



REFERENCE

Appendices

IN THE APPENDICES

A	Definitions
B	Indicator Inventory
C	Tables and the Indicators Behind Them
D	Method
E	Companion Indicator Statistical Analyses
F	Index
G	Disclosure of the Use of Artificial Intelligence

“

The thing the numbers cannot show is that all of this is untested. On paper the contracts with our model and data providers cover changes on their side, and I scored that well. In practice, the one time a provider moved a model version underneath us, the contract did nothing for me on the day, and we found out from a broken overnight refresh rather than from anybody's notice. And the licences were written for named humans, so nobody can tell me whether the agent counts as one user or as ten. None of it has cost us anything yet, because our volumes are small and we are still pre-POC. But a rating between one and ten cannot say fine until it is not, and that is the honest position.”

ENGAGEMENT MANAGER

KPO, Corporate and back-office functions

Appendices

The appendices provide definitions, the indicator and exhibit inventories, methods, a guide to the companion analyses, an index, and the AI-use disclosure. The companion analyses provide more detailed results than the main text.

Appendices A–C identify terms, indicators, and exhibits. Appendix D explains the methods and limitations; Appendix E describes and provides access to the 216 companion analyses. Appendices F and G contain the index and the AI-use disclosure.

1 Appendix A: Definitions

Appendix A defines the study's principal terms using the Study Taxonomy. Survey questions and respondent quotations retain their original wording, including terminology that differs from the report's usage.

Table 33. Terms used in the report and their definitions

<i>Term</i>	<i>Definition</i>
Financial institution	Tier 1 firm buying the outsourced service, of 10 types that include insurers and private equity and private capital firms.
Vendor	Outsourcing firm delivering the service and employing every respondent.
The two firms	Financial institution and the vendor considered together, where a weakness belongs to their relationship rather than to either firm alone.
Model provider	Supplier of the underlying AI models and the platforms they run on.
Third-party provider	Any other supplier to the engagement, such as a supplier of tools, data feeds, or platforms.
Engagement	One contracted piece of outsourced work performed by one vendor for one financial institution. The engagement is the primary unit of the study.
Agent	One autonomous AI system operating within an engagement.
Agentic AI outsourcing	Outsourcing in which agents perform part of the contracted work alongside the vendor's delivery team.
Study set	150 engagements and 850 respondents the study covers.
Respondent	A member of the vendor's delivery team who answered the survey about one engagement.
Individual contributor	Vendor role that performs the engagement's tasks. The study set holds 667.
Engagement manager	Vendor role that runs a single engagement. The study set holds 150, one on each engagement.
Portfolio director	Vendor role that oversees several engagements and their engagement managers. The study set holds 31.
Customer management executive	Vendor role that owns the relationship with the financial institution. The study set holds two, reported as counts only.
Delivery team	All vendor staff on one engagement, across roles.
Delivery workforce	All 850 respondents across the study set.
Financial institution staff	People employed by the financial institution who work with the engagement, including named subject-matter experts.
Outsourcing model	Form of outsourcing under which an engagement was contracted, customer experience management (CXM), information technology outsourcing (ITO), business process outsourcing (BPO) or knowledge process outsourcing (KPO).
Class of work	Whether an engagement served a line of business the financial institution runs or a corporate or back-office function.
Process family	One of eight groups of use cases. Each engagement was assigned to one process family from the use case it recorded.
Type of financial institution	One of 10 types recorded for each engagement, from asset management to private equity and private capital.
Vendor type	Respondent's employer type. The five types are the IT services major, the BPO specialist, the consulting firm, the pure-play agentic AI firm, and the boutique or specialist provider.
Build-Out	First stage measured, covering planning, team build-out, and knowledge transfer.
Transition/Cut-Over	Second stage measured, covering development, proof of concept, and engagements on hold.
Production	Third stage measured, covering engagements in production and delivery.
Canceled	An engagement terminated as of the study date.
Active	An engagement still running as of the study date.
Failure category	One of five groups of 20 rated indicators, each describing the conditions of an engagement by where a weakness lies.
Financial institution readiness	Conditions inside the financial institution, covering its data, its processes, its legacy systems, the skills of its people, its governance, and its responsiveness.
Vendor capability	Conditions inside the vendor, covering leadership, industry knowledge, agentic AI knowledge, the delivery playbook, platforms, talent, training, and the commercial model.
Co-management	Conditions arising from the two firms working together, covering decision rights, joint management of risks and issues, change control, human review, and the fit of governance.
AI technology	Conditions of the agentic tools themselves, covering reliability, non-determinism, memory, context, security, authorized action, token economics, and observability.
Third-party ecosystem	Conditions arising from third-party providers and the chain of dependencies beneath them.
Indicator	A survey question, engagement field or study measure analyzed in its own companion analysis.
Rated indicator	One of the 100 indicators, 20 in each failure category, that make up the engagement indicator mean.
10-point scale	Scale from 1 to 10 on which respondents rated the indicators and the other rated questions.
Rating	A respondent's answer from 1 to 10 on a rated question.
Favorable-direction indicator	An indicator on which favorability increases across the 10-point scale, where 1 is the most unfavorable rating and 10 the most favorable.
Favorable-direction conversion	Presentation of an unfavorable-direction indicator's rating in the favorable direction, made by subtracting each rating from 11, so that a higher figure denotes the more favorable outcome.
Unfavorable-direction indicator	An indicator on which 1 is the most favorable rating and 10 the least favorable. The eight are C019, C021, C023, C025, C039, C051, C053 and C054, among them intent to leave and cancellation likelihood.
Question without a favorable end	A rated question with neither a favorable nor an unfavorable end. The eight are C027, C028, C066, C067, C068, C203, C209 and C215.
Distressed	A rating at the bad end of the question's own scale, 3 or below where a low rating is bad and 8 or above where a high rating is bad. The rule applies to an engagement's mean rating in the same way.
Indicators distressed	Share of an engagement's ratings on the rated indicators that are distressed, computed for each respondent, averaged within the engagement and then across engagements.
Indicator quintile	One of five bands of 30 engagements, ranked by engagement indicator mean.
Indicator class	Level at which an indicator was recorded, as respondent level, engagement attribute, or perceived fact.

Term	Definition
Engagement attribute	One fact about the engagement, repeated across its respondents and never counted as independent observations.
Perceived fact	A fact about the engagement as the respondent understood it, such as the asset size of the financial institution. An answer of "I do not know" is reported as a finding.
Structured answers	Answers to the 195 structured survey questions, 172 numeric and 23 categorical. The study holds 149,619.
Written answers	Answers to the study's 16 open questions. The study holds 8,918.
Written response	A written answer about the respondent's own engagement that addresses a given indicator. Non-answers and answers about another engagement are excluded.
Primary and secondary response	A primary response answered the open question of the indicator's own survey section. A secondary response answered another open question and addressed the indicator in its content.
Sentiment score	Score of a written response on five values, from -10, a subject described at or near its worst, through -5, 0 and +5 to +10, at or near its best.
Theme	A recurring point in the written answers that no survey question captures, such as the frequency of manual work.
Emotional state	A state the respondent reported of himself or herself in a written answer, such as stressed, tired, or frustrated.
Non-answer	Answer "I do not know" or "I prefer not to answer," counted separately and excluded from a mean.
Four core study measures	Cancellation rate, customer satisfaction, engagement indicator mean, and intent to leave, applied to every indicator in that order.
Cancellation rate	Share of a named group of engagements canceled by the study date, computed at engagement level.
Customer satisfaction	Satisfaction figure recorded against each engagement on a 100-point scale, as reported by the delivery team. The measure is never abbreviated.
Engagement indicator mean	Mean of an engagement's ratings on the 100 rated indicators, computed for each respondent and averaged within the engagement.
Intent to leave	A respondent's rating, from 1 to 10, of how likely he or she is to leave the current job in the next 12 months.
Measured return	Return on an engagement, recorded for the 30 engagements in Production and printed with that base.
Cancellation likelihood	Delivery team's rating, from 1 to 10, of how likely the financial institution was to cancel an active engagement.
High risk	An active engagement whose mean cancellation likelihood rating was 7 or above.
Moderate risk	An active engagement whose mean cancellation likelihood rating was 5 or above and below 7.
Canceled or at risk	Canceled engagements together with the active engagements at high or moderate risk.
Manual alteration	A change made by hand by the delivery team to financial institution material, agent output, or reporting, counted by five methods.
Manual alteration of financial-institution-provided data	Changing by hand the data the financial institution supplied to the engagement.
Manual alteration of financial institution processes	Changing by hand a process of the financial institution, such as deciding a process step the financial institution had not decided.
Manual alteration of AI outputs	Adjusting by hand the output of an agent before it was used or sent.
Reporting alteration	Vendor's alteration of the performance reports and dashboards the financial institution received.
Third-party work done by hand	Delivery team performing by hand the tasks a third-party system did not deliver, such as keying in data a broken feed did not supply.
Internal disclosure	Telling the vendor's own management about a manual alteration.
Financial institution disclosure	Respondent or the respondent's management discussing a manual alteration with the financial institution.
Reported experience	Respondents' account of their own engagements, the evidence class of the whole study.
Engagement level	A figure computed with one value per engagement, on a base of up to 150 engagements.
Respondent level	A figure computed across respondents, whose answers cluster within engagements, on a base of up to 850 respondents.
Base	Count of respondents, answers, or engagements on which a figure was computed, being those who answered rather than those who were asked.
Confidence interval	Range around an estimate that contains the true value at a stated level of confidence, 95% unless stated otherwise.
p value	Probability of a result at least as extreme as the one observed if no association existed.
Pearson's r	Correlation between two measured quantities, from -1 to 1.
Spearman's rho	A correlation computed on the ranks of the values, used as a check on Pearson's r.
Chi-square test	A test of whether a split across categories departs from what independence would produce.
Welch's analysis of variance	A test of whether a mean differs across the levels of an attribute, without assuming equal spread across the levels.
Holm correction	An adjustment of each p value in a family of tests for the number of tests in the family.
False discovery rate	Expected share of false positives among the results that pass the test, controlled across the 300 partial tests.
Detection floor	Smallest correlation the tests could reliably detect, 0.23 at 150 engagements, a test level of 0.05, and 80% power.
Equivalence bound	Bound of 0.30 in absolute value of r, below which an association is treated as negligible.
Equivalence test	A pair of one-sided tests of whether an association lies within the equivalence bound.
Held	State of an association that had p below 0.05, survived correction for multiple comparisons, and kept its sign and its p below 0.05 under the robustness checks.
Did not hold	State of an association that had p below 0.05 on the primary test but failed correction or one or more robustness checks.
Null equivalent to zero	A result that had p of 0.05 or above and whose interval lay within the equivalence bound.
Inconclusive null	A result that had p of 0.05 or above and could not be shown equivalent to zero.
Clean estimate	An association computed against customer satisfaction among the 94 active engagements, where no cancellation had occurred for a respondent to react to.
Descriptive contrast	An association computed against realized cancellation across all 150 engagements, whose canceled engagements were rated after cancellation.
Leave-one-out	Recomputation of the engagement indicator mean without the indicator under test.
Partial test	A test of whether an indicator was associated with a measure beyond what the engagement indicator mean accounts for.
Intraclass correlation (ICC)	Share of the variance in an indicator that lies between engagements rather than within them.
Design effect	Factor by which clustering within engagements inflates variance, equal to one plus the ICC multiplied by the mean cluster size less one.
Effective sample size	Number of respondents who answered divided by the design effect, the number of independent answers the clustered answers are worth.

<i>Term</i>	<i>Definition</i>
Cluster bootstrap	An interval computed by resampling whole engagements rather than respondents. The answers from one engagement stay together.
Robustness check	A repeat of a primary test on a subset of the study set or under another method.
Study norm	Median association of the study's indicators with each of the four core study measures, against which each indicator is compared.
Thin base	A base under 30, printed with its count and without a confidence interval.
Anecdote strength	A base under 10, reported as counts only and never as a rate.
Companion analysis	Full statistical analysis of one indicator, engagement field, study measure, or derived analysis. The study holds 216.

2 Appendix B: Indicator Inventory

Appendix B lists the analyzed indicators by code, with the question wording, survey section, analytical class, report location, and companion-analysis filename. A perceived-fact classification identifies information reported from the respondent's perspective; it does not establish independent verification.

Table 34. Indicator inventory

<i>Code</i>	<i>Question as asked</i>	<i>Survey section</i>	<i>Class</i>	<i>Polarity</i>	<i>Where reported</i>	<i>Companion analysis</i>
C005	What is your role on this engagement?	1.1 Role and Identity	Respondent level	Not rated	Part 4, 1.1.1	Part 4, C005
C006	What country are you located in?	1.1 Role and Identity	Respondent level	Not rated	Part 1, 5.3.4	None
C007	How many years of experience do you have in your career?	1.2 Experience and Expertise	Respondent level	Not rated	Part 4, 1.2.1	Part 4, C007
C008	How many years of experience do you have with agentic AI?	1.2 Experience and Expertise	Respondent level	Not rated	Part 4, 1.3.1	Part 4, C008
C009	How many years have you been working in the outsourcing industry?	1.2 Experience and Expertise	Respondent level	Not rated	Part 4, 1.2.2	Part 4, C009
C010	How many years have you been working within the financial industry?	1.2 Experience and Expertise	Respondent level	Not rated	Part 4, 1.2.3	Part 4, C010
C011	How many agentic AI outsourcing engagements have you worked on across your career, including this one?	1.2 Experience and Expertise	Respondent level	Not rated	Part 4, 1.3.2	Part 4, C011
C012	How well do you know the financial industry?	1.2 Experience and Expertise	Respondent level	Higher is better	Part 4, 1.4.1	Part 4, C012
C013	How well do you know agentic AI specific to the financial industry?	1.2 Experience and Expertise	Respondent level	Higher is better	Part 4, 1.4.2	Part 4, C013
C014	Rate your overall agentic AI expertise.	1.2 Experience and Expertise	Respondent level	Higher is better	Part 4, 1.4.3	Part 4, C014
C015	What professional certifications do you have?	1.2 Experience and Expertise	Respondent level	Not rated	Part 4, 1.5.1	Part 4, C015
C016	What AI certifications do you have?	1.2 Experience and Expertise	Respondent level	Not rated	Part 4, 1.5.2	Part 4, C016
C017	Where did you gain your AI knowledge?	1.2 Experience and Expertise	Respondent level	Not rated	Part 4, 1.5.3	Part 4, C017
C018	How happy are you with your current job?	1.3 Sentiment and Retention	Respondent level	Higher is better	Part 4, 3.1.1	Part 4, C018
C019	How likely are you to leave your current job in the next 12 months?	1.3 Sentiment and Retention	Respondent level	Lower is better	Part 4, 5.1.1; Introduction, 4.5.4	Part 4, C019
C020	How happy is your extended team with the current engagement?	1.3 Sentiment and Retention	Respondent level	Higher is better	Part 4, 3.1.2	Part 4, C020
C021	How likely are members of your current team to leave their jobs in the next 12 months?	1.3 Sentiment and Retention	Respondent level	Lower is better	Part 4, 5.1.2	Part 4, C021
C022	Compared to 5 years ago, I am happier with the work I am doing now.	1.3 Sentiment and Retention	Respondent level	Higher is better	Part 4, 3.2.1	Part 4, C022
C023	Compared to 5 years ago, I am more likely to seek a new job.	1.3 Sentiment and Retention	Respondent level	Lower is better	Part 4, 5.2.1	Part 4, C023
C024	Compared to 5 years ago, the team I work with seems to be happier.	1.3 Sentiment and Retention	Respondent level	Higher is better	Part 4, 3.2.2	Part 4, C024
C025	Compared to 5 years ago, the team I work with is more likely to seek a new job.	1.3 Sentiment and Retention	Respondent level	Lower is better	Part 4, 5.2.2	Part 4, C025
C026	I want to continue working on projects in the global financial industry.	1.3 Sentiment and Retention	Respondent level	Higher is better	Part 4, 5.3.1	Part 4, C026
C027	I would prefer to work directly for a bank in one of their global capability centers.	1.3 Sentiment and Retention	Respondent level	Directional	Part 4, 5.3.2	Part 4, C027
C028	Working directly for a bank would be better than working for an outsourcing vendor.	1.3 Sentiment and Retention	Respondent level	Directional	Part 4, 5.3.3	Part 4, C028
C029	What are your top 3 challenges that reduce your job satisfaction?	1.3 Sentiment and Retention	Respondent level	Not rated	Part 4, 3.3	None
C030	What are the top 3 things that you love most about your job?	1.3 Sentiment and Retention	Respondent level	Not rated	Part 4, 3.5	None
C031	What type of a bank is the customer?	2.1 Engagement Identification	Engagement attribute	Not rated	Part 1, 5.1.1	Part 1, C031
C032	What is the outsourcing type?	2.1 Engagement Identification	Engagement attribute	Not rated	Part 1, 5.3.1	Part 1, C032

<i>Code</i>	<i>Question as asked</i>	<i>Survey section</i>	<i>Class</i>	<i>Polarity</i>	<i>Where reported</i>	<i>Companion analysis</i>
C033	Is the outsourcing engagement Corporate Functional, or Line of Business Operational?	2.1 Engagement Identification	Engagement attribute	Not rated	Part 1, 5.3.2	Part 1, C033
C034	What is the specific use case for this engagement?	2.1 Engagement Identification	Engagement attribute	Not rated	Part 1, 5.3.3	None
C035	How long ago did this specific engagement start, in months?	2.2 Timeline and Status	Engagement attribute	Not rated	Part 1, 1.1.2	Part 1, C035
C036	What stage is the engagement in?	2.2 Timeline and Status	Engagement attribute	Not rated	Part 1, 1.1.1; Introduction, 4.5.2	Part 1, C036
C037	How many months has the engagement been in its current stage?	2.2 Timeline and Status	Engagement attribute	Not rated	Part 1, 1.1.3	Part 1, C037
C038	Per the current plan, how many months from today until the engagement reaches Production/Delivery?	2.2 Timeline and Status	Engagement attribute	Not rated	Part 1, 1.1.4	Part 1, C038
C039	If the engagement has not been canceled, how likely is the customer to cancel the engagement?	2.2 Timeline and Status	Perceived fact	Lower is better	Part 1, 2.2.1	Part 1, C039
C040	How many people are on the vendor team for this engagement today?	2.3 Scale and Team	Engagement attribute	Not rated	Part 1, 5.6.1	Part 1, C040
C041	What was the peak vendor team headcount on this engagement to date?	2.3 Scale and Team	Engagement attribute	Not rated	Part 1, 5.6.2	Part 1, C041
C042	Per the current plan, how many people will be on the vendor team 12 months from now?	2.3 Scale and Team	Perceived fact	Not rated	Part 1, 5.6.3	Part 1, C042
C043	How many AI agents are running in production on this engagement today?	2.3 Scale and Team	Engagement attribute	Not rated	Part 1, 1.2.1	Part 1, C043
C044	What is the pricing model for this engagement?	2.4 Commercial Structure	Perceived fact	Not rated	Part 1, 5.4.1	Part 1, C044
C045	What is the contract term of this engagement, in months?	2.4 Commercial Structure	Perceived fact	Not rated	Part 1, 5.4.2	Part 1, C045
C046	What is the annual contract value of this engagement, in USD?	2.4 Commercial Structure	Perceived fact	Not rated	Part 1, 5.4.3	Part 1, C046
C047	I know the KPIs the customer uses to measure success.	2.5 Performance and Delivery	Respondent level	Higher is better	Part 1, 3.2.4	Part 1, C047
C048	The customer is happy with the performance of the engagement at this stage.	2.5 Performance and Delivery	Respondent level	Higher is better	Part 1, 3.1.2	Part 1, C048
C049	The engagement is meeting delivery dates successfully.	2.5 Performance and Delivery	Respondent level	Higher is better	Part 1, 3.2.1	Part 1, C049
C050	The engagement is within budget.	2.5 Performance and Delivery	Respondent level	Higher is better	Part 1, 3.2.2	Part 1, C050
C051	The engagement is not achieving stated KPIs.	2.5 Performance and Delivery	Respondent level	Lower is better	Part 1, 3.2.3	Part 1, C051
C052	The dashboard and performance reports given to the customer are accurate.	2.5 Performance and Delivery	Respondent level	Higher is better	Part 5, 1.4.1	Part 5, C052
C053	The vendor manipulates the dashboard and performance reports given to the customer.	2.5 Performance and Delivery	Respondent level	Lower is better	Part 5, 1.4.2	Part 5, C053
C054	Compared to 5 years ago, is manipulation of the dashboards and performance reports given to the customer more or less common now?	2.5 Performance and Delivery	Respondent level	Lower is better	Part 5, 1.4.3	Part 5, C054
C055	How does the vendor manipulate performance reports and dashboards?	2.5 Performance and Delivery	Respondent level	Not rated	Part 5, 1.4	None
C056	What laws, regulations, and similar govern this outsourcing engagement?	2.6 Regulatory Context	Respondent level	Not rated	Part 5, 6.1	None
C057	I fully understand how these laws, regulations, and similar affect the outsourcing engagement.	2.6 Regulatory Context	Respondent level	Higher is better	Part 5, 7.1.1	Part 5, C057
C058	My management team fully understands these laws, regulations, and similar, and how they affect the engagement.	2.6 Regulatory Context	Respondent level	Higher is better	Part 5, 7.1.2	Part 5, C058
C059	The customer's team actively discusses these laws, regulations, and similar with our team.	2.6 Regulatory Context	Respondent level	Higher is better	Part 5, 7.2.1	Part 5, C059
C060	The training available to me effectively teaches me about these laws, regulations, and similar.	2.6 Regulatory Context	Respondent level	Higher is better	Part 5, 7.3.1	Part 5, C060
C062	What is the customer's approximate total asset size, in USD?	3.1 Institution Profile	Perceived fact	Not rated	Part 1, 5.1.2	Part 1, C062
C063	Does the customer operate its own global capability centers?	3.1 Institution Profile	Perceived fact	Not rated	Part 1, 5.1.3	Part 1, C063

Code	Question as asked	Survey section	Class	Polarity	Where reported	Companion analysis
C064	How many months has your employer been serving this customer across all work, not just this engagement?	3.1 Institution Profile	Perceived fact	Not rated	Part 1, 5.2.1	Part 1, C064
C065	How many agentic AI outsourcing engagements does your employer run concurrently with this customer today?	3.1 Institution Profile	Perceived fact	Not rated	Part 1, 5.2.2	Part 1, C065
C066	The customer is scaling their agentic AI outsourcing engagements.	3.2 Customer AI Strategy	Perceived fact	Directional	Part 1, 5.1.5	Part 1, C066
C067	The customer is increasing the number of agentic AI outsourcing engagements.	3.2 Customer AI Strategy	Perceived fact	Directional	Part 1, 5.1.6	Part 1, C067
C068	The customer is shifting more work to their global capability centers.	3.2 Customer AI Strategy	Perceived fact	Directional	Part 1, 5.1.4	Part 1, C068
C069	I trust the customer.	3.3 Relationship and Communication	Respondent level	Higher is better	Part 4, 4.1.1	Part 4, C069
C070	The customer knows and cares about me.	3.3 Relationship and Communication	Respondent level	Higher is better	Part 4, 4.1.2	Part 4, C070
C071	I talk with peers inside the customer regularly.	3.3 Relationship and Communication	Respondent level	Higher is better	Part 4, 4.2.1	Part 4, C071
C072	I talk with engagement management inside the customer regularly.	3.3 Relationship and Communication	Respondent level	Higher is better	Part 4, 4.2.2	Part 4, C072
C073	I talk with the customer's leadership team regularly.	3.3 Relationship and Communication	Respondent level	Higher is better	Part 4, 4.2.3	Part 4, C073
C074	I know how the customer feels about my team.	3.3 Relationship and Communication	Respondent level	Higher is better	Part 4, 4.1.3	Part 4, C074
C075	I know how the customer feels about the engagement I am on.	3.3 Relationship and Communication	Respondent level	Higher is better	Part 4, 4.1.4	Part 4, C075
C076	What type of company is your employer?	4.1 Vendor Profile	Engagement attribute	Not rated	Part 1, 5.5.1	Part 1, C076
C077	Approximately how many employees does your employer have worldwide?	4.1 Vendor Profile	Perceived fact	Not rated	Part 1, 5.5.2	Part 1, C077
C078	How many years has your employer been delivering agentic AI outsourcing?	4.1 Vendor Profile	Perceived fact	Not rated	Part 1, 1.2.2	Part 1, C078
C079	How many agentic AI engagements does your employer run for banks today?	4.1 Vendor Profile	Perceived fact	Not rated	Part 1, 5.5.3	Part 1, C079
C080	My manager is effective.	4.2 My Manager	Respondent level	Higher is better	Part 4, 2.1.1	Part 4, C080
C081	My manager knows the financial industry well.	4.2 My Manager	Respondent level	Higher is better	Part 4, 2.1.2	Part 4, C081
C082	My manager knows agentic AI well.	4.2 My Manager	Respondent level	Higher is better	Part 4, 2.1.3	Part 4, C082
C083	My manager cares about me.	4.2 My Manager	Respondent level	Higher is better	Part 4, 4.3.1	Part 4, C083
C084	My manager cares about the larger team.	4.2 My Manager	Respondent level	Higher is better	Part 4, 4.3.2	Part 4, C084
C085	My manager knows how to manage agentic AI outsourcing engagements.	4.2 My Manager	Respondent level	Higher is better	Part 4, 2.1.4	Part 4, C085
C086	My manager knows how to manage agentic AI outsourcing engagements for global banks.	4.2 My Manager	Respondent level	Higher is better	Part 4, 2.1.5	Part 4, C086
C087	My employer cares about me.	4.3 Employer Relationship	Respondent level	Higher is better	Part 4, 4.4.1	Part 4, C087
C088	My employer cares about the team.	4.3 Employer Relationship	Respondent level	Higher is better	Part 4, 4.4.2	Part 4, C088
C089	My employer's top leadership team cares about my wellbeing and development.	4.3 Employer Relationship	Respondent level	Higher is better	Part 4, 4.4.3	Part 4, C089
C090	I trust my employer.	4.3 Employer Relationship	Respondent level	Higher is better	Part 4, 4.4.4	Part 4, C090
C091	I trust my manager.	4.3 Employer Relationship	Respondent level	Higher is better	Part 4, 4.3.3	Part 4, C091
C092	I am being used effectively by my employer.	4.3 Employer Relationship	Respondent level	Higher is better	Part 4, 4.4.5	Part 4, C092
C093	Compared to 5 years ago, is my employer more or less prepared to manage outsourcing engagements?	4.3 Employer Relationship	Respondent level	Higher is better	Part 2, 4.8.1	Part 2, C093
C094	Rate the quality of the customer's data across the systems this engagement depends on.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.1.1	Part 2, C094
C095	Rate how accessible the customer's data is to this engagement when and where the work requires it.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.1.2	Part 2, C095

<i>Code</i>	<i>Question as asked</i>	<i>Survey section</i>	<i>Class</i>	<i>Polarity</i>	<i>Where reported</i>	<i>Companion analysis</i>
C096	Rate how well the customer's data is normalized, with consistent formats and definitions across source systems.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.1.3	Part 2, C096
C097	Rate the maturity of the customer's data governance over the data this engagement uses.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.1.4	Part 2, C097
C098	Rate how well the customer's legacy systems support the integration this engagement requires.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.2.1	Part 2, C098
C099	Rate the readiness of the customer's technical infrastructure to support the agentic AI this engagement deploys.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.2.2	Part 2, C099
C100	Rate how completely the customer's processes are mapped and documented to the depth agentic AI requires.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.3.1	Part 2, C100
C101	Rate the maturity of the customer's processes in scope, meaning how standardized and repeatable they are.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.3.2	Part 2, C101
C102	Rate how consistently the customer performs the same processes across its business units and regions.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.3.3	Part 2, C102
C103	Rate the customer's willingness to redesign processes for agentic AI rather than automate them as they are.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.3.4	Part 2, C103
C104	Rate the agentic AI knowledge and skills of the customer staff working with this engagement.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.4.2	Part 2, C104
C105	Rate the effectiveness of the customer's management of its side of this engagement.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.5.1	Part 2, C105
C106	Rate the availability of the customer's subject matter experts when this engagement needs them.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.4.3	Part 2, C106
C107	Rate the customer organization's readiness to absorb the changes to daily work that agentic AI introduces.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.4.1	Part 2, C107
C108	Rate the customer's ability to provision the system access and identities that AI agents require.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.7.1	Part 2, C108
C109	Rate how quickly the customer responds when this engagement raises issues that need customer action.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.5.2	Part 2, C109
C110	Rate how well the customer's change control process handles the changes this engagement requires.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.6.3	Part 2, C110
C111	Rate the customer's capacity to govern and oversee the risks of AI agents doing this work.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.6.2	Part 2, C111
C112	Rate how clearly ownership and accountability for AI are assigned inside the customer organization.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.6.1	Part 2, C112
C113	Rate the stability of the customer's business case and ROI definition for this engagement.	5.1 Customer	Respondent level	Higher is better	Part 2, 2.8.1	Part 2, C113
C114	Add any comments on the customer organization that the ratings in this section do not capture.	5.1 Customer	Respondent level	Not rated	Part 2, 2	None
C115	Rate the overall effectiveness of your employer's top leadership team.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.1.1	Part 2, C115
C116	Rate how well your employer is transforming its commercial model as agentic AI reduces the billable human effort it sells today.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.1.2	Part 2, C116
C117	Rate how well your employer's internal management processes have been redesigned for agentic AI delivery rather than traditional staffed delivery.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.2.1	Part 2, C117
C118	Rate the effectiveness of your employer's process and leadership development programs.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.5.2	Part 2, C118
C119	Rate your employer's ability to manage agentic AI outsourcing engagements end to end.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.2.2	Part 2, C119
C120	Rate the depth of your employer's knowledge of the financial industry.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.3.1	Part 2, C120

Code	Question as asked	Survey section	Class	Polarity	Where reported	Companion analysis
C121	Rate your employer's ability to manage agentic AI outsourcing engagements for global banks in particular.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.2.3	Part 2, C121
C122	Rate your employer's current supply of qualified agentic AI talent against what its engagements need.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.5.1	Part 2, C122
C123	Rate the adequacy of your employer's investment in training its workforce.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.5.3	Part 2, C123
C124	Rate how well your employer's training programs build practical agentic AI skills.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.5.4	Part 2, C124
C125	Rate how well your employer's training programs build financial industry knowledge.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.5.5	Part 2, C125
C126	Rate the maturity of your employer's playbook for taking agentic use cases from concept to governed production.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.6.1	Part 2, C126
C127	Rate your employer's quality assurance capability for agents whose outputs are not identical every run.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.6.2	Part 2, C127
C128	Rate your employer's ability to price agentic work profitably when billable hours no longer measure the value delivered.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.1.3	Part 2, C128
C129	Rate how well the technical platforms your employer builds or licenses are optimized for agentic AI outsourcing.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.6.3	Part 2, C129
C130	Rate your employer's readiness to take accountability when an agent it operates makes a costly error.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.7.1	Part 2, C130
C131	Rate the depth of your employer's knowledge of agentic AI.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.3.2	Part 2, C131
C132	Rate the depth of your employer's experience running agentic AI systems in live production, beyond pilots.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.4.1	Part 2, C132
C133	Rate your employer's ability to operate large numbers of agents in production, including monitoring and exception handling.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.4.2	Part 2, C133
C134	Rate your employer's overall qualification to deliver elite agentic AI outsourcing services to banks.	5.2 Vendor	Respondent level	Higher is better	Part 2, 4.2.4	Part 2, C134
C135	Add any comments on your employer that the ratings in this section do not capture.	5.2 Vendor	Respondent level	Not rated	Part 2, 4	None
C136	Rate how effectively the customer and vendor jointly manage this engagement now that agents perform part of the work.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.1.1	Part 2, C136
C137	Rate how clearly the two firms have settled who decides what agents may do without human approval.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.1.2	Part 2, C137
C138	Rate how clearly ownership of each workflow step is assigned when agents, vendor staff, and customer staff work in the same process.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.1.3	Part 2, C138
C139	Rate how clearly accountability lands on one firm when an agent's action causes a problem in this engagement.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.1.4	Part 2, C139
C140	Rate how well the two firms maintain a single shared view of the risks that agent work introduces.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.2.1	Part 2, C140
C141	Rate how effectively the two firms manage issues arising from agent work as one team rather than through two separate processes.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.2.2	Part 2, C141
C142	Rate how well the speed of joint issue resolution matches the speed at which agent-caused problems can spread.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.2.3	Part 2, C142
C143	Rate how quickly the two firms together can stop an agent that is behaving badly in live operations.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.2.4	Part 2, C143
C144	Rate how well the joint change control process keeps pace with the speed of agent behavior changes.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.3.1	Part 2, C144
C145	Rate how well agent updates made by one firm become visible to the other firm before they take effect.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.3.2	Part 2, C145

<i>Code</i>	<i>Question as asked</i>	<i>Survey section</i>	<i>Class</i>	<i>Polarity</i>	<i>Where reported</i>	<i>Companion analysis</i>
C146	Rate how well the two firms jointly govern changes agents make to their own behavior between formal releases.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.3.3	Part 2, C146
C147	Rate how well human review points are placed across the two firms at the moments where agent errors would matter most.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.4.1	Part 2, C147
C148	Rate how well joint human review capacity keeps up with the volume of output the agents produce.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.4.2	Part 2, C148
C149	Rate how much genuine scrutiny agent outputs receive at human review points rather than routine approval.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.4.3	Part 2, C149
C150	Rate how reliably the joint quality assurance loop catches agent errors before they reach live operations.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.4.4	Part 2, C150
C151	Rate how well the quality measures both firms use capture the kinds of errors agents actually make.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.4.5	Part 2, C151
C152	Rate how fully each firm can see what the other firm's agents actually did when questions arise.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.5.1	Part 2, C152
C153	Rate how closely the speed of joint decision making matches the speed at which the agents act.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.5.2	Part 2, C153
C154	Rate how well today's joint governance model fits work performed by agents rather than work performed by people.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.5.3	Part 2, C154
C155	Rate how ready the joint management model is for the level of agent autonomy you expect in this engagement within two years.	5.3 Co-Management	Respondent level	Higher is better	Part 2, 3.5.4	Part 2, C155
C156	Add any comments on the joint management of this engagement that the ratings in this section do not capture.	5.3 Co-Management	Respondent level	Not rated	Part 2, 3	None
C157	Rate the factual reliability of the AI tools' outputs on this engagement before any human review.	5.4 AI	Respondent level	Higher is better	Part 2, 5.2.1	Part 2, C157
C158	Rate the accuracy of citations, references, and quotations the tools produced against the real sources.	5.4 AI	Respondent level	Higher is better	Part 2, 5.2.2	Part 2, C158
C159	Rate the consistency of the tools' outputs when given the same input and instructions more than once.	5.4 AI	Respondent level	Higher is better	Part 2, 5.2.3	Part 2, C159
C160	Rate the end-to-end reliability of the tools on multi-step tasks, from first step to finished output.	5.4 AI	Respondent level	Higher is better	Part 2, 5.2.4	Part 2, C160
C161	Rate the predictability of the tools' token consumption for a given volume of work.	5.4 AI	Respondent level	Higher is better	Part 2, 5.4.1	Part 2, C161
C162	Rate the efficiency of the tools' token use, considering retries, loops, and reprocessing needed to complete work.	5.4 AI	Respondent level	Higher is better	Part 2, 5.4.2	Part 2, C162
C163	Rate how completely the tools kept each client's and each session's information isolated from every other client and session.	5.4 AI	Respondent level	Higher is better	Part 2, 5.5.1	Part 2, C163
C164	Rate how reliably the tools retained instructions and constraints given earlier in long tasks and sessions.	5.4 AI	Respondent level	Higher is better	Part 2, 5.3.1	Part 2, C164
C165	Rate how completely the tools captured material facts located anywhere in long documents, including mid-document content.	5.4 AI	Respondent level	Higher is better	Part 2, 5.3.2	Part 2, C165
C166	Rate the integrity of the tools' persistent memory against contamination by incorrect or malicious content encountered during work.	5.4 AI	Respondent level	Higher is better	Part 2, 5.3.3	Part 2, C166
C167	Rate the tools' resistance to hostile instructions embedded in the content they read and process.	5.4 AI	Respondent level	Higher is better	Part 2, 5.5.3	Part 2, C167
C168	Rate how consistently the agents took only intended and authorized actions when operating connected tools and systems.	5.4 AI	Respondent level	Higher is better	Part 2, 5.6.1	Part 2, C168

<i>Code</i>	<i>Question as asked</i>	<i>Survey section</i>	<i>Class</i>	<i>Polarity</i>	<i>Where reported</i>	<i>Companion analysis</i>
C169	Rate how safely the tools handled sensitive data, with no exposure through outputs, logs, or connected systems.	5.4 AI	Respondent level	Higher is better	Part 2, 5.5.2	Part 2, C169
C170	Rate how reliably the tools' built-in guardrails and safety controls held during production use.	5.4 AI	Respondent level	Higher is better	Part 2, 5.6.2	Part 2, C170
C171	Rate the overall production readiness of the AI tools for bank-grade work as experienced on this engagement.	5.4 AI	Respondent level	Higher is better	Part 2, 5.1.1	Part 2, C171
C172	Rate how accurately the tools' own reports of success or failure matched the actual outcomes of their work.	5.4 AI	Respondent level	Higher is better	Part 2, 5.8.1	Part 2, C172
C173	Rate the tools' ability to explain and evidence how each output was produced, at the depth an audit requires.	5.4 AI	Respondent level	Higher is better	Part 2, 5.8.2	Part 2, C173
C174	Rate how well the tools supported reliable testing, meaning correct behavior could be proven before work went live.	5.4 AI	Respondent level	Higher is better	Part 2, 5.7.1	Part 2, C174
C175	Rate the stability of the tools' behavior over the engagement when nothing changed on the engagement side.	5.4 AI	Respondent level	Higher is better	Part 2, 5.7.2	Part 2, C175
C176	Rate the quality of the tools' logs, traces, and usage telemetry for detecting and diagnosing failures.	5.4 AI	Respondent level	Higher is better	Part 2, 5.8.3	Part 2, C176
C177	Add any comments on the agentic AI tools that the ratings in this section do not capture.	5.4 AI	Respondent level	Not rated	Part 2, 5	None
C178	Rate how well the agents on this engagement are protected from breaking silently when a third party changes something without warning.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.2.1	Part 2, C178
C179	Rate the quality of the fallback the engagement has ready when a critical third-party feed, platform, or model stops working.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.7.1	Part 2, C179
C180	Rate the security posture of the third-party tools, platforms, and sub-processors the engagement's agents connect to.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.6.1	Part 2, C180
C181	Rate the quality of advance notice third-party providers give before changes that can affect the agents on this engagement.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.1.1	Part 2, C181
C182	Rate how reliably word of an upcoming third-party change reaches the people who operate the affected agents in time to act.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.1.2	Part 2, C182
C183	Rate the engagement's monitoring for unannounced changes in the third-party APIs, data feeds, and models its agents depend on.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.2.2	Part 2, C183
C184	Rate how well the contracts with third-party providers deal with provider changes that can break the agents on this engagement.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.4.1	Part 2, C184
C185	Rate how well the engagement absorbs model version deprecations and retirements forced by model providers.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.3.2	Part 2, C185
C186	Rate the availability of true machine interfaces, rather than screens built for humans, across the third-party systems the agents must use.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.5.1	Part 2, C186
C187	Rate the stability of the third-party APIs and data structures the agents depend on, from one provider release to the next.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.3.1	Part 2, C187
C188	Rate the quality of the data arriving from third-party feeds, judged by whether agents can use it without human cleanup.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.5.2	Part 2, C188
C189	Rate how well third-party rate limits and quotas hold up under the volume the agents on this engagement generate.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.5.3	Part 2, C189
C190	Rate the engagement's ability to still perform by hand the manual workarounds that once absorbed third-party quirks and breakage.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.7.2	Part 2, C190

Code	Question as asked	Survey section	Class	Polarity	Where reported	Companion analysis
C191	Rate how completely the engagement knows where humans used to quietly fix third-party data and interface problems before agents took over.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.7.3	Part 2, C191
C192	Rate how deliberately the quiet human fixes for third-party quirks were replaced with explicit checks when agents took over the work.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.7.4	Part 2, C192
C193	Rate how well the engagement manages its dependence on a small number of model and cloud providers.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.6.2	Part 2, C193
C194	Rate how well this engagement could keep operating if it had to switch its primary model or cloud provider.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.6.3	Part 2, C194
C195	Rate how well third-party licenses on this engagement cover work done by agents rather than named human users.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.4.2	Part 2, C195
C196	Rate how completely the engagement has mapped every third and fourth party its agents depend on, down the full chain.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.8.1	Part 2, C196
C197	Rate how well the engagement manages agent dependence on gated external rails such as payment networks and regulatory submission systems.	5.5 3rd Party Ecosystem	Respondent level	Higher is better	Part 2, 6.8.2	Part 2, C197
C198	Add any comments on the third-party ecosystem that the ratings in this section do not capture.	5.5 3rd Party Ecosystem	Respondent level	Not rated	Part 2, 6	None
C199	I have altered client-provided data	5.6 Alterations	Respondent level	Not rated	Part 5, 1.1.1	Part 5, C199
C200	I discussed the data alteration with internal management	5.6 Alterations	Respondent level	Not rated	Part 5, 4.1.1	Part 5, C200
C201	I or my management discussed data alteration with the client	5.6 Alterations	Respondent level	Not rated	Part 5, 4.2.1	Part 5, C201
C202	Why did you alter client provided data?	5.6 Alterations	Respondent level	Not rated	Part 5, 3.1	None
C203	My greater team regularly alters client provided data.	5.6 Alterations	Respondent level	Directional	Part 5, 1.1.2	Part 5, C203
C204	Comment	5.6 Alterations	Respondent level	Not rated	Part 5, 1.1	None
C205	I have altered client processes	5.6 Alterations	Respondent level	Not rated	Part 5, 1.2.1	Part 5, C205
C206	I discussed the process alteration with internal management	5.6 Alterations	Respondent level	Not rated	Part 5, 4.1.2	Part 5, C206
C207	I or my management discussed process alteration with the client	5.6 Alterations	Respondent level	Not rated	Part 5, 4.2.2	Part 5, C207
C208	Why did you alter client processes?	5.6 Alterations	Respondent level	Not rated	Part 5, 3.2	None
C209	My greater team regularly alters client processes.	5.6 Alterations	Respondent level	Directional	Part 5, 1.2.2	Part 5, C209
C210	Comment	5.6 Alterations	Respondent level	Not rated	Part 5, 1.2	None
C211	I have altered AI outputs	5.6 Alterations	Respondent level	Not rated	Part 5, 1.3.1	Part 5, C211
C212	I discussed the AI output alteration with internal management	5.6 Alterations	Respondent level	Not rated	Part 5, 4.1.3	Part 5, C212
C213	I or my management discussed AI output alteration with the client	5.6 Alterations	Respondent level	Not rated	Part 5, 4.2.3	Part 5, C213
C214	Why did you alter AI outputs?	5.6 Alterations	Respondent level	Not rated	Part 5, 3.3	None
C215	My greater team regularly alters AI outputs.	5.6 Alterations	Respondent level	Directional	Part 5, 1.3.2	Part 5, C215
C216	Comment	5.6 Alterations	Respondent level	Not rated	Part 5, 1.3	None
Eng: Delivery location	Engagements sheet field: Delivery Location	Engagements sheet	Engagement field	Not rated	Part 1, 5.3.4	None
Field: Bank Type	Engagements sheet field: Bank Type	Engagements sheet	Engagement field	Not rated	Part 1, 5.1.1	Part 1, Field: Bank Type
Field: Outsourcing Type	Engagements sheet field: Outsourcing Type	Engagements sheet	Engagement field	Not rated	Part 1, 5.3.1	Part 1, Field: Outsourcing Type
Field: Class of Work	Engagements sheet field: Corporate Functional or LOB Operational	Engagements sheet	Engagement field	Not rated	Part 1, 5.3.2	Part 1, Field: Class of Work

Code	Question as asked	Survey section	Class	Polarity	Where reported	Companion analysis
Field: Use Case	Engagements sheet field: Use Case	Engagements sheet	Engagement field	Not rated	Part 1, 5.3.3	Part 1, Field: Use Case
Field: Engagement Age	Engagements sheet field: Age of Engagement (Months)	Engagements sheet	Engagement field	Not rated	Part 1, 1.1.2	Part 1, Field: Engagement Age
Field: Engagement Phase	Engagements sheet field: Engagement Phase	Engagements sheet	Engagement field	Not rated	Part 1, 1.1.1	Part 1, Field: Engagement Phase
Field: Headcount	Engagements sheet field: Headcount	Engagements sheet	Engagement field	Not rated	Part 1, 5.6.1	Part 1, Field: Headcount
Field: IC Respondents	Engagements sheet field: IC Respondent Count	Engagements sheet	Engagement field	Not rated	Part 4, 1.1.1	Part 4, Field: IC Respondents
Field: PD Respondents	Engagements sheet field: PD Respondent Count	Engagements sheet	Engagement field	Not rated	Part 4, 1.1.1	Part 4, Field: PD Respondents
Field: CME Respondents	Engagements sheet field: CME Respondent Count	Engagements sheet	Engagement field	Not rated	Part 4, 1.1.1	Part 4, Field: CME Respondents
Field: Total Respondents	Engagements sheet field: Total Respondents	Engagements sheet	Engagement field	Not rated	Part 4, 1.1.1	Part 4, Field: Total Respondents
Measure: Cancellation	Measure: Cancellation	Study measure	Study measure	Not rated	Part 1, 2.1.1	Part 1, Measure: Cancellation
Measure: Satisfaction	Measure: Satisfaction	Study measure	Study measure	Not rated	Part 1, 3.1.1; Introduction, 4.5.3	Part 1, Measure: Satisfaction
Measure: Indicator Mean	Measure: Indicator Mean	Study measure	Study measure	Not rated	Part 1, 4.1.1; Introduction, 4.5.1	Part 1, Measure: Indicator Mean

3 Appendix C: Tables and the Indicators Behind Them

Each table and figure of the report is listed in Appendix C with the indicators it draws on and its base. Where one base covers the whole exhibit, Appendix C gives it. Where the base differs from row to row, the exhibit prints it beside each figure. From any table or figure, a reader can move to the companion analyses that hold the full statistics behind it. Where a table combines several indicators, each indicator is listed.

Table 35. Tables and figures of the report and the indicators behind them

<i>Exhibit</i>	<i>Title</i>	<i>Where reported</i>	<i>Indicators</i>	<i>Base</i>
Figure 1	Mean indicator rating of the five failure categories	Introduction, 3	Study: cancellation, Analysis: Active Engagements, Study: customer satisfaction, Analysis: Measured Return, C051, Study: disclosure, Study: intent to leave, C148, C144, C150, C078, C008, C126, C104, C103, Study: engagement indicator mean	Given in the exhibit
Figure 2	Respondents by role and engagements by lifecycle stage	Introduction, 4.1	Study: roles, Study: lifecycle, Study: outsourcing type, Study: class of work, Study: institution type, Study: vendor type, Study: delivery location	n = 850 respondents
Figure 3	Engagements by status and lifecycle stage	Introduction, 4.1.3	Study: lifecycle	n = 150 engagements
Figure 4	Engagements by outsourcing model	Introduction, 4.1.4	Study: outsourcing type	n = 150 engagements
Figure 5	Engagements by type of financial institution	Introduction, 4.1.5	Study: institution type	n = 150 engagements
Figure 6	Engagements by process family	Introduction, 4.1.6	Study: process families	n = 150 engagements
Table 1	The five manual alteration methods	Introduction, 4.6.1	C199, C205, C211, C055, C191, C192	Given in the exhibit
Figure 1.1	Cancellation rate by lifecycle stage against the study-set rate	Part 1, 1.1	Study: lifecycle, Field: Engagement Phase, C036	n = 150 engagements
Figure 1.2	Composition of the 150 engagements by cancellation and rated risk	Part 1, 2	Study: cancellation, Analysis: Active Engagements, C039, Measure: Cancellation	Given in the exhibit
Figure 1.3	Engagements by recorded customer satisfaction	Part 1, 3.1	Measure: Satisfaction, C048	Given in the exhibit
Figure 1.4	Cancellation rate by engagement indicator quintile	Part 1, 4.2	Analysis: Four Measures, Analysis: Mean Quintiles, Analysis: Active Engagements	n = 150 engagements
Table 2	The study measures by engagement indicator quintile	Part 1, 4.2.2	Analysis: Mean Quintiles	Given in the exhibit
Table 3	Structural attributes against the four core study measures	Part 1, 5	C031, Field: Bank Type, C032, C033, Field: Use Case, C044, C045, C046, C062, C063, C064, C065, C066, C067, C068, C076, C077, C079, C040, C041, C042	Given in the exhibit
Table 4	The four core study measures by type of financial institution	Part 1, 5.1.1	Field: Bank Type, Study: cancellation, Study: customer satisfaction, Study: engagement indicator mean, Study: intent to leave	n = 150 engagements
Table 5	The four core study measures by outsourcing model	Part 1, 5.3.1	C032, Field: Outsourcing Type, Study: cancellation, Study: customer satisfaction, Study: engagement indicator mean, Study: intent to leave	n = 150 engagements
Table 6	The four core study measures by class of work	Part 1, 5.3.2	C033, Field: Class of Work, Study: cancellation, Study: customer satisfaction, Study: engagement indicator mean, Study: intent to leave	n = 150 engagements
Table 7	The four core study measures by process family	Part 1, 5.3.3	Field: Use Case (process family), Study: cancellation, Study: customer satisfaction, Study: engagement indicator mean, Study: intent to leave	n = 150 engagements
Table 8	The four core study measures by pricing model	Part 1, 5.4.1	C044, Study: cancellation, Study: customer satisfaction, Study: engagement indicator mean, Study: intent to leave	n = 107 engagements
Table 9	The four core study measures by vendor type	Part 1, 5.5.1	C076, Study: cancellation, Study: customer satisfaction, Study: engagement indicator mean, Study: intent to leave	n = 150 engagements
Table 10	The five failure categories ranked by mean rating	Part 2, 1	Study: engagement indicator mean	Given in the exhibit
Figure 2.1	Share of answers in the distressed range, by failure category	Part 2, 1.1	Study set (no single indicator)	Given in the exhibit
Table 11	Mean rating of each failure category by engagement status, lifecycle stage and respondent role	Part 2, 1.1	Study set (no single indicator)	Given in the exhibit
Table 12	The indicators of financial institution readiness, from the lowest mean rating	Part 2, 2	C114	Given in the exhibit
Figure 2.2	Share of answers in the distressed range, the four data indicators	Part 2, 2.1	C094, C095, C096, C097	Given in the exhibit
Table 13	The indicators of co-management, from the lowest mean rating	Part 2, 3	C156	Given in the exhibit
Figure 2.3	Written responses on the scrutiny of agent output, by sentiment score	Part 2, 3.4.3	C149	Given in the exhibit

Exhibit	Title	Where reported	Indicators	Base
Table 14	The indicators of vendor capability, from the lowest mean rating	Part 2, 4	C093, C135	Given in the exhibit
Table 15	The indicators of AI technology, from the lowest mean rating	Part 2, 5	C177	Given in the exhibit
Table 16	The indicators of the third-party ecosystem, from the lowest mean rating	Part 2, 6	C198	Given in the exhibit
Figure 3.1	Share of answers in the distressed range, the 19 governance indicators	Part 3, 1	C110, C111, C112, C136, C137, C138, C139, C140, C141, C142, C143, C144, C145, C146, C147, C148, C149, C150, C154	Given in the exhibit
Table 17	Respondents by role	Part 4, 1.1.1	C005, Study: roles, Study: intent to leave	Given in the exhibit
Figure 4.1	Mean years of career experience and of agent AI experience, by respondent role	Part 4, 1.3	C007, C008, C011	Given in the exhibit
Table 18	Professional certifications reported by respondents	Part 4, 1.5.1	C015	Given in the exhibit
Table 19	AI certifications reported by respondents	Part 4, 1.5.2	C016	Given in the exhibit
Table 20	Sources of AI knowledge reported by respondents	Part 4, 1.5.3	C017	Given in the exhibit
Table 21	The five manager indicators	Part 4, 2.1	C080, C081, C082, C085, C086	Given in the exhibit
Table 22	Trust in and care from the financial institution	Part 4, 4.1	C069, C070, C074, C075	Given in the exhibit
Table 23	Trust in and care from the employer	Part 4, 4.4	C087, C088, C089, C090, C092	Given in the exhibit
Figure 4.2	Share of intent-to-leave ratings in the distressed range, by engagement status and lifecycle stage	Part 4, 5.1	C019, C021	Given in the exhibit
Figure 5.1	Respondents reporting each directly counted manual alteration method	Part 5, 1	C199, C205, C211, C203, C209, C215	Given in the exhibit
Table 24	Cancellation by the answer assigned to each engagement on each counted method	Part 5, 1	C199, C205, C211, C203, C209, C215	Given in the exhibit
Figure 5.2	Share of respondents reporting each manual alteration, by respondent role	Part 5, 2	Analysis: Alterations by Role, Analysis: Manager Alterations, Study: disclosure	Given in the exhibit
Table 25	Respondents answering True on each counted question, by role	Part 5, 2.1.1	Analysis: Alterations by Role	Given in the exhibit
Table 26	Disclosure of the alterations by role, summed across the three paired questions	Part 5, 2.2.1	Analysis: Manager Alterations	Given in the exhibit
Table 27	Respondents answering True on each counted method, by indicator quintile	Part 5, 2.3	Study: alteration by indicator quintile	Given in the exhibit
Table 28	Indicators named most in the reasons for each counted method	Part 5, 3.4	C094, C108, C109, C100, C159, C157	Given in the exhibit
Figure 5.3	Manual alterations discussed with vendor management and with the financial institution	Part 5, 4	Study: disclosure, C200, C206, C212, C201, C207, C213	Given in the exhibit
Table 29	Disclosure of manual alteration to the financial institution	Part 5, 4.2	Study: disclosure, C201, C207, C213	Given in the exhibit
Table 30	Obligations named in the most answers as governing the engagement	Part 5, 6.1	C056	Given in the exhibit
Table 31	The seven immediate actions and the primary findings on which each rests	Part 6, 1	Study set (no single indicator)	Given in the exhibit
Table 32	The program at a glance, with what reaches the board at the end of each block	Part 6, 2	Study set (no single indicator)	Given in the exhibit

4 Appendix D: Method

The study analyzed responses from 850 vendor delivery staff associated with 150 agentic AI outsourcing engagements serving institutions classified in the study as Tier 1. The reported fieldwork period was January–June 2026. A common analytical framework was applied, with tests selected according to variable type, analytical unit, and available sample size.

4.1 Survey Design

Each respondent answered about the one engagement on which they worked. Five sections made up the survey. Its first section covered the respondent's role, experience, and intent to stay in the job. A second covered the engagement, including its timeline, scale, commercial structure, performance, and governing regulations. Questions on the financial institution followed, on its profile, its AI strategy, and its relationship with the vendor. A fourth section asked about the vendor, the respondent's manager, and the respondent's relationship with the employer.

In the fifth section, the survey examined the challenges of the engagement in depth. Each of the five failure categories, financial institution readiness, vendor capability, co-management, AI technology, and the third-party ecosystem, received 20 rated indicators. An open question closed each category, inviting comment on what its ratings did not capture.

A final group of 18 questions covered manual alteration. Three methods were asked about directly, manual alteration of financial-institution-provided data, of financial institution processes, and of AI outputs. Each method received six questions. A yes-or-no question asked whether the respondent had made the alteration, and a written question asked why. Two further yes-or-no questions asked whether the respondent had discussed it with the vendor's management and whether the respondent or the respondent's management had discussed it with the financial institution. A rating of the wider team's practice and an open question for comment completed the set.

Three types of question appeared in the survey. Rated questions used a scale from 1 to 10. Categorical questions recorded attributes such as role or outsourcing model, together with the yes-or-no answers on manual alteration and its disclosure. Open questions asked for a written answer. In place of a rating, respondents could answer “I do not know” or “I prefer not to answer,” and both answers were counted apart from the ratings.

Rated questions ran in both directions. Of the 146 rated questions with a favorable and an unfavorable end, 138 were worded so that a low rating is bad and eight so that a high rating is bad. Eight further rated questions had neither end. Another 18 numeric questions recorded typed numbers, such as months of experience, team size, or contract term.

Survey answers fill 216 columns of the study data. Five were set aside before analysis, namely the respondent and engagement identifiers, two codes that repeated another answer, and the location of the

financial institution's headquarters. What remains is 211 survey questions, of which 172 are numeric, 23 categorical, and 16 open. Analysis also covers 12 engagement fields, recorded once for each engagement.

4.2 Sampling and Respondents

Respondents were vendor employees answering about a sampled engagement. The study records state that financial institutions, vendors, model providers, and other third parties did not supply independent validation. The analytical files available to the author contained deidentified records. This does not establish that identifying information was never collected during recruitment or administration.

The reported recruitment process used professional networks, published profiles, vetted panels, and referrals through a specialist provider. Participation was compensated. More than 900 completed responses were reportedly screened for working knowledge of banking, outsourcing, and agentic AI, leaving 850 respondents. Study documentation also describes telephone validation with 122 respondents and reports no material discrepancies. These are descriptions of the administration process; the files supporting the present analysis do not provide a complete recruitment disposition, response rate, screening instrument, or validation record.

Engagement selection sought variation in institutional type, outsourcing model, and use case. The sample is not established as probability-based or representative of the industry. Estimates describe the sampled engagements; confidence intervals and significance tests do not remove this selection limitation.

Three lifecycle stages are represented, with 60 engagements in Build-Out, 60 in Transition/Cut-Over, and 30 in Production. By the study date, 56 engagements had been canceled and 94 remained active. Customer experience management accounted for 42 engagements, information technology outsourcing for 40, business process outsourcing for 37, and knowledge process outsourcing for 31. Line-of-business operations held 76 engagements and corporate functions 74. Engagement age had a mean of 5.69 months and a median of 5 months ($n = 150$ engagements).

Individual contributors made up 667 of the 850 respondents (78.47%). Alongside them were 150 engagement managers, one on each engagement, 31 portfolio directors, and two customer management executives.

Among the 793 respondents with an originally recorded vendor type, 316 reported an IT services major, 198 a BPO specialist, 117 a consulting firm, 87 a boutique or specialist provider, and 75 a pure-play agentic AI firm.

Vendor type was originally blank for 57 respondents. The analytical preparation derived these values from the engagement-level classification, which was internally consistent within each engagement. Derived

values remain distinguishable from originally reported values; the preceding counts describe the original answers.

The valid base varies by question and analysis. Numeric summaries exclude nonnumeric answers unless a stated sensitivity analysis specifies otherwise. Categorical analyses identify which answers are retained as levels. Counts of written answers, respondents, and engagements are reported separately.

4.3 The Study Data

Two sheets and the written answers make up the study data. One row for each respondent and one column for each survey question form the Data sheet. On the Engagements sheet, each of the 150 engagements has one row of 17 columns. Those columns record the delivery location, the type of financial institution, the outsourcing model, the class of work, the use case, the age of the engagement, and its lifecycle stage. Cancellation, the return on the engagement, customer satisfaction, headcount, and the number of respondents in each role fill the remaining columns.

The analysis uses the Engagements sheet as the governing source when it conflicts with respondent-level records. The two sources agreed on lifecycle stage in all 501 comparable cases. Cancellation, customer satisfaction, return, and lifecycle stage were taken from the engagement record; this source hierarchy does not independently validate the recorded values.

Data preparation excluded one of 851 rows because it contained only a column mean, leaving 850 respondent records. Zero values in career and tenure fields were interpreted under the study rules as less than one year. Return was blank for the 120 engagements outside Production; among the 30 production records, zero and positive values were retained as recorded.

The 150 engagement records contain 121 distinct use-case strings. The computational engine matches each string exactly to the crosswalk in Study Meta.md, section 9, dated September 21, 2026, which assigns one of eight process families. The supplied metadata states that these assignments agree with the process_family values in the coded-answer records. This is an exact lookup against an established crosswalk, not a new first-match keyword classification during statistical computation.

Analytical classes determine the unit and treatment of a variable. Respondent-level indicators describe individual answers. Engagement attributes use one engagement-level value. Perceived facts describe the respondent's understanding of engagement conditions. "I do not know" identifies unavailable or uncertain information for that respondent; it does not establish that the organization lacked the information.

The statistical record identifies compute_statistics.py as the computational engine and records input fingerprints. Its v16 layer specifies random seed 20261006, sequential bootstrap settings, and Monte Carlo settings. The older base layer records different settings. These records support traceability but do not establish independent reproducibility in a separate environment.

4.4 Construction of the Four Core Study Measures

The four core measures are cancellation, customer satisfaction, engagement indicator mean, and intent to leave. Their original directions are retained in ordinary association reporting: cancellation and intent to leave are adverse outcomes. For formal comparisons of association strength, the measures are reoriented so that higher values consistently indicate more favorable outcomes. Ratings converted to the favorable direction are identified where used.

Cancellation is recorded as one binary value per engagement; 56 of 150 engagements were canceled (37.33%). Engagement-level analysis gives each engagement equal weight unless a sensitivity analysis specifies otherwise. Respondent-weighted comparisons answer a different question and are identified separately. Ratings on canceled engagements were collected after cancellation, so associations with cancellation are retrospective descriptive contrasts.

Customer satisfaction is a recorded engagement score on a 100-point scale. Its mean was 43.85 points across all 150 engagements and 64.10 points across the 94 active engagements. The source field is labeled as a percentage, but these values should not be interpreted as the percentage of customers satisfied. Primary satisfaction associations use active engagements; all-engagement comparisons include post-cancellation scores and are descriptive.

The engagement indicator mean is constructed by averaging each respondent's available favorable-direction ratings across the 100 rated indicators, then averaging respondent means within each engagement. Study-level summaries give each engagement equal weight. The mean was 4.02 on the 10-point scale ($n = 150$ engagements), with a standard deviation of 1.48, median of 3.81, and range of 1.16–7.28. Differences in missing-item coverage can affect the composition of respondent means.

Each rated indicator is one of the 100 components of that mean. Tested against a mean that contains it, an indicator would be correlated in part with itself. When a rated indicator is tested, the mean is recomputed without it. Its companion analysis prints the correlation with the indicator included and the change. An indicator outside the 100 components is tested against the full mean.

Intent to leave is the answer to the question of how likely the respondent is to leave the current job in the next 12 months, rated from 1 to 10. Measured at respondent level, it is reported on its original scale. Its mean was 6.77 ($n = 798$ respondents). A rating of 7 or above was given by 430 of the 798 respondents (53.88%). An engagement-level form, the mean of each engagement's respondents, is reported alongside.

The four measures were strongly interrelated: all pairwise engagement-level correlations had absolute values of at least 0.72, and one component accounted for 88.23% of their variance ($n = 150$ engagements). The analysis also examined 300 partial associations between individual rated indicators and cancellation, satisfaction, or intent to leave, controlling for the engagement indicator mean. Thirty-three unadjusted p values were below 0.05, but none remained significant after false-discovery-rate correction. This does not establish that every independent effect is zero.

Supplementary results include recorded return, the share of distressed indicator ratings, and indicator quintiles. Of 30 production engagements, 17 had positive recorded returns, averaging 24.71%; the other 13 had recorded zero values. The reporting period, financial denominator, and treatment of costs were not established in the supplied analytical records. The mean distressed-rating share was 42.49%, calculated within respondents, then engagements, and then across engagements. Quintiles contain 30 engagements each, ordered by indicator mean with identifier-based tie breaking; differences in that mean across quintiles are true by construction.

4.5 Tests of Association and Difference

Association and difference tests generally use a two-sided significance threshold of 0.05; equivalence tests use two one-sided tests as described below. Under the specified assumptions, the approximate detectable correlation at 80% power is $|r| = 0.23$ for 150 independent engagements. This is a design calculation, not a guarantee of sensitivity in every analysis. Companion results report effect sizes, intervals, p values, and applicable bases; the main text summarizes selected results.

Numeric indicators were tested against each measure in the form that fits the measure. Against cancellation, the indicator's engagement mean was correlated with the cancellation flag across the 150 engagements as a point-biserial r . An odds ratio per rating point and the area under the curve accompany it. Fisher's exact test compared the cancellation rate of engagements distressed on the indicator with that of the other engagements.

Primary customer-satisfaction analysis correlates the indicator's engagement mean with satisfaction across 94 active engagements. Companion results also report a slope and squared correlation where applicable. Analysis against the engagement indicator mean uses all eligible engagements and excludes the indicator itself from that mean when it is a component.

Intent to leave is correlated with the indicator across the respondents who answered both questions. Its p value uses standard errors that allow for clustering by engagement, and its interval comes from a cluster bootstrap. Respondents distressed on the indicator are compared with the other respondents on the share at 7 or above. Beside the respondent-level result sits the engagement-level correlation, computed on engagement means.

Categorical indicators are compared across response levels. Engagement assignments use the unique most frequent response, with ties excluded. Cancellation comparisons use contingency-table tests, with Fisher's exact test for sparse two-by-two tables or Monte Carlo inference for larger sparse tables. Continuous-outcome comparisons use Welch's analysis of variance. Omega squared and its interval are calculated separately from the conventional analysis-of-variance decomposition. Levels with fewer than 10 eligible engagements are excluded from inferential comparisons and retained descriptively.

Segment comparisons cover outsourcing model, institutional type, process family, work class, engagement status and stage, respondent role, and vendor type. Numeric comparisons use Welch and

Kruskal–Wallis tests; categorical comparisons use contingency-table methods. Engagement-defined families use engagement-level values. Some role and vendor-type comparisons use respondents without clustering adjustments, whereas specified respondent-level outcome tests use cluster-robust methods. The analytical unit and adjustment must be read for each result.

Segment analyses use four status-and-stage groups: active engagements in Build-Out, Transition/Cut-Over, and Production, plus canceled engagements. The protocol suppresses segment intervals below 30 observations and reports groups below 10 descriptively. The supplementary return analysis is a separate analysis class and reports intervals for 17 positive-return engagements; those exploratory intervals require particular caution and should not be treated as meeting the segment-reporting threshold.

Formal pairwise comparisons of numeric-indicator associations use engagement-level versions of all four measures on a common available base, oriented so higher values are favorable. Steiger’s test compares dependent correlations sharing the indicator. These tests concern associations, not differences in causal importance. Categorical effect sizes are not compared with Steiger’s test.

The computational protocol also compares oriented correlations with the study’s median correlation using Fisher-z calculations. Because the indicator and reference correlations arise from shared data and correlated variables, these comparisons should be treated as descriptive benchmarks rather than evidence of a reliable ranking of indicators.

The v16 statistical layer specifies sequential simulation targets rather than a fixed number of draws for every result.

Bootstrap sampling uses batches of 10,000, a maximum of 500,000, and a target standard error of 0.0025 for each interval endpoint. Monte Carlo testing uses batches of 10,000, a maximum of 2,000,000, and a target simulation error of 0.0005 for p. Reaching a draw limit does not itself establish that the precision target was met; analysis-specific computation records govern that determination.

4.6 Correction for Multiple Testing

Multiple related tests increase the opportunity for false-positive findings. The study uses Holm adjustment within specified families, rather than treating every unadjusted p value below 0.05 as confirmatory evidence. Holm’s procedure orders the p values and applies sequential multiplicity factors while preserving monotonic adjusted values.

Four families receive the correction in each companion analysis.

1. The tests of difference across the segment families.
2. The association with one measure across the study set and every segment level, run separately for each of the four measures.
3. The six pairwise comparisons of the four associations.

4. The four comparisons with the study norm.

Holm adjustment applies to these defined families. The statistical record also retains unadjusted primary p values, and some main-text results report those values. An unadjusted p value should not be interpreted as adjusted unless explicitly identified. These within-analysis corrections do not control multiplicity across every indicator, written-response comparison, and analysis in the report.

For the 300 partial tests of the rated indicators net of the engagement indicator mean, the study controlled the false discovery rate, the expected share of false positives among the results that pass the test. Its smallest adjusted value was 0.323. The largest partial correlation in absolute value was 0.25.

4.7 Robustness Checks

The numeric-indicator protocol contains up to 22 checks distributed across the four measures. It does not apply 22 separate checks to every individual association. Checks are omitted or reported as noncomputable where an identity, unavailable base, or other constraint prevents estimation.

For a respondent-level numeric indicator with all checks computable, the protocol assigns five checks to cancellation, five to customer satisfaction, six to the engagement indicator mean, and six to intent to leave. The seven check types and their allocation are as follows.

1. Rank correlation in place of Pearson's r , applied to all four measures.
2. Non-answers excluded and then imputed at the engagement mean, applied to intent to leave.
3. Active engagements only, applied to the engagement indicator mean and intent to leave.
4. Removal of the most influential engagement, the one with the largest Cook's distance, applied to all four measures.
5. Engagements weighted by the number of their respondents, applied to cancellation, customer satisfaction, and the engagement indicator mean.
6. A bootstrap of engagements in place of the analytic interval, applied to all four measures.
7. Removal of the thin-base segments, applied to all four measures.

A check is recorded as holding when its result falls on the same side of the specified significance threshold as the primary test and, for a significant primary association, retains its direction. Interval-based checks assess whether zero is included. This criterion concerns consistency under the specified alternatives; it is not proof that an estimate is unbiased or that an association is causal. A bootstrap check is noncomputable if fewer than 90% of resamples are valid.

Thin-base segments are the levels of the outsourcing model or the process family with fewer than 30 engagements. Type of financial institution is not used for that check, since every one of its levels holds fewer than 30 engagements.

Categorical indicators received checks suited to a comparison of levels. In one, the Kruskal-Wallis test replaced Welch's analysis of variance. Another repeated the test on active engagements only. A third repeated it at respondent level, as a Wald F test with standard errors clustered by engagement.

4.8 Equivalence Tests

A nonsignificant association test does not establish absence of an effect. The study's equivalence analysis asks whether an association is sufficiently small to fall within a prespecified region. For correlations, the region is $-0.30 < r < 0.30$. This is a study convention; effects within that interval may still matter in particular operational contexts.

For correlations, both one-sided tests must reject their respective boundary nulls at 0.05 to support equivalence within the stated interval. Equivalence does not mean a literal zero correlation. Categorical comparisons use an omega-squared bound of 0.09; for cancellation, this is calculated from the binary flag and is distinct from the contingency-table significance test. A small association can be statistically different from zero and still meet the equivalence criterion.

Each equivalence table prints the achieved power of the test and the smallest effect it could detect at 80% power, beside the detection floor of 0.23 at 150 engagements. Intent to leave was tested on its effective sample size rather than its count of respondents, since clustered answers carry less information than independent ones.

When an association is nonsignificant and does not meet the equivalence criterion, it is inconclusive. Limited sample size can produce this result. Ten of the 19 structural attributes examined met the study's equivalence criteria for all four core measures; the remaining results require measure-specific interpretation.

4.9 Clustering Within Engagements

Respondents within an engagement share circumstances, so their observations may be correlated. The protocol estimates an intraclass correlation and approximates the design effect as $1 + (\text{mean cluster size} - 1) \times \max(\text{ICC}, 0)$. Nominal sample size divided by this design effect gives an approximate effective sample size. The calculation summarizes information loss under its assumptions; it does not replace a cluster-adjusted model or fully capture unequal cluster sizes and higher-level dependence.

Clustering more than halved the information in respondent answers on the two indicators that follow. On whether the engagement was within budget, the intraclass correlation was 0.61 and the design effect

2.63. Its 554 respondent answers carried the information of about 211 independent answers. For the quality of the financial institution's data across the systems the engagement depends on, the intraclass correlation was 0.57. Those 795 answers carried the information of about 230.

Aggregation removes duplication of respondent observations within an engagement but does not establish independence between engagements sharing institutions or vendors. Respondent-level intent-to-leave analyses use cluster-robust standard errors and cluster bootstrap intervals. Other intervals and tests follow their stated analytical units and methods.

Some segment comparisons by role and vendor type treat respondents as independent. Effective-sample-size summaries describe the extent of clustering but do not correct those tests. Separately reported cluster-robust outcome comparisons, including the role analysis, should not be confused with these unadjusted segment tests.

4.10 Treatment of the Written Answers

The data contain 8,918 written entries across the open-response and use-case fields. This is distinct from the broader total of 12,088 entries, which also includes 3,170 numbers entered in text boxes.

Under written coding rules, 6,908 were coded. Another 615 were stock non-answers, such as “none,” “n/a,” or “same as above,” classified by exact text. Answers on the laws and regulations governing the engagement, 548 in all, were not coded, and the obligations each named were extracted into a list. Use-case answers, 847 in all, were classified into the eight process families.

Rulebook version 0.5 linked answers to 211 measures, including 149 scored measures, 52 additional themes, seven expressed emotional states, and four non-answer types. Other-engagement accounts were flagged and excluded from links to the respondent's own engagement. Initial rules were fixed before production coding, but later documented revisions expanded frequency-theme wording and added sentiment scores under version 0.6.

A mark needs the answer's own words to state it, and a meaning that could be inferred was not enough. Where the words fit two measures without settling which one the respondent meant, neither was marked.

AI language-model coders using Claude processed batches of answers under the rulebook. The documented workflow required five steps.

1. The question and the whole answer were read.
2. The answer was split into statements, each the shortest run of the respondent's own words that makes one point. Together the statements rebuild the answer word for word.
3. Each statement was marked with every measure, theme, and emotional state its words state.
4. Each marked measure with a favorable and an unfavorable end was scored.

5. A second pass checked each assigned mark against the rulebook. This was an AI coding procedure, not independent human review of every answer.

Scores take five values. A score of -10 describes a subject at or near its worst, and a score of -5 a clear but partial problem. Zero marks no position, or good and bad in balance. A score of +5 describes a clear but partial strength, and a score of +10 a subject at or near its best. Either extreme needs words of totality or constancy, serious harm, or the respondent's strongest language.

Coding produced 28,513 statements and 22,024 marks across 6,908 answers. Automated validation checked permitted codes and scores, placement of marks, and exact reconstruction of the source answer; all 153 output files passed. These structural checks do not establish semantic accuracy of every code.

A separate AI coder, without access to the first coding, recoded 681 production answers and 100 calibration answers. Across the 781 double-coded answers, Krippendorff's alpha was 0.919 for topical marks, 0.926 for themes, 0.904 for emotional states, and 0.969 for scores, exceeding the prespecified 0.80 threshold. Production topical agreement was 0.916. These results measure consistency between AI coding runs; they are not independent human validation or proof of construct validity.

The two AI codings differed on at least one mark or score in 288 of the 681 production answers. High aggregate agreement can coexist with substantial record-level disagreement.

An author-ruled set of 48 difficult answers served as an additional check, and the final coding reproduced all 48 rulings. This targeted check does not establish accuracy across the full corpus.

On October 9, 2026, rulebook version 0.6 added 3,599 sentiment scores to previously unscored marks on 34 measures and an overall score to each of 6,874 substantive answers. Separate AI recoding of 777 answers yielded alpha values of 0.945 for overall scores and 0.930 for 429 paired additional measure scores. These additions were documented after the initial coding and should not be described as part of an unchanged original protocol.

A written response is an answer about the respondent's own engagement coded to an indicator. Primary responses come from the designated section's open question; secondary responses come from other questions. Counts describe the coded material and may include several answers per respondent. Structured ratings measure responses to their own wording, not the prevalence of every practice described in the comments. Fewer than 10 written responses support illustrative description only; larger qualitative samples remain self-selected.

Respondent quotations retain the supplied spelling and grammar. Ellipses identify omissions, and the supplied material includes the name substitutions described below. Attributions identify role, outsourcing model, and process family.

The study's documented preparation process states that a third party replaced financial-institution and vendor names before the material reached the author. The quotations retain those supplied

substitutions. This describes deidentification of the analytical material, not a finding that no identifying information existed at any earlier stage.

4.11 Limits of the Evidence

The study primarily records vendor staff members' reported experiences, supplemented by the engagement-level fields in the analytical record. It does not include independent observation of delivery, inspection of institutional records, or validation by the financial institutions served.

One source supplied the evidence. No financial institution, model provider, or third-party provider contributed. Customer satisfaction and the return on the engagement are the figures recorded for the engagement, and the study holds no record from the financial institution against which to confirm them.

The analysis is cross-sectional. Associations do not establish temporal order or causal direction. Retrospective five-year items are perceptions collected at one time, not repeated observations; their interpretation is further limited by the short careers reported by many respondents.

Selection covered a range of types with partial intent. Within the study set, the share of any type of financial institution, outsourcing model, or use case does not reflect its frequency in the industry. Findings describe the 150 engagements studied.

Ratings from the 56 canceled engagements were collected after cancellation and may reflect it. Cancellation associations are retrospective contrasts. Primary satisfaction associations use the 94 active engagements, which avoids combining those scores with post-cancellation scores but introduces its own restriction to surviving engagements.

The measures and indicators were strongly interrelated. None of the 300 partial indicator–outcome associations remained significant after false-discovery-rate correction. This limits claims of distinct explanatory contributions; it does not establish that independent effects are absent. Descriptive rankings of means or coefficients should not be interpreted as tested rankings of importance.

Role and experience were closely related. Mean career experience was 1.9 years for individual contributors, 10.0 for engagement managers, and 15.1 for portfolio directors. The reported unadjusted comparisons do not isolate their respective contributions. Other shared respondent and engagement characteristics may also affect associations.

Missing and nonnumeric responses alter analytical coverage. For example, 554 respondents rated whether the engagement was within budget, while 242 answered “I do not know.” Such an answer may reflect lack of information or uncertainty, rather than organizational ignorance. The study does not establish that excluded responses are missing at random.

Several groups have limited coverage. The two customer management executives support counts only. Portfolio-director analyses include at most 31 respondents, with clustering reducing effective

information. Every institutional-type group contains fewer than 30 engagements, and three process families contain fewer than 10. Study thresholds determine which results are descriptive or eligible for inference.

Restricted variation limits inference. For example, every recorded answer on vendor experience in agentic AI outsourcing was two years or less ($n = 571$ respondents). Such a narrow range can reduce the ability to identify an association and cannot establish the value of substantially greater experience.

Respondent-level figures carry less information than their counts suggest. Clustering reduced 554 answers to the equivalent of about 211 independent answers, and 795 answers to about 230, on the two indicators of section 4.9.

Written answers are self-selected and influenced by question framing. Frequency-theme counts depend on documented wording rules, including a post-coding expansion, and may miss descriptions outside those rules. AI coding agreement measures consistency, while structural validation checks format and source reconstruction; neither establishes independent semantic accuracy for the full corpus.

Disclosure is measured through respondents' accounts of their own or management's discussion. Structured answers and coded written accounts have different bases and may differ. Neither establishes all information available to a financial institution through other channels.

No regulatory or legal breach can be determined from the study. Obligations the respondents named are reported as context for an audit and never as a legal conclusion.

The sample covers Build-Out, Transition/Cut-Over, and Production, not directly observed preparation or procurement stages. Return data comprise 30 production records, including 17 positive and 13 zero values, with incomplete documentation of the financial measure. Findings cannot establish industry-wide returns, prospective cancellation risk, or the effectiveness of the proposed executive actions.

5 Appendix E: Companion Indicator Statistical Analyses

The 216 companion analyses provide indicator-level statistics and related derived analyses. Each follows a common structure, adapted to the type of measure and available evidence.

Analyses are grouped by the Part of the report that summarizes them. A first group of 46 analyses (Part 1) covers the three study measures computed from the engagement record, seven engagement fields, 31 survey questions on the engagement and the financial institution, and five derived analyses. Those derived analyses treat the active engagements, the four measures together, the indicator quintiles, measured return, and the associations net of the engagement indicator mean.

The second group of 101 analyses (Part 2) covers the 100 rated indicators of the five failure categories and one question on the vendor's preparedness to manage outsourcing engagements. The third group holds 48 analyses (Part 4), 47 on questions about the delivery workforce and one a derived analysis of experience by intent to leave. The fourth group holds 21 analyses (Part 5), 19 on manual alteration, its disclosure, and reporting alteration, and two derived analyses of manual alteration by role.

A companion analysis identifies the measure, wording, analytical unit, scale, valid responses, and non-answers. It then provides descriptive results, eligible segment comparisons, associations with the four core measures, written-evidence summaries, and applicable diagnostic checks.

Statistical sections include clustering, equivalence, robustness, and comparisons with the study norm where applicable. Exclusion of an item from a composite reduces mechanical self-inclusion but does not make the variables independent. Co-occurrence and written-sentiment comparisons provide further descriptive context.

Each companion analysis includes interpretation, limitations, definitions, and source information. Source sections identify input files and computational records, with fingerprints, computation dates, and template versions where available. These identifiers support traceability; they do not by themselves establish public availability of every input or independent reproducibility.

Appendix B maps each indicator to its main-text section and companion-analysis filename. The online index below provides access to the published pages.

The 216 companion analyses are available through Akholi's "Agentic AI Outsourcing Study Data" index at <https://akholi.com/outsourcing-2026-agentic-ai-tier-1-banks/study-data/> (accessed October 16, 2026). The index also links related derived analyses. Online access to analysis pages should be distinguished from access to the underlying respondent data, computational environment, and all materials needed for independent reproduction.

Table 36. Register of the companion indicator statistical analyses

<i>Part</i>	<i>Companion analysis</i>	<i>Subject</i>	<i>Where reported</i>
Part 1	Analysis: Active Engagements	Analysis: Active Engagements	Part 1, 4.2.3; Introduction, 4.4.2; Part 1, 2
Part 1	Analysis: Four Measures	Analysis: Four Measures	Part 1, 4.2.1
Part 1	Analysis: Mean Quintiles	Analysis: Mean Quintiles	Part 1, 4.2.2

<i>Part</i>	<i>Companion analysis</i>	<i>Subject</i>	<i>Where reported</i>
Part 1	Analysis: Measured Return	Analysis: Measured Return	Part 1, 3.3.1; Introduction, 4.5.5
Part 1	Analysis: Net Associations	Analysis: Net Associations	Part 1, 4.3.1
Part 1	C031	What type of a bank is the customer?	Part 1, 5.1.1
Part 1	C032	What is the outsourcing type?	Part 1, 5.3.1
Part 1	C033	Is the outsourcing engagement Corporate Functional, or Line of Business Operational?	Part 1, 5.3.2
Part 1	C035	How long ago did this specific engagement start, in months?	Part 1, 1.1.2
Part 1	C036	What stage is the engagement in?	Part 1, 1.1.1; Introduction, 4.5.2
Part 1	C037	How many months has the engagement been in its current stage?	Part 1, 1.1.3
Part 1	C038	Per the current plan, how many months from today until the engagement reaches Production/Delivery?	Part 1, 1.1.4
Part 1	C039	If the engagement has not been canceled, how likely is the customer to cancel the engagement?	Part 1, 2.2.1
Part 1	C040	How many people are on the vendor team for this engagement today?	Part 1, 5.6.1
Part 1	C041	What was the peak vendor team headcount on this engagement to date?	Part 1, 5.6.2
Part 1	C042	Per the current plan, how many people will be on the vendor team 12 months from now?	Part 1, 5.6.3
Part 1	C043	How many AI agents are running in production on this engagement today?	Part 1, 1.2.1
Part 1	C044	What is the pricing model for this engagement?	Part 1, 5.4.1
Part 1	C045	What is the contract term of this engagement, in months?	Part 1, 5.4.2
Part 1	C046	What is the annual contract value of this engagement, in USD?	Part 1, 5.4.3
Part 1	C047	I know the KPIs the customer uses to measure success.	Part 1, 3.2.4
Part 1	C048	The customer is happy with the performance of the engagement at this stage.	Part 1, 3.1.2
Part 1	C049	The engagement is meeting delivery dates successfully.	Part 1, 3.2.1
Part 1	C050	The engagement is within budget.	Part 1, 3.2.2
Part 1	C051	The engagement is not achieving stated KPIs.	Part 1, 3.2.3
Part 1	C062	What is the customer's approximate total asset size, in USD?	Part 1, 5.1.2
Part 1	C063	Does the customer operate its own global capability centers?	Part 1, 5.1.3
Part 1	C064	How many months has your employer been serving this customer across all work, not just this engagement?	Part 1, 5.2.1
Part 1	C065	How many agentic AI outsourcing engagements does your employer run concurrently with this customer today?	Part 1, 5.2.2
Part 1	C066	The customer is scaling their agentic AI outsourcing engagements.	Part 1, 5.1.5
Part 1	C067	The customer is increasing the number of agentic AI outsourcing engagements.	Part 1, 5.1.6
Part 1	C068	The customer is shifting more work to their global capability centers.	Part 1, 5.1.4
Part 1	C076	What type of company is your employer?	Part 1, 5.5.1
Part 1	C077	Approximately how many employees does your employer have worldwide?	Part 1, 5.5.2
Part 1	C078	How many years has your employer been delivering agentic AI outsourcing?	Part 1, 1.2.2
Part 1	C079	How many agentic AI engagements does your employer run for banks today?	Part 1, 5.5.3
Part 1	Field: Bank Type	Field: Bank Type	Part 1, 5.1.1
Part 1	Field: Class of Work	Field: Class of Work	Part 1, 5.3.2
Part 1	Field: Engagement Age	Field: Engagement Age	Part 1, 1.1.2
Part 1	Field: Engagement Phase	Field: Engagement Phase	Part 1, 1.1.1
Part 1	Field: Headcount	Field: Headcount	Part 1, 5.6.1
Part 1	Field: Outsourcing Type	Field: Outsourcing Type	Part 1, 5.3.1
Part 1	Field: Use Case	Field: Use Case	Part 1, 5.3.3
Part 1	Measure: Cancellation	Measure: Cancellation	Part 1, 2.1.1
Part 1	Measure: Indicator Mean	Measure: Indicator Mean	Part 1, 4.1.1; Introduction, 4.5.1
Part 1	Measure: Satisfaction	Measure: Satisfaction	Part 1, 3.1.1; Introduction, 4.5.3
Part 2	C093	Compared to 5 years ago, is my employer more or less prepared to manage outsourcing engagements?	Part 2, 4.8.1
Part 2	C094	Rate the quality of the customer's data across the systems this engagement depends on.	Part 2, 2.1.1
Part 2	C095	Rate how accessible the customer's data is to this engagement when and where the work requires it.	Part 2, 2.1.2
Part 2	C096	Rate how well the customer's data is normalized, with consistent formats and definitions across source systems.	Part 2, 2.1.3
Part 2	C097	Rate the maturity of the customer's data governance over the data this engagement uses.	Part 2, 2.1.4
Part 2	C098	Rate how well the customer's legacy systems support the integration this engagement requires.	Part 2, 2.2.1
Part 2	C099	Rate the readiness of the customer's technical infrastructure to support the agentic AI this engagement deploys.	Part 2, 2.2.2
Part 2	C100	Rate how completely the customer's processes are mapped and documented to the depth agentic AI requires.	Part 2, 2.3.1
Part 2	C101	Rate the maturity of the customer's processes in scope, meaning how standardized and repeatable they are.	Part 2, 2.3.2
Part 2	C102	Rate how consistently the customer performs the same processes across its business units and regions.	Part 2, 2.3.3
Part 2	C103	Rate the customer's willingness to redesign processes for agentic AI rather than automate them as they are.	Part 2, 2.3.4
Part 2	C104	Rate the agentic AI knowledge and skills of the customer staff working with this engagement.	Part 2, 2.4.2
Part 2	C105	Rate the effectiveness of the customer's management of its side of this engagement.	Part 2, 2.5.1

Part	Companion analysis	Subject	Where reported
Part 2	C106	Rate the availability of the customer's subject matter experts when this engagement needs them.	Part 2, 2.4.3
Part 2	C107	Rate the customer organization's readiness to absorb the changes to daily work that agentic AI introduces.	Part 2, 2.4.1
Part 2	C108	Rate the customer's ability to provision the system access and identities that AI agents require.	Part 2, 2.7.1
Part 2	C109	Rate how quickly the customer responds when this engagement raises issues that need customer action.	Part 2, 2.5.2
Part 2	C110	Rate how well the customer's change control process handles the changes this engagement requires.	Part 2, 2.6.3
Part 2	C111	Rate the customer's capacity to govern and oversee the risks of AI agents doing this work.	Part 2, 2.6.2
Part 2	C112	Rate how clearly ownership and accountability for AI are assigned inside the customer organization.	Part 2, 2.6.1
Part 2	C113	Rate the stability of the customer's business case and ROI definition for this engagement.	Part 2, 2.8.1
Part 2	C115	Rate the overall effectiveness of your employer's top leadership team.	Part 2, 4.1.1
Part 2	C116	Rate how well your employer is transforming its commercial model as agentic AI reduces the billable human effort it sells today.	Part 2, 4.1.2
Part 2	C117	Rate how well your employer's internal management processes have been redesigned for agentic AI delivery rather than traditional staffed delivery.	Part 2, 4.2.1
Part 2	C118	Rate the effectiveness of your employer's process and leadership development programs.	Part 2, 4.5.2
Part 2	C119	Rate your employer's ability to manage agentic AI outsourcing engagements end to end.	Part 2, 4.2.2
Part 2	C120	Rate the depth of your employer's knowledge of the financial industry.	Part 2, 4.3.1
Part 2	C121	Rate your employer's ability to manage agentic AI outsourcing engagements for global banks in particular.	Part 2, 4.2.3
Part 2	C122	Rate your employer's current supply of qualified agentic AI talent against what its engagements need.	Part 2, 4.5.1
Part 2	C123	Rate the adequacy of your employer's investment in training its workforce.	Part 2, 4.5.3
Part 2	C124	Rate how well your employer's training programs build practical agentic AI skills.	Part 2, 4.5.4
Part 2	C125	Rate how well your employer's training programs build financial industry knowledge.	Part 2, 4.5.5
Part 2	C126	Rate the maturity of your employer's playbook for taking agentic use cases from concept to governed production.	Part 2, 4.6.1
Part 2	C127	Rate your employer's quality assurance capability for agents whose outputs are not identical every run.	Part 2, 4.6.2
Part 2	C128	Rate your employer's ability to price agentic work profitably when billable hours no longer measure the value delivered.	Part 2, 4.1.3
Part 2	C129	Rate how well the technical platforms your employer builds or licenses are optimized for agentic AI outsourcing.	Part 2, 4.6.3
Part 2	C130	Rate your employer's readiness to take accountability when an agent it operates makes a costly error.	Part 2, 4.7.1
Part 2	C131	Rate the depth of your employer's knowledge of agentic AI.	Part 2, 4.3.2
Part 2	C132	Rate the depth of your employer's experience running agentic AI systems in live production, beyond pilots.	Part 2, 4.4.1
Part 2	C133	Rate your employer's ability to operate large numbers of agents in production, including monitoring and exception handling.	Part 2, 4.4.2
Part 2	C134	Rate your employer's overall qualification to deliver elite agentic AI outsourcing services to banks.	Part 2, 4.2.4
Part 2	C136	Rate how effectively the customer and vendor jointly manage this engagement now that agents perform part of the work.	Part 2, 3.1.1
Part 2	C137	Rate how clearly the two firms have settled who decides what agents may do without human approval.	Part 2, 3.1.2
Part 2	C138	Rate how clearly ownership of each workflow step is assigned when agents, vendor staff, and customer staff work in the same process.	Part 2, 3.1.3
Part 2	C139	Rate how clearly accountability lands on one firm when an agent's action causes a problem in this engagement.	Part 2, 3.1.4
Part 2	C140	Rate how well the two firms maintain a single shared view of the risks that agent work introduces.	Part 2, 3.2.1
Part 2	C141	Rate how effectively the two firms manage issues arising from agent work as one team rather than through two separate processes.	Part 2, 3.2.2
Part 2	C142	Rate how well the speed of joint issue resolution matches the speed at which agent-caused problems can spread.	Part 2, 3.2.3
Part 2	C143	Rate how quickly the two firms together can stop an agent that is behaving badly in live operations.	Part 2, 3.2.4
Part 2	C144	Rate how well the joint change control process keeps pace with the speed of agent behavior changes.	Part 2, 3.3.1
Part 2	C145	Rate how well agent updates made by one firm become visible to the other firm before they take effect.	Part 2, 3.3.2
Part 2	C146	Rate how well the two firms jointly govern changes agents make to their own behavior between formal releases.	Part 2, 3.3.3
Part 2	C147	Rate how well human review points are placed across the two firms at the moments where agent errors would matter most.	Part 2, 3.4.1
Part 2	C148	Rate how well joint human review capacity keeps up with the volume of output the agents produce.	Part 2, 3.4.2

<i>Part</i>	<i>Companion analysis</i>	<i>Subject</i>	<i>Where reported</i>
Part 2	C149	Rate how much genuine scrutiny agent outputs receive at human review points rather than routine approval.	Part 2, 3.4.3
Part 2	C150	Rate how reliably the joint quality assurance loop catches agent errors before they reach live operations.	Part 2, 3.4.4
Part 2	C151	Rate how well the quality measures both firms use capture the kinds of errors agents actually make.	Part 2, 3.4.5
Part 2	C152	Rate how fully each firm can see what the other firm's agents actually did when questions arise.	Part 2, 3.5.1
Part 2	C153	Rate how closely the speed of joint decision making matches the speed at which the agents act.	Part 2, 3.5.2
Part 2	C154	Rate how well today's joint governance model fits work performed by agents rather than work performed by people.	Part 2, 3.5.3
Part 2	C155	Rate how ready the joint management model is for the level of agent autonomy you expect in this engagement within two years.	Part 2, 3.5.4
Part 2	C157	Rate the factual reliability of the AI tools' outputs on this engagement before any human review.	Part 2, 5.2.1
Part 2	C158	Rate the accuracy of citations, references, and quotations the tools produced against the real sources.	Part 2, 5.2.2
Part 2	C159	Rate the consistency of the tools' outputs when given the same input and instructions more than once.	Part 2, 5.2.3
Part 2	C160	Rate the end-to-end reliability of the tools on multi-step tasks, from first step to finished output.	Part 2, 5.2.4
Part 2	C161	Rate the predictability of the tools' token consumption for a given volume of work.	Part 2, 5.4.1
Part 2	C162	Rate the efficiency of the tools' token use, considering retries, loops, and reprocessing needed to complete work.	Part 2, 5.4.2
Part 2	C163	Rate how completely the tools kept each client's and each session's information isolated from every other client and session.	Part 2, 5.5.1
Part 2	C164	Rate how reliably the tools retained instructions and constraints given earlier in long tasks and sessions.	Part 2, 5.3.1
Part 2	C165	Rate how completely the tools captured material facts located anywhere in long documents, including mid-document content.	Part 2, 5.3.2
Part 2	C166	Rate the integrity of the tools' persistent memory against contamination by incorrect or malicious content encountered during work.	Part 2, 5.3.3
Part 2	C167	Rate the tools' resistance to hostile instructions embedded in the content they read and process.	Part 2, 5.5.3
Part 2	C168	Rate how consistently the agents took only intended and authorized actions when operating connected tools and systems.	Part 2, 5.6.1
Part 2	C169	Rate how safely the tools handled sensitive data, with no exposure through outputs, logs, or connected systems.	Part 2, 5.5.2
Part 2	C170	Rate how reliably the tools' built-in guardrails and safety controls held during production use.	Part 2, 5.6.2
Part 2	C171	Rate the overall production readiness of the AI tools for bank-grade work as experienced on this engagement.	Part 2, 5.1.1
Part 2	C172	Rate how accurately the tools' own reports of success or failure matched the actual outcomes of their work.	Part 2, 5.8.1
Part 2	C173	Rate the tools' ability to explain and evidence how each output was produced, at the depth an audit requires.	Part 2, 5.8.2
Part 2	C174	Rate how well the tools supported reliable testing, meaning correct behavior could be proven before work went live.	Part 2, 5.7.1
Part 2	C175	Rate the stability of the tools' behavior over the engagement when nothing changed on the engagement side.	Part 2, 5.7.2
Part 2	C176	Rate the quality of the tools' logs, traces, and usage telemetry for detecting and diagnosing failures.	Part 2, 5.8.3
Part 2	C178	Rate how well the agents on this engagement are protected from breaking silently when a third party changes something without warning.	Part 2, 6.2.1
Part 2	C179	Rate the quality of the fallback the engagement has ready when a critical third-party feed, platform, or model stops working.	Part 2, 6.7.1
Part 2	C180	Rate the security posture of the third-party tools, platforms, and sub-processors the engagement's agents connect to.	Part 2, 6.6.1
Part 2	C181	Rate the quality of advance notice third-party providers give before changes that can affect the agents on this engagement.	Part 2, 6.1.1
Part 2	C182	Rate how reliably word of an upcoming third-party change reaches the people who operate the affected agents in time to act.	Part 2, 6.1.2
Part 2	C183	Rate the engagement's monitoring for unannounced changes in the third-party APIs, data feeds, and models its agents depend on.	Part 2, 6.2.2
Part 2	C184	Rate how well the contracts with third-party providers deal with provider changes that can break the agents on this engagement.	Part 2, 6.4.1
Part 2	C185	Rate how well the engagement absorbs model version deprecations and retirements forced by model providers.	Part 2, 6.3.2
Part 2	C186	Rate the availability of true machine interfaces, rather than screens built for humans, across the third-party systems the agents must use.	Part 2, 6.5.1

<i>Part</i>	<i>Companion analysis</i>	<i>Subject</i>	<i>Where reported</i>
Part 2	C187	Rate the stability of the third-party APIs and data structures the agents depend on, from one provider release to the next.	Part 2, 6.3.1
Part 2	C188	Rate the quality of the data arriving from third-party feeds, judged by whether agents can use it without human cleanup.	Part 2, 6.5.2
Part 2	C189	Rate how well third-party rate limits and quotas hold up under the volume the agents on this engagement generate.	Part 2, 6.5.3
Part 2	C190	Rate the engagement's ability to still perform by hand the manual workarounds that once absorbed third-party quirks and breakage.	Part 2, 6.7.2
Part 2	C191	Rate how completely the engagement knows where humans used to quietly fix third-party data and interface problems before agents took over.	Part 2, 6.7.3
Part 2	C192	Rate how deliberately the quiet human fixes for third-party quirks were replaced with explicit checks when agents took over the work.	Part 2, 6.7.4
Part 2	C193	Rate how well the engagement manages its dependence on a small number of model and cloud providers.	Part 2, 6.6.2
Part 2	C194	Rate how well this engagement could keep operating if it had to switch its primary model or cloud provider.	Part 2, 6.6.3
Part 2	C195	Rate how well third-party licenses on this engagement cover work done by agents rather than named human users.	Part 2, 6.4.2
Part 2	C196	Rate how completely the engagement has mapped every third and fourth party its agents depend on, down the full chain.	Part 2, 6.8.1
Part 2	C197	Rate how well the engagement manages agent dependence on gated external rails such as payment networks and regulatory submission systems.	Part 2, 6.8.2
Part 4	Analysis: Experience by Intent	Analysis: Experience by Intent	Part 4, 6.1.1
Part 4	C005	What is your role on this engagement?	Part 4, 1.1.1
Part 4	C007	How many years of experience do you have in your career?	Part 4, 1.2.1
Part 4	C008	How many years of experience do you have with agentic AI?	Part 4, 1.3.1
Part 4	C009	How many years have you been working in the outsourcing industry?	Part 4, 1.2.2
Part 4	C010	How many years have you been working within the financial industry?	Part 4, 1.2.3
Part 4	C011	How many agentic AI outsourcing engagements have you worked on across your career, including this one?	Part 4, 1.3.2
Part 4	C012	How well do you know the financial industry?	Part 4, 1.4.1
Part 4	C013	How well do you know agentic AI specific to the financial industry?	Part 4, 1.4.2
Part 4	C014	Rate your overall agentic AI expertise.	Part 4, 1.4.3
Part 4	C015	What professional certifications do you have?	Part 4, 1.5.1
Part 4	C016	What AI certifications do you have?	Part 4, 1.5.2
Part 4	C017	Where did you gain your AI knowledge?	Part 4, 1.5.3
Part 4	C018	How happy are you with your current job?	Part 4, 3.1.1
Part 4	C019	How likely are you to leave your current job in the next 12 months?	Part 4, 5.1.1; Introduction, 4.5.4
Part 4	C020	How happy is your extended team with the current engagement?	Part 4, 3.1.2
Part 4	C021	How likely are members of your current team to leave their jobs in the next 12 months?	Part 4, 5.1.2
Part 4	C022	Compared to 5 years ago, I am happier with the work I am doing now.	Part 4, 3.2.1
Part 4	C023	Compared to 5 years ago, I am more likely to seek a new job.	Part 4, 5.2.1
Part 4	C024	Compared to 5 years ago, the team I work with seems to be happier.	Part 4, 3.2.2
Part 4	C025	Compared to 5 years ago, the team I work with is more likely to seek a new job.	Part 4, 5.2.2
Part 4	C026	I want to continue working on projects in the global financial industry.	Part 4, 5.3.1
Part 4	C027	I would prefer to work directly for a bank in one of their global capability centers.	Part 4, 5.3.2
Part 4	C028	Working directly for a bank would be better than working for an outsourcing vendor.	Part 4, 5.3.3
Part 4	C069	I trust the customer.	Part 4, 4.1.1
Part 4	C070	The customer knows and cares about me.	Part 4, 4.1.2
Part 4	C071	I talk with peers inside the customer regularly.	Part 4, 4.2.1
Part 4	C072	I talk with engagement management inside the customer regularly.	Part 4, 4.2.2
Part 4	C073	I talk with the customer's leadership team regularly.	Part 4, 4.2.3
Part 4	C074	I know how the customer feels about my team.	Part 4, 4.1.3
Part 4	C075	I know how the customer feels about the engagement I am on.	Part 4, 4.1.4
Part 4	C080	My manager is effective.	Part 4, 2.1.1
Part 4	C081	My manager knows the financial industry well.	Part 4, 2.1.2
Part 4	C082	My manager knows agentic AI well.	Part 4, 2.1.3
Part 4	C083	My manager cares about me.	Part 4, 4.3.1
Part 4	C084	My manager cares about the larger team.	Part 4, 4.3.2
Part 4	C085	My manager knows how to manage agentic AI outsourcing engagements.	Part 4, 2.1.4
Part 4	C086	My manager knows how to manage agentic AI outsourcing engagements for global banks.	Part 4, 2.1.5
Part 4	C087	My employer cares about me.	Part 4, 4.4.1
Part 4	C088	My employer cares about the team.	Part 4, 4.4.2
Part 4	C089	My employer's top leadership team cares about my wellbeing and development.	Part 4, 4.4.3
Part 4	C090	I trust my employer.	Part 4, 4.4.4
Part 4	C091	I trust my manager.	Part 4, 4.3.3
Part 4	C092	I am being used effectively by my employer.	Part 4, 4.4.5
Part 4	Field: CME Respondents	Field: CME Respondents	Part 4, 1.1.1
Part 4	Field: IC Respondents	Field: IC Respondents	Part 4, 1.1.1
Part 4	Field: PD Respondents	Field: PD Respondents	Part 4, 1.1.1

<i>Part</i>	<i>Companion analysis</i>	<i>Subject</i>	<i>Where reported</i>
Part 4	Field: Total Respondents	Field: Total Respondents	Part 4, 1.1.1
Part 5	Analysis: Alterations by Role	Analysis: Alterations by Role	Part 5, 2.1.1
Part 5	Analysis: Manager Alterations	Analysis: Manager Alterations	Part 5, 2.2.1
Part 5	C052	The dashboard and performance reports given to the customer are accurate.	Part 5, 1.4.1
Part 5	C053	The vendor manipulates the dashboard and performance reports given to the customer.	Part 5, 1.4.2
Part 5	C054	Compared to 5 years ago, is manipulation of the dashboards and performance reports given to the customer more or less common now?	Part 5, 1.4.3
Part 5	C057	I fully understand how these laws, regulations, and similar affect the outsourcing engagement.	Part 5, 7.1.1
Part 5	C058	My management team fully understands these laws, regulations, and similar, and how they affect the engagement.	Part 5, 7.1.2
Part 5	C059	The customer's team actively discusses these laws, regulations, and similar with our team.	Part 5, 7.2.1
Part 5	C060	The training available to me effectively teaches me about these laws, regulations, and similar.	Part 5, 7.3.1
Part 5	C199	I have altered client-provided data	Part 5, 1.1.1
Part 5	C200	I discussed the data alteration with internal management	Part 5, 4.1.1
Part 5	C201	I or my management discussed data alteration with the client	Part 5, 4.2.1
Part 5	C203	My greater team regularly alters client provided data.	Part 5, 1.1.2
Part 5	C205	I have altered client processes	Part 5, 1.2.1
Part 5	C206	I discussed the process alteration with internal management	Part 5, 4.1.2
Part 5	C207	I or my management discussed process alteration with the client	Part 5, 4.2.2
Part 5	C209	My greater team regularly alters client processes.	Part 5, 1.2.2
Part 5	C211	I have altered AI outputs	Part 5, 1.3.1
Part 5	C212	I discussed the AI output alteration with internal management	Part 5, 4.1.3
Part 5	C213	I or my management discussed AI output alteration with the client	Part 5, 4.2.3
Part 5	C215	My greater team regularly alters AI outputs.	Part 5, 1.3.2

6 Appendix F: Index

The index lists each term defined in Appendix A with the Part and section in which it appears. A term used in more than 40 sections is listed by Part.

Table 37. Index of defined terms

<i>Term</i>	<i>Where it appears</i>
10-point scale	Introduction: 3, 4, 4.1.2, 4.3.1, 4.5.1; Part 1: 4; Part 2: 1, 2, 4; Part 4: 1.1.1, 5
Active	Introduction; Part 1; Part 2; Part 4; Part 5; Part 6
Agent	Introduction; Part 1; Part 2; Part 3; Part 4; Part 5; Part 6
Agentic AI outsourcing	Introduction: 3, 4, 4.3.1, 4.7.5; Part 1: 1, 1.2, 1.2.2, 5.1, 5.1.5, 5.1.6, 5.2.2; Part 2: 4.2.2, 4.2.3, 4.2.4, 4.6.3; Part 4: 1.3, 1.3.2, 2.1.4, 2.1.5; Part 6: 1, 2.1
AI technology	Introduction: 3; Part 1: 4; Part 2: 1, 1.1, 5, 5.2.2, 5.3, 5.3.2, 5.6, 5.6.2, 6.8.2; Part 5: 3.4
Anecdote strength	Introduction: 4.1.1; Part 2: 3.5.4, 4.5.2, 4.8.1, 5.3.3, 5.5.3, 6.6.1, 6.7.4; Part 3: 5.1; Part 4: 1.2.1, 1.3.2, 1.4.2, 4.3.3, 4.4.1, 4.4.2; Part 5: 7.3.1
Base	Introduction; Part 1; Part 2; Part 3; Part 4; Part 5; Part 6
Build-Out	Introduction: 4.1.3; Part 1: 1, 1.1, 1.1.2, 1.1.3, 1.1.4, 1.2.1; Part 2: 1.1, 2.1.4, 3, 3.4.3; Part 4: 5.1; Part 5: 1.2.1
Canceled	Introduction; Part 1; Part 2; Part 4; Part 5
Canceled or at risk	Part 1
Cancellation likelihood	Introduction: 4.4.1, 4.4.2; Part 1: 2, 2.2, 2.2.1
Cancellation rate	Introduction; Part 1; Part 2; Part 4; Part 5
Chi-square test	Introduction: 4.7.3
Class of work	Part 1: 2.1.1, 5, 5.3, 5.3.2, 5.6.1
Clean estimate	Introduction: 4.2.5, 4.5.3, 4.7.1, 4.8.3
Cluster bootstrap	Part 1; Part 2; Part 4
Co-management	Introduction: 3; Part 1: 4; Part 2: 1, 1.1, 2.6, 3, 3.2.2, 3.5.4, 5; Part 3: 1, 1.2; Part 6: 1, 1.2, 2, 2.3, 2.4, 2.5
Companion analysis	Introduction: 4, 4.7, 4.8.1, 4.8.2; Part 1: 5.3.4
Confidence interval	Introduction: 4.7.1
Customer management executive	Part 4: 1.1.1
Customer satisfaction	Introduction; Part 1; Part 2; Part 4; Part 5
Delivery team	Introduction; Part 1; Part 2; Part 3; Part 4; Part 5; Part 6
Delivery workforce	Introduction: 3; Part 1: 5.6; Part 2: 4.5; Part 4: 1, 1.2.1
Descriptive contrast	Introduction: 4.2.5, 4.7.1
Design effect	Appendix A only
Detection floor	Appendix A only
Did not hold	Introduction: 4.2.5, 4.3.3, 4.7.2; Part 1: 1.1.4, 5, 5.2.1, 5.6.2; Part 5: 4.1.1
Distressed	Introduction; Part 1; Part 2; Part 3; Part 4; Part 5; Part 6
Effective sample size	Appendix A only
Emotional state	Part 4: 5.4
Engagement	Introduction; Part 1; Part 2; Part 3; Part 4; Part 5; Part 6
Engagement attribute	Appendix A only
Engagement indicator mean	Introduction; Part 1; Part 2; Part 4; Part 5
Engagement level	Introduction: 4.5.2; Part 1: 5.4.3, 5.6.2; Part 2: 4.2.4, 5.2.1, 5.8.2; Part 4: 1, 1.5.2, 1.5.3; Part 5: 1.1.1, 4.1, 4.1.1, 4.1.3
Engagement manager	Part 1; Part 2; Part 3; Part 4; Part 5
Equivalence bound	Introduction: 4.7.2
Equivalence test	Appendix A only
Failure category	Part 1: 4.1; Part 2: 1, 1.1; Part 5: 3.4; Part 6: 1
False discovery rate	Introduction: 4.7.1; Part 1: 4.3.1
Favorable-direction conversion	Introduction: 4, 4.3, 4.3.1, 4.3.2, 4.4.1, 4.8.3; Part 1: 2.2.1, 3, 3.2, 3.2.3; Part 4: 5.1, 5.1.1, 5.1.2, 5.2, 5.2.1, 5.2.2; Part 5: 1.4.2, 1.4.3, 7, 7.3
Financial institution	Introduction; Part 1; Part 2; Part 3; Part 4; Part 5; Part 6
Financial institution disclosure	Introduction: 4.6.3
Financial institution readiness	Introduction: 3, 4.2.3; Part 1: 4; Part 2: 1, 1.1, 2, 2.3.1, 2.6.1, 5; Part 3: 1; Part 5: 1.1, 3.4; Part 6: 1, 1.1
Financial institution staff	Introduction: 3; Part 2: 2, 2.4.2; Part 6: 2
Four core study measures	Introduction; Part 1; Part 2; Part 4
Held	Introduction; Part 1; Part 2; Part 3; Part 4; Part 5; Part 6
High risk	Introduction: 4.4.2
Holm correction	Part 1: 2.1.1, 5.6.1; Part 2: 3.1.3, 3.3.1; Part 5: 4.1.1, 4.1.2, 4.1.3
Inconclusive null	Introduction: 4.7.2
Indicator	Introduction; Part 1; Part 2; Part 3; Part 4; Part 5
Indicator class	Appendix A only
Indicator quintile	Part 1: 4.2.2; Part 5: 2.3
Indicators distressed	Introduction: 4.3.2, 4.5, 4.5.5
Individual contributor	Part 1; Part 2; Part 3; Part 4; Part 5
Intent to leave	Introduction; Part 1; Part 2; Part 4; Part 5; Part 6
Internal disclosure	Introduction: 4.6.3; Part 5: 4, 4.1, 4.1.1, 4.1.2, 4.1.3
Intraclass correlation (ICC)	Appendix A only

<i>Term</i>	<i>Where it appears</i>
Leave-one-out	Introduction: 4.5.1
Manual alteration	Introduction; Part 1; Part 2; Part 3; Part 4; Part 5; Part 6
Manual alteration of AI outputs	Introduction: 3, 4.6.1; Part 3: 2.3; Part 5: 1, 1.3.1
Manual alteration of financial institution processes	Introduction: 3, 4.6.1; Part 5: 1, 1.2.1
Manual alteration of financial-institution-provided data	Introduction: 3, 4.6.1; Part 2: 2.1, 2.1.4; Part 5: 1.1.1
Measured return	Introduction: 3, 4.5, 4.5.5; Part 1: 3, 3.3, 3.3.1
Model provider	Introduction: 4.2.1; Part 2: 6.6, 6.6.2, 6.6.3, 6.8.1; Part 6: 1.4
Moderate risk	Introduction: 4.4.2
Non-answer	Appendix A only
Null equivalent to zero	Introduction: 4.7.2
Outsourcing model	Introduction: 4.1.2, 4.7.3, 4.7.5, 4.7.6; Part 1: 2, 2.1.1, 5, 5.3, 5.3.1; Part 5: 1.2.1
p value	Introduction: 4.7
Partial test	Introduction: 4.7.1
Pearson's r	Introduction: 4.7.1
Perceived fact	Introduction: 4.3.3
Portfolio director	Part 1; Part 2; Part 3; Part 4; Part 5
Primary and secondary response	Appendix A only
Process family	Introduction: 4.1.3, 4.7.6; Part 1: 2.1.1, 5, 5.3, 5.3.3, 5.6.1, 5.6.3; Part 2: 3.1.3, 3.3.1
Production	Introduction; Part 1; Part 2; Part 3; Part 4; Part 6
Question without a favorable end	Appendix A only
Rated indicator	Part 2: 3.5; Part 4: 1.4
Rating	Introduction; Part 1; Part 2; Part 4; Part 5
Reported experience	Introduction: 4, 4.2; Part 1: 1; Part 3
Reporting alteration	Introduction: 3, 4.6.1, 4.6.2; Part 5: 1.4, 1.4.3; Part 6: 1.3
Respondent	Introduction; Part 1; Part 2; Part 4; Part 5; Part 6
Respondent level	Introduction: 4.5.4; Part 1: 5.4.3; Part 4: 1.1, 1.1.1, 1.2, 1.2.1, 1.2.2, 1.2.3, 1.3.1, 1.3.2; Part 5: 2.3
Robustness check	Introduction: 4.8.2; Part 1: 2.1.1, 3.2; Part 2: 3.1.2, 3.2.2, 3.3.3, 3.4, 4.4, 5.5, 5.6.2, 6.4, 6.5, 6.6, 6.7.1; Part 4: 1.3.1, 1.3.2, 1.5.3, 2; Part 5: 1.1.2, 1.2.1, 1.3.1
Sentiment score	Introduction: 4.1.2, 4.7.6
Spearman's rho	Appendix A only
Structured answers	Introduction: 4.1.2, 4.8.1
Study norm	Appendix A only
Study set	Introduction: 4.1, 4.1.2, 4.2.1, 4.3.2, 4.5.1, 4.5.2, 4.5.5, 4.7.1, 4.7.2; Part 1: 1, 1.2, 2, 2.1.1, 3.1, 3.3, 4, 4.2, 5.3, 5.3.4; Part 2: 1, 1.1; Part 3; Part 4: 5.1; Part 5: 7
The two firms	Part 1: 5.2, 5.2.2; Part 2: 1, 1.1, 3, 3.1, 3.1.1, 3.1.2, 3.1.4, 3.2, 3.2.1, 3.2.2, 3.2.3, 3.2.4, 3.3, 3.3.3, 3.4, 3.4.1, 3.5.1, 3.5.2, 5, 5.6, 5.6.1, 5.8, 6.8.2; Part 3: 1, 2, 2.1, 2.2; Part 6: 1, 1.2, 2.4
Theme	Introduction: 4.8.3
Thin base	Introduction: 4.1.1
Third-party ecosystem	Introduction: 3; Part 1: 4; Part 2: 1, 1.1, 6, 6.1.1, 6.1.2, 6.2.1, 6.2.2, 6.3.1, 6.3.2, 6.4.1, 6.4.2, 6.5.1, 6.5.2, 6.5.3, 6.6.2, 6.6.3, 6.7.1, 6.7.2, 6.7.3, 6.7.4, 6.8.1, 6.8.2; Part 5: 1.5; Part 6: 2.3, 2.4
Third-party provider	Introduction: 4.2.1; Part 2: 6.4, 6.4.1, 6.5.3; Part 6: 1.4
Third-party work done by hand	Introduction: 3, 4.6.1; Part 2: 6; Part 5: 1.5
Transition/Cut-Over	Introduction: 4.1.3; Part 1: 1, 1.1, 1.1.4, 1.2.1; Part 2: 1.1; Part 4: 5.1; Part 5: 1.1.1, 1.3.1
Type of financial institution	Introduction: 4.1.2, 4.1.3, 4.7.3, 4.7.5; Part 1: 2, 2.1.1, 5, 5.1.1, 5.6.3
Unfavorable-direction indicator	Appendix A only
Vendor	Introduction; Part 1; Part 2; Part 3; Part 4; Part 5; Part 6
Vendor capability	Introduction: 3; Part 1: 4; Part 2: 1, 1.1, 4, 4.2.2, 4.3, 4.3.2, 4.5.3, 4.5.5, 5; Part 6: 2
Vendor type	Introduction: 4.7.3; Part 1: 5, 5.5.1, 5.5.2; Part 4: 1.5.2, 1.5.3
Welch's analysis of variance	Introduction: 4.7.3
Written answers	Introduction: 4.1.2, 4.6.1, 4.6.2, 4.7.6, 4.8.1; Part 2: 6; Part 4: 3, 3.4; Part 5: 1, 1.4, 1.5, 3.4, 4.2, 5, 5.1, 6.2; Part 6: 1.3, 1.6
Written response	Introduction: 4.1.2, 4.2, 4.7.6; Part 1: 1.1.4, 1.2.2, 3.2.4, 5.1.4, 5.3.1; Part 2: 2.5.1, 2.7.1, 3.1.1, 3.1.3, 4.1.1, 4.2.3, 4.5.3, 5.2.3, 5.4.2, 5.8.3, 6.1.1, 6.2.1, 6.5.1, 6.8.1; Part 4: 1.2.3, 2.1.1, 2.1.2, 2.1.4, 3.1.1, 3.2.1, 3.2.2, 4.2.2, 5.3, 5.3.3

7 Appendix G: Disclosure of the Use of Artificial Intelligence

The report was prepared with assistance from Claude, an Anthropic AI system, for analysis, qualitative coding, organization, and drafting. OpenAI Codex assisted with the version 17 editorial revision, including review of prose, internal consistency, and supporting analytical records. These systems contributed to the preparation of the report; their outputs are not independent evidence validating its findings.

The statistical workflow used a computational program and recorded input fingerprints and analytical settings. Companion analyses incorporated figures programmatically. The author determined substantive conventions through documented rulings, some of which revised earlier decisions. Independent reproduction of the complete computational workflow has not been established.

Claude-based coders applied the documented qualitative rulebook. Separate AI runs recoded samples without access to the first coding, and an author-ruled set of 48 difficult answers supplied an additional check. Appendix D reports the agreement results and subsequent rulebook changes. These procedures should not be described as independent human validation of every coded answer.

AI systems assisted with drafting and editorial revision under the author's standards. The author retains responsibility for reviewing and approving the final wording, interpretations, and recommendations. Inserted respondent quotations were preserved during the version 17 editorial revision.

Responsibility for the study's design, analytical decisions, interpretations, and publication rests with the author. AI assistance does not remove the need to verify factual claims, assess methodological limitations, or distinguish recommendations from findings.