Which Go-to-Market Motions Produce a First Paying Buyer? An Evidence Review and a Weekly Operating System for Early Products

Jacob Krajewski
In Motion Automation, Tampa Bay, Florida

26 October 2026

Abstract

Most new products stall before their first paying customer, and the founder has to choose among many ways of reaching buyers: sales calls, cold outreach, founder posts, answers from AI assistants, directories, referrals, free product tiers and paid ads. Public benchmarks for these choices are scattered, define their rates differently and often come from vendors. We reviewed primary sources for each of eight motions, including buyer surveys, platform rules, statutes, benchmark reports and pricing pages, and labelled every number by confidence. Business buyers form a shortlist before they speak to a seller: in one survey of 2,509 buyers, 81% had a preferred vendor before first contact, and in another of 1,169, AI chatbots were the most often chosen outside influence on the shortlist. Average cold email reply rates published for 2025 range from 0.45% to 3.43% depending on what is counted, so a team has to rely on its own counts. Google states that its AI answers need no special files or markup, and three independent analyses found no citation gain from an llms.txt file. No public source gives a comparable conversion rate for founder calls, founder posts or referrals. From this evidence we built an operating system that works for any product: five stages with exit tests counted in buyers, a buyer list in which every row links proof that the person already pays for the job, a five-day weekly cadence with one number, kill, hold and double-down rules, a compliance checklist for US, EU and UK outreach, and tool stacks at three budgets.

1. Introduction

A new product earns its first money from a stranger long after its first line of code. In between, the founder decides each week where to spend a limited number of hours: calling people, writing to them, posting, getting listed, building a free tier, or buying ads. Each of these is a *motion*, one repeatable way of reaching buyers. The work of choosing and running motions is go-to-market (GTM) work. In this paper a *paying buyer* is a person or firm outside the founder's own circle who has paid money, and every rule we propose is counted in paying buyers.

Two changes make the choice harder in 2026. First, business buyers do much of their comparing before they contact any vendor, and they increasingly ask AI assistants which vendors to consider. The few vendors a buyer compares are the *shortlist*. An AI answer can *mention* a firm by name or *cite* it by linking one of its pages. Second, mailbox providers and social platforms have tightened written rules on bulk and automated outreach, so the cheapest way to send a thousand messages is also the riskiest.

An earlier audit of a widely viewed twelve-move B2B playbook, published on X by the founder of a GTM agency, confirmed most of the playbook's cited facts, reproduced its claim that a self-published "best agencies" list was named first in Google's AI answers (but not in ChatGPT), and found the most written legal and platform exposure in high-volume cold email and in scraping LinkedIn engagers (Krajewski, 2026; Vacca, 2026). Two gaps remained. The playbook assumed a firm that already sells, with an account list, a CRM and ad budget, and it gave no rule for when to stop a channel. A founder with a product and no buyers needs a different starting point.

We asked two questions. What does public evidence say about how well each motion produces early buyers, and how strong is that evidence? What weekly operating system follows from it, one that works for any product and any founder with ten or more hours a week?

Evidence is strongest where buyers are surveyed and where platforms publish rules, and weakest where founders spend most of their early time. Contributions:

1. An evidence table by motion with ranges, sources and confidence labels (Section 4, Table 2).
2. A product-agnostic operating system with stages, a buyer-list method, a weekly cadence, decision rules, a compliance checklist and tool stacks at three budgets (Section 5).

2. Terms

Table 1 defines the terms used in the rest of the paper.

Table 1: Terms used in this paper.

Term	Meaning here
Ideal customer profile (ICP)	The kind of buyer who pays and stays, written as one sentence naming role, firm type and job.
Proof row	One row of the buyer list with a public URL showing that this person or firm already pays, in money or staff time, for the job the product does.
Will-pay statement	A buyer's own words naming a price or terms they would pay, recorded verbatim in a call.
Positive reply	A reply that asks for a call, a price, a trial or more detail. Auto-replies and refusals are not positive.
Reply rate	Replies divided by a denominator. Sources use emails sent, recipients or contacts; Section 4.2 shows why this changes the number.
Free-to-paid conversion	Share of new free accounts or trials that become paying within a stated window.
Stage and exit test	One of five stages of Figure 4. A stage ends when its exit test, a count, is met.
Weekly number	Single count reviewed every Friday: unique paying buyers, later paying buyers per week by channel.
Leading indicators	Counts that move before the weekly number: proof rows, conversations, will-pay statements, positive replies and AI-answer mentions.

3. Methods

3.1. Sources and reading

We read every source between 00:00 and 01:00 Eastern time on 26 October 2026, except those reused from the earlier audit of the playbook, read on 25 October 2026. We preferred, in order: statutes and regulator guidance; platform documentation and terms; independent research organizations and government statistics; surveys with a stated sample; vendor benchmark reports with a stated sample and definition; and vendor or

agency pricing pages. Where a primary page did not load, we used an archived copy or a secondary report and labelled the claim partly confirmed. Appendix A lists each claim with its URL.

3.2. Confidence labels

Each evidence row carries one label. *High*: official or independent source with a stated method and sample, or a written rule. *Medium*: survey or vendor dataset with a stated sample and definition, where the publisher may benefit from the result. *Low*: a single firm's data, an anecdote, or no public benchmark. Each verified claim also carries C (confirmed on the primary page), P (partly confirmed, for example only on a secondary page or with a missing term) or U (unconfirmed).

3.3. Design of the operating system

We derived the stages, cadence and rules from the evidence and from the earlier audit. Thresholds in the system, such as thirty proof rows or ten calls, are design defaults chosen to fit a founder working ten to twenty hours a week. They are not benchmarks, and each one can be changed to fit a different product or schedule.

3.4. Tool prices

Tool prices are list prices for month-to-month billing on the vendors' own pages.

4. Results: evidence by motion

Table 2 summarizes the evidence. The subsections below explain the rows that shape the operating system.

Table 2: Evidence by motion. Ranges are as published; confidence follows Section 3. Sources in Appendix A.

Motion	Best available numbers	Confidence	Weakness
Founder calls and discovery	No public conversion rate. Bridge Group SDRs hold a median 4.1 quality conversations a day in funded teams; Superhuman used 40% "very disappointed" as a fit threshold	Low	Team and survey context, not founder conversion
Outbound email	Average reply rate 0.45% (unique replies per email, Belkins 2025) to 3.43% (all replies per email, Instantly 2025); top 10% of Instantly senders 10.7%+; 58% of replies come from step one	Medium	Vendor data; different denominators; no link to revenue
Founder posts and content	No public post-to-buyer rate	Low	Anecdotes only
AI-answer visibility	AI chatbots were 17.1% of shortlist-influence responses, the top source (G2, 1,169 buyers); with an AI summary, users clicked a result in 8% of visits vs 15% without (Pew)	Medium	Survey; correlational studies on what earns mentions
Directories and review sites	Review sites 15.1% of shortlist-influence responses (G2)	Medium	No listing-to-paid rate
Partnerships and referrals	Peers 8.9% of shortlist-influence responses (G2)	Low	No referral conversion rate
Product-led loops	Free-to-paid: freemium 3 to 5% good, 8 to 12% great; trial with card 25 to 35% good, 50 to 60% great (200 products)	Medium	Needs sign-ups first; vendor and consultant data
Paid ads	Google search median conversion rate 7.52% and cost per lead $70.11 (16,446 US campaigns)	Medium	All industries; leads, not buyers

4.1. Buyers pick a shortlist before they talk to a seller, and AI chatbots now shape it

In 6sense's 2024 survey of 2,509 recent B2B buyers, 81% had chosen a preferred vendor before speaking with sales, and about 69% of the buying process happened before buyers engaged a seller (6sense, 2024). G2's 2025 survey of 1,169 B2B software decision makers asked which sources influence the vendors on a shortlist. Generative AI chatbots were the most common answer at 17.1% of responses, ahead of review sites (15.1%), vendor websites (12.8%), market research (10.6%), independent content (9.7%), peers (8.9%) and vendor salespeople (8.8%) (G2, 2025). Figure 1 plots these shares. The LinkedIn B2B Institute, working with the Ehrenberg-Bass Institute, estimates that about 95% of a firm's potential business buyers are not in the market at any given time (LinkedIn, n.d.a).

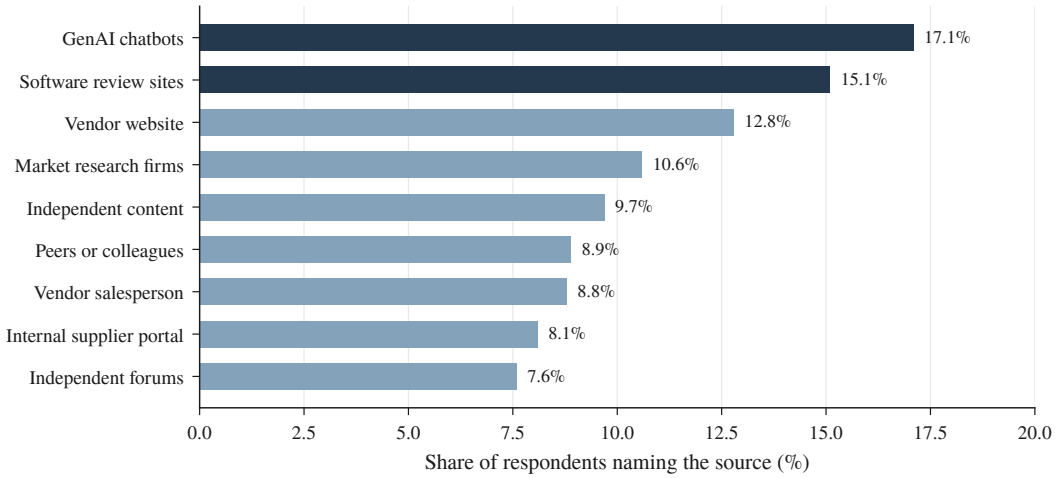


Figure 1: Sources that influence which vendors reach a B2B software shortlist, share of responses (G2, 2025). Survey of 1,169 decision makers; vendor-published.

Both surveys are run by firms that sell to marketers, so we rate them Medium. Taken together they mean a founder's direct outreach usually reaches a buyer who has already formed a view, and that a buyer who is not in the market will not reply however good the message is. The system therefore keeps two kinds of work running in parallel from the start: direct conversations with people who pay for the job today, and a small amount of work that helps the product appear where buyers form shortlists.

4.2. Published cold email reply rates differ by a factor of seven because they count different things

Instantly's 2026 benchmark, built from its platform's sends between 1 January and 18 December 2025, reports an average reply rate of 3.43%, a top quartile at 5.5% or more and a top decile at 10.7% or more; 58% of replies came from the first email in a sequence (Instantly, 2026). Its rate counts all replies, including replies to follow-ups, divided by emails sent. Belkins analyzed 7,530,489 emails sent in 2025 with 34,393 replies and reports 0.45%, counting unique replies with auto-replies removed (Belkins, 2026). Figure 2 shows both.

Neither rate counts paying buyers, and both come from senders who use bulk tooling. Gmail's sender guidelines ask senders to keep the user-reported spam rate below 0.3%, to authenticate with SPF and DKIM aligned for DMARC, and to offer one-click unsubscribe on marketing messages (Google, 2024). Google's acceptable use policy for Workspace prohibits unsolicited mass email (Google, 2026a). For an early product, outbound in this system means individually written messages to proof rows, at about twenty a day per sender,

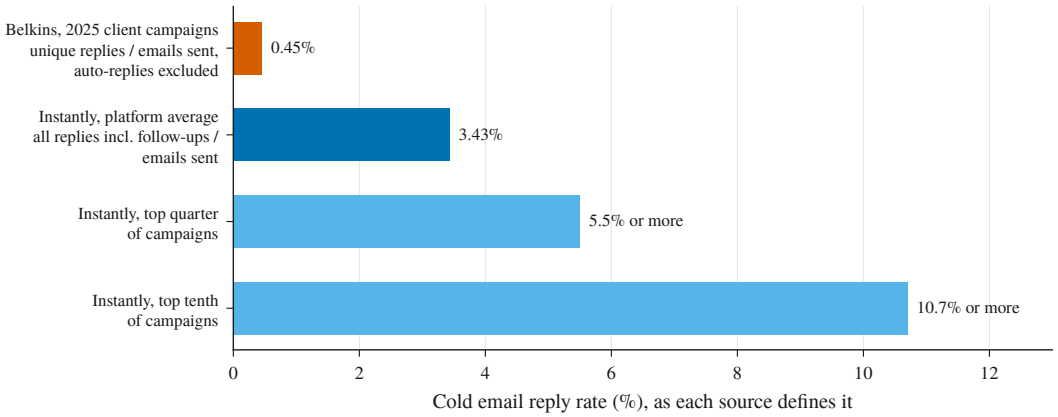


Figure 2: Published 2025 cold email reply rates under two definitions (Belkins, 2026; Instantly, 2026).

from the founder's own mailbox. At that volume the team's own reply count after fifty contacts tells it more than either benchmark.

4.3. Google says its AI answers need no special files, and three studies found no llms.txt effect

Google's guide to AI features in Search states that site owners do not need llms.txt files or special markup for these features, that seeking inauthentic mentions is not helpful, and that structured data is not required but can help rich results (Google, 2026b). SE Ranking's analysis of about 300,000 domains found no measurable relationship between having an llms.txt file and AI citations, with 10.13% adoption (SE Ranking, 2025). Geoa compared 1,912 Nordic companies and found 43.6% of sites with the file cited against 43.3% without ($p = 0.94$) (Geoa, 2026). Trakkr's comparison of 37,894 domains found the same median citation count with and without the file (Trakkr, 2026).

What does correlate with AI mentions is being talked about elsewhere. Ahrefs found brand web mentions correlated with visibility in ChatGPT, AI Mode and AI Overviews at 0.664, 0.709 and 0.656, against 0.218 for backlinks, and found that "best X" lists made up 43.8% of page types cited by ChatGPT in its sample (Allsopp, 2025; Linehan, 2025). These are correlations from a tool vendor. Pew Research found that Google users who saw an AI summary clicked a traditional result in 8% of visits compared with 15% without one, and clicked a link inside the summary in 1% of visits (Chapekis and Lieb, 2025). Visits from AI answers are also undercounted: one agency found that 189 leads said an AI assistant sent them while its CRM credited 28 to AI (Omniscient Digital, 2026).

For the system, AI-answer work means honest pages that answer buyer questions, listings in places buyers and AI answers read, a monthly check of whether answers name the product, and a "How did you hear about us?" field on every form and call. Google's spam policies on scaled content abuse and site reputation abuse rule out mass-produced pages (Google, 2026c).

4.4. Product trials that ask for a card convert far better, after people sign up

Kyle Poyar's 2026 analysis with ChartMogul and ProductLed of about 200 products gives good and great free-to-paid ranges by model (Figure 3): freemium 3 to 5% good and 8 to 12% great; free trial without a card 4 to 6% and 10 to 15%; free trial with a card 25 to 35% and 50 to 60% (ChartMogul, 2026; Poyar, 2026). ChartMogul's report gives a median of about 8% across models and about 30% for card-required

trials. These rates start from a sign-up, so they only help once the product has traffic. For a product with no buyers, a free tier is a loop to add once paying buyers exist and can invite others.

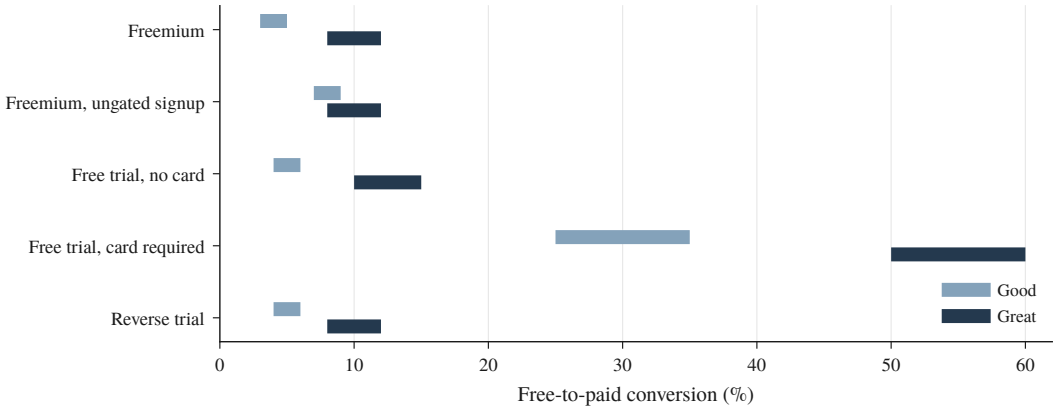


Figure 3: Good and great free-to-paid conversion ranges by model (Poyar, 2026).

4.5. Paid search leads cost about $70 each at the median, and LinkedIn ads need a list and a daily floor

WordStream's 2025 analysis of 16,446 US Google search campaigns reports a median conversion rate of 7.52% and a median cost per lead of $70.11 across industries (WordStream, 2025). A lead here is a form fill; the report does not track purchases. LinkedIn requires a daily budget of at least $10 and a matched company list of at least 300 companies (LinkedIn, n.d.b,n). Paid ads therefore fit a product that already knows which message converts, which is Stage 4 below.

4.6. Founder calls, posts, referrals and directories have no comparable public benchmark

We found no primary source with a denominator for founder calls to paying buyers, founder posts to buyers, referral asks to buyers, or directory listings to buyers. What exists is practitioner guidance. Paul Graham's essay advises founders to recruit users manually (Graham, 2013). Superhuman's founder used the share of users who would be "very disappointed" without the product, with 40% as the threshold, to judge fit (First Round Review, n.d.). The Bridge Group's 2025 survey of 351 companies reports a median of 4.1 quality conversations a day for a full-time sales development rep (Bridge Group, 2025), which gives an upper bound on what one person can hold, not a conversion rate.

Because these motions lack benchmarks, the system measures them by counts that the founder controls (calls held, will-pay statements, referral asks made) and judges them against the weekly number.

5. An operating system built on this evidence

5.1. Five stages, each ended by a count of buyers

Figure 4 shows the stages and Table 3 states their rules. A stage adds motions only after the previous exit test is met, so early hours go to the motions that produce conversations, while one post and one AI-answer item a week start the slower motions.

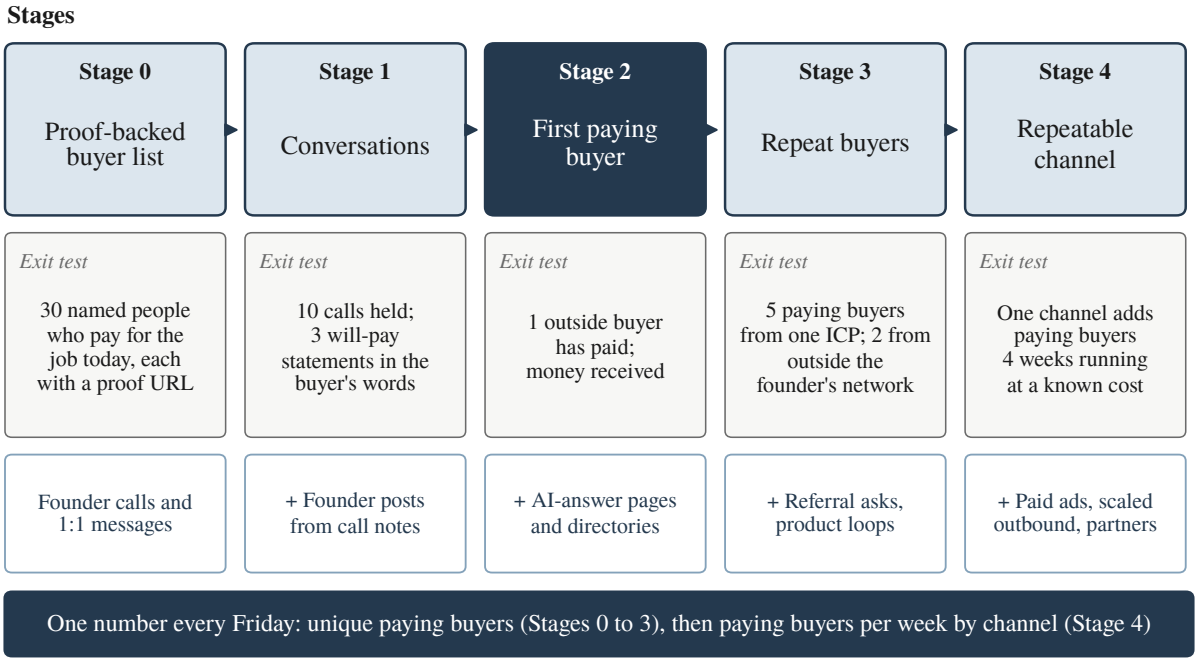


Figure 4: Five stages from no buyers to a repeatable channel, with exit tests and motions added at each stage. Thresholds are editable design defaults.

Table 3: Stage rules. Thresholds are design defaults.

Stage	Name	Exit test	Motions in use
0	Proof-backed buyer list	30 named people who pay for the job today, each with a proof URL; one-sentence ICP	Founder research; 1:1 messages to the first 10
1	Conversations	10 calls held; 3 will-pay statements in the buyer's words	Founder calls, 1:1 outreach, founder posts from call notes
2	First paying buyer	1 outside buyer has paid; money received	Adds AI-answer pages and directory listings
3	Repeat buyers	5 paying buyers from one ICP, 2 of them outside the founder's network	Adds referral asks and product loops
4	Repeatable channel	One channel adds paying buyers 4 weeks running at a known cost per buyer	Adds paid ads, scaled compliant outbound, partners

5.2. Buyer list with a proof requirement

A row enters the list only with a proof URL: a job post for the role the product replaces, a public case-study quote naming a competing tool, a pricing page of a tool the firm uses, a forum post asking for the job to be done, or a marketplace listing that charges for the same job. Each row records name, role, firm, country (for the compliance column), channel, proof URL and proof type, date found, status, last touch, next step, will-pay statement and source of first contact. The list is built by hand from public pages; scraping LinkedIn and bought lists are excluded (Section 5.5). The proof requirement keeps out people who fit a description but have never paid for anything like the product.

5.3. Weekly cadence and one number

Figure 5 shows the week. Monday adds twenty proof rows and checks each row's region and opt-out status. Tuesday and Wednesday send individually written messages and one follow-up, matching Instantly's finding that most replies come from the first message and that Tuesday and Wednesday lead (Instantly, 2026). Thursday holds calls and logs will-pay statements verbatim. Friday is the weekly review. Two assets come from call notes each week: one founder post and one AI-answer item (an answer page, a comparison, or a listing).

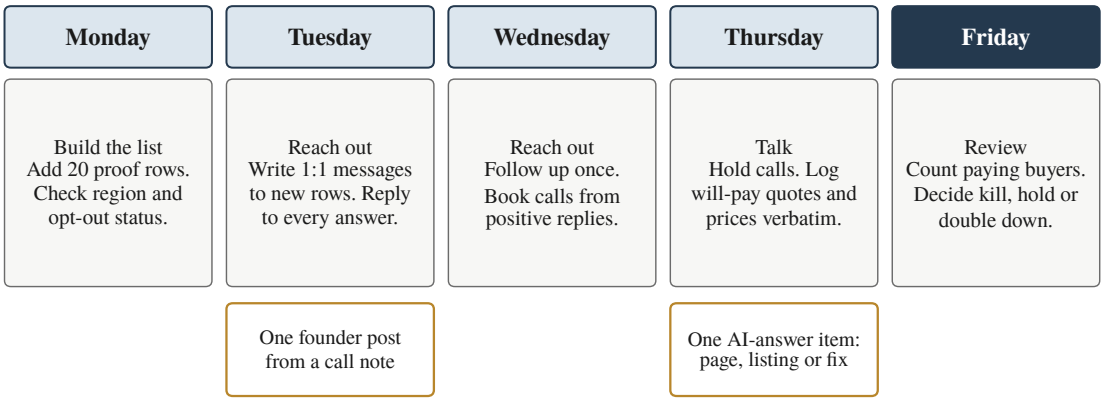


Figure 5: Weekly cadence for a founder with ten to twenty hours a week.

The weekly number is unique paying buyers until Stage 3 ends, then paying buyers per week by channel. Leading indicators are proof rows added, conversations held, will-pay statements, positive replies, and AI-answer mentions across a fixed set of twenty buyer prompts checked monthly.

5.4. Decision rules

Each Friday, every running channel and every ICP gets one of three decisions (Table 4). The rules use minimum samples so that a channel is not judged on a handful of messages.

Table 4: Kill, hold and double-down rules (design defaults).

Decision	Rule
Hold	Fewer than 50 proof-backed contacts reached on the channel, or fewer than 14 days since it started. Keep running unchanged.
Kill a message	50 contacts reached, zero positive replies. Rewrite once from call notes; a second zero kills the message.
Kill a channel	100 contacts reached across two messages, zero positive replies, for this ICP.
Kill an ICP	10 calls held, zero will-pay statements. Rewrite the ICP from the calls that went furthest.
Kill a product path	30 days in Stage 1 with zero will-pay statements, or 30 days after a paid offer goes live with zero outside paying buyers.
Double down	A channel produced 2 or more paying buyers in 30 days at a cost per buyer below one month of the price. Move half of next week's hours to it.

5.5. Compliance checklist

The checklist applies to every message and every list (Table 5). In the United States, CAN-SPAM covers business-to-business email with no exception; each commercial email must identify itself as an advertise-

ment, carry a valid postal address and a working opt-out honored within ten business days, and use accurate headers and subject lines, with penalties up to $53,088 per violating email (Federal Trade Commission, n.d.). In the EU, a person whose data was not collected from them must receive the Article 14 notice, at the latest at first communication when the data is used to contact them (European Union, 2016a); Article 21 gives a right to object to direct marketing that must be brought to their attention at first communication, and processing for direct marketing stops on objection (European Union, 2016b); Recital 47 allows that direct marketing may be a legitimate interest (European Union, 2016c). In the UK, the ICO allows marketing email to corporate subscribers without consent, but sole traders and some partnerships need consent (Information Commissioner's Office, n.d.). LinkedIn's User Agreement prohibits bots, automated messaging and scraping (LinkedIn, 2025), and hiQ's dispute with LinkedIn ended in a consent judgment (hiQ Labs v. LinkedIn, 2022). X's automation rules prohibit automated unsolicited replies, mentions and direct messages (X, 2026). The FTC's rule on consumer reviews and testimonials prohibits fake, bought and undisclosed insider reviews (Federal Trade Commission, 2024).

Table 5: Compliance checklist applied to every list and message.

Area	Check	Source
US email	Ad disclosure; postal address; opt-out honored within 10 business days; true headers and subject	CAN-SPAM
EU contacts	Article 14 notice at first message; objection offered at first message; legitimate-interest note on file	GDPR Art. 14, 21; Rec. 47
UK contacts	Corporate address only without consent; sole traders need consent	ICO, PECR
Gmail and Workspace	SPF, DKIM, DMARC aligned; spam rate under 0.3%; no mass unsolicited mail from Workspace	Gmail, Google AUP
LinkedIn	Manual messages only; no automation, no scraping, no bought engager lists	User Agreement 8.2
X	No automated unsolicited replies, mentions or DMs	X automation rules
Search and AI answers	No mass-produced pages; no paid or fake mentions	Google spam policies
Reviews	No paid, fake or undisclosed insider reviews	16 CFR Part 465

5.6. Tool stacks at three budgets

Figure 6 and Table 6 give three stacks at list price for month-to-month billing, read on 26 October 2026. The $0 stack covers Stages 0 to 2. The under-$200 stack adds a sending tool, an AI-answer tracker and call notes. The funded stack fits Stage 4.

5.7. Templates

Appendix B gives three templates for running the system: a call script outline, a thirty-day plan and a one-page experiment brief.

6. Discussion

6.1. Why the system starts with conversations

The surveys in Section 4.1 say buyers arrive with a preference formed elsewhere, and the motions that shape that preference (AI answers, review sites, peers) take months to show results. The motions with the best

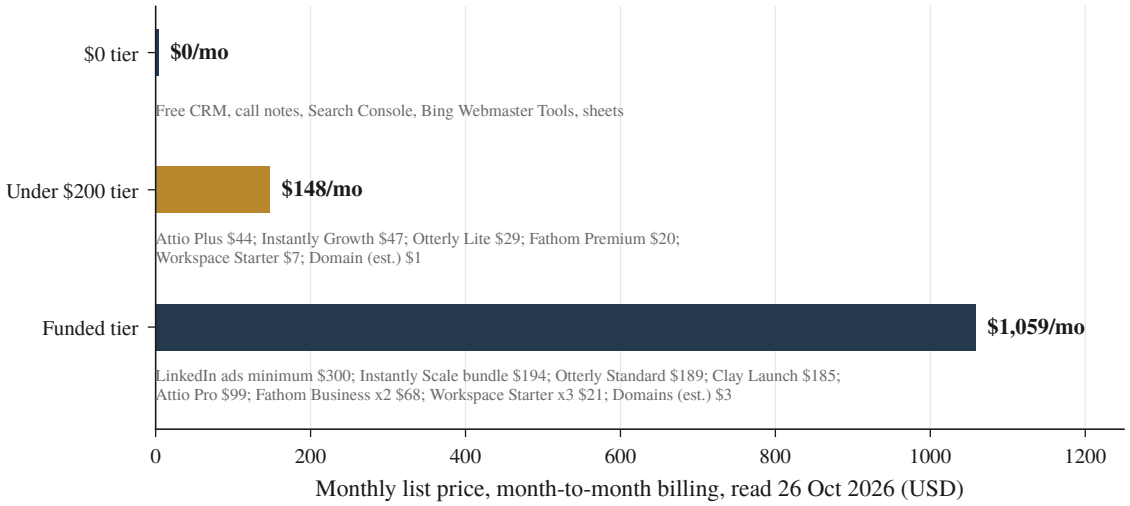


Figure 6: Monthly list price of three tool stacks, month-to-month billing, read 26 October 2026.

Table 6: Tool list prices used in the stacks (monthly billing, USD). Stack: 0 = $0, 200 = under $200, F = funded.

Tool	Plan and price	Stack
Attio (CRM)	Free up to 3 seats; Plus $44; Pro $99 per user	0, 200, F
Fathom (call notes)	Free; Premium $20; Business $34 per user, 2-seat minimum	0, 200, F
Instantly (sending)	Growth $47; Hypergrowth $97; Scale bundle $194	200, F
Otterly (AI-answer tracking)	Lite $29 (15 prompts); Standard $189; Premium $489	200, F
Google Workspace	Business Starter $7 list per user; Standard $14	200, F
Clay (research tables)	Launch $185; Growth $495	F
LinkedIn ads	$10 a day minimum, about $300 a month	F
HubSpot (alternative CRM)	Starter $20 per seat list	Option
Free tools	Google Search Console, Bing Webmaster Tools, spreadsheets	0

public numbers, product-led conversion and paid search, both start from traffic an early product does not have. Founder calls have no benchmark but give the one input every later motion needs: the buyer's own words about the job and the price. Stages 0 and 1 are therefore almost entirely conversations, with one post and one AI-answer item each week to start the slower motions early.

6.2. Why reply benchmarks are not targets

Two vendors' 2025 averages differ sevenfold because of the choice of denominator, so neither works as a target. The system replaces them with the team's own counts against a minimum sample, and counts positive replies rather than all replies.

6.3. Limitations

Most benchmark sources are vendors who benefit from the result. The surveys measure stated influence, not purchases. The stage thresholds and decision rules are our design defaults and have not been tested across products.

6.4. Directions

Two tests would strengthen the system. First, running it on several products for thirty days and publishing the counts at each stage, including zeros, would show whether the thresholds hold across products. Second, checking AI-answer mentions on a fixed prompt set each month and comparing them with the source buyers report in the CRM would show how often AI answers send buyers who are credited elsewhere.

7. Related work

Graham's essay on doing things that do not scale is the standard argument for manual early sales (Graham, 2013), and the Superhuman account gives a measurable fit threshold (First Round Review, n.d.). Buyer surveys from 6sense and G2 describe how B2B buyers shortlist (G2, 2025; 6sense, 2024), and the LinkedIn B2B Institute's 95-5 rule explains why most outreach finds no one in market (LinkedIn, n.d.a). Vendor benchmarks from Instantly, Belkins, WordStream and ChartMogul provide channel rates (Belkins, 2026; ChartMogul, 2026; Instantly, 2026; WordStream, 2025). Ahrefs and Pew study AI answers from the publisher and the user side (Chapekis and Lieb, 2025; Linehan, 2025). An earlier audit checked one agency playbook claim by claim (Krajewski, 2026). This paper adds stage exit tests counted in buyers, a proof requirement for buyer lists, and minimum-sample decision rules.

8. Conclusion

Public evidence on early go-to-market is uneven. It is fairly strong on how buyers shortlist and on platform rules, moderate on outbound, product-led and paid search rates, and absent on founder calls, posts and referrals. A founder can still run a disciplined week: build a list of people who already pay for the job, talk to them, count will-pay statements and paying buyers, apply written kill and double-down rules every Friday, and add motions only as each stage is met.

Reproducibility

All sources are public and listed with URLs in Appendix A. The figures plot values from those sources, and the stack totals in Figure 6 add the list prices printed in the figure.

References

Allsopp, G. (2025). Do self-promotional "best" lists boost ChatGPT visibility? Ahrefs blog, 4 December 2025. https://ahrefs.com/blog/best-lists-research/

Belkins (2026). Cold email response rates: 2026 benchmarks. https://belkins.io/blog/cold-email-response-rates

The Bridge Group (2025). 2025 SDR models and metrics report. https://www.bridgegroupinc.com/research/2025-sdr-models-metrics-report-the-bridge-group

ChartMogul (2026). SaaS conversion report. https://chartmogul.com/reports/saas-conversion-report/

Chapekis, A. and Lieb, A. (2025). Google users are less likely to click on links when an AI summary appears in the results. Pew Research Center, 22 July 2025. https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/

European Union (2016a). Regulation (EU) 2016/679, Article 14. https://eur-lex.europa.eu/eli/reg/2016/679/art_14/oj

European Union (2016b). Regulation (EU) 2016/679, Article 21. https://gdpr-info.eu/art-21-gdpr/

European Union (2016c). Regulation (EU) 2016/679, Recital 47. https://gdpr-info.eu/recitals/no-47/

Federal Trade Commission (2024). Federal Trade Commission announces final rule banning fake reviews and testimonials, 14 August 2024 (16 CFR Part 465). https://www.ftc.gov/news-events/news/press-releases/2024/08/federal-trade-commission-announces-final-rule-banning-fake-reviews-testimonials

Federal Trade Commission (n.d.). CAN-SPAM Act: a compliance guide for business. https://www.ftc.gov/business-guidance/resources/can-spam-act-compliance-guide-business

First Round Review (n.d.). How Superhuman built an engine to find product/market fit. https://review.firstround.com/how-superhuman-built-an-engine-to-find-product-market-fit/

G2 (2025). 2025 buyer behavior report. https://learn.g2.com/hubfs/G2-2025-Buyer-Behavior-Report-AI-Always-Included.pdf

Geoa (2026). Effect of llms.txt on AI citations. https://geoa.app/en/data/llms-txt-effekt

Google (2024). Email sender guidelines. Gmail Help. https://support.google.com/mail/answer/81126

Google (2026a). Google Cloud acceptable use policy (applies to Google Workspace). https://workspace.google.com/terms/use_policy.html

Google (2026b). AI features and your website: optimization guide. Google Search Central. https://developers.google.com/search/docs/fundamentals/ai-optimization-guide

Google (2026c). Spam policies for Google web search. https://developers.google.com/search/docs/essentials/spam-policies

Graham, P. (2013). Do things that don't scale. https://paulgraham.com/ds.html

hiQ Labs, Inc. v. LinkedIn Corp., No. 17-cv-03301-EMC (N.D. Cal.), consent judgment, 8 December 2022. https://storage.courtlistener.com/recap/gov.uscourts.cand.312704/gov.uscourts.cand.312704.405.0.pdf

Information Commissioner's Office (n.d.). Business-to-business marketing. https://ico.org.uk/for-organisations/direct-marketing-and-privacy-and-electronic-communications/business-to-business-marketing

Instantly (2026). Cold email benchmark report 2026. Updated 12 January 2026. https://instantly.ai/cold-email-benchmark-report-2026

Krajewski, J. (2026). A self-published shortlist leads Google's AI answers but not ChatGPT's: an audit of a 2026 B2B go-to-market playbook. In Motion Automation Research, 26 October 2026.

Linehan, L. (2025). Top brand visibility factors in ChatGPT, AI Mode, and AI Overviews (75k brands studied). Ahrefs blog, 12 December 2025. https://ahrefs.com/blog/ai-brand-visibility-correlations/

LinkedIn (2025). User Agreement, Section 8.2. https://www.linkedin.com/legal/user-agreement

LinkedIn B2B Institute (n.d.a). The 95-5 rule. https://business.linkedin.com/advertise/resources/b2b-institute/b2b-research/trends/95-5-rule

LinkedIn (n.d.b). Maximize your budget. https://business.linkedin.com/advertise/ads/best-practices/maximize-your-budget

LinkedIn (n.d.c). Company list upload requirements. https://www.linkedin.com/help/lms/answer/a423102

Omniscient Digital (2026). First-touch attribution captures 15% of our AI-sourced leads. https://beomniscient.com/blog/first-touch-vs-self-reported-attribution-aeo/

Poyar, K. (2026). Free-to-paid conversion benchmarks by model. Substack note, 4 February 2026. https://substack.com/@kylepoyar/note/c-209940962

SE Ranking (2025). llms.txt: does it affect AI citations? https://seranking.com/blog/llms-txt/; summary in Search Engine Journal, 20 November 2025.

6sense (2024). 2024 B2B buyer experience report. https://6sense.com/science-of-b2b/2024-buyer-experience-report/

Trakkr (2026). Does llms.txt increase AI citations? https://trakkr.ai/trakkr-research/llmstxt-effect/answers/does-llms-txt-increase-ai-citations

Vacca, A. (2026). The only GTM playbook you'll need in 2026. X Article, 23 October 2026. https://x.com/itsalexvacca/status/2113670303842963562

WordStream (2025). Google Ads benchmarks 2025. https://www.wordstream.com/blog/2025-google-ads-benchmarks

X (2026). Automation rules. Archived 12 October 2026. https://web.archive.org/web/20261012153529/https://help.x.com/en/rules-and-policies/x-automation

A. Verification table

Table 7: Claims used in this paper, labels and sources (read 26 October 2026 unless marked). C = confirmed on the primary page, P = partly confirmed, U = unconfirmed.

ID	Claim	Label	Finding and source
Buyer behavior			
V1	81% of buyers had a preferred vendor before speaking with sales; about 69% of the process happens before engaging sellers; 2,509 buyers	C	https://6sense.com/science-of-b2b/2024-buyer-experience-report/
V2	Shortlist influence: AI chatbots 17.1%, review sites 15.1%, vendor sites 12.8%, peers 8.9%, salespeople 8.8%; 1,169 decision makers	C	G2 report PDF. https://learn.g2.com/hubfs/G2-2025-Buyer-Behavior-Report-AI-Always-Included.pdf
V3	About 95% of potential B2B buyers are out of market at a given time	C	https://business.linkedin.com/advertise/resources/b2b-institute/b2b-research/trends/95-5-rule
Outbound			
V4	Average reply rate 3.43%; top quartile 5.5%+; top decile 10.7%+; 58% of replies from step one; data 1 Jan to 18 Dec 2025; all replies per email sent	C	https://instantly.ai/cold-email-benchmark-report-2026
V5	0.45% reply rate; 7,530,489 emails; 34,393 replies; unique replies, auto-replies removed	C	https://belkins.io/blog/cold-email-response-rates
V6	Gmail: spam rate under 0.3%; SPF and DKIM aligned for DMARC; one-click unsubscribe for marketing	C	https://support.google.com/mail/answer/81126
V7	Google acceptable use policy prohibits unsolicited mass email	C	Read 25 Oct 2026. https://workspace.google.com/terms/use_policy.html
AI answers			
V8	Google: no llms.txt or special markup needed for AI features; seeking inauthentic mentions not helpful; structured data not required	C	https://developers.google.com/search/docs/fundamentals/ai-optimization-guide
V9	About 300,000 domains: no measurable llms.txt effect; 10.13% adoption	C	https://seranking.com/blog/llms-txt/
V10	1,912 companies: 43.6% vs 43.3% cited, $p = 0.94$	C	https://geoa.app/en/data/llms-txt-effekt
V11	37,894 domains: same median citations with and without llms.txt	C	https://trakkr.ai/trakkr-research/llmstxt-effect/answers/does-llms-txt-increase-ai-citations

ID	Claim	Label	Finding and source
V12	With an AI summary 8% clicked a result vs 15% without; 1% clicked a summary link	C	https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/
V13	Brand mentions correlate 0.664, 0.709, 0.656 with AI visibility; backlinks 0.218; "best" lists 43.8% of ChatGPT-cited page types	C	Read 25 Oct 2026. https://ahrefs.com/blog/ai-brand-visibility-correlations/; https://ahrefs.com/blog/best-lists-research/
V14	189 leads said AI sent them; CRM first-touch credited 28	C	Read 25 Oct 2026. https://beomniscient.com/blog/first-touch-vs-self-reported-attribution-aeo/
V15	Spam policies cover scaled content abuse and site reputation abuse	C	https://developers.google.com/search/docs/essentials/spam-policies
Product-led and paid			
V16	Good and great free-to-paid ranges by model, about 200 products	P	Seen in the author's note, not a full report. https://substack.com/@kylepoyar/note/c-209940962
V17	Median free-to-paid about 8%; card-required trials about 30%	C	https://chartmogul.com/reports/saas-conversion-report/
V18	Median conversion rate 7.52%, cost per lead $70.11, 16,446 US search campaigns	C	https://www.wordstream.com/blog/2025-google-ads-benchmarks
V19	LinkedIn: $10 daily minimum; company list of 300 or more	C	Read 25 Oct 2026. https://business.linkedin.com/advertise/ads/best-practices/maximize-your-budget; https://www.linkedin.com/help/lms/answer/a423102
Founder-led			
V20	Recruit users manually	C	https://paulgraham.com/ds.html
V21	40% "very disappointed" threshold	C	https://review.firstround.com/how-superhuman-built-an-engine-to-find-product-market-fit/
V22	SDR median 4.1 quality conversations a day; 351 companies	C	https://www.bridgegroupinc.com/research/2025-sdr-models-metrics-report-the-bridge-group
Compliance			
V23	CAN-SPAM: no B2B exception; ad disclosure, postal address, opt-out within 10 business days; up to $53,088 per email	C	https://www.ftc.gov/business-guidance/resources/can-spam-act-compliance-guide-business
V24	GDPR Art. 21(2) to (3): right to object to direct marketing, stated at first communication	C	https://gdpr-info.eu/art-21-gdpr/
V25	GDPR Recital 47: direct marketing may be a legitimate interest	C	https://gdpr-info.eu/recitals/no-47/
V26	GDPR Art. 14 notice when data not collected from the person	C	Read 25 Oct 2026. https://eur-lex.europa.eu/eli/reg/2016/679/art_14/oj
V27	UK: corporate subscribers may be emailed without consent; sole traders need consent	C	Read 25 Oct 2026. https://ico.org.uk/for-organisations/direct-marketing-and-privacy-and-electronic-communications/business-to-business-marketing
V28	LinkedIn prohibits bots, automated messaging and scraping	C	Read 25 Oct 2026. https://www.linkedin.com/legal/user-agreement

ID	Claim	Label	Finding and source
V29	X prohibits automated unsolicited replies, mentions and DMs	P	Live page blocked the fetch; read from an archive of 12 Oct 2026. https://web.archive.org/web/20261012153529/https://help.x.com/en/rules-and-policies/x-automation
V30	FTC rule prohibits fake, bought and undisclosed insider reviews	C	https://www.ftc.gov/news-events/news/press-releases/2024/08/federal-trade-commission-announces-final-rule-banning-fake-reviews-testimonials
Tool prices (monthly billing)			
V31	Clay Launch $185; Growth $495	C	https://www.clay.com/pricing
V32	Instantly Growth $47; Hypergrowth $97; Scale bundle $194	C	https://instantly.ai/pricing
V33	Otterly Lite $29; Standard $189; Premium $489	C	https://otterly.ai/pricing
V34	Attio Free; Plus $44; Pro $99 per user	C	https://attio.com/pricing
V35	Fathom Free; Premium $20; Business $34, 2-seat minimum	C	https://fathom.ai/pricing
V36	Google Workspace Business Starter $7 list; Standard $14	C	https://workspace.google.com/pricing
V37	HubSpot Starter $20 per seat list	C	https://www.hubspot.com/products/crm/starter
V38	Apollo Basic $65 a month	U	Pricing page did not render; only third-party pages. Not used.

B. Templates

B.1. Call script outline (30 minutes)

1. Permission and purpose (2 min): who you are, why them, ask to take notes.
2. Last time they did the job (8 min): what happened, who was involved, what it cost in money and hours.
3. What they use and pay now (5 min): tools, people, contracts, renewal dates.
4. What breaks (5 min): the last failure and its cost.
5. Show, if a product exists (5 min): one flow on their case.
6. Price question (3 min): “If this did that for you, would you pay [price] a month starting [date]?” Record the answer verbatim.
7. Next step with a date (1 min).
8. Referral and source (1 min): “Who else pays for this?” and “How did you hear about us?”

B.2. Thirty-day plan

Table 8: Thirty-day plan for any product starting at Stage 0.

Week	Work	Exit count
1	Write the one-sentence ICP; build 30 proof rows; set up CRM fields and opt-out handling; record an AI-answer baseline on 20 buyer prompts	30 proof rows
2	Send 1:1 messages to 20 rows a day for four days, one follow-up; first founder post from research notes	5 calls booked
3	Hold calls; log will-pay statements; second post and first AI-answer item; ask for a paid pilot where a will-pay statement exists	10 calls held; 3 will-pay statements
4	Follow up on pilots; collect first payment; Friday review with kill, hold or double-down per channel and ICP	1 paying buyer or a written kill

B.3. One-page experiment brief

Each experiment fills seven fields before it starts: hypothesis in one sentence; channel and ICP; minimum sample (from Table 4); start and review dates; owner; cost in hours and money; and the decision each result will trigger.