Akholi

PART 1 / 2026 RESEARCH STUDY

Engagement Outcomes and Their Correlates

Agentic AI Outsourcing at Tier 1 Financial Institutions

Phil Hatch

150 engagements · 850 delivery-team respondents

Observational, cross-sectional study of a nonrepresentative sample. Financial institutions did not independently validate the reported experience. The findings do not establish causation or legal breach.

Version 1.8.2 | November 1, 2026

REPORT INFORMATION

Agentic AI Outsourcing at Tier 1 Financial Institutions

Full Study Report

Author	Phil Hatch
Publication date	November 1, 2026
Version	1.8.2
ISBN	979-8-234-26882-2
SSRN DOI	https://doi.org/10.2139/ssrn.7164559
Figshare DOI	https://doi.org/10.6084/m9.figshare.34036890
Contact	info@akholi.com
Website	akholi.com

AUTHOR ONLINE

Akholi author page | ORCID | LinkedIn | ResearchGate

SUGGESTED CITATION

Hatch, Phil. 2026. *Agentic AI Outsourcing at Tier 1 Financial Institutions: Full Study Report*. Version 1.8.2. Akholi. https://doi.org/10.2139/ssrn.7164559.

COPYRIGHT

© 2026 Phil Hatch and Akholi. All rights reserved.

No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the copyright holders, except for brief quotations with full attribution in reviews, research, and other noncommercial uses permitted by copyright law. Requests for permission should be sent to info@akholi.com.

USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence tools were used to prepare this report. Appendix G discloses how they were used.

Contents

Engagement Outcomes and Their Correlates **5**
The Immaturity of Agentic AI Outsourcing **7**
1 Lifecycle Position of the Engagements 9
2 Depth of Agentic Operation and Vendor Experience 19

Cancellation and the Risk of Cancellation **25**
1 Realized Cancellation 27
2 Perceived Cancellation Likelihood 29

Customer Satisfaction, Delivery Performance, and Measured Return **33**
1 Recorded and Perceived Customer Satisfaction 35
2 Delivery Against Dates, Budget, and Key Performance Indicators 39
3 Measured Return in Production 49

The Engagement Indicator Mean and the Interdependence of the Four Measures **51**
1 The Engagement Indicator Mean as a Summary Measure 53
2 Shared Structure Across the Four Study Measures 55
3 Indicators Net of the Engagement Indicator Mean 59

Structural Attributes and Outcome Associations **61**
1 The Financial Institution 65
2 The Relationship Between the Two Firms 77
3 The Work Outsourced 83
4 Commercial Terms 91
5 The Vendor 97
6 Size of the Delivery Team 105

About the Author **113**
About Akholi **115**

“

We are four months into this engagement and customer satisfaction has collapsed. My team corrects client data almost every cycle before it can be used, legal research inputs arrive incomplete and inconsistent, and by the time it reaches the model it is not fit to work from.”

ENGAGEMENT MANAGER

KPO, Regulatory, audit, tax and legal

Engagement Outcomes and Their Correlates

Cancellation, lower reported customer satisfaction, less favorable indicator ratings, and greater intent to leave occurred together across the engagements studied. These relationships describe substantial delivery difficulties, but they do not identify independent causes. Structural attributes provided mixed evidence: some estimated relationships fell within the study's prespecified small-association bounds, others were supported, and others remained inconclusive. The analysis examines these findings alongside engagement maturity and commercial arrangements, distinguishing full-sample cancellation contrasts from estimates restricted to active engagements.

The four core measures were strongly correlated. The engagement indicator mean summarizes respondents' ratings across the 100 engagement indicators, averaged within each engagement. Indicator quintiles are five equal groups ranked by that mean. However, each describes a different aspect of the engagement. The section “The Engagement Indicator Mean and the Interdependence of the Four Measures” examines their shared variance, the pattern across indicator quintiles, and the limits of attributing outcomes to individual indicators.

“

Half of these questions assume there was a defined vendor master to work against. There wasn't. The documentation exists, i will give them that, which is the only reason process readiness got a 3 from me, but we never got in front of anyone who could say which vendor record was the authoritative one. Asking how quickly they govern AI risk when nobody has been named to govern it is not a question anyone here can answer.”

ENGAGEMENT MANAGER

CXM, Corporate and back-office functions

The Immaturity of Agentic AI Outsourcing

Agentic AI outsourcing is a new form of delivery, and the engagements in the study were young. The mean is the arithmetic average, and n identifies the number of observations included in a result. Their mean age at the time of data collection was 5.69 months, and the oldest had run for 22 months (n = 150 engagements). Production had been reached by 30 of the 150 engagements, and the remaining 120 were in Build-Out or Transition/Cut-Over.

The vendors were new to agentic AI outsourcing as well. Every respondent who answered placed the vendor's experience of delivering agentic AI outsourcing at two years or less, with a mean of 1.0 year (n = 571 respondents). No respondent reported more than one year of personal experience with agentic AI (n = 793 respondents). Engagement managers who had run earlier outsourcing programs for the same vendor described what that inexperience meant on the engagement.

> *"The last program I ran for this employer was a contact center transition, and we were good at that, because we had done forty of them and there was a playbook that actually worked. Here there was no playbook, and the people who wrote the proposal had not run one of these either…"*
>
> **ENGAGEMENT MANAGER**
>
> CXM, Client servicing and communications

> *…we were sold into this as if the firm had done it before. It had not. There was no runbook for taking a screening agent from pilot into governed production, so we wrote the runbook ourselves at night while the alert backlog grew."*
>
> **ENGAGEMENT MANAGER**
>
> ITO, Financial crime and client onboarding

The associations of engagement age, time in the current stage, and production-agent count with the four core measures met the study's equivalence criteria. Equivalence testing assesses whether a relationship lies within prespecified bounds treated as small under the study's criterion. It does not establish a zero relationship. Vendor experience also met those criteria, although its restricted range limits interpretation. These singular cross-sectional results do not demonstrate whether individual engagements improved over time.

Limited experience provides context for the findings. Written accounts described vendors developing delivery practices during paid engagements. Respondents also rated the agentic AI knowledge and skills of financial institution staff poorly: 418 of 793 ratings (52.71%) were

distressed. These accounts identify perceived capability gaps without establishing their causal contribution to outcomes.

> *"I have been in the industry a long time. The last six months have been rough. … We are creating a new reality every day. AI behaviour changes without explanation, data feeds are not consistent. Our team has to make updates to data, the process, and the AI rules every day."*
>
> **PORTFOLIO DIRECTOR**
>
> BPO, Payments, cards and disputes

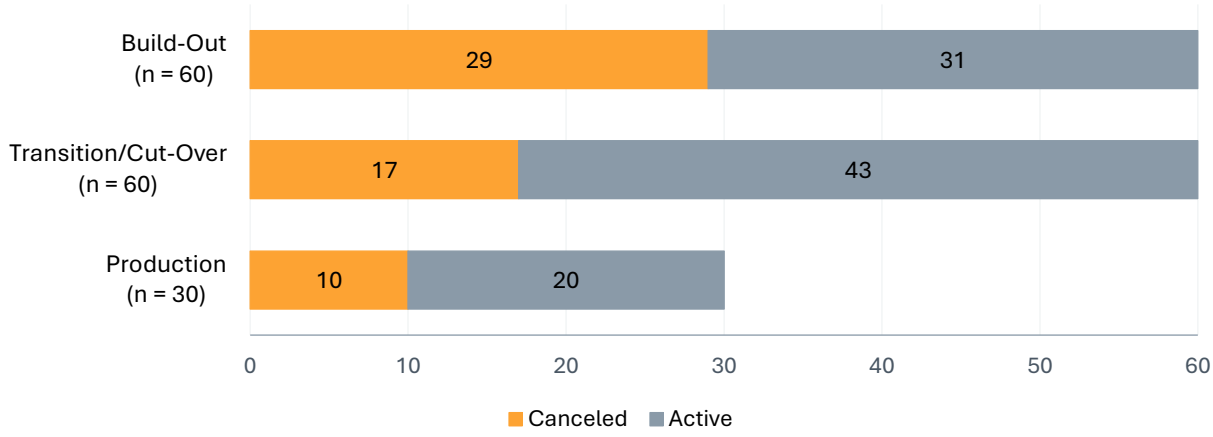
> *Eight months in, this is better than it was. When we started, nobody had written down which of the feeds into the repository actually mattered, and we found out the hard way when one of them changed its date format over a weekend. Since then we've mapped the ones that touch classification and we put explicit checks where the old team used to just eyeball it and fix it themselves… that part I'd call solved. What hasn't moved is warning. We still find out that a provider changed something because the queue looks wrong on Monday morning, not because anyone told us. And the model retirement notices go to our platform group, not to us, so we get maybe a week. For a POC that's survivable. If this goes to production on the same footing I'd be nervous about it."*
>
> **ENGAGEMENT MANAGER**
>
> BPO, Corporate and back-office functions

1 Lifecycle Position of the Engagements

Cancellation occurred at every measured lifecycle stage. The sample included 60 engagements in Build-Out, 60 in Transition/Cut-Over, and 30 in Production. The observed rate differences must be distinguished from statistically supported differences and estimates meeting the study's equivalence criteria.

Cancellation was recorded for 29 of 60 Build-Out engagements (48.33%), 17 of 60 Transition/Cut-Over engagements (28.33%), and 10 of 30 Production engagements (33.33%). The chi-square test compares cancellation frequencies across stages. The stage difference did not reach statistical significance (chi-square 5.39, p = 0.068, n = 150). The p value is the probability, under the null model, of a result at least as extreme as the observed result. Among active engagements, the stage contrasts for customer satisfaction, engagement indicator mean, and intent to leave met the study's equivalence criteria. These results do not establish that the stage is irrelevant outside the observed sample.

Figure 1.1. Engagement counts by cancellation status and lifecycle stage



Note: The groups are ordered categories. The connecting lines do not represent elapsed time.

Canceled engagements: 29 of 60 in Build-Out, 17 of 60 in Transition/Cut-Over, and 10 of 30 in Production. The difference across stages did not reach significance (chi-square 5.39, p = 0.068, n = 150 engagements).

Counts are engagements. Cancellation rates were 48.33% in Build-Out, 28.33% in Transition/Cut-Over, and 33.33% in Production. Overall, 56 of 150 engagements were canceled (37.33%). The lifecycle stages are categories, not elapsed-time observations.

Engagements were canceled early as well as late. Respondents described an engagement that ended within weeks of starting.

> *"We were canceled at eight weeks, which means most of what this section asks about was never actually tested..."*
>
> **ENGAGEMENT MANAGER**
> CXM, Client servicing and communications

> *"On the Tuesday we agreed how vendor queries should route between their team and mine, by the Friday the client had pulled the work."*
>
> **ENGAGEMENT MANAGER**
> CXM, Corporate and back-office functions

1.1 Lifecycle Stage and Outcome Comparisons

Indicator C036 · What stage is the engagement in? · Full statistical analysis ›

Cancellation was the most frequent respondent-reported lifecycle answer, although cancellation is recorded separately from the engagement's measured stage. The most frequent answer was Engagement Canceled, given by 309 of 810 respondents (38.15%). Where both a respondent's answer and the Engagements sheet recorded a stage, the two agreed in 501 of 501 cases.

Reported stage was associated with the engagement indicator mean (Welch F 40.50, $p < 0.001$, $n = 141$ engagements) and intent to leave (Welch F 71.51, $p < 0.001$, $n = 141$). The canceled category had a mean indicator score of 2.6 and a mean intent to leave of 9.3 ($n = 56$ engagements). Among active engagements, neither stage comparison reached statistical significance, and the customer-satisfaction contrast met the equivalence criterion. The full-sample result reflects the inclusion of cancellation as a response category.

The written responses that addressed the stage numbered 162, from 139 respondents on 69 engagements. 139 came from canceled engagements. Neutral scores were given to 124 of the 162 responses (76.54%), and the neutral responses recorded what the respondent had valued in an engagement that had ended. Among 61 responses on what respondents loved most about the job, 57 referred to the end of the engagement, each from a canceled engagement. An individual contributor named the end of the engagement itself as what they loved most about the job.

> *...the fact that it is finally over"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Technology engineering and operations

> *to be candid i felt relieved when they told us it is stopping. before this i was on a card disputes queue, small account, but the process was written down, one way to do the thing, and if you follow it the case closes. here it was never like that."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Client servicing and communications

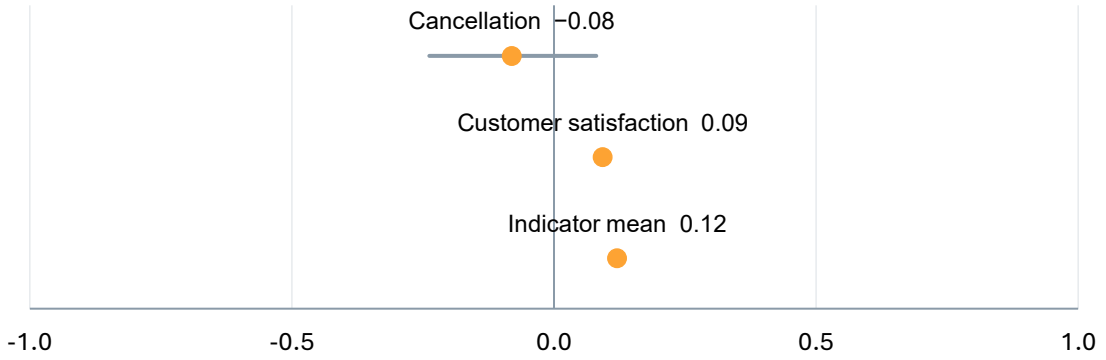
1.2 Engagement Age and the Four Study Measures

Indicator C035 · How long ago did this specific engagement start, in months? · Full statistical analysis ›

The engagements were generally recent. The mean age was 5.69 months and the median 5.0 months, from 1.0 to 22.0 months (n = 150 engagements). The median is the middle value in an ordered distribution. Every respondent's answer on the age of the engagement matched in the study.

Figure 1.2. Associations of engagement age with cancellation, customer satisfaction, and the engagement indicator mean

Associations of engagement age with cancellation, customer satisfaction, and the engagement indicator mean



Comparison	Estimate and analysis base
Cancellation	r = −0.08 [−0.24, 0.08] 150 engagements p = 0.327
Customer satisfaction	r = 0.09 94 active engagements p = 0.374
Indicator mean	r = 0.12 150 engagements p = 0.142

Note: Mean engagement age in months: 5.6 (n = 804 respondents).

Group means and numeric-answer bases: Individual contributor (n = 621): 5.52. Engagement manager (n = 150): 5.69. Portfolio director (n = 31): 5.97. Build-Out (n = 161): 2.70. Transition/Cut-Over (n = 241): 5.68. Production (n = 98): 11.40. Canceled (n = 304): 5.17.

r is Pearson's correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational. A confidence interval expresses uncertainty under the stated method. A p value is the probability, under the null model, of a result at least as extreme as the observed result.

The respondent-level estimate for intent to leave is reported in the accompanying analysis.

The estimated associations of engagement age with all four measures met the equivalence criterion. Correlation, denoted r, expresses the direction and strength of a linear association. A confidence interval (CI) expresses uncertainty around an estimate under the stated method. The correlation with cancellation was −0.08 (95% CI −0.24 to 0.08, p = 0.327, n = 150 engagements). The correlation with customer satisfaction among active engagements was 0.09 (p = 0.374, n = 94 engagements). It was 0.12 with the engagement indicator mean (p = 0.142, n = 150 engagements). At respondent level, the correlation with intent to leave was −0.08 (p = 0.360, n = 798 respondents). Mean age ranged from 2.7 months in active Build-Out engagements (n = 31) to 11.5 months in active Production engagements (n = 20). Age varied with stage without a detected outcome association under these tests.

The written responses that addressed age numbered 73, from 69 respondents on 48 engagements. In 15 of the 73, the engagement was too young for the survey's questions to be answered as asked. In 14, all from canceled engagements, it had ended within weeks or months. Seven said a problem had persisted for six months or more without improving.

> *"We are two month in and there is no platform yet, so what am i rating… i send a request for a schema and it sit for eleven days, then the reply is ask someone else."*
>
> **INDIVIDUAL CONTRIBUTOR**
> ITO, Technology engineering and operations

> *"Hard to answer these when engagement lasted four weeks. There was no joint model to speak of, only two kickoff calls and a governance pack nobody signed."*
>
> **ENGAGEMENT MANAGER**
> KPO, Corporate and back-office functions

> *"I feel these questions are asking about things which did not happen yet for us. We are two months in, nothing is in Production, so how do I answer for running many agents in live."*
>
> **INDIVIDUAL CONTRIBUTOR**
> KPO, Corporate and back-office functions

> *"The scores look good because we are only one month in and the volume is small."*
>
> **INDIVIDUAL CONTRIBUTOR**
> KPO, Credit, underwriting and suitability

> *19 months and nobody from their side came to sit with the queue for one hour."*

INDIVIDUAL CONTRIBUTOR

CXM, Client servicing and communications

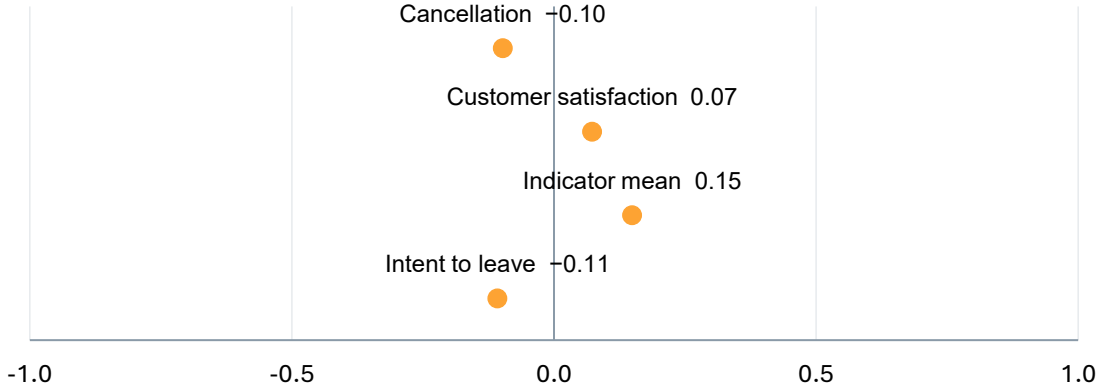
1.3 Time in the Current Stage and Outcome Associations

Indicator C037 · How many months has the engagement been in its current stage? · Full statistical analysis ›

Reported time in the current stage averaged 3.1 months, with a median of 2.0 (n = 794 respondents). Its estimated associations met the equivalence criterion for cancellation (r = −0.10, p = 0.233, n = 150 engagements), active-engagement customer satisfaction (r = 0.07, p = 0.488, n = 94), engagement indicator mean (r = 0.15, p = 0.069, n = 150), and intent to leave (r = −0.11, p = 0.166, n = 745 respondents).

Figure 1.3. Time in the Current Stage and Outcome Associations: associations with the study measures

Time in the current stage and outcome associations



Comparison	Estimate and analysis base
Cancellation	r = −0.10 150 engagements p = 0.233
Customer satisfaction	r = 0.07 94 active engagements p = 0.488
Indicator mean	r = 0.15 150 engagements p = 0.069
Intent to leave	r = −0.11 745 respondents p = 0.166

Note: Mean months in the current stage: 3.1 (n = 794 respondents).

Group means and numeric-answer bases: Individual contributor (n = 614): 3.08. Engagement manager (n = 150): 3.18. Portfolio director (n = 28): 3.04. Build-Out (n = 167): 1.68. Transition/Cut-Over (n = 228): 3.12. Production (n = 98): 6.48. Canceled (n = 301): 2.81.

r is Pearson's correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The time in stage varied with the lifecycle stage, as the age of the engagement did. It ranged from 1.7 months among respondents on active engagements in Build-Out (n = 167 respondents) to 6.5 months among those in Production (n = 98 respondents). Three written responses addressed it, each describing a stalled engagement or a request that went months without an adequate answer.

> *"We asked in March for role catalogue, we asked again in May, and what came back was a spreadsheet exported from an old attestation cycle, nine hundred groups, no description against most of them."*
>
> **ENGAGEMENT MANAGER**
>
> ITO, Technology engineering and operations

> *"Once the pilot stalled, the vendor kept the infrastructure-readiness dashboard green by reporting only the cloud environments that had passed setup, quietly excluding the legacy on-prem systems that never cleared readiness testing."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Client servicing and communications

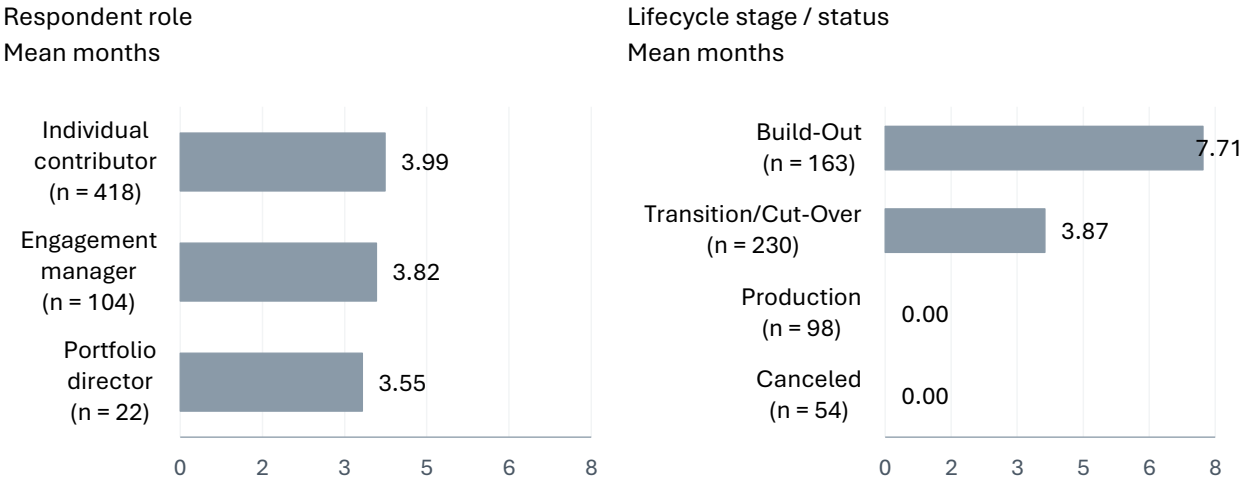
1.4 Planned Time to Production

Indicator C038 · Per the current plan, how many months from today until the engagement reaches Production/Delivery? · Full statistical analysis ›

Respondents reported how many months remained, under the current plan, until the engagement reached Production. The mean answer was 3.9 months (n = 545 respondents on 104 engagements), and 251 respondents answered “I do not know.” Respondents on canceled engagements and on engagements already in Production reported zero. Active engagements in Build-Out reported a mean of 7.7 months (n = 163 respondents), and those in Transition/Cut-Over reported 3.9 months (n = 230 respondents).

Figure 1.4. Mean months remaining to production, by respondent role and engagement status

Respondents reported how many months remained, under the current plan, until the engagement reached Production.



Note: Mean months remaining to production: 3.9 (n = 545 respondents).
n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

A longer planned time to Production was associated with less cancellation (r = −0.37, 95% CI −0.53 to −0.19, p < 0.001, n = 104 engagements). The association rests on the zero answers of canceled engagements and did not hold with the thin-base segments removed. Among active engagements, none of the tested associations reached statistical significance. The customer-satisfaction estimate was within the study’s equivalence bounds (r = −0.04, p = 0.680, n = 94 engagements).

> *There are two issue trackers. We log in ours, they log in theirs, and when same issue is in both nobody merge them, so one problem has two statuses and two owners. I raise this in stand up in first month, nothing change.”*

INDIVIDUAL CONTRIBUTOR

ITO, Technology engineering and operations

2 Depth of Agentic Operation and Vendor Experience

Few engagements were running agents in live operation, and few vendors had delivered agentic work for long. The mean number of agents running in production was 1.2, and the median was 0 (n = 787 respondents), and every respondent on an active engagement before Production reported none. The vendor's experience of agentic AI outsourcing was placed at two years or less by every respondent who answered (n = 571 respondents).

> *We are not having many live systems running yet, which is why the low marks from me... the teams here are still learning on this one."*
>
> **INDIVIDUAL CONTRIBUTOR**
> KPO, Corporate and back-office functions

Production-agent count and reported vendor experience had estimated associations within the study's equivalence bounds for all four measures. These findings do not establish the adequacy of delivery methods or the effect of acquiring more experience. Both questions describe a sample with limited exposure to sustained agentic operation.

> *We sold a data platform build with agents and we did not have anyone in the firm who had actually stood one up end to end, so my engineers were learning the orchestration layer at the same time as they were meant to be delivering it."*
>
> **ENGAGEMENT MANAGER**
> ITO, Technology engineering and operations

> *for me personally it has been good learning, i knew nothing about custody six month back and now i can read a statement properly. only thing, we are the first team doing this so every mistake is ours to find... that is tiring some weeks"*
>
> **INDIVIDUAL CONTRIBUTOR**
> BPO, Trade, settlement and reconciliation

2.1 Agents in Production

Indicator C043 · How many AI agents are running in production on this engagement today? · Full statistical analysis ›

Respondents on active Production engagements reported more agents operating in production than respondents on canceled engagements. Respondents on active engagements in Production reported a mean of 6.1 agents (n = 97 respondents) and a median of 0 (n = 787 respondents). Canceled engagements reported a mean of 1.2 (n = 304 respondents).

Figure 1.5. Agents in Production: associations with the study measures

Agents in production: associations with the study measures

Comparison	Estimate and analysis base
Cancellation	r = −0.01 150 engagements p = 0.922
Customer satisfaction	r = 0.06 94 active engagements p = 0.573
Indicator mean	r = 0.03 150 engagements p = 0.746
Intent to leave	r = 0.01 742 respondents p = 0.913

Note: Mean AI agents in production: 1.2 (n = 787 respondents).

Group means and numeric-answer bases: Individual contributor (n = 612): 1.24. Engagement manager (n = 150): 1.21. Portfolio director (n = 25): 0.88. Build-Out (n = 161): 0.00. Transition/Cut-Over (n = 225): 0.00. Production (n = 97): 6.11. Canceled (n = 304): 1.21.

r is Pearson's correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The estimated associations of production-agent count met the equivalence criterion for cancellation (r = −0.01, p = 0.922, n = 150 engagements), active-engagement customer satisfaction (r = 0.06, p = 0.573, n = 94), engagement indicator mean (r = 0.03, p = 0.746, n = 150), and intent to leave (r = 0.01, p = 0.913, n = 742 respondents). The median count of zero limits conclusions about engagements operating larger numbers of agents.

Four written responses addressed the number of agents, each in a comment on the vendor. They described counts of five to about 40 agents and treated the capacity to run more as a gap in the vendor's readiness.

> *"What has not moved is running agents at any real number... we are fine with twelve, I would not want to be asked for two hundred tomorrow."*
>
> **ENGAGEMENT MANAGER**
>
> KPO, Credit, underwriting and suitability

> *"Running many agents at once is not there yet, we have five and it is already manual"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Technology engineering and operations

2.2 Vendor Tenure in Agentic Delivery Confined to Two Years or Less

Indicator C078 · How many years has your employer been delivering agentic AI outsourcing? · Full statistical analysis ›

No respondent reported that the vendor had delivered agentic AI outsourcing for more than two years. Every answer was zero, one, or two years, with a mean of 1.0 year and a median of 1.0 year (n = 571 respondents). Another 220 respondents answered, “I do not know.”

Figure 1.6. Vendor Tenure in Agentic Delivery Confined to Two Years or Less: associations with the study measures

Vendor tenure in agentic delivery confined to two years or less: associations with the study measures

Comparison	Estimate and analysis base
Cancellation	r = −0.12 150 engagements p = 0.136
Customer satisfaction	r = −0.09 94 active engagements p = 0.407
Indicator mean	r = 0.04 150 engagements p = 0.633
Intent to leave	r = −0.02 543 respondents p = 0.726

Note: Mean years the vendor has delivered agentic AI outsourcing: 1.0 (n = 571 respondents).

Group means and numeric-answer bases: Individual contributor (n = 389): 0.98. Engagement manager (n = 150): 0.94. Portfolio director (n = 30): 0.97. Build-Out (n = 120): 1.05. Transition/Cut-Over (n = 177): 0.99. Production (n = 68): 1.12. Canceled (n = 206): 0.85.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Estimated associations of vendor experience met the equivalence criterion for all four core measures. The correlation with cancellation was r = −0.12 (p = 0.136, n = 150 engagements). It was r = −0.09 with customer satisfaction among active engagements (p = 0.407, n = 94 engagements) and r = 0.04 with the engagement indicator mean (p = 0.633, n = 150 engagements). At respondent level, the correlation with intent to leave was r = −0.02 (p = 0.726, n = 543 respondents). The restricted range of zero to two years limits inference about longer experience. It does not force a null result or establish that experience is unimportant.

Written responses described firms delivering agentic work for the first time.

> *"But nobody here has done this before, there is no playbook. Every time we ask how a agent goes to production the answer is different from each person."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Technology engineering and operations

> *"Playbooks are better than three months back, hence i escalate less. But nobody senior here really knows the agent side, we learn it on shift."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Technology engineering and operations

“

We base the dashboards on real data, but AI still creates statistics that are not correct. We try to fix these before the customer can see but the dashboards are real time. If they look at the dashboard during their work hours, I don't know if what they see is good.”

INDIVIDUAL CONTRIBUTOR

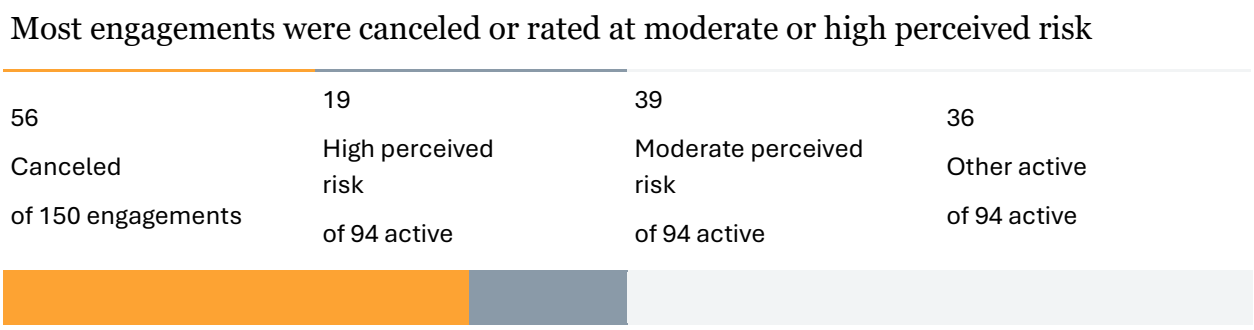
KPO, Regulatory, audit, tax and legal

Cancellation and the Risk of Cancellation

The study recorded cancellation on 56 of 150 engagements (37.33%). Comparisons by institution type, outsourcing model, and vendor type did not detect statistically significant rate differences after the applicable adjustments. This does not establish equal rates across those groups.

Among the 94 active engagements, 19 had a mean original-scale cancellation-likelihood rating of at least 7 out of 10, and another 39 had a mean of at least 5 but below 7. Combining these 58 engagements with the 56 canceled engagements yields 114 of 150 (76.00%). The combined measure describes realized cancellation and perceived risk, not a forecast of eventual cancellations.

Figure 1.7. Composition of the 150 engagements by cancellation and rated risk of cancellation



Note: Risk bands use the engagement mean of cancellation likelihood among the 94 active engagements: high risk 7 or above, moderate risk 5 or above and below 7. Canceled or at risk: 114 of 150 engagements (76.00%).

High perceived risk means a mean cancellation-likelihood rating of 7 or above. Moderate perceived risk means a mean rating of at least 5 and below 7. These ratings express respondents' perceptions. The strip partitions all 150 engagements. Risk categories use the 94 active engagements as their base.

Cancellation contrasts compare ratings recorded after the outcome. Active-engagement analyses avoid that particular temporal problem but remain cross-sectional and subject to confounding and selection. The cancellation-likelihood item requires an additional distinction because respondents on canceled engagements also answered it.

> *"If I am honest we were flying blind. We asked for custodian access in week two and still didn't have it when the plug was pulled."*
>
> **ENGAGEMENT MANAGER**
>
> BPO, Trade, settlement and reconciliation

“

We asked the client's collections team for their hardship scripts in week two and got four different versions from four regions. In fairness that is not carelessness on their part. Collections practice in this bank grew up branch by branch over twenty years and every region has its own regulator conversation behind it, so of course the documentation is uneven. Their change board has been quick with us, and the account data itself is clean, which is more than I usually see at three months. What they do not have is one person who can sit with us for a full day and settle what a compliant arrangement offer looks like. That person exists somewhere. We cannot get her diary.”

ENGAGEMENT MANAGER

CXM, Payments, cards and disputes

1 Realized Cancellation

Realized cancellation is the engagement status recorded in the study dataset. Unlike a rating of future likelihood, it identifies an event reported as already having occurred. Its associations describe differences between canceled and active engagements. The status record does not independently verify the sequence or causes of their operational difficulties.

1.1 Realized Cancellation and Its Association With Indicator Ratings and Intent to Leave

The Engagements sheet records 56 of the 150 engagements as canceled (37.33%). Canceled engagements had a markedly lower engagement indicator mean ($r = -0.72$, 95% CI −0.79 to −0.63, $p < 0.001$, $n = 150$ engagements). Respondents on canceled engagements reported markedly higher intent to leave ($r = 0.76$, cluster bootstrap 95% CI 0.70 to 0.81, $p < 0.001$, $n = 798$ respondents). The cluster bootstrap resamples whole engagements to retain the grouping of respondents within engagements. Both associations held in every robustness check computed.

No tested segment difference in cancellation rate remained statistically significant after Holm adjustment. That procedure limits false-positive conclusions within a specified family of comparisons. This includes institution type, outsourcing model, process family, class of work, and vendor type. The satisfaction distributions for canceled and active engagements did not overlap, as detailed in section 1.1 of “Customer Satisfaction, Delivery Performance, and Measured Return”. Their correlation is a full-sample descriptive contrast rather than an active-engagement estimate.

The ratings of canceled engagements were given after the cancellation, and the associations describe how those engagements were rated, not what preceded the decision. The written responses from canceled engagements describe requests that went unanswered until the engagement ended.

> *“I flagged forty odd vendor records with two different bank details back in March, no answer, chased it twice, still nothing when the account got pulled. There is no one on their side who is allowed to say yes to anything, so the queue just grows and we work round it.”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Corporate and back-office functions

> *Before that I would put a surveillance rule change in our tracker and it would just sit. … We do not have a shared risk list either. We keep ours in jira and they keep theirs in whatever they use, and when an alert set went wrong last month it took two days to even agree it was a problem."*

INDIVIDUAL CONTRIBUTOR

ITO, Regulatory, audit, tax and legal

> *The vendor's monthly report aggregated error rates across all payroll cycles run that month, diluting the visibility of the specific cycle where a paycheck error triggered the cancellation, so the dashboard never flagged the real risk building underneath."*

INDIVIDUAL CONTRIBUTOR

BPO, Corporate and back-office functions

2 Perceived Cancellation Likelihood

The active-engagement risk bands use only the 94 engagements still operating. The broader indicator analysis includes numeric answers from all 150 engagements, despite wording that restricted the question to engagements not yet canceled. The distinction is necessary when interpreting its full-sample correlations.

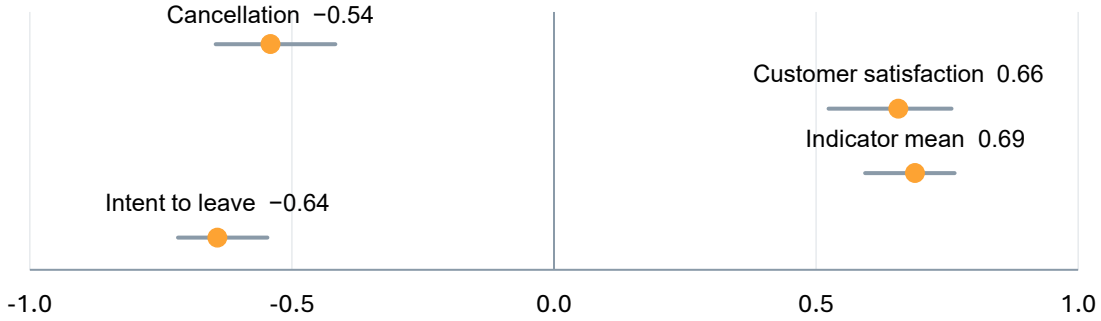
2.1 Delivery Team Estimates of Cancellation Likelihood

Indicator C039 · If the engagement has not been canceled, how likely is the customer to cancel the engagement? · Full statistical analysis ›

The cancellation-likelihood item received 328 numeric responses across 150 engagements and 269 “I do not know” responses. Its favorable-direction mean and median were both 5.0. Higher values indicate lower perceived likelihood of cancellation. These pooled summaries include responses from canceled engagements and should not be read as prospective ratings for active engagements alone.

Figure 1.8. Delivery Team Estimates of Cancellation Likelihood: associations with the study measures

Delivery team estimates of cancellation likelihood: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.54 [−0.65, −0.42] 150 engagements p < 0.001
Customer satisfaction	r = 0.66 [0.52, 0.76] 94 active engagements p < 0.001
Indicator mean	r = 0.69 [0.59, 0.76] 150 engagements p < 0.001
Intent to leave	r = −0.64 [−0.72, −0.55] 316 respondents p < 0.001

Note: Survey question: “If the engagement has not been canceled, how likely is the customer to cancel the engagement?”.

Distressed ratings: 76 of 328 answers (23.17%).

Group means and numeric-answer bases: Individual contributor (n = 145): 4.72. Engagement manager (n = 150): 5.06. Portfolio director (n = 31): 6.26. Build-Out (n = 76): 5.59. Transition/Cut-Over (n = 91): 5.91. Production (n = 40): 5.92. Canceled (n = 121): 3.70.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The indicator is converted to the favorable direction as 11 minus the original rating before these comparisons.

The favorable-direction rating was associated with lower realized cancellation (r = −0.54, 95% CI −0.65 to −0.42, n = 150 engagements), higher customer satisfaction among active engagements (r = 0.66, 95% CI 0.52 to 0.76, n = 94), and a higher engagement indicator mean (r = 0.69, 95% CI 0.59 to 0.76, n = 150). At the respondent level, it was associated with lower original-scale intent to leave (r = −0.64, cluster bootstrap 95% CI −0.72 to −0.55, n = 316). These results passed the applicable checks in the 22-check protocol. The cancellation contrast includes retrospective responses and cannot establish predictive accuracy. Distressed ratings came from 76 of the 328 answers (23.17%), and the engagement mean was distressed on 26 of the 150 engagements.

Individual contributors judged cancellation more likely than portfolio directors, at a mean of 4.7 (n = 145 respondents) against 6.3 (n = 31 respondents). Nine written responses addressed the likelihood of cancellation. Seven, all from canceled engagements, described an end the respondent had expected.

> *“...this whole engagement already feel like it is ending badly so one more issue doesnt change much”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Regulatory, audit, tax and legal

> *“i dont know if anyone above me even knows, happens almost every day, i just override it and keep moving because nobody will fix this before it gets canceled”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Client servicing and communications

“

When the agent mis-cleared a set of cash breaks in March, our own side took two weeks to decide who owned the fix. The platform is good, the answer on accountability was slower than the client's patience.”

ENGAGEMENT MANAGER

BPO, Trade, settlement and reconciliation

Customer Satisfaction, Delivery Performance, and Measured Return

Recorded customer satisfaction was substantially lower on canceled engagements, with no overlap between the canceled and active distributions. Section 1.1 reports their ranges. Both the recorded score and the separate perception rating were supplied by vendor personnel. Neither was verified against financial institution records.

Delivery-date performance, budget performance, and progress against key performance indicators received mean ratings between 4.8 and 4.9. These indicator-level results and their associations are reported in section 2. They describe perceived delivery performance rather than independently verified service outcomes.

> *KPIs keep slipping and nobody agrees on why"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Payments, cards and disputes

> *ROI case for this was shaky from the start"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Corporate and back-office functions

Return analysis is restricted to Production engagements and, for the reported correlations, to the 17 with positive recorded returns. The small selected sample and unspecified financial measurement conventions limit comparison and generalization (section 3).

> *The problem is we raise the same funding cutoff issue on every status call and get nowhere, we learned the actual process just by watching their ops team work it live. And any decision that needs sign off stalls enough to blow the settlement window."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

“

What would actually help is having one person on the team from day one who has run an internal service desk. We have people who know the AI side well, that part is fine. What we don't have is anyone who can tell a desk owner what a realistic deflection number looks like, so I end up guessing. The other piece is pricing. We are still selling this shaped like a staffing deal, and it gets awkward when the agent takes the volume and the invoice looks the same.”

ENGAGEMENT MANAGER

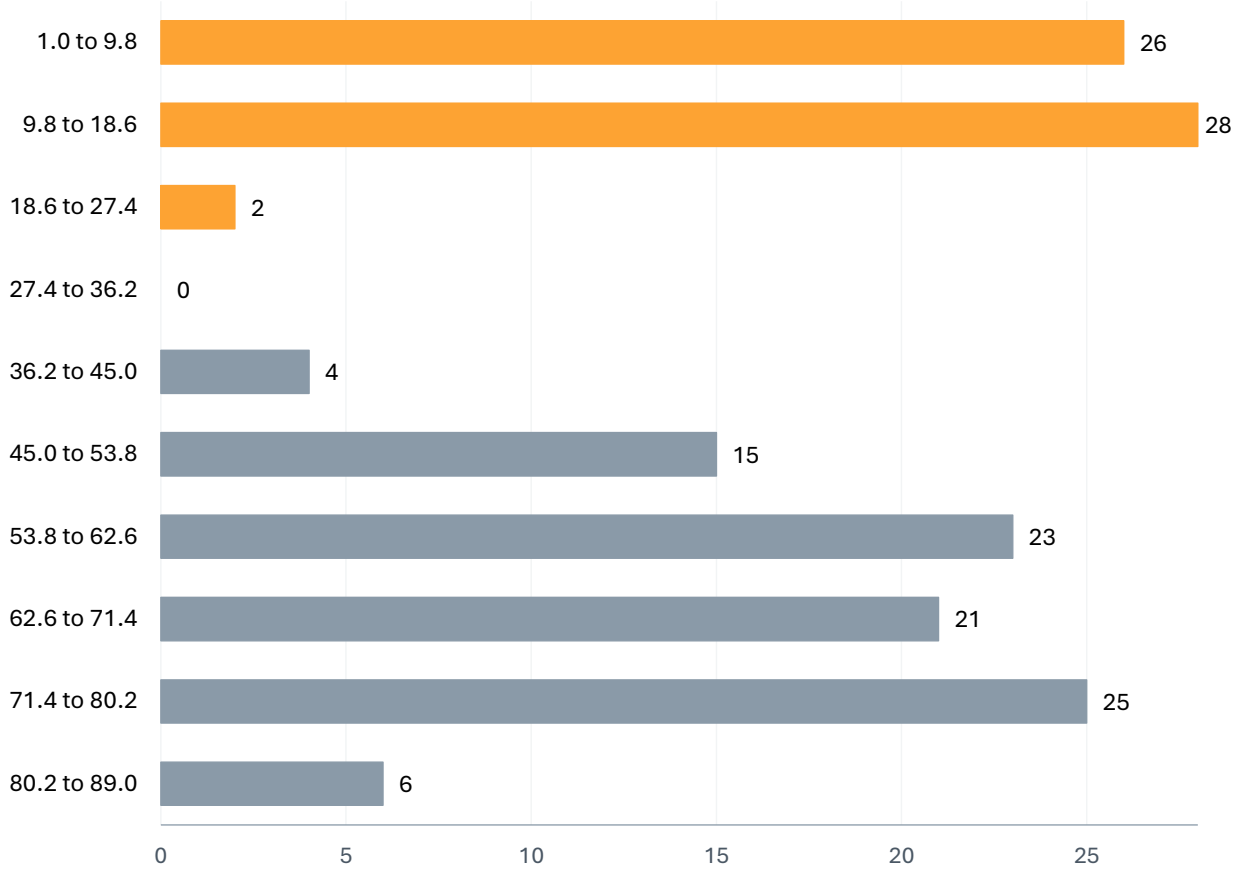
CXM, Client servicing and communications

1 Recorded and Perceived Customer Satisfaction

The study distinguishes a 100-point satisfaction score recorded for each engagement from a separate 10-point item asking whether the financial institution was satisfied with performance. Figure 1.9 presents the distribution of recorded scores.

Figure 1.9. Engagements by recorded customer satisfaction, in ten equal bands, 150 engagements

Recorded customer satisfaction varied widely across engagements



Note: The three lowest bands hold 56 engagements, exactly the canceled engagements (amber). No engagement recorded a value from 27.4 to 36.2.

1.1 Non-Overlapping Satisfaction of Canceled and Active Engagements

Recorded customer satisfaction averaged 43.85 points, with a median of 53.50, across 150 engagements. Active engagements averaged 64.10 points and ranged from 40.00 to 89.00 (n =

94). Canceled engagements averaged 9.88 points and ranged from 1.00 to 20.00 (n = 56). The two distributions did not overlap.

Satisfaction was associated with every other study measure, and the associations were retained under the applicable checks in the 16-check measure protocol. Its association with cancellation (r = −0.94, 95% CI −0.96 to −0.92, n = 150 engagements) measures the separation between the two groups. Higher satisfaction was associated with a higher engagement indicator mean (r = 0.85, 95% CI 0.80 to 0.89, n = 150 engagements). It was also associated with lower intent to leave (r = −0.84, cluster bootstrap 95% CI −0.87 to −0.81, n = 798 respondents).

The associations with the engagement indicator mean and with intent to leave did not depend on the canceled engagements. Among the 94 active engagements alone, they held at r = 0.84 and r = −0.63, respectively. Among the engagements still running, those delivered under better circumstances recorded higher satisfaction, and their delivery teams were less inclined to leave.

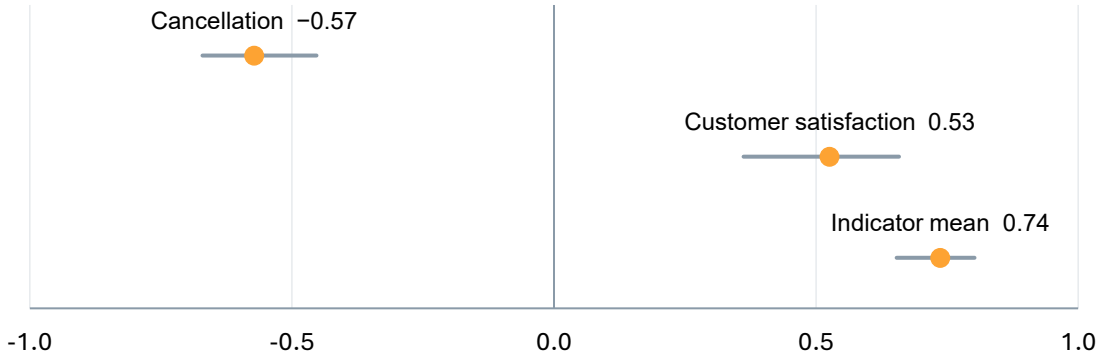
1.2 Delivery Team Perception of Financial Institution Satisfaction

Indicator C048 · The customer is happy with the performance of the engagement at this stage · Full statistical analysis ›

Individual contributors assessed financial institution satisfaction less favorably than portfolio directors, and ratings were lowest on canceled engagements. The mean rating was 5.0, and the median was 5.0 (n = 379 respondents). 244 respondents answered: “I do not know.”

Figure 1.10. Delivery Team Perception of Financial Institution Satisfaction: associations with the study measures

Delivery team perception of financial institution satisfaction: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.57 [−0.67, −0.45] 150 engagements p < 0.001
Customer satisfaction	r = 0.53 [0.36, 0.66] 94 active engagements p < 0.001
Indicator mean	r = 0.74 [0.65, 0.80] 150 engagements p < 0.001

Note: Survey question: “The customer is happy with the performance of the engagement at this stage”.

Distressed ratings: 89 of 379 answers (23.48%).

Group means and numeric-answer bases: Individual contributor (n = 196): 4.76. Engagement manager (n = 150): 5.00. Portfolio director (n = 31): 5.97. Build-Out (n = 80): 5.88. Transition/Cut-Over (n = 110): 5.65. Production (n = 46): 5.61. Canceled (n = 143): 3.69.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The graphic includes the associations reproduced in the section. Additional results remain in the text.

Perceived financial institution satisfaction was associated with all four core measures and passed the applicable checks in the 22-check protocol. Its correlation with cancellation was −0.57 (95% CI −0.67 to −0.45, n = 150 engagements). Among active engagements, its correlation with recorded customer satisfaction was 0.53 (95% CI 0.36 to 0.66, n = 94 engagements). Its correlation with the engagement indicator mean was 0.74 (95% CI 0.65 to 0.80, n = 150 engagements), the largest observed coefficient among these comparisons. Coefficient ordering alone does not establish a difference in the underlying association strengths. Distressed ratings came from 89 of the 379 answers (23.48%), and the engagement mean was distressed on 29 of the 150 engagements.

The written responses that addressed the financial institution's satisfaction numbered 13, from 13 respondents on 12 engagements, and 9 of the 13 were negative.

> *...I sat between the client asking why supplier records were still wrong and my own management asking me for a green status. Both directions, every day. ...”*
>
> **ENGAGEMENT MANAGER**
>
> CXM, Corporate and back-office functions

> *Not everything here is same, main thing is templates get changed their side, we find out when client complains”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Client servicing and communications

2 Delivery Against Dates, Budget, and Key Performance Indicators

Reported delivery assessments were concentrated in a narrow range, with favorable-direction means from 4.8 to 4.9 (n = 554–798 respondents). The four questions addressed delivery dates, budget adherence, progress toward key performance indicators, and respondents' knowledge of those indicators. The measures and the practices described in written answers remain distinct subjects.

> *“sign off arrive after the filing deadline it was even meant for, the Review Board sit too rarely to help, so i decide myself and kindly hope it hold up later”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Regulatory, audit, tax and legal

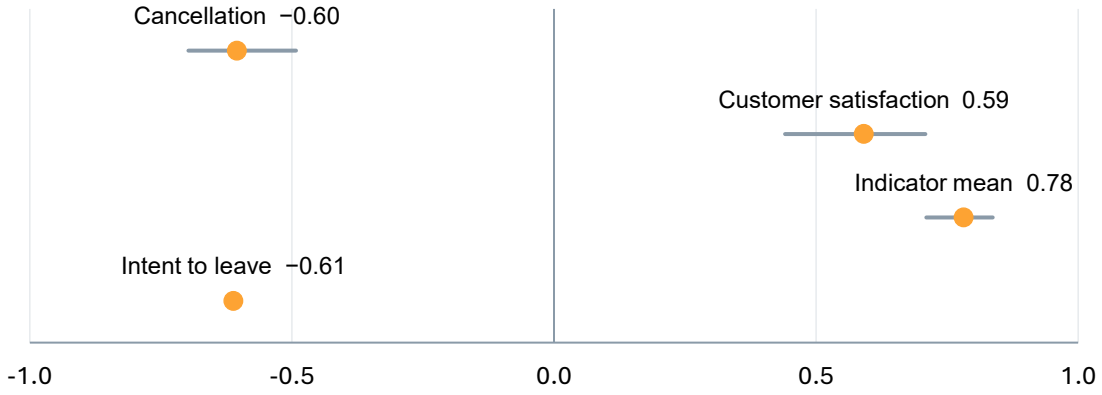
2.1 Reported Meeting Delivery Dates

Indicator C049 · The engagement is meeting delivery dates successfully · Full statistical analysis ›

Ratings of meeting delivery dates were lowest on canceled engagements and less favorable among individual contributors than portfolio directors. The mean rating was 4.8 and the median 5.0 (n = 798 respondents). Canceled engagements rated lowest, at 3.4 (n = 304 respondents), and individual contributors rated lower than portfolio directors.

Figure 1.11. Reported Meeting Delivery Dates: associations with the study measures

Reported meeting delivery dates: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.60 [−0.70, −0.49] 150 engagements p < 0.001
Customer satisfaction	r = 0.59 [0.44, 0.71] 94 active engagements p < 0.001
Indicator mean	r = 0.78 [0.71, 0.84] 150 engagements p < 0.001
Intent to leave	r = −0.61 751 respondents p < 0.001

Note: Survey question: “The engagement is meeting delivery dates successfully”.

Distressed ratings: 234 of 798 answers (29.32%).

Group means and numeric-answer bases: Individual contributor (n = 624): 4.61. Engagement manager (n = 150): 5.19. Portfolio director (n = 24): 6.46. Build-Out (n = 161): 5.55. Transition/Cut-Over (n = 236): 5.56. Production (n = 97): 5.79. Canceled (n = 304): 3.43.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Ratings of delivery against dates were associated with all four study measures, each association holding under the applicable robustness checks. Engagements rated as meeting their dates had a lower cancellation rate (r = −0.60, 95% CI −0.70 to −0.49, n = 150 engagements). They also recorded higher customer satisfaction among active engagements (r = 0.59, 95% CI 0.44 to 0.71, n = 94 engagements). The largest observed absolute coefficient was for the engagement indicator mean (r = 0.78, 95% CI 0.71 to 0.84, n = 150 engagements). Respondents who rated delivery higher reported lower intent to leave (r = −0.61, n = 751 respondents). Distressed ratings came from 234 of the 798 answers (29.32%), and the engagement mean was distressed on 30 of the 150 engagements.

The written responses that addressed delivery dates numbered 57, from 54 respondents on 42 engagements, and 42 of the 57 were negative. In 14 responses, approvals arrived after the deadline, and the respondent decided without them or let the date slip.

> *"I just made the call myself, the change approval group meets barely at all and the decision always lands after the deadline its needed for, so I tweak the account data to match what the workflow expects and push it through myself before anyone signs off."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Client servicing and communications

> *"In June we had a release window on a Friday evening, and the change record needed an approver from their platform team, who had gone on leave without delegating, so the deploy sat until Monday even though everything else was ready."*
>
> **ENGAGEMENT MANAGER**
>
> ITO, Technology engineering and operations

Another 14 written responses described meeting deadlines by bypassing a process or approval, and seven described reports that concealed delays from the financial institution. These are respondent accounts, not independently inspected delivery or reporting records.

> *...change process is too slow for how fast we work, so I catch it myself and keep report on time"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Regulatory, audit, tax and legal

> *The delivery methodology gaps got hidden by scoping every missed milestone as a change request, which reset the baseline schedule each time and kept the reported on-time delivery rate looking healthy."*

INDIVIDUAL CONTRIBUTOR

ITO, Technology engineering and operations

> *And do note it happen most weeks, i tell my lead about the same issue every time, nothing change so i just fix it and the payroll goes out on time"*

INDIVIDUAL CONTRIBUTOR

BPO, Corporate and back-office functions

> *The vendor reported fleet-scaling milestones as met by counting agents deployed to a test environment rather than production, so the status report showed steady progress for months while zero tickets were actually being routed live."*

INDIVIDUAL CONTRIBUTOR

CXM, Technology engineering and operations

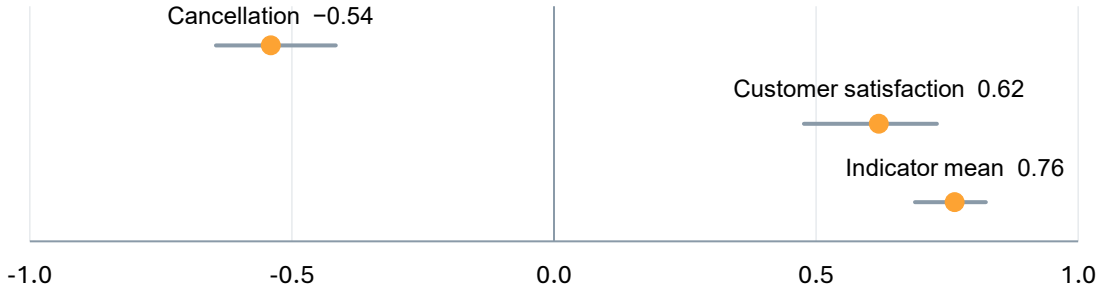
2.2 Budget Pressure From Effort the Plan Did Not Price

Indicator C050 · The engagement is within budget · Full statistical analysis ›

More favorable assessments of delivery within budget were associated with all four core study measures. The mean rating was 4.9, and the median was 5.0 (n = 554 respondents), and 242 respondents answered: “I do not know.” Delivery within budget was associated with all four study measures, and none of the applicable robustness checks overturned an association. Engagements rated within budget had a lower cancellation rate (r = −0.54, 95% CI −0.64 to −0.42, n = 150 engagements). Their customer satisfaction among active engagements was higher (r = 0.62, 95% CI 0.48 to 0.73, n = 94 engagements). The largest observed absolute coefficient was for the engagement indicator mean (r = 0.76, 95% CI 0.69 to 0.82, n = 150 engagements). Distressed ratings came from 151 of the 554 answers (27.26%), and the engagement mean was distressed on 31 of the 150 engagements.

Figure 1.12. Budget Pressure From Effort the Plan Did Not Price: associations with the study measures

Budget pressure from effort the plan did not price: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.54 [−0.64, −0.42] 150 engagements p < 0.001
Customer satisfaction	r = 0.62 [0.48, 0.73] 94 active engagements p < 0.001
Indicator mean	r = 0.76 [0.69, 0.82] 150 engagements p < 0.001

Note: Survey question: “The engagement is within budget”.

Distressed ratings: 151 of 554 answers (27.26%).

Group means and numeric-answer bases: Individual contributor (n = 375): 4.79. Engagement manager (n = 150): 5.10. Portfolio director (n = 28): 6.07. Build-Out (n = 117): 5.72. Transition/Cut-Over (n = 164): 5.58. Production (n = 70): 5.96. Canceled (n = 203): 3.62.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The graphic includes the associations reproduced in the section. Additional results remain in the text.

Seven written responses addressed the budget, each negative. Six described effort or cost that the engagement absorbed outside its plan or fee, in the form of rework, staff added under a fixed fee, repeated runs of the AI tools, and manual checks.

> *In June we were asked to add two more allocation reviewers and were told the fee is fixed on the numbers agreed last year, which nobody has revisited since agents took half the run. So we absorbed it."*
>
> **ENGAGEMENT MANAGER**
>
> BPO, Trade, settlement and reconciliation

> *on my last project the bot was a script, you knew the cost. this one burns through budget re-runing the same patch check four times, nobody warned us about that part"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Client servicing and communications

> *We check output every run now, it move on its own sometimes, feels like extra work nobody planned for but we do it anyway."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Corporate and back-office functions

> *Scope-creep tickets tied to mandate changes were logged under a generic 'change request' bucket, which kept the reported schedule variance small while the real rework hours piled up underneath it."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Payments, cards and disputes

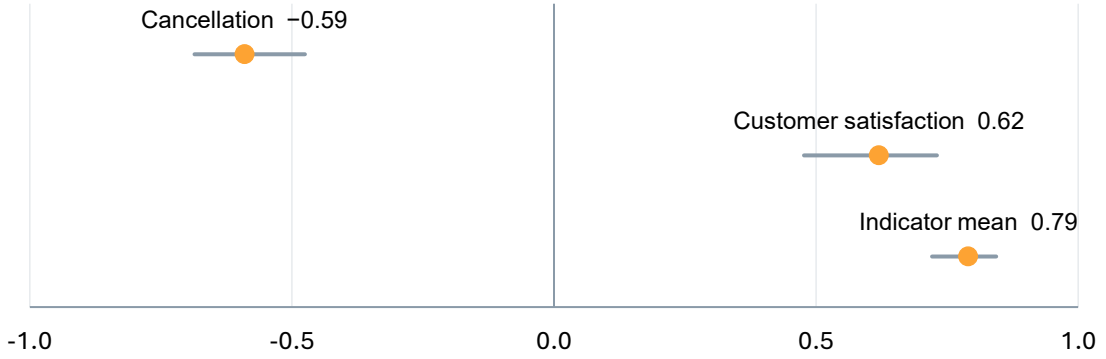
2.3 Engagements Not Achieving Stated Key Performance Indicators

Indicator C051 · The engagement is not achieving stated KPIs · Full statistical analysis ›

Individual contributors assessed progress toward the engagement's stated key performance indicators less favorably than portfolio directors. Converted to the favorable direction, on which a higher figure marks an engagement closer to its targets, the mean rating was 4.8 (n = 782 respondents). Individual contributors rated the engagement furthest from its targets, at 4.7 (n = 608 respondents), and portfolio directors closest, at 5.8 (n = 24 respondents).

Figure 1.13. Engagements Not Achieving Stated Key Performance Indicators: associations with the study measures

Engagements not achieving stated key performance indicators: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.59 [−0.69, −0.48] 150 engagements p < 0.001
Customer satisfaction	r = 0.62 [0.48, 0.73] 94 active engagements p < 0.001
Indicator mean	r = 0.79 [0.72, 0.84] 150 engagements p < 0.001

Note: Survey question: “The engagement is not achieving stated KPIs”.

Distressed ratings: 241 of 782 answers (30.82%).

Group means and numeric-answer bases: Individual contributor (n = 608): 4.65. Engagement manager (n = 150): 5.06. Portfolio director (n = 24): 5.83. Build-Out (n = 166): 5.63. Transition/Cut-Over (n = 229): 5.53. Production (n = 93): 5.83. Canceled (n = 294): 3.35.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The indicator is converted to the favorable direction as 11 minus the original rating before these comparisons.

The graphic includes the associations reproduced in the section. Additional results remain in the text.

Progress toward the key performance indicators was associated with all four study measures, and the associations held under every one of the applicable robustness checks. Engagements rated further from their targets had a higher cancellation rate (r = −0.59, 95% CI −0.69 to −0.48, n = 150 engagements). Engagements rated closer to their targets recorded higher customer satisfaction among active engagements (r = 0.62, 95% CI 0.48 to 0.73, n = 94 engagements). They also had a higher engagement indicator mean (r = 0.79, 95% CI 0.72 to 0.84, n = 150 engagements). Distressed ratings came from 241 of the 782 answers (30.82%), and the engagement mean was distressed on 32 of the 150 engagements.

The written responses that addressed the indicator numbered 14, from 13 respondents on 12 engagements, and 12 of the 14 were negative. Respondents described key performance indicators that were missed and service levels that declined once agents took on the tasks. Four responses, given to the question on how the vendor altered its reports, described missed targets presented to the financial institution under other headings. Part 5 examines the alteration of performance reporting.

> *“Missed KPI targets get reclassified as data-quality exceptions rather than performance misses, which keeps the official miss rate lower than what the team actually experiences month to month.”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Client servicing and communications

> *“SLAs collapsed once agents touched the work”*
>
> **ENGAGEMENT MANAGER**
>
> ITO, Client servicing and communications

> *“KPIs are missed almost every cycle and nobody above me seems alarmed enough”*
>
> **ENGAGEMENT MANAGER**
>
> CXM, Corporate and back-office functions

> *“SLA and KPI targets collapsed the moment agent output stopped matching human pace”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

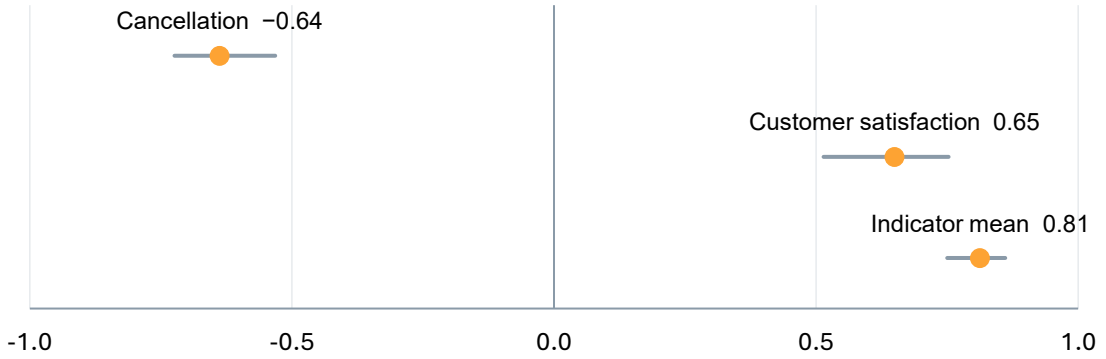
2.4 Delivery Team Knowledge of Financial Institution Key Performance Indicators

Indicator C047 · I know the KPIs the customer uses to measure success · Full statistical analysis ›

Respondents rated the statement that they knew the key performance indicators the financial institution used to measure success, with a mean rating of 4.8 and a median of 5.0 (n = 787 respondents). Canceled engagements rated lowest, and individual contributors rated lower than portfolio directors.

Figure 1.14. Delivery Team Knowledge of Financial Institution Key Performance Indicators: associations with the study measures

Delivery team knowledge of financial institution key performance indicators: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.64 [−0.72, −0.53] 150 engagements p < 0.001
Customer satisfaction	r = 0.65 [0.51, 0.75] 94 active engagements p < 0.001
Indicator mean	r = 0.81 [0.75, 0.86] 150 engagements p < 0.001

Note: Survey question: “I know the KPIs the customer uses to measure success”.

Distressed ratings: 226 of 787 answers (28.72%).

Group means and numeric-answer bases: Individual contributor (n = 609): 4.65. Engagement manager (n = 150): 5.05. Portfolio director (n = 26): 6.08. Build-Out (n = 162): 5.59. Transition/Cut-Over (n = 234): 5.56. Production (n = 95): 5.85. Canceled (n = 296): 3.36.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The graphic includes the associations reproduced in the section. Additional results remain in the text.

Self-rated knowledge of performance targets was associated with lower cancellation (r = −0.64, 95% CI −0.72 to −0.53, n = 150 engagements), higher active-engagement customer satisfaction (r = 0.65, 95% CI 0.51 to 0.75, n = 94), and a higher engagement indicator mean (r = 0.81, 95% CI 0.75 to 0.86, n = 150). The applicable sensitivity checks did not overturn these associations. The coefficient for the indicator mean was numerically largest among the four delivery indicators, without establishing an independent contribution or a statistically confirmed ranking. Distressed ratings came from 226 of the 787 answers (28.72%), and the engagement mean was distressed on 30 of the 150 engagements.

The single written response on the indicator described expectations the financial institution had not stated to the delivery team.

3 Measured Return in Production

The recorded-return field applies to the 30 Production engagements. It is not populated for the other 120 engagements. A full-sample mean would mix observed values with structural absence. The financial definition and source limitations are stated in the Introduction, section 5.5 of “Study Design, Evidentiary Basis, and Definitions.”

3.1 Measured Return Confined to Active Production Engagements

ANALYSIS · REPORTED RETURN ON THE ENGAGEMENT, ROI % · FULL STATISTICAL ANALYSIS ›

3.1 Positive Measured Return Confined to Active Production Engagements

All 30 Production engagements had a return value recorded. Seventeen had positive values, averaging 24.71% and ranging from 6.00% to 47.00%. The other 13 were recorded at zero: 10 canceled engagements and three active engagements. All positive-return engagements were active. These values describe the study record, not an independent financial verification. The recorded zeros do not establish whether the respondent lacked the information, the engagement had achieved no return, or a return had not yet been realized.

Within the 17 engagements that reported a return, a higher return was associated with higher customer satisfaction (r = 0.71, 95% CI 0.35 to 0.89, p = 0.001, n = 17 engagements). A higher return was also associated with lower intent to leave (r = −0.71, 95% CI −0.89 to −0.34, p = 0.002, n = 17 engagements). The association with the engagement indicator mean was positive and inconclusive (r = 0.47, 95% CI −0.01 to 0.78, p = 0.056, n = 17 engagements), a test with power of 0.48 at this base. Every result rests on 17 engagements, and the intervals are wide.

The return is provided by the vendor delivery team. The study had no method to validate the provided figures or the level of first-hand knowledge by the respondents. The study holds no record from the financial institution against which to confirm it.

Two engagement managers, writing on the vendor's reporting, described return calculations that overstated the return. Individual contributors described efficiency figures computed without the hours staff spent checking or completing the agents' work.

> *“The vendor's ROI tracker recalculated the baseline savings assumption each quarter to match whatever the actual result was, so the reported ROI never fell short even as the underlying business case kept eroding.”*
>
> **ENGAGEMENT MANAGER**
>
> KPO, Corporate and back-office functions

> *The vendor's monthly value report counted every hour the agent spent on forecasting as fully productive, with no line for the hours controllers then spent reconciling those forecasts against the ledger, which made the ROI case look real when it was not."*

ENGAGEMENT MANAGER

KPO, Corporate and back-office functions

> *The vendor's efficiency report calculated hours saved against a fully manual baseline without netting out the additional hours analysts spent verifying agent-flagged outputs, inflating the apparent productivity gain shown to the desk."*

INDIVIDUAL CONTRIBUTOR

KPO, Corporate and back-office functions

> *The vendor reported time saved based on tasks routed to the agent, without considering the extra time staff spent manually completing the same tasks after the agent hit an API wall, which overstated the actual efficiency gain."*

INDIVIDUAL CONTRIBUTOR

CXM, Corporate and back-office functions

> *The vendor's client-facing performance dashboard rolls up billing accuracy as a single blended percentage across all fee types, which lets a persistent error rate on complex tiered fee calculations get diluted by the much larger volume of simple flat fees the agent handles cleanly, so the dashboard never isolates the tiered-fee problem that audit would actually care about."*

PORTFOLIO DIRECTOR

BPO, Payments, cards and disputes

The Engagement Indicator Mean and the Interdependence of the Four Measures

The engagement indicator mean summarizes ratings across the five indicator categories. This section examines its distribution, shared variation with the other core measures, and the extent to which individual indicators retain associations after adjustment for that mean.

The indicators should be interpreted as correlated assessments of engagement conditions. None of the 300 partial tests met the multiplicity-adjusted significance criterion. This limits claims about indicator-specific associations but does not prove that individual conditions have no independent effect.

> *“The vendor's reporting rolled up successful completions across all payroll runs company-wide, which buried the specific business unit where repeated errors were occurring until the pattern was undeniable and the contract had already soured.”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Corporate and back-office functions

“

Honestly it comes down to the deadline being fixed while everything else moves. Model output shifts between runs, nobody on the client side is willing to own the call, and change control cannot keep pace, so I sign off on adjusted numbers myself rather than let the deadline slip again. I know it is not strictly how this is meant to work, and I have said so to both sides without anything changing.”

ENGAGEMENT MANAGER

KPO, Corporate and back-office functions

1 The Engagement Indicator Mean as a Summary Measure

The summary averages respondents' ratings across the 100 indicators and then aggregates them by engagement. Its construction and the leave-one-out procedure are defined in the Introduction, section 5.1 of “Study Design, Evidentiary Basis, and Definitions.”

Across the 150 engagements, the engagement indicator mean was 4.02 and the median 3.8, from 1.2 to 7.3. The rating behind it was given by all 850 respondents, each of whom rated at least one of the 100 indicators. Canceled engagements recorded a mean of 2.6 (n = 56 engagements), compared with 4.8 to 4.9 in the three active stages.

A higher engagement indicator mean was associated with less cancellation (r = −0.72, 95% CI −0.79 to −0.63, n = 150 engagements). It was associated with higher customer satisfaction among the active engagements (r = 0.84, 95% CI 0.77 to 0.89, n = 94 engagements). The mean was also associated with lower intent to leave (r = −0.78, cluster bootstrap 95% CI −0.82 to −0.74, n = 798 respondents). The associations were retained under the applicable checks in the 15-check measure protocol.

“

To be fair the outside systems were not what killed this. The card platform worked, the statement print vendor did what they said. We just never got far enough to test any of it properly, two months and it was over. What broke was on our side and their side, the feeds we build ourselves. I would not put the outside people in the same bucket, they held up fine in the little we did.”

INDIVIDUAL CONTRIBUTOR

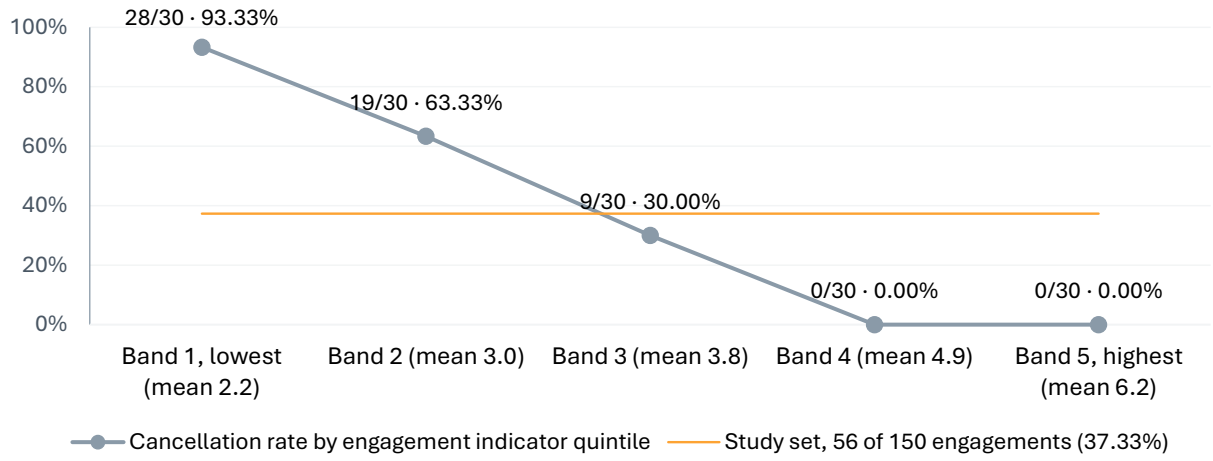
BPO, Payments, cards and disputes

2 Shared Structure Across the Four Study Measures

The four core study measures shared substantial variation across engagements, and the other three measures differed systematically across the engagement indicator quintiles. Pairwise correlations and principal component analysis describe the shared structure. Principal component analysis identifies statistical combinations that capture shared variation among the measures. Figure 1.15 shows the observed cancellation gradient across those quintiles.

Figure 1.15. Cancellation rate by engagement indicator quintile, five bands of 30 engagements

Cancellation declined across the engagement indicator bands



Note: The groups are ordered categories. The connecting lines do not represent elapsed time.

Canceled engagements: 28, 19, 9, 0, and 0 of 30 in bands 1 to 5. The bands differed in cancellation (chi-square 85.32, df = 4, $p < 0.001$, n = 150 engagements).

The measures share substantial variance. An indicator associated with several outcomes does not supply several independent confirmations of the same finding. Common respondent perceptions and engagement circumstances may contribute to the pattern.

2.1 A Single Dominant Component Across the Four Study Measures

The four core measures were correlated with one another in every pair across the 150 engagements. The absolute correlations ranged from 0.72 to 0.94. The strongest pair was cancellation with customer satisfaction ($r = -0.94$, 95% CI −0.96 to −0.92, $p < 0.001$, n = 150 engagements). The weakest was cancellation with the engagement indicator mean ($r = -0.72$, 95% CI −0.79 to −0.63, $p < 0.001$, n = 150 engagements).

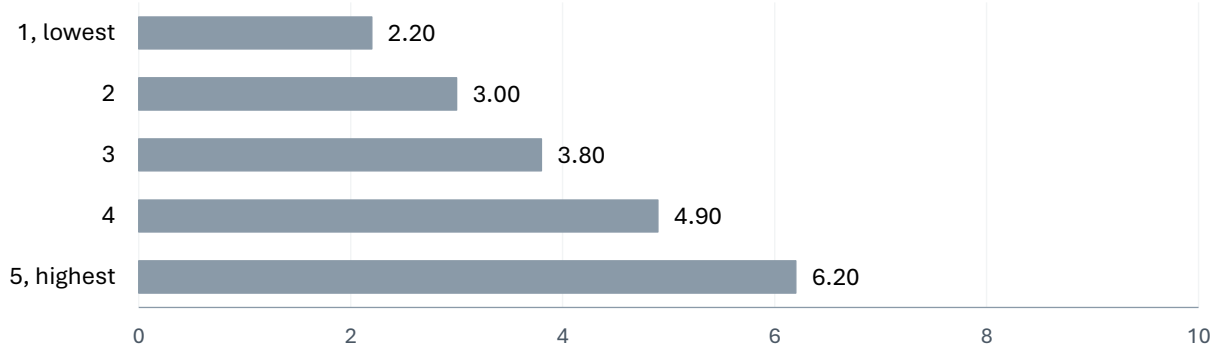
Principal component analysis identified one dominant component, accounting for 88.23% of variance across the 150 engagements. An eigenvalue is the variance represented by a component. Its eigenvalue was 3.53, compared with 0.30 for the second component. A loading expresses how strongly a measure relates to a component. Absolute loadings on the first component were at least 0.91. Cancellation and intent to leave loaded negatively. The result describes shared variation, not a causal mechanism or the interchangeability of the four measures.

2.2 Differences in the Study Measures Across Indicator Quintiles

The 150 engagements were ranked by engagement indicator mean and divided into five bands of 30. No study measure was worse in any band than in the band before it.

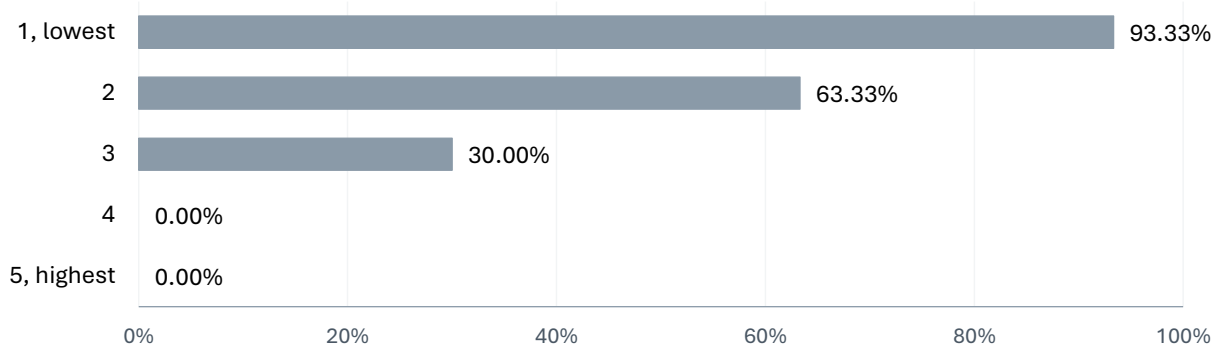
Table 1.1. The study measures by engagement indicator quintile

Engagement indicator mean



Band	Engagements	Engagement indicator mean	Canceled
1, lowest	30	2.2	28
2	30	3.0	19
3	30	3.8	9
4	30	4.9	0
5, highest	30	6.2	0

Cancellation rate



Band	Cancellation rate	Customer satisfaction (mean)	Intent to leave (engagement mean)
1, lowest	93.33%	13.0	9.1
2	63.33%	24.3	8.5
3	30.00%	43.0	7.0
4	0.00%	63.2	5.4
5, highest	0.00%	75.8	3.8

The bands differed in cancellation (chi-square 85.32, df = 4, $p < 0.001$, n = 150 engagements). They differed in customer satisfaction ($F = 84.65$, df = 4 and 145, $p < 0.001$, omega squared 0.690, n = 150 engagements). The bands also differed in intent to leave ($F = 76.29$, df = 4 and 145, $p < 0.001$, omega squared 0.668, n = 150 engagements). Here F is the test statistic for differences in means, df denotes degrees of freedom, and omega squared estimates the proportion of variation associated with the groups. Degrees of freedom describe the information available to estimate variation. The rise of the engagement indicator mean from band to band follows from how the bands were built and is not a finding. The behavior of the other three measures is. Engagements in the lowest indicator quintile were almost all canceled, at 28 of 30. None of the 60 engagements in the two highest indicator quintiles had been canceled.

2.3 Persistence of the Associations Among Active Engagements

Among the 94 engagements not canceled by the study date, the three remaining core measures were correlated in every pair. Higher customer satisfaction was associated with a higher engagement indicator mean ($r = 0.84$, 95% CI 0.77 to 0.89, $p < 0.001$, n = 94 engagements). It was associated with lower intent to leave ($r = -0.73$, 95% CI −0.81 to −0.62, $p < 0.001$, n = 94 engagements). A higher engagement indicator mean was associated with lower intent to leave ($r = -0.69$, 95% CI −0.78 to −0.57, $p < 0.001$, n = 94 engagements).

Among active engagements, higher perceived cancellation likelihood was associated with lower customer satisfaction ($r = -0.66$, 95% CI −0.76 to −0.52, $p < 0.001$, n = 94). This is an additional subjective-risk measure. Realized cancellation is constant in the active group, so the active-only analysis supports associations among the three remaining core measures rather than reproducing the original four-measure structure.

3 Indicators Net of the Engagement Indicator Mean

Partial correlations examine whether an indicator retains an association after accounting for the engagement indicator mean. These tests inform interpretation of individual indicators. They do not identify causal effects or demonstrate that all conditions operate as a single mechanism.

Each of the 100 rated indicators was correlated with cancellation, customer satisfaction, and the engagement mean of intent to leave, with the engagement indicator mean held constant. The mean was recomputed without the indicator under test. In 33 of the 300 tests, p was under 0.05, against 15 expected by chance alone.

None of the 300 partial tests remained statistically significant after false-discovery-rate adjustment. The smallest adjusted q value was 0.323. The largest absolute partial correlation was 0.25. Its magnitude was below the study's 0.30 equivalence bound, but a point estimate alone does not establish equivalence.

Accordingly, the report does not attribute outcomes independently to an individual indicator. Differences in observed coefficients are descriptive unless a specified comparison test supports a stronger conclusion. The written accounts help distinguish the conditions respondents described without establishing their separate causal contributions.

“

Merchant onboarding volume went up nearly double after client launched new acquiring product, and our agents absorbed it without adding headcount. Good news, except that our invoice to client is still built on seats. So we did more work and billed less, and i had to explain to my own director why revenue on account went flat in a month where we handled forty percent more cases.”

ENGAGEMENT MANAGER

CXM, Payments, cards and disputes

Structural Attributes and Outcome Associations

The structural analysis covers institution type, outsourcing model, work characteristics, commercial terms, vendor characteristics, and team size. Ten of 19 attributes met the study's equivalence criteria against all four core measures. Other results included supported associations and inconclusive estimates. This mixed pattern requires attribute-specific interpretation rather than a general conclusion that structural characteristics do not matter.

Each attribute's result against each core measure is stated in one of the four states of a reported result.

Table 1.2. Structural attributes against the four core study measures

Attribute	Cancellation rate	Customer satisfaction	Engagement indicator mean	Intent to leave
Type of financial institution	Equivalent to zero	Inconclusive	Equivalent to zero	Equivalent to zero
Asset size of the financial institution	No difference	No difference	No difference	No difference
Financial institution operates capability centers	Equivalent to zero	Inconclusive	Equivalent to zero	Equivalent to zero
Financial institution scaling agentic outsourcing	Equivalent to zero	Equivalent to zero	Equivalent to zero	Equivalent to zero
Financial institution adding agentic engagements	Equivalent to zero	Equivalent to zero	Equivalent to zero	Equivalent to zero
Financial institution shifting work to its capability centers	Associated	Associated	Associated	Associated
Length of the vendor's relationship	Equivalent to zero	Inconclusive	Equivalent to zero	Equivalent to zero
Concurrent agentic engagements with the financial institution	Equivalent to zero	Equivalent to zero	Equivalent to zero	Equivalent to zero
Outsourcing model	Equivalent to zero	Equivalent to zero	Equivalent to zero	Equivalent to zero
Class of work	Equivalent to zero	Equivalent to zero	Equivalent to zero	Equivalent to zero

Table 1.2. Structural attributes against the four core study measures (continued)

Attribute	Cancellation rate	Customer satisfaction	Engagement indicator mean	Intent to leave
Process family	Inconclusive	Equivalent to zero	Equivalent to zero	Inconclusive
Pricing model	Inconclusive	Not tested	Inconclusive	Equivalent to zero
Contract term	Equivalent to zero	Equivalent to zero	Equivalent to zero	Equivalent to zero
Annual contract value (31 engagements)	Inconclusive	Inconclusive	Inconclusive	Inconclusive
Vendor type	Equivalent to zero	Equivalent to zero	Equivalent to zero	Equivalent to zero
Vendor size	Equivalent to zero	Equivalent to zero	Equivalent to zero	Equivalent to zero
Vendor's agentic engagements for financial institutions	Equivalent to zero	Equivalent to zero	Equivalent to zero	Equivalent to zero
Current size of the delivery team	Equivalent to zero	Equivalent to zero	Equivalent to zero	Equivalent to zero
Peak size of the delivery team	Associated	Equivalent to zero	Associated, not among active engagements	Associated, not among active engagements

Two attributes are exceptions, each with a limit. Engagements whose respondents agreed that the financial institution was shifting more work to its own global capability centers recorded worse outcomes on every measure, and the noncancellation associations were retained in the active-engagement analyses. The rating records the respondents' view of the financial institution's plans, and the study holds no evidence on which came first. Engagements that had reached a larger peak team had a higher cancellation rate. Two of the three associations of peak team size did not hold among active engagements, and the data do not show whether teams grew in response to problems.

Many respondents selected “I do not know” for commercial attributes: 228 of 575 for pricing model, 258 of 597 for contract term, and 243 of 597 for relationship length. All 150 engagement managers and 31 portfolio directors answered whether the institution operated capability centers, compared with 168 of 667 individual contributors (25.19%). These differences

concern the information respondents supplied and may reflect role-specific access or responsibilities.

> *The playbook that exists is the old staffed-delivery playbook with the word agent inserted. Pricing was still constructed on headcount and hours."*
>
> **PORTFOLIO DIRECTOR**
> KPO, Corporate and back-office functions

Written accounts described preparation and pricing concerns that differ from the attributes tested statistically. Such accounts can identify possible mechanisms for further investigation, but they do not resolve inconclusive comparisons or demonstrate why outcomes differed.

> *The vendor priced the expense agent's throughput against the old FTE headcount baseline instead of actual hours saved, so the billed 'efficiency gain' stayed flat no matter how many expense reports the agent cleared on its own."*
>
> **INDIVIDUAL CONTRIBUTOR**
> BPO, Corporate and back-office functions

> *Pricing model still assumes FTE hours nobody is spending anymore"*
>
> **INDIVIDUAL CONTRIBUTOR**
> CXM, Technology engineering and operations

“

If you only read one line of this: we don't have the people. I was given three analysts for a diligence workstream that needs six, and two of them had never seen a data room before this.”

ENGAGEMENT MANAGER

KPO, Corporate and back-office functions

1 The Financial Institution

Institution type, asset size, capability-center presence, and reported expansion plans produced a mixture of equivalent and inconclusive estimates. Reported movement of work to capability centers was associated with weaker outcomes. Each result below is interpreted according to its own measure and sample.

Information availability varied by role. “I do not know” responses exceeded numeric ratings on reported growth in agentic engagements, and 334 of 576 respondents (57.99%) selected that response for institutional asset size. The pattern does not establish whether respondents needed that information to perform their duties.

1.1 Institution Type and Outcome Comparisons

Indicator C031 · What type of a bank is the customer? · Full statistical analysis ›

Institution-type comparisons supported equivalence for three core measures, while the active-engagement customer-satisfaction comparison remained inconclusive. The Engagements sheet classifies the financial institution served by each engagement into 10 types, from 24 engagements in asset management to nine in private equity and private capital. The type of financial institution showed no statistically significant association with cancellation (chi-square 5.71, df = 8, p = 0.693, n = 141 engagements). It showed no statistically significant association with the engagement indicator mean (Welch F 1.03, p = 0.425, n = 141 engagements) or with intent to leave (Welch F 0.83, p = 0.583, n = 141 engagements). Each of the three results was within the study's equivalence bounds. The test of customer satisfaction, on the four types with 10 or more active engagements, did not reach significance and was inconclusive (Welch F 2.39, p = 0.097, n = 46 engagements).

Table 1.3. The four core study measures by type of financial institution, 150 engagements

Type of financial institution	Engagements	Cancellation rate	Customer satisfaction
Asset management	24	13 of 24 (54.17%)	37.83
Wealth management and private banking	23	8 of 23 (34.78%)	44.74
Investment banking (advisory and capital-markets origination)	19	9 of 19 (47.37%)	34.84
Insurance (life, property and casualty, health)	14	5 of 14 (35.71%)	47.00
Commercial and business banking	13	4 of 13 (30.77%)	46.54
Payments and merchant acquiring (networks, acquiring, processing)	13	3 of 13 (23.08%)	49.77
Credit card and payments issuing	12	5 of 12 (41.67%)	39.33
Retail and consumer banking	12	4 of 12 (33.33%)	50.50
Markets and trading (sales and trading, global markets)	11	3 of 11 (27.27%)	49.91
Private equity and private capital	9	2 of 9	49.11

Indicator mean

Table 1.3. The four core study measures by type of financial institution, 150 engagements (continued)

Type of financial institution	Indicator mean	Intent to leave
Asset management	3.85	7.21
Wealth management and private banking	4.09	6.49
Investment banking (advisory and capital-markets origination)	3.42	7.66
Insurance (life, property and casualty, health)	4.00	6.91
Commercial and business banking	4.46	6.50
Payments and merchant acquiring (networks, acquiring, processing)	4.20	6.34
Credit card and payments issuing	3.69	6.47
Retail and consumer banking	4.43	5.76
Markets and trading (sales and trading, global markets)	4.38	6.71
Private equity and private capital	4.09	7.11

Note: Cancellation rate is canceled engagements over engagements of the type. Customer satisfaction, indicator mean, and intent to leave are means of the engagement values across all engagements of the type, canceled and active. Every type rests on fewer than 30 engagements. Private equity and private capital rest on nine engagements, below the base of 10 at which the study prints a rate. Its cancellation is shown as a count.

Investment banking engagements recorded the lowest customer satisfaction among active engagements (mean 56.40 points, n = 10 engagements). They also recorded the lowest engagement indicator mean (3.4, n = 19 engagements) and the highest intent to leave (7.7, n = 19 engagements). None of the three comparisons reached statistical significance.

Six written responses addressed the type of financial institution, and four were negative. Each negative response placed the difficulty in the vendor's preparation for that type of institution, not in the type itself. No written response provided additional comparative performance analysis between financial institution types.

> *"The spend on training is real, I will grant that, but the courses are generic and nothing in them touches an asset manager's book."*
>
> **PORTFOLIO DIRECTOR**
>
> KPO, Trade, settlement and reconciliation

> *"On the estate they are no better or worse than most asset managers we deal with at this stage. Their people answered when we raised things, and the business case never wobbled."*
>
> **PORTFOLIO DIRECTOR**
>
> CXM, Corporate and back-office functions

1.2 Reported Knowledge of Financial Institution Asset Size

Indicator C062 · What is the customer's approximate total asset size, in USD? · Full statistical analysis ›

Most respondents answering the question about the financial institution's approximate total asset size did not know it. The question was answered by 576 of the 850 respondents, and 334 of the 576 (57.99%) answered "I do not know." Those answers came from 141 of the 150 engagements. Among the 242 respondents who gave a listed size, the most frequent answer was $250 billion to under $1 trillion (90 respondents, 37.19%).

The proportion supplying an asset-size category increased with seniority: 146 of 667 individual contributors (21.89%), 64 of 150 engagement managers (42.67%), and 30 of 31 portfolio directors (96.77%). With "I do not know" retained as a response category, the answer showed no statistically significant association with cancellation (chi-square 3.96, df = 4, p = 0.411, n = 115 engagements). It showed no statistically significant association with customer satisfaction (F 1.48, df = 4 and 110, p = 0.215, n = 115 engagements), the engagement indicator mean (F 0.91, df = 4 and 110, p = 0.460, n = 115 engagements), or intent to leave (F 0.93, df = 4 and 110, p = 0.449, n = 115 engagements). The results for cancellation, the engagement indicator mean, and intent to leave were within the study's equivalence bounds, and the customer-satisfaction

result was inconclusive. Only one listed size category held 10 or more engagements, so the listed categories were not compared with one another.

1.3 Financial Institution Capability Centers as Reported by the Delivery Team

Indicator C063 · Does the customer operate its own global capability centers? · Full statistical analysis ›

Knowledge of whether the financial institution operated global capability centers varied substantially by respondent role. The question was answered by 596 of the 850 respondents, and 247 of the 596 (41.44%) answered: “I do not know.” The financial institution operated its own capability centers on the account of 190 of the 349 respondents who answered yes or no (54.44%).

Figure 1.16. Share answering Yes to “Does the customer operate its own global capability centers?”, by respondent role and engagement status

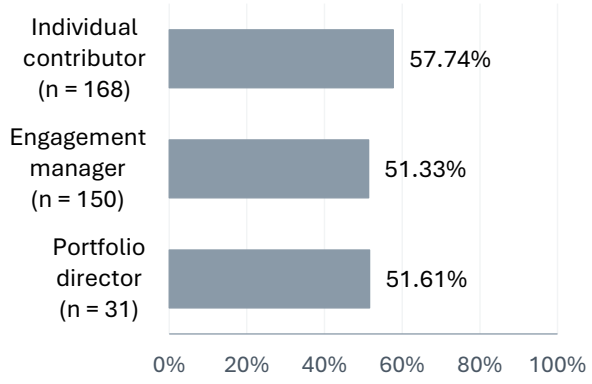
Availability of Yes/No answers varied by respondent role

596 of 850 respondents answered the question. Of these answers, 247 were “I do not know” and 349 were Yes or No.

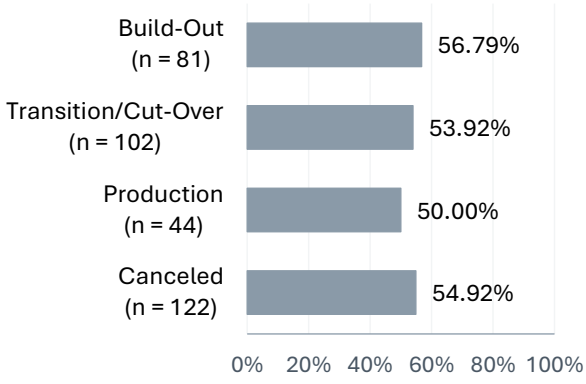
The percentage panels below concern Yes answers among valid Yes/No answers. They do not measure the proportion of each role able to answer.

Respondent role	Valid Yes/No answers	All respondents in role
Individual contributor	168	667
Engagement manager	150	150
Portfolio director	31	31

Respondent role
Share answering Yes



Lifecycle stage / status
Share answering Yes



Note: True or Yes answers: 190 of 349 answers (54.44%).

n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

All 150 engagement managers and 31 portfolio directors gave a yes-or-no answer, compared with 168 of 667 individual contributors (25.19%). Among the 78 engagements assigned a yes or no answer, reported capability-center presence showed no statistically significant association with cancellation (chi-square 0.60, df = 1, p = 0.439, n = 78 engagements). It showed no statistically significant association with the engagement indicator mean (Welch F 0.14, p = 0.711, n = 78 engagements) or with intent to leave (Welch F 0.02, p = 0.892, n = 78 engagements). Each of the three results was within the study's equivalence bounds. The comparison of customer satisfaction among active engagements did not reach significance and was inconclusive (Welch F 2.15, p = 0.148, n = 53 engagements). Presence and reported transfer of work are separate constructs and should not be conflated.

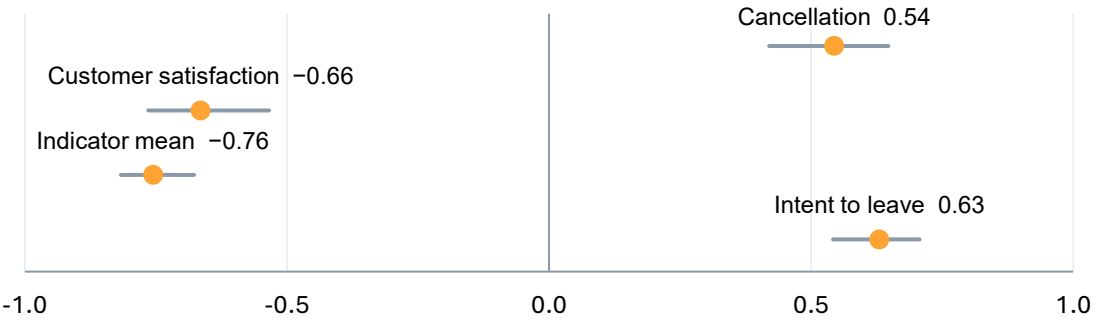
1.4 Shift of Work to Financial Institution Capability Centers and Uniformly Weaker Outcomes

Indicator C068 · The customer is shifting more work to their global capability centers · Full statistical analysis ›

Stronger agreement that the financial institution was shifting work to its own global capability centers was associated with less favorable outcomes on every core measure. The question's scale had neither a favorable nor an unfavorable end. The mean rating was 6.1, and the median 6.0 (n = 341 respondents), and 246 respondents answered: “I do not know.”

Figure 1.17. Shift of Work to Financial Institution Capability Centers and Uniformly Weaker Outcomes: associations with the study measures

Shift of work to financial institution capability centers and uniformly weaker outcomes: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = 0.54 [0.42, 0.65] 150 engagements p < 0.001
Customer satisfaction	r = −0.66 [−0.76, −0.53] 94 active engagements p < 0.001
Indicator mean	r = −0.76 [−0.82, −0.68] 150 engagements p < 0.001
Intent to leave	r = 0.63 [0.54, 0.71] 323 respondents p < 0.001

Note: mean rating: 6.1 (n = 341 respondents).

Survey question: “The customer is shifting more work to their global capability centers”.

Group means and numeric-answer bases: Individual contributor (n = 160): 6.53. Engagement manager (n = 150): 5.93. Portfolio director (n = 30): 4.83. Build-Out (n = 73): 5.42. Transition/Cut-Over (n = 101): 5.62. Production (n = 46): 5.13. Canceled (n = 121): 7.31.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Stronger agreement was associated with worse outcomes on every measure, beginning with more cancellations (r = 0.54, 95% CI 0.42 to 0.65, n = 150 engagements). Agreement was associated with lower customer satisfaction among active engagements (r = −0.66, 95% CI −0.76 to −0.53, n = 94 engagements). It was also associated with a lower engagement indicator mean (r = −0.76, 95% CI −0.82 to −0.68, n = 150 engagements). Respondents who agreed more strongly reported higher intent to leave (r = 0.63, cluster bootstrap 95% CI 0.54 to 0.71, n = 323 respondents). The applicable robustness checks retained the associations, including active-only checks for the indicator mean and intent to leave. Cancellation itself is constant within active engagements. Agreement was strongest on canceled engagements and among individual contributors, and weakest among portfolio directors.

The finding has a clear limit. The rating records how respondents perceived the financial institution's plans, not a measured movement of work. The study holds no evidence on which came first. Work may have moved in-house before the engagement weakened, or after it. No written response addressed the indicator.

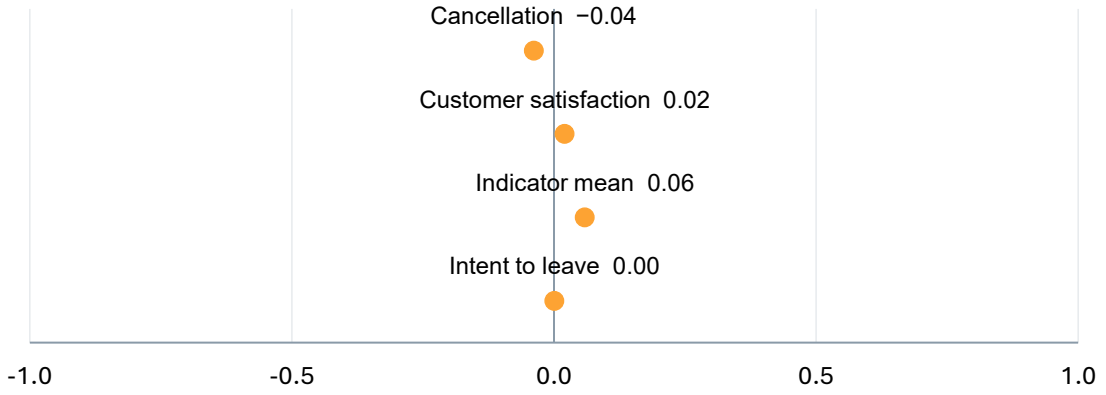
1.5 Reported Scaling of Agentic Outsourcing

Indicator C066 · The customer is scaling their agentic AI outsourcing engagements · Full statistical analysis ›

The estimated relationships between reported scaling of the financial institution's agentic AI outsourcing engagements and the core measures met the study's equivalence criteria. The question's scale had neither a favorable nor an unfavorable end. The mean rating was 5.5, and the median 5.0 (n = 364 respondents), and 256 respondents answered: “I do not know.”

Figure 1.18. Reported Scaling of Agentic Outsourcing: associations with the study measures

Reported scaling of agentic outsourcing: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.04 150 engagements p = 0.640
Customer satisfaction	r = 0.02 94 active engagements p = 0.846
Indicator mean	r = 0.06 150 engagements p = 0.476
Intent to leave	r = 0.00 352 respondents p = 0.990

Note: mean rating: 5.5 (n = 364 respondents).

Survey question: “The customer is scaling their agentic AI outsourcing engagements”.

Group means and numeric-answer bases: Individual contributor (n = 183): 5.48. Engagement manager (n = 150): 5.53. Portfolio director (n = 29): 5.55. Build-Out (n = 75): 5.29. Transition/Cut-Over (n = 115): 5.68. Production (n = 46): 5.48. Canceled (n = 128): 5.49.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Estimated associations of reported scaling met the equivalence criterion for all four core measures. The correlation with cancellation was r = −0.04 (p = 0.640, n = 150 engagements). It was r = 0.02 with customer satisfaction among active engagements (p = 0.846, n = 94 engagements) and r = 0.06 with the engagement indicator mean (p = 0.476, n = 150 engagements). At respondent level, it was r = 0.00 with intent to leave (p = 0.990, n = 352 respondents). These estimates concern perceived financial institution plans, not measured changes in outsourcing volume.

1.6 Reported Growth in the Number of Engagements Unknown to Most Respondents

Indicator C067 · The customer is increasing the number of agentic AI outsourcing engagements · Full statistical analysis ›

More respondents did not know whether the financial institution was increasing the number of its agentic AI outsourcing engagements than supplied a rating, 345 against 247. The question's scale had neither a favorable nor an unfavorable end. The mean rating was 5.3 and the median 5.0 (n = 247 respondents).

Figure 1.19. Reported Growth in the Number of Engagements Unknown to Most Respondents: associations with the study measures

Reported growth in the number of engagements unknown to most respondents: associations with the study measures

Comparison	Estimate and analysis base
Cancellation	r = 0.04 127 engagements p = 0.635
Customer satisfaction	r = −0.07 77 active engagements p = 0.563
Indicator mean	r = −0.08 127 engagements p = 0.392
Intent to leave	r = 0.05 235 respondents p = 0.442

Note: mean rating: 5.3 (n = 247 respondents).

Survey question: "The customer is increasing the number of agentic AI outsourcing engagements".

Group means and numeric-answer bases: Individual contributor (n = 162): 5.36. Engagement manager (n = 52): 5.38. Portfolio director (n = 31): 5.29. Build-Out (n = 59): 5.27. Transition/Cut-Over (n = 63): 5.19. Production (n = 29): 5.66. Canceled (n = 96): 5.40.

r is Pearson's correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The estimated associations of reported growth met the equivalence criterion for all four core measures. The correlation with cancellation was r = 0.04 (p = 0.635, n = 127 engagements). It

was r = −0.07 with customer satisfaction among active engagements (p = 0.563, n = 77 engagements) and r = −0.08 with the engagement indicator mean (p = 0.392, n = 127 engagements). At respondent level, it was r = 0.05 with intent to leave (p = 0.442, n = 235 respondents). The high number of unknown responses limits conclusions about the financial institution's wider program.

2 The Relationship Between the Two Firms

Relationship length and the number of concurrent agentic engagements describe different aspects of the relationship between the firms. Most estimated associations met the equivalence criteria, but the customer-satisfaction result for relationship length was inconclusive. Conflicting accounts of relationship duration also limit interpretation.

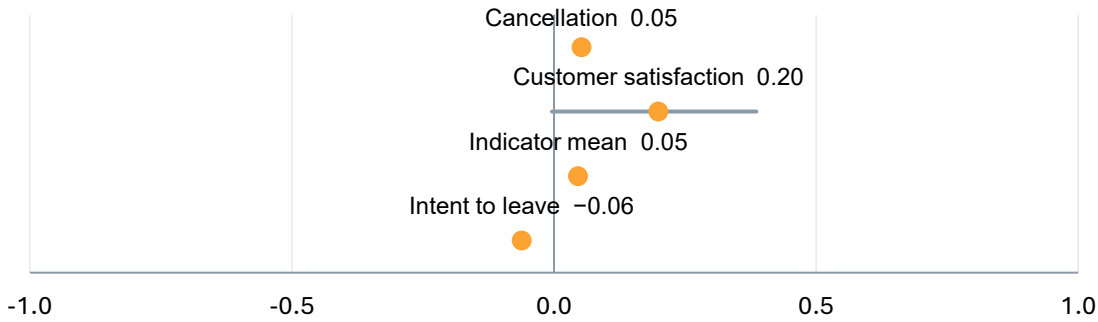
2.1 Relationship Length and Outcome Comparisons

Indicator C064 · How many months has your employer been serving this customer across all work, not just this engagement? · Full statistical analysis ›

The length of the vendor's relationship with the financial institution met the study's equivalence criteria for cancellation, engagement indicator mean, and intent to leave, while the active-engagement customer-satisfaction result was inconclusive. The question covered the vendor's relationship across all services and was answered in months. The mean answer was 32.6 months and the median 33.0 months, ranging from 3.0 to 67.0 months (n = 354 respondents), and 243 respondents answered: “I do not know.”

Figure 1.20. Relationship Length and Outcome Comparisons: associations with the study measures

Relationship length and outcome comparisons: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = 0.05 150 engagements p = 0.524
Customer satisfaction	r = 0.20 [−0.00, 0.39] 94 active engagements p = 0.055
Indicator mean	r = 0.05 150 engagements p = 0.576
Intent to leave	r = −0.06 340 respondents p = 0.467

Note: Mean months the vendor has served the institution: 32.6 (n = 354 respondents).

Group means and numeric-answer bases: Individual contributor (n = 172): 33.38. Engagement manager (n = 150): 31.79. Portfolio director (n = 31): 32.06. Build-Out (n = 69): 32.13. Transition/Cut-Over (n = 101): 28.83. Production (n = 48): 37.35. Canceled (n = 136): 34.07.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The length of the relationship showed no statistically significant association with cancellation (r = 0.05, p = 0.524, n = 150 engagements). The engagement indicator mean showed no

statistically significant association with it either (r = 0.05, p = 0.576, n = 150 engagements). No statistically significant association was detected for intent to leave (r = −0.06, p = 0.467, n = 340 respondents), and each of the three results was within the study's equivalence bounds. The association with customer satisfaction among active engagements was positive and did not reach significance (r = 0.20, 95% CI 0.00 to 0.39, p = 0.055, n = 94 engagements). Two robustness checks did not hold, and the result is inconclusive.

Five written responses addressed relationship length. Three described relationships lasting 14 years, 15 years, and two decades, exceeding the maximum structured answer of 67 months. The difference remains unresolved and may reflect inconsistent units or interpretations of the relationship. The structured estimates are retained as recorded, and the longer written accounts are not treated as corroborating the same duration measure.

> *"Fourteen years of relationship equity gone in one incident"*
>
> **PORTFOLIO DIRECTOR**
> CXM, Client servicing and communications

> *"Fifteen years of a dependable client relationship"*
>
> **ENGAGEMENT MANAGER**
> CXM, Corporate and back-office functions

> *"two decades of relationship history to draw on"*
>
> **ENGAGEMENT MANAGER**
> KPO, Trade, settlement and reconciliation

2.2 Concurrent Agentic Engagements With the Same Financial Institution

Indicator C065 · How many agentic AI outsourcing engagements does your employer run concurrently with this customer today? · Full statistical analysis ›

Most respondents answering the question about concurrent agentic AI outsourcing engagements between the vendor and the financial institution did not know the number. More respondents answered “I do not know” than gave a number, 362 against 254. Among those who gave a number, the mean was 17.2 and the median 18.0, from 2.0 to 30.0 (n = 254 respondents on 129 engagements).

Figure 1.21. Concurrent Agentic Engagements With the Same Financial Institution: associations with the study measures

Concurrent agentic engagements with the same financial institution: associations with the study measures

Comparison	Estimate and analysis base
Cancellation	r = 0.10 129 engagements p = 0.267
Customer satisfaction	r = 0.03 83 active engagements p = 0.804
Indicator mean	r = 0.03 129 engagements p = 0.716
Intent to leave	r = 0.04 243 respondents p = 0.691

Note: Mean concurrent agentic engagements with the institution: 17.2 (n = 254 respondents).

Group means and numeric-answer bases: Individual contributor (n = 163): 17.61. Engagement manager (n = 58): 16.36. Portfolio director (n = 31): 16.26. Build-Out (n = 56): 15.75. Transition/Cut-Over (n = 78): 17.55. Production (n = 26): 14.00. Canceled (n = 94): 18.65.

r is Pearson’s correlation coefficient, measuring signed linear association from –1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The count of concurrent engagements was within the study’s equivalence bounds against each of the four study measures. It recorded r = 0.10 with cancellation (p = 0.267, n = 129

engagements) and r = 0.03 with customer satisfaction among active engagements (p = 0.804, n = 83 engagements). With the engagement indicator mean it recorded r = 0.03 (p = 0.716, n = 129 engagements), and with intent to leave r = 0.04 (p = 0.691, n = 243 respondents). The scale of the two firms' agentic work together did not distinguish the engagements that held from those that were canceled.

“

Funding-cutoff timing reports showed settlements completing on schedule because the metric only counted the final submission step, not the manual keyboard intervention required upstream every time the platform rejected an agent session. Exception-queue counts were reported net of items a human quietly resolved off the books before the daily report ran.”

ENGAGEMENT MANAGER

BPO, Trade, settlement and reconciliation

3 The Work Outsourced

The sample covered contact-center work, trade settlement, identity management, financial-crime screening, investor communications, and other processes. Outsourcing model and class-of-work comparisons met the study's equivalence criteria. Process-family results were mixed, including inconclusive cancellation and intent-to-leave comparisons. Delivery location is descriptive only.

Written responses also describe why respondents valued particular tasks. Their positive content must be interpreted in light of the question asking what they valued most. It is not an independent comparison of all possible sources of job satisfaction.

3.1 Outsourcing Model and Outcome Comparisons

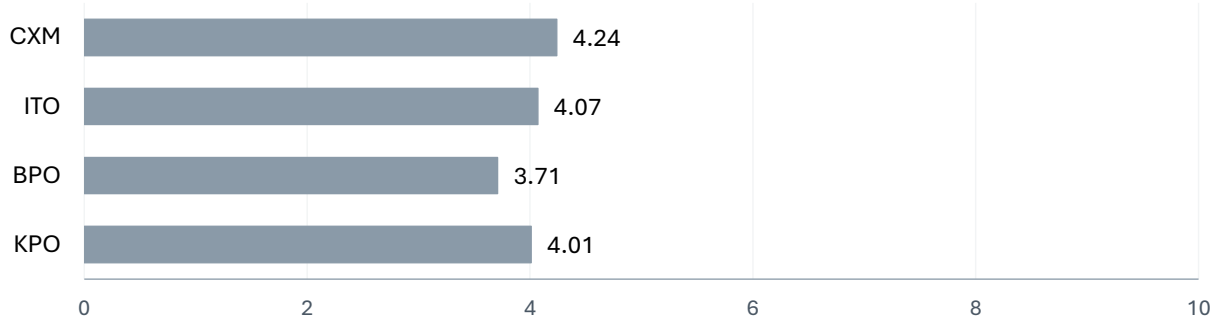
Indicator C032 · What is the outsourcing type? · Full statistical analysis ›

The Engagements sheet classifies each engagement by outsourcing model, with CXM on 42 engagements, ITO on 40, BPO on 37, and KPO on 31. Every respondent's answer matched The study. No statistically significant differences by outsourcing model were detected for the four study measures, and each result was within the study's equivalence bounds. Cancellation recorded chi-square 0.08 (p = 0.994, n = 150 engagements), and customer satisfaction among active engagements Welch F 0.78 (p = 0.510, n = 94 engagements). The engagement indicator mean recorded Welch F 0.90 (p = 0.444, n = 150 engagements), and intent to leave Welch F 0.47 (p = 0.703, n = 150 engagements).

Table 1.4. The four core study measures by outsourcing model, 150 engagements

Outsourcing model	Engagements	Cancellation rate	Customer satisfaction
CXM	42	15 of 42 (35.71%)	44.43
ITO	40	15 of 40 (37.50%)	44.20
BPO	37	14 of 37 (37.84%)	45.57
KPO	31	12 of 31 (38.71%)	40.58

Indicator mean



Outsourcing model	Indicator mean	Intent to leave
CXM	4.24	6.48
ITO	4.07	7.09
BPO	3.71	6.72
KPO	4.01	6.81

Note: Cancellation rate is canceled engagements over the engagements in the row. Customer satisfaction, indicator mean, and intent to leave are means of the engagement values across all engagements in the row, canceled and active.

The models differed in vendor and process composition, while the tested outcome contrasts met the equivalence criteria. One engagement manager on a canceled engagement described a mismatch between the commercial model and staffing arrangements.

> *"Training on funds and NAV work is decent, I will give them that. On agents the knowledge is thin. We were sold as an AI delivery and staffed as an ordinary BPO."*
>
> **ENGAGEMENT MANAGER**
>
> BPO, Trade, settlement and reconciliation

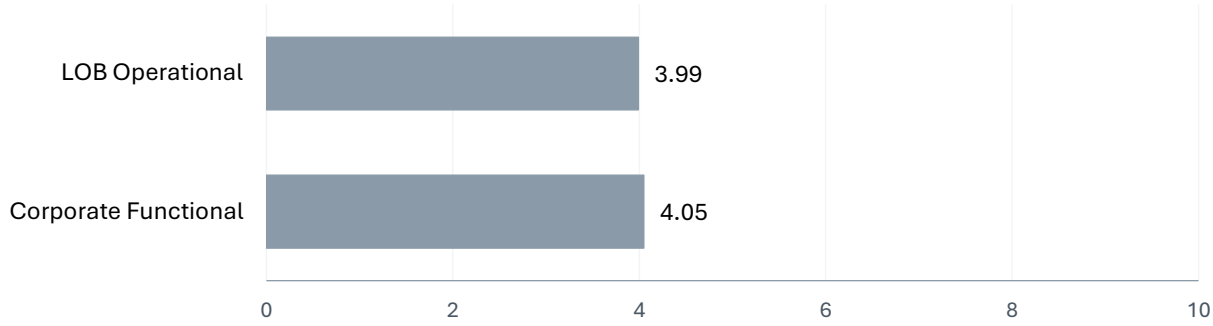
3.2 Class of Work and Outcome Comparisons

Indicator C033 · Is the outsourcing engagement Corporate Functional, or Line of Business Operational? · Full statistical analysis ›

Table 1.5. The four core study measures by class of work, 150 engagements

Class of work	Engagements	Cancellation rate	Customer satisfaction
LOB Operational	76	30 of 76 (39.47%)	42.54
Corporate Functional	74	26 of 74 (35.14%)	45.20

Indicator mean



Class of work	Indicator mean	Intent to leave
LOB Operational	3.99	6.85
Corporate Functional	4.05	6.69

Note: Cancellation rate is canceled engagements over the engagements in the row. Customer satisfaction, indicator mean, and intent to leave are means of the engagement values across all engagements in the row, canceled and active.

The Engagements sheet classifies each engagement as serving a line-of-business operation (76 engagements) or a corporate function (74 engagements). The respondents' answers on the class of work agreed with the study in every case. No statistically significant differences between the two classes were detected on the four study measures, and every result was within the study's equivalence bounds. The test of cancellation gave chi-square 0.30 (p = 0.583, n = 150 engagements), and the test of customer satisfaction among active engagements gave Welch F 0.05 (p = 0.828, n = 94 engagements). The engagement indicator mean gave Welch F 0.07 (p = 0.795, n = 150 engagements), and intent to leave Welch F 0.17 (p = 0.678, n = 150 engagements).

The class followed the process served. Corporate and back-office and technology engagements were mostly corporate functions, while servicing, trade, and payments engagements were mostly line-of-business operations. Whether an engagement served the business directly or a function behind it showed no statistically significant association with its outcomes.

3.3 Study Use Cases

ENGAGEMENT FIELD · USE CASE · FULL STATISTICAL ANALYSIS ›

Each engagement's use case is assigned to one of eight process families, from 47 engagements in corporate and back-office functions to seven in financial crime and client onboarding. The process family showed no statistically significant association with customer satisfaction among active engagements (Welch F 0.16, p = 0.850, n = 63 engagements). It showed no statistically significant association with the engagement indicator mean either (Welch F 1.20, p = 0.325, n = 125 engagements), and both results were within the study's equivalence bounds. The test of cancellation (chi-square 7.48, df = 4, p = 0.113, n = 125 engagements) did not reach significance and was inconclusive. The test of intent to leave (Welch F 1.77, p = 0.155, n = 125 engagements) was inconclusive as well.

Table 1.6. The four core study measures by process family, 150 engagements

Process family	Engagements	Cancellation rate	Customer satisfaction
Corporate and back-office functions	47	19 of 47 (40.43%)	41.68
Technology engineering and operations	35	10 of 35 (28.57%)	49.91
Client servicing and communications	16	9 of 16 (56.25%)	34.25
Trade, settlement and reconciliation	15	8 of 15 (53.33%)	35.53
Payments, cards and disputes	12	2 of 12 (16.67%)	56.08
Credit, underwriting and suitability	9	1 of 9	57.78
Regulatory, audit, tax and legal	9	4 of 9	37.33
Financial crime and client onboarding	7	3 of 7	37.43

Indicator mean

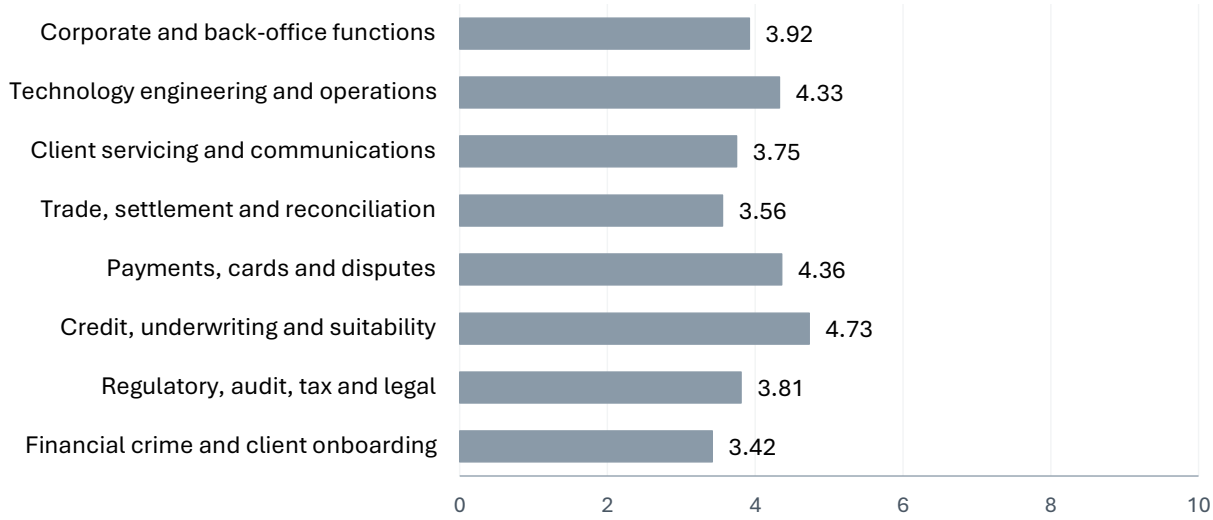


Table 1.6. The four core study measures by process family, 150 engagements (continued)

Process family	Indicator mean	Intent to leave
Corporate and back-office functions	3.92	6.90
Technology engineering and operations	4.33	6.55
Client servicing and communications	3.75	7.40
Trade, settlement and reconciliation	3.56	7.54
Payments, cards and disputes	4.36	5.25
Credit, underwriting and suitability	4.73	6.15
Regulatory, audit, tax and legal	3.81	7.18
Financial crime and client onboarding	3.42	6.78

Note: Cancellation rate is canceled engagements over the engagements in the row. Customer satisfaction, indicator mean, and intent to leave are means of the engagement values across all engagements in the row, canceled and active. Credit, underwriting and suitability, Regulatory, audit, tax and legal and Financial crime and client onboarding rest on fewer than 10 engagements, the base below which the study prints no rate. Cancellation is shown as a count.

The written responses that addressed the use case numbered 33, and 27 of the 33 responses were positive (81.82%), among the most favorable written evidence in the study. Most answered the question on what the respondent loved most about the job and named the use case or the problem it addressed. Respondents on canceled engagements valued the problem they had been asked to solve as well.

> *The fraud detection problem itself was worth solving"*
>
> **ENGAGEMENT MANAGER**
> ITO, Financial crime and client onboarding

> *Solving identity sprawl at this scale is real work worth doing"*
>
> **INDIVIDUAL CONTRIBUTOR**
> ITO, Technology engineering and operations

> *Genuinely hard problem, keeps me sharp"*
>
> **INDIVIDUAL CONTRIBUTOR**
> ITO, Technology engineering and operations

3.4 Delivery Location

Delivery was located in India on 92 of the 150 engagements and in the United States on 19, with the remaining 39 delivered from other locations. Delivery location showed no statistically significant association with cancellation (chi-square 0.47, df = 2, p = 0.790, n = 150 engagements). It showed no statistically significant association with customer satisfaction (F 0.34, df = 2 and 147, p = 0.711, n = 150 engagements), the engagement indicator mean (F 0.46, df = 2 and 147, p = 0.630, n = 150 engagements), or intent to leave (F 0.40, df = 2 and 147, p = 0.670, n = 150 engagements). Each of the four results was within the study's equivalence bounds, so outcomes did not differ by delivery location.

“

...nobody in my batch knew anything about agents when we joined, and the company did put us on a course. Two weeks. It was not enough for AML work but it was something, and the trainer did answer questions after also. The problem is nobody above the trainer has done this before either.”

INDIVIDUAL CONTRIBUTOR

ITO, Financial crime and client onboarding

4 Commercial Terms

No statistically significant association was detected in the reported tests of pricing model, contract term, or annual value, but several estimates were inconclusive and customer satisfaction could not be tested for pricing model. Written pricing accounts address perceived incentives rather than providing an alternative statistical test of those associations.

4.1 Pricing on Human Effort for Work Performed by Agents

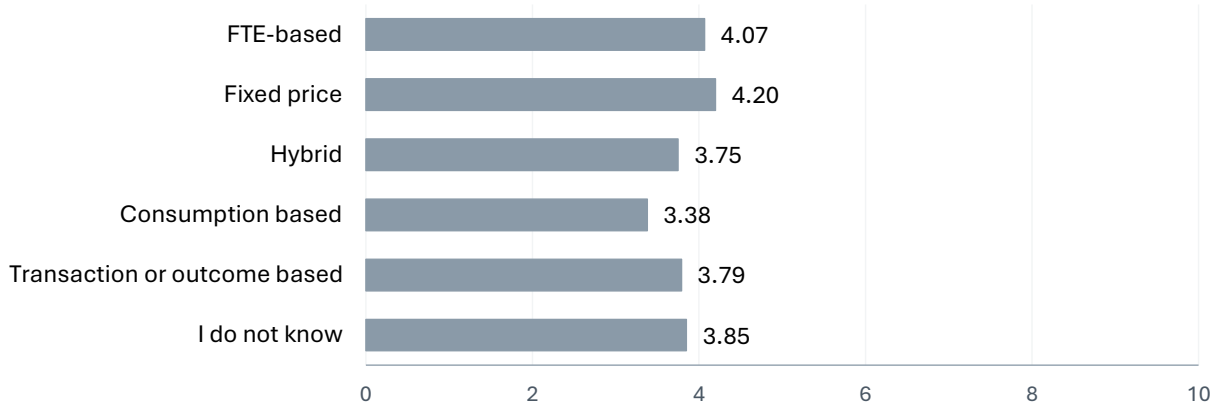
Indicator C044 · What is the pricing model for this engagement? · Full statistical analysis ›

Of 575 respondents answering the pricing question, 228 selected “I do not know.” Among the 347 naming a model, 140 reported full-time-equivalent staffing as the pricing basis (40.35%). Tested differences did not reach statistical significance for cancellation (chi-square 2.97, simulated p = 0.413, n = 71 engagements), engagement indicator mean (Welch F 1.31, p = 0.290, n = 71), or intent to leave (Welch F 0.79, p = 0.511, n = 71). Only the last met the equivalence criterion. Customer satisfaction was not tested.

Table 1.7. The four core study measures by pricing model, 107 engagements

Pricing model	Engagements	Cancellation rate	Customer satisfaction
FTE-based	35	11 of 35 (31.43%)	47.80
Fixed price	13	5 of 13 (38.46%)	44.00
Hybrid	13	4 of 13 (30.77%)	43.77
Consumption based	10	6 of 10 (60.00%)	29.00
Transaction or outcome based	8	6 of 8	25.00
I do not know	28	12 of 28 (42.86%)	41.68

Indicator mean



Pricing model	Indicator mean	Intent to leave
FTE-based	4.07	6.50
Fixed price	4.20	6.63
Hybrid	3.75	6.59
Consumption based	3.38	7.71
Transaction or outcome based	3.79	8.44
I do not know	3.85	6.82

Note: Cancellation rate is canceled engagements over the engagements in the row. Customer satisfaction, indicator mean, and intent to leave are means of the engagement values across all engagements in the row, canceled and active. Each engagement carries the pricing answer assigned to it from its respondents' answers. Transaction or outcome based rests on fewer than 10 engagements, the base below which the study prints no rate. Cancellation is shown as a count.

All 33 written pricing responses were negative. Twenty-one described pricing or billing based on headcount, seats, or hours while agents performed some tasks. At least three respondents described a perceived incentive conflict in which automation reduced billable human effort without providing a corresponding commercial benefit to the vendor. These accounts concern particular arrangements. They do not establish a general effect of staffing-based pricing.

> *"As long as this is billed on heads, every hour an agent saves is an hour I am asked to explain, and on that basis there is no incentive anywhere in my chain to let the thing work properly."*
>
> **ENGAGEMENT MANAGER**
>
> BPO, Trade, settlement and reconciliation

> *"We are still sold and staffed as a headcount deal, forty seats on a monitoring queue, while the deck that won the work promises the agents do the first pass. So when the agents do less than that, the only lever I have is more analysts, which I cannot get, and when they do more, we are billing for seats nobody needs."*
>
> **ENGAGEMENT MANAGER**
>
> BPO, Financial crime and client onboarding

> *The vendor billed the engagement on a per-FTE-equivalent basis calculated from pre-agent staffing levels, so the invoice looked justified on paper even though the agent had already absorbed most of the manual expense-processing work months earlier."*

INDIVIDUAL CONTRIBUTOR

BPO, Corporate and back-office functions

> *We must stop pricing this as if it were bodies on a floor."*

PORTFOLIO DIRECTOR

ITO, Technology engineering and operations

4.2 Contract Term and Outcome Associations

Indicator C045 · What is the contract term of this engagement, in months? · Full statistical analysis ›

Reported contract terms averaged 30.9 months, with a median of 36 months, or three years, and a range of 12–60 months (n = 339 respondents on 150 engagements). Another 258 respondents selected “I do not know.”

Figure 1.22. Contract Term and Outcome Associations: associations with the study measures

Contract term and outcome associations

Comparison	Estimate and analysis base
Cancellation	r = −0.10 150 engagements p = 0.223
Customer satisfaction	r = 0.03 94 active engagements p = 0.744
Indicator mean	r = 0.08 150 engagements p = 0.302
Intent to leave	r = −0.02 327 respondents p = 0.781

Note: Mean contract term in months: 30.9 (n = 339 respondents).

Group means and numeric-answer bases: Individual contributor (n = 156): 31.38. Engagement manager (n = 150): 30.48. Portfolio director (n = 31): 30.58. Build-Out (n = 73): 32.55. Transition/Cut-Over (n = 95): 32.08. Production (n = 43): 29.30. Canceled (n = 128): 29.62.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Each association of the contract term with the four study measures was within the study's equivalence bounds. Its correlation with cancellation was r = −0.10 (p = 0.223, n = 150 engagements), and with customer satisfaction among active engagements r = 0.03 (p = 0.744, n = 94 engagements). With the engagement indicator mean it was r = 0.08 (p = 0.302, n = 150 engagements), and with intent to leave r = −0.02 (p = 0.781, n = 327 respondents). A longer contract term showed no statistically significant association with less cancellation.

4.3 Contract Value Unknown to Most of the Delivery Team

Indicator C046 · What is the annual contract value of this engagement, in USD? · Full statistical analysis ›

Most respondents answering the question about annual engagement contract value did not know it, and knowledge increased with seniority. The question was answered by 601 of the 850 respondents, and 353 of the 601 (58.74%) answered "I do not know." Among the 248 who gave a listed value, 106 (42.74%) answered $1 million to under $5 million and 102 (41.13%) $5 million to under $10 million. The share who gave a value was greater in each more senior role, from 152 of the 667 individual contributors (22.79%) to 30 of the 31 portfolio directors (96.77%).

At engagement level, no statistically significant association was detected for contract value, and most results were inconclusive on a base of 31 engagements. Respondent-level comparisons retained "I do not know" as a category and differed for cancellation (F = 3.74, p = 0.006, n = 601) and the respondent indicator mean (F = 3.06, p = 0.018, n = 601). Respondents selecting the largest contract-value category were most often on canceled engagements, but both extreme categories contained fewer than 30 respondents. The results require caution and do not establish a relationship between contract value and performance.

5 The Vendor

Three survey questions described the firm that delivered each engagement. They asked for the type of company, the number of its employees worldwide, and the number of agentic AI engagements it runs for financial institutions. No statistically significant cancellation differences were detected for these attributes. Each was tested against cancellation, customer satisfaction among active engagements, the engagement indicator mean, and intent to leave, and all 12 results were within the study's equivalence bounds. The results do not establish identical outcomes across vendor types. They showed no statistically significant association with the size of the vendor or with the breadth of its agentic work for financial institutions.

For vendor size, 260 of 798 respondents answering the question selected "I do not know" (32.58%). For the vendor's number of financial-institution agentic engagements, 328 of all 850 respondents supplied a number (38.59%). These figures describe response coverage, not verified ignorance of the vendor's business.

5.1 Vendor Type and Outcome Comparisons

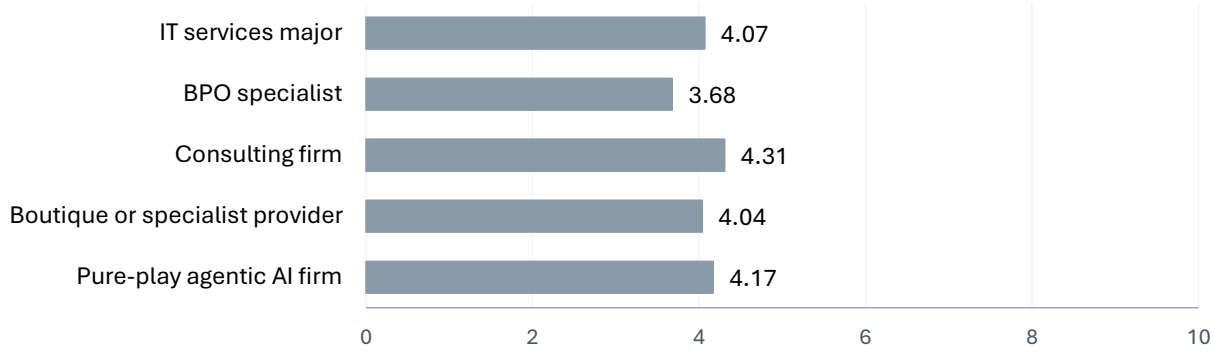
Indicator C076 · What type of company is your employer? · Full statistical analysis ›

IT services majors accounted for the largest reported share of vendors, followed by BPO specialists. An IT services major was named in 316 of the 793 answers (39.85%), and a BPO specialist in 198 (24.97%). A consulting firm was named in 117 of the 793 answers (14.75%), a boutique or specialist provider in 87 (10.97%), and a pure-play agentic AI firm in 75 (9.46%). The respondents on each engagement gave the same answer, and the question classifies the 150 engagements as well. IT services majors delivered 59 of the 150 engagements (39.33%) and pure-play agentic AI firms 14 of the 150 (9.33%).

Table 1.8. The four core study measures by vendor type, 150 engagements

Vendor type	Engagements	Cancellation rate	Customer satisfaction
IT services major	59	19 of 59 (32.20%)	46.92
BPO specialist	37	17 of 37 (45.95%)	37.22
Consulting firm	23	8 of 23 (34.78%)	44.04
Boutique or specialist provider	17	7 of 17 (41.18%)	44.41
Pure-play agentic AI firm	14	5 of 14 (35.71%)	47.50

Indicator mean



Vendor type	Indicator mean	Intent to leave
IT services major	4.07	6.62
BPO specialist	3.68	7.11
Consulting firm	4.31	6.59
Boutique or specialist provider	4.04	6.99
Pure-play agentic AI firm	4.17	6.57

Note: Cancellation rate is canceled engagements over the engagements in the row. Customer satisfaction, indicator mean, and intent to leave are means of the engagement values across all engagements in the row, canceled and active. Each engagement carries the vendor type assigned to it from its respondents' answers.

The engagement-level comparisons did not detect statistically significant differences by vendor type, and each met the equivalence criterion: cancellation chi-square 2.02 ($p = 0.731$, $n = 150$ engagements), active-engagement customer satisfaction Welch F 0.76 ($p = 0.524$, $n = 85$), engagement indicator mean Welch F 0.74 ($p = 0.573$, $n = 150$), and engagement-level intent to leave Welch F 0.31 ($p = 0.872$, $n = 150$). These tests differ from respondent-level vendor-type comparisons used in individual indicator analyses. Vendor types differed in their outsourcing models and process families without a detected outcome difference under the reported tests.

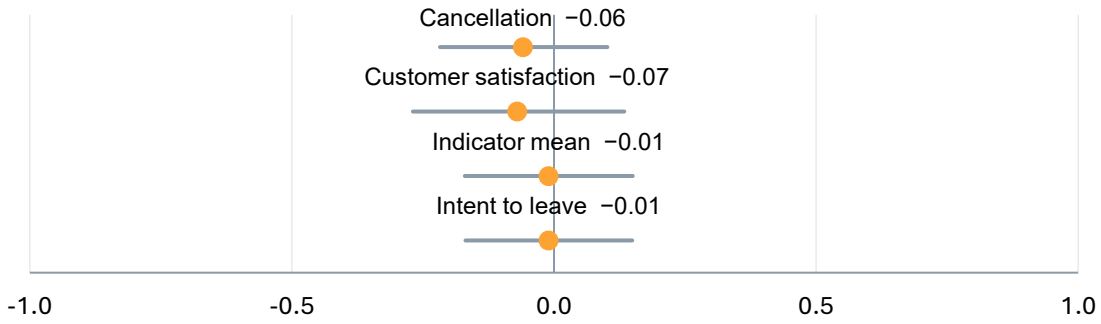
5.2 Vendor Size and Outcome Associations

Indicator C077 · Approximately how many employees does your employer have worldwide? · Full statistical analysis ›

Respondents described vendors with widely differing worldwide headcounts, but their estimates were not independently verified. The mean was approximately 110,955 employees, the median 90,854, and the range 321–340,114 (n = 538 respondents on 150 engagements). Another 260 respondents selected “I do not know,” including at least one respondent on 127 of 150 engagements (84.67%). The estimates were not independently verified against employer records.

Figure 1.23. Vendor Size and Outcome Associations: associations with the study measures

Vendor size and outcome associations



Comparison	Estimate and analysis base
Cancellation	r = −0.06 [−0.22, 0.10] 150 engagements p = 0.471
Customer satisfaction	r = −0.07 [−0.27, 0.13] 94 active engagements p = 0.500
Indicator mean	r = −0.01 [−0.17, 0.15] 150 engagements p = 0.901
Intent to leave	r = −0.01 [−0.17, 0.15] 508 respondents p = 0.901

Note: Mean vendor employees worldwide: 110,955 (n = 538 respondents).

Group means and numeric-answer bases: Individual contributor (n = 356): 111,081.72. Engagement manager (n = 150): 108,780.67. Portfolio director (n = 31): 122,090.48. Build-Out (n = 121): 105,430.31. Transition/Cut-Over (n = 152): 98,560.89. Production (n = 70): 161,431.73. Canceled (n = 195): 105,924.09.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Vendor size recorded results within the study's equivalence bounds on all four study measures. Against cancellation, it recorded r = −0.06 (95% CI −0.22 to 0.10, p = 0.471, n = 150 engagements). Against customer satisfaction among active engagements, it recorded r = −0.07 (95% CI −0.27 to 0.13, p = 0.500, n = 94 engagements). Against the engagement indicator mean, it recorded r = −0.01 (95% CI −0.17 to 0.15, p = 0.901, n = 150 engagements). Against intent to leave it recorded r = −0.01 (cluster bootstrap 95% CI −0.17 to 0.15, p = 0.901, n = 508 respondents). The typed sizes differed markedly by vendor type, from the IT services majors to the pure-play agentic AI firms, consistent with the vendor types respondents named.

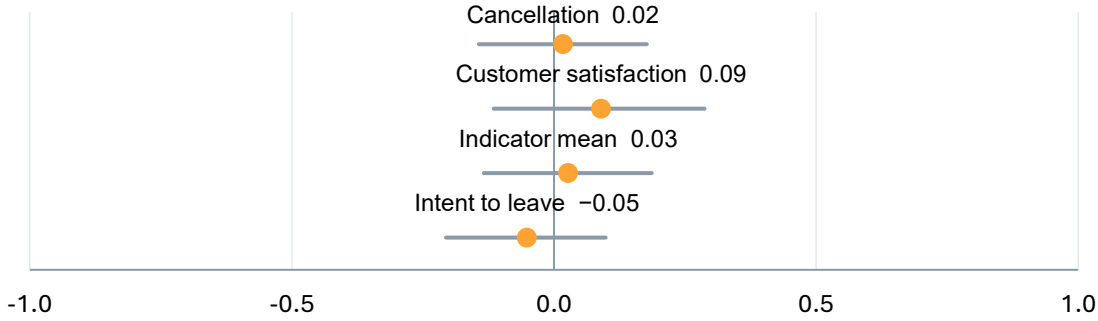
5.3 Breadth of Vendor Agentic Work for Financial Institutions

Indicator C079 · How many agentic AI engagements does your employer run for banks today? · Full statistical analysis ›

Fewer than half the respondents supplied a number for the vendor's agentic AI engagements serving financial institutions. A number was typed by 328 of the 850 respondents (38.59%), 242 (28.47%) answered “I do not know,” and 280 (32.94%) left the question blank. Among those who gave a number, the mean was 24.6 engagements and the median 25.0, from 5.0 to 40.0 (n = 328 respondents on 150 engagements).

Figure 1.24. Breadth of Vendor Agentic Work for Financial Institutions: associations with the study measures

Breadth of vendor agentic work for financial institutions: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = 0.02 [−0.14, 0.18] 150 engagements p = 0.833
Customer satisfaction	r = 0.09 [−0.11, 0.29] 94 active engagements p = 0.390
Indicator mean	r = 0.03 [−0.13, 0.19] 150 engagements p = 0.743
Intent to leave	r = −0.05 [−0.21, 0.10] 312 respondents p = 0.501

Note: Mean agentic AI engagements the vendor runs for financial institutions: 24.6 (n = 328 respondents).

Group means and numeric-answer bases: Individual contributor (n = 145): 24.17. Engagement manager (n = 150): 24.39. Portfolio director (n = 31): 27.65. Build-Out (n = 72): 24.43. Transition/Cut-Over (n = 98): 25.20. Production (n = 40): 24.00. Canceled (n = 118): 24.42.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The breadth of the vendor's agentic work showed no statistically significant association with the four study measures, each result within the study’s equivalence bounds. The correlation

with cancellation was r = 0.02 (95% CI −0.14 to 0.18, p = 0.833, n = 150 engagements). With customer satisfaction among active engagements, it was r = 0.09 (95% CI −0.11 to 0.29, p = 0.390, n = 94 engagements). With the engagement indicator mean, it was r = 0.03 (95% CI −0.13 to 0.19, p = 0.743, n = 150 engagements). With intent to leave, it was r = −0.05 (cluster bootstrap 95% CI −0.21 to 0.10, p = 0.501, n = 312 respondents).

The count records the breadth of a vendor's agentic work for financial institutions. It does not record whether the vendor had delivered the same service before. An engagement manager described a vendor that had sold a service it had never delivered.

> *...this is not about the engineers... they were good, they were simply sent in without a pipeline of our own to deploy from. We sold release automation to a global bank and we do not have a working reference of it anywhere."*
>
> **ENGAGEMENT MANAGER**
>
> ITO, Technology engineering and operations

“

We are the first team in our unit to run agents on an investment platform build, so we are working out the method as we go, that is the honest position. Leadership are supportive and the people we are given are capable enough, but ownership when an agent puts a wrong assumption into the pricing module is still not settled anywhere above me.”

ENGAGEMENT MANAGER

ITO, Technology engineering and operations

6 Size of the Delivery Team

The survey distinguished current, peak, and planned team size. Current headcount had estimated associations within the equivalence bounds for all four measures. Peak headcount was higher on canceled engagements. Planned headcount was associated with engagement status and, among active engagements, with customer satisfaction. These measures do not establish whether staffing changes preceded or followed performance changes.

Written accounts addressed staffing sufficiency rather than headcount alone. That distinction is retained below. Part 4 examines its relationship to workforce experience.

6.1 Current Team Size and Outcomes

Indicator C040 · How many people are on the vendor team for this engagement today? · Full statistical analysis ›

Reported current delivery-team headcounts agreed with the engagement-level data on every engagement. The mean was 8.4 people and the median 8.0, from 3.0 to 16.0 (n = 150 engagements). Respondents were also asked how many people were on the delivery team today, and every answer equaled the headcount recorded for the engagement (n = 789 respondents on 150 engagements).

Figure 1.25. Current Team Size and Outcomes: associations with the study measures

Current team size and outcomes: associations with the study measures

Comparison	Estimate and analysis base
Cancellation	r = 0.12 [−0.05, 0.27] 150 engagements p = 0.159
Customer satisfaction	r = 0.02 [−0.18, 0.22] 94 active engagements p = 0.851
Indicator mean	r = −0.10 [−0.26, 0.06] 150 engagements p = 0.226
Intent to leave	r = 0.06 742 respondents p = 0.442

Note: Mean delivery team headcount today: 8.8 (n = 789 respondents).

Group means and numeric-answer bases: Individual contributor (n = 617): 8.93. Engagement manager (n = 150): 8.35. Portfolio director (n = 22): 8.86. Build-Out (n = 157): 8.49. Transition/Cut-Over (n = 232): 9.11. Production (n = 96): 7.71. Canceled (n = 304): 9.12.

r is Pearson's correlation coefficient, measuring signed linear association from –1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Against the four study measures, every association of the headcount was within the study's equivalence bounds. Its correlation with cancellation was r = 0.12 (95% CI −0.05 to 0.27, p = 0.159, n = 150 engagements). With customer satisfaction among active engagements, it was r = 0.02 (95% CI −0.18 to 0.22, p = 0.851, n = 94 engagements). With the engagement indicator mean, it was r = −0.10 (95% CI −0.26 to 0.06, p = 0.226, n = 150 engagements). With intent to leave, it was r = 0.09 (cluster bootstrap 95% CI −0.07 to 0.24, p = 0.289, n = 798 respondents).

The headcount followed the type of engagement. After Holm correction, it differed by process family (Welch F 4.28, Holm p = 0.022, n = 150 engagements) and by class of work (Welch F 12.32, Holm p = 0.003, n = 150 engagements). Corporate and back-office engagements and corporate functional engagements had the smaller teams.

The written responses addressed whether the team was large enough for its tasks. There were 33 written responses on the size of the team, from 33 respondents on 29 engagements, and 23 of the 33 (69.70%) were negative. The most common statement, made in 17 of the 33, was that the team was too small for its tasks. Six described members who had left, had been moved, or were looking to leave.

> *…my team is only four people and two already looking for other job."*
>
> **INDIVIDUAL CONTRIBUTOR**
> CXM, Corporate and back-office functions

> *My lead is asking for more people every week, nobody give him…"*
>
> **INDIVIDUAL CONTRIBUTOR**
> KPO, Corporate and back-office functions

> *Talent shortage means we're always short-staffed"*
>
> **INDIVIDUAL CONTRIBUTOR**
> BPO, Corporate and back-office functions

The voluntary accounts concern perceived sufficiency for assigned tasks. A headcount measure does not directly test that judgment, so its null associations do not establish that staffing was adequate.

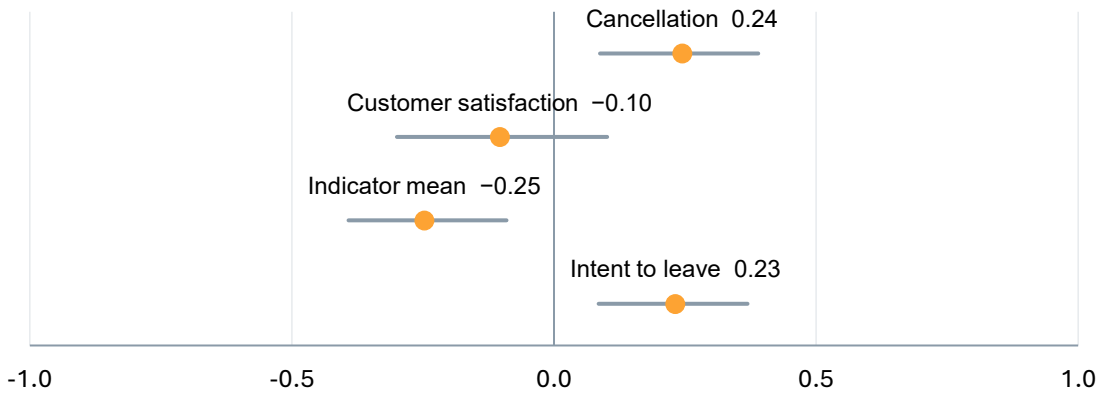
6.2 Team Size at Peak

Indicator C041 · What was the peak vendor team headcount on this engagement to date? · Full statistical analysis ›

Mean peak delivery-team headcount exceeded mean current headcount at engagement level. The mean answer was 10.5 people, and the median 10.0, from 3.0 to 18.0 (n = 787 respondents on 150 engagements), and no respondent answered: “I do not know.” At engagement level, the mean peak team was 10.0 people against a mean current headcount of 8.4 (n = 150 engagements).

Figure 1.26. Team Size at Peak: associations with the study measures

Team size at peak: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = 0.24 [0.09, 0.39] 150 engagements p = 0.003
Customer satisfaction	r = −0.10 [−0.30, 0.10] 94 active engagements p = 0.322
Indicator mean	r = −0.25 [−0.39, −0.09] 150 engagements p = 0.002
Intent to leave	r = 0.23 [0.09, 0.37] 739 respondents p = 0.002

Note: Mean peak delivery team headcount: 10.5 (n = 787 respondents).

Group means and numeric-answer bases: Individual contributor (n = 611): 10.66. Engagement manager (n = 150): 10.01. Portfolio director (n = 25): 10.88. Build-Out (n = 167): 10.02. Transition/Cut-Over (n = 230): 10.39. Production (n = 91): 8.81. Canceled (n = 299): 11.46.

r is Pearson’s correlation coefficient, measuring signed linear association from –1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Engagements with a larger peak team had a higher cancellation rate (r = 0.24, 95% CI 0.09 to 0.39, p = 0.003, n = 150 engagements). They also had a lower engagement indicator mean (r = −0.25, 95% CI −0.39 to −0.09, p = 0.002, n = 150 engagements). Respondents on those engagements reported higher intent to leave (r = 0.23, cluster bootstrap 95% CI 0.09 to 0.37, p = 0.002 with standard errors clustered by engagement, n = 739 respondents). The peak team showed no statistically significant association with customer satisfaction among active engagements, and that association was within the study's equivalence bounds (r = −0.10, 95% CI −0.30 to 0.10, p = 0.322, n = 94 engagements). Canceled engagements reported the largest peak teams.

Two of the three associations did not hold among active engagements alone. Within the 94 active engagements, the association with the engagement indicator mean was r = −0.15 (p = 0.136, n = 94 engagements). Among respondents on those engagements, the association with intent to leave was r = 0.09 (p = 0.358). Those two associations reflect the difference between canceled and active engagements more than a gradient within active work.

The data record the largest size the team reached. They do not record the course of the team before or after that point, or whether it grew in response to problems on the engagement. One individual contributor described a team reduced in a single day, partway through a task.

> *"in March they moved two people off my team to another account in one day, no handover. And we were mid way through tuning the screening rules. My lead found out from the rota, not from anyone above him. After that backlog went to nine days and the same managers sat in review asking why disposition quality dropped... nobody joined those two things."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Financial crime and client onboarding

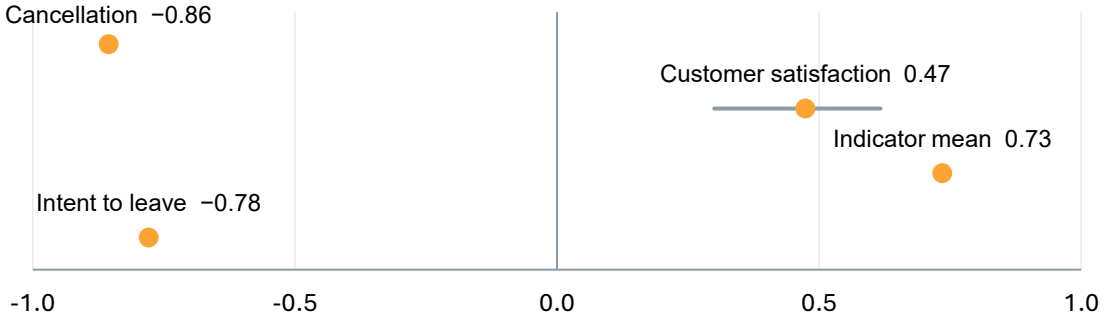
6.3 Planned Team Size as a Marker of Continuation

Indicator C042 · Per the current plan, how many people will be on the vendor team 12 months from now? · Full statistical analysis ›

Knowledge of the planned delivery-team headcount 12 months ahead was limited across the engagements. A number was given by 339 of the 850 respondents (39.88%). The question was answered by 589 respondents, and 250 of the 589 (42.44%) answered: “I do not know.” On 131 of the 150 engagements (87.33%), at least one respondent did not know the plan for the team. Among the respondents who gave a number, the mean was 5.6 people and the median 6.0 (n = 339 respondents).

Figure 1.27. Planned Team Size as a Marker of Continuation: associations with the study measures

Planned team size as a marker of continuation: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.86 150 engagements p < 0.001
Customer satisfaction	r = 0.47 [0.30, 0.62] 94 active engagements p < 0.001
Indicator mean	r = 0.73 150 engagements p < 0.001
Intent to leave	r = −0.78 323 respondents p < 0.001

Note: Mean planned delivery team headcount in 12 months: 5.6 (n = 339 respondents).

Group means and numeric-answer bases: Individual contributor (n = 157): 5.46. Engagement manager (n = 150): 5.71. Portfolio director (n = 31): 6.29. Build-Out (n = 63): 9.70. Transition/Cut-Over (n = 100): 9.50. Production (n = 41): 8.46. Canceled (n = 135): 0.00.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

All 135 numeric answers from respondents on canceled engagements reported a planned headcount of zero. The resulting cancellation association (r = −0.86, p < 0.001, n = 150

engagements) largely reflects reported engagement status. The full-sample associations with engagement indicator mean ($r = 0.73$, $n = 150$) and intent to leave ($r = -0.78$, $n = 323$ respondents) require the same qualification. They do not establish that staffing plans caused continuation or cancellation.

Among active engagements, where the plan is a forecast, the 204 respondents gave a number typed from 2.0 to 19.0 people. A larger planned team was associated with higher customer satisfaction ($r = 0.47$, 95% CI 0.30 to 0.62, $p < 0.001$, $n = 94$ engagements). It was associated with a higher engagement indicator mean ($r = 0.35$, 95% CI 0.16 to 0.52, $p < 0.001$, $n = 94$ engagements). Among respondents on the active engagements, it was associated with lower intent to leave ($r = -0.35$, 95% CI −0.47 to −0.22, $p < 0.001$). The association is descriptive. The data do not show whether the plan for the team followed the state of the engagement or anticipated it.

The structural results provide limited support for attributing outcomes to engagement type, vendor scale, or current headcount. Several estimates remain inconclusive, and associations involving peak and planned staffing have unresolved temporal direction. Part 2 examines the reported delivery conditions that covaried with outcomes, without treating those associations as causal explanations.

“

I rated my own company low, and I will still say this in its defense, nobody has this figured out yet, and the ones who claim they do are selling something. We were asked to staff a capital markets platform with agent-assisted delivery eighteen months after the tools existed, and we did staff it, and the platform is live. Where we are short is testing work that lands different on every run, so my leads test it by reading it, at night, on top of their own sprint work. That is a real gap, and it is not a gap anyone trained us to fill.”

ENGAGEMENT MANAGER

ITO, Technology engineering and operations

About the Author

Phil Hatch

Outsourcing and workforce development

Akholi author page

LinkedIn

ORCID 0000-0003-1709-8860

SSRN research paper

Phil Hatch is an outsourcing and workforce-development executive with more than 30 years of experience. He previously led the global outsourcing business of ManpowerGroup/Experis. His work includes service delivery, workforce strategy, and organizational development across industries and countries.

Hatch has also advised governments on labor and economic development, including workforce strategies intended to support investment and employment.

Following Russia's 2022 invasion of Ukraine, Hatch was asked by the U.S. Department of State and USAID to advise Ukraine on workforce and economic development. His work has addressed technology, outsourcing, skills development, and international investment in support of economic recovery.

Through Akholi, Hatch advises organizations and governments on outsourcing and the implications of AI for work and service delivery.

“

Decision rights and change control sit at rock bottom here… the joint sign-off process for exception handling takes weeks and breaks cannot wait that long, so the team ends up making the call itself under real time pressure. I have flagged this upward more than once and it has not moved, and with cancellation now confirmed, I am fairly sure it never will.”

ENGAGEMENT MANAGER

BPO, Trade, settlement and reconciliation

About Akholi

Akholi, established in 2003, is a research and advisory firm focused on outsourcing, service delivery, and global capability centers. Its work includes the 2026 study of 150 agentic AI outsourcing engagements described in this report.

The study examines relationships among financial institution readiness, co-management, vendor capability, AI technology, the third-party ecosystem, and reported engagement outcomes. Its findings are observational and do not establish that any one category is the sole or primary cause of those outcomes.

Akholi applies this research through advisory and educational work with institutional leaders. Recommendations derived from the study require assessment against the circumstances and evidence of each engagement.

Akholi

PART 1 / 2026 RESEARCH STUDY

Engagement Outcomes and Their Correlates

Part 1 examines engagement maturity, cancellation, perceived cancellation risk, customer satisfaction, delivery performance, recorded return, and structural attributes associated with the four study measures.

Phil Hatch | Akholi

Study ISBN 979-8-234-26882-2
SSRN DOI: https://doi.org/10.2139/ssrn.7164559
Figshare DOI: https://doi.org/10.6084/m9.figshare.34036890

Full study and supporting data

akholi.com/outsourcing-2026-agentic-ai-tier-1-banks

Version 1.8.2 | November 1, 2026

The evidence reflects respondents' reported experience. It is not an independent finding of regulatory breach.