ZTS ADVISORY · FIELD STUDY · JULY 2026

The Front Door Integration Test

What happens when a homeowner with money calls a PE-backed service brand

55 calls · 48 brands · about 20 sponsors · July 16–25, 2026 · US and Canada · nobody booked

ZAHA AL-HMOUD · ZTSADVISORY.COM/REPORT

What I did

I called the number each brand's marketing budget exists to make ring. I played the caller every one of these businesses says it wants: a real problem, money to spend, two or three companies on the list. Then I did the one thing that makes a phone call worth watching — I hesitated at the end, the way a shopper does.

Fifty-five calls, forty-eight brands, one scoring sheet.

The phones get answered. The selling stops there.

Reached a live person	49	of 55
Never reached anyone — every one at a brand paying for ads that week	6	of 55
Couldn't tell me what a visit costs	9	of 49
Reached someone who could only take a message	11	of 49
Tried to keep me when I hesitated	8	of 49
Asked for my email	13	of 49
Sent me anything in writing	1	of 55
Offered me a sister brand	0	of 55

The shared software arrived at most of these platforms. The shared habits did not. The caller is the only person in a position to notice.

Observed, July 16–25, 2026, US and Canada. Denominators move because a step only happens if the call gets that far. Nine couldn't quote at all. Another eleven could read the fee card but couldn't book, so they count inside the 40 who priced a visit. Confidence intervals: answered 78–96%, kept me 7–30%, email 15–41%, in writing 0.0–9.7%, sister brand 0–6.5%.

How to read this

Four labels, on every number that matters.

Observed — found in the 55 calls, always with a denominator. **Benchmark** — from a named outside source. **Interpretation** — my read, with the other read beside it. **Modelled** — arithmetic from assumptions you should replace with your own.

Nobody booked a job, so nothing here is a close rate, and I never claim a missed step cost anyone work. What I can show is narrower: whether the things that turn a call into a job happened when a caller set them up perfectly.

What this can’t tell you

It can’t tell you a close rate — nobody booked. It can’t tell you about any one brand — most got a single call. The brands weren’t drawn at random, so there’s no industry average in here.

Two zeros deserve different weight. Zero out of fifty-five means something: if brands really passed callers along even seven percent of the time, seeing it never in 55 tries would be unlucky. Zero out of five means very little on its own, so I treat it as a question worth asking and publish no rate.

July is the busy season, and a full schedule is a real reason to let a hesitant caller go. It isn’t a reason to skip asking for an email. I raised second problems on purpose, so those come up more often here than they would on their own. One person scored the calls: me.

Nobody was booked, no truck rolled, and no free work was taken. Brands and sponsors stay anonymous. Quotes are word for word.

The one I didn't publish

Eleven calls ended with someone promising to call me back. Two did, both the same day. The other nine landed after the study closed.

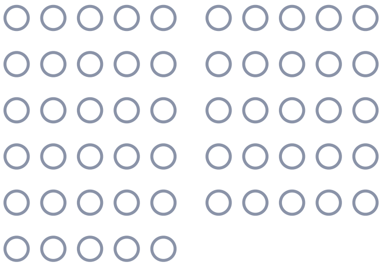
So I have no callback rate to give you, and I'd rather leave the gap visible than fill it with a number I can't stand behind.

If you can pull your own callback-kept rate out of your CRM this week, you can answer a question this study could not.

Observed. 11 promised, 2 arrived, 9 unobserved. No rate published.

Nobody passed me to a sister brand

0 of 55



Calls where anyone offered me a sister brand: none. Thirteen of the 55 had a second problem in the house, unprompted.

WORTH READING IF YOU CLOSED AN ADD-ON IN THE LAST YEAR.

I mentioned the water heater as well as the AC, the roof as well as the gutters. Three brands handled both problems on one trip. The rest quoted a second visit, a second fee, or told me to call somebody else.

> *“We only handle hot water tanks or tankless water heaters, but I don’t have any plumbers on staff for anything else. So you’d have to call a plumber for that, unfortunately.”*

The brand that said this belongs to a platform that sells plumbing under other names.

Cross-selling is about a fifth of the revenue synergy in a deal like this, and most buyers come in about a quarter short of what they promised. When it misses, it misses here, at the phone.

Some platforms keep their brands apart on purpose — licensing, service areas, brand equity. If that’s you, the zero is a decision, and the deck should say so the next time it claims cross-sell. Either way the caller hangs up thinking you sell one trade, so somebody should choose which of the two is true.

Worth an hour. Pull a week of calls from two sister brands in one city. Count how often a second need comes up, and how often anyone does something about it.

Observed, 0/55 (95% CI upper bound 6.5%); 3 of 13 one-trip (CI 5–54%). Benchmark: McKinsey revenue-synergy research [M1][M2]. Interpretation on the cause.

One platform, four different front doors

Four brands, forty-eight hours

WORTH READING IF YOU'RE INTEGRATING AN ADD-ON, OR YOU'VE JUST INHERITED THE PLATFORM.

Four sister brands under one sponsor, all heating and cooling, same story, inside two days. Every one quoted the same visit fee and offered the same member discount. After that they had nothing in common.

	BRAND 1	BRAND 2	BRAND 3	BRAND 4
Soonest visit	Tomorrow	Friday	Same evening	“Nothing after 4”
Second problem	Second fee	“Call a plumber”	Separate booking	Turned down
Membership	Full pitch	If asked	Offered	Never came up
Anything in writing	No	“No way to send it”	Texted during the call	“Check the website”
What they wrote down	Name, address, phone	Address	Phone	Nothing

> *“I don’t have any way to send you the information, but you can check it out on our website.”*

That was brand four, the same week brand three was texting me a summary mid-call.

The fee is a setting. The follow-up is a habit.

One call each, so take it as an existence proof: the gap inside one platform can run as wide as the gap across the whole industry.

Worth an hour. Call your own brands with the same two-problem story in the same week. The distance between your best and your worst is your integration list, already sorted.

Observed, n=4 brands, one call each. Interpretation on the cause.

The machines tried harder to keep me

MACHINES 6 of 13 · 46%

PEOPLE 2 of 33 · 6%

Offered me something when I hesitated: hold the slot, book a backup, a callback at a stated time. Fisher's exact $p = 0.004$.

WORTH READING IF YOU HAVE A VOICE AGENT RUNNING AND NO NUMBER ATTACHED TO IT.

The gap holds up statistically, which is unusual at this sample size. How big the gap is, I can't say.

Three of the 49 answered calls opened with an agent and passed me to a person. They sit in neither group, which is why the two add to 46 rather than 49.

> *"Would it help to just go ahead and get us on the schedule as a free backup comparison?"*

An after-hours agent at a roofing brand, right after I said I needed to talk to my husband.

Trying is all I could see. Nobody booked on any call in this study, so no close was observed from anyone, human or machine.

The two human saves were the best selling I heard anywhere. One rep held that evening's slot with a text-to-cancel and no commitment. Another waived the visit fee without being asked. Both are one sentence long. Neither showed up again in 55 calls.

The fair objection: agents mostly answer after hours, so I wasn't comparing like with like. It still doesn't explain why the daytime team almost never says the twenty words the platform already wrote down for its machine.

Worth an hour. Pull ten evening calls and ten daytime ones and count who offers something. Then find that line in your agent’s script and ask when a person last said it.

Observed. 6/13 versus 2/33; 7.6× rate ratio; Fisher’s exact $p = 0.004$; odds ratio 13.3. Intervals stay wide — agents 19–75%, people 1–20%. Direction well supported, size uncertain.

The $8,500 caller got the service-call process

WORTH READING IF THE AD SPEND CLEARS EVERY MONTH AND BOOKED JOBS STAY FLAT.

None of the five. Twice I offered to email the quote over; nobody asked to see it, or what was quoted, or by whom, or when I was deciding. The soonest anyone could come out was two to ten days.

Five calls can't give you a rate: the true figure could run as high as half. Here is the part I would bet on anyway. The same queue, the same questions and the same calendar handled an $8,500 job and a $432 one, and nobody in the chain had a reason to tell them apart. That showed up on all five, and it is the cheapest thing in this report to check.

> *"I was hoping to get an answer before Monday because I'm wanting to make up my mind this weekend about the purchase. So, I guess we're out of luck here."*

That's me, saying my deadline out loud, to an after-hours agent.

The same brands quoted me around $432 for a service call. So the job worth fifteen to twenty times more went through the same queue, the same questions, and the same calendar as a $432 service call.

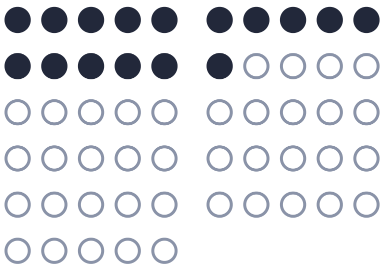
Answer inside the hour and the odds of qualifying the lead go up about sevenfold. That research was done on web leads, so the multiplier doesn't transfer — but the shape does, and two to ten days sits at the wrong end of every version of that curve.

Worth an hour. Pull ninety days of replacement calls, measure the hours from call to visit, and count how many have the competitor's quote saved on the record.

Observed, 0/5 (CI 0–52%) — suggestive only, stated as a question. Ticket sizes observed on these calls. Benchmark: Oldroyd/InsideSales and the 2011 HBR audit of 2,241 firms [O1][O2].

A machine repeats its mistakes exactly

16 of 55



Calls answered by a scripted agent. Thirteen of the 16 ran start to finish. That is 29% of the sample.

WORTH READING IF THE AI LINE ITEM IS APPROVED AND NOTHING MEASURES IT.

Among those sixteen: one required an email before it would schedule, then took one I announced out loud was fake. One confirmed a town about 150 miles from the one I'd said, and carried on. One walked its own empty calendar forward into September while a replacement buyer waited, and never fetched a person. One greeting promised emergency service the voice behind it could neither price nor book.

> *"Sure, I will just give you, like, a fake email. Marcus 1234 at gmail dot com." "Thanks, Marcus. And just to double check, about how old is your roof?"*

The booking flow carried on exactly as designed.

Not one platform could have told me what its agent books, on the same definition it uses for people. The spending is approved. The number that would say whether it worked was never attached to it.

The same agents did the best job of keeping me on the line, back in the last finding. Both facts come from the same place: a setting.

Worth an hour. Ask each company what its agent books, on the human definition, and don't accept a different metric. The pause is the answer. Then find out who can change the script — if only the vendor can, you're renting the channel.

Four things worth writing down, borrowed from the NIST framework: what the agent is allowed to promise, every flow it runs, the booking number it reports, and a handoff to a person after two failed tries.

Observed within n=16, by absence for the measurement. Benchmark used as governance only: NIST AI Risk Management Framework and its Generative AI Profile [N1][N2].

What one of these might be worth

Take the replacement calls. This prices only the extra jobs better handling would win. Bad handling doesn't take a close rate to zero, so pricing the whole pipeline would be theatre.

Eligible calls × average value × conversion lift × margin – what it costs to fix

Calls a year	50,000	Your offered-call count
Share holding a quote	5%	Assumption — use your own call reasons
Average value	$8,500	Observed on these calls
Lift	3 / 7 / 12 points	Assumption — a pilot answers this in 60 days
Margin	25 / 27.5 / 30%	Assumption — below typical trade margins
Cost to fix, year one	$40K / $75K / $120K	Assumption — no software to buy

	REVENUE	EBITDA	AFTER COST
Low	$637K	$159K	~$40–119K
Middle	**$1.49M**	**$409K**	**~$290–335K**
High	$2.55M	$765K	~$645–725K

Don't add the leaks together. The same caller sits in two of them at once, so a total double-counts. Rank them, fix one, and let a measured number replace a modelled one.

Modelled throughout. Four of the six inputs are assumptions, marked above.

Six fixes, and how you'd know they worked

Every one of these runs inside software you already own.

FIX	WHAT IT CLOSES	SCALE IT WHEN
A written summary at the end of the call	1 of 55 sent anything	The summary group recovers 3+ more points of unbooked calls over 8 weeks
The save line, said in daylight	8 of 49 tried to keep me	The trained group beats the other by 15 points, cancellations flat
A separate lane for quote-in-hand callers	0 of 5 asked to see the quote	Those jobs book better and your estimators aren't drowning
Working the callbacks and the dropped calls	11 promised, 2 arrived	Kept rate holds above 80% for a month
A way to hand a caller to a sister brand	0 of 55 offered one	Referred jobs clear 5% of eligible calls over 8 weeks
Somebody owning the agent's settings	No platform knew what its agent books	The agent books more without flooding your people

Two guardrails on all of it. Watch cancellations and completed jobs, because a booking that never happens is worse than no booking. And settle who gets credit for a cross-brand job before the first referral — unpaid credit kills this faster than any technical problem.

In order: thirty days measuring, thirty days piloting two things, thirty days scaling what worked.

Seven questions and three reports

An hour of listening, and three pulls.

	LISTEN TO FIVE CALLS PER BRAND	HERE
1	Does anyone ask for an email address?	13 of 49
2	When the caller hesitates, what's the next sentence?	8 of 49 offered something
3	Can whoever answers say what a visit costs?	9 of 49 couldn't; 11 more could only take a message
4	Does a second problem get one trip, or a second bill?	3 of 13 got one trip
5	Does anything arrive in writing?	1 of 55
6	If the caller has a competitor's quote, does anyone ask to see it?	0 of 5
7	Call your own after-hours line tonight. Can it quote, book, or promise a time?	16 of 55 calls were answered by a machine

	PULL	HOW
8	Calls to booked jobs, by brand	Jobs booked within 48 hours ÷ qualified calls, one definition everywhere. If your brands define it differently, that's the finding.
9	Speed to answer, and real abandonment	Hang-ups past your threshold ÷ calls offered. Speed measured on answered calls only.
10	Membership attach	Sold ÷ eligible, with eligibility written down and existing members left out.

If the first seven come back clean, close this and keep the hour. Of the 48 brands I called, none would have.

What belongs in the board pack

One page, next to what you spend on marketing. If your brands define a line differently today, agreeing the definition is the first job.

	MEASURE	NOTES
1	Qualified calls ÷ all inbound	Leaves out internal, vendor, spam
2	Answered ÷ offered	Quick hang-ups excluded
3	Real abandonment	Past your threshold only
4	**Booked jobs ÷ qualified calls**	**48 hours. One definition, every brand.**
5	Cancelled before dispatch ÷ booked	14 days
6	Completed ÷ booked	30 days
7	Hours from quote-in-hand call to visit	Median
8	Competitor quotes saved ÷ quotes mentioned	—
9	Booked within 14 days ÷ unbooked calls	—
10	Sister-brand referrals done ÷ second needs heard	Sampled until you have a call reason for it
11	Memberships sold ÷ eligible	Eligibility in writing
12	The same lines, machine against person	Same definitions both sides
13	The same lines, by brand and by hour of day	Business hours against after hours

The bottom line

Across 55 calls to 48 PE-backed brands, the things that turn an answered phone into a booked job — keeping the caller, writing down who they are, sending something, calling back, passing them along, noticing a big job when it arrives — happened rarely, unevenly, or never. The software mostly worked. Where a machine was answering, it followed its instructions more faithfully than the people did.

None of that is a verdict on anyone's business. I called from the outside and I only saw the front door.

The next move is small and it fits in this quarter: one definition of calls-to-booked-jobs across every brand, the seven questions above, and one pilot with a number attached to it.

The front door is the integration test. Most of the platforms I called would not pass it today, and the person who finds that out is the customer you paid to reach.

IF THE SEVEN QUESTIONS COME BACK THE WAY THEY DID HERE

A two-week read on your own numbers

I run the three pulls on your ServiceTitan and phone data, listen to a sample of your own recorded calls, and hand back this same arithmetic done on your numbers instead of mine.

It ends with a short list of what to fix first, in the order I would fix it. Your team runs all of it, and there is no software to buy.

TWO WEEKS · FIXED FEE · NOTHING TO INSTALL

If you want it, reply with the name of your busiest brand. I will tell you inside a day whether your own data can answer the seven questions. If it cannot, that is worth knowing on its own, and it costs you nothing to find out.

ZAHA AL-HMOUD · ZAHA@ZTSADVISORY.COM · ZTSADVISORY.COM