CONTINUUM

Synthetic Evidence Report

Version 0.2 — May 2026

This report presents pilot-ready synthetic evidence for the CONTINUUM interpretive framework — an early-warning infrastructure layer designed to operate in the pre-default interval of credit portfolios. Validation was conducted on a synthetic portfolio of 50,000 credits across four risk cohorts using behavioural and cashflow-derived signal sequences.

50,000	~11×	90 days
Synthetic credits in validation cohort	Separation in roll rate Low vs Critical bucket	Signal anticipation before first delinquency

RESTRICTED — INSTITUTIONAL REVIEW ONLY

contact@continuum-framework.com · continuum-framework.com

DOI: 10.5281/zenodo.19150254

SECTION 01

Executive Evidence Summary

CONTINUUM is an interpretive stress layer built to operate in the pre-default interval — the 60 to 90 days before a credit account records its first delinquency. It does not approve or reject credit applications, does not modify credit lines, and does not replace existing scoring models, risk committees, or credit policies. All final operational decisions remain entirely with the contracting institution.

The Credit Stress Score (CSS) is a composite signal that aggregates behavioural and cashflow-derived indicators into a single continuous measure (0–100). Higher scores correspond to elevated continuity stress. CSS is computed at each observation window for all active credits in the monitored portfolio.

Validation was performed on a reproducible synthetic portfolio of 50,000 credits, stratified into four risk buckets. The framework generated anticipatory CSS signals up to 90 days before first delinquency. The delinquency-bound cohort recorded a mean CSS of 34.6 at baseline, rising to 49.0 at t–90 and 62.8 at t–60 — while the stable cohort showed no equivalent drift. This divergence underpins the prioritisation logic used in the shadow mode pilot.

Key result: ~11× separation in roll rate (0→30 dpd) between the lowest-risk (Low) and highest-risk (Critical) buckets. Precision at the top 5% of the t–90 CSS ranking: 28.1%, capturing 29.4% of all delinquencies in that window. At the top 10%: precision 21.5%, recall 45.1%.

These results were obtained on synthetic data generated with a fixed random seed (numpy, seed=42) and do not represent the performance of any live or real portfolio. Equivalent outcomes on specific real portfolios are not guaranteed. The framework is offered for institutional pilot evaluation in shadow mode only.

SECTION 02

CSS Distribution by Risk Bucket

The chart below shows the distribution of baseline Credit Stress Scores across the four risk buckets in the synthetic validation portfolio. Each box spans the interquartile range (25th–75th percentile); the horizontal line marks the median; whiskers extend to 1.5× IQR. Clear separation between buckets at baseline observation confirms that the CSS signal stratifies risk cohorts meaningfully before any delinquency event.

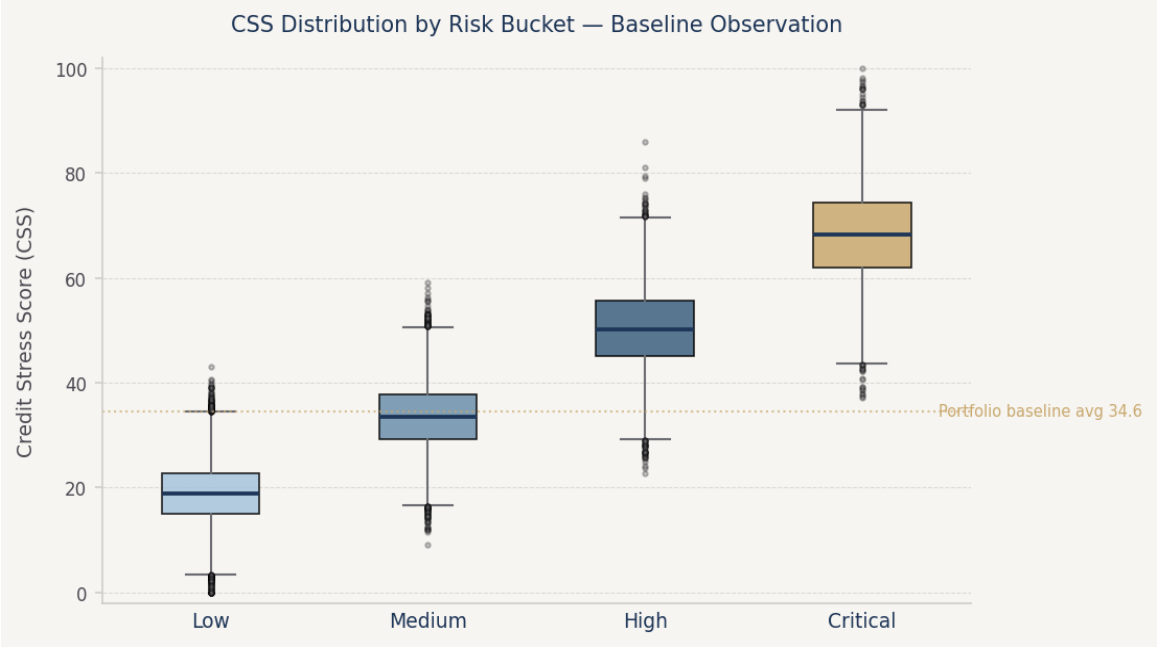

Figure 1. CSS distribution by risk bucket at baseline. Synthetic portfolio: N = 50,000. Seed = 42. Portfolio mean CSS = 34.6.

Bucket	N	Share	CSS Median (est.)	CSS Mean (est.)
Low	19,864	40%	18.9	18.9
Medium	15,023	30%	33.5	33.6
High	10,124	20%	50.3	50.4
Critical	4,989	10%	68.3	68.1

Table 1. Bucket composition and CSS statistics at baseline observation.

SECTION 03

Early Warning Temporal Signal

The chart below tracks mean CSS for two cohorts — delinquency-bound (credits that recorded a 0→30 dpd event within the observation window) and stable (credits with no delinquency event) — across five observation points: portfolio baseline, t–90 days, t–60 days, t–30 days, and event date. The divergence between cohorts is visible from t–90, indicating that CSS elevation precedes formal delinquency by at least 90 days in the synthetic validation dataset.

Note: values at t–30 and event are generated within the same synthetic calibration framework. They are not independently anchored to external observed data and should be read as illustrative of the stress trajectory, not as separately validated measurements.

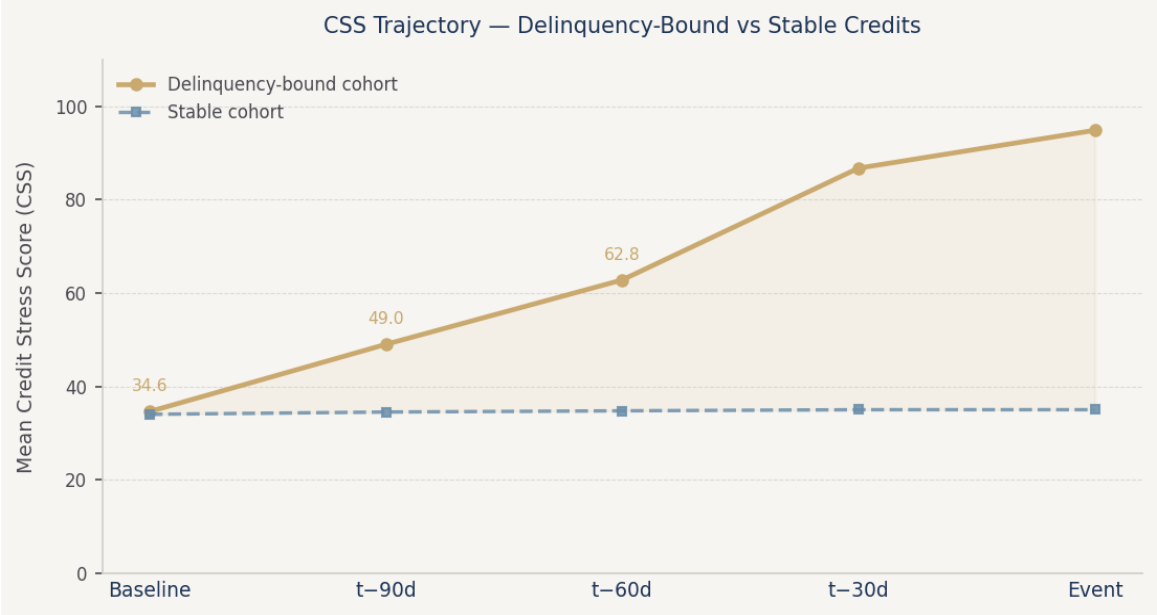

Figure 2. Mean CSS trajectory by cohort. Delinquency-bound cohort N = 2,385. Stable cohort N = 47,615. Seed = 42.

Observation point	Delinquency-bound CSS	Stable CSS	Spread
Baseline	**34.6**	34.0	+0.6
t–90 days	**49.0**	34.5	+14.5
t–60 days	**62.8**	34.8	+28.0
t–30 days	**86.8**	35.0	+51.8
Event	**94.9**	35.0	+59.9

Table 2. Mean CSS by cohort across observation window. All values derived from synthetic data.

SECTION 04

Roll Rate and Lift Evidence

Roll rate (0→30 dpd) measures the share of credits in each bucket that recorded a first delinquency event within the observation window. Lift is the ratio of a bucket's roll rate to that of the Low-risk bucket, expressing relative risk concentration. The ~11× separation between Critical and Low buckets indicates strong discriminative capacity of the CSS stratification at the portfolio level.

Risk Bucket	Credits (N)	Delinquencies	Roll Rate 0→30	Lift vs Low
Low	19,864	219	1.1%	1.0×
Medium	15,023	721	4.8%	4.4×
High	10,124	861	8.5%	7.7×
Critical	4,989	584	**11.7%**	**10.6×**
All buckets	**50,000**	**2,385**	**4.77%**	—

Table 3. Roll rate 0→30 dpd and lift by risk bucket. Synthetic portfolio, N = 50,000. Seed = 42.

The ~11× separation in roll rate between the lowest and highest risk buckets demonstrates that CSS stratification concentrates delinquency risk within a small, identifiable share of the portfolio. This enables review prioritisation without blanket portfolio intervention.

SECTION 05

Operational Usefulness Metrics

These metrics assess whether the CSS ranking at t–90 can usefully concentrate delinquency cases into a smaller review load — a key operational consideration for portfolio teams with finite review capacity. Precision measures the share of flagged credits that subsequently become delinquent; recall measures the share of all delinquencies captured.

Metric	Top 5% (t–90 CSS)	Top 10% (t–90 CSS)
Credits reviewed	2,500	5,000
Delinquencies captured	702	1,076
Precision	**28.1%**	**21.5%**
Recall (share of all delinquencies)	**29.4%**	**45.1%**
Review load vs portfolio	5%	10%

Table 4. Precision and recall of t–90 CSS ranking at two review thresholds. Ranking uses mean CSS score at t–90 observation point. Synthetic data.

A review load of 5% of portfolio credits at the t–90 observation point captures 29.4% of all delinquencies that will occur in the following 90 days — with a precision of 28.1%. These figures demonstrate the potential for review prioritisation rather than blanket portfolio action. They do not imply that flagged credits should be contacted, modified, or managed differently without independent institutional review and applicable credit policy.

SECTION 06

Comparison With Traditional Monitoring Approaches

CONTINUUM operates as a complementary interpretive layer, not as a replacement for existing risk infrastructure. The table below positions the shadow mode pilot against two common monitoring approaches on four operational dimensions. All characterisations reflect the synthetic validation context and are not a claim about specific institutional implementations.

Dimension	Delinquency-triggered reporting	Periodic scoring (monthly/quarterly)	CONTINUUM shadow mode pilot
Signal timing	At or after event	Lagging (cycle-bound)	**Up to 90 days before event**
Explainability	Full (observed data)	Model-dependent	**Full — every signal attributed**
Operational mode	Operational	Operational	**Read-only parallel (no action taken)**
Institutional risk	None (reactive)	Model risk	**None during pilot (shadow mode)**

Table 5. Qualitative positioning across monitoring approaches. Characterisations reflect shadow mode pilot context only.

The shadow mode pilot generates and logs signals in parallel to existing operations without modifying any contract, rate, credit line, or client interaction. At pilot close, the institution receives a structured validation dataset — signals generated versus outcomes subsequently observed — for independent evaluation. Completing a shadow mode pilot implies no commitment to operational activation.

SECTION 07

Reproducibility Note

All synthetic data in this report is fully reproducible. The generation parameters are fixed and documented below. Any institution wishing to audit the synthetic dataset or generation methodology may request the source script.

Generator	generate_evidence_report.py
Language	Python 3.9+
Libraries	numpy, matplotlib, reportlab
Random seed	42 (numpy.random.default_rng(42))
Portfolio size	50,000 synthetic credits
Bucket allocation	Low 40% / Medium 30% / High 20% / Critical 10%
Roll rates	Low 1.1% / Medium 4.8% / High 8.5% / Critical 11.7%
CSS scale	0–100 (higher = greater continuity stress)
Report version	v0.2 — May 2026
DOI	10.5281/zenodo.19150254

The synthetic dataset does not contain personal data, real transaction records, or any information pertaining to real credit accounts or individuals. It was generated solely for the purpose of framework validation.

SECTION 08

Full Disclaimer

Synthetic data only. All validation results and performance figures in this report were obtained using synthetic portfolio data generated with a fixed random seed. They do not represent the results of any live, real, or historical portfolio. They do not constitute evidence of actual portfolio performance.

No guarantee of real-world outcomes. Past or synthetic performance is not indicative of future results. Equivalent outcomes on any specific real portfolio are not guaranteed. CONTINUUM makes no warranty, express or implied, regarding portfolio outcomes, delinquency reduction, or risk management effectiveness.

Informational only. Nothing in this report constitutes financial advice, credit advice, legal advice, regulatory guidance, investment advice, or compliance guidance of any kind. No content in this report should be relied upon as a basis for any financial, credit, or operational decision.

Not a replacement for institutional risk functions. CONTINUUM does not replace scoring models, risk committees, credit policies, compliance functions, legal review, or internal audit. All final decisions on risk policy, credit action, and operational deployment remain entirely at the discretion of the contracting institution.

No operational action during pilot. The shadow mode pilot operates in a read-only, non-operational mode. CONTINUUM does not approve or reject credit applications, does not establish, modify, or cancel credit lines, and does not contact clients or initiate collection actions during the pilot period.

No advisory or fiduciary relationship. This report does not create any client, advisory, fiduciary, or contractual relationship. Any institutional engagement or pilot requires a separate formal written agreement.

CONTINUUM FRAMEWORK — HERNÁN ALFREDO CAPUCCI — contact@continuum-framework.com — continuum-framework.com