LocaPerk

LocaPerk Labs Working Paper Series

Working Paper No. 002

LocaPerk Labs Research Note

MBD-mini Temporal Robustness, Rebalancing, and Probability Guardrails

---

Author: Damilola Olayemi

Founder, LocaPerk

Lead Researcher, LocaPerk Labs

Date: 2026

Public Research Note

These working papers document production-facing methods, deployment use cases, and governance positions from the LocaPerk Labs programme. The views expressed are those of the author and do not necessarily represent the views of partner institutions.

locaperk.com/labs

dami@locaperk.com

Executive Summary

This updated research note documents how the MBD-mini behavioural pipeline has progressed from compact transaction-led prototyping toward a more realistic temporal robustness programme. The latest reports show stronger timestamp coverage, improved rebalance performance under longer-horizon testing, and a clearer separation between ranking utility and probability certainty. At the same time, the probability-engine report still marks the final engine as not yet reliable enough for universal watchlist or triage probability claims, reinforcing a key deployment lesson: behavioural systems are currently strongest when used for relative ranking, concern tiers, and governed review-first workflows rather than absolute probability automation.

Keywords

Behavioural financial intelligence • Temporal robustness • Probability guardrails • Temporal strategy • Concern tiers • Governed triage

Table of Contents

1. Why The Temporal Question Matters

1.1 From Fast Prototyping To Longer-Horizon Validation

Transaction data offers one of the clearest windows into how financial life is evolving over time, yet it is still underused in operational financial intelligence. Earlier MBD-mini work focused on compact engineering validation. The latest temporal rebalance reports now extend that work into a longer-horizon setting covering 725 days from 2024-01-01 to 2025-12-27, making the temporal question more realistic even though cohort fallback logic is still needed.

This shift matters because the central production question is no longer just whether behavioural signals can rank emerging concern. It is whether they remain useful when cohorts move, class balance changes, and probability calibration comes under pressure. That is the real deployment boundary for live decision-support infrastructure.

2. Behavioural Signal Engineering

2.1 From Raw Transactions To Interpretable Features

The first stage of the pipeline still converts raw transaction activity into monthly behavioural snapshots. These snapshots summarise recurring financial behaviour through interpretable concepts rather than raw ledger history. Liquidity resilience approximates how much buffer appears to exist relative to typical outflows. Income stability describes the rhythm and consistency of inflows. Spending pressure captures whether spending appears to be intensifying relative to historical norms or incoming resources.

Additional features describe obligation coverage, cashflow volatility, payment consistency, and irregularity in recurring commitments. These features remain useful because they are legible to operators and governance teams as well as to model training pipelines. Their value lies in operational readability as much as in predictive utility.

2.2 Why Feature Legibility Still Matters

- Liquidity resilience is easier to explain operationally than a latent hidden factor.
- Spending pressure and cashflow volatility align with how support and relationship teams already think about strain.
- Payment consistency and obligation coverage create a bridge between behavioural interpretation and risk oversight.

3. Temporal Model Selection

3.1 What The Latest Rebalance Tests Show

The temporal rebalance summary identifies a strongest temporal candidate in the latest holdout tests, together with a recency-aware strategy and hybrid balancing as the best class-balance approach. These results matter because they indicate that the strongest temporal candidate is not simply the strongest static setup repeated over time; it is a recency-aware configuration designed for evolving cohorts.

- A recency-aware temporal strategy performed best in the latest rebalance report.
- Best balancing strategy: `hybrid`.
- Current temporal readiness in the report: `not_ready`.

3.2 Why Probability Stabilisation Remains A Separate Layer

A central lesson from behavioural modelling is that strong ranking does not automatically imply stable probabilities. The probability-engine report still recommends a leading final engine, yet explicitly marks it as not reliable enough for broad watchlist or triage probability claims. This is not a contradiction. It shows that ranking utility can mature before universal probability certainty does, which is exactly why probability stabilisation must remain a separate governed layer.

4. Operational Interpretation

4.1 What Operational Success Looks Like

For this kind of system, the most important practical question is whether it can concentrate later concern into relatively small review groups. If an institution can review only a limited share of customers, does the behavioural ranking help those reviews capture a disproportionate share of future concern? This ranking-focused framing is more relevant than raw accuracy because the product is intended for queue prioritisation and triage.

The latest MBD-mini temporal work supports that framing. Behavioural transaction signals, recency-aware ranking strategies, and guarded recalibration can improve the focus of small review buckets. That makes them useful for watchlisting and proactive review, even where absolute probability claims still need caution.

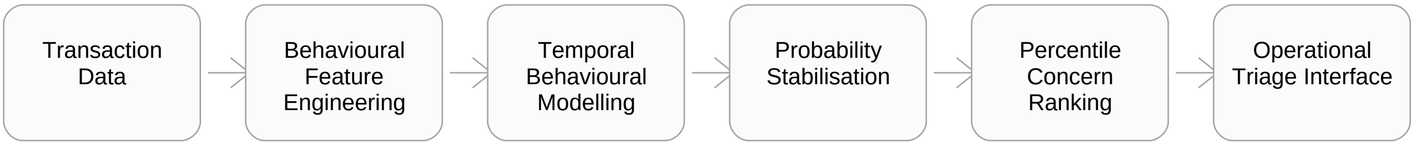

Figure 1: Behavioural Intelligence Pipeline

4.2 Concern Tiers Remain More Defensible Than Hard Probability Claims

The most responsible conclusion remains that probabilities are most useful when translated into percentile concern bands rather than treated as rigid universal cutoffs. Relative tiering is better aligned with real-world drift. It also reduces the temptation to over-claim precision in settings where the evidence is still developing.

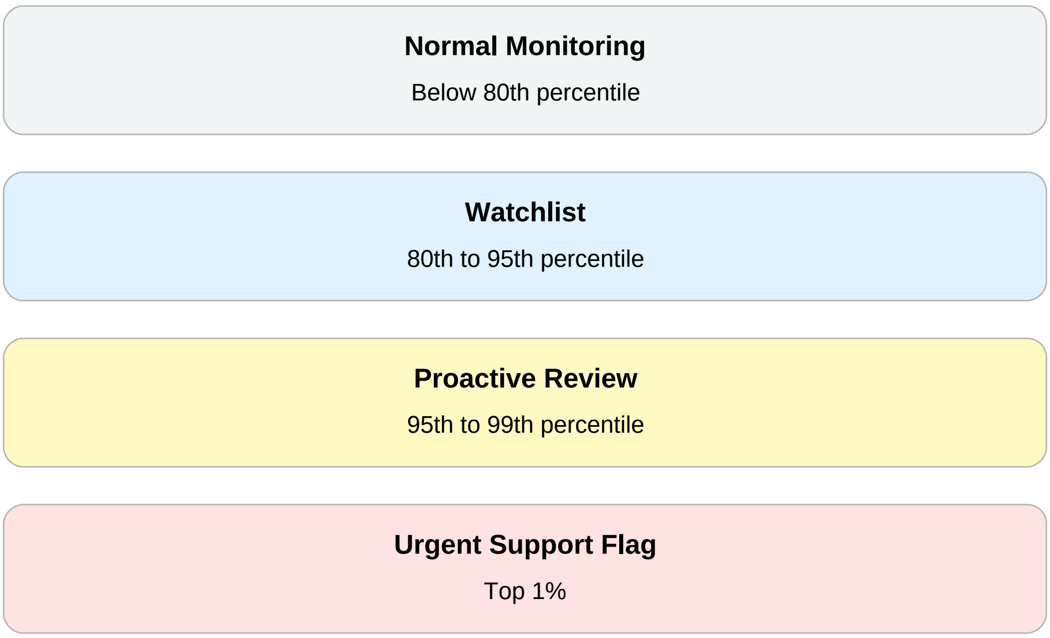

Figure 2: Example Concern Tier Structure

5. Governance And Deployment Boundaries

5.1 What The Reports Support Today

- Use behavioural models for prioritisation, monitoring, and review support before using them for stronger automated claims.
- Treat ranking and concern tiers as more mature than universal probability promises.
- Keep explainability and operator review central when temporal reliability remains incomplete.
- Use live deployment in monitored partner workflows as evidence-building infrastructure, not as the end of validation.

This is why the current product posture is governed deployment. The system can already support live review-first workflows, and the partner-case scoring path can call a remote production scoring service. But the temporal reports still justify caution around stronger autonomous interpretations, which is exactly the kind of boundary mature Labs publications should make explicit.

6. Conclusion

6.1 A Stronger But Still Guarded Research Position

The updated MBD-mini note shows that the behavioural intelligence programme has moved beyond a small prototype into a more credible temporal robustness phase. Recency-aware rebalancing strategies perform best in the rebalance report, and the system is increasingly useful for live review and triage support.

At the same time, the probability reports still justify caution. The strongest current operational claim is therefore not universal probability precision, but governed decision support: relative ranking, concern tiers, watchlisting, proactive review, and monitored deployment while the evidence base continues to mature.

7. References

• Bazarbash, M., and Beaton, K. (2020). Filling the gap: Digital credit and alternative data. IMF Working Paper.

• Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., and Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. Journal of Finance, 77(1), 5–47.

• Makridakis, S., Spiliotis, E., and Assimakopoulos, V. (2022). Forecasting and temporal instability in applied predictive systems. International Journal of Forecasting.

• Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1, 206–215.

• LocaPerk Labs. (2026). Temporal probability engine report and temporal rebalance summary.