DOCUMENT 1 OF 3 · 18 PAGES

SR 11-7 Model Validation Workbook

Model-risk validation pack aligned to Federal Reserve SR 11-7 / OCC 2011-12 with effective challenge, ongoing monitoring, finding ratings and three-lines-of-defense sign-off.

Audience · Model Risk Management (2nd line), Internal Audit (3rd line), Model Owner (1st line), CRO, regulators
Companion documents · Fair-Lending Bias-Testing Pack · Adverse-Action XAI & Model Card Templates
Use as · Audit-grade artefact requiring named human sign-off
Version · v1.0 · 2026-05

§ 1

Contents & how to use this workbook

This workbook codifies the documentation, evidence and sign-off artefacts required for SR 11-7 / OCC 2011-12 compliant model validation. It is designed to drop into a regulated bank's Model Risk Management (MRM) workflow without modification.

How to use. Each section is self-contained. Treat the boxes labelled "evidence required" as a checklist — every item must be linked to a specific document, table or screenshot before the model is approved for production use.

§ 2

SR 11-7 framework & three lines of defense

SR 11-7 (Federal Reserve, 2011) and the parallel OCC 2011-12 establish the supervisory expectation that banks manage model risk through a documented framework covering development, implementation, use, validation, and governance. This workbook organises evidence against those expectations.

Three lines of defense — explicit responsibilities

Line	Function	Owns	Cannot also do
1st line	Model owner / business	Development, use, monitoring, remediation	Independent validation of own model
2nd line	Model Risk Management (MRM)	Independent validation, effective challenge, monitoring oversight, MRM policy	Development of the model
3rd line	Internal Audit	Periodic review of MRM framework, evidence sampling, board reporting	Validation execution

Independence boundary

Validators in the 2nd line must not have contributed to development, calibration, or feature selection of the model under review. Independence is documented in the validation report cover with a positive attestation.

> **Common finding.** "Validator independence not evidenced." Cure: every validation report must list the validators, their prior involvement (none), and the date their independence was verified by MRM leadership.

§ 3

Model inventory & risk tiering

Model identification

Field	Description	Source of truth
Model ID	Permanent identifier in the model inventory	Model registry
Model name & version	Human-readable name + semantic version	Model registry
Business use	Decision the model supports	Owner attestation
Quantitative method	e.g. logistic regression, GBM, neural net, LLM-assisted	Development doc
Data sources	Internal / external / vendor / synthetic	Data lineage
Materiality	$ exposure, customers affected, regulatory bearing	1st line + MRM

Risk-tier matrix

Tier	Criteria	Validation cadence	Board reporting
Tier 1 (high)	Material financial exposure, regulatory model, customer-facing credit/AML decisions	Pre-prod + annual	Each cycle
Tier 2 (medium)	Material decision support, internal capital / liquidity inputs	Pre-prod + biennial	On material change
Tier 3 (low)	Operational analytics, marketing scoring	Pre-prod, light annual review	By exception

Decision note. When risk tier is ambiguous, assign the higher tier and reassess on next review. Tier downgrades require MRM Committee approval and documented rationale.

§ 4

Development documentation requirements

SR 11-7 expects development documentation thorough enough that a competent third party can replicate the model. The minimum bar:

#	Element	Evidence required
1	Statement of intended use	Decision, population, region, exclusion rules
2	Theoretical basis / hypothesis	Literature references, internal rationale
3	Data sources & vintages	Source, extract date, period, refresh cadence
4	Data quality assessment	Completeness, accuracy, timeliness, representativeness
5	Feature engineering log	Transformations, missing-value handling, derivations
6	Sampling and partitioning	Train / validation / test / OOT splits, stratification
7	Algorithm choice rationale	Why this method beats benchmarks; what was rejected
8	Hyperparameter search	Search space, criterion, seed control, reproducibility
9	Performance metrics	Discrimination, calibration, stability — primary + secondary
10	Sensitivity & stress testing	Behaviour under adverse and outlier inputs
11	Limitations	Where the model should not be used; mitigation
12	Implementation plan	Production environment, monitoring, rollback

> **Reproducibility test.** The validator independently re-runs item 5–9 using only the development documentation. A failure to reproduce constitutes a finding of insufficient development documentation.

§ 5

Conceptual soundness review

Conceptual soundness asks: is the chosen approach theoretically appropriate for the intended use, given the data and the decision?

Review dimensions

Dimension	Validator checks	Common failure modes
Theoretical fit	Match between method and the problem (binary / continuous / sequence / interaction-heavy)	Linear model on highly non-linear relationship
Data sufficiency	Sample size relative to feature dimensionality, class balance, time coverage	Rare-event modelling without class adjustment
Variable selection	Predictive value vs. multicollinearity vs. fair-lending risk	Proxy variables for protected characteristics
Functional form	Splines, interactions, monotonicity constraints where business sense requires	Non-monotonic risk score in regulated decision
Assumptions	Independence, stationarity, representativeness	Pandemic-era data assumed representative of future
Boundary behaviour	Behaviour at extreme inputs, missingness, new categorical values	Silent fallback to global mean for unseen categories

Validator opinion

The validator records a written opinion on conceptual soundness with one of: **Acceptable**, **Acceptable with remediation**, or **Not acceptable**. "Not acceptable" blocks production use.

> **Anti-pattern.** "Black-box justification by performance." A model that scores well in-sample but cannot be explained is not conceptually sound for credit, AML or capital decisions; document the theoretical basis explicitly.

§ 6

Outcomes analysis & benchmarking

Performance evidence required

Class	Metric	Required for
Discrimination	AUC / KS / Gini · with confidence intervals	All credit / AML scoring
Calibration	Hosmer-Lemeshow, reliability curves, calibration-in-the-large	PD, LGD, PPNR, anything where the score is interpreted
Rank-order stability	PSI (Population Stability Index) by segment over time	All production scoring
Segment performance	By geography, channel, product, vintage	Tier 1 + 2
Benchmarking	Versus a documented challenger (simple, transparent baseline)	All
Stress / sensitivity	Extreme inputs, macro shock scenarios, adversarial perturbation	Tier 1

Benchmark selection

The challenger must be transparent (logistic regression, scorecard) and use the same target and population. The validator reports lift over benchmark in the same metric system as production decisions.

> **Decision note.** If the candidate model fails to beat a documented benchmark on the primary metric by a material margin, production approval should be denied unless an explicit business case justifies the additional complexity.

§ 7

Ongoing monitoring contract

The ongoing-monitoring contract is a written agreement between model owner and MRM stating what is monitored, at what frequency, and what triggers escalation.

Monitor	Frequency	Alert threshold	Action on breach
Score distribution (PSI)	Monthly	PSI > 0.25 (material)	Re-validation review
Feature drift (per-feature PSI)	Monthly	PSI > 0.25 on any top-10 feature	Feature-level investigation
Performance (AUC / KS)	Quarterly (when labels arrive)	≥ 10% relative decline vs baseline	Re-calibration / re-fit
Calibration drift	Quarterly	Reliability curve deviation by segment	Re-calibration
Override rate (1st line)	Monthly	> 5% sustained	Override-pattern analysis
Fair-lending disparity (AIR)	Quarterly	AIR < 0.80 in any protected class	Fair-lending review (see companion doc)
Data quality (input)	Daily	Missingness > X%, schema drift	Pipeline incident
System SLA	Continuous	Latency / availability breach	SRE on-call

The monitoring report is filed monthly with MRM and quarterly with the MRM Committee. Repeated yellow alerts auto-escalate to a re-validation trigger (§ 13).

§ 8

Effective challenge protocol

"Effective challenge" is SR 11-7's expectation that critical assumptions, data choices, and methodologies be tested by a competent, independent and informed party. This section codifies how the challenge is run.

Challenge sessions

1. **Pre-development challenge.** Validator reviews the development plan before modelling starts; can require changes to scope, data or method.
2. **Mid-development challenge.** Validator inspects feature engineering, sampling and benchmark choices.
3. **Pre-implementation challenge.** Validator independently reproduces key results, runs sensitivity tests, writes a finding-rated report.
4. **Annual challenge.** Validator re-tests assumptions against current data; may require re-validation.

Challenge log

Session	Date	Challenger	Items raised	Disposition
Pre-development				
Mid-development				
Pre-implementation				
Annual				

> **Quality marker.** Empty disposition columns are a red flag in audit. Every item must record: accepted (with change), accepted with rationale (no change), or rejected — and the named decision-maker.

§ 9

Limitations & assumptions register

Every model embeds assumptions and has limitations. Failure to document them is a recurring finding. The register below is the canonical place.

#	Assumption / limitation	Source	Impact if violated	Mitigation	Owner
1	Population stationarity over training horizon	Development	Performance decline	Monthly PSI	Owner
2	Independence of training observations	Sampling	Optimistic performance	Block resampling test	Owner
3	No protected-class proxies among inputs	Variable selection	Fair-lending risk	Companion doc tests	Fair-lending officer
4	Adequate sample in tail segments	Data	Tail-segment bias	Tail-segment metrics	Validator
5	Vendor data refresh on contract cadence	Vendor	Stale features	Freshness SLO	Vendor mgr
6	Override rate below X%	Use	Erosion of model authority	Override review	Business
7	No silent fallback for unseen categories	Implementation	Hidden errors	Schema validation	Engineering

The register is reviewed at every annual challenge and updated whenever an assumption is challenged in production.

§ 10

Finding ratings & remediation tracking

Finding-rating scale

Rating	Definition	Production impact	Remediation SLA
CRITICAL	Material risk that blocks safe use	Cannot deploy / must withdraw	Immediate
HIGH	Significant weakness with compensating controls available	Deploy with conditions	≤ 90 days
MEDIUM	Documentation or rigour gap, not blocking	Deploy with action plan	≤ 180 days
OBSERVATION	Improvement opportunity	No restriction	By next cycle

Remediation log template

Finding #	Rating	Description	Owner	Due	Status	Verified by
F-001						
F-002						
F-003						
F-004						
F-005						

Decision note. Critical and unremediated High findings must be reported to the Model Risk Committee at every cycle until cleared. Internal Audit independently samples remediations.

§ 11

GenAI / foundation-model addendum

Generative AI and foundation-model usage requires controls that go beyond classical SR 11-7. This addendum extends the standard validation pack.

Control	What to evidence	Owner
Use-case classification	Decision-support / drafting / autonomous; risk tier	1st line
Prompt & system-message registry	Versioned, signed, change-controlled	Engineering
Retrieval grounding	Sources, freshness SLO, citation policy	Engineering
Hallucination & refusal monitoring	Sampled offline + online evals with thresholds	MRM
Toxicity / PII leakage tests	Pre-prod + ongoing	MRM + Privacy
Vendor model controls	Indemnification, update notice, exit plan	Procurement
Human-in-the-loop contract	Where the human must sign off; logged	Business
Audit trail	Prompt, response, model version, confidence, override	Engineering

Forbidden by default. Autonomous credit, AML, fair-lending, or capital decisions without a human in the loop. Re-enabling requires explicit written sign-off from MRM and the responsible business head.

§ 12

Vendor / third-party model controls

Stage	Control	Evidence
Pre-contract	Vendor due diligence (financial, security, model risk)	VDD report
Pre-contract	SOC 2 / ISO 27001 / SR 11-7 alignment statement	Vendor attestation
Contract	Right-to-audit, model documentation access, update notice	Master agreement
Contract	Indemnification, IP warranty, data-use restrictions	Master agreement
Onboarding	Independent validation of vendor model on bank data	Validation report
Ongoing	Vendor update review board (any version change triggers re-test)	Change log
Exit	Data return / destruction; portability plan	Exit playbook

Vendor models are treated as in-scope for SR 11-7 to the same standard as internal models; lack of vendor cooperation is itself a finding.

§ 13

Change management & re-validation triggers

Triggers requiring re-validation

- Material change to the modelling methodology (algorithm class, target definition, training population)
- Material change to inputs (new vendor data, deprecated feature, refresh cadence)
- Material change to use (new product, new geography, new threshold)
- Sustained breach of any monitoring threshold for ≥ 2 cycles
- Fair-lending finding of adverse-impact ratio < 0.80 in any protected class
- Regulatory change (new guidance, MRA, Consent Order)
- Annual review cadence (Tier 1 + 2)

Change-control gates

Change type	1st line	MRM	MRM Committee	Audit
Re-calibration (no structural change)	Plan	Review & sign	Notified	Sampled
Re-fit (same algorithm, new data)	Develop	Validate	Approve	Sampled
New version (structural)	Develop	Full validation	Approve	Yes
Retirement	Propose	Sign	Notified	Notified

§ 14

MRM Committee briefing template

One-page briefing that travels to the Model Risk Committee. Fill in for every approval decision.

Model identity

ID: Name: Version:
Tier: Owner: Validator:

Decision requested

[ ] New approval [ ] Re-approval [ ] Conditional approval [ ] Retirement

Conceptual soundness

Acceptable · Acceptable with remediation · Not acceptable

Performance vs benchmark

Primary metric · point + CI · lift over benchmark

Open findings

Critical: __ · High: __ · Medium: __ · Observations: __

Fair-lending status

AIR by protected class · LDA search complete (Y/N)

Limitations & mitigations

Top 3 limitations · compensating controls

Monitoring contract

Filed (Y/N) · review cadence · escalation owner

Committee decision. Approve · Approve with conditions · Reject · Defer — recorded in committee minutes with named voting members.

§ 15

Audit evidence index

One row per evidence artefact. Internal Audit samples this index to test the validation.

Evidence #	Description	Section	Location / link	Owner
E-01	Validator independence attestation	§ 2	MRM repo	MRM lead
E-02	Model inventory entry	§ 3	Model registry	Owner
E-03	Development documentation	§ 4	Project workspace	Owner
E-04	Reproducibility report	§ 4	MRM repo	Validator
E-05	Conceptual soundness opinion	§ 5	Validation report	Validator
E-06	Outcomes & benchmarking report	§ 6	Validation report	Validator
E-07	Ongoing monitoring contract	§ 7	MRM repo	Owner + MRM
E-08	Effective challenge log	§ 8	MRM repo	MRM
E-09	Limitations register	§ 9	MRM repo	Owner + MRM
E-10	Findings & remediation tracker	§ 10	MRM repo	MRM
E-11	GenAI addendum (if applicable)	§ 11	MRM repo	Owner
E-12	Vendor due-diligence pack	§ 12	Procurement	Vendor mgr
E-13	Change-control records	§ 13	Registry	Owner
E-14	MRM Committee minutes	§ 14	Committee repo	Secretary

§ 16

Glossary, dependencies & references

Glossary

Term	Meaning
AIR	Adverse-Impact Ratio (fair lending)
BISG	Bayesian Improved Surname Geocoding (proxy)
KS	Kolmogorov–Smirnov statistic
LDA	Less-Discriminatory Alternative (fair-lending search)
MRA	Matter Requiring Attention (regulatory finding)
MRM	Model Risk Management (the 2nd-line function)
OOT	Out-of-Time validation sample
PSI	Population Stability Index
SaMD	Software as a Medical Device (cross-sector reference)

Reference framework

- Federal Reserve SR 11-7 (Guidance on Model Risk Management)
- OCC Bulletin 2011-12 (parallel guidance)
- FFIEC Examination Handbooks (IT, BSA/AML)
- Companion: Fair-Lending Bias-Testing Pack (ECOA / FHA mapping)
- Companion: Adverse-Action XAI & Model Card Templates

Dependencies on the broader bank

- Authoritative model inventory and registry
- Data governance with documented lineage
- MRM independence and resourcing
- Internal Audit access to MRM artefacts

§ 17

Sign-off & attestation

Sign-off block

Role	Name	Date	Signature
Model owner (1st line)			
Model risk lead (2nd line, validator)			
MRM head			
Fair-lending officer			
Internal audit (3rd line)			
CRO (or delegate)			

Submission to the Model Risk Committee requires every row above to be completed. Wet signatures or controlled e-signatures are equally acceptable; both must be archived alongside the validation report.