Who You Reach and Who Writes: First-Party Evidence on Targeting and Message Authorship in AI-Assisted B2B Outbound

Namanyay Goel
PumpGTM
n@pumpgtm.com

Technical report v1.0 November 9, 2026

Abstract

Which parts of an AI-assisted outbound program move response rates: who is contacted, or who writes the message? We study this with first-party production data from PumpGTM, a platform that finds B2B prospects and contacts them over LinkedIn, email and X from the customer's own accounts. The frozen extract covers 127,551 prospect records in 723 customer workspaces (March to November 2026), after removing internal, demo and test workspaces. No new experiments were run. **Targeting.** Prospects drawn from people who had engaged with a related social post accepted LinkedIn connection requests at 57.1% (95% workspace-bootstrap interval 31.5–68.9) against 20.3% (18.2–22.4) for prospects found by model-planned filter search, a difference of +36.9 percentage points (10.3 to 49.1). This rests on 126 prospects in 7 workspaces and should be read as a strong but small-sample signal. Lists compiled by customers' general-purpose AI assistants were not distinguishable from in-product search within the same workspaces. **A public case.** One customer, Supermemory, received replies from 23.7% of its connection requests by November 3 and 27.2% at the freeze. **Message authorship.** Among 7,175 connections who received a first message, the reply rate was 14.7% (12.4–17.0) for PumpGTM's in-product writer and 13.2% (8.6–18.5) for general-purpose assistants writing through PumpGTM's MCP interface with PumpGTM's style rules: a difference of +1.6 points (−3.6 to +6.1), which we read as a tie. **Channel mix.** Whether adding email to LinkedIn raises replies is not yet measurable in this data; we will rerun it once email reply capture has a full period behind it. This is a first-party report by an interested party and has not been peer reviewed. Aggregated data and analysis code are released for reproduction.

1 Introduction

AI writing tools have made cold outreach cheap to produce. Vendors of specialized go-to-market software, including the author's company, routinely claim that their systems find better prospects and write better messages than a general-purpose assistant would. Such claims are rarely tested against the vendor's own production data with the uncertainty reported.

This report asks three questions that PumpGTM's existing data can address without new experiments:

1. **List source.** Do acceptance and reply rates differ by how the prospect list was built: model-planned filter search, engagement with a related social post, lists compiled by a general-purpose assistant, curated lists, or lists the customer uploaded?

2. **Message authorship.** Holding the sending infrastructure fixed, does the reply rate differ when the first message was written by PumpGTM's in-product writer, by a general-purpose AI assistant the customer connected over the Model Context Protocol (MCP), by a fixed template, or by the customer?

3. **Channel mix.** Do prospects contacted on LinkedIn and email reply more often than prospects contacted on LinkedIn alone?

We follow the structure of recent industrial technical reports that pair a stated research question with a frozen dataset, named treatments, paired comparisons with clustered bootstrap intervals and an explicit statement of where the evidence stops [Marquis, 2026]. We report every comparison defined in the analysis code, including those that do not favour PumpGTM.

Summary of findings. The largest differences in the data are between list sources: audiences with an engagement signal accepted far more often than searched lists, in the few workspaces that used them. Who wrote the first message made no detectable difference: PumpGTM's specialized writer and guided general-purpose assistants tied, with intervals that exclude advantages larger than about six points either way. The channel-mix question cannot be answered yet.

2 Related work

Lead scoring and prospect selection. Prior work models which B2B leads convert [Monat, 2011], uses iterative customer-acquisition models to predict prospects [D'Haen and Van den Poel, 2013], and automates lead scoring with machine learning [Nygård and Mezei, 2020]. Research on B2B social selling describes how sellers use LinkedIn to identify and engage prospects [Ancillai et al., 2019], and a large randomized study on LinkedIn found that moderately weak ties are the most useful for job mobility, which bears on the value of connections to strangers [Rajkumar et al., 2022].

Cold outreach and generative AI. Field evidence on email marketing shows that even non-informative personalization, such as adding the recipient's name, raises response [Sahni et al., 2018]. Large language models can write messages tailored to psychological profiles that are more persuasive than generic ones [Matz et al., 2024], although microtargeted messages from language models did not outperform a single well-written generic message in a political setting [Hackenburg and Margetts, 2024]. Workplace studies find that generative AI raises writing productivity and helps less experienced workers most [Noy and Zhang, 2023, Brynjolfsson et al., 2025], with uneven effects across tasks [Dell'Acqua et al., 2023].

Channels and methods. Multichannel marketing research warns that customers reached on more channels self-select, so naive comparisons overstate channel effects [Venkatesan et al., 2007]. We use workspace-clustered bootstrap intervals [Efron and Tibshirani, 1993, Cameron et al., 2008], paired within-workspace contrasts where possible, and a negative-control outcome to detect confounding [Lipsitch et al., 2010]. Online controlled experiments remain the standard for causal claims about product changes [Kohavi et al., 2020]; this study is observational and should be read accordingly [Rosenbaum, 2002].

Table 1: Activity in customer workspaces after exclusions, by month of the event. Earlier months (March to June) together hold 20 connection requests.

Month	Workspaces sending	LinkedIn requests	Accepted	First messages	Emails sent	Replies (any channel)
2026-07	10	123	8	8	0	0
2026-08	45	1,144	209	42	0	10
2026-09	125	5,966	1,120	415	7	76
2026-10	465	46,251	10,946	9,883	2,874	1,553
2026-11	181	7,854	2,117	2,537	1,771	384
All (Mar–Nov)	–	61,358	14,400	12,885	4,652	2,023

3 Data

3.1 Source and freeze

All data come from PumpGTM's production database, read-only, frozen at 2026-11-07 07:00 UTC. The platform records each prospect with the sequence that contacted them, timestamps for the LinkedIn connection request, acceptance, first message, reply and any recorded meeting, the sourcing strategy or import batch that produced them, and per-channel delivery tracks. Every revision of every sequence is stored with its full message text and an edit source (in-product system, dashboard, MCP or pre-revision backfill). Inbound replies since mid-October carry a model classification such as interested, question, objection or not interested.

3.2 Exclusions and privacy

We removed 85 internal, demo and test workspaces by name pattern, including PumpGTM's own outbound workspace and a sales-demo workspace that contains synthetic, time-shifted activity. The remaining 127,551 prospect records belong to 723 workspaces, of which 618 sent at least one connection request. To give replies time to arrive, outcome analyses include only prospects contacted at least 14 days before the freeze, that is before 2026-10-24. Only aggregate counts leave the database: no prospect names, message text, customer names or workspace identifiers appear in this report or its dataset, with the exception of one customer that is already a public case. PumpGTM's privacy policy covers using customer data to provide and improve the service and retaining anonymous, aggregated usage data for analytics. Neither it nor the terms of service addresses research publication, so this report publishes aggregates only.

3.3 Data audit

Table 1 summarises activity after exclusions and Table 2 what the data can support. Volume is concentrated in October 2026. Two features of the data shape what can be studied. First, meetings are almost never recorded in the database (three customer meetings in the whole period), so we use acceptances and replies instead. Second, email sending began at scale only in October, and automatic capture of email replies inside multi-channel sequences began on 2026-10-15.

Table 2: What existing data can and cannot support.

Question	Status
Target account mapping, signal quality	Partly supported: acceptance and reply by list source, customer review of proposed prospects, time to build a list.
Specialized writer vs. general-purpose AI	Supported as an observational comparison. No unguided generic-AI baseline exists: assistants writing over MCP receive PumpGTM's style rules and a campaign check. Copy written with an outside assistant and pasted in is not identifiable.
Multi-channel vs. single-channel	Not yet measurable. Email reply capture was incomplete, and the email group is strongly selected (Section 5.4).
Meetings booked	Not supported: almost no meetings are recorded.
Speed vs. a manual baseline	Not supported: there is no record of how long manual list building takes. PumpGTM's own build times are reported descriptively.

4 Methods

4.1 List source

A prospect who received a connection request before 2026-10-24, with acceptance and reply per request as outcomes. Sources are: *model-planned filter search* (the reference), where a language model turns the customer's description into structured people-search filters; *engagement signal*, people who reacted to or commented on a related social post; *list built by a general-purpose assistant*, prospects the customer's assistant compiled and imported over MCP; *in-app assistant list*, prospects proposed by PumpGTM's in-app assistant; *curated person list*, prebuilt lists such as accelerator founder maps; *customer list*, uploaded or exported lists; and *account signals*, competitor customers and lookalike accounts. We also report how customers reviewed proposed prospects and how long completed list-building runs took.

4.2 Message authorship

Unit and outcome. A prospect who accepted a connection request and received a first LinkedIn message before 2026-10-24. The outcome is whether the prospect replied on any channel by the freeze. A secondary outcome is the share of model-classified replies that were positive (meeting intent, interested, question, relevant later or referral).

Attributing authorship. For each sequence we walk its revisions in order and assign the copy to the author of the last substantive change. A revision counts as a new author's copy when the word-bigram Jaccard similarity of all message text to the previous copy falls below 0.6; smaller edits keep the previous author. The author of a substantive change is given by the edit source: MCP edits are the *general-purpose assistant*; dashboard edits are the *customer*; in-product system writes are the *in-product writer* when the sequence was generated by PumpGTM's language-model writer and a *fixed template* when it came from a template, a fallback or a fixed onboarding message. Each prospect is assigned the author whose copy was live when their first message was sent.

Treatments.

- **In-product writer.** PumpGTM's sequence generator, prompt families `cardless_sequence_prompt` v6–v8 and `campaign_lanes` v1–v2, running on `gemini-3.6-flash` until 2026-10-22 and `gemini-3.5-flash-lite` afterwards, given a confirmed analysis of the customer's company. Outputs that fail an unsupported-claims check are replaced by a fixed fallback, which is counted as a template.
- **General-purpose assistant (baseline).** The customer's own AI assistant (desktop and web chat assistants and coding agents) writing the sequence through PumpGTM's MCP server. The assistant receives PumpGTM's written style and follow-up rules and can call a campaign check before saving. The underlying model and version are chosen by the customer and are not recorded. This is therefore a *guided* general-purpose baseline, not an unguided one.
- **Fixed template.** Product templates, fallbacks and fixed onboarding messages.
- **Customer-written.** Copy written or substantially rewritten in the dashboard.

In-product writes whose generator could not be identified (1,157 prospects) are reported but not used in contrasts.

4.3 Channel mix

A prospect who received a LinkedIn connection request before 2026-10-24. We compare prospects in LinkedIn-only sequences with prospects in LinkedIn-plus-email sequences, and within LinkedIn-plus-email sequences, prospects who were sent at least one email with those for whom no address was found. The outcome is a reply on any channel. LinkedIn acceptance serves as a negative-control outcome: the connection request precedes the first email, so email cannot cause acceptance, and a difference in acceptance shows that the groups differ for other reasons.

4.4 Statistics

For each arm we report the pooled rate (total successes over total prospects) with a 95% percentile interval from 10,000 bootstrap replicates that resample whole workspaces with replacement, since prospects within a workspace share a product, sender and audience [Cameron et al., 2008]. Differences between arms are computed within each replicate from the same resampled workspaces. Where at least three workspaces have 10 or more prospects in both arms, we also report a *paired* contrast: the unweighted mean of within-workspace rate differences, with a workspace bootstrap interval, and how many workspaces favour each arm. The seed is 20261106. No multiplicity correction is applied, so intervals should be read as descriptive. Intervals for groups with fewer than about 50 prospects or 5 workspaces are unreliable and are shown for completeness.

5 Results

5.1 Where the list came from

List source produces the largest differences in the data (Tables 3 and 4, Figure 1). People who had engaged with a related post accepted at 57.1% (31.5–68.9) against 20.3% (18.2–22.4) for filter search, a pooled difference of +36.9 pp (10.3 to 49.1). In the three workspaces that used both, engagement audiences accepted more often in all three (+36.0 pp, 27.4 to 48.8). Their reply rate

Table 3: Acceptance and reply per connection request by list source.

Source	Workspaces	Requests	Accepted %	Replied %
Engagement signal (post engagers)	7	126	57.1 [31.5, 68.9]	27.8 [1.8, 51.6]
List built by a general-purpose assistant	37	2,050	27.7 [18.7, 36.7]	6.5 [2.0, 12.0]
Model-planned filter search	315	27,099	20.3 [18.2, 22.4]	2.7 [2.2, 3.3]
In-app assistant list	324	2,420	19.5 [17.1, 21.8]	2.4 [1.6, 3.4]
Curated person list	39	4,550	16.5 [13.3, 20.6]	1.1 [0.5, 1.6]
Customer upload / exported list	3	402	15.9 [10.5, 16.8]	2.2 [0.0, 4.7]
Account signals (competitor customers, lookalikes)	5	20	30.0 [0.0, 38.9]	0.0 [0.0, 0.0]

Table 4: Each source minus model-planned filter search, percentage points. Paired columns use workspaces with at least 10 prospects from both sources.

	Pooled [95% CI]			Paired [95% CI]	
Source	Accepted	Replied	Paired workspaces	Accepted	Replied
Engagement signal (post engagers)	+36.9 [+10.3, +49.1]	+25.1 [-0.8, +50.1]	3	+36.0 [+27.4, +48.8]	+36.8 [-2.1, +66.4]
List built by a general-purpose assistant	+7.4 [-1.9, +16.2]	+3.8 [-0.8, +9.2]	15	-0.5 [-8.9, +7.2]	+1.7 [-2.0, +6.8]
In-app assistant list	-0.8 [-3.5, +1.9]	-0.3 [-1.1, +0.6]	84	-1.0 [-4.5, +2.5]	+0.3 [-0.9, +1.7]
Curated person list	-3.7 [-7.6, +0.8]	-1.6 [-2.5, -0.9]	12	+6.5 [-3.2, +17.5]	+1.8 [-1.7, +5.7]
Customer upload / exported list	-4.4 [-10.4, -2.2]	-0.5 [-2.8, +2.2]	1	–	–
Account signals (competitor customers, lookalikes)	+9.7 [-20.0, +19.0]	-2.7 [-3.3, -2.2]	0	–	–

per request was 27.8% against 2.7%, but with only 126 prospects in 7 workspaces the interval for that difference runs from −0.8 to +50.1 pp.

Lists compiled by general-purpose assistants show higher pooled rates than in-product filter search (acceptance 27.7% against 20.3%, reply 6.5% against 2.7%), but within the 15 workspaces that used both the differences are small and uncertain (acceptance −0.5 pp, −8.9 to +7.2; reply +1.7 pp, −2.0 to +6.8), so the pooled gap mostly reflects which customers import their own lists. Lists from the in-app assistant perform like filter search (84 paired workspaces, reply +0.3 pp, −0.9 to +1.7). Curated lists have a lower pooled reply rate but no difference within workspaces.

Customer review and speed. Of 103,374 prospects proposed by filter search, 69,949 were kept by the customer or by automatic import and 6,778 were explicitly rejected (8.8% of those decided). Most rejections carry no reason; the most common stated reasons are wrong company segment (413) and wrong job function (285). Completed filter-search runs took a median of 7 seconds (90th percentile 45 seconds) and returned a median of 12 new candidates, because most runs top up an existing list. Searches over competitors' customers took a median of 473 seconds. There is no baseline for manual list building, so these numbers describe PumpGTM only.

5.2 A public customer case

Supermemory, whose results PumpGTM has described publicly, illustrates what a narrowly targeted list can do and what the aggregate rates hide. By 2026-11-03 its workspace had sent 524 connection requests and received replies from 23.7% of those prospects. At the freeze the figures were 588 requests, 56.6% accepted and 27.2% replied (49.1% of the connections who received a message). PumpGTM staff helped operate this account, and it is one workspace, so it is a case, not evidence about the population.

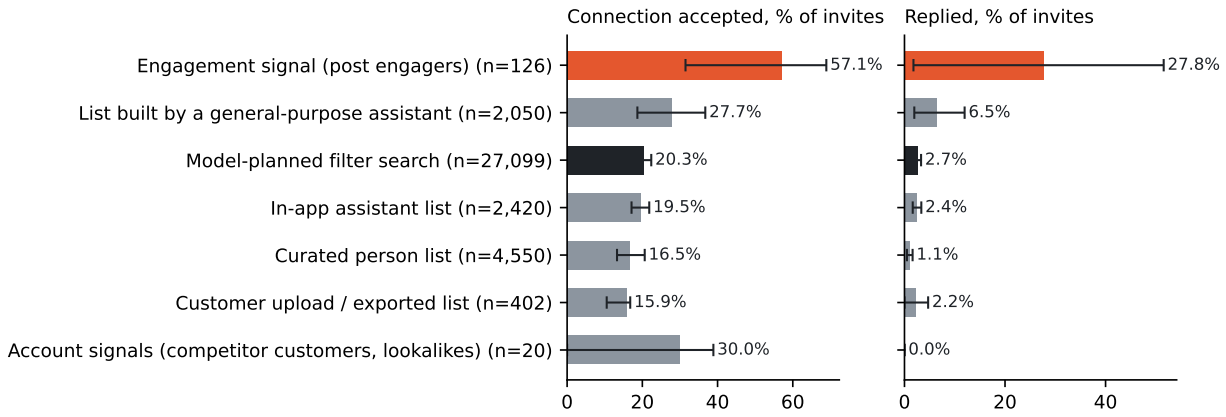


Figure 1: Acceptance and reply per connection request by list source.

Table 5: Reply rate per messaged connection by copy author. Brackets are 95% workspace-bootstrap intervals. *Search audiences* restricts to prospects found by model-planned filter search. Positive share is shown where at least 20 replies were classified.

		All audiences		Search audiences		Positive share
Author	Workspaces	Messaged	Reply %	Messaged	Reply %	of replies %
In-product writer	194	2,699	14.7 [12.4, 17.0]	2,216	13.9 [11.7, 15.8]	49 [42, 55]
General-purpose assistant (MCP)	74	1,351	13.2 [8.6, 18.5]	796	15.1 [9.7, 21.1]	60 [48, 71]
Fixed template	85	1,857	15.8 [12.2, 19.5]	107	23.4 [10.3, 37.6]	60 [50, 71]
Customer-written	6	54	38.9 [0.0, 57.1]	3	0.0 [0.0, 0.0]	–
In-product, unclassified	49	1,157	14.8 [10.4, 18.5]	–	–	–

5.3 Who wrote the message

Tables 5 and 6 and Figure 2 show the results. The in-product writer and the guided general-purpose assistant are statistically indistinguishable on reply rate in every cut. The pooled difference is +1.6 pp (−3.6 to +6.1); among prospects from the same audience source it is −1.2 pp (−7.7 to +4.6). Only five workspaces used both authors on at least ten messaged prospects, so the paired contrast is uninformative. On the secondary outcome, replies to assistant-written copy were more often positive (60% against 49%, difference −11.0 pp, −22.2 to +0.9), a difference that falls just short of conventional significance and points against the specialized writer.

Fixed templates did about as well as either model overall (15.8%). Their audience mix differs: a fixed message sent to people who had engaged with a related post had the highest rate of any template group, consistent with the list-source results. Customer-written copy shows a high rate (38.9%) but comes from 54 prospects in 6 workspaces, and its interval spans 0 to 57%.

Interpretation. The data do not show that PumpGTM's specialized writer beats general-purpose assistants. Both arms, however, operate under PumpGTM's rules: the writer by construction and the assistant through MCP instructions and a campaign check. The result is compatible with two readings that this data cannot separate: that the shared rules and context matter more than which model applies them, or that copy differences in this range matter little next to audience.

Table 6: In-product writer minus general-purpose assistant, percentage points.

Contrast	Difference	95% interval	Workspaces favouring writer / assistant
All messaged leads, pooled	+1.6	[-3.6, +6.1]	–
Model-planned search audiences only, pooled	-1.2	[-7.7, +4.6]	–
October 2026 only, pooled	+2.2	[-3.7, +7.1]	–
Within workspace, all audiences	-1.1	[-19.1, +16.9]	2 / 3 of 5
Within workspace, search audiences	-2.0	[-16.7, +16.7]	1 / 2 of 3
Positive share of classified replies	-11.0	[-22.2, +0.9]	–

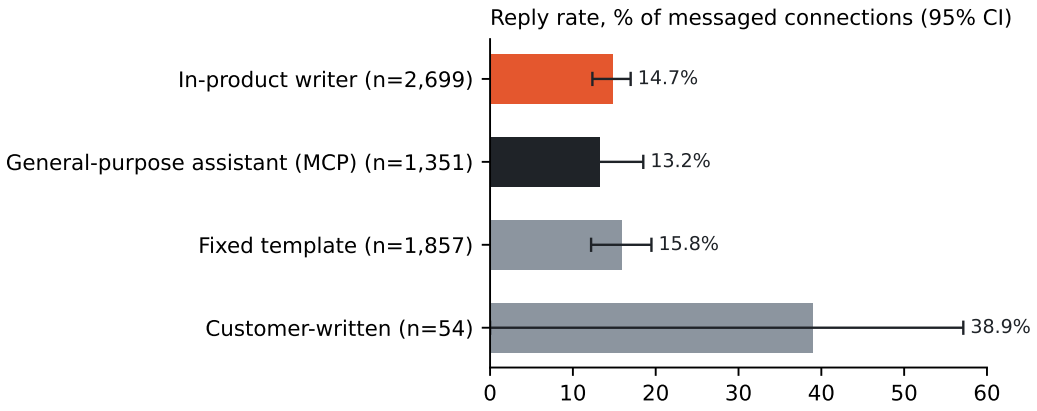


Figure 2: Reply rate per messaged connection by who wrote the copy live at send time.

5.4 Channel mix: not yet measurable

Naively, LinkedIn-only sequences show a higher reply rate per request (2.8%) than LinkedIn-plus-email sequences (2.2%), and within those sequences prospects who were sent email replied at 1.7% against 2.2% for those without an address (Table 7, Figure 3). None of this is evidence about email. In the six workspaces with enough prospects in both groups, prospects who later received email had accepted the LinkedIn request at a rate 32.1 pp higher (5.7 to 63.0, higher in all six), although acceptance happens before any email is sent. The groups differ in who they are and when they were contacted. In addition, automatic capture of email replies inside sequences began on 2026-10-15, and only 42 email replies were logged across the whole platform during the period. Reply drafts, which record the channel of each inbound reply, list 3,099 LinkedIn replies, 31 on X and none by email in customer workspaces. We therefore treat the channel-mix question as not yet measurable and will rerun it once email replies have been captured for a full period.

6 Where the evidence stops

This is observational first-party data. No treatment was randomized. Customers chose where to find prospects and how to write copy, and those choices correlate with product, market, sender seniority and skill. Paired within-workspace contrasts remove stable workspace differences but rest on few workspaces. The engagement-signal result in particular comes from 7 workspaces.

The general-purpose baseline is guided. Assistants writing over MCP received PumpGTM's style rules and could run its campaign check, and their underlying models are not recorded. We

Table 7: Acceptance and reply per connection request by sequence design. Acceptance is a negative control for the email contrast.

Group	Workspaces	Requests	Accepted %	Replied %
LinkedIn-only sequence	414	32,946	20.8 [18.9, 22.6]	2.8 [2.2, 3.5]
LinkedIn + email, email sent	13	478	19.5 [8.5, 43.4]	1.7 [0.3, 3.9]
LinkedIn + email, no address found	57	4,226	17.3 [12.0, 23.7]	2.2 [1.0, 3.8]

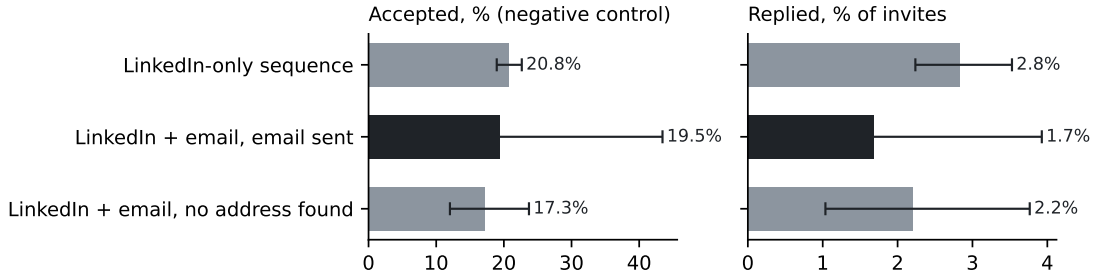


Figure 3: Channel mix. Acceptance precedes any email, so the acceptance gap within paired workspaces shows the email group was different before any email was sent.

cannot say how an unguided general-purpose model would perform. Copy that customers wrote with an outside assistant and pasted into the dashboard is indistinguishable from customer-written copy.

Outcomes are partial. Replies are measured reliably on LinkedIn and on email only from mid-October. Meetings and revenue are not measured. Reply rates are read at the freeze; prospects contacted in late October had two to three weeks to reply.

Attribution is heuristic. Authorship uses a similarity threshold of 0.6 on message text. Arm sizes move with the threshold, and some in-product writes cannot be attributed.

Customer mix and time. Volume grew roughly eightfold from September to October as new customers arrived, and the writer's model changed on 2026-10-22. Restricting the authorship comparison to October does not change the conclusion (+2.2 pp, −3.7 to +7.1). Workspaces that churned remain in the data; workspaces deleted before the freeze do not.

What the results do not show. They do not show that any list source raises meetings or revenue, that PumpGTM's writer is better or worse than general-purpose assistants, or that multi-channel outreach works or does not.

7 Conflict of interest

The author is the founder and CEO of PumpGTM and holds equity in it. PumpGTM sells the product studied here and has a commercial interest in favourable results. The analysis was designed and run by PumpGTM with AI assistance. No external party funded, reviewed or approved this

report. To limit selective reporting, every comparison defined in the analysis code is reported, including those that do not favour PumpGTM.

8 Data and code availability

The reproduction package at `https://pumpgtm.com/research/who-you-reach-and-who-writes` contains the extraction definitions, the authorship and aggregation code, the anonymised aggregate dataset (counts by randomly coded workspace, arm and month, with no prospect-level rows, names, message text or workspace identifiers), the statistics and figure code, and these LaTeX sources. Re-running the statistics from the released aggregates reproduces every number in this report. Re-running the extraction requires access to PumpGTM's production database and is not possible for outside readers.

Acknowledgements

Thanks to the PumpGTM customers whose aggregated, anonymised activity makes up this dataset.

References

Chiara Ancillai, Harri Terho, Silvio Cardinali, and Federica Pascucci. Advancing social media driven sales research: Establishing conceptual foundations for B-to-B social selling. *Industrial Marketing Management*, 82:293–308, 2019. doi: 10.1016/j.indmarman.2019.01.002.

Erik Brynjolfsson, Danielle Li, and Lindsey Raymond. Generative AI at work. *The Quarterly Journal of Economics*, 140(2):889–942, 2025. doi: 10.1093/qje/qjae044.

A. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427, 2008. doi: 10.1162/rest.90.3.414.

Fabrizio Dell'Acqua, Edward McFowland, Ethan R. Mollick, Hila Lifshitz-Assaf, Katherine Kellogg, Saran Rajendran, Lisa Krayer, François Candelon, and Karim R. Lakhani. Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality. Working Paper 24-013, Harvard Business School, 2023.

Jeroen D'Haen and Dirk Van den Poel. Model-supported business-to-business prospect prediction based on an iterative customer acquisition framework. *Industrial Marketing Management*, 42(4): 544–551, 2013. doi: 10.1016/j.indmarman.2013.03.006.

Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 1993.

Kobi Hackenburg and Helen Margetts. Evaluating the persuasive influence of political microtargeting with large language models. *Proceedings of the National Academy of Sciences*, 121(24): e2403116121, 2024. doi: 10.1073/pnas.2403116121.

Ron Kohavi, Diane Tang, and Ya Xu. *Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing*. Cambridge University Press, 2020.

Marc Lipsitch, Eric Tchetgen Tchetgen, and Ted Cohen. Negative controls: A tool for detecting confounding and bias in observational studies. *Epidemiology*, 21(3):383–388, 2010. doi: 10.1097/EDE.0b013e3181d61eeb.

Guillaume Marquis. Domain-specialized company memory for grounded marketing recommendations. Technical report, Pancake, 2026. URL `https://pancake.ai/research/marketing-memory`. v1.0.0.

S. C. Matz, J. D. Teeny, S. S. Vaid, H. Peters, G. M. Harari, and M. Cerf. The potential of generative AI for personalized persuasion at scale. *Scientific Reports*, 14:4692, 2024. doi: 10.1038/s41598-024-53755-0.

Jamie P. Monat. Industrial sales lead conversion modeling. *Marketing Intelligence & Planning*, 29 (2):178–194, 2011. doi: 10.1108/02634501111117610.

Shakked Noy and Whitney Zhang. Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192, 2023. doi: 10.1126/science.adh2586.

Robert Nygård and József Mezei. Automating lead scoring with machine learning: An experimental study. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020. doi: 10.24251/HICSS.2020.177.

Karthik Rajkumar, Guillaume Saint-Jacques, Iavor Bojinov, Erik Brynjolfsson, and Sinan Aral. A causal test of the strength of weak ties. *Science*, 377(6612):1304–1310, 2022. doi: 10.1126/science.abl4476.

Paul R. Rosenbaum. *Observational Studies*. Springer, 2nd edition, 2002.

Navdeep S. Sahni, S. Christian Wheeler, and Pradeep Chintagunta. Personalization in email marketing: The role of noninformative advertising content. *Marketing Science*, 37(2):236–258, 2018. doi: 10.1287/mksc.2017.1066.

Rajkumar Venkatesan, V. Kumar, and Nalini Ravishanker. Multichannel shopping: Causes and consequences. *Journal of Marketing*, 71(2):114–132, 2007. doi: 10.1509/jmkg.71.2.114.