

MUDITA STUDIOS RESEARCH — THE STUDIO SERIES

Kill Fast, Scale Smart

The Economics of Validating Before You Build

Our expected-value model of a venture portfolio finds that cheap early screens cut the capital needed per winner by about two-thirds, mostly by keeping doomed ideas away from scale funding. It also finds when screening stops paying: when good ideas are common, when winners are enormous, or when the screen is barely better than a coin flip.

PUBLISHED
October 2026

BY
Mudita Studios Research

LENGTH
11 min · 5 exhibits

EXECUTIVE SUMMARY

We modelled the economics of validating venture ideas before building an MVP, and when that step stops paying. The model shows studio operators, corporate innovators and investors where capital is lost and what a cheap early screen must achieve to add value.

$4.1M
capital consumed by one false positive in the base case

The costly mistake is scaling a bad idea

That is 54 times a typical studio MVP. Under build first, 89% of all capital goes to scaling ventures that will not win.

65%
cut in capital per winner from validating first, from $82M to $29M

Validating first saves capital

That frees $6.5M of every $10M deployed in the base case. Validation saved capital in all 20,000 simulations, by $4.4M to $7.3M per $10M (P10 to P90).

61%
screen accuracy needed to create net value from a fixed pipeline of ideas

Saving capital and creating value are different tests

A one-week, $2,500 screen pays for itself in capital at 50% accuracy. 61% is roughly Y Combinator-level selection, and no study yet measures how accurate real screens are.

25%
of simulations in which validation lost on value

Validation is not always the right call

Building first wins once more than 3.9% of ideas are winners, once winners return more than 49x, or once the MVP test already stops more than 84% of bad ideas.

01 The expensive mistake is not the MVP

Startups mostly die of market risk. In CB Insights' 2026 analysis of 431 VC-backed shutdowns, poor product-market fit was a factor in 43%. The median company had raised **$11M** before it closed. Building is comparatively cheap: in Inniches' 2024 survey of 123 venture studios, about 63% spend $100K or less on an MVP, and half reach one in 3–4 months.

The question is how much validating *before* the MVP saves, and when that step costs more than it returns.

> **$11M** median capital raised by VC-backed startups before they shut down (CB Insights, 2026)

02 How the model works

Each raw idea is either a **winner**, meaning it would return at least about 10x if built and scaled, or not. In the base case 2% of ideas are winners, close to the 1.9% random-pick success rate in the VCBench

founder dataset. Every stage is a noisy test that passes some winners and lets some non-winners through.

- **Build first:** every idea gets an MVP; those that look good in the market get scale capital.
- **Validate first:** ideas pass a desk/AI screen and then a demand test (landing pages, fake doors, paid traffic, interviews) before any MVP.

Validation has two costs. Screens wrongly kill some winners, and the 1.75 months of extra validation costs winners 1% of their value per month in the base case. We score each strategy two ways:

- **Cost lens:** capital per winner, assuming ideas are plentiful.
- **Value lens:** net value from a *fixed* pipeline of ideas, where every winner killed by mistake is gone.

Every input has a low, base and high value and is either sourced or labeled as an assumption (Table 4). The Monte Carlo draws all inputs from those ranges.

---

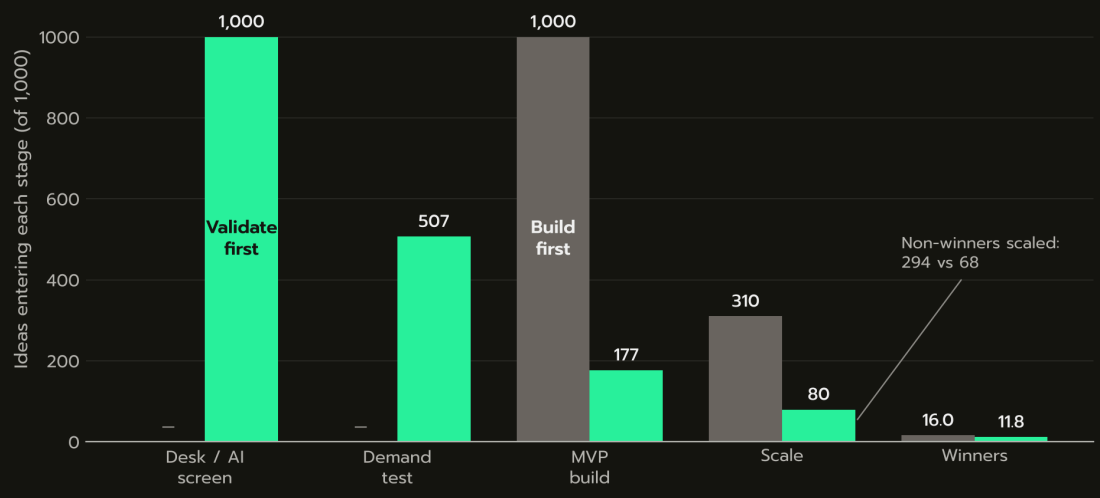
03 Finding 1: The money is lost at scale

Exhibit 1 follows 1,000 raw ideas. Under build first, all 1,000 get MVPs, 310 get scale funding and 16 become winners, so **294 scaled ventures were never going to win**. Under validate first, 177 ideas reach an MVP and only **68 non-winners reach scale, 4.3 times fewer**. The price is winners: the screens keep 11.8 per 1,000 ideas against 16.

EXHIBIT 1

Validating first sends 4.3x fewer doomed ideas to the scale stage

Ideas entering each stage per 1,000 raw ideas, base case



Source: Mudita Studios expected-value model, base-case inputs (sources and ranges in Table 4). Winners: 16 vs 11.8 per 1,000 ideas.

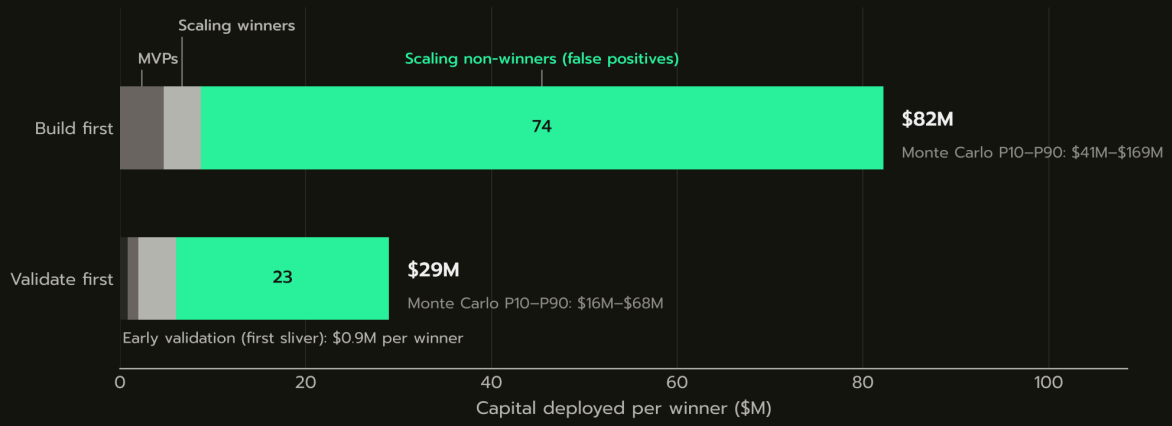
MUDITA STUDIOS

Exhibit 2 shows where the capital goes. Build first spends $4.7M per winner on MVPs and **$74M on scaling ventures that fail**. Validate first spends $0.9M on early validation, $1.1M on MVPs and $23M on failed scale-ups. Capital per winner falls from **$82M to $29M**, so $10M buys about 2.8 times as many winners.

EXHIBIT 2

Validating first cuts capital per winner by 65%, mostly by not scaling losers

Capital deployed per winner, $M, base case; mint = scale capital spent on ventures that fail



Source: Mudita Studios expected-value model, base-case inputs (sources and ranges in Table 4). Monte Carlo: 20,000 runs over the low/base/high ranges.

MUDITA STUDIOS

As a calibration check, about 5% of scaled ventures under build first are winners, close to Correlation Ventures' finding that 4% of financings return 10x or more.

> **4.3x** fewer non-winners reach scale funding when ideas are validated first

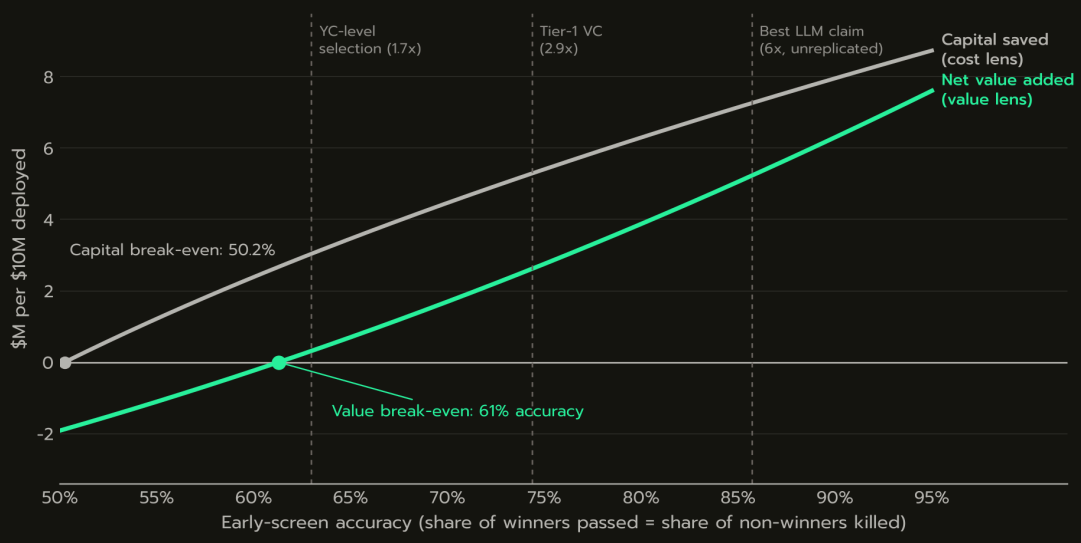
04 Finding 2: A screen can save capital without creating value

Exhibit 3 places a single $2,500, one-week screen before the MVP and varies its accuracy. Here accuracy means the screen passes the same share of winners as the share of non-winners it kills.

EXHIBIT 3

A cheap screen saves capital above 50% accuracy and adds value above 61%

Gain from one $2,500, one-week (0.25-month) screen before the MVP vs building every idea, base case



Source: Mudita Studios model. Reference lines map VCBench (2025, preprint) selection lifts to accuracy as lift/(1+lift); the mapping is approximate.

MUDITA STUDIOS

- **Cost lens: about 50%.** A screen barely better than a coin flip still saves capital. It costs $2,500, a false positive costs $4.1M, and a wrongly killed winner can be replaced by screening another idea.
- **Value lens: about 61%.** When the pipeline is fixed, every false negative is a winner lost. The screen then needs roughly the discrimination of Y Combinator's selection, which the VCBench preprint puts at 1.7 times random (about 63% on our scale). Tier-1 VC selection (2.9 times random, about 74%) clears the bar comfortably. Claims of LLMs reaching six times random would clear it easily, but they come from VC-affiliated preprints and have not been replicated.

Table 1: Even an expensive screen breaks even on capital near 54% accuracy; creating value takes 61–64%

Early screen	Cost per idea	Time	Break-even accuracy, cost lens	Break-even accuracy, value lens
Desk / AI screen (base cost)	$2.5K	0.25 mo	50.2%	61.3%
Demand test (base cost)	$15K	1.5 mo	51.4%	62.3%
Expensive screen	$50K	2 mo	53.5%	64.0%

Source: Mudita Studios model, base-case inputs. VCBench selection lifts converted to accuracy as lift/(1+lift), an approximation.

Accuracy matters far more than price: a screen 20 times costlier raises the value break-even by under three points.

To bound the demand test's accuracy we used the best causal evidence available. In Camuffo et al.'s randomized trial of 116 startups, founders taught to test hypotheses pivoted away from their original idea 67% more often than controls (19 of 59 against 11 of 57), without dropping out more. Koning, Hasan and Chatterji found that startups adopting A/B testing improved performance by 30–100% and "failed faster" on bad ideas. Neither study measures gate accuracy, so we use them only to bound how much structured testing improves the kill rate of bad ideas: +16%, +67% and +100% in the low, base and high cases.

05 Finding 3: The result holds across most of the uncertainty, not all of it

Across 20,000 simulated portfolios (Exhibit 4), validation saved capital in every run, with a median of $6.0M per $10M. It added value in **75%** of runs, with a median gain of $1.8M per $10M. In the worst tenth of runs it destroyed $2.0M or more.

Table 2: Validation wins in the base and favorable cases, and loses when every input goes against it

Scenario	Capital per winner: build first	Capital per winner: validate first	Change	Capital saved per $10M	Net value change per $10M	Team-months per winner (build / validate)
Unfavorable to validation	$5.4M	$8.9M	+64%	-$6.4M	-$37.7M	75 / 230
Base case	$82.2M	$29.0M	-65%	$6.5M	$3.1M	250 / 145
Favorable to validation	$716.6M	$36.6M	-95%	$9.5M	$9.3M	1,125 / 111

Source: Mudita Studios model. The outer scenarios set every input to the end that hurts or helps validation at once; they are extreme brackets, not forecasts.

Validation also saves team time (145 team-months per winner against 250), though each winner waits an extra 1.75 months for its MVP.

06 Where validation does not pay

1. When good ideas are common. The winner base rate is the biggest driver of the value result. With other inputs at base, building first creates more value once more than **3.9%** of ideas are winners. A studio with privileged deal flow, such as deeply experienced domain teams or a corporate builder with a proven channel, should screen less.

2. When winners are enormous. If a winner returns more than about **49x** its capital, losing even a few to false negatives costs more than screening saves. On the value lens, the power-law objection holds.

3. When the MVP is already a good test. If the post-MVP market test already stops more than **84%** of bad ideas, an early screen adds little. Cheaper MVPs alone do not flip the result, because the money is lost at scale; a much *sharper* MVP test would.

4. When speed is everything. Delay alone rarely flips the base case: validation keeps its edge until each month costs a winner about 22% of its value. But delay compounds with a high base rate (Exhibit 5). At a 3.5% winner rate the break-even falls to about 3.5% of value lost per month, inside our modelled range.

Table 3: Validation stops adding value when winners are common or huge, or the MVP test already filters well

Input (others at base)	Base value	Validation stops saving capital at	Validation stops adding value at
Share of ideas that are winners	2%	92%	3.9%
Winner value multiple	25x	never	49x
MVP test: share of non-winners stopped	70%	97%	84%
Value lost per month of delay	1%/mo	45%/mo	22%/mo
MVP cost ($1K–$20M tested)	$75K	never	never

Source: Mudita Studios model, one input varied at a time.

A fifth failure sits outside the model: a good screen used badly. In a field trial with 640 Kenyan entrepreneurs (Otis et al.), AI advice helped high performers by about 15% and hurt low performers by about 8%.

> **8%** drop in performance for low performers given AI advice in a field trial with 640 Kenyan entrepreneurs, while high performers gained about 15%

07 Where AI screens fit

An AI desk screen, such as Mudita's Miles, is the kind of cheap first gate this model describes. The 61% accuracy bar in Exhibit 3 is the test any such tool has to pass before it adds value rather than just saving capital. Research on LLM evaluation (Doshi et al., 2025) suggests aggregating many AI judgments rather than relying on a single one.

08 What this means

Studio operators should measure their screens' accuracy against later outcomes, not just their cost. Corporate innovators get the biggest savings from a hard gate before scale funding. LPs should ask studios how often their screens turn out wrong; without that, "kill fast" may save capital without creating value.

IMPLICATIONS

What this means for studio operators, corporate innovators and LPs

01 Measure screens' accuracy against later outcomes as well as their cost. A screen needs about 61% accuracy to create value.

02 Put a hard gate before scale funding. That is where most capital is lost on ventures that will not win.

03 Screen less when good ideas are common or winners are enormous. Building first can then create more value.

04 Ask studios how often their screens turn out wrong. Without that, kill fast may save capital without creating value.

APPENDIX

Method, data and sources

Methodology and limitations –

Each gate passes winners with one probability (sensitivity) and non-winners with another (1 - specificity). Capital is counted across all investors, from screens to post-MVP scale capital ($1M–$11M, bounded by CB Insights' median equity raised before shutdown).

A winner is worth 10–40 times its capital: the low end is Correlation Ventures' 10x bucket, the high end our inference from Horsley Bridge data (about 4.5% of dollars produced about 60% of returns). Non-winners that fool the early screens are assumed more likely to fool the MVP test too. The Monte Carlo draws 20,000 input sets from independent triangular distributions. The model inputs and code are available on request.

Limitations.

- **Gate accuracies are assumptions.** No study measures how well desk screens, demand tests or MVP tests predict venture outcomes. The studies we cite bound our ranges; translating them into gate accuracy is our own judgment.
- **Outcomes are binary.** Grouping 1–5x outcomes with failures overstates the cost of false positives, which favors validation.
- **Inputs are drawn independently.** In reality fast markets may also produce bigger winners, which would fatten the downside for validation.
- **Pivots and founder effects are not modelled.**

- **The evidence on studios cuts both ways.** Independent PitchBook-based data (Inniches, 2024) found studio-backed pre-seed deals exited 24% of the time, against 38% for traditional VC-backed deals. This model does not show that studios outperform. It shows that validating first, done well, uses capital more efficiently than building every idea.
- **Some inputs are self-reported or secondary:** MVP cost and timing come from studio survey answers; Correlation Ventures and Horsley Bridge figures from secondary write-ups.

Appendix

Table 4: Every model input is sourced, derived from a source, or labeled as an assumption

Input	Low	Base	High	Basis
Share of raw ideas that are winners	1%	2%	5%	Derived: VCBench 1.9%; Correlation 4% ≥10x; Inniches funnel
Winner value (multiple of capital consumed)	10x	25x	40x	Derived: Correlation Ventures; Horsley Bridge
Capital recovered from a failed scaled venture	0%	15%	35%	Assumption, bounded by Ghosh; Correlation
Desk / AI screen: cost per idea	$500	$2.5K	$10K	Assumption
Desk / AI screen: time	0.1 mo	0.25 mo	1 mo	Assumption
Desk / AI screen: winners passed	75%	85%	92%	Assumption, bounded by VCBench lifts
Desk / AI screen: non-winners killed	30%	50%	70%	Assumption, bounded by VCBench lifts
Demand test: cost per idea	$5K	$15K	$40K	Assumption
Demand test: time	1 mo	1.5 mo	3 mo	Assumption
Demand test: non-winners killed	35%	67%	90%	Derived: Camuffo et al.; Koning et al.
Demand test: winners passed	75%	85%	92%	Assumption
MVP: cost	$35K	$75K	$200K	Inniches survey (industry survey)
MVP: time	3 mo	4 mo	9 mo	Inniches survey (industry survey)
MVP test: winners passed	70%	80%	90%	Assumption
MVP test: non-winners stopped	50%	70%	85%	Assumption, checked against Inniches funnel
Scale capital per venture	$1M	$4M	$11M	CB Insights 2026
Winner value lost per month of delay	0%	1%	4%	Assumption
Error correlation, early screens vs MVP test	0	0.3	0.8	Assumption

Source: Mudita Studios; sources listed below.

Table 5: "Success" falls from a coin flip to about 2% as the definition tightens

Measure	Rate	Source
New U.S. establishments surviving 5 years	49.6%	BLS
VC-backed pre-seed deals that exited (traditional VC / studio)	38% / 24%	Inniches (PitchBook data)
VC-backed companies never returning cash to investors	~75%	Ghosh, HBS
Venture financings returning less than 1x / 10x or more	65% / 4%	Correlation Ventures
Founders successful at inception (random pick)	1.9%	VCBench (preprint)

Source: as listed.

Sources

- BLS (2018). Establishment survival. https://www.bls.gov/opub/ted/2018/18-percent-of-businesses-opening-in-march-1994-remained-open-as-of-march-2017.htm
- Camuffo, Cordova, Gambardella, Spina (2020). *Management Science* 66(2). https://Www.Gwern.net/doc/economics/2019-camuffo.pdf
- CB Insights (2026). The top reasons startups fail. https://www.cbinsights.com/research/report/startup-failure-reasons-top/
- Correlation Ventures, via Seth Levine (2014). https://sethlevine.com/wp/2014/08/venture-outcomes-are-even-more-skewed-than-you-think
- Doshi, Bell, Mirzayev, Vanneste (2025). *Strategic Management Journal* 46(3). https://ideas.repec.org/a/bla/stratm/v46y2025i3p583-610.html
- Ghosh, via HBS Working Knowledge (2012). https://www.library.hbs.edu/working-knowledge/why-companies-failand-how-their-founders-can-bounce-back
- Horsley Bridge, via a16z (secondary, c. 2015). https://marginalfutility.substack.com/p/the-power-law-in-venture-capital
- Inniches (2024). Big Venture Studio Research. https://inniches.com/big-venture-studio-research
- Koning, Hasan, Chatterji (2022). *Management Science* 68(9). https://www.nber.org/papers/w26278
- Otis, Clarke, Delecourt, Holtz, Koning (2026). https://www.hks.harvard.edu/centers/cid/publications/uneven-impact-generative-ai-entrepreneurial-performance
- VCBench (2025, preprint). https://arxiv.org/abs/2509.14448