companies for a specific quantitative pattern, a period of ordinary performance followed by fifteen years of exceptional returns, and found eleven that qualified.[3]

THE ELEVEN, 2001	WHAT HAPPENED
Circuit City	The book's best performer at 18.5 times the market. Filed for bankruptcy in 2008
Fannie Mae	Federal bailout in the same crisis, having lost the overwhelming majority of its value
Wells Fargo	A decade of conduct scandals
Abbott, Gillette, Kimberly-Clark, Kroger, Nucor, Philip Morris, Pitney Bowes, Walgreens	An investor buying all eleven on publication day would have underperformed the S&P 500

Circuit City is the detail worth sitting with. **It appears in both books.** It was selected independently, by two different teams, using two different methodologies, nineteen years apart, as an exemplar of what makes a company durable. It then went bankrupt.

What Rosenzweig named

The sharpest methodological criticism came from Phil Rosenzweig in *The Halo Effect*, and it goes further than survivorship.[4] When a company is performing well, everything about it is described in flattering terms: its culture is focused, its chief executive decisive, its strategy bold. When the same company falters, the same culture is insular, the same executive autocratic, the same strategy reckless. Nothing changed except the results.

So a study that gathers evidence from press coverage and retrospective interviews at successful companies is not measuring the causes of success. It is measuring the vocabulary that success generates. The independent variable has been contaminated by the dependent one.

Every startup category list is built from exactly this material: what successful founders said afterwards, and what journalists wrote while they were winning.

The corporate version of the genre at least gets audited, because public companies have share prices and somebody eventually checks. The startup version does not. The companies are private, the lists live on the internet where they can be quietly edited, and the audience turns over every few years.

03 Nobody involved knows the denominator

Before you can evaluate advice about reaching a billion dollar outcome, you need to know how often that outcome happens. It is unknown, and the way in which it is unknown is instructive.

THE FIGURE	AS PUBLISHED	THE ODDS IT IMPLIES
0.07%	of venture-backed software and internet companies funded in the US since 2003	1 in 1,538
about 1%	of startups	1 in 100
0.00006%	of startups	1 in 1,700,000

They differ by a factor of roughly seventeen thousand.

The oldest of the three is by far the most honest, and it is worth reading at source rather than in citation. The 0.07 per cent figure comes from the 2013 analysis that coined the word unicorn.[5] Its numerator is thirty-nine companies. Its denominator is around sixty thousand, and the article says openly how that was reached: three available estimates of how many companies venture capital had funded, one of sixteen thousand, one of twelve thousand two hundred and ninety-one, and one of ten to fifteen thousand a year, aggregated to sixty thousand as a reasonable figure for the decade.

> The canonical base rate of this field is thirty-nine divided by an educated guess, and its author said so. Every citation since has kept the number and dropped the caveat.

The same thing happened to the most-used validation test

The standard instrument for deciding whether you have product-market fit is the forty per cent test: survey your users, ask how they would feel if they could no longer use the product, and if forty per cent or more say very disappointed, you have it.

Its origin is Sean Ellis, in 2009, from observation across roughly a hundred startups he had worked with.[6] Ellis had unusually good access, having been the first marketer at Dropbox, LogMeIn, Eventbrite and Lookout. It is a genuine observation from a real practitioner.

It has also never been independently validated. Search for the evidence behind the forty per cent threshold and what comes back is dozens of pages explaining how to run the test and none reporting an attempt to check it. The number has been in continuous use for seventeen years on the strength of one person's private sample.

And there is a subtler problem, which is why the test belongs in this paper rather than in a footnote. **The survey goes to current users.** Everyone who tried the product and left is definitionally excluded from the sample.

It is Wald's bombers again, one level down. You are measuring the enthusiasm of the people who did not churn, and using it to predict whether people will churn.

04 The idea, stated as generously as it deserves

Here is the thesis in its strongest form, because a paper that argues with a weak version of an idea has not argued with anything.

> Demand is not evenly distributed across product types. Some products are obviously valuable the moment you hear about them, and some require you to convince the buyer that a problem exists. The first kind spreads on its own and the second has to be pushed uphill forever. Look at the companies that got enormous and you find the first kind vastly over-represented. **Therefore, when choosing what to build, choose a shape that has historically pulled.**

Almost every clause of this is defensible. Demand really is unevenly distributed. Products really do differ in how much explaining they need, and the difference compounds through everything from marketing cost to sales cycle to churn.

The failure is entirely in the last sentence, and specifically in the word *therefore*.

The observation is about **which shapes appear among winners**. The conclusion is about **which shapes are more likely to win**. Those are different statements, and getting from the first to the second requires one further piece of information that no version of this advice has ever supplied: **how common each shape is among everything that was ever attempted.**

If two thirds of everything anyone ever founded is a marketplace, a cheaper substitute or an earning platform, then finding those shapes among the winners tells you precisely nothing. It is fully explained by the shapes being common. And no amount of additional winners can separate the two explanations, because each new winner adds to both sides of the ratio at once.

A list drawn from winners cannot be tested on winners. Not because the evidence is thin, but because it is evidence about a different question. The only way to settle it is to go and look at the companies that did not come back.

05 What the record shows

Two public bodies of evidence, put side by side. The first is the complete list of private companies currently valued at a billion dollars or more, 1,404 of them, with valuation, country, industry and the year each crossed the line. The second is a compilation of startup failure post-mortems, 410 companies, of which 266 say clearly enough what the product did to be classified.[10]

Then seven product shapes, deliberately mine rather than borrowed from any published list, arrived at by reading the failure descriptions until the recurring shapes stopped changing: **two-sided marketplace, earning platform, labour substitution, cheaper or easier substitute, localised replication, invitation software, and access to something previously gated.** Anyone who has read this literature will recognise all seven, which is the point. If the argument here is right, the result should not depend on whose seven you use.

Sixty failures were drawn at random with a fixed seed, so the sample could not be adjusted after I saw what it said, and each was classified from its own description alone, with no reference to the company name and no looking up what became of it.

Finding one: a quarter of the answer is out of reach before you start

All seven shapes describe consumer or business software. So does every published version of this list. None has a category for a rocket, a battery, a semiconductor or a monoclonal antibody, because those companies do not win by being immediately appealing to a consumer.

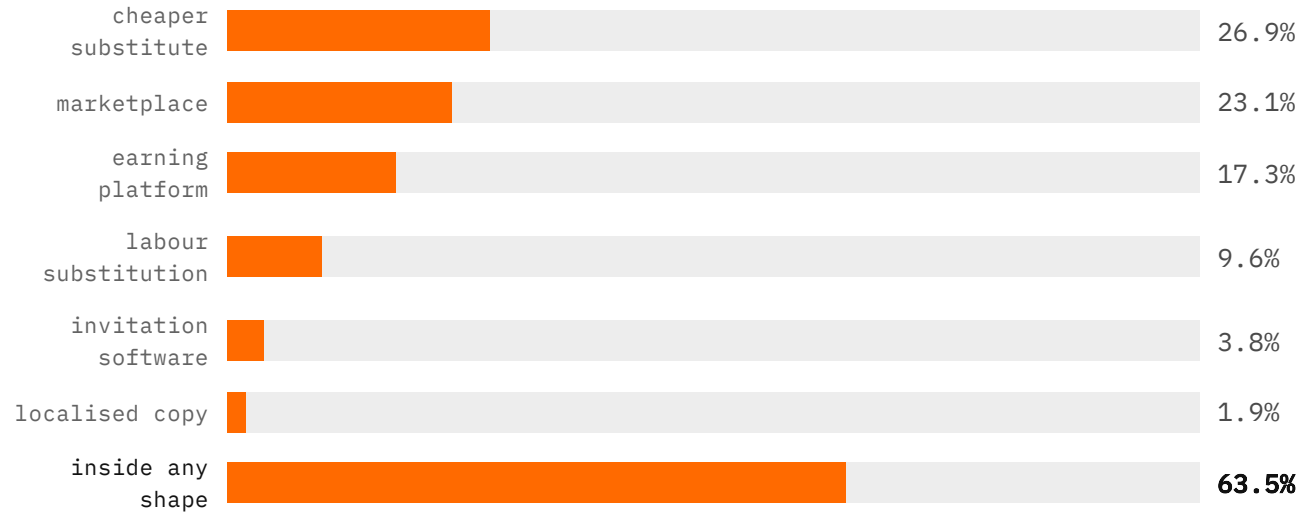
INDUSTRY	SHARE OF THE BILLION-DOLLAR LIST	CAN A SOFTWARE SHAPE DESCRIBE IT?
Enterprise Tech	37.0%	yes
Financial Services	16.4%	yes
Industrials	**15.2%**	no
Consumer & Retail	14.9%	yes
Healthcare & Life Sciences	**9.1%**	no
Media & Entertainment	5.6%	yes
Insurance	**1.6%**	no

366 of 1,404 companies, 26.1 per cent, sit in industries no product-shape scheme can reach. Whatever share of winners the shapes describe, it cannot be higher than 73.9 per cent. That is a ceiling, and it matters for what comes next.

Finding two: the same shapes describe most of the dead

Of the sixty failures sampled, eight turned out to quote the founder's farewell letter without ever saying what the product was, so they could not be classified either way and were set aside rather than guessed at.

Of the fifty-two that could be classified, thirty-three fall inside at least one of the seven shapes. That is 63.5 per cent.



Now put that against the ceiling from finding one.

At most 73.9 per cent of the winners can be inside the scheme. 63.5 per cent of the dead already are. The entire space in which this idea can be true is about ten percentage points wide.

And ten points is the generous reading, because 73.9 is a ceiling rather than a measurement. The true figure for winners sits somewhere below it, and every point it falls comes straight out of that ten. Throw out a third of my classifications as too generous and the picture barely moves: 42.3 per cent of the dead are still inside.

The scheme is not identifying companies that will succeed. It is identifying software companies. Most software companies die.

Finding three: what wins does not hold still

Every such list is fitted to the winners of a particular period. So does the population hold still long enough for a fitted rule to survive?

ENTERED THE LIST	HOW MANY	ENTERPRISE	CONSUMER	MEDIA	US	CHINA
2011 to 2017	**94**	**14.9%**	**41.5%**	**16.0%**	**29.8%**	**46.8%**
2018 to 2020	256	28.9%	21.9%	5.5%	52.0%	21.1%
2021	443	33.2%	13.5%	5.6%	59.1%	7.4%
2022 to 2026	**610**	**46.6%**	**8.7%**	**4.1%**	**62.8%**	**5.7%**

The winners of 2011 to 2017 were 41.5 per cent consumer companies and 46.8 per cent Chinese. That is the cohort every familiar example is drawn from. A taxonomy read off it is a taxonomy of consumer marketplaces, built during a period of historically cheap capital and a liberalising Chinese internet.

The winners since 2022 are 46.6 per cent enterprise and 5.7 per cent Chinese. Consumer fell by a factor of nearly five. China fell by a factor of eight.

A rule fitted to the first cohort and applied to the second is not slightly out of date. It is describing a population that has been replaced.

06 The graveyard is full of category members

The numbers above are abstract. The record is not, and it is more persuasive.

The on-demand wave

Around 2014 the single most confidently recommended shape in the world was the one later summarised as Uber for X. It combined three of the seven shapes at once: a marketplace, an earning platform for the supply side, and a cheaper and more convenient substitute for the buyer. Score that era's winners and this shape tops the list.

What happened to the cohort that built it:[8]

COMPANY	WHAT IT WAS	OUTCOME
Homejoy	on-demand home cleaning	shut down July 2015 after $39.7M raised, felled by worker classification suits
Shyp	on-demand shipping	reached a $250M valuation in 2015, never scaled past San Francisco, shut down
Sprig	on-demand prepared meals	shut down
Munchery	on-demand meals, better funded	bankrupt, March 2019
Zirtual	on-demand assistants	suspended abruptly, staff learned by email
Good Eggs	on-demand groceries	survived, after collapsing to a fraction of its footprint

Homejoy is in my failure sample, classified as an earning platform and a marketplace before I knew or checked what it was. It fits the category beautifully. It is dead. The category was not wrong about these companies. It described them accurately. It simply carried no information about whether they would live.

The same idea, opposite outcomes, ten years apart

The cleanest disproof of shape as destiny is the pair of cases where the shape was held constant and the outcome flipped.[9]

SAME CATEGORY	EARLIER ATTEMPT	LATER ATTEMPT
Groceries ordered online	**Webvan**: spent roughly $1.2B, bankrupt 2001	**Instacart**: one of the most valuable private companies in the world
Pet supplies online	**Pets.com**: raised about $300M, liquidated 2000	**Chewy**: acquired for $3.35B in 2017

If the shape of the product decided the outcome, these pairs are impossible. What differed is everything the category does not record: when it happened, what infrastructure already existed, whether the company had to build the pipeline or could rent it, and what a customer had already been trained to do by the time they arrived.

Webvan built its own warehouses and delivery fleet before knowing whether anyone wanted the service at that scale. Instacart rented the shelves of shops that already existed and the cars of drivers who already had them. That is not a difference of category. It is a difference of capital structure and timing, and no list of product shapes has a column for either.

And the counter-examples run the other way too

The advice also says, in most versions, that you want to arrive when the pull first appears, because whoever is first when conditions turn is the one who wins.

Google was not an early search engine. It launched in 1998 into a thoroughly crowded field, behind Yahoo, AltaVista, Excite, Lycos, Infoseek and Netscape, all of which had defined the category and several of which were considered to have won it. Facebook was not an early social network either. It arrived after Friendster and while MySpace was dominant. Both entered late, into categories held to be finished, and won on execution inside a shape that had already been proven to guarantee nothing.

The category told you nothing in either direction. It did not save Webvan and it did not stop Google.

07 What survives contact with the evidence

A paper that only demolishes wastes the reader's time. Several things in this literature hold up, and they share one property: **each names a mechanism, and each can be checked against something other than a list of winners.**

Switching costs are the best-supported idea in the field

The argument that accumulated reputation, an accumulated audience, locked-in data and a consolidated social graph keep customers in place is not pattern matching. It names a specific cost the customer would have to pay to leave, it explains why mediocre products survive for decades, and it makes a prediction you can test on your own company this afternoon: a product with none of these will lose users to a slightly better competitor.

The venture arithmetic is correct, and explains most of the rest

The standard explanation of why investors need billion-dollar outcomes checks out when you run it. A $1M cheque at a $15M post-money valuation buys 6.67 per cent of a company. Sixty per cent dilution across later rounds leaves 2.67 per cent. On a $1B exit that returns $26.7M, which against a $20M fund is 1.33 times the fund.

One billion-dollar outcome barely returns the fund. That single fact explains the entire incentive structure the advice sits inside: why investors push growth over profitability, why they prefer a small chance of an enormous outcome to a large chance of a good one, and why advice optimised for their portfolio is not automatically optimised for you.

> **It also implies something the advice rarely says out loud.** If your realistic best case is a profitable company worth twenty million dollars, that is an excellent outcome for you and a rounding error for a fund. Taking their money converts your good outcome into their failure.

What a properly done study looks like

The contrast is worth seeing, because the difference is not intelligence or effort. It is access to the denominator.

In 2020 four economists published an analysis of founder age using US Census administrative records.[7] Because they used government data covering essentially every new firm, they had the entire population, not the winners. They could therefore compute what nobody in the advice literature can: the characteristics of the top performers *relative to everyone who tried*.

> **The mean age of the founders of the fastest-growing one in a thousand new ventures is 45. Not 25.**

The popular image of the young founder is not merely unsupported, it is inverted, and the effect holds in high-technology sectors and in entrepreneurial hubs specifically. The same study found that prior experience in the specific industry strongly predicts success.

That is what a finding looks like when someone has the denominator. It is specific, counter-intuitive, replicable, and it contradicts folk wisdom assembled by looking at famous young founders. Note which famous young founders that folk wisdom was assembled from: the ones who came back.

08 So what should you actually do

The point of taking something apart is to say what to do instead. Seven suggestions, in the order you would use them.

1. Demote the category from a prediction to a prior

A shape is not useless. It tells you where to look first and which failure modes have historically eaten companies like yours: marketplaces have a cold start problem, earning platforms have a supply quality problem, cheaper substitutes have a margin problem, localised replication has a why-has-the-incumbent-not-expanded-here problem. That is genuinely valuable, and it is roughly the whole legitimate content of the taxonomy. What it cannot do is tell you that your company will work.

2. Ask for money instead of asking for opinions

The standard validation exercise asks a hundred people to rate their likelihood of buying and counts the top scores. Do not run it in that form. The failure record is full of companies that had a positive read on demand. One in my own sample is a banking platform whose post-mortem says, in as many words, that the company believed it had established product-market fit and it was not enough to keep it from going under.

A stated intention costs the respondent nothing, and it is systematically inflated when they know the person asking built the thing. Raising the bar to nine out of ten corrects the scale, not the mechanism. Replace the rating with something that costs them something: a refundable deposit, a signed letter of intent, a paid pilot. **The number that matters is conversion from approach to payment, and it is meaningless unless you publish how many people you approached.**

3. Count the refusals, and write down the reason for each

The people who say no are the more valuable half of the exercise and the half everybody throws away. Sort refusals into price, trust, switching cost, and genuine absence of need. Those four have completely different implications, and only one of them means the idea is wrong. This is the cheapest way to stop being the person who studies the returning planes.

4. Write down what would make you abandon this, before you look

Decide in advance what result would tell you the idea is wrong. Not a feeling, a number, with a date attached, kept somewhere you cannot quietly edit. Without this step every result is confirmatory, because a sufficiently motivated founder can reinterpret any outcome. This is what separates a test from a ritual, and almost nobody does it.

5. Prefer iteration over prediction whenever you can afford it

Prediction is a substitute for iteration, and it is the expensive substitute. It was worth doing when building took two years and money you did not have. The more cheaply you can put a real thing in front of a real buyer, the less your judgement about which idea is best is worth, and the more your ability to run a clean test is worth.

6. Take the base rate seriously enough to choose your game

Nobody knows the true odds within four orders of magnitude, but every estimate is punishing. That is not an argument for despair, it is an argument for asking what you actually want. If a profitable, independent, twenty-million-dollar company would be a triumph for you, then most of this literature is written about a different objective and will push you into decisions that are wrong for yours.

7. Buy your own control group

Whatever you conclude about your market, find five companies that tried something adjacent and failed, and read their post-mortems before you commit. They are freely available and almost nobody reads them. You will not enjoy it, and it is the highest-return hour in the whole process, because it is the only part of your research that samples the planes that did not come home.

09 Two steps forward: why this gets more wrong, not less

Everything above rests on two conditions that were true when this advice was written and are both now weakening. Neither was ever stated, because at the time neither needed stating. What follows is reasoning about a direction of travel, not a measurement, and should be read as such.

The first condition: that building is the expensive part

The entire apparatus of choosing correctly before you build exists because building used to consume the scarce resource. Two years of engineering spent on the wrong idea was the catastrophic error, and avoiding it justified almost any amount of upfront analysis.

As the cost of producing a working product falls, the value of choosing correctly in advance falls with it, and it falls faster than habits adjust. When you can afford one attempt, you must predict. When you can afford twenty, prediction is the wrong tool, because the market's judgement is better than yours and you can now afford to consult it directly.

If that continues, the scarce input moves. It stops being the ability to build, and it stops being the taste to pick. It becomes **distribution**, because twenty products with no route to a buyer is worse than one with a route, and **the quality of the selection process itself**, because twenty badly run experiments produce twenty ambiguous results instead of one clear one.

The founder skill that rewards is not vision. It is designing a test whose result you cannot argue with. That is a research skill, and essentially nothing in the current literature teaches it.

The second condition: that the buyer is a person who must understand you

Every version of this thesis contains, usually as its first clause, some form of: the customer must grasp the value immediately, without a long explanation. That is a claim about human cognition at the moment of discovery, and it has been correct.

Discovery is moving. A growing share of buying decisions now begins with a question put to a model rather than a search of a page, and the thing that has to comprehend the product first is not the buyer. It is a system summarising the category on the buyer's behalf, and it fails differently from a person.

A human skims your page and forms an impression. A model reads whatever about you is machine-legible, weighs it against everything else it has read about your category, and hands the buyer a comparison instead of your page. What makes a product legible to that process is not what makes it charming to a person: structured and specific claims, an unambiguous category label, corroboration from sources with no stake in you, and specificity that survives being paraphrased by something indifferent to your framing.

If that continues, "must be immediately comprehensible" splits into two requirements that can pull against each other. A product can be instantly clear to a human being and invisible to the system that would have introduced them.

Whether that becomes the dominant channel is exactly the kind of forward claim section 03 warns about. It is offered as the thing to go and measure, not as a finding.

10 The one sentence

Advice assembled from the companies that came back tells you where a company can be hit and survive. It cannot tell you where the fatal shots land, because the companies that took those are not in the sample, and they are the ones you need to hear from.

Go and read the post-mortems. Ask somebody for money instead of an opinion. Write down in advance what would change your mind. Everything else in this literature, including this paper, is commentary.

HOW THE NUMBERS WERE MADE

The two corpora. The billion-dollar population is the complete published list of private companies valued at $1B or more, 1,404 companies with valuation, entry date, country and industry, retrieved 1 September 2026. The failure record is a published compilation of startup post-mortems, 410 companies parsed, of which 266 describe the product clearly enough to classify.

The classification. Seven product shapes, written from the failure descriptions rather than borrowed from a published taxonomy. Sixty companies drawn at random with a fixed seed, each classified from its own description alone, with no reference to the company name and no looking up what happened to it. Eight could not be classified and were excluded rather than guessed at. Every company, its description, its classification and the reason are published in data/paper-11/, so anyone who disagrees can re-classify them and see what changes.

WHAT THIS CANNOT DO, STATED PLAINLY

The winners were **not** classified for product shape, because the list carries no product descriptions and classifying fourteen hundred companies from their names would import exactly the outcome knowledge the method exists to keep out. So the comparison runs against a ceiling of 73.9 per cent rather than a measured figure, and the true headroom could be narrower than ten points. It cannot be much wider. Closing this properly needs product descriptions for a matched sample of winners, classified by somebody who does not know which pile each company came from.

There was also only one classifier, me, where the honest standard asks for two working blind. Both gaps are real and neither is hidden.

> TWO CLAIMS THAT DID NOT SURVIVE THEIR OWN CHECK
>
> **A quotation that was not in the corpus.** An earlier draft attributed a post-mortem line to this failure record from memory. When I searched the corpus for it, it was not there. It has been replaced with one that is, quoted in section 08.
>
> **"Google was the fourteenth search engine."** A figure I have seen repeated for years. I went looking for its source before publishing and could not find one. No contemporaneous count of search engines in 1998 appears to exist. Section 06 now makes the weaker claim the record actually supports.
>
> Both are recorded here rather than quietly fixed, because a paper about unverified claims that did not check its own would deserve nothing.

REFERENCES

1 **Wald, A. (1943)** Work on aircraft survivability, Statistical Research Group, Columbia University.
Widely reproduced; the armour example is the standard account.

2 **Peters, T. and Waterman, R. (1982)** *In Search of Excellence*. New York: Harper & Row.
Subsequent performance of the forty-three companies compiled from later retrospectives, including McKinsey's own review of the cohort.

3 **Collins, J. (2001)** *Good to Great*. New York: HarperBusiness.
Subsequent outcomes for the eleven companies from public filings and reporting on the 2008 crisis.

4 **Rosenzweig, P. (2007)** *The Halo Effect*. New York: Free Press.

5 **Lee, A. (2013)** 'Welcome To The Unicorn Club: Learning From Billion-Dollar Startups', *TechCrunch*, 2 November. https://techcrunch.com/2013/11/02/welcome-to-the-unicorn-club/ *Read at primary. Numerator 39 companies; denominator around 60,000, aggregated by the author from three conflicting estimates which the article names. Accessed 1 September 2026.*

6 **Ellis, S. (2009)** The "very disappointed" survey and the forty per cent threshold. *Origin and current usage traced across practitioner sources. No independent validation of the threshold located.*

7 **Azoulay, P., Jones, B., Kim, J. D. and Miranda, J. (2020)** 'Age and High-Growth Entrepreneurship', *American Economic Review: Insights*, 2(1). https://www.aeaweb.org/articles?id=10.1257/aeri.20180582 *US Census administrative data. Mean founder age 45.0 for the top one in a thousand fastest-growing new ventures.*

8 **On-demand cohort outcomes**, from contemporaneous reporting on Homejoy (TechCrunch and BuzzFeed News, July 2015), Shyp (TechCrunch, 2018), Munchery, Sprig, Zirtual and Good Eggs.

9 **Webvan, Instacart, Pets.com and Chewy**, from contemporaneous reporting and the PetSmart acquisition of Chewy at $3.35B (Axios, December 2017).

10 **CB Insights** 'The Complete List Of Unicorn Companies' and '483 startup failure post-mortems'. *Both retrieved 1 September 2026 and preserved with this paper's data.*

Where sources disagree, as they do by four orders of magnitude on the base rate, all figures are shown and none is averaged.

ABOUT THE RECORD

1,404 companies currently valued at a billion dollars or more, and 410 startup failure post-mortems, both parsed to structured data and published with this paper.

Sixty of the failures were drawn with a fixed seed and classified blind to their outcome. Every row is published with its reason.

WHAT IT SHOWS

The product categories said to predict success already describe 63.5 per cent of the companies that went to zero, against a ceiling of 73.9 per cent on the winners.

And the winning population turns over faster than the advice is revised: consumer companies fell from 41.5 per cent of new entrants to 8.7 between cohorts.

CORRECTIONS AND CONFLICTS

Two claims in this paper failed their own checks before publication and are named in the text rather than quietly removed.

Funding: none. No sponsor, client or vendor paid for, reviewed or approved this work.

broikos.gr