Honeypot Buyer

An inside look at the coordinated inauthentic behaviour economy, from the buyer's seat of a paid-review vendor

Mike Wright · ORCID 0009-0008-6998-9917

Independent researcher (mikesploit), Summerside, Prince Edward Island, Canada · contributor, Trust Report Canada

Preprint · 9 November 2026 · DOI 10.5281/zenodo.24120754 · https://mikesploit.com/papers/honeypot-buyer

Abstract. Most research on coordinated inauthentic behaviour (CIB) studies networks after they have been detected: takedown datasets, leaked tasking documents and platform reports. Less is known about the commercial edge where CIB is sold, one customer at a time, to small businesses. This paper reports a fourteen-day engagement in which the author, operating a small investigative Facebook page, accepted an unsolicited offer of reviews, likes and follows and played the vendor's buyer. The buyer was a honeypot: operator-only contact, no payment, a decoy target the operator controlled, and a stopping rule. In that time the vendor delivered a sample review through a separate persona account, quoted a price curve from USD 3 to USD 0.50 per review, offered five payment rails and disclosed two of them, a reused Bitcoin address and a UK e-money account, stated the account's registered holder in response to a screenshot pretext, and finally asked the buyer to create and hand over a new Facebook page or account. We describe the vendor's tactics, techniques and procedures (TTPs) and map them to the DISARM framework. We analyse the social engineering on both sides through persuasion and negotiation psychology, and release the pretext artifact as an open-source tool. The case shows that the buyer role is a productive collection position. It yields delivery, pricing, payment and asset-acquisition indicators that passive observation cannot reach, and it shows the retail CIB economy recruiting its own customers as a source of accounts.

Keywords: coordinated inauthentic behaviour; fake reviews; engagement economy; social engineering; elicitation; honeypot; DISARM; counter-fraud; OSINT

1. Introduction

On 25 October 2026 a Facebook profile sent a direct message to a small Facebook page the author runs to examine marketing claims. The message offered "Google review / Facebook page review / trustpilot review / yelp review / BB review all review services", plus page likes, follows, Instagram engagement and YouTube "seo". Messages like this reach small-business pages constantly. Recipients ignore them, block the sender, or, often enough to sustain the trade, buy.

The CIB literature mostly sees this trade from the far end. Platform transparency reports describe networks after removal. Academic work reconstructs markets from recruitment groups and from product-review data. Regulators describe the harm in aggregate. What is rarely documented is the transaction itself: how a vendor sells, what it delivers first, how it prices, how it wants to be paid, and what it asks of the customer once the customer is engaged.

This paper documents one such transaction from the inside, using the buyer's seat as a collection position. Its contributions are:

- a primary-source account of a retail CIB sale from first contact to the vendor's request for account assets, with every message captured and hashed;
- a TTP description of the vendor, mapped to the DISARM Red framework;
- an analysis of the elicitation techniques the honeypot buyer used, and the vendor's own persuasion techniques, grounded in established persuasion and negotiation research;
- an open-source, configurable pretext artifact for authorized elicitation research, with rules of use;
- a discussion of the ethics and limits of buyer-side engagement.

The paper is deliberately a TTP and method write-up. Enforcement outcomes are out of scope; the case's public record and reporting are maintained separately by Trust Report Canada (2026).

2. Background

2.1 Coordinated inauthentic behaviour as a market

Meta introduced "coordinated inauthentic behaviour" for coordinated efforts to manipulate public debate in which inauthentic accounts are central to the operation (Meta, 2018 onward). The term grew out of political influence operations, but the same machinery of persona accounts, coordinated engagement and paid amplification is sold commercially. A review posted by a persona account for a paying business is CIB at retail scale: coordinated (vendor, persona and seeding accounts act together), inauthentic (the reviewer is not a customer) and behavioural (the harm is in what the accounts do, not in any single piece of content).

He, Hollenbeck and Proserpio (2022) documented a large market for fake product reviews organised through Facebook groups, in which sellers recruit reviewers and compensate them, and showed that purchased reviews raise ratings and sales in the short term. Regulators have since moved: the US Federal Trade Commission's 2024 rule on consumer reviews and testimonials (16 CFR Part 465) prohibits buying or selling fake reviews, and in Canada the Competition Act's deceptive marketing provisions, enforced by the Competition Bureau, apply to reviews and testimonials that misrepresent customer experience.

2.2 DISARM

The DISARM framework catalogues influence-operation behaviour as tactics (phases) and techniques, in the style of MITRE ATT&CK (DISARM Foundation, n.d.). Its Red framework includes techniques that fit the review trade directly, notably *T0109 Consumer Review Networks* (TA07 Select Channels and Affordances) and the asset techniques under *TA15 Establish Assets*. Section 5 uses DISARM identifiers as published in the framework's repository, checked on 9 November 2026.

2.3 Social engineering and elicitation

Elicitation, obtaining information without asking for it directly, is a core social-engineering skill (Hadnagy, 2018). Its mechanics draw on well-studied effects: reciprocity and commitment (Cialdini, 2021), the foot-in-the-door effect (Freedman & Fraser, 1966), anchoring (Tversky & Kahneman, 1974), the sunk-cost effect (Arkes & Blumer, 1985) and the need for cognitive closure (Kruglanski & Webster,

1996). Section 6 shows both parties using these effects, and the engagement turning on who controls time.

3. Method

3.1 Setting and roles

The vendor contacted the page unprompted. The author (the "operator") replied as a buyer acting for an unnamed manager who approves spending. Every outbound message was typed by the operator; no automated or agent-generated messages were sent to the subject. Analysis, capture handling and enrichment were done separately from the conversation.

3.2 Rules of engagement

Table 1. Rules of engagement in force for the engagement.

Rule	In practice
No payment	No money, voucher or proof of payment was ever sent. Any purchase would have needed a separate written approval and spending cap; none was given.
Decoy target	The only property offered for delivery was the operator's own page. No third-party business received a review on our behalf.
No asset handover	No page, account, login or access was created for or given to the vendor.
Transaction-bounded asks	Every detail elicited is one a real payment would require (a payment route, a beneficiary name). Nothing personal beyond that was sought.
Screenshot, never link	The vendor was never sent a link to anything we host and never entered data into our systems.
Third parties	Accounts and businesses touched by the vendor's delivery were recorded but not contacted or profiled beyond what the delivery itself displayed.
Stopping rule	The engagement ends when the vendor asks for something the rules forbid us to give. The vendor's request for a page or account (section 4) triggered it.

3.3 Capture and enrichment

Each turn of the conversation was screenshotted on the operator's device and filed the same day under a UTC-timestamped name containing the first eight characters of its SHA-256 hash, then recorded in a case manifest. Times quoted here are the operator's device clock (Atlantic Canada). Payment identifiers were enriched passively: the Bitcoin address through public block explorers and wallet-clustering services; the bank details through offline IBAN checksum validation and public BIC and sort-code registries. No account, service or person was contacted for enrichment. Where third-party data appears below (the delivery persona and its reactions), it is what the vendor's own delivery displayed to the operator.

3.4 Presentation

Figures 1 to 9 are re-typeset from the filed screenshots. Message text is verbatim, including spelling; payment details, the stated account holder and the persona's name are redacted. Some intervening messages are omitted for length.

4. The engagement

Table 2. Timeline (operator device clock, 2026).

When	Event
25 Oct	Vendor's unsolicited offer (Fig. 1).
26 Oct	Operator replies "Yes we do", then asks for a trial; vendor agrees (Fig. 2).
27 Oct	Operator names the decoy target; vendor asks for quantity and keywords and drafts review text for approval.
28 Oct	Vendor reports delivery with a screenshot of the target's Reviews tab (Fig. 3).
31 Oct	Volume anchor (10,000); price falls from USD 3 to USD 0.50 per unit (Fig. 4). Operator defers to a "manager"; vendor lists payment rails and follows up (Fig. 5). Operator states what "his side" will need.
1 Nov	Vendor sends a Bitcoin address and "I am ready to post review" (Fig. 6).
2–3 Nov	Operator reports that Bitcoin "can't work"; vendor sends UK bank details (Fig. 7).
3 Nov	Operator sends a screenshot of a transfer form flagging the account name; vendor states a company name (Fig. 8). Vendor asks for proof of payment, then reports the target's review option as disabled.
7–8 Nov	Vendor asks the operator to create a new Facebook page, then a new account, and send it to him (Fig. 9). Stopping rule met.
9 Nov	Operator disengages without reply.

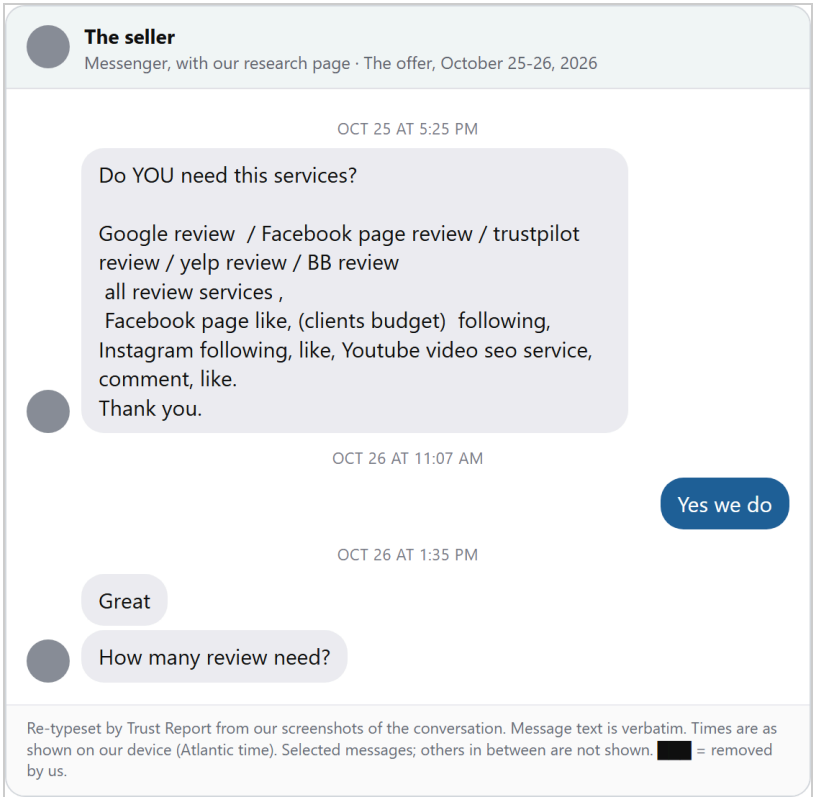


Figure 1. First contact: a cross-platform engagement menu, unsolicited.

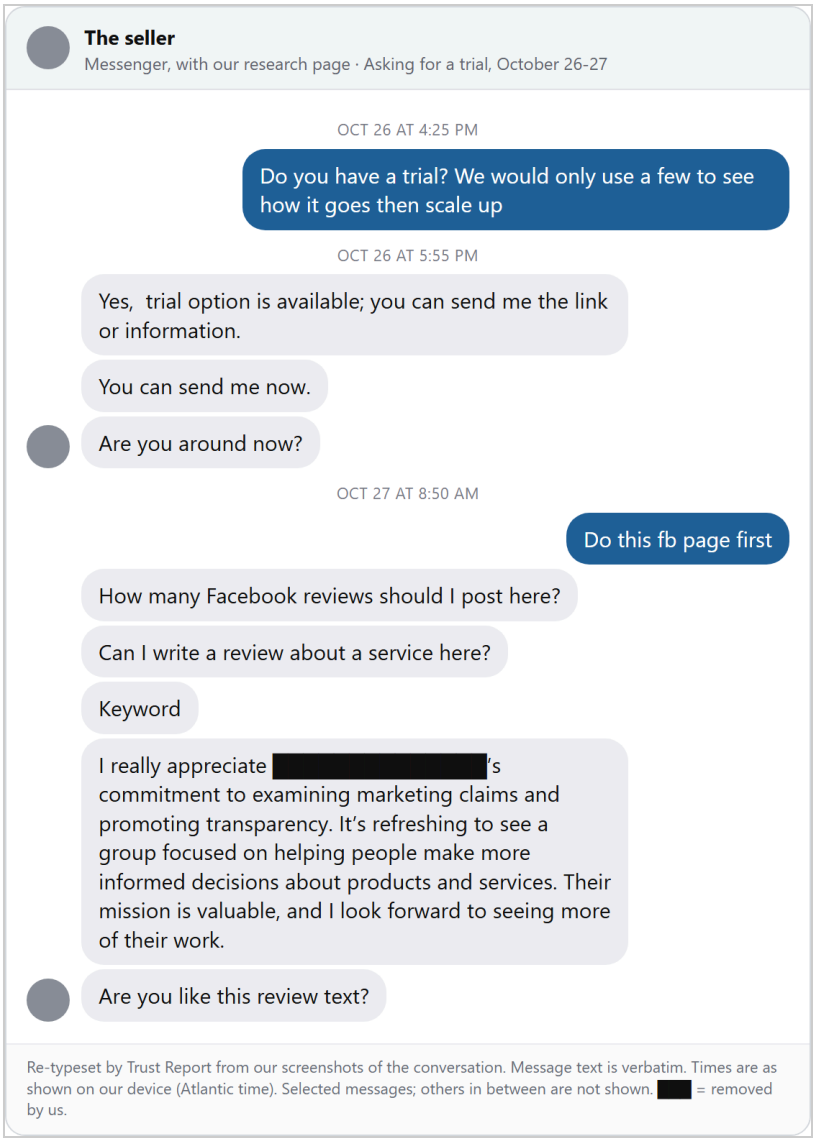


Figure 2. The trial. The vendor asks for quantity and keywords, then drafts the review for the buyer's approval.

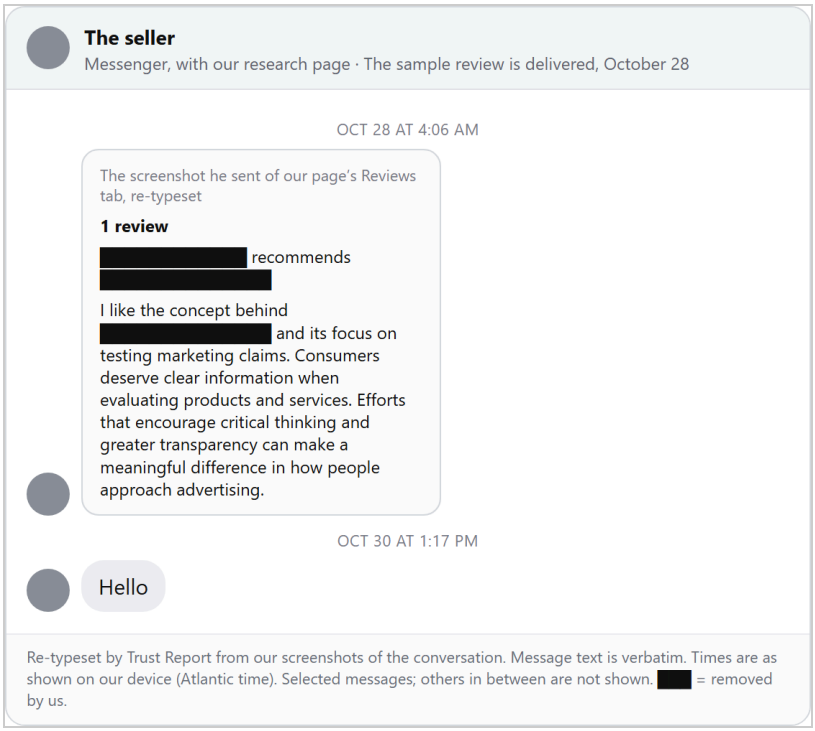


Figure 3. Delivery report. The posted text differs from the draft but makes the same points; it was posted by a separate persona account (redacted).

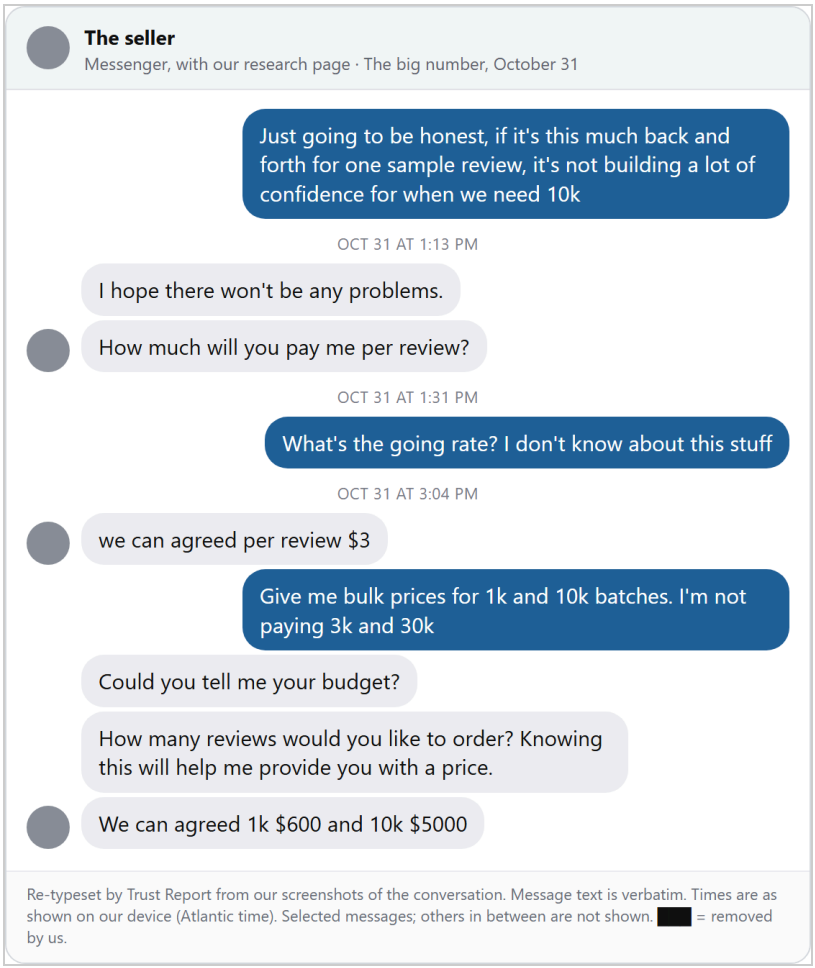


Figure 4. The volume anchor and the price curve.

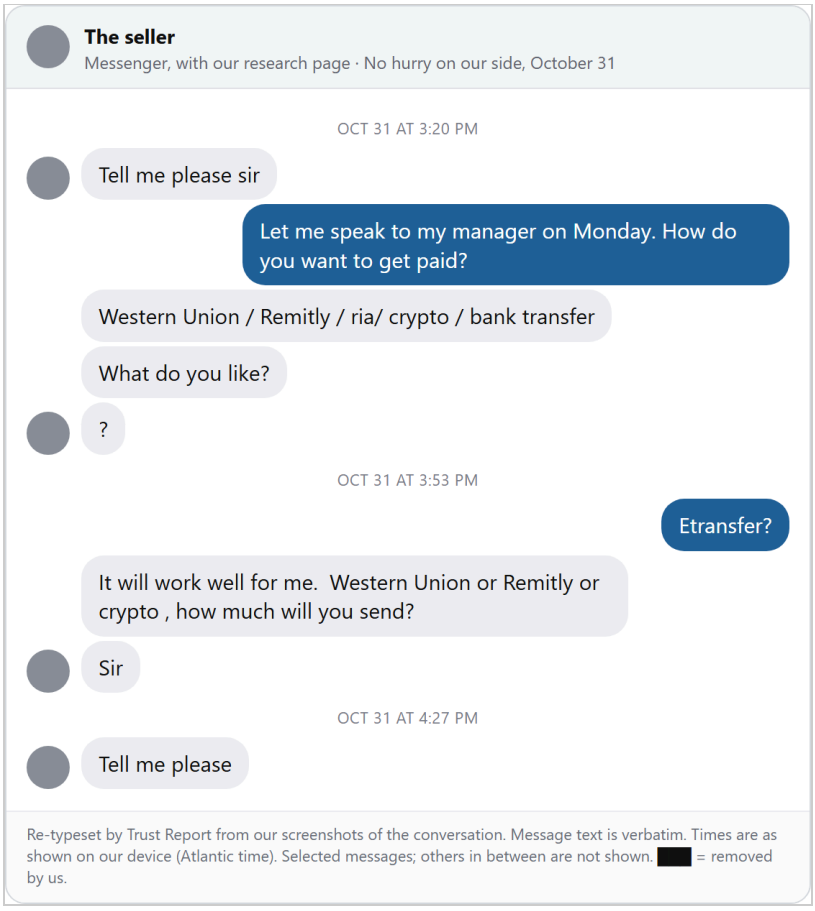


Figure 5. The absent principal. The vendor lists five payment rails and starts chasing.

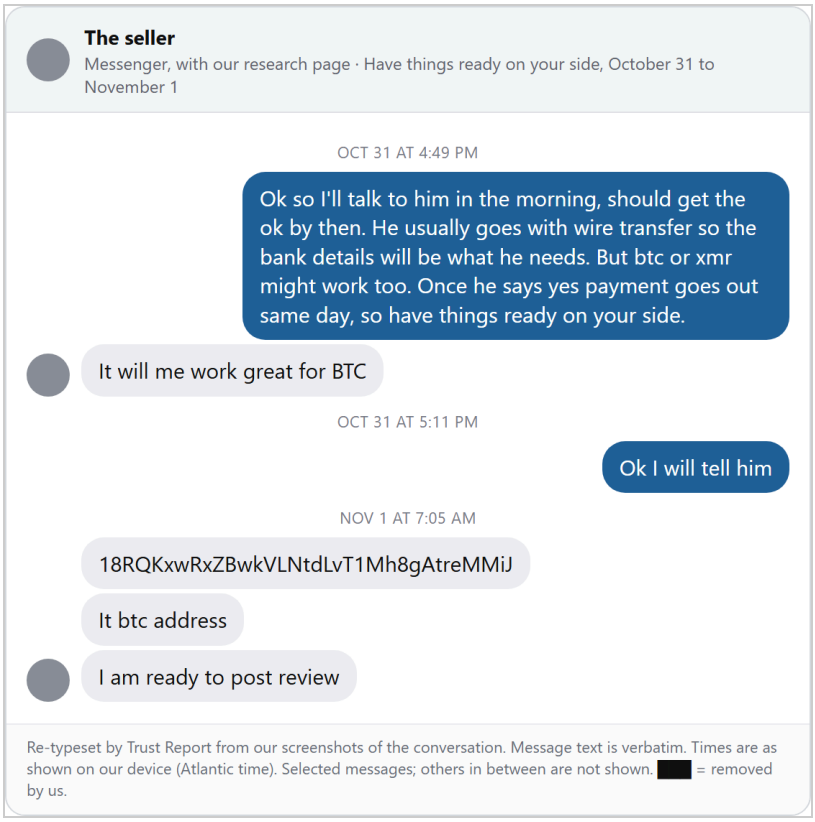


Figure 6. The implied requirement ("have things ready on your side") and the vendor's unprompted disclosure.

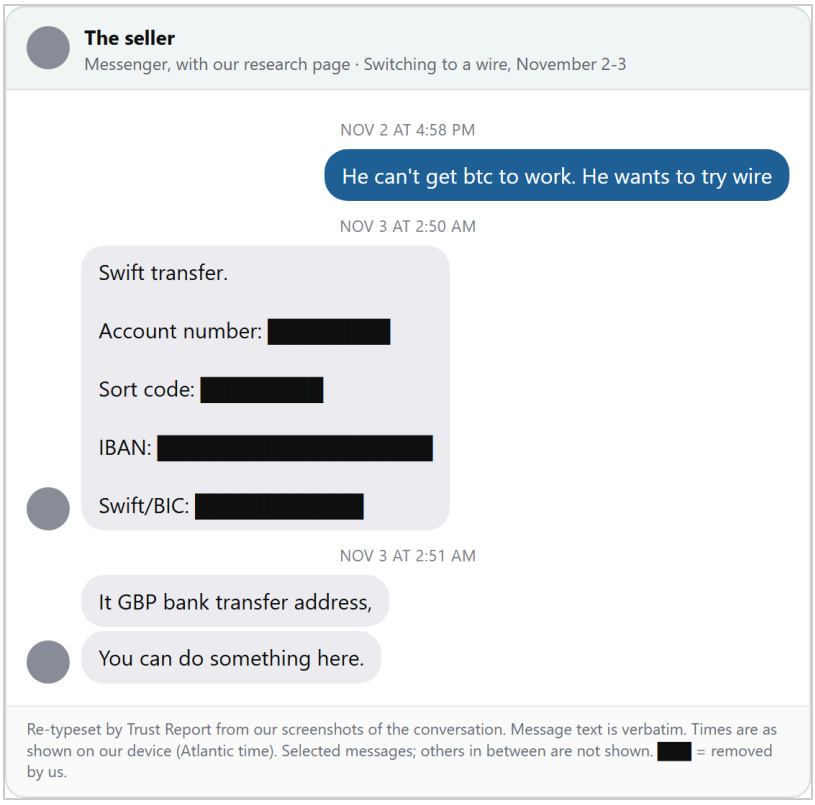


Figure 7. A one-line friction pretext moves the vendor to a regulated rail (details redacted).

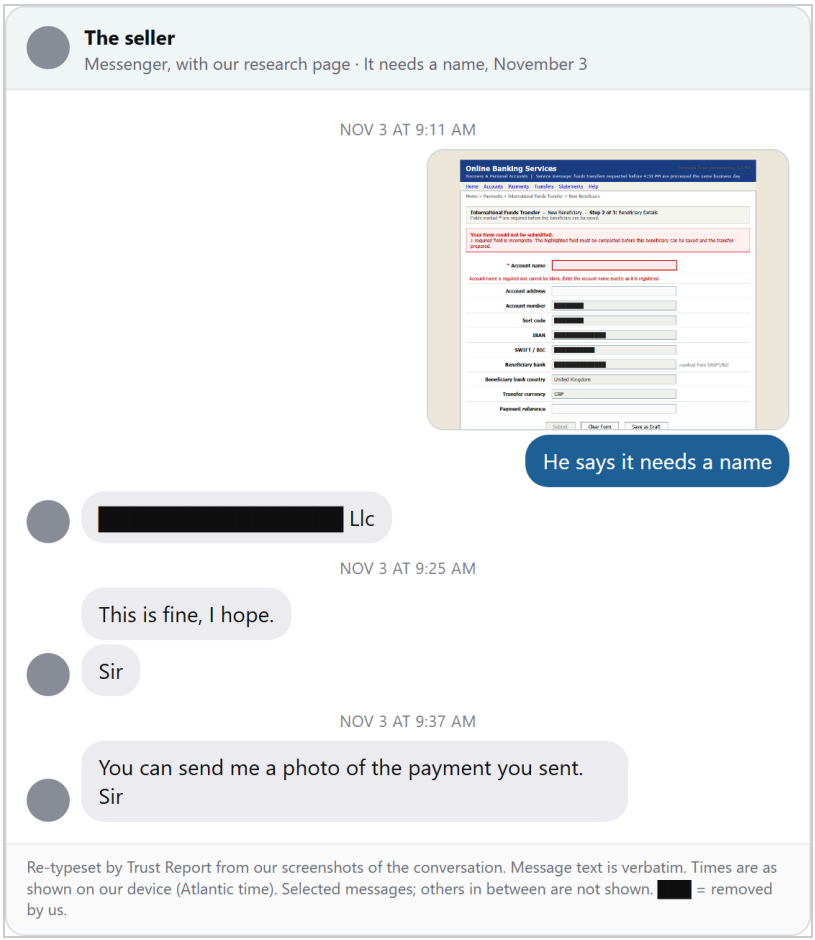


Figure 8. The screenshot pretext and the stated holder (redacted). The form shown is the published, redacted version; the vendor's copy carried his own details.

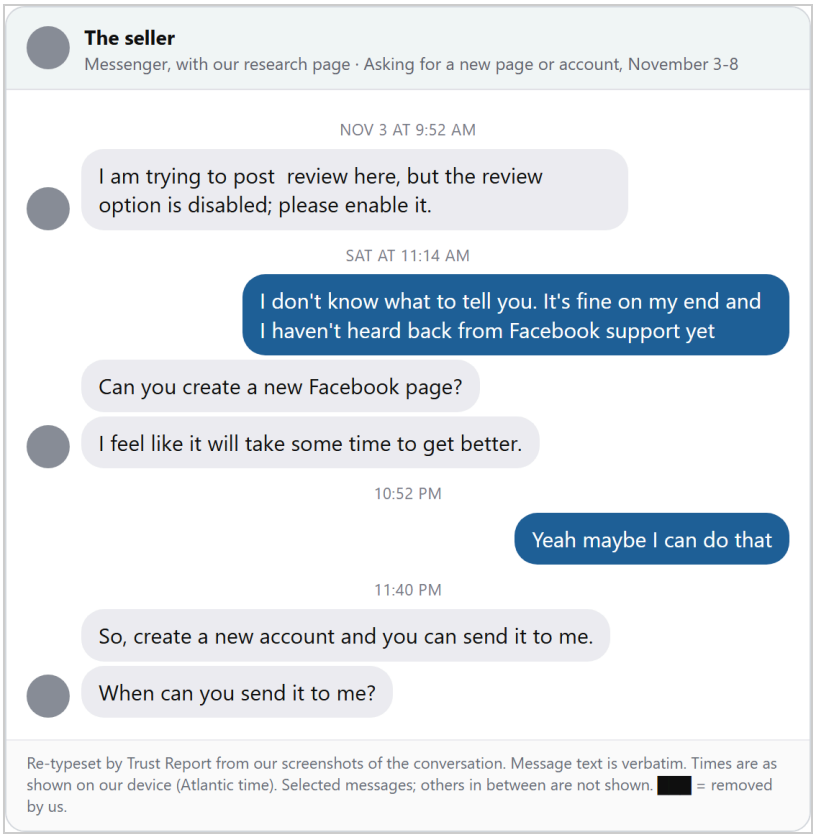


Figure 9. Escalation to asset acquisition. The operator's "Yeah maybe I can do that" kept the thread open; nothing was created or sent.

5. Vendor TTPs

Table 3. Vendor behaviour observed, with DISARM Red framework mapping where one fits.

ID	Behaviour observed	Evidence	DISARM
V1	Unsolicited, cross-platform menu sold to a page by direct message	Fig. 1	From the buyer's side, the sale is the recruitment of a contractor: T0091.001 Recruit Contractors (TA15)
V2	Reviews as the core product across five review platforms	Fig. 1	T0109 Consumer Review Networks (TA07)
V3	Free trial on a target of the buyer's choosing, with keyword intake and vendor-drafted copy	Fig. 2	None specific; a sales technique (section 6.1)
V4	Delivery through a separate persona account: London on every location field; seven recommendations (our page and six UK small businesses in unrelated trades and regions); generic service praise	Fig. 3; persona captures	Consistent with T0143.002 Fabricated Persona (TA16) on a T0146 Account Asset (TA15); fabrication not proven
V5	Engagement seeding: the review's only two reactions came from two other accounts, one stating a Bangladesh location	Reactor captures	No exact technique; supports the coordination element of CIB
V6	Volume pricing on request: USD 3, then USD 600 per 1,000 and USD 5,000 per 10,000; budget elicitation first	Fig. 4	Commercial; see 6.1
V7	Rail menu weighted to irreversible or low-friction rails (cash remittance, crypto); steering away from an accepted domestic e-transfer	Fig. 5	Commercial
V8	Bitcoin address with two years of prior receipts, all swept, tagged by a clustering service as part of a larger aggregation wallet	Fig. 6; enrichment	Infrastructure indicator
V9	UK e-money (GBP) receiving account; stated holder a company name with a US-style "Llc" suffix	Figs. 7–8	Infrastructure indicator
V10	Proof-of-payment demands and high-frequency follow-ups (at least 25 unanswered nudges such as "Sir", "?" and "Tell me please")	Figs. 5, 8	Pressure; see 6.1
V11	Blame shifted to the target's settings ("the review option is disabled; please enable it")	Fig. 9	Pressure
V12	Request that the buyer create a new Facebook page, then a new account, and send it to the vendor	Fig. 9	Asset acquisition: T0146 Account Asset with T0150.006 Purchased Asset or T0150.007 Rented Asset (TA15), sourced from the customer

V12 is the most important finding. Once the decoy target stopped accepting reviews, the vendor did not walk away from a five-figure order; he asked for a fresh asset. A page or account created by a real buyer and handed to a vendor carries the buyer's authentic signals (a real person's age, location, device and payment history) and costs the vendor nothing. The retail CIB economy does not just sell inauthentic engagement to its customers. It can harvest its customers for the accounts that produce it.

6. Social engineering on both sides

6.1 The vendor's techniques

Reciprocity and the foot in the door. The free trial is the vendor's investment in the relationship. It creates a felt obligation (Cialdini, 2021) and a small first commitment that makes a larger one easier (Freedman & Fraser, 1966). Asking the buyer to choose the target, approve keywords and approve draft text extends that commitment: the buyer co-authors the first deliverable.

Budget-first pricing. "Could you tell me your budget?" and "How many reviews would you like to order?" are the vendor's own elicitation: price discrimination by asking the buyer to anchor first.

Deference and persistence. Honorifics ("sir") paired with a steady cadence of nudges keep the vendor in a service role while sustaining pressure. Unanswered nudges never escalate to anger; they repeat.

Rail steering. The vendor accepted the buyer's suggested e-transfer ("It will work well for me") and in the same message steered back to Western Union, Remitly or crypto. The preference order favours rails with weak recourse for the payer.

Verification pressure and blame transfer. Proof-of-payment requests ask the buyer to perform commitment before delivery. When delivery failed, the failure was located in the buyer's settings, which set up the asset request.

6.2 The buyer's techniques

Table 4. Buyer-side techniques, their mechanism and their yield.

ID	Technique	Mechanism	Yield
B1	Decoy target	Contains the trial to an asset we control	Delivery method, persona account, seeding accounts, with no third-party harm
B2	Trial request	Low-commitment entry that a seller cannot refuse without losing the sale	Proof of capability before any money
B3	Confidence challenge with a volume anchor ("for when we need 10k")	Anchoring (Tversky & Kahneman, 1974); raises the stakes of the vendor's next answer	Vendor patience; a whole price curve
B4	Refusal of the implied total ("I'm not paying 3k and 30k")	The vendor's own unit price, multiplied out, becomes the anchor to beat	An 80 to 83% unit-price concession in one reply
B5	Absent principal ("let me speak to my manager on Monday")	The limited-authority gambit: delay without refusal; no one in the conversation can be pressured into deciding	Time
B6	Urgency inversion	The vendor's sunk cost in the trial (Arkes & Blumer, 1985) and the size of the order make waiting costly for him, not us	Follow-ups that volunteer information
B7	Implied requirement ("have things ready on your side")	Elicitation by stating a need instead of asking a question (Hadnagy, 2018); the vendor closes the loop himself	Bitcoin address, unprompted
B8	Friction pretext ("he can't get btc to work")	A plausible failure on our side moves the vendor to the next rail on his own menu	A KYC'd bank rail, with full details
B9	Screenshot pretext (section 7)	Need for closure (Kruglanski & Webster, 1996): one red blank between the vendor and a transaction he believes is real	The stated account holder, in 14 minutes
B10	Silent disengagement	No refusal or reveal that might prompt clean-up	A stable record for reporting

The engagement turned on who controlled time. Classic pressure tactics rely on urgency, and the vendor used them: a steady stream of nudges and demands for proof of payment. The buyer never matched that urgency. By placing the decision with an absent manager and the timing on "Monday", the buyer left the vendor holding both the sunk cost of the trial and the pending reward of a large order. Each disclosure followed a gap the vendor filled himself. In none of the key turns did the operator ask for the identifier that was disclosed.

6.3 Honesty in the pretext

The buyer persona was thin by design: a person with a manager, interested in volume, unfamiliar with prices. Each claim was either true (we did want to see what he would deliver) or a framing whose falsity harms no one but the transaction he was offering. The operator never claimed to be anyone real, and never impersonated an institution.

7. The elicitation artifact

Payment details identify where money goes, not whose account it is. The registered holder's name is the pivot that links an account to a person or entity, and every real transfer flow asks for it. We built a single static page in the style of a generic, dated online-banking "new beneficiary" form, prefilled with the details the vendor had already sent, with one field (Account name) blank, outlined in red and carrying an inline required-field error, and the Submit button disabled. The operator sent a screenshot with one line: "He says it needs a name."

Design constraints: a generic masthead with no real institution's name, logo or colours; no back end, links or data collection; sent only as an image; and it asks only for what the transaction would require anyway. The artifact succeeds because it is consistent with the vendor's model of the situation. It is not a request from the buyer; it is the bank's form refusing to proceed.

We have generalised it into **pretext-form**, an open-source, offline tool (Fig. 10). Presets cover UK, EU, US and Canadian beneficiary forms. Each field can be shown, flagged (required, blank, red) or locked (greyed out as verified), and all page text is editable. A capture mode hides the controls for a clean screenshot. The page's Content Security Policy blocks network connections and form submission. Source, rules of use and a live copy:

- Source: https://github.com/mikesploitsec/pretext-form
- Live: https://mikesploitsec.github.io/pretext-form/

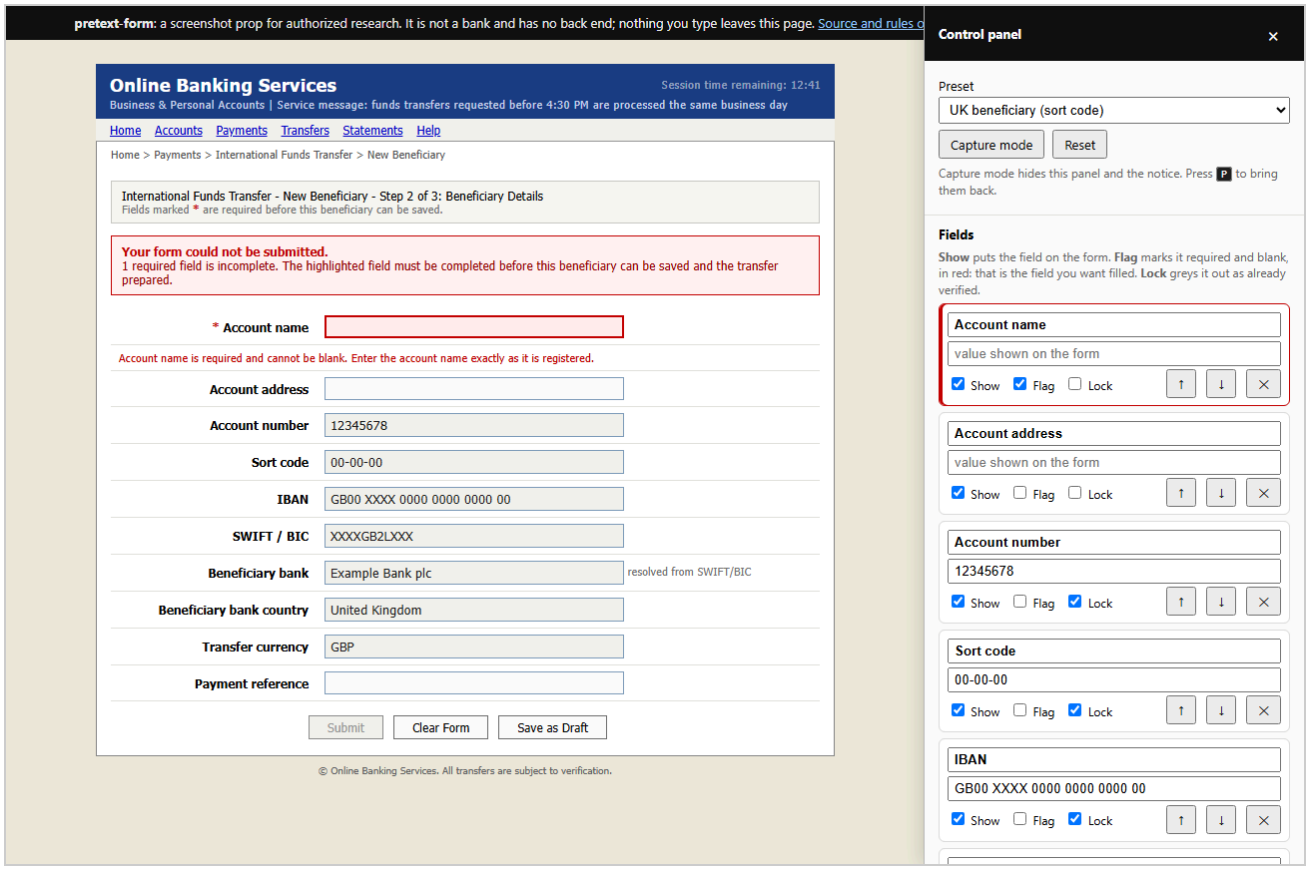


Figure 10. pretext-form with its control panel open: the UK preset, with the account name flagged. Placeholder values only.

8. Indicators

Delivery persona. Profile fields all set to one foreign city; a "Reviews given" list spanning unrelated trades in distant regions; generic, service-agnostic praise; reactions only from a small set of other accounts; follows that include a page tied to the vendor's stated country. None of these alone is conclusive; together they make a useful triage pattern for platforms and for business owners reading their own reviews.

Payment. A cryptocurrency address reused across years of receipts and swept into a larger cluster indicates that payments to the vendor are pooled. A per-user e-money receiving account in one jurisdiction, with a holder whose name carries another jurisdiction's entity suffix, is a mismatch worth an issuer's attention. Both indicators were reached only by entering the transaction.

Asset demand. A request to create and hand over pages or accounts is a strong behavioural indicator of a vendor sourcing assets, whatever product it sells.

9. Discussion

The retail edge is reachable. Influence operations are usually studied through the content they publish. Retail CIB vendors cannot sell without talking to customers, and a customer can be a researcher. The buyer's seat gives access to the parts of the operation that never appear in public: the sales process, the price curve, the payment rails and the asset pipeline.

Customers as supply. V12 suggests a loop. Vendors sell inauthentic engagement to small businesses, and when delivery fails they ask those businesses for accounts and pages, which become delivery assets for the next customer. Advice for small businesses usually says "don't buy reviews". It should also say "never create or hand over an account for someone else".

Competition harm. At USD 0.50 per review, a reputation that takes an honest business years to earn costs a competitor less than a month's advertising. The harm falls on consumers, who cannot tell the reviews apart, and on competitors, who are outranked. That is why the trade sits under competition law and not only platform policy.

Detection. Platforms hold signals the vendor cannot hide: locale consistency across a reviewer's recommendations, reaction graphs on new reviews, and messages that offer engagement for payment. Issuers hold the other half: receiving accounts that a counter-party describes as payment for reviews.

10. Ethics and limitations

Ethics. The subject initiated a commercial offer. The engagement stayed within what a real buyer of that offer would do, with three exceptions that each reduced harm: no payment was made, the only delivery target was our own page, and no account was handed over. Deception was limited to the buyer persona and a screenshot of a form. Third parties touched by the delivery (the persona, the reacting accounts and the businesses the persona reviewed) are not named here; some may be unaware of the activity. The full record went to the platform, the payment issuer and law enforcement.

Limitations. This is a single case (n = 1). The identity of the person or people behind the vendor account is unknown. Whether the delivery persona is fabricated or a compromised or rented real account

is not established. Times rely on the operator's device clock. The inference that some accounts were created in the same period, from Facebook's rising numeric IDs, is low confidence. The price curve is one vendor's offer to one buyer, not a market rate.

11. Conclusion

Playing the customer turned a routine unsolicited message into a view of the retail CIB economy: how it delivers, prices, collects and replenishes. The techniques that made it work are old ones: anchoring, deferral to an absent authority, implied requirements and a single blank on a form. The new observation is the vendor's turn from selling engagement to acquiring accounts from the buyer. We release the pretext artifact so others can run the same play under the same rules.

Data availability

The evidence set (screenshots, enrichment captures and a SHA-256 manifest) is held by the author and has been provided to the platform, the payment issuer and law enforcement. It is available to platforms, regulators and researchers on request. The public record of the vendor account is maintained by Trust Report Canada (2026).

References

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes, 35*(1), 124–140.

Cialdini, R. B. (2021). *Influence, new and expanded: The psychology of persuasion*. Harper Business.

DISARM Foundation. (n.d.). *DISARM frameworks*. https://github.com/DISARMFoundation/DISARMframeworks (accessed 9 November 2026).

Federal Trade Commission. (2024). Trade regulation rule on the use of consumer reviews and testimonials. 16 CFR Part 465.

Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology, 4*(2), 195–202.

Hadnagy, C. (2018). *Social engineering: The science of human hacking* (2nd ed.). Wiley.

He, S., Hollenbeck, B., & Proserpio, D. (2022). The market for fake reviews. *Marketing Science, 41*(5), 896–921.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing". *Psychological Review, 103*(2), 263–283.

Meta. (2018 onward). Coordinated inauthentic behavior reports. https://transparency.meta.com/metasecurity/threat-reporting

Trust Report Canada. (2026). Review seller (TRC-015) [wiki entry]. https://trustreport.ca/wiki/accounts/review-seller-trc-015

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.