Open Future Forum

OPEN FUTURE FORUM

YC Founder AI Report

NOVEMBER 2026

November 2026 | Edition 4 | Research report | Murray Newlands | Open Future Forum

Prepared 2 November 2026. Evidence cutoff: 31 October 2026, 23:59 America/Los_Angeles. Pricing and buyer-path measures are carried forward, with new commercial analysis. This independent OFF report is not affiliated with or endorsed by Y Combinator.

About this publication: Open Future Forum publishes research for executive and investor audiences. Murray Newlands is its founder. This report separates OFF survey observations, external research, author analysis and illustrative scenarios. Vendor-authored guidance is identified as such and does not constitute an independent product assessment. The collection and supporting methods are available through the OFF research index: https://openfutureforum.com/research/

4,497 substantive body words | 5,527 total words
Tables, FAQs, methods and references are excluded from the body count.

In this report

01 The decision in brief
02 Executive summary
03 What November adds
04 Evidence and sample boundaries
05 1. Start with the promise the buyer can verify
06 2. Separate the billable unit from the cost-bearing work
07 3. Make acceptance a product capability
08 4. Price outcomes only when attribution is workable
09 5. Identify the sponsor, the signer and the operating owner
10 6. Make expansion a new commercial decision
11 External evidence and its limits
12 An illustrative pricing sensitivity and buyer review
13 The next customer decision
14 Founder questions
15 Citable facts
16 Methods and limitations
17 References
18 Related research and citation
19 Editable working tool

The decision in brief

Choose a pricing promise that the customer can understand and the company can deliver profitably under the actual contract. Define the billable unit, acceptance conditions, cost exposure and buyer authority before treating a pilot as evidence of scalable demand. OFF's founder data supports several commercial models; it does not identify one universally superior pricing strategy.

Executive summary

Founders face two connected tests when selling enterprise AI: whether a buyer values the proposed result and whether the supplier can deliver that result at an acceptable cost. A pilot can pass the first test while concealing the second. Intensive founder support, generous retries, manual corrections and custom integrations may make the demonstration compelling without establishing repeatable delivery economics.

OFF's carried-forward pricing question has 148 responses and permits multiple selections. Usage-based pricing was selected by 63 respondents, flat subscription by 36, outcome-based pricing by 35 and per-seat pricing by 29; 12 selected not charging yet. These are reported approaches, not mutually exclusive company categories or evidence of realized revenue. Hybrid arrangements are possible, but the distribution alone does not reveal the precise contract structures behind the selections. [PRICE]

The same-sized buyer-path question records business-unit leaders in 58 responses, CTO or CIO in 53 and finance in 41. It also permits multiple selections. Those responses describe reported entry points, not a measured sequence from sponsor to signer. Separately, the broader signing-authority instrument has 467 responses. Its CEO and CFO selections cannot be used as the next stage of a founder conversion funnel. [DOOR; SIGN]

November's contribution is a commercial design method that joins those questions. The report examines what is sold, what triggers a charge, what constitutes acceptance and which party carries the cost of retries, exceptions and integration. An illustrative pricing model shows how a plausible headline margin can disappear when acceptance rates fall and manual review rises. The example is a sensitivity calculation, not an observed startup result.

An early company needs to learn what customers want before it can settle every contract term or build extensive reporting. The answer is to narrow the promise and the evidence requirement to the next paid commitment. A founder should be able to explain one billable unit, one delivery boundary and one decision to expand. That is often more useful than a large evidence room assembled before the commercial question is clear.

What November adds

The October founder material developed an enterprise evidence-room handoff. That remains useful once a buyer knows what it is considering. This edition addresses the commercial assumptions that should be settled before the evidence room can support a meaningful decision. A complete set of documents cannot rescue an undefined unit of value or a price that shifts an unbounded delivery risk to the supplier.

The first addition is a distinction between the charging unit and the economic unit. A customer may pay for a resolved request while the supplier incurs costs for attempts, tool calls, review and exceptions. Those units can differ legitimately, but the company must understand the conversion between them. A change in customer behavior can affect the conversion even when the model's token price falls.

The second addition is an acceptance rule that connects product performance to billing. A system declaring a task complete is insufficient when the customer disputes the result or the external system never changed. The contract and product should agree on what counts, what is excluded, how a correction is handled and what record resolves a dispute. This is especially important when the founder offers outcome-based pricing.

The third addition is a buyer-path interview designed to disconfirm the apparent sale. It asks who benefits, who bears the operating burden, who can authorize spending and who can stop the deployment. A supportive business-unit sponsor may have no authority to accept a data-processing obligation or an uncapped consumption bill. Discovering that distinction early is useful commercial learning rather than a procurement failure.

No new YC batch comparison is claimed. Recent founder-related rosters do not supply new answers to the pricing and buyer-path instruments used here. November therefore retains the existing distributions and identifies its original contribution as analysis and practical tools. The report's audience label describes the community it serves; it does not make the sample representative of all YC companies or imply YC endorsement.

Evidence and sample boundaries

PRICE and DOOR each have 148 eligible unique respondents after per-question deduplication. Both are multiple-selection questions. Counts can sum above the denominator, and selecting usage-based pricing does not rule out a subscription minimum or another component. A response is a self-report from registration data, not a contract inspection, revenue audit or verified business-model classification.

The broader SIGN question has a different population and a base of 467. It is used only to remind founders that purchasing authority can sit beyond the initial sponsor. There is no linked longitudinal sales dataset connecting the founder's entry point to a later executive signature. The report does not calculate a conversion rate, average sales cycle or probability of winning through a particular function.

The source selection excludes invited-only records and byte-identical exports, retaining the latest eligible nonblank answer per respondent per question through the cutoff. Registration timestamps do not establish when an answer was edited. No causal claim is made about pricing and growth, and no small title or batch subgroup is promoted into a benchmark.

External engineering and commercial guidance informs the proposed method. The worked economics use invented round numbers that are explicitly labelled as illustrative. A company should replace them with its own delivery records and accounting treatment before using the model to make a pricing or investment decision.

Exhibit 1. Reported pricing components

Selection	Count	Base	Rounded share
Usage-based	63	148	43%
Flat subscription	36	148	24%
Outcome-based	35	148	24%
Per seat	29	148	20%
Not charging yet	12	148	8%

Exhibit note: PRICE, carried forward; multiple selections, any mention. These are components respondents selected, not mutually exclusive company types or audited revenue mixes.

1. Start with the promise the buyer can verify

A useful pricing discussion begins with the customer decision the product is intended to improve. For a support product, the buyer might value an accepted resolution, a shorter wait or additional service capacity. Those outcomes are related but not interchangeable. A founder who promises all three without defining the relevant workload may struggle to explain whether the product succeeded when one improves and another deteriorates.

The commercial promise should identify the eligible task, the expected result and the conditions under which it applies. A promise to resolve routine account questions is narrower than a promise to handle customer support. The narrower version can specify the knowledge sources, exception categories and escalation route. It gives the buyer a basis for testing performance and gives the supplier a boundary for estimating cost.

The charging unit then needs a clear relationship to that promise. Per-seat pricing can fit a product whose value grows with access and repeated use by employees. Usage pricing can fit a service with a measurable consumption unit. Outcome pricing can fit a result whose acceptance and attribution are sufficiently clear. None of these statements establishes that a particular model is best; each identifies a condition under which the model becomes easier to explain and administer.

Hybrid pricing can address different obligations within one relationship. A platform fee may support availability, integration and ongoing administration, while a variable component covers volume. A minimum commitment may compensate for reserved capacity or a substantial onboarding effort. Those arrangements should be explained as separate commercial obligations. Hiding fixed costs inside an optimistic variable price can make early growth look attractive while leaving the business exposed when usage remains low.

The buyer should be able to reproduce the invoice from an agreed record. If a billable unit depends on an internal score that the customer cannot inspect, the supplier has created a trust problem. A definition such as

accepted case can be useful only when the acceptance event, exclusions and dispute process are clear. The product's telemetry and the contract's language should describe the same event.

Founders should also test whether the unit creates the right incentive. Charging for attempts may encourage the supplier to tolerate retries, while charging only for accepted outcomes may encourage avoidance of difficult cases. A good design identifies those incentives and uses eligibility rules, service commitments or reporting to make them visible. Pricing is part of product behavior because it influences what both parties optimize.

2. Separate the billable unit from the cost-bearing work

The supplier's cost is rarely limited to the final successful model call. A task can involve retrieval, orchestration, tool use, repeated attempts, review and support. Some of those costs scale with requests, some with complexity and some with the number of customers. A single average cost per task can conceal an expensive subgroup that becomes more important as the product expands.

Build the initial model around a small number of observable drivers. For a service sold per accepted request, useful drivers include attempts per request, cost per attempt, review frequency, review time and the proportion of submitted requests accepted. The company can then see how an operating change affects contribution before assuming that a higher usage volume will improve the economics.

The denominator matters. Cost per attempted task can fall while cost per accepted task rises if acceptance deteriorates. Similarly, a reduction in model cost can be outweighed by more expensive retrieval or manual correction. The founder should preserve both the technical and commercial units so that an engineering improvement is not automatically presented as an improvement in customer economics.

Customer-specific work deserves separate treatment. An integration that can be reused across many customers may support product investment. A custom workflow maintained for one customer can be a continuing service obligation. Both may be commercially sensible, but they imply different pricing and capacity decisions. A founder should identify which work is repeatable, which is bespoke and which remains uncertain before describing the contract as scalable software revenue.

The timing of cost also matters. A prepaid annual contract may generate cash before the delivery effort occurs. That cash can fund operations, but it does not remove the obligation to serve the customer later. Conversely, a paid pilot may look unprofitable because it includes reusable development. The internal model should distinguish cash timing, delivery contribution and investment in the product, rather than selecting whichever view makes the sale look strongest.

Sensitivity analysis is more useful than a single forecast at this stage. Show what happens if review doubles, acceptance falls or the customer uses a more demanding task mix. A founder does not need a precise probability for every scenario to learn which assumption dominates the result. The pricing promise should be reconsidered when a plausible operating change makes the contract unattractive, especially if the customer has the unilateral ability to cause that change.

3. Make acceptance a product capability

An acceptance rule is often treated as legal wording added near signature. For an AI service, it should influence product design earlier. The company needs a reliable way to know whether the agreed result occurred, whether it occurred within the permitted scope and whether an exception requires correction. Without that information, billing, quality management and customer success may each maintain a different version of performance.

For an action-taking product, completion should be checked in the relevant external state. A message saying that a refund was issued is different from a verified refund in the authorized payment system. For a drafting product, the relevant result may be acceptance by a qualified reviewer under a defined standard. The test depends on the task. It should not be replaced by a generic model confidence score that the buyer has no reason to treat as commercial acceptance.

Disputed outcomes should have a clear treatment. The parties need to know whether the unit is temporarily excluded from billing, credited after review or replaced with a corrected result. The answer can vary by contract, but ambiguity creates unnecessary operational cost. A founder who cannot reconcile credits and disputes to the original task record will struggle to understand both customer satisfaction and realized revenue.

Acceptance also creates a sampling problem. If only easy tasks are evaluated, a strong result may not support the intended rollout. The pilot should include the kinds of requests the buyer expects to submit, including difficult but legitimate cases. Exclusions should be explicit. A product can be valuable within a restricted workload; it becomes misleading when the restricted result is used to promise unrestricted performance.

The company should retain enough evidence to explain the result without collecting unnecessary sensitive information. The required record may include an event identifier, task category, authorized action, acceptance state and relevant version. Raw customer content may require different access and retention arrangements. The commercial need to resolve a dispute does not justify unlimited collection or indefinite retention. Product, security and legal teams should agree the minimum useful record for the actual contract.

Finally, acceptance should support learning. A rejected result can reveal an unclear task definition, a product defect, an integration problem or a difference in customer expectations. Those causes require different responses. Treating every rejection as model failure can waste engineering effort; treating every rejection as customer misunderstanding can conceal a weak product. A structured reason code and a short review of recurring disputes can improve both delivery and the next contract.

4. Price outcomes only when attribution is workable

Outcome pricing is attractive because it appears to align supplier payment with buyer value. The difficulty is defining which part of the outcome the supplier controls. A product may help a salesperson prepare better outreach, while conversion also depends on the offer, account quality, sales execution and market conditions. A payment formula based on total revenue can create disputes if those influences are not considered.

The founder should distinguish an operational outcome from a business outcome. An accepted support resolution is closer to the product's direct activity than a change in company retention. Both may be valuable, but they require different evidence. The more distant the outcome, the more important the baseline, attribution rule and treatment of changes outside the supplier's control become.

A workable agreement can specify a narrow eligible population and a defined measurement window. It can exclude cases that the buyer changes materially after assignment or identify how mixed human and automated work is treated. These choices should be negotiated openly because they affect incentives and reported performance. A supplier that claims credit for every improvement but excludes every adverse case has not created a credible outcome model.

The customer also needs to understand its own obligations. Data quality, timely review, integration availability and employee adoption can affect delivery. If the buyer does not provide agreed inputs, the supplier may be unable to meet the promised result. That does not mean the contract should shift all risk back to the buyer. It means the operating dependencies should be explicit enough that both parties can distinguish product failure from a missing prerequisite.

A cap or minimum can make a pilot easier to understand while the parties learn the conversion between effort and value. The pilot should have an explicit purpose: estimate the relevant drivers, test acceptance or determine whether the result matters to the customer. A capped learning contract should not be described as proof that the eventual uncapped commercial model is viable. The next commitment requires its own assessment.

The QJE study of generative AI in customer support is a useful reminder that effects can vary across workers and tasks. It does not provide a universal uplift that a startup can insert into a sales proposal. A founder should use external research to sharpen the pilot question, then establish the result in the buyer's own workflow. A credible promise is one whose evidence, population and economic consequence the buyer can inspect. [E-QJE]

5. Identify the sponsor, the signer and the operating owner

The person who first responds enthusiastically may be the best source of insight into the problem. That does not establish their authority to purchase the product or accept the responsibilities it creates. The founder should understand the sponsor's desired outcome and then map the other decisions required for a paid deployment. This is a conversation about the actual purchase, not an exercise in collecting a long list of stakeholders.

OFF's buyer-path distribution suggests several plausible entry points. Business-unit leaders, technology leaders and finance all appear in the responses. Because the question permits multiple selections, the data does not identify one exclusive route or prove that a route performs better. The practical implication is to ask which function owns the problem in this customer and which function can authorize the proposed commitment. [DOOR]

The operating owner is especially important. A buyer can approve spending while another team absorbs review work, integration maintenance or exception handling. If the operating team has no capacity or

incentive to perform that work, the pilot can succeed only while the founder supplies extraordinary support. A paid contract may then create dissatisfaction even though the original sponsor remains enthusiastic.

Useful disconfirmation questions are concrete. Who can approve the first invoice? Who must accept a variable bill above the pilot cap? Who will maintain the integration after launch? Who decides whether a result is accepted? Which function can stop the workflow? A clear answer identifies a real path to deployment. An unresolved answer identifies the next commercial risk to address.

The founder should avoid treating every objection as friction to be overcome. A finance question about variable cost may reveal an unbounded contract. A security question about downstream credentials may reveal a product limitation. An operating team's concern about review time may reveal that the apparent benefit has been transferred rather than created. Those objections can improve the offering if they are addressed before the company commits to broad delivery.

The broader OFF signing-authority data provides context, not a conversion model. Its respondents are different from the founder sample, and its answers are not the next observed step in the same deals. The founder should therefore use it to prepare questions about authority, not to forecast a percentage of sponsor conversations that will become CEO or CFO approvals. Actual sales forecasting requires records from the company's own opportunities and defined stages. [SIGN]

Exhibit 2. Reported buyer entry points

Selection	Count	Base	Rounded share
Business-unit leader	58	148	39%
CTO or CIO	53	148	36%
Finance	41	148	28%
Other	18	148	12%
Bottom-up adoption	16	148	11%
Still figuring it out	6	148	4%

Exhibit note: DOOR, carried forward; multiple selections. These are reported entry points, not verified stages or conversion rates.

6. Make expansion a new commercial decision

A successful pilot should answer a defined question. It should not automatically authorize every larger use case. Expansion can change task difficulty, review requirements, integration cost, customer concentration and the authority granted to the product. The founder should identify what changes between the pilot and the proposed contract, then determine which pricing and delivery assumptions need to be tested again.

Volume alone can alter the economics. A small pilot may be handled by a founder who understands every exception. At higher volume, the company needs repeatable support, documentation and escalation. The

cost of that transition may be justified, but it belongs in the expansion decision. Treating founder effort as free can make the first contract appear profitable while obscuring the resources needed to serve the next ten customers.

Task mix can be more important than volume. A buyer may begin with routine requests and later submit complex cases because the service appears convenient. If the commercial unit remains unchanged, the supplier can experience higher cost without higher revenue. The contract should identify the relevant task classes or provide a process for revisiting scope when the workload changes materially. A price increase is easier to explain when it follows a visible change in obligations.

Expansion also creates a negotiation about learning. The supplier may have developed reusable capabilities through the pilot, while the customer may have contributed substantial domain knowledge and review effort. The parties should clarify the rights and obligations associated with that work. A founder should not assume that access to customer data or feedback implies unrestricted rights to reuse it for every purpose. The relevant terms require review in the actual relationship.

Renewal should be based on realized performance under the agreed definition. The company should reconcile billed units, accepted outcomes, credits, support effort and the customer's continuing business need. A high usage count can coexist with dissatisfaction if the customer experiences surprise bills or recurring correction work. A lower usage count can still support a valuable contract if the service addresses a consequential problem with reliable delivery.

The commercial record should preserve what was learned. Which assumptions held, which failed and which changed because the customer expanded the scope? That information improves pricing for subsequent customers and prevents the sales team from repeating a promise that operations cannot sustain. The goal is a repeatable relationship between value, delivery and payment, supported by evidence that becomes more precise as the business grows.

External evidence and its limits

Stripe's current documentation describes usage-based billing infrastructure and distinguishes its Metronome platform from the lower-level Billing Meters implementation. This is relevant to recording and charging for consumption. It does not establish that a particular unit represents customer value, that a contract is profitable or that invoiced amounts meet the applicable requirements for revenue recognition. Those are separate commercial and accounting questions. [E-STRIPE]

FinOps Foundation guidance distinguishes technical and business measures in unit economics. That distinction informs the model below: attempts and review effort drive cost, while accepted requests drive the illustrative charge. The proposed calculations are the author's application, not a FinOps benchmark. A startup should select units that match its product and customer agreement rather than adopting a framework's terminology without the necessary operational records. [E-FINOPS]

Anthropic's engineering guidance on agent evaluations distinguishes task outcomes from the agent's own account of what happened. This supports checking the relevant external result when a product takes action. It is vendor-authored engineering guidance, not independent evidence that one vendor's agents are reliable

or that the proposed acceptance process will work in every setting. The customer contract and operating context still determine the relevant result. [E-ANTHROPIC-EVAL]

The published QJE study, Generative AI at Work, examines a specific customer-support setting and finds heterogeneous effects. It provides evidence that assistance can improve performance in that setting, while also showing why a single average should not be treated as a universal sales claim. The report does not apply its estimated effect to OFF respondents or use it to forecast a startup's revenue. [E-QJE]

These sources address different questions: how usage can be billed, how costs can be related to a business unit, how an action outcome can be evaluated and how assistance affected one studied workflow. None chooses the founder's price. That decision requires evidence about willingness to pay, delivery obligations, competitive alternatives and the company's own costs. External research is most useful when it helps specify what must be learned from the next customer.

Pricing should remain flexible while the product develops, with experimental terms made explicit. A time-limited pilot price, a defined cap or a narrow eligibility rule can support learning without pretending to be a permanent model. Problems arise when temporary concessions become hidden assumptions in the forecast. A founder should know which terms are experimental and what evidence will justify changing them before the next commitment.

A founder can turn these distinctions into a customer experiment before committing to an enduring price. Define an eligible request and the observable acceptance event, then record attempts, review minutes, credits and payment separately. Ask the customer to confirm that the acceptance definition represents useful work. Compare the result with the agreed pilot terms rather than a broad research average. If the customer changes the eligible workload, treat the next period as a new commercial test. This gives the pricing discussion an explicit source of evidence while preserving room to learn; it does not require a small company to claim that a short pilot proves retention or lifetime margin.

Evidence comparison

Source	Contribution to this decision	Limit on interpretation
Stripe [E-STRIPE]	Records and charges for usage	Billing infrastructure does not establish recognition or value
FinOps [E-FINOPS]	Relates technical costs to business units	No prescribed price or benchmark contribution margin
Anthropic [E-ANTHROPIC-EVAL]	Checks external task outcomes	Vendor guidance; acceptance must fit the customer contract
QJE study [E-QJE]	Shows effects in a studied support workflow	Do not import the average into a startup revenue forecast

An illustrative pricing sensitivity and buyer review

Consider a hypothetical service charging $2 for each accepted request. The customer submits 10,000 eligible requests per month. In the initial scenario, 80% are accepted, producing 8,000 billable units and $16,000 of monthly charges. Each submitted request requires an average of 1.5 attempts at $0.04 per attempt. Attempt cost is therefore $600. These figures are invented for the calculation; they are not OFF observations or a market price recommendation.

Assume 20% of submitted requests require three minutes of human review at a fully loaded rate of $40 per hour. Review cost is 2,000 multiplied by three minutes, divided by 60, multiplied by $40: $4,000. Add $2,000 of allocated monthly service and support cost. Total modeled delivery cost is $6,600, leaving $9,400, or 58.75% of charges, before omitted expenses such as sales, general administration, financing and taxes. This is modeled delivery contribution, not accounting gross margin or company profit.

Now change only the operating assumptions. Acceptance falls to 60%, attempts rise to 2.5 per submitted request and review rises to 40% of submissions. Charges fall to $12,000. Attempt cost rises to $1,000 and review cost to $8,000. With the same $2,000 service allocation, modeled delivery cost becomes $11,000 and contribution falls to $1,000, or 8.33%. A lower model price would not necessarily compensate for that deterioration in acceptance and review effort.

Exhibit 3. Illustrative economics per month

Driver or result	Initial scenario	More demanding workload
Submitted requests	10,000	10,000
Accepted share and billable units	80%; 8,000	60%; 6,000
Charge per accepted unit	$2	$2
Monthly charges	$16,000	$12,000
Attempts per submission; attempt cost	1.5; $600	2.5; $1,000
Review share; review cost	20%; $4,000	40%; $8,000
Allocated service and support	$2,000	$2,000
Modeled delivery contribution	$9,400; 58.75%	$1,000; 8.33%

Exhibit note: Hypothetical calculation. Three minutes of review at $40/hour; $0.04 per attempt. Acceptance and review are separate assumptions, with no claim that reviewed requests always succeed. Contribution excludes other company expenses and is not an audited margin.

The point is not that outcome pricing is unattractive. It is that the founder must understand which operating changes affect payment and cost. The initial contract might remain sensible with a defined eligible workload, a review limit or a different fee structure. The company could also improve the product or decline a customer

whose requirements are uneconomic. Each response should be based on the actual driver rather than an assumption that more usage automatically creates scale benefits.

Before expanding, the founder and buyer should review a compact commercial record. It should define the billable event, the acceptance evidence, the treatment of retries and credits, the eligible task population, the operating owner and the spending authority. The founder should also record which cost assumptions have been measured and which remain estimates. This makes the next negotiation a discussion of known obligations rather than a contest between optimistic forecasts.

The buyer interview should end with a decision. Is the next step a paid learning period, a restricted production contract or a broader commitment? What would disconfirm the expected value? Who can approve the change? If those answers remain unclear, the founder has identified work required before treating the opportunity as repeatable revenue. That is commercially useful even when it delays signature, because it reduces the risk of winning a contract that the company cannot deliver well.

The next customer decision

Select one active enterprise opportunity and write the commercial promise in a sentence that a buyer can verify. Then specify the unit that triggers payment, the work that drives cost and the conditions under which a result is accepted. If those definitions differ across the sales proposal, product telemetry and contract, resolve the discrepancy before using the opportunity as evidence of scalable demand.

Use the pilot to learn the assumptions that matter most. A founder may need evidence about review effort, the buyer's operating capacity or willingness to pay more than another demonstration of model capability. The next experiment should answer the question that constrains the commercial decision. A strong product demonstration is useful, but it is only one part of a viable enterprise relationship.

OFF's distributions show a range of pricing components and buyer entry points. They do not select a winning formula. November's recommendation is to make the relationship between value, obligation, cost and authority explicit enough that both parties can decide whether to expand. That discipline gives the founder a clearer basis for pricing, a more credible conversation with the buyer and better evidence about what the business can repeat.

Founder questions

Is usage pricing the best model because it has the largest selection?

No. The 63/148 result measures reported use, not comparative performance. Multiple components can be selected. [PRICE]

Can the buyer-path and signer questions form a sales funnel?

No. They have different populations and are not linked stages in observed deals. [DOOR; SIGN]

Is the worked contribution an accounting gross margin?

No. It is an illustrative delivery calculation with stated omissions. Accounting classification requires the company's actual policies and facts.

Does outcome pricing require a distant business outcome?

No. A clearly accepted operational result may be easier to attribute and administer than a company-wide revenue change.

Is an agent's completion message sufficient for billing?

Only if it is the agreed and valid evidence of the billable event. For external actions, verify the relevant result rather than relying solely on the message. [E-ANTHROPIC-EVAL; recommendation]

Does this edition contain a new YC cohort or batch analysis?

No. The relevant measures are carried forward, and the report makes no YC batch comparison or endorsement claim.

Citable facts

- Usage-based pricing was selected by 63/148 respondents, 43%; multiple selections permitted. [PRICE]

- Business-unit entry points were selected by 58/148, 39%; CTO/CIO by 53/148, 36%; finance by 41/148, 28%. [DOOR]

- These measures do not establish sales conversion, realized revenue or comparative pricing success.

Methods and limitations

Calculation-v1.1 uses authorized 1 November exports through the stated cutoff. Byte-identical exports and invited-only records are excluded. The latest eligible nonblank answer is selected per normalized respondent identifier per question. The pricing and buyer-path distributions reconcile to the published October baseline and are carried forward.

The sample is selective, self-reported and not representative of all startups or YC companies. Multiple selections are reported as any mention. No contract, revenue, profitability, batch or longitudinal sales analysis is claimed. The pricing model and buyer review are original illustrations and recommendations.

References

[OFF-DATA] Open Future Forum. Executive AI Statistics, November 2026. Question IDs PRICE, DOOR and SIGN.

[E-STRIPE] Stripe. Basic usage-based billing, current documentation consulted 2 November 2026. https://docs.stripe.com/billing/subscriptions/usage-based

[E-FINOPS] FinOps Foundation. Unit Economics capability. https://www.finops.org/framework/capabilities/unit-economics/

[E-ANTHROPIC-EVAL] Anthropic Engineering. Demystifying evals for AI agents. 9 January 2026. https://www.anthropic.com/research/demystifying-evals-for-ai-agents

[E-QJE] Brynjolfsson, E., Li, D. and Raymond, L. Generative AI at Work. Quarterly Journal of Economics, 140(2), 889-942, 2025. https://academic.oup.com/qje/article/140/2/889/7990658

[OFF-PRIOR] Open Future Forum. AI Startup Enterprise Evidence Room: A 30-Day Handoff, October 2026. https://openfutureforum.com/blog/ai-startup-enterprise-evidence-room-30-day-handoff

Related research and citation

The November CFO report qualifies buyer benefits, AI Leaders examines production expansion, VC and CVC tests investment theses, and Investment Banking distinguishes usage, contracts, revenue and cash.

Suggested citation: Newlands, Murray. YC Founder AI Report, November 2026, Edition 4. Open Future Forum. Evidence cutoff 31 October 2026.

Open Future Forum research: https://openfutureforum.com/research/

Editable working tool

Acceptance-based pricing experiment: download the editable CSV worksheet. The worksheet implements the report's recommendation; it is not a validated score, mandatory standard or completed assessment.