Which Performing CMBS Loans Break Next?

Research note · Gaurav Gupta, CFA, FRM, CAIA · 20 October 2026 · Independent research on public SEC data; not deployed; not investment advice

36.5%
of distressed balance caught by reviewing the riskiest 10% of loans (watchlist rule 28.8%)

0.770
out-of-time AUC, refit quarterly (0.629 if never refit; rule 0.647)

51.8%
of performing loans within 12 months of maturity entered distress within a year

6% → 39%
office share of distress events, 2019-20 vs 2022-23

Question and answer

Question. Of conduit CMBS loans paying on time today, which enter special servicing, 60+ day delinquency or maturity default within 12 months?

Answer. A gradient-boosted model refit every quarter on resolved outcomes ranks the book materially better than a servicer-style watchlist. The same model never refit is worse than the watchlist: the stress moved from hotels and retail (2020) to office and refinancing (2022-23).

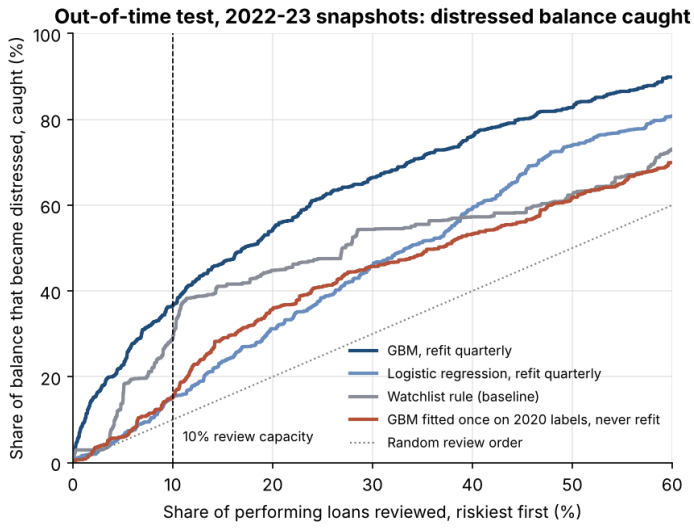
Results: walk-forward test, 2022Q2-2023Q2

Model	AUC	Top 10%: events	Top 10%: balance
Watchlist rule	0.647	26.1%	28.8%
Logistic (selected on dev)	0.653	18.8%	15.2%
GBM, fundamentals only	0.762	37.9%	38.4%
GBM, with property type	**0.770**	**38.1%**	**36.5%**
GBM fit once on 2020, never refit	0.629	18.1%	15.3%

38,765 snapshots, 8,143 loans, 755 events (1.95%). AUC 95% CI (loan-clustered bootstrap): GBM 0.744-0.793; rule 0.618-0.678.

Business outcome (illustrative)

Per $10bn performing book: 2.74% ($274M) entered distress within a year. Reviewing 10% of loans reaches ~$100M with the model vs ~$79M with the rule. At an *assumed* 35% severity and 15% loss avoided by acting a year earlier: ~$5.2M vs ~$4.1M, about $1.1M more per $10bn per year. Review hit rate 7.4% vs 5.1%.



Methodology

Sources. SEC EDGAR Form ABS-EE exhibit 102 (Reg AB II loan-level CMBS data) for 190 conduit trusts issued 2016-2021; copy used: per-trust parquet mirrors by DenyTranDFW on Hugging Face (GPL), each tape pinned by dataset revision and sha256. Market context: Trepp via Scotsman Guide (6 Oct 2026).
Panel. 3,280 quarterly tapes, Dec 2019-Jun 2024; 10,531 loans; 109,543 performing loan-quarter snapshots; 2,376 distress events (2.17%).
Label. Special-servicing transfer, 60+ days delinquent, non-performing matured balloon or workout code within the next four quarterly tapes.
Features (27, same tape only). DSCR and NOI vs underwriting, debt yield, occupancy and change, months to maturity, coupon, loan age and size, largest-tenant share and lease term, financials staleness, prior distress on earlier tapes, property type.
Models. Watchlist trigger count (DSCR<1.10x, occupancy<80%, NOI -25%, maturity<12m, large tenant rolling); logistic regression; monotone LightGBM with and without property type.
Validation. Walk-forward: each quarter scored by models fitted only on snapshots 4+ quarters older (asserted). Selection on 2020Q4-2021Q2 only. Checks: shuffled labels (mean AUC 0.47, five runs), history-flag ablation (0.765), frozen model, selection-gap bootstrap.
Key calcs. Capture = distressed balance in top 10% by score ÷ all distressed balance. PSI dev→test 0.92. Mean predicted 0.95% vs 1.95% observed.

What we learned that matters

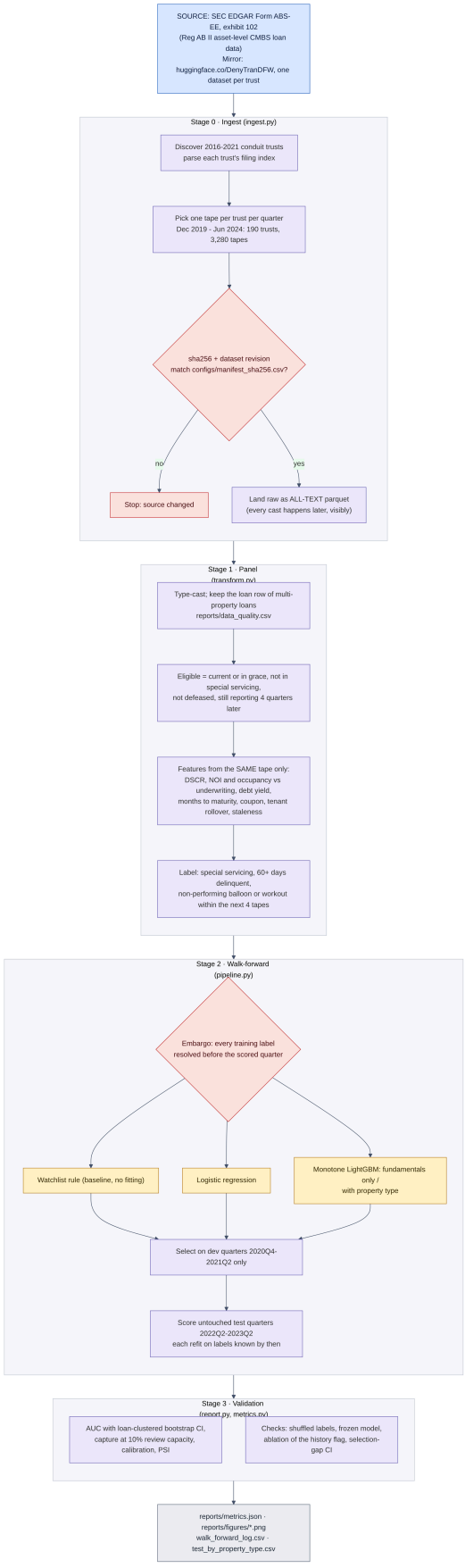
- **Cadence beats algorithm.** Frozen on 2020 outcomes, the model loses to the rule.
- **Selection risk is real.** Dev window picked logistic (gap CI -0.031 to +0.064); on test it caught 15.2% vs 36.5%.
- **Refinancing wall.** 51.8% distress within a year inside 12 months of maturity; 0.45% beyond seven years. Distress by coupon quintile: 0.8% to 3.4%.

Recommendations

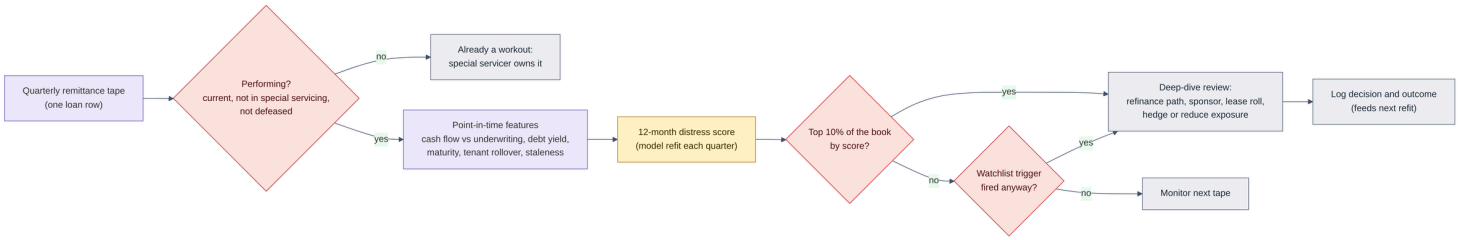
1. Order the surveillance queue by the refit score; keep watchlist triggers as a floor.
2. Refit quarterly with a one-year embargo; fall back to the rule after two quarters of underperformance.
3. Run champion and challenger in parallel; promote on resolved evidence.
4. Rank with the score; set reserves with a stressed scenario.
5. Review every loan inside 24 months of maturity regardless of score.

Code: github.com/guptagaurav-info/cmbs-distress-early-warning · Article: guptagaurav.info/articles/cmbs-distress-early-warning · Methods: Ke et al. 2017 (LightGBM); Lundberg et al. 2020 (TreeSHAP); Hanley & McNeil 1982 (AUC); Efron & Tibshirani 1993 (bootstrap). Loss severity and mitigation are illustrative assumptions.

Pipeline


SOURCE: SEC EDGAR Form ABS-EE, exhibit 102 (Reg AB II asset-level CMBS loan data) Mirror: huggingface.co/DenyTranDFW, one dataset per trust
Stage 0 · Ingest (ingest.py)
Discover 2016-2021 conduit trusts parse each trust's filing index
Pick one tape per trust per quarter Dec 2019 - Jun 2024: 190 trusts, 3,280 tapes
sha256 + dataset revision match configs/manifest_sha256.csv?
no
yes
Stop: source changed
Land raw as ALL-TEXT parquet (every cast happens later, visibly)
Stage 1 · Panel (transform.py)
Type-cast; keep the loan row of multi-property loans reports/data_quality.csv
Eligible = current or in grace, not in special servicing, not defeased, still reporting 4 quarters later
Features from the SAME tape only: DSCR, NOI and occupancy vs underwriting, debt yield, months to maturity, coupon, tenant rollover, staleness
Label: special servicing, 60+ days delinquent, non-performing balloon or workout within the next 4 tapes
Stage 2 · Walk-forward (pipeline.py)
Embargo: every training label resolved before the scored quarter
Watchlist rule (baseline, no fitting)
Logistic regression
Monotone LightGBM: fundamentals only / with property type
Select on dev quarters 2020Q4-2021Q2 only
Score untouched test quarters 2022Q2-2023Q2 each refit on labels known by then
Stage 3 · Validation (report.py, metrics.py)
AUC with loan-clustered bootstrap CI, capture at 10% review capacity, calibration, PSI
Checks: shuffled labels, frozen model, ablation of the history flag, selection-gap CI
reports/metrics.json · reports/figures/*.png walk_forward_log.csv · test_by_property_type.csv

Surveillance decision logic for one loan



Production design (proposed, not implemented)

