Time to resolution is -6.8 hours

Time to resolution changed by -6.8 hours in the rollout group relative to the comparison group (95% range -7.1 to -6.5 hours). The rollout group started at 18.1 hours.

- Time to resolution improved, and the 95% range does not include zero. The comparison group is not
- a candidate for this rollout.

These are sample numbers. Replace them with an admin export or CSV.

Time to resolution

Change	**-6.8 hours**
95% range	**-7.1 to -6.5 hours**

Change in the rollout group minus change in the comparison group. Each week counts equally. Inside a week, each team counts by the number of observations on that row. The range is a 95% interval from week-to-week variation.

The other metrics

The agreed metric is the one in the reading above. These are the other metrics on file.

- CSAT: +0.01, range -0.04 to +0.06, range includes no change
- Reopen rate: 0 pp, range -0.4 pp to +0.4 pp, range includes no change
- Tickets resolved: +3.6 tickets per person per week, range +3.3 to +3.9 tickets per person per week,
- agreed direction

What it cost

Loaded pay is an assumption you can edit: $1,900 per person per week. Cost per resolved ticket uses the rollout group's before and after. It is not adjusted by the comparison group. The agent's own cost, when shown, leaves human pay out and is not comparable to the other rows.

Seat spend	**$0.00**
Token and API spend	**$0.00**
Agent spend	**$14,000**
Other spend	**$0.00**
In this window	**$14,000**

- People, before: $94.24 per resolved ticket. Loaded pay in the rollout group, divided by weekly volume
- before the rollout.
- People plus the tool, after: $78.88 per resolved ticket. Loaded pay plus seat, token, and other tool
- spend, divided by weekly volume after.
- The agent's own work: $3.98 per resolved ticket. Agent spend divided by the units the agent
- completed. This is not a fully loaded human cost.

Adoption

Who used it, at the team. This explains the rollout. It is not a second result.

100% of job titles were recognized without an edit.

Baseline response 71% of 78 people invited.

Reading

The primary metric moved in the agreed direction. The 95% range does not include zero. The comparison group is not a candidate for this rollout.

Who was in it

- Billing: First wave, 24 people
- Orders: First wave, 20 people
- Trust: Hold back, 16 people
- Logistics: Hold back, 18 people

How this was calculated

The number is the first wave's change minus the held-back group's change over the same weeks. Each week counts equally. Inside a week, each team counts by the number of observations on that row. A row under the privacy floor is left out, and a week is left out when the included observations add up to less than the floor. The range is how much those weekly results bounced. It is not proof the tool caused the change.

The reading stays a description when the metric or the groups were changed after they were confirmed, when everyone got it at once, when the range still includes both no change and an improvement that would matter, or when the groups are too small to see that improvement. Groups under 5 are hidden before any of this runs.

The smallest change these groups can reliably detect is about 0.43 hours. You said 1.8 hours would matter, and this study can detect a change of that size.

The calculations workbook repeats every row. On its Estimate sheet, column B is a formula and column C is the number in this report. They should match.