

Model validation memo — Behavioural NMD & Prepayment Suite v4

IRRBB — behavioural deposit and prepayment assumptions

Interest rate risk in the banking book

MODEL

Behavioural NMD & Prepayment Suite v4

ENGAGEMENT

Sample — behavioural assumption review

PREPARED

2026-10-07

PREPARED BY

AlphaIQ CertifyIQ, computed by the product

RECOMMENDATION



May be relied upon, with conditions

12 sections · 5 traced figures · 3 finding(s) · 4 test(s) named as not performed

SYNTHETIC SAMPLE

Every figure in this report was computed by CertifyIQ from a synthetic book — Deposit and mortgage cohorts, modelled against realised balances over 36 months. No real institution, model or customer is described. The book is `irrbb-cohort-book.csv`; how it was generated, and the defects planted in it, are documented in `irrbb-cohort-book-dgp.json`.

AT A GLANCE · SYNTHETIC SAMPLE

Behavioural NMD & Prepayment Suite v4

IRRBB — behavioural deposit and prepayment assumptions · Interest rate risk in the banking book

May be relied upon, with conditions

1 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

5TRACED
FIGURES**1**

MATERIAL

2

NOTABLE

0

OBSERVATIONS

4NOT
PERFORMED

Supervisory frame: this memo is built to "IRRBB Review Pack", published under EU · EBA / CRD. 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 EU'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 (irrbb-cohort-book.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 reports an independent review of Behavioural NMD & Prepayment Suite v4 against the irrbb metric pack, over 216 observed cohort-month(s) across 6 cohort(s) and 3 product class(es), over 36 month(s) from 2023-07-01 to 2026-06-01.

WHAT IS UNDER TEST IS THE INSTITUTION'S OWN ASSUMPTIONS, and there is no alternative to that here. This pack ships no threshold: the framework governing interest rate risk in the banking book places obligations on how behavioural assumptions are set, reviewed and evidenced, and publishes no error tolerance for any of them.

So every figure below is measured against a number AlphaIQ sample engagement declared in Sample ALCO behavioural assumptions paper on 2026-03-18 — a core share of 75.0%, a pass-through of 0.45, and a prepayment rate of 4.0% a year — and the memo names whose claim each bar is.

The estimation window the institution declares is 36 months. The core-share figure is measured at exactly that horizon and at no other: a survival rate means nothing apart from the horizon it was measured at, and this book was refused at import if it could not reach it.

2 Model description and provenance

Three behavioural models feed the banking book under review: the non-maturing deposit core and non-core split, the pass-through of a reference-rate move to the rate actually paid, and prepayment or early redemption. Each is estimated from history and each is capable of being wrong in a way a number shows, which is what makes them models rather than conventions.

Evidence: "Deposit and mortgage cohort book — synthetic", AlphaIQ synthetic cohort book (irrbb-cohort-book-dgp.json). Content hash

5c359c5dfbbccf9a9fc77e3a9a4f06749734d5d65a774990d3c7c68e23aa5ccf. 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.

Product classes present: retail-current, retail-savings, sme-current. Cohorts: 6.

3 Independent replication

NO REPLICA WAS BUILT AND NONE WAS EXECUTED, and the reason is a scope decision rather than an impossibility — which is a different answer from the one this section gives for a machine-learning family, and the difference matters to what a reader should conclude.

A behavioural model IS declarable. A pass-through has coefficients; a core and non-core split has a rule; a prepayment model has a functional form. So unlike a boosted ensemble, there is something here a specification could state and a replica could reproduce. What this pack does not rebuild is the ALM ENGINE those assumptions feed — the machinery that turns them into an economic-value-of-equity or net-interest-income figure under prescribed shocks.

That boundary is deliberate and it is where the value is. What is unvalidated at almost every institution is the behavioural INPUT, not the arithmetic stacked on top of it: the engine is specified by the regulation and audited as a calculation, while the assumptions feeding it are estimated in-house, revised infrequently, and rarely tested against what actually happened. This review tests them.

A reader should therefore NOT conclude that the institution's reported rate-risk figures have been independently reproduced. They have not been, and section 11 names that omission with the limits it applies to.

4 Data and evidence reviewed

216 observed cohort-month(s) across 6 cohort(s) and 3 product class(es), over 36 month(s) from 2023-07-01 to 2026-06-01. 216 row(s) were read and 0 dropped.

A row is kept when it carries a cohort, a product class, a month this parser can read, a positive opening balance, a realised closing balance, a modelled closing balance, both rates and a prepaid amount. A row missing any of those is DROPPED WITH ITS REASON rather than defaulted, because a modelled balance compared against a defaulted realised one is a comparison of this platform's arithmetic with itself.

5 Conceptual soundness

For a behavioural model, conceptual soundness is whether the shape the institution assumed is the shape the book exhibits. This section reports the core and non-core split, which is the assumption the largest number of downstream figures rest on.

Measured over 6 cohort(s) observed for the full 36-month window. Declared 75.0%, realised 72.9%, a drift of 2.10 percentage points.

Core share realised at the declared horizon

MEASURED 72.90%

Core share realised at the declared horizon: 72.90% measured on 216 observed cohort-months carrying 216 realised closing balances. Benchmark source: Declared by AlphaIQ sample engagement in Sample ALCO behavioural assumptions paper, 2026-03-18.

Core share declared

MEASURED 75.00%

Core share declared: 75.00% measured on 216 observed cohort-months carrying 216 realised closing balances. Benchmark source: Declared by AlphaIQ sample engagement in Sample ALCO behavioural assumptions paper, 2026-03-18.

NOTABLE `irrbb-core-share` The realised core share differs from the declared one by more than the declared tolerance

The institution declared a core share of 75.0% and the book delivers 72.9% at the 36-month horizon. A core share set too high understates how much of the balance reprices, and therefore understates rate sensitivity in every figure computed from it.

Evidence to re-run: Balance-weighted over 6 cohort(s), content hash 5c359c5dfbbc.

6 Outcomes analysis

Outcomes analysis for this family asks one question of each model: did what it projected match what happened.

Runoff error is balance-weighted rather than cohort-weighted, deliberately: a model accurate on many small cohorts and wrong on the one holding most of the book has not been accurate, and an unweighted mean would report the opposite. Measured over 216 cohort-month(s).

The reference rate moved 172 bp across this book and the balance-weighted paid rate moved 57.2 bp, a realised pass-through of 0.333 against a declared 0.45.

Runoff error (balance-weighted)

MEASURED 1.47%

Runoff error (balance-weighted): 1.47% measured on 216 observed cohort-months carrying 216 realised closing balances. Benchmark source: Declared by AlphaIQ sample engagement in Sample ALCO behavioural assumptions paper, 2026-03-18.

Pass-through, realised

MEASURED 0.333

Pass-through, realised: 0.333 measured on 216 observed cohort-months carrying 216 realised closing balances. Benchmark source: Declared by AlphaIQ sample engagement in Sample ALCO behavioural assumptions paper, 2026-03-18.

MATERIAL irrbb-pass-through **The realised pass-through is materially below the declared assumption**

The institution assumed it would pass through 0.45 of a reference-rate move and passed through 0.333. An assumed pass-through that is too HIGH makes deposits look more rate-sensitive than they are, which understates the economic value at risk from a rate move. This is the most common finding of its kind and it is a finding about the ASSUMPTION rather than about the institution's pricing decision, which may well have been the right one.

Evidence to re-run: Rate path from 2023-07-01 to 2026-06-01, content hash 5c359c5dfbbc.

7 Scenario and stress testing

NO SCENARIO WAS PROJECTED AND NO SHOCK WAS APPLIED BY THIS REVIEW, and this is the section where that matters most, so it is stated rather than left to the appendix.

The supervisory outlier tests — a decline in economic value of equity of more than 15% of Tier 1 capital, and a one-year decline in net interest income of more than 5% of Tier 1 — are the only two numeric pass/fail bars published anywhere in this domain. Both are properties of the institution's own ALM engine under six prescribed shock scenarios, and this pack computes neither.

The institution's own submitted results belong in this review as a declared claim reconciled against those limits. They are deliberately NOT reported as a figure here: a number that came out of a submission must not sit in a figures table beside figures this pack computed, under one heading, with one kind of authority, where a reader has no way to tell them apart.

NOT PERFORMED — AND WHY

Economic value of equity and net interest income under the six prescribed shock scenarios

Computed by the institution's ALM engine, which this platform does not rebuild. What this review tests is the behavioural input to that engine, which is the part that is estimated and the part that is rarely tested.

8 Ongoing monitoring

Ongoing monitoring asks whether the model still works, and for which parts of the book.

Worst: sme-current at 8.64%. Best: retail-current at 0.30%. Across 3 product class(es).

A per-class figure is reported because an aggregate error inside tolerance while one class misses badly is exactly the case a single pooled number hides — and the two rows disagreeing is information rather than noise.

Runoff-error spread across product classes

MEASURED 8.34%

Runoff-error spread across product classes: 8.34% measured on 216 observed cohort-months carrying 216 realised closing balances. Benchmark source: Declared by AlphaIQ sample engagement in Sample ALCO behavioural assumptions paper, 2026-03-18.

NOTABLE irrbb-segment-spread **The model tracks retail-current and does not track sme-current**

Runoff error differs by 8.34 percentage points between the best and worst product class. This is a finding about where the model applies rather than about whether it works: a behavioural model calibrated on retail behaviour and applied to a class that behaves differently will be accurate in aggregate for as long as the aggregate is dominated by the class it was fitted to.

Evidence to re-run: Per-class runoff over 216 cohort-month(s), content hash 5c359c5dfbbc.

9 Findings

3 finding(s) were raised, 1 of them material. Each carries the evidence a reviewer would re-run to check it.

Against the findings register: 3 finding(s) open, 1 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.

MATERIAL irrbb-pass-through **The realised pass-through is materially below the declared assumption**

The institution assumed it would pass through 0.45 of a reference-rate move and passed through 0.333. An assumed pass-through that is too HIGH makes deposits look more rate-sensitive than they are, which understates the economic value at risk from a rate move. This is the most common finding of its kind and it is a finding about the ASSUMPTION rather than about the institution's pricing decision, which may well have been the right one.

Evidence to re-run: Rate path from 2023-07-01 to 2026-06-01, content hash 5c359c5dfbbc.

NOTABLE irrbb-core-share **The realised core share differs from the declared one by more than the declared tolerance**

The institution declared a core share of 75.0% and the book delivers 72.9% at the 36-month horizon. A core share set too high understates how much of the balance reprices, and therefore understates rate sensitivity in every figure computed from it.

Evidence to re-run: Balance-weighted over 6 cohort(s), content hash 5c359c5dfbbc.

NOTABLE irrbb-segment-spread **The model tracks retail-current and does not track sme-current**

Runoff error differs by 8.34 percentage points between the best and worst product class. This is a finding about where the model applies rather than about whether it works: a behavioural model calibrated on retail behaviour and applied to a class that behaves differently will be accurate in aggregate for as long as the aggregate is dominated by the class it was fitted to.

Evidence to re-run: Per-class runoff over 216 cohort-month(s), content hash 5c359c5dfbbc.

10 Third-party dependencies and residual risk

No third-party dependency is declared for this engagement. The behavioural models under review are estimated in-house from the institution's own balance history, and the pack reads only the book the institution supplied.

Where an institution's behavioural parameters are supplied or calibrated by a vendor ALM system, that is a dependency and it belongs on this page — the assumptions would then be somebody else's estimates carried on the institution's balance sheet, and the review would need to say who is accountable for revising them.

11 Limitations and tests not performed

Every check a reviewer might expect of this family and this pack does not perform is named here with its reason, so that what WAS performed means something.

THE PACK SHIPS NO THRESHOLD, which is the single most important limitation and is a property of the framework rather than of this platform: the governing guidelines place obligations on how behavioural assumptions are set, reviewed and evidenced, and publish no numeric tolerance for any of them. Each figure above is therefore judged against a number the institution declared, or is reported and not judged at all.

NOT PERFORMED — AND WHY

The supervisory outlier tests on economic value of equity and net interest income

THESE ARE THE ONLY TWO NUMERIC BARS IN THE DOMAIN and this pack computes neither, deliberately. A decline in EVE of more than 15% of Tier 1 capital (Directive 2013/36/EU Article 98(5)(a)) and a one-year decline in NII of more than 5% of Tier 1 (Commission Delegated Regulation (EU) 2024/856 Article 5(1)) are properties of the institution's ALM engine under six prescribed shock scenarios. Rebuilding that engine is the simulation this platform does not do, and a validator does not need it here: what is unvalidated is the BEHAVIOURAL INPUT, not the arithmetic stacked on it. The institution's own submitted results belong in the review as a declared claim reconciled against those limits — a number that came out of a submission must not sit in a figures table beside four this pack computed, under one heading, with one kind of authority

Credit spread risk in the banking book

EBA/GL/2022/14 covers CSRBB alongside IRRBB and this pack reads neither its evidence nor its assumptions. A cohort book of balances, rates and prepayments says nothing about the spread component, and a pack that quietly covered one of the guideline's two halves would be read as covering both. Named here so the half that is absent is visible

Basis risk between reference rates

This book carries ONE reference rate per observed month. Basis risk is the divergence BETWEEN reference rates, so it is not a figure a single-rate book can produce however it is read. A second rate column would be a different evidence shape and a different declaration, not a computation added to this one

Appendix A Reproducibility

Engine irrbb-metrics/1.0.0, metric pack "irrbb", thresholds from AlphaIQ IRRBB behavioural pack — NO threshold is shipped. Every row reads a tolerance the institution declares, because the framework publishes obligations for behavioural models and no numbers.

The cohort book is stored as supplied and content-hashed over its raw bytes; it is re-hashed on every read and a disagreement refuses the run rather than repairing it. The declaration is stored beside it and validated on resolution with the same function the parser uses, so a book cannot be measured against a tolerance the parser would have refused.

Every figure in this memo is a pure function of the stored book and the stored declaration. No model was executed, no replica was built and no random draw was made, so a rerun over the same two artefacts reproduces this document exactly.



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 2c21e071f554992354fba137d11bba50826542f343cb37b1a92661149ae1ea50

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