

Model validation memo — Corporate PD Model v3.2

Credit risk — corporate PD

SR 26-2 model validation memo

MODEL

Corporate PD Model v3.2

ENGAGEMENT

Sample — corporate PD annual validation

PREPARED

2026-10-07

PREPARED BY

AlphaIQ CertifyIQ, computed by the product

RECOMMENDATION



May be relied upon, with conditions

12 sections · 14 traced figures · 3 finding(s) · 26 test(s) named as not performed

SYNTHETIC SAMPLE

Every figure in this report was computed by CertifyIQ from a synthetic book — 4,000 corporate obligors, replicated from the declared specification and stressed across four illustrative scenarios. No real institution, model or customer is described. The book is `corporate-pd-population.csv`; how it was generated, and the defects planted in it, are documented in `corporate-pd-dgp.json`.

AT A GLANCE · SYNTHETIC SAMPLE

Corporate PD Model v3.2

Credit risk — corporate PD · SR 26-2 model validation memo

May be relied upon, with conditions

3 material finding(s) require remediation or an accepted compensating control before unqualified reliance. Each is listed in section 9 with the evidence that produced it.

12

SECTIONS

14TRACED
FIGURES**3**

MATERIAL

0

NOTABLE

0

OBSERVATIONS

26NOT
PERFORMED

Supervisory frame: this memo is built to "SR 26-2 Model Validation", published under US · Federal Reserve / OCC / FDIC. No supervisory frame is declared for this engagement, so the frame above was chosen by the model's family rather than by its supervisor. That is not a statement that the two agree: nobody has recorded who supervises this model, and a reader in another jurisdiction should treat the section structure, the citations and the recommendation vocabulary as US's.

Evidence manifest: none sealed. The documents behind each figure are named where the figure is; the set of documents on record when this memo was written has not been attested by anyone, so it cannot be cited here.

How to read this sample. Every number below was computed by CertifyIQ from a synthetic book (corporate-pd-population.csv), through the same resolver, metric pack and report builder a signed-in engagement uses. Each figure carries its own trace — where it came from and what it was judged against. Tests that could not be performed are named with the reason, never left out. Where no threshold was declared, a metric is reported but not judged, because this report does not invent an institution's tolerances.

CONTENTS

[1 Purpose and scope](#)[2 Model description and provenance](#)[3 Independent replication](#)[4 Data and evidence reviewed](#)[5 Conceptual soundness](#)[6 Outcomes analysis](#)[7 Scenario and stress testing](#)[8 Ongoing monitoring](#)[9 Findings](#)[10 Third-party dependencies and residual risk](#)[11 Limitations and tests not performed](#)[Appendix A Reproducibility](#)

1 Purpose and scope

This memo records an independent validation of Corporate PD Model v3.2, operated by AlphaIQ (internal sample), performed by AlphaIQ (lead: Ilan) under engagement "Sample — corporate PD annual validation".

Scope covers independent replication of the declared specification, outcomes analysis against the realised population, counterfactual and boundary testing, scenario and stress simulation, and migration analysis where a transition matrix is declared. Section 10 records what was not covered and why.

This memo is a recommendation to the model owner and its second-line function. It is not a certification, and it does not constitute supervisory approval.

Reference framework: this memo is shaped to SR 26-2, Revised Guidance on Model Risk Management (17 April 2026), issued jointly by the Federal Reserve, the OCC and the FDIC, which replaced SR 11-7 (4 April 2011) and SR 21-8 (9 April 2021). SR 26-2 names three validation components — conceptual soundness, outcomes analysis and ongoing model monitoring — and each has a section here: 5, 6 and 8 respectively. The ten-section shape is AlphaIQ's convention, not the letter's; SR 26-2 states what validation must cover and does not prescribe a document layout.

What this reference does NOT claim. Rendering a section is not the same as satisfying a component: section 8 reports a reading only when the recorded runs carry one and reports a refusal otherwise, and section 10 lists what was not performed. SR 26-2 "does not set forth enforceable standards or prescriptive requirements; accordingly, non-compliance with this guidance will not result in supervisory criticism" — so neither conformance nor non-conformance with it is a finding AlphaIQ can make, and none is made here. The letter is also stated to be "most relevant to banking organizations with over \$30 billion in total assets regulated by the Federal Reserve"; whether it is the right frame for this institution is the model owner's judgement, not ours.

Scope of the guidance itself: SR 26-2 records that "[g]enerative AI and agentic AI models are novel and rapidly evolving. As such, they are not within the scope of this guidance", and that its principles apply to traditional statistical and quantitative models and to non-generative, non-agentic AI models. This memo is therefore not offered as a validation framework for generative or agentic systems, and AlphaIQ refuses to issue it under this template for a model family of that kind.

2 Model description and provenance

The model is a logit-link specification over 5 declared features, read from Corporate PD Model v3.2 — Model Development Document (pp. 28–34 (specification), p. 44 (severe-scenario claim)) as declared by Model owner, first line on 2026-08-25.

Model access: AlphaIQ's independent replica of the declared model specification; the replica reproduces the model owner's scored population within the tolerance in force (maximum deviation 0.00 bp against 0.00 bp).

Where AlphaIQ authored the machine-readable specification by reading the owner's documentation, that specification is an AlphaIQ artefact until the owner attests to it. Attestation status is recorded on the model specification record and is a precondition for any certification.

3 Independent replication

Replica reproduces the model owner's scored population to a maximum absolute PD deviation of 5e-9 bp over 4,000 obligors, within the 0.00001 bp tolerance in force.

Tolerance in force: 0.00001 bp. AlphaIQ's convention (1e-9 on the PD scale, floating-point noise); no owner tolerance is declared and none could be inferred from the evidence. State a "Replication tolerance" row in the model documentation's portfolio assumptions table — as a fraction of the PD scale or in basis points, with its source — and the gate holds the replica to the precision the documentation actually carries instead.

Maximum absolute deviation, replica vs owner - scored population

MEASURED

5e-9 bp

Maximum absolute deviation, replica vs owner - scored population: 5e-9 bp measured on 4,000 obligors scored by both the owner and the replica. Benchmark source: Tolerance in force: 0.00001 bp. AlphaIQ's convention (1e-9 on the PD scale, floating-point noise); no owner tolerance is declared and none could be inferred from the evidence. State a "Replication tolerance" row in the model documentation's portfolio assumptions table — as a fraction of the PD scale or in basis points, with its source — and the gate holds the replica to the precision the documentation actually carries instead.

Rank correlation, replica vs owner - scored population

MEASURED

1.000

Rank correlation, replica vs owner - scored population: 1.000 measured on 4,000 obligors scored by both the owner and the replica. Benchmark source: Spearman rank correlation; ranking agreement matters even where levels drift.

4 Data and evidence reviewed

Evidence: "Corporate PD sample book — synthetic", AlphaIQ synthetic corporate population, generated for the sample engagement. Not a client extract, observed as at 2026-06-30. Content hash bad08b4e32245aa34da5f61343d9e4404355c89626b755d89943ee3d901f5a63. Every figure in this memo derived from realised data is computed on this file and no other; re-running the validation against a file with a different hash produces a different memo.

The metrics in section 6 rest on 4,000 obligors carrying 66 realised defaults (1.65%): 0 of 4,000 rows read carried no readable score or outcome and were excluded rather than imputed.

The replication gate in section 3 and the simulation layers in section 7 rest on the same 4,000 obligors carrying 66 realised defaults (1.65%): every row read also passed the specification's declared input-validity bounds and missing-value policy.

With 4,000 obligors and 66 defaults (1.65%), the 95% bootstrap interval on AUC is [0.706, 0.813] around a point estimate of 0.764. A challenger model whose AUC differs by less than 0.107 could not be distinguished from this one on this population. This quantifies the power of the available evidence; it is not evidence about calibration, and comparing two models on the same obligors requires a paired test that this layer does not perform. no obligor identifier was supplied to this layer, so every draw below is addressed by the obligor's own recorded values instead. That is still content addressing and these figures are still invariant to the order the rows arrived in; what it cannot do is tell apart two obligors identical in every field this layer reads. No two of these 4,000 obligors are identical in that sense, so nothing here rests on it.

Obligors in the scored population

MEASURED

4,000

Obligors in the scored population: 4,000 measured on 4,000 obligors carrying 66 realised defaults. Benchmark source: Corporate PD sample book — synthetic — AlphaIQ synthetic corporate population, generated for the sample engagement. Not a client extract.

Realised defaults

MEASURED

66

Realised defaults: 66 measured on 4,000 obligors carrying 66 realised defaults. Benchmark source: Corporate PD sample book — synthetic — AlphaIQ synthetic corporate population, generated for the sample engagement. Not a client extract.

Smallest AUC difference this population can resolve

MEASURED

0.107

Smallest AUC difference this population can resolve: 0.107 measured on 4,000 obligors carrying 66 realised defaults. Benchmark source: Layer 2 percentile bootstrap on the replica's scores over Corporate PD sample book — synthetic.

MATERIAL F-POWER The population cannot decide discriminatory power against the 0.712 floor

The interval [0.706, 0.813] contains the declared floor of 0.712. The point estimate of 0.764 clears it, but this population cannot distinguish a model that meets the floor from one that does not. The pass/fail is not supported by the evidence, whichever way it points. Threshold source: Declared by Ilan on 2026-08-30, from Model development document, p. 31.

Evidence to re-run: Layer 2 probe, 400 bootstrap resamples on 4,000 obligors with 66 defaults.

5 Conceptual soundness

Sweeping each declared feature across its domain found 1 exceedance region(s) above the stated detection limits, over 1 grid location(s) in total. Each region is ONE finding below: a cliff spans many adjacent grid points and reporting one finding per point would raise the same defect a hundred times and re-raise it as new whenever the grid or the sample moved it.

Each feature was swept across its declared support at a fixed step — `ebitdaMargin` in steps of 0.0015; `interestCoverage` in steps of 0.0550; `leverage` in steps of 0.0323; `revenueGrowth` in steps of 0.0025; `workingCapitalRatio` in steps of 0.0025 — and that step is the detection limit every finding below is measured against.

MATERIAL F-CF1 grade-flip — `ebitdaMargin`, `interestCoverage`, `leverage`, `revenueGrowth`, `workingCapitalRatio`

grade-flip on `ebitdaMargin`, `interestCoverage`, `leverage`, `revenueGrowth`, `workingCapitalRatio`, at one grid location, at . The measured step is 0. At most 9.90% of the population sits within one grid step of any single location in this region; that is the largest share at one location and deliberately not a sum, because adjacent locations overlap. THIS IS ONE REGION REPORTED ONCE, not 1 defect(s): the location of a cliff moves with the grid and with the sample, so a register that raised a finding per grid step would re-raise this defect as new on every run and track nothing. The sweep's own sentence for the worst location: 99 of 1000 obligors (9.9%) have at least a 1-in-4 chance of landing in a different rating grade under the measurement uncertainty declared in the model documentation. 6 are closer to a coin flip than to a rating. Grade assignment carries a capital consequence, so this is a provisioning question, not a modelling nicety.

Evidence to re-run: Layer 3 sweep over 240 grid point(s) per feature on 1,000 obligor(s), root seed 20260825; the step size behind the detection limit is stated in the paragraph above, once, rather than repeated on every finding.

MATERIAL F-FLIP Rating grades are unstable under declared measurement uncertainty

9.9% of obligors change grade with probability above 25% and 0.6% above 50%, using the measurement uncertainty declared per feature.

Evidence to re-run: Layer 3 grade-flip simulation, 200 draws over 1,000 obligors.

NOT PERFORMED — AND WHY

Gradient exceedance — leverage

No policy sensitivity limit has been declared for this feature (`spec.features[].sensitivityLimit`). Inventing one would make the finding an artefact of AlphaIQ's assumption rather than the bank's policy, so the gradient is measured and reported but judged against nothing.

Gradient exceedance — interestCoverage

No policy sensitivity limit has been declared for this feature (`spec.features[].sensitivityLimit`). Inventing one would make the finding an artefact of AlphaIQ's assumption rather than the bank's policy, so the gradient is measured and reported but judged against nothing.

Gradient exceedance — ebitdaMargin

No policy sensitivity limit has been declared for this feature (`spec.features[].sensitivityLimit`). Inventing one would make the finding an artefact of AlphaIQ's assumption rather than the bank's policy, so the gradient is measured and reported but judged against nothing.

Gradient exceedance — workingCapitalRatio

No policy sensitivity limit has been declared for this feature (`spec.features[].sensitivityLimit`). Inventing one would make the finding an artefact of AlphaIQ's assumption rather than the bank's policy, so the gradient is measured and reported but judged against nothing.

Gradient exceedance — revenueGrowth

No policy sensitivity limit has been declared for this feature (`spec.features[].sensitivityLimit`). Inventing one would make the finding an artefact of AlphaIQ's assumption rather than the bank's policy, so the gradient is measured and reported but judged against nothing.

Feature -interaction testing

No feature pairs have been declared for interaction testing (`spec.interactionPairs`). Unlike a policy limit this is a scope decision rather than an assumption — testing a pair cannot manufacture a finding — but which pairs are worth testing follows from the model's use and its development documentation, and AlphaIQ has not been told. No pair was tested, and this entry exists so that is visible rather than inferred from an absent section.

6 Outcomes analysis

Outcomes analysis was computed directly from the realised population using credit. Thresholds are taken from AlphaIQ credit pack — each threshold cited to its own instrument and are not adjusted to the result.

Out-of-sample AUC

MEASURED 0.764

Out-of-sample AUC: 0.764 measured on 4,000 obligors carrying 66 realised defaults, bootstrap interval [0.706, 0.813].

Benchmark source: The owner's documented champion-model AUC, declared as a claim. Judged against the owner's declared 0.712 (Model development document, p. 31).

KS statistic

MEASURED 0.37

KS statistic: 0.37 measured on 4,000 obligors carrying 66 realised defaults. Benchmark source: Judged against the KS floor the institution declares in its own model risk policy. No supervisory instrument publishes a KS floor, and this platform ships a threshold only when a named published instrument sets it; until the floor is declared the statistic is reported unjudged.

Brier score

MEASURED 0.016

Brier score: 0.016 measured on 4,000 obligors carrying 66 realised defaults. Benchmark source: Judged against the Brier ceiling the institution declares. A Brier score is bounded below by the book's own base rate — $p(1-p)$ for a perfectly calibrated model — so any ceiling is a statement about this portfolio and its owner's appetite, which no published instrument sets; until it is declared the score is reported unjudged.

Calibration (observed / predicted)**MEASURED 1.33**

Calibration (observed / predicted): 1.33 measured on 4,000 obligors carrying 66 realised defaults. Benchmark source: The institution's declared calibration tolerance.

Calibration bias (mean predicted PD – observed default rate)**MEASURED -40 bp**

Calibration bias (mean predicted PD – observed default rate): -40 bp measured on 4,000 obligors carrying 66 realised defaults. Benchmark source: Judged against the over-prediction floor the institution declares for this portfolio, in basis points. No published instrument sets one; until it is declared the bias is reported unjudged. This is a CALIBRATION BIAS and NOT the EBA/GL/2017/16 margin of conservatism, which is a documentary decomposition into categories A (data and method deficiencies), B (relevant changes to standards or processes) and C (general estimation error), applied per risk parameter. That decomposition is declared not performed by this pack.

NOT PERFORMED — AND WHY**PSI (population stability)**

PSI compares a reference window with a recent one, and this evidence declares no observation-date column to window on. The old row-position split stood in for a time order nothing guaranteed — the same book re-sorted moved the figure 49% — so it is refused rather than guessed. Map a vintage or observation-date column.

Override analysis — the institution's rating override register

No override log is stored for this engagement, so nothing is known here about how often, in which direction, by whom or with what stated reason the model's ratings were overridden before use. A model that is routinely overridden is two models, and this memo has validated only the one whose ratings it was given. Import the register at /override-log-import and this section reports it.

7 Scenario and stress testing

Scenario results were produced by AlphaIQ's independent replica of the declared model specification; the replica reproduces the model owner's scored population within the tolerance in force (maximum deviation 0.00 bp against 0.00 bp). Intervals are Monte Carlo error only: they capture simulation noise and not parameter, specification or scenario uncertainty.

Scenario provenance, run by run: "Baseline" — Drawn by AlphaIQ to the shape of published supervisory scenarios. Illustrative only — it is not a supervisor's scenario and cannot support a certified figure. (Internal planning baseline — flat deviations by construction.) "Adverse — moderate recession" — Drawn by AlphaIQ to the shape of published supervisory scenarios. Illustrative only — it is not a supervisor's scenario and cannot support a certified figure. (Illustrative trajectory calibrated to the shape of published supervisory severely-adverse scenarios. NOT a transcription — replace with the client's supervisory scenario file before regulated use.) "Severely adverse — deep recession with credit crunch" — Drawn by AlphaIQ to the shape of published supervisory scenarios. Illustrative only — it is not a supervisor's scenario and cannot support a certified figure. (Illustrative trajectory calibrated to the shape of published supervisory severely-adverse scenarios. NOT a transcription — replace with the client's supervisory scenario file before regulated use.) "Stagflation — weak growth with persistent funding costs" — Drawn by AlphaIQ to the shape of published supervisory scenarios. Illustrative only — it is not a supervisor's scenario and cannot support a certified figure. (Illustrative trajectory calibrated to the shape of published supervisory severely-adverse scenarios. NOT a transcription — replace with the client's supervisory scenario file before regulated use.)

Portfolio PD — Baseline**SIMULATED 1.286%**

Portfolio PD — Baseline: 1.286% under Baseline (Internal planning baseline — flat deviations by construction.), transmitted via 5 declared macro sensitivities mapping gdpGrowth, hySpread, unemployment onto ebitdaMargin, interestCoverage, leverage, revenueGrowth (Corporate PD Model v3.2 — Model Development Document, pp. 28–34 (specification), p. 44 (severe-scenario claim)), Monte Carlo interval [1.286%, 1.286%], provenance replica, run 42ae7ced634c.

Portfolio PD — Adverse — moderate recession**SIMULATED 1.783%**

Portfolio PD — Adverse — moderate recession: 1.783% under Adverse — moderate recession (Illustrative trajectory calibrated to the shape of published supervisory severely-adverse scenarios. NOT a transcription — replace with the client's supervisory scenario file before regulated use.), transmitted via 5 declared macro sensitivities mapping gdpGrowth, hySpread, unemployment onto ebitdaMargin, interestCoverage, leverage, revenueGrowth (Corporate PD Model v3.2 — Model Development Document, pp. 28–34 (specification), p. 44 (severe-scenario claim)), Monte Carlo interval [1.773%, 1.792%], provenance replica, run 42ae7ced634c.

Portfolio PD — Severely adverse — deep recession with credit crunch**SIMULATED 3.102%**

Portfolio PD — Severely adverse — deep recession with credit crunch: 3.102% under Severely adverse — deep recession with credit crunch (Illustrative trajectory calibrated to the shape of published supervisory severely-adverse scenarios. NOT a transcription — replace with the client's supervisory scenario file before regulated use.), transmitted via 5 declared macro sensitivities mapping gdpGrowth, hySpread, unemployment onto ebitdaMargin, interestCoverage, leverage, revenueGrowth (Corporate PD Model v3.2 — Model Development Document, pp. 28–34 (specification), p. 44 (severe-scenario claim)), Monte Carlo interval [3.058%, 3.146%], provenance replica, run 42ae7ced634c.

Portfolio PD — Stagflation — weak growth with persistent funding costs**SIMULATED 1.919%**

Portfolio PD — Stagflation — weak growth with persistent funding costs: 1.919% under Stagflation — weak growth with persistent funding costs (Illustrative trajectory calibrated to the shape of published supervisory severely-adverse scenarios. NOT a transcription — replace with the client's supervisory scenario file before regulated use.), transmitted via 5 declared macro sensitivities mapping gdpGrowth, hySpread, unemployment onto ebitdaMargin, interestCoverage, leverage, revenueGrowth (Corporate PD Model v3.2 — Model Development Document, pp. 28–34 (specification), p. 44 (severe-scenario claim)), Monte Carlo interval [1.904%, 1.933%], provenance replica, run 42ae7ced634c.

NOT PERFORMED — AND WHY**Comparison of the documented severe-scenario claim against the replicated projection**

Not made. The severe-scenario projection in this section was computed on "Severely adverse — deep recession with credit crunch", whose provenance is declared illustrative — a trajectory drawn by AlphaIQ to the shape of published supervisory scenarios, not a supervisor's scenario. The owner's claim of 2.10% (Model development document, p. 44) was made about their own scenario of record. Differencing the two would report a gap between answers to different questions, and would raise it as a finding against the institution on the strength of a trajectory this platform drew. Import the supervisor's or the owner's own scenario file and the comparison is made automatically.

Stress projection on a supervisor's or the owner's own scenario

Not performed. 3 of the 3 stressed scenario(s) relied on here ("Adverse — moderate recession", "Severely adverse — deep recession with credit crunch", "Stagflation — weak growth with persistent funding costs") are declared illustrative — drawn by AlphaIQ to the shape of published supervisory scenarios, and not a supervisor's scenario. The projections computed on them describe how the model responds to a shape this platform chose. They are not evidence about the model's behaviour under any scenario the institution or its supervisor has adopted, and a certificate will refuse them (VAL-68). Importing the supervisory scenario file replaces them; it is a data task and not a code change.

Systematic parameter shock ranking (sensitivity)

No shock size is declared for this engagement, so no parameter sensitivity was computed. This platform does not choose one: plus or minus ten per cent is a convention of practice and no authority publishes a sensitivity shock size, so a ranking computed against a size AlphaQ picked would be AlphaQ's methodological judgement printed under the institution's name. Declaring the size, cited to the policy that sets it, is what produces the ranking.

8 Ongoing monitoring

Ongoing monitoring is a named validation component in both SR 11-7 and SR 26-2, and it asks a question no single run can answer: whether the model still behaves as it did, as conditions change. It is not performed for this engagement, for the reason below.

Only one distinct scored population is on record. Ongoing monitoring compares performance across cycles, and one cycle is not a series — reporting it as stable would be a claim, not a reading.

NOT PERFORMED — AND WHY

Ongoing monitoring

Only one distinct scored population is on record. Ongoing monitoring compares performance across cycles, and one cycle is not a series — reporting it as stable would be a claim, not a reading. 1 distinct scored population(s) and 0 rerun(s) are on record. Recording a metric run each cycle, on that cycle's population, is what makes this section computable.

9 Findings

3 finding(s) were raised, 3 of them material. Each carries the evidence a reviewer would re-run to check it.

Against the findings register: 3 finding(s) open, 3 at grade 1, 3 carrying a grade the engine PROPOSED and nobody has owned, 3 newly raised. Oldest open finding is 0 day(s) old.

Remediation is recorded as verified only when a rerun no longer raises the finding; it is never closed on assertion. An acceptance carries the value it was granted at and a tolerance, and lapses automatically when the evidence moves beyond those terms.

3 live finding(s) have no owner assigned. An unowned finding has no one accountable for closing it and will age without anyone noticing.

MATERIAL F-CF1 grade-flip — ebitdaMargin, interestCoverage, leverage, revenueGrowth, workingCapitalRatio

grade-flip on ebitdaMargin, interestCoverage, leverage, revenueGrowth, workingCapitalRatio, at one grid location, at . The measured step is 0. At most 9.90% of the population sits within one grid step of any single location in this region; that is the largest share at one location and deliberately not a sum, because adjacent locations overlap. THIS IS ONE REGION REPORTED ONCE, not 1 defect(s): the location of a cliff moves with the grid and with the sample, so a register that raised a finding per grid step would re-raise this defect as new on every run and track nothing. The sweep's own sentence for the worst location: 99 of 1000 obligors (9.9%) have at least a 1-in-4 chance of landing in a different rating grade under the measurement uncertainty declared in the model documentation. 6 are closer to a coin flip than to a rating. Grade assignment carries a capital consequence, so this is a provisioning question, not a modelling nicety.

Evidence to re-run: Layer 3 sweep over 240 grid point(s) per feature on 1,000 obligor(s), root seed 20260825; the step size behind the detection limit is stated in the paragraph above, once, rather than repeated on every finding.

MATERIAL F-FLIP Rating grades are unstable under declared measurement uncertainty

9.9% of obligors change grade with probability above 25% and 0.6% above 50%, using the measurement uncertainty declared per feature.

Evidence to re-run: Layer 3 grade-flip simulation, 200 draws over 1,000 obligors.

MATERIAL F-POWER The population cannot decide discriminatory power against the 0.712 floor

The interval [0.706, 0.813] contains the declared floor of 0.712. The point estimate of 0.764 clears it, but this population cannot distinguish a model that meets the floor from one that does not. The pass/fail is not supported by the evidence, whichever way it points. Threshold source: Declared by Ilan on 2026-08-30, from Model development document, p. 31.

Evidence to re-run: Layer 2 probe, 400 bootstrap resamples on 4,000 obligors with 66 defaults.

10 Third-party dependencies and residual risk

No third-party dependency is declared for this engagement. That is a statement by the institution that the model embeds no vendor or open-source component it cannot validate inside — not an absence of the question.

11 Limitations and tests not performed

The following checks were not performed, each with the reason. A validation report's most dangerous property is the test it quietly omitted, so this register is part of the memo rather than an appendix to it.

Monte Carlo and bootstrap intervals throughout this memo capture simulation and sampling error only. They do not capture parameter uncertainty, specification uncertainty or scenario uncertainty, and must not be read as confidence intervals on the underlying quantity.

NOT PERFORMED — AND WHY

Gradient exceedance — leverage

No policy sensitivity limit has been declared for this feature (`spec.features[].sensitivityLimit`). Inventing one would make the finding an artefact of AlphaIQ's assumption rather than the bank's policy, so the gradient is measured and reported but judged against nothing.

Gradient exceedance — interestCoverage

No policy sensitivity limit has been declared for this feature (`spec.features[].sensitivityLimit`). Inventing one would make the finding an artefact of AlphaIQ's assumption rather than the bank's policy, so the gradient is measured and reported but judged against nothing.

Gradient exceedance — ebitdaMargin

No policy sensitivity limit has been declared for this feature (`spec.features[].sensitivityLimit`). Inventing one would make the finding an artefact of AlphaIQ's assumption rather than the bank's policy, so the gradient is measured and reported but judged against nothing.

Gradient exceedance — workingCapitalRatio

No policy sensitivity limit has been declared for this feature (`spec.features[].sensitivityLimit`). Inventing one would make the finding an artefact of AlphaIQ's assumption rather than the bank's policy, so the gradient is measured and reported but judged against nothing.

Gradient exceedance — revenueGrowth

No policy sensitivity limit has been declared for this feature (`spec.features[].sensitivityLimit`). Inventing one would make the finding an artefact of AlphaIQ's assumption rather than the bank's policy, so the gradient is measured and reported but judged against nothing.

Feature-interaction testing

No feature pairs have been declared for interaction testing (`spec.interactionPairs`). Unlike a policy limit this is a scope decision rather than an assumption — testing a pair cannot manufacture a finding — but which pairs are worth testing follows from the model's use and its development documentation, and AlphaIQ has not been told. No pair was tested, and this entry exists so that is visible rather than inferred from an absent section.

PSI (population stability)

PSI compares a reference window with a recent one, and this evidence declares no observation-date column to window on. The old row-position split stood in for a time order nothing guaranteed — the same book re-sorted moved the figure 49% — so it is refused rather than guessed. Map a vintage or observation-date column.

Override analysis — the institution's rating override register

No override log is stored for this engagement, so nothing is known here about how often, in which direction, by whom or with what stated reason the model's ratings were overridden before use. A model that is routinely overridden is two models, and this memo has validated only the one whose ratings it was given. Import the register at `/override-log-import` and this section reports it.

Comparison of the documented severe-scenario claim against the replicated projection

Not made. The severe-scenario projection in this section was computed on "Severely adverse — deep recession with credit crunch", whose provenance is declared illustrative — a trajectory drawn by AlphaIQ to the shape of published supervisory scenarios, not a supervisor's scenario. The owner's claim of 2.10% (Model development document, p. 44) was made about their own scenario of record. Differencing the two would report a gap between answers to different questions, and would raise it as a finding against the institution on the strength of a trajectory this platform drew. Import the supervisor's or the owner's own scenario file and the comparison is made automatically.

Stress projection on a supervisor's or the owner's own scenario

Not performed. 3 of the 3 stressed scenario(s) relied on here ("Adverse — moderate recession", "Severely adverse — deep recession with credit crunch", "Stagflation — weak growth with persistent funding costs") are declared illustrative — drawn by AlphaIQ to the shape of published supervisory scenarios, and not a supervisor's scenario. The projections computed on them describe how the model responds to a shape this platform chose. They are not evidence about the model's behaviour under any scenario the institution or its supervisor has adopted, and a certificate will refuse them (VAL-68). Importing the supervisory scenario file replaces them; it is a data task and not a code change.

Systematic parameter shock ranking (sensitivity)

No shock size is declared for this engagement, so no parameter sensitivity was computed. This platform does not choose one: plus or minus ten per cent is a convention of practice and no authority publishes a sensitivity shock size, so a ranking computed against a size AlphaIQ picked would be AlphaIQ's methodological judgement printed under the institution's name. Declaring the size, cited to the policy that sets it, is what produces the ranking.

Ongoing monitoring

Only one distinct scored population is on record. Ongoing monitoring compares performance across cycles, and one cycle is not a series — reporting it as stable would be a claim, not a reading. 1 distinct scored population(s) and 0 rerun(s) are on record. Recording a metric run each cycle, on that cycle's population, is what makes this section computable.

Row-order invariance — established for all four layers, the replication gate and the run hash

this platform addressed Layer 2's bootstrap resample and Layer 4's idiosyncratic asset shock by the obligor rather than by the row it sits on. did the same for Layer 1, whose interactive sample took every k-th ROW and now takes every k-th obligor of a canonical order derived from the obligors themselves, and for Layer 3, which additionally addressed each obligor's measurement-error draws by that obligor's position in that sample. this platform closed the last one, which was never in a layer: the replication gate reported the ten obligors the replica disagrees most about by RAW ROW INDEX, and accumulated its error sums in arrival order, inside the same content hash. It now names them by ADDRESS and sums in canonical obligor order. EVERY FIGURE ALL FOUR LAYERS PUBLISH, AND EVERY FIELD OF THIS RUN'S CONTENT HASH, IS UNCHANGED TO THE LAST BIT IF THIS SAME PORTFOLIO IS EXPORTED AGAIN IN A DIFFERENT ROW ORDER, and a test computes all of them over a shuffled book and asserts exactly that BEFORE the hash is allowed to depend on it. Two differently sorted exports of one book are now reported as ONE run. WHAT THIS IS NOT: it is not a claim that a re-sort can hide a change. The addresses are content addresses over every field the evidence carries, so an obligor that changes changes its own address; a shuffled book with one obligor moved still diverges, and that is asserted in both directions rather than argued. WHAT IT STILL COSTS: this run addressed its obligors by obligor-content. No obligor identifier column was supplied, so each obligor is addressed by its own recorded values — still content addressing, still order-invariant, but unable to tell apart two obligors identical in every field a layer reads. No two of these 4,000 obligors are identical in that sense, so nothing here rests on it. The manifest no longer records the byte order of the uploaded file, which is deliberate — the order of an export is not a fact about a book — and the "obligor addressing" line carries the order-free identity of the obligors the run was computed on.

Layer 4 — transition matrix & migration Monte Carlo

The specification declares no rating transition matrix. Migration and capital figures are not derived from an invented matrix.

Appendix A Reproducibility

Simulation engine alphaiq-sim/1.6.0, PRNG prng/1.0.0 (self-check 54ed466ec60f), root seed 20260825, specification hash 1fa221c3be64, runner 1fa221c3be64, run content hash 42ae7ced634cad6595f06a96c77241d15b48344ba8008b6e6cb6ba7fd8becaf9.

Metric pack credit (engine credit-metrics/1.0.0), thresholds from AlphaIQ credit pack — each threshold cited to its own instrument.

Layers executed: layer1.macro-stress@sim.macro-stress/1.2.0, layer3.counterfactual@sim.counterfactual/1.3.0, layer2.ldp@sim.ldp/1.1.0. Scenarios: Baseline (Internal planning baseline — flat deviations by construction.); Adverse — moderate recession (Illustrative trajectory calibrated to the shape of published supervisory severely-adverse scenarios. NOT a transcription — replace with the client's supervisory scenario file before regulated use.); Severely adverse — deep recession with credit crunch (Illustrative trajectory calibrated to the shape of published supervisory severely-adverse scenarios. NOT a transcription — replace with the client's supervisory scenario file before regulated use.); Stagflation — weak growth with persistent funding costs (Illustrative trajectory calibrated to the shape of published supervisory severely-adverse scenarios. NOT a transcription — replace with the client's supervisory scenario file before regulated use.).

Re-running the same specification, population, scenarios and root seed reproduces this memo's figures byte for byte. A rerun whose PRNG self-check differs is reported as non-reproducible rather than as agreement.



AlphaIQ CertifyIQ — independent validation of models and the decisions they drive. This sample was produced by the product itself; nothing in it was written by hand.

Memo content hash 3606d16b231438c9f3712c1c14e102ddf4048ed41009d3dda7639168404d7db2

This document is a recommendation, not a certification. A certificate is issued only on attested evidence, which a synthetic sample does not have.