

FIELDWORK · CORPORATE-STARTUP COMMERCIALIZATION

What Makes a Startup-to-Corporate Pilot Turn Into a Customer?

We looked at what actually happened inside 26 corporate startup programs across defense, healthcare, heavy industry, government procurement, telecoms, consumer goods and the accelerator business, covering roughly 1982 to 2026, to work out why some pilots became real customers while others stopped at the end of the experiment.

26

PROGRAMS EXAMINED

7

INDUSTRIES

1

APPROXIMATE COUNTERFACTUAL

THE ANSWER

The buyer existed before the pilot did.

Across the strongest converting programs, four things were usually in place before the startup arrived: what the problem was, who inside the company owned it, how a purchase would happen if the pilot worked, and which number would decide.

The startups in those programs were not running an experiment and then hunting for a budget. They were running an experiment for someone who already had one.

Where those four things were missing, the pattern is remarkably consistent. Introductions happen, meetings happen, technically successful pilots happen, and then nothing happens. The European Commission's own May 2025 assessment of the European market puts a number on the endpoint: fewer than 1 percent of startup projects submitted to corporates make it to market. **FACT**

The tempting shorthand for all of this is "paid pilots convert better." That is close, and it is wrong in a way that matters. Payment turned out to be the most reliable *signal* of the condition rather than the condition itself, because a purchase order forces budget, vendor onboarding, security review and named ownership to become real at the same moment. One program in the study converted at 40 percent with an entirely unpaid pilot, and it did so by clearing the procurement route in advance instead. That case is covered below, and it is the reason the finding is stated the way it is:

What predicts conversion is a cleared path to purchase, owned by the person accountable for the problem, established before the pilot begins.

TWO SEQUENCES

The same activities in a different order produce very different outcomes.

Almost everything comes down to the order

Both models involve a corporation, a startup, an introduction and a pilot. The difference is what exists before the introduction.

INTRODUCTION-LED

Startup → Meeting → Pilot → ~~Find budget~~ → ~~Find buyer~~

The two hardest steps sit at the end, after the momentum is spent and the budget cycle has moved on.

BUYER-LED

Problem → Owner → Purchase path → Startup → Pilot → Rollout

The hard steps are resolved first, when they are cheap, and the pilot only tests the one thing still genuinely unknown.

THE EVIDENCE

Company-reported figures are marked. Independent and audited findings are marked separately.

What happened when organizations tried this

BMW built its unit around the purchase order

BMW set up the Startup Garage in 2015 on a premise that sounds obvious and almost nobody was following: rather than invest in a startup, become its customer. Gregor Gimmy, who created the unit, called the resulting transaction a "minimum viable purchase." A business unit brings a problem it already owns. The unit finds a company that solves it. BMW buys a prototype.

The mechanics are stated plainly on BMW's own site, in language aimed at founders: "BMW becomes your first automotive client, not your shareholder. Once accepted into the program, you get a supplier number and a purchase order. The money you receive when accepted to the program is a payment for a prototype of your technology and the costs associated with its integration and validation. This payment accounts as revenue." FACT

Notice how much that one sentence does. The supplier number is procurement onboarding, completed before the work rather than after it. The purchase order requires a budget line, which requires a business unit, which requires a real problem. Pilots run four months, conducted by the actual problem owner.

BMW reports 6,000 or more startups assessed, more than 280 pilot projects with companies from 26 countries, and over 50 percent of alumni securing paid follow-up projects. INTERESTED-PARTY SOURCE Those are the company's own figures with no

published methodology. Harder to dismiss is the operating detail Alexandra Renner, who runs the unit, gave in a 2025 interview: surface cybersecurity and certification requirements early, design contracts a startup can actually sign, and budget with funnel math in mind at roughly a 20 percent hit rate. SECONDARY She described the work as first and foremost a people business requiring credibility with purchasing.

Holcim runs a near-identical process in construction materials and gives a soberer number. Its unit head, Bengt Steinbrecher, said in early 2026 that about half of pilots succeeded and about half of those led to some adoption, which is roughly a quarter reaching deployment across 25 or more pilots a year. SECONDARY His example is worth keeping because it is unglamorous. Plant managers were climbing 20-metre silos to check cement levels, so the unit found a company making external fill-level sensors, now deployed in more than 10 countries. That pilot did not begin with a startup. It began with a safety problem a named plant manager owned.

Europe's largest matchmaker published its own conversion rate, then went shopping for a fix

The European Innovation Council is the cleanest contrast available, because it is a public body with no commercial incentive to flatter itself.

Between 2017 and 2025 its Corporate Partnership Programme ran 91 Corporate Days and Multi-Corporate Days, facilitated 2,493 one-to-one meetings between EIC-backed companies and more than 100 large European corporations including Airbus, BMW, Roche, Shell, Siemens Energy and Telefónica, and produced more than 100 business deals. Satisfaction among participating companies ran at 92 percent. FACT

That is roughly 4 percent of meetings becoming deals, with nearly everyone reporting themselves pleased. The EIC's published denominators vary between meetings and engagements and are not reconciled in the source material, so the precise figure moves. UNRESOLVED The order of magnitude does not.

What the EIC did next is the most useful single fact in this study. In June 2026 it issued a tender for the next version of the program. The new services are intended "to increase business conversion and long-term impact," supporting companies "throughout the entire collaboration journey, from initial matchmaking and pilot projects to commercial adoption and long-term business partnerships," including promoting venture clienting. The selected contractor will be responsible for "post-matching follow-up and business conversion." FACT

The organization that ran the introduction model at greater scale than anyone in Europe, for eight years, went to market to buy the part that happens after the introduction. Its own July 2025 report is also unusually candid about format: in-person single-corporate engagements deliver the highest number of business deals, while online and multi-corporate formats increase reach and visibility. FACT Those are two different products, and the EIC says so.

The pilot that worked and still lost

Joshua Liu founded SeamlessMD and has written a first-hand account of the pattern from the vendor side. Clinical teams told his company to prove ROI with a pilot before involving hospital executives at all. So they ran pilots of six to twelve months, invested heavily in implementation, and produced results the clinical teams described as excellent, including reductions of around 50 percent in

emergency department visits or readmissions, often statistically significant. Then came the meeting with leadership.

Josh, these results are fantastic. And I can see how SeamlessMD is great for patients too. But actually, reducing readmissions isn't a priority for us.

A HOSPITAL CEO, QUOTED BY JOSHUA LIU · SECONDARY, OPERATOR ACCOUNT

Liu writes that he first assumed it was an anomaly, and then recognized it as a pattern. Nothing about that pilot failed except the part that decided the outcome, and that part was decided months earlier, when people who could not authorize a purchase chose what the pilot would measure.

This is the difference between technical success and buyer success, and it is not a question of rigor. It is a question of who picked the metric. The academic literature reaches the same place: the NASSS framework, published in 2017 after studying why health technologies fail to move beyond early adoption, concludes that sustained adoption depends less on the technology than on the organization surrounding it. **FACT**

Same government, same startups, two different mechanisms

The strongest natural experiment in the study sits between two United States programs that fund small companies to build things the government might buy.

The Defense Innovation Unit takes a capability gap from a military sponsor and defines success as transition to a production contract. From fiscal 2016 through 2023 it awarded 450 prototype agreements and transitioned 62 to production, about 51 percent of projects with a completed prototype. **FACT, GAO-25-106856** The feature that explains the number is who paid. The sponsoring military organization co-funds its own prototype, and outside analysis of DIU's published funding chart suggested nearly 90 percent of prototype money came from partner organizations.

SECONDARY Transition is not a handoff to a new budget. It is the same budget continuing.

SBIR funds a superficially similar activity and produces a very different result. Phase I funds feasibility, Phase II funds development, and Phase III has no SBIR funding at all. It is a contracting status rather than a check. **FACT** The gap is not a bug that emerged in practice, it is written into the design, and the SBA's commercialization benchmark for repeat awardees is an average of \$100,000 in sales or investment per Phase II award over ten years.

The Air Force then ran the experiment on itself. Its Open Topic SBIR had a documented failure mode: a firm could win Phase I, finish it, and find no program office ready to fund Phase II, because the original pitch had no program office customer. **SECONDARY** AFWERX made customer co-investment and end-user letters effectively mandatory. On the other side of that gate it reports 84 percent of STRATFI companies and 69 percent of TACFI companies transitioning to Phase III.

FACT, PROGRAM-PUBLISHED Selection there is plainly endogenous, and that is the point.

The gate is the mechanism, and the gate is made of the conditions this study keeps finding. **EVIDENCE-BACKED INFERENCE**

Published outcomes, not directly comparable.

PROGRAM	WHO FUNDED THE PILOT	PUBLISHED CONVERSION
---------	----------------------	----------------------

PROGRAM	WHO FUNDED THE PILOT	PUBLISHED CONVERSION
AFWERX STRATFI	Customer, matched	84% to Phase III
AFWERX TACFI	Customer, 1:1 cost share	69% to Phase III
Maersk Growth	Business unit	>60% of pilots scaled
BMW Startup Garage	Business unit, purchase order	>50% paid follow-on
Defense Innovation Unit	Military sponsor, co-funded	51% to production
Unilever Foundry	Brand budgets	~48% to partnership
Startup in Residence	Nobody. Unpaid residency	40% issued contracts
Holcim MAQER	Business unit	~25% to adoption
EIC Corporate Partnership	Nobody. Free matchmaking	~4% meeting to deal
European corporates, all routes	Various	<1% projects to market

BMW, Maersk, Holcim and Unilever figures are self-reported by interested parties without published methodology. AFWERX and EIC figures are program-published. DIU's is audited by the GAO. The European Commission figure is from a May 2025 staff working document. "Scaled," "adoption," "follow-on" and "deal" are defined differently by each organization and the percentages are not directly comparable.

The best-evaluated program in the study is the one that fell short

SBRI Healthcare is worth more than any success story here, because RAND Europe evaluated it independently under a Department of Health commission, and because it is the only program in the study where anyone bothered to survey the companies that were rejected.

The design is good on paper and good in practice, upstream. NHS clinicians identify unmet needs. Themed competitions invite companies to propose solutions. Phase 1 awards up to £100,000 over six months for feasibility, Phase 2 up to £1 million over 12 months. Between 2013 and 2016 the program awarded £57 million across 168 contracts. **FACT** RAND found the identification and articulation of needs to be a genuine strength.

Then look at who owns the purchase. The money is a national R&D contract, not the operating budget of the hospital trust that might buy the thing. No individual trust is obligated to purchase anything. Selection is by expert panel rather than by the buyer. Procurement is engaged at no point in the process.

The results, from the independent evaluation: of 45 successful applicants surveyed, 13 reported any product sales at all, totalling £4 million, of which £3 million was to the NHS, and one company accounted for 37 percent of the total. 57 percent of awardees named resistance to innovation within the NHS as a barrier, describing a lack of motivation and accountability for uptake. 25 percent cited procurement processes. 37 percent said they received no help at all accessing NHS customers. **FACT** Program stakeholders were explicit that supporting adoption sat outside the program remit, with 11 of 16 interviewees pointing to regional networks as responsible instead. Eight Phase 3 awards aimed specifically at accelerating adoption were made in 2013/14. None have been made since.

For all intents and purposes, [the programme] operates like a grant.

AN SBRI HEALTHCARE AWARDEE, QUOTED IN THE RAND EUROPE EVALUATION

Another awardee described the commercial reality of a decentralized health system: instead of ten representatives of all trusts, "250 representatives of individual Trusts to sell to." A third said the procurement cycle runs two to three years and that it was easier to focus on faster-adopting markets abroad. FACT

The lesson is not that the program was badly run. RAND found the opposite. The lesson is that a well-defined problem and generous funding produce development rather than deployment when nobody owns the purchase.

THE USEFUL CONTRADICTION

The case that broke the first version of the finding.

An unpaid pilot that converted at 40 percent

The working conclusion, halfway through this research, was that paid pilots convert and unpaid ones do not. Then Startup in Residence turned up.

STIR is a 16-week residency in which a startup works with a government agency on a stated civic challenge. The residency is pro bono. The startups volunteer their time and receive nothing. In 2018, roughly 40 percent of participating companies were issued contracts after the challenge period. FACT, PROGRAM-REPORTED By 2019 the program spanned 22 governments, with around 700 applicants competing for 40 places.

Forty percent from an unpaid engagement sits inside the range the paid programs produce. The explanation is in the program design rather than in the goodwill of the participants. San Francisco treats the STIR application itself as a competitive bidding process, so a successful startup moves directly into contract negotiation at the end of the residency with no separate procurement to run. FACT In the case of Binti, which built a foster care application tool for the city's Human Services Agency, the contract sat below the agency's \$100,000 threshold for competitive sourcing, so no bidding process was required at all.

The money was not spent up front. The procurement was. The agency that wrote the challenge was the agency that would sign, and the route between the two had been cleared before the work began.

That is why the finding in this study is worded around a cleared path to purchase rather than around payment. Payment remains the single most useful early test, because it is difficult to fake and it forces four other conditions into existence simultaneously. It is a proxy for the mechanism and not the mechanism itself. Where an organization can clear the route another way, it gets the same result without the check.

The same case forces a second correction. SBRI Healthcare and the UK's GovTech Catalyst both define problems exceptionally well and both struggled at adoption. A problem defined by a clinician, an engineer or a policy team is not the same artifact as a problem owned by someone with a budget. These are two conditions, not one, and the second is the one that predicts conversion.

Where the chain breaks

The sequence that appears repeatedly in the strongest converting cases is short enough to hold in your head:

THE CHAIN

Problem an operator owns → Named owner → Path to purchase → Pilot on the buyer's metric → Rollout

Six failure mechanisms recur across the cases, roughly in order of how often they appear.

- 01

An innovation mandate with no operating owner.

Most of the failing programs never produce the artifact at all. There is a cohort, a category or a mandate, and no written problem attached to a person. Coca-Cola Founders selected entrepreneurs in ten cities, in several cases before they had a startup idea, and closed within weeks of the CEO who blessed it announcing his departure.
- 02

The pilot proves the wrong thing.

SeamlessMD's readmissions result was real, rigorous and irrelevant to the person holding the budget. This happens whenever the pilot is designed by people who cannot buy.
- 03

Procurement and security arrive after technical validation.

BMW issues the supplier number at acceptance. The failing pattern brings legal, IT and procurement in at the end, when integration or compliance objections can still kill a finished pilot.
- 04

The rollout needs a second, separate funding decision.

Pilots funded from a discretionary innovation line have to win money again, in a different budget cycle, from a different decision-maker, after the momentum has gone. SBIR institutionalizes this gap by barring its own funds from Phase III.
- 05

The champion leaves.

Techstars' \$80 million J.P. Morgan-backed program ended when the bank did not recommit, taking 17 percent of the company's staff with it. Programs that survive turnover are the ones where a business unit inherited an operating line rather than an individual's enthusiasm.
- 06

Nobody owns the relationship after the introduction.

This is the mechanism the EIC identified in its own program and then went to market to purchase.

WHAT MATTERED LESS

Five factors that looked important going in and did not survive the cases.

What did not matter as much as expected

EXECUTIVE SPONSORSHIP

This is where the evidence diverges most from conventional advice. Senior sponsorship reliably predicts that a program gets funded and launched. It did not predict that individual pilots converted, and in two cases it predicted the program's death. Coca-Cola Founders had a VP of Innovation and Entrepreneurship,

presence in ten cities and CEO-level blessing. Techstars had an \$80 million institutional commitment. Both ended on a change of mind at the top.

Plug and Play's own partner materials draw the distinction without meaning to. The corporate annual fee, they write, "guarantees that we have C-level buy in and an SVP champion to help promote the efforts from within the corporation itself." FACT The C-level authorizes the program. The SVP has to make something happen inside a business unit. Only one of those predicts a purchase.

EVIDENCE-BACKED INFERENCE

PILOT LENGTH

Converting pilots in this study ran ten weeks, three to six months, four months, and up to twelve months. The pilot that failed most instructively ran six to twelve months. No relationship is visible in either direction. NOT SUPPORTED

DEMO DAYS

The EIC ran the largest documented test of this and found that in-person single-corporate formats deliver the most business deals, while online and multi-corporate formats increase reach and visibility. FACT Demo Days are a discovery and visibility instrument and they work at that. The evidence does not support treating them as a reliable commercial-conversion mechanism, and the recurring error is pricing a reach product as a deals product.

TECHNICAL SUCCESS ON ITS OWN

Covered above, and worth restating in one line: a pilot that hits its numbers against criteria the buyer did not choose is not a partial win. In the cases here it is indistinguishable from a loss.

THE INTERMEDIARY'S PRESENCE, AS SUCH

The most reliable converters here are not intermediaries at all. BMW, Holcim, Unilever, Wayra and DIU run the process in-house. The two independent intermediaries that do produce pilot-level outcomes, Plug and Play and L Marks, are paid by the corporate, which makes them an extension of the buyer. Intermediaries that stay on the seller's side do not produce contract-level outcomes and mostly do not claim to. EVIDENCE-BACKED INFERENCE The intermediary has to be working for the buyer, and from the outside the two arrangements look identical.

HISTORICAL DRIFT

Everyone is moving the same direction

Across the cases there is one directional trend and no documented counter-example. Organizations move from showcase toward purchase. None moved back.

WAYRA

Founded 2011 as a Telefónica accelerator. Repositioned around 2018 into a venture client operation connecting portfolio companies into business units and the customer base. Now reports revenue on both sides: over €1.06 billion for Telefónica and €615 million for the startups. INTERESTED-PARTY SOURCE

EIC	Eight years of matchmaking, 2,493 meetings, then a 2026 tender for post-matching follow-up, business conversion and venture clienting.
CEDARS-SINAI	The original accelerator ran three months with a \$100,000 to \$120,000 investment and a Demo Day at a conference. Its replacement, Accelerator+, runs twelve months and describes its own purpose as fast-tracking "paid pilots linked to enterprise contracts." No conversion data has been published for either.
UNRESOLVED	

WHY IT MATTERS NOW

AI is creating a lot more opportunities to repeat the same mistake

Enterprises evaluating AI vendors are running pilots in unusual volume, which makes the historical pattern unusually relevant right now.

What the record says is that pilot volume and pilot conversion are close to unrelated. The EIC generated 2,493 high-quality meetings and a 4 percent deal rate. SBRI Healthcare deployed £57 million across 168 contracts and 13 of 45 surveyed companies reported any sales. Both were competently run. Both were missing the same thing.

An organization about to run fifty AI pilots this year can find out which ones will matter without waiting for the results. Ask who owns each problem, whether that person has budget, what happens contractually if the pilot succeeds, and who chose the number that decides. Where those answers do not exist yet, the pilot is a research project, and it should be funded and staffed as one.

RESEARCH LIMITATIONS

What this study cannot tell you

The percentages are not a benchmark. The gap between the buyer-funded programs and the matchmaking programs is large and consistent, and it should be read as a difference in kind rather than as a calibrated ratio. "Scaled" means at least one repeat project at Unilever, "some adoption in the organisation" at Holcim, "paid follow-up project" at BMW, and "production or service contract" at DIU. These are different things measured different ways. Do not treat any range in this study as a standardized figure.

Most conversion rates are self-reported by parties selling something. BMW, Maersk, Holcim, Wayra and L Marks all publish their own numbers without methodology, and all of them use those numbers commercially. They are used here as evidence of what these organizations choose to measure, which is the finding that matters, rather than as verified quantities.

Programs publish when the numbers are good. Quiet successes inside introduction-driven programs are invisible in this study, which means it probably overstates how reliably matchmaking fails. The failure cases are documented precisely because they failed publicly, and each had company-specific causes that are not reducible to pilot conversion.

Industries are not interchangeable. Automotive and construction pilots test equipment in operating environments where the business metric is obvious. Healthcare involves clinical workflow change, regulation and a payer whose incentives differ from the provider's. Defense operates under contracting law with no commercial equivalent. The mechanism appears to transfer. The base rates almost certainly do not.

The counterfactual has been approximated once. This is the caveat that should travel furthest. RAND Europe also surveyed 173 companies whose SBRI Healthcare applications were rejected. 55 percent developed their ideas anyway, and of the 88 who reported sales data, 29 reported sales totalling roughly £4 million, which is similar to the total from the awardees. **FACT** RAND flags every reason not to over-read this: the groups are not comparable, unsuccessful applicants arrived with more mature ideas, response rates were 22 and 45 percent, and all of it is self-reported. **UNRESOLVED**

Those caveats are real. It remains the case that across 26 programs the counterfactual was approximated once, and when it was, the program's effect on commercial outcomes did not show up in the numbers. Every conversion rate in this study measures what happened to companies that were selected. None of them isolates what the program contributed.

PASTBEHAVIOR

Fieldwork No. 039 · August 2026 · Approximately 26 programs examined across defense, healthcare, construction, telecommunications, consumer goods, public procurement and the accelerator industry.

Evidence labels used throughout: **FACT** denotes a claim traceable to a primary or audited source. **SECONDARY** denotes reporting at one remove. **INTERESTED-PARTY SOURCE** denotes a figure published by an organization with a commercial interest in it. **EVIDENCE-BACKED INFERENCE** denotes an interpretation drawn from the cases rather than stated by a source. **UNRESOLVED** denotes a conflict or gap in the public record that has been reported rather than reconciled.

Principal sources include GAO-25-106856 on the Defense Innovation Unit, the RAND Europe evaluation of SBRI Healthcare commissioned by the UK Department of Health, European Innovation Council programme publications and its 2026 Corporate Partnership Programme tender, European Commission staff working document SWD(2025) 138, and program materials published by BMW Startup Garage, AFWERX, Holcim MAQER, Unilever Foundry, Wayra, Plug and Play, L Marks and City Innovate.