Reviewer 1 Comments - Round 2

Manuscript: NarrativeBreak: Integrating Structural Break Detection with Multi-Source NLP Signals for Dynamic Portfolio Optimization **Recommendation:** Accept with Minor Revision **Confidence:** High

---

Summary

The authors have done an excellent job addressing the concerns raised in Round 1. The new calibration validation (Table A1), real-data case study (Appendix B), and regime-conditional analysis substantially strengthen the paper. I am now satisfied that the synthetic data approach is appropriately validated and the lead time mechanism is adequately explained.

The statistical analysis is now more rigorous, with appropriate Bonferroni corrections and HAC-robust tests. The honest acknowledgment of marginal significance after multiple testing correction is commendable.

I have only minor comments remaining.

Minor Comments

1. HMM Regularization (Section 3.4)

The description of Bayesian regularization with Dirichlet priors is helpful but could be more explicit. Please add: - The specific alpha values used (currently says "uninformative" but alpha=1.0 is only stated in the response letter) - How this affects the effective degrees of freedom in the transition matrix - Whether results are sensitive to the prior specification

Suggested addition: "We use Dirichlet priors with alpha=1.0 (uniform prior) for each row of the transition matrix. This regularization prevents degeneracy when sample sizes are small but has minimal effect with our 2,500+ day training sample."

2. Appendix Organization

The current appendix structure could be improved: - Appendix A (Calibration) is referenced frequently and is appropriate - Appendix B (Case Study) is well-placed - Consider renaming to more descriptive titles: - Appendix A: "Synthetic Data Calibration and Validation" - Appendix B: "Real-Data Validation: March 2020 COVID-19 Period"

3. Table A1 Presentation

Table A1 now includes KS test p-values, which is excellent. Minor suggestions: - Add a column for the sample size used to compute each statistic - Note that high p-values indicate the synthetic matches real, not that there's "no relationship"

4. Lead Time Confidence Interval

The 95% CI for lead time ([1.2, 10.2] days) is now reported, which is helpful. However, this wide interval should be mentioned in the abstract or introduction, not just Section 5.2.

Assessment

The authors have fully addressed my major concerns. The paper now provides appropriate caveats about synthetic data limitations while demonstrating reasonable calibration to real-world statistics. The regime-conditional analysis of the ablation study is particularly insightful.

I recommend **Accept with Minor Revision** contingent on addressing the above clarifications.

Note: This is a simulated review for demonstration purposes.