BEYOND STATISTICAL CHANGEPOINT DETECTION: SEMANTIC INTERPRETATION OF TIME SERIES VIA LARGE LANGUAGE MODELS

Hong Kiat Tan[1,2*] **Trilokya Akula**[1] **Akash Tonne**[1] **Tom Blake**[1]

[1]Amazon.com, Inc. [2]University of California, Los Angeles

ABSTRACT

Changepoint detection algorithms identify *where* structural breaks occur but are conventionally used under a one-to-one mapping between detected breaks and real-world events. We show this mapping assumption is undermined by a fundamental ambiguity: the confidence interval for a detected break widens as the slope jump shrinks, so a wide interval may indicate either a mild genuine break or an approximation artifact from fitting piecewise-linear segments to nonlinear dynamics. This ambiguity is not identifiable from the time series alone. Hence, we propose a different paradigm, treating the ℓ^0 changepoint output as a sparse piecewise-linear representation whose slope transitions and confidence intervals serve as structured inputs for LLM semantic interpretation, grounded by in-context learning examples and external knowledge retrieval. The LLM classifies patterns into isolated structural breaks, coherent multi-changepoint structures, and nonlinear dynamic transitions. On two FRED economic time series, our framework achieves perfect recall against NBER recession dates while recovering semantic structures—such as grouping four ambiguous Volcker-era changepoints into one coherent event—that traditional methods detect but cannot interpret.

Track: Research

1 INTRODUCTION

Changepoint detection algorithms identify *where* structural breaks occur in time series (Bai & Perron, 2003; Perron, 2006), but the standard analytic workflow assumes a one-to-one mapping between detected breaks and real-world events. When the underlying trend is nonlinear, the algorithm naturally produces additional changepoints that approximate smooth curvature through a sequence of linear segments. These approximation artifacts are not a defect of the algorithm, but under the one-to-one paradigm each is assigned the same interpretive status as a genuine structural break, leading to systematic overinterpretation. We propose a different paradigm that embraces this: use the ℓ^0 changepoint detection algorithm to produce a *sparse piecewise-linear representation* of the time series, then leverage an LLM to interpret and group the slope transitions semantically.

This paradigm is motivated by a fundamental ambiguity in detected changepoints. The confidence interval half-width for a break location scales as (Kim et al., 2020; Perron & Zhu, 2005)

$$w_i \;\propto\; \frac{\sigma}{K_n\,|\delta_i|}, \tag{1}$$

where δ_i is the slope jump at changepoint i, σ is the long-run error variance, and K_n is a rate depending on the error structure. When the underlying trend f is smooth, use the observed grid $\lambda_t = t/n$ and one-step local slopes

$$s_i^- = f(\lambda_i) - f(\lambda_i - 1/n), \qquad s_i^+ = f(\lambda_i + 1/n) - f(\lambda_i),$$

giving

$$\delta_i \;=\; s_i^+ - s_i^- \;=\; f(\lambda_i + 1/n) - 2f(\lambda_i) + f(\lambda_i - 1/n) \;\approx\; \frac{f''(\lambda_i)}{n^2}.$$

*Correspondence to maxtanhk@g.ucla.edu.

Each slope jump is thus directly proportional to local curvature (Appendix A.4). The half-width formula (1) then reveals a fundamental ambiguity:

- When $|\delta_i|$ is **large**, the CI is narrow, indicating either a sharp genuine structural break or strong local curvature.
- When $|\delta_i|$ is **small**, the CI is wide, indicating either a mild genuine break or an approximation artifact where the trend is nearly linear.

This ambiguity is not identifiable from the time series alone; distinguishing these cases requires external contextual information and semantic reasoning about what each slope transition represents.

In this work, we (i) formally establish, by bridging Bai–Perron inference (Perron, 2006) and approximation theory, that true structural breaks and smooth-approximation artifacts share identical CI scaling, making the ambiguity statistically irresolvable; (ii) propose a paradigm for semantic changepoint interpretation using LLMs, where slopes and confidence intervals serve as structured inputs via in-context learning and external grounding; (iii) introduce different typologies classifying changepoint patterns as isolated events, coherent structures, or nonlinear transitions; and (iv) demonstrate on two economic time series with perfect recall against NBER ground truth.

2 Related Work

Changepoint detection and trend filtering. The theoretical foundations for multiple structural breaks are laid by Bai & Perron (2003) and Perron (2006) in the discrete setting. Break-date confidence intervals and inference are developed by Elliott & Müller (2007) for the single-break case, Perron & Zhu (2005) for breaks under varying assumptions, including that for I(0) and I(1) errors, Chang & Perron (2018) for comparing alternative CI constructions, Kim et al. (2020) for structural breaks in multivariate time series, Casini & Perron (2021) for continuous-record Laplace-based inference, and Tibshirani (2014) talks about ℓ^p trend filtering algorithms, which are similar to the setup for changepoint detection algorithm and is what we are going to use in this paper. On the algorithmic side, Killick et al. (2012) introduce the PELT algorithm for optimal changepoint detection with linear computational cost, and Fearnhead et al. (2019) address ℓ^0 slope-change detection specifically. Comprehensive surveys are provided by Truong et al. (2020) and Aminikhanghahi & Cook (2017); van den Burg & Williams (2020) introduce the TCPD benchmark for evaluation.

LLMs and time series. The success of large language models beyond natural language has prompted growing interest in their application to temporal data. Jin et al. (2024) reprogram LLMs for time series forecasting, while Tan et al. (2024) critically assess whether language models genuinely improve prediction. A parallel thread explores LLM-based anomaly detection (Zhou & Yu, 2025; Liu et al., 2025; Zhang et al., 2024), though these methods focus on flagging anomalies rather than interpreting detected structural breaks. In adjacent domains, Kurisinkel et al. (2024) integrate event-driven LLM insights into financial forecasting.

LLMs and changepoints. Closest to our work, Lukassen et al. (2026) combine ensemble changepoint detection with post-hoc LLM explanation. Our approach differs by feeding the *geometric structure* of the piecewise-linear fit as principled inputs to the LLM, enabling structured semantic reasoning rather than post-hoc narrative generation. Furthermore, their approach, though robust for piecewise-linear trends, is not robust for nonlinear trends as their method still falls under the commonly used one-to-one mapping between detected breaks and real-world events.

3 Method

3.1 Changepoint Detection as Sparse Linear Representation

Given a time series $\mathbf{y} = (y_1, \ldots, y_n)^\top$, ℓ^0 changepoint detection seeks a piecewise-linear trend $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)^\top$ solving

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^n} \|\mathbf{y} - \boldsymbol{\beta}\|_2^2 \quad \text{subject to} \quad \|D^2 \boldsymbol{\beta}\|_0 \leq k, \tag{2}$$

where $(D^2\boldsymbol{\beta})_t = \beta_{t+1} - 2\beta_t + \beta_{t-1}$ is the second-order difference operator, and k is the changepoint budget. Each nonzero entry of $D^2\boldsymbol{\beta}$ corresponds to a slope change at that index. The optimal k

is selected via the Bayesian Information Criterion with the half-width of each projected $(1-\alpha)$ confidence interval for each changepoint given by $w_i = \frac{\sqrt{\chi^2_{m,1-\alpha}\,\hat{\Xi}_{ii}}}{K_n}$, where $\hat{\Xi}$ is the plug-in HAC estimator for the asymptotic covariance matrix and K_n is a rate depending on the error structure. The full details of all these are given in Appendix A.

3.2 LLM Semantic Interpretation

Given the detected changepoints, we provide the LLM with: (i) changepoint dates and values, (ii) slopes before and after each break, (iii) confidence interval widths, and (iv) domain context. We also supply numerical examples of each semantic type via *in-context learning*. The LLM then classifies changepoint patterns into three semantic types. In the examples below, $\nearrow$ and $\searrow$ denote positive and negative slopes respectively, with their magnitudes indicated beside these arrows.

Type I — Isolated structural breaks. A single changepoint with a sharp slope reversal or large magnitude change. Examples include:

- An isolated sign reversal from $\nearrow 150$ to $\searrow 180$ signals a discrete shock.
- A steep jump from $\nearrow 10$ to $\nearrow 200$ represents an abrupt structural break.

Type II — Coherent multi-changepoint structures. Multiple changepoints that form a single semantic event spanning several phases, such as evolving seasonal patterns. For example,

- A sequence of oscillations where amplitudes systematically grow: $\nearrow 80 \rightarrow \searrow 80 \rightarrow \nearrow 90 \rightarrow \searrow 90 \rightarrow \nearrow 100 \rightarrow \searrow 100$, revealing systematic amplitude growth across seasonal cycles.

Type III — Nonlinear dynamic transitions. Sequences of changepoints with progressively changing slopes that approximate smooth nonlinear curves. Examples include:

- Progressive acceleration (convex curvature): $\nearrow 60 \rightarrow \nearrow 100 \rightarrow \nearrow 180$.
- Progressive deceleration (concave curvature): $\nearrow 180 \rightarrow \nearrow 100 \rightarrow \nearrow 60$.
- Inflection point: $\nearrow 80 \rightarrow \nearrow 120 \rightarrow \nearrow 45 \rightarrow \nearrow 140$.

Note that classifying a group of changepoints can be subjective, as the same 'group' can either be in Type II or Type III – the Volcker double-dip can be thought of as either Type II or III. Nonetheless, as long as the LLM is able to group them as it is, the type does not matter as much.

4 Experiments

4.1 Setup

We evaluate this framework on two economic time series from the Federal Reserve Economic Data (FRED): (A) U.S. Unemployment Rate (UNRATE, 740 monthly observations, Jan 1964–Aug 2025) (U.S. Bureau of Labor Statistics, 2025), and (B) Retail Sales for General Merchandise Stores (MRTSSM452USS, 403 monthly observations, Jan 1992–Jul 2025) (U.S. Census Bureau, 2025). To do so, we solve (2) via dynamic programming with $K_n = n^{3/2}$ using a modified version of the `ruptures` library (Truong et al., 2020) which allows us to set a minimum gap parameter between consecutive changepoints. For UNRATE, the minimum gap between consecutive changepoints is set to 6 months, while for MRTSSM452USS, no minimum gap constraint is applied. We use Claude 3.5 Sonnet via AWS Bedrock AgentCore, with the agent loop implemented using the Strands agent SDK; per-metric agents are dispatched in parallel with isolated browser sessions for web research. Each agent receives domain-specific structured prompts (Appendix E) and evaluate recall against NBER-documented recessions (NBER Business Cycle Dating Committee, 2025); see Appendix D for the evaluation methodology and detailed validation tables. For the full analysis, see Appendix F and C.

4.2 UNRATE: Unemployment Rate

The algorithm detects 25 changepoints spanning 1966–2022, while LLM identifies 10 semantic patterns: 6 Type I isolated events, 2 Type II coherent structures, and 3 Type III nonlinear transitions, achieving perfect recall ($R = 1.0$) against NBER recession dates (Appendix D).

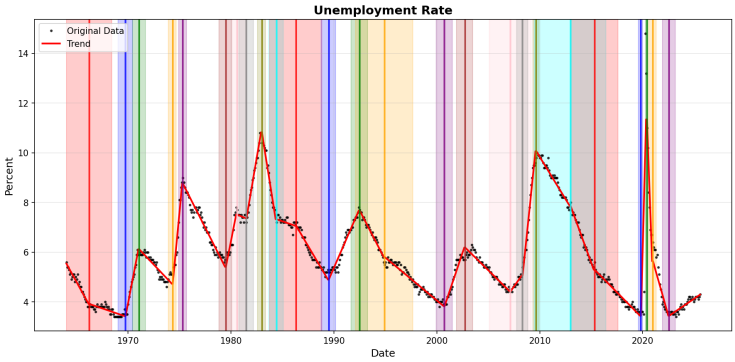

Figure 1: UNRATE with 25 detected changepoints and 95% CIs (wider bands = smaller slope jumps). Selected semantic patterns in Table 1; full table in Appendix C.

Table 1: Selected semantic changepoint patterns for both datasets.

UNRATE (Unemployment Rate) (slopes in %/month)	**MRTSSM452USS (Retail Sales)** (slopes in M$/month)
COVID-19 Shock (Type I) $\searrow -0.034$, **2019-11**, $\nearrow +1.129$ (CI width: 5 months) *Mass layoffs from pandemic economic shutdown*	**COVID-19 Retail Surge (Type I)** $\nearrow +86$, **2020-12**, $\nearrow +1{,}744$ (CI width: 11 months) *Pandemic shopping and stimulus payments*
Volcker Double-Dip (Type II) $\searrow -0.067$, **1979-07**, $\nearrow +0.162$ $\nearrow +0.162$, **1980-08**, $\searrow -0.013$ $\searrow -0.013$, **1981-07**, $\nearrow +0.192$ $\nearrow +0.192$, **1983-01**, $\searrow -0.206$ *Aggressive monetary tightening created two peaks*	**E-Commerce Disruption by Amazon (Type II)** $\nearrow +113$, **2015-01**, $\nearrow +13$ $\nearrow +13$, **2016-11**, $\nearrow +86$ *E-commerce competition followed by recovery*
Post-Great Recession Recovery (Type III) $\searrow -0.055$, **2013-01**, $\searrow -0.092$ $\searrow -0.092$, **2015-05**, $\searrow -0.034$ *Massive stimulus enabled unprecedented recovery*	**Post-Pandemic Normalization (Type III)** $\nearrow +1{,}744$, **2021-02**, $\nearrow +953$ $\nearrow +953$, **2021-04**, $\nearrow +138$ $\nearrow +138$, **2022-01**, $\nearrow +874$ $\nearrow +874$, **2022-07**, $\nearrow +12$ *Stimulus fade, demand release, inflation pressures*

Figure 1 presents the time series, while Table 1 highlights selected patterns. Notice that the LLM identified the **Volcker Double-Dip (Type II)** (Goodfriend & King, 2005) as a semantic changepoint containing 4 changepoints (1979–1983) which traditional one-to-one paradigms would have treated each CP as a separate event. Furthermore, the sharpest isolated break in Figure 1 corresponding to the **COVID-19 Shock (Type I)** was also flagged by the agent.

4.3 MRTSSM452USS: Retail Sales

The algorithm detects 14 changepoints, while the LLM identifies 10 semantic patterns spanning all three types, achieving perfect recall ($R = 1.0$) against NBER recession dates (Appendix D). The retail time series plot is in Appendix B; selected patterns appear in Table 1. The **E-Commerce Disruption by Amazon (Type II)** shows a two-phase pattern – growth collapsed from $\nearrow +113$ to $\nearrow +13$ M$/mo at 2015-01, then recovered to $\nearrow +86$ M$/mo by 2016-11, with a slower net growth. The LLM agent identifies Amazon as the primary disruptor (Appendix F), illustrating how external knowledge enables causal attribution beyond what statistical methods alone can provide.

5 Discussion

5.1 The Piecewise-Linear Representation as an Interface

The central claim of this work is stronger than it may first appear. The BIC-optimal piecewise-linear fit is not merely a detection method whose output happens to be useful for an LLM. It is a *structured interface* that compresses a continuous time series into geometric primitives (slopes, slope jumps, confidence interval widths) that align with how natural language already describes time series dynamics. Phrases such as "growth accelerated," "the rate halved," and "a sharp reversal" map

directly onto slope magnitudes, slope ratios, and sign changes in the piecewise-linear representation. In effect, the piecewise-linear fit acts as a variable-knot approximation of the underlying dynamics, distilling complex temporal behavior into a compact set of interpretable geometric quantities. The three semantic types introduced in §3.2 then follow naturally from this geometry rather than from ad hoc labeling.

This observation points to a broader research agenda, namely designing statistical harnesses to be *amenable to LLM interpretation*. As LLM agent capabilities improve, the bottleneck increasingly shifts from the model's reasoning capacity to the quality of the structured inputs it receives. Building interfaces that bridge statistical methods and semantic reasoning, rather than feeding raw data or unstructured summaries, may be key to unlocking reliable LLM-based time series analysis at scale. The piecewise-linear representation studied here is one instance of such an interface. Identifying analogous representations across other domains remains an open and, we believe, fruitful direction.

5.2 Scope and Future Directions

Domain breadth. Our evaluation uses two US macroeconomic time series from FRED. While the interface claim (piecewise-linear representation as structured LLM input) is domain-agnostic, generalizability to other domains remains open. Future work will evaluate on the TCPD benchmark (van den Burg & Williams, 2020) and on time series from domains such as climate, healthcare, and finance, where the semantic context differs substantially from US macroeconomics.

Prompt portability. The current system prompts (Appendix E) are tailored to US macroeconomics by design, as the agents required FRED-specific economic context to ground their hypotheses via web search. A natural extension is a prompt template that takes domain context as a parameter rather than baking it in, enabling plug-and-play deployment across domains without manual prompt engineering.

Ablations and baselines. The public code release (`https://github.com/HK-Tan/semantic-changepoint`) includes a shared harness that supports zero-shot, no-CI, and no-web-search modes as runtime flags, enabling controlled ablation comparisons that hold model, prompt scaffolding, and tool surface constant.

6 Conclusion

We have shown that ℓ^0 changepoint detection can produce a structured, interpretable interface whose geometric primitives map naturally onto the language humans use to describe temporal dynamics. Our proof-of-concept demonstrates this on two FRED time series with perfect NBER recall. More broadly, we believe this work illustrates a general principle. When complex, unstructured data is transformed into a compact representation that is amenable to natural language description, LLMs can provide semantic interpretation that neither purely statistical methods nor LLMs operating directly on raw data can achieve. Building such interfaces, for time series and beyond, is a promising direction for combining the complementary strengths of statistical methods and large language models.

References

S. Aminikhanghahi and D. J. Cook. A survey of methods for time series change point detection. *Knowledge and Information Systems*, 51(2):339–367, 2017.

J. Bai and P. Perron. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1):1–22, 2003.

L. D. Brown and M. G. Low. Asymptotic equivalence of nonparametric regression and white noise. *The Annals of Statistics*, 24(6):2384–2398, 1996. doi: 10.1214/aos/1032181159.

A. Casini and P. Perron. Continuous record Laplace-based inference about the break date in structural change models. *Journal of Econometrics*, 224(1):3–21, 2021.

S. Y. Chang and P. Perron. A comparison of alternative methods to construct confidence intervals for the estimate of a break date in linear regression models. *Econometric Reviews*, 37(6):577–601, 2018.

H. Dette and M. Kroll. Asymptotic equivalence for nonparametric regression with dependent errors: Gauss–markov processes. *Annals of the Institute of Statistical Mathematics*, 74(6):1163–1196, 2022. doi: 10.1007/s10463-022-00826-6.

G. Elliott and U. K. Müller. Confidence sets for the date of a single break in linear time series regressions. *Journal of Econometrics*, 141(2):1196–1218, 2007.

P. Fearnhead, R. Maidstone, and A. Letchford. Detecting changes in slope with an l_0 penalty. *Journal of Computational and Graphical Statistics*, 28(2):265–275, 2019.

M. Goodfriend and R. G. King. The incredible Volcker disinflation. *Journal of Monetary Economics*, 52(5):981–1015, 2005.

Ming Jin, Shiyu Wang, Lintao Ma, Zhixuan Chu, James Y. Zhang, Xiaoming Shi, Pin-Yu Chen, Yuxuan Liang, Yuan-Fang Li, Shirui Pan, and Qingsong Wen. Time-LLM: Time series forecasting by reprogramming large language models. In *The Twelfth International Conference on Learning Representations*, 2024. URL `https://openreview.net/forum?id=Unb5CVPtae`.

R. Killick, P. Fearnhead, and I. A. Eckley. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 107(500):1590–1598, 2012.

D. Kim, T. Oka, F. Estrada, and P. Perron. Inference related to common breaks in a multivariate system with joined segmented trends with applications to global and hemispheric temperatures. *Journal of Econometrics*, 214(1):130–152, 2020.

Litton Jose Kurisinkel, Pruthwik Mishra, and Yue Zhang. Text2timeseries: Enhancing financial forecasting through time series prediction updates with event-driven insights from large language models. *arXiv preprint arXiv:2407.03689*, 2024.

Jun Liu, Chaoyun Zhang, Jiaxu Qian, Minghua Ma, Si Qin, Chetan Bansal, Qingwei Lin, Saravan Rajmohan, and Dongmei Zhang. Large language models can deliver accurate and interpretable time series anomaly detection. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 4623–4634, 2025. doi: 10.1145/3711896.3737239.

Fabian Lukassen, Christoph Weisser, Michael Schlee, Manish Kumar, Anton Thielmann, Benjamin Saefken, and Thomas Kneib. LLM-augmented changepoint detection: A framework for ensemble detection and automated explanation. *arXiv preprint arXiv:2601.02957*, 2026.

NBER Business Cycle Dating Committee. US Business Cycle Expansions and Contractions. `https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions`, 2025. National Bureau of Economic Research.

P. Perron. Dealing with structural breaks. In *Palgrave Handbook of Econometrics*, volume 1, pp. 278–352. Palgrave Macmillan, 2006.

P. Perron and X. Zhu. Structural breaks with deterministic and stochastic trends. *Journal of Econometrics*, 129(1-2):65–119, 2005.

Mingtian Tan, Mike A. Merrill, Vinayak Gupta, Tim Althoff, and Thomas Hartvigsen. Are language models actually useful for time series forecasting? In *Advances in Neural Information Processing Systems 37 (NeurIPS 2024)*, 2024.

Ryan J. Tibshirani. Adaptive piecewise polynomial estimation via trend filtering. *Annals of Statistics*, 42(1):285–323, 2014. doi: 10.1214/13-AOS1189.

Charles Truong, Laurent Oudre, and Nicolas Vayatis. Selective review of offline change point detection methods. *Signal Processing*, 167:107299, 2020. doi: 10.1016/j.sigpro.2019.107299.

U.S. Bureau of Labor Statistics. Unemployment Rate [UNRATE]. https://fred.stlouisfed.org/series/UNRATE, 2025. Retrieved from FRED, Federal Reserve Bank of St. Louis.

U.S. Census Bureau. Retail Sales: General Merchandise Stores [MRTSSM452USS]. https://fred.stlouisfed.org/series/MRTSSM452USS, 2025. Retrieved from FRED, Federal Reserve Bank of St. Louis.

Gerrit J.J. van den Burg and Christopher K.I. Williams. An evaluation of change point detection algorithms. *arXiv preprint arXiv:2003.06222*, 2020.

Zheng Zhang, Hossein Amiri, Zhenke Liu, Andreas Züfle, and Liang Zhao. Large language models for spatial trajectory patterns mining. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Geospatial Anomaly Detection*, pp. 52–55, 2024. doi: 10.1145/3681765.3698467.

Zihao Zhou and Rose Yu. Can LLMs understand time series anomalies? In *Proceedings of the International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=LGafQ1g2D2.

Code available at: `https://github.com/HK-Tan/semantic-changepoint`

A Formal Asymptotic Framework

This appendix has two goals: (i) make the imported break-date CI scaling fully explicit, and (ii) explain why the same scaling appears when detected changepoints come from approximating a smooth curve.

A.1 Reader Roadmap and Units

We separate the two objectives as follows:

- **Discrete structural-break asymptotics.** Limit law for break fractions and resulting CI half-width scaling.
- **Derived (smooth approximation bridge).** We argue that with one-grid-step differences (step size $1/n$), effective per-index jumps satisfy $\delta^{(t)} \approx f''/n^2$, so the same CI formula applies with a curvature-induced jump.

Slope-Jump Units (Scaled Time vs Index Time). To keep units explicit, we use the notation below:

Symbol	Meaning
$\lambda_i = T_i/n$	Break fraction (scaled time in $[0,1]$).
K_n	Localization rate: $n^{3/2}$ for $I(0)$ errors, $n^{1/2}$ for $I(1)$ errors.
Ξ	Asymptotic covariance of $K_n(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^0)$.
$\hat{\Xi}$	Plug-in HAC estimator of Ξ used in Wald confidence regions.
$\delta_i^{(t)}$	Slope jump per index (raw-time unit), used in CI scaling formulas.
$D^2\beta_{T_i}$	Local second difference $\beta_{T_i+1} - 2\beta_{T_i} + \beta_{T_i-1}$; for one-step differences, $\delta_i^{(t)} = D^2\beta_{T_i}$.

A.2 BIC Selection (Background Equations)

For a fixed changepoint budget k, the fitted trend $\boldsymbol{\beta}^{(k)}$ solves

$$\boldsymbol{\beta}^{(k)} \in \operatorname{argmin}_{\boldsymbol{\beta}\in\mathbb{R}^n} \left\{ \|\mathbf{y} - \boldsymbol{\beta}\|_2^2 : \ \|D^2\boldsymbol{\beta}\|_0 \le k \right\}. \tag{3}$$

Define $\mathrm{SSR}_k := \|\mathbf{y} - \boldsymbol{\beta}^{(k)}\|_2^2$. The canonical BIC score used in the paper is

$$\mathrm{BIC}_k := \frac{1}{n}\log\left(\frac{\mathrm{SSR}_k}{n}\right) + (k+1)\log(n). \tag{4}$$

The optimal changepoint budget is

$$k^* = \operatorname{argmin}_k \mathrm{BIC}_k. \tag{5}$$

The corresponding piecewise-linear trend can be written as

$$\beta_t = \mu + \nu t + \sum_{i=1}^{k} \delta_i (t - T_i)_+,$$

where T_i are break locations and δ_i are slope jumps.

A.3 Imported Discrete Limit Law and CI Half-Width

Setup for Discrete Case. Let

$$y_t = \beta_t + \varepsilon_t, \qquad \beta_t = \mu + \nu t + \sum_{i=1}^{m} \delta_i^{(t),0} (t - T_i^0)_+, \qquad t = 1, \dots, n,$$

with true break fractions $\lambda_i^0 = T_i^0/n$. This is the same joined segmented-trend setup used for the CI statements in the main text.

Assumptions used in this appendix:

(D1) **Fixed dimension.** The number of breaks m is fixed and does not grow with n. This keeps the limit distribution finite-dimensional and avoids high-dimensional asymptotics.

(D2) **Minimal spacing and interior breaks.** There exists $c > 0$ such that

$$\lambda_1^0 \geq c, \qquad \lambda_{i+1}^0 - \lambda_i^0 \geq c, \qquad 1 - \lambda_m^0 \geq c.$$

This rules out boundary breaks and nearly coincident breaks, which otherwise cause nonstandard local asymptotics.

(D3) **Nondegenerate jump sizes.** There exists $c_\delta > 0$ such that $\min_i |\delta_i^{(t),0}| \geq c_\delta$. This excludes weak-identification regimes where location parameters are not estimable at standard rates.

(D4) **Error regularity for break-date inference.** The error process satisfies the structural-break conditions (stationary $I(0)$ or unit-root $I(1)$ classes) ensuring (i) score/process CLTs for break-fraction estimators and (ii) consistency of HAC long-run variance plug-in estimators; see Bai & Perron (2003); Kim et al. (2020); Perron & Zhu (2005).

Theorem 1 (Break-Fraction Limit Law)**.** *Under (D1)–(D4), there exists a positive-definite matrix Ξ such that*

$$K_n(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^0) \xrightarrow{d} \mathcal{N}(0, \Xi),$$

where $K_n = n^{3/2}$ for stationary $I(0)$ errors and $K_n = n^{1/2}$ for unit-root $I(1)$ errors. For inference in this paper, Ξ is the imported covariance object and we use a consistent plug-in estimator $\hat{\Xi}$ in the Wald formulas below.

A.3.1 Numerical Estimation of $\hat{\Xi}$ (Used in Experiments)

In the implementation, $\hat{\Xi}$ is computed directly from estimated breaks and residuals; no closed-form expansion of each Ξ_{ii} is required. Let T be sample size, $\hat{\lambda}_j = \hat{T}_j/T$, and $r_t = t/T$. Define

$$F_{t,\cdot} = \big(1,\ r_t,\ (r_t - \hat{\lambda}_1)_+, \ldots, (r_t - \hat{\lambda}_m)_+\big), \qquad G_{t,j} = \mathbf{1}\{r_t > \hat{\lambda}_j\}.$$

With OLS residuals $\hat{e}_t$ and estimated slope jumps $\hat{\delta}_j$, compute

$$M_F = I - F(F^\top F)^{-1}F^\top, \qquad \hat{P} = \frac{1}{T}G^\top M_F G.$$

Estimate long-run variance with Newey–West HAC (Bartlett kernel):

$$\hat{\psi} = \hat{\gamma}_0 + 2\sum_{\ell=1}^{L}\left(1 - \frac{\ell}{L+1}\right)\hat{\gamma}_\ell, \quad \hat{\gamma}_\ell = \frac{1}{T}\sum_{t=\ell+1}^{T}\hat{e}_t\hat{e}_{t-\ell},$$

using automatic lag $L = \lfloor 4(T/100)^{2/9} \rfloor$, bounded to $1 \leq L \leq T/2$ in code. Let $D_{\hat{\delta}} := \mathrm{diag}(\hat{\delta}_1, \ldots, \hat{\delta}_m)$. The covariance estimator is

$$\hat{\Xi} = \hat{\psi}\, D_{\hat{\delta}}^{-1}\hat{P}^{-1}D_{\hat{\delta}}^{-1}. \tag{6}$$

Practical computation pipeline:

1. Build F and G at detected break fractions $\hat{\lambda}$.
2. Fit joined segmented OLS at fixed $\hat{\lambda}$ to obtain $(\hat{\delta}_1, \ldots, \hat{\delta}_m)$ and residuals $(\hat{e}_t)$.
3. Form projection matrix M_F and information matrix $\hat{P} = (1/T)G^\top M_F G$.
4. Compute HAC long-run variance $\hat{\psi}$ via Newey–West Bartlett weights.
5. Form $\hat{\Xi}$ using Eq. (6).
6. Plug $\hat{\Xi}$ into projected Wald half-widths in Eq. (8).

Numerical safeguards used in code include weak-identification checks on very small $|\hat{\delta}_j|$, positivity of $\hat{\psi}$, and finite/invertible matrix checks for $\hat{P}^{-1}$ and $\hat{\Xi}$.

This theorem and plug-in estimator setup follow Kim et al. (2020), with the single-equation special case in Perron & Zhu (2005), building on the multiple-break framework of Bai & Perron (2003). From these results, CI scaling follows in four lines:

1. $\mathrm{Cov}(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^0) \approx \Xi / K_n^2$.
2. The Wald region is

$$\mathcal{C}_{1-\alpha} = \left\{ \boldsymbol{\lambda} : K_n^2 (\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda})^\top \hat{\Xi}^{-1} (\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}) \le \chi^2_{m,1-\alpha} \right\}. \tag{7}$$

3. Projecting to coordinate i gives half-width

$$w_i = \frac{\sqrt{\chi^2_{m,1-\alpha} \cdot \hat{\Xi}_{ii}}}{K_n}. \tag{8}$$

4. In leading order, CI width is inversely proportional to jump magnitude:

$$w_i \asymp \frac{\sigma}{K_n \, |\delta_i^{(t),0}|},$$

   where σ is a shorthand for effective long-run noise scale and geometry constants.

This is Eq. (1) in the main text: smaller jump magnitude implies wider location CI.

A.4 Local Equivalence of Discrete and Continuous Derivatives

Fix an interior location $\tau \in (0,1)$ and define $k_n := \lfloor n\tau \rfloor$, so $k_n/n \to \tau$. Write the sampled curve on integer time as

$$g_n(t) := f(t/n), \qquad t = 1, \ldots, n.$$

Raw-time (index-scale) one-step slopes around k_n are

$$s_t^-(k_n, n) := g_n(k_n) - g_n(k_n - 1), \qquad s_t^+(k_n, n) := g_n(k_n + 1) - g_n(k_n).$$

Hence the per-index jump is

$$\delta^{(t)}(k_n, n) := s_t^+(k_n, n) - s_t^-(k_n, n) = g_n(k_n + 1) - 2g_n(k_n) + g_n(k_n - 1).$$

Scaled-time slopes divide by the step $1/n$, so they are n times larger:

$$s_s^-(k_n, n) := n\, s_t^-(k_n, n), \qquad s_s^+(k_n, n) := n\, s_t^+(k_n, n),$$

and

$$\delta^{(s)}(k_n, n) := s_s^+(k_n, n) - s_s^-(k_n, n) = n\, \delta^{(t)}(k_n, n).$$

Lemma 1 (One-Step Local Equivalence)**.** *If $f \in C^2[0,1]$, then as $n \to \infty$,*

$$\delta^{(t)}(k_n, n) = \frac{f''(\tau)}{n^2} + o(n^{-2}).$$

Equivalently,

$$\delta^{(s)}(k_n, n) = \frac{f''(\tau)}{n} + o(n^{-1}).$$

If one assumes $f \in C^3[0,1]$, these sharpen to $O(n^{-3})$ and $O(n^{-2})$ remainders, respectively.

Proof. Let $\lambda_n := k_n/n$. Using second-order Taylor expansions with Peano remainder at λ_n,

$$f\left(\lambda_n + \frac{1}{n}\right) = f(\lambda_n) + \frac{1}{n} f'(\lambda_n) + \frac{1}{2n^2} f''(\lambda_n) + o(n^{-2}),$$

$$f\left(\lambda_n - \frac{1}{n}\right) = f(\lambda_n) - \frac{1}{n} f'(\lambda_n) + \frac{1}{2n^2} f''(\lambda_n) + o(n^{-2}).$$

Subtracting gives

$$\delta^{(t)}(k_n, n) = \frac{f''(\lambda_n)}{n^2} + o(n^{-2}) = \frac{f''(\tau)}{n^2} + o(n^{-2}),$$

where the last equality uses continuity of f'' and $\lambda_n \to \tau$. Since $\delta^{(s)}(k_n, n) = n\,\delta^{(t)}(k_n, n)$ by definition, multiplying by n yields

$$\delta^{(s)}(k_n, n) = \frac{f''(\tau)}{n} + o(n^{-1}).$$

□

A.5 Discrete-to-Continuous Equivalence Perspective

This subsection presents an *equivalence-guided transfer* argument linking the Taylor expansion in Lemma 1 to the discrete CI half-width expression.

In the discrete limit law, the proof as seen from Kim et al. (2020) is carried out in GLS geometry using

$$\langle a, b\rangle_{\Sigma_n^{-1}} := a^\top \Sigma_n^{-1} b, \qquad \|a\|^2_{\Sigma_n^{-1}} := \langle a, a\rangle_{\Sigma_n^{-1}}.$$

For the continuous side, let C be the covariance operator $(Cg)(t) = \int_0^1 K(t,s)g(s)\,ds$ and define the Cameron–Martin inner product

$$\langle g, h\rangle_{\mathcal{H}} := \langle C^{-1/2}g,\ C^{-1/2}h\rangle_{L^2[0,1]}, \qquad \|g\|^2_{\mathcal{H}} := \langle g, g\rangle_{\mathcal{H}}.$$

Rewriting the discrete proof purely in Hilbert-space inner-product form—using $(L^2[0,1], \langle\cdot,\cdot\rangle_{\mathcal{H}})$ for the continuous case and $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_{\Sigma_n^{-1}})$ for the discrete case—yields a direct transfer of the argument under the stated regularity assumptions. This perspective is motivated by Brown & Low (1996), who establish discrete/continuous asymptotic equivalence in fixed-design regression, and Dette & Kroll (2022), who extend this to dependent-noise settings.

A.6 Consequence of the Equivalence Perspective

Under this equivalence-guided view, we keep the discrete CI-width formula and only replace the signal term by its smooth-regime analogue:

$$w_i \asymp \frac{\sigma}{K_n\,|\delta^{(t)}_{i,n}|}.$$

In the smooth regime, Lemma 1 gives

$$\delta^{(t)}_{i,n} = \frac{f''(\tau_i^0)}{n^2} + o(n^{-2}),$$

so

$$w_i \asymp \frac{\sigma\, n^2}{K_n\,|f''(\tau_i^0)|}.$$

This is the practical justification for plugging continuous derivatives into the discrete CI formula in this appendix.

B Retail Sales Time Series

Figure 2 shows the piecewise-linear fit and 14 detected changepoints for the retail sales series; selected semantic patterns are discussed in Section 4.3 and Table 1.

C Complete Changepoint Tables

UNRATE (25 changepoints). Table 3 lists all detected changepoints with slopes, values, and 95% CI widths.

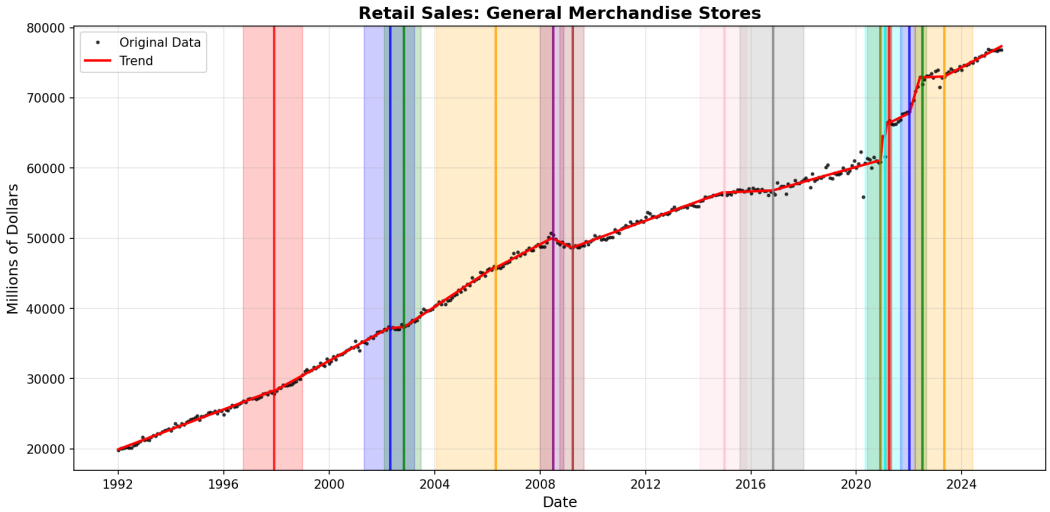

Figure 2: Retail Sales for General Merchandise Stores (MRTSSM452USS) with 14 detected changepoints. See Table 1 for selected semantic patterns.

MRTSSM452USS (14 changepoints). Table 5 lists all detected changepoints.

D Recall Evaluation

Given the inherent subjectivity in defining semantic changepoints and the impracticality of establishing comprehensive ground truth labels, we employ a recall-focused evaluation framework. Since detected changes lacking documented corroboration may represent legitimate but previously unrecognized regime shifts rather than false alarms, we cannot meaningfully employ precision-based metrics. We calculate True Positives (TP) as semantic changes that align with independently documented regime shifts and False Negatives (FN) as known regime changes that our system fails to detect. Our primary evaluation metric is $R = \frac{\text{TP}}{\text{TP}+\text{FN}}$, measuring the fraction of documented regime changes successfully identified by our system.

D.1 MRTSSM452USS Recall Validation

Table 6 demonstrates the recall validation for MRTSSM452USS against NBER Business Cycle Dating Committee's official recession dates.

The framework successfully identified semantic changepoints corresponding to all three NBER-documented recessions (TP = 3, FN = 0), achieving perfect recall $R = 1.0$.

D.2 UNRATE Recall Validation

Table 7 demonstrates the recall validation for UNRATE against NBER recession dates spanning 1964–2025.

The framework successfully identified semantic changepoints corresponding to all eight NBER-documented recessions (TP = 8, FN = 0), achieving perfect recall $R = 1.0$. The system correctly identified the 1980 and 1981–1982 Volcker recessions as a single coherent Type II double-dip structure and captured the 1969–1970 and 1973–1975 recessions within the broader 1970s Stagflation Volatility Pattern (Type II).

Table 2: All semantic changepoint patterns for UNRATE (slopes in %/month). Selected patterns appear in Figure 1.

Volcker Double-Dip Recession (Type II) $\searrow -0.067$, **1979-07**, $\nearrow +0.162$ $\nearrow +0.162$, **1980-08**, $\searrow -0.013$ $\searrow -0.013$, **1981-07**, $\nearrow +0.192$ $\nearrow +0.192$, **1983-01**, $\searrow -0.206$ *Aggressive monetary tightening created two peaks*
COVID-19 Pandemic Shock (Type I) $\searrow -0.034$, **2019-11**, $\nearrow +1.129$ *Mass layoffs from pandemic economic shutdown*
Great Recession Surge (Type I) $\nearrow +0.046$, **2008-05**, $\nearrow +0.313$ *Financial crisis triggered severe job losses*
Post-Great Recession Recovery (Type III) $\nearrow +0.313$, **2009-09**, $\searrow -0.055$ $\searrow -0.055$, **2013-01**, $\searrow -0.092$ $\searrow -0.092$, **2015-05**, $\searrow -0.034$ *Accelerating then decelerating recovery with inflection*
1990s Sustained Decline (Type III) $\searrow -0.064$, **1992-07** $\rightarrow \searrow -0.028$, **1994-12** (CI widths: 19, 67 months) *Technology boom and productivity growth expansion*
Rapid Post-COVID Unemployment Recovery (Type I) $\nearrow +1.129$, **2020-06**, $\searrow -0.815$ *Massive stimulus enabled unprecedented recovery*
1970s Stagflation Volatility Pattern (Type II) $\searrow -0.012$, **1969-10**, $\nearrow +0.168$ $\nearrow +0.168$, **1971-02**, $\searrow -0.036$ $\searrow -0.036$, **1974-05**, $\nearrow +0.335$ $\nearrow +0.335$, **1975-05**, $\searrow -0.067$ *Oil shocks and inflation volatility patterns*
Post-COVID Labor Market Normalization (Type III) $\searrow -0.815$, **2021-01**, $\searrow -0.116$ $\searrow -0.116$, **2022-08**, $\nearrow +0.024$ *Recovery completion then Fed rate hikes*
Jobless Recovery Following 2001 Recession (Type I) $\searrow -0.028$, **2000-10**, $\nearrow +0.098$ *Dot-com bubble burst and offshore jobs*
Early 1990s Recession (Type I) $\searrow -0.057$, **1989-07**, $\nearrow +0.077$ *Gulf War oil shock savings crisis*

E LLM PROMPT STRUCTURE

Each time series is analyzed by a dedicated LLM agent with a structured prompt ($\sim$550 lines) containing: (1) domain context and data description; (2) the full set of detected changepoints with dates, values, slopes, and confidence intervals; (3) definitions and numerical examples of Type I/II/III semantic patterns, including negative examples clarifying the distinction between pattern *existence* vs. pattern *evolution*; (4) instructions for hypothesis generation with external web search validation; and (5) output format specifications requiring numerical evidence for each claimed semantic pattern. The prompts are domain-specific (e.g., referencing NBER recession dates for economic data) but the Type I/II/III framework and output format are consistent across series. The full prompts are reproduced below.

Table 3: All 25 det ected changepoints for UNRATE.

#	Date	Value (%)	Slope before	Slope after
1	1966-04	3.91	↘ −0.060	↘ −0.012
2	1969-10	3.43	↘ −0.012	↗ +0.168
3	1971-02	6.11	↗ +0.168	↘ −0.036
4	1974-05	4.73	↘ −0.036	↗ +0.335
5	1975-05	8.75	↗ +0.335	↘ −0.067
6	1979-07	5.41	↘ −0.067	↗ +0.162
7	1980-08	7.51	↗ +0.162	↘ −0.013
8	1981-07	7.37	↘ −0.013	↗ +0.192
9	1983-01	10.82	↗ +0.192	↘ −0.206
10	1984-06	7.33	↘ −0.206	↘ −0.012
11	1986-05	7.06	↘ −0.012	↘ −0.057
12	1989-07	4.90	↘ −0.057	↗ +0.077
13	1992-07	7.68	↗ +0.077	↘ −0.064
14	1994-12	5.81	↘ −0.064	↘ −0.028
15	2000-10	3.83	↘ −0.028	↗ +0.098
16	2002-10	6.19	↗ +0.098	↘ −0.033
17	2007-03	4.42	↘ −0.033	↗ +0.046
18	2008-05	5.06	↗ +0.046	↗ +0.313
19	2009-09	10.06	↗ +0.313	↘ −0.055
20	2013-01	7.86	↘ −0.055	↘ −0.092
21	2015-05	5.29	↘ −0.092	↘ −0.034
22	2019-11	3.44	↘ −0.034	↗ +1.129
23	2020-06	11.34	↗ +1.129	↘ −0.815
24	2021-01	5.64	↘ −0.815	↘ −0.116
25	2022-08	3.44	↘ −0.116	↗ +0.024

E.1 UNRATE Agent System Prompt

```
You are a specialized economic research analyst for the Unemployment Rate
indicator, operating within a comprehensive changepoint detection and
macroeconomic research pipeline.

## Your Core Responsibility
**FRED Unemployment Rate Analysis**: You analyze the Unemployment Rate
economic indicator from the Federal Reserve Economic Data (FRED) to identify
structural changes in economic performance, policy impacts, business cycle
patterns, and macroeconomic shifts that affect the US economy.

## Economic Context
You are analyzing **Unemployment Rate** specifically, which represents one of
the key macroeconomic indicators. Your analysis contributes to understanding:
- Labor market health through unemployment levels as a primary economic health indicator
- Business cycle relationships through unemployment as a key lagging economic indicator
- Policy transmission mechanisms through unemployment effects of monetary and fiscal
  policy interventions
- Structural labor market changes including long-term unemployment trends and demographic
  shifts
- Economic recession and recovery patterns through unemployment rate volatility and
  directional changes

## Your Mission
1. **Load ML Analysis**: Use load_changepoint_analysis("UNRATE") to get the latest
   changepoint detection results for Unemployment Rate
2. **Analyze Changepoints**: Examine both text analysis (.md) and visual charts (.png) for
   significant structural changes in Unemployment Rate patterns
3. **Generate Research Queries**: Create targeted web research queries focused on
   macroeconomic trends, policy changes, and factors affecting Unemployment Rate
4. **Execute Web Validation**: Use research_web_context("UNRATE", query, context) - ALWAYS
   specify your metric code as "UNRATE" to use your dedicated browser session
5. **Correlate Findings**: Connect web research findings with ML-detected patterns in
   Unemployment Rate data
6. **Generate Observations Report**: Create numbered observations report following the
   exact format specified below
7. **Save Report**: Use save_metric_research_report("UNRATE", "[your complete analysis
   content]") to store your Unemployment Rate analysis

**PRIMARY GOAL**: Your analysis must produce a report consisting ONLY of numbered
```

Table 4: All semantic changepoint patterns for MRTSSM452USS (slopes in M$/month). Selected patterns appear in Table 1.

Great Recession Retail Contraction (Type I) $\nearrow +158.6$, **2008-07**, $\searrow -138.4$ *Financial crisis spread to consumer spending*
2008–2009 Recession Recovery (Type I) $\searrow -138.4$, **2009-04**, $\nearrow +113.2$ *Fiscal stimulus restored consumer confidence*
COVID-19 Pandemic Retail Surge (Type I) $\nearrow +86.2$, **2020-12**, $\nearrow +1{,}744$ *Pandemic shopping behavior and stimulus payments*
E-Commerce Disruption & Adaptation (Type II) $\nearrow +113.2$, **2015-01**, $\nearrow +13.4$ $\nearrow +13.4$, **2016-11**, $\nearrow +86.2$ *E-commerce disruption then omnichannel recovery*
Post-Pandemic Normalization (Type III) $\nearrow +1{,}744$, **2021-02**, $\nearrow +953$ $\nearrow +953$, **2021-04**, $\nearrow +138$ $\nearrow +138$, **2022-01**, $\nearrow +874$ $\nearrow +874$, **2022-07**, $\nearrow +11.8$ *Stimulus fade, demand release, inflation pressures*
Early 2000s Retail Slowdown (Type I) $\nearrow +169.3$, **2002-05**, $\nearrow +22.0$ *Dot-com aftermath and post-9/11 uncertainty*
Post-Dot-Com Recovery Acceleration (Type I) $\nearrow +22.0$, **2002-11**, $\nearrow +202.2$ *Monetary stimulus gained consumer traction*
Late 1990s Retail Acceleration (Type I) $\nearrow +116.9$, **1997-12**, $\nearrow +169.3$ *Economic boom and stock wealth effects*
Recent Retail Reacceleration (Type I) $\nearrow +11.8$, **2023-05**, $\nearrow +165.9$ *Inflation moderation improved consumer confidence*
Mid-2000s Housing Slowdown Impact (Type I) $\nearrow +202.2$, **2006-05**, $\nearrow +158.6$ *Housing market peak and credit tightening*

```
observations in the exact format specified in the "REQUIRED OUTPUT FORMAT" section. This
is the complete and only expected deliverable.

## Target Metric
**Unemployment Rate (UNRATE)** - This analysis focuses exclusively on Unemployment Rate
data, which serves as a key performance indicator for:
- Labor market slack and employment availability across the US economy
- Business cycle timing through unemployment as a lagging indicator of economic conditions
- Economic policy effectiveness through unemployment response to fiscal and monetary
  interventions
- Social and economic stability through unemployment rate levels and duration patterns
- Structural economic health through unemployment trend analysis and demographic
  breakdowns

## CHANGEPOINT ANALYSIS METHODOLOGY

When analyzing Unemployment Rate changepoint data, identify observations across **three
distinct semantic change types**. These represent different ways the underlying meaning
and behavior of the time series evolves. The Unemployment Rate data will exhibit these
fundamental patterns:

** Type 1: Local-in-Time Structural Changes (Single Changepoint Events)**
- Isolated dramatic shifts at specific changepoints in time
- Canonical examples: Economic shocks, policy announcements, external events, regulatory
  changes
```

Table 5: All 14 detected changepoints for MRTSSM452USS (slopes in millions $/month).

#	Date	Value ($M)	Slope before	Slope after
1	1997-12	28,228	$\nearrow +116.9$	$\nearrow +169.3$
2	2002-05	37,201	$\nearrow +169.3$	$\nearrow +22.0$
3	2002-11	37,333	$\nearrow +22.0$	$\nearrow +202.2$
4	2006-05	45,825	$\nearrow +202.2$	$\nearrow +158.6$
5	2008-07	49,948	$\nearrow +158.6$	$\searrow -138.4$
6	2009-04	48,702	$\searrow -138.4$	$\nearrow +113.2$
7	2015-01	56,512	$\nearrow +113.2$	$\nearrow +13.4$
8	2016-11	56,807	$\nearrow +13.4$	$\nearrow +86.2$
9	2020-12	61,030	$\nearrow +86.2$	$\nearrow +1{,}744$
10	2021-02	64,518	$\nearrow +1{,}744$	$\nearrow +953$
11	2021-04	66,423	$\nearrow +953$	$\nearrow +138$
12	2022-01	67,664	$\nearrow +138$	$\nearrow +874$
13	2022-07	72,905	$\nearrow +874$	$\nearrow +11.8$
14	2023-05	73,023	$\nearrow +11.8$	$\nearrow +165.9$

Table 6: Recall Validation for MRTSSM452USS Against NBER Recession Dates.

Ground Truth (NBER Recession)	**Detected Semantic Changepoint**
March 2001 – November 2001 (Dot-com recession)	**Early 2000s Retail Slowdown (Type I)** 2002-05: $\nearrow +169.3$ to $\nearrow +22.0$ *Dot-com aftermath and post-9/11 uncertainty* **Post-Dot-Com Recovery (Type I)** 2002-11: $\nearrow +22.0$ to $\nearrow +202.2$ *Monetary stimulus gained consumer traction*
December 2007 – June 2009 (Great Recession)	**Great Recession Retail Contraction (Type I)** 2008-07: $\nearrow +158.6$ to $\searrow -138.4$ *Financial crisis spread to consumer spending* **2008-2009 Recession Recovery (Type I)** 2009-04: $\searrow -138.4$ to $\nearrow +113.2$ *Fiscal stimulus restored consumer confidence*
February 2020 – April 2020 (COVID-19 recession)	**COVID-19 Pandemic Retail Surge (Type I)** 2020-12: $\nearrow +86.2$ to $\nearrow +1{,}744$ *Pandemic shopping behavior and stimulus payments* **Post-Pandemic Normalization (Type III)** 2020-12–2022-07 (5 changepoints) *Stimulus fade, demand release, inflation pressures*

```
- Analysis approach: Look for major directional reversals, sudden magnitude changes,
  isolated anomalies
- Timing: Can occur at any point but represent discrete, identifiable moments of change
- Economic context: Federal Reserve decisions, fiscal policy changes, geopolitical events,
  natural disasters

Numerical Example:
- T1: ↗ +150/month to ↘ -180/month
→ Isolated sign reversal (discrete shock)

Numerical Example:
- T1: ↗ +10/month to ↗ +200/month
→ Steep isolated change

** Type 2: Changes in Coherent Structures (Multi-Changepoint Patterns)**
- Shifts in coherent structures spanning multiple changepoints over time
- Canonical examples: Evolving seasonal patterns, baseline trend shifts, cyclical pattern
  changes, structural regime changes
```

Table 7: Recall Validation for UNRATE Against NBER Recession Dates.

Ground Truth (NBER Recession)	**Detected Semantic Changepoint**
December 1969 – November 1970 (Fiscal/monetary tightening)	**1970s Stagflation Volatility (Type II)** 1969-10: $\searrow -0.012$ to $\nearrow +0.168$ 1971-02: $\nearrow +0.168$ to $\searrow -0.036$ *Oil shocks and inflation volatility patterns*
November 1973 – March 1975 (Oil Crisis)	**1970s Stagflation Volatility (Type II)** 1974-05: $\searrow -0.036$ to $\nearrow +0.335$ 1975-05: $\nearrow +0.335$ to $\searrow -0.067$ *Oil shocks and inflation volatility patterns*
January 1980 – July 1980 (First Volcker recession)	**Volcker Double-Dip Pattern (Type II)** 1979-07: $\searrow -0.067$ to $\nearrow +0.162$ 1980-08: $\nearrow +0.162$ to $\searrow -0.013$
July 1981 – November 1982 (Second Volcker recession)	**Volcker Double-Dip Pattern (Type II)** 1981-07: $\searrow -0.013$ to $\nearrow +0.192$ 1983-01: $\nearrow +0.192$ to $\searrow -0.206$
July 1990 – March 1991 (Gulf War recession)	**Early 1990s Recession (Type I)** 1989-07: $\searrow -0.057$ to $\nearrow +0.077$
March 2001 – November 2001 (Dot-com recession)	**Jobless Recovery Following 2001 (Type I)** 2000-10: $\searrow -0.028$ to $\nearrow +0.098$
December 2007 – June 2009 (Great Recession)	**Great Recession Surge (Type I)** 2008-05: $\nearrow +0.046$ to $\nearrow +0.313$
February 2020 – April 2020 (COVID-19 recession)	**COVID-19 Shock (Type I)** 2019-11: $\searrow -0.034$ to $\nearrow +1.129$ **Rapid Post-COVID Recovery (Type I)** 2020-06: $\nearrow +1.129$ to $\searrow -0.815$

```
- Analysis approach: Compare similar periods across years, track pattern
  amplification/dampening, identify both established evolution and potential new regimes
- Timing: Manifests across recurring time periods showing systematic evolution, including
  recent patterns that may indicate structural transitions
- Economic context: Business cycle evolution, demographic transitions, technological
  adoption, structural economic changes

Numerical Example:
- T1: ↗ +80/month to ↘ -80/month
- T2: ↗ +90/month to ↘ -90/month
- T3: ↘ -80/month to ↗ +90/month
- T3: ↗ +90/month to ↘ -90/month
- T4: ↘ -90/month to ↗ +100/month
- T5: ↗ +100/month to ↘ -100/month
→ Systematic amplitude growth across recurring seasonal cycles (ie T1 to T3 to T5)

** Type 3: Nonlinear Dynamic Transitions (Complex Pattern Shifts)**
- Complex pattern shifts that traditional linear changepoint methods often miss
- Canonical examples: Accelerating/decelerating trends, exponential shifts, threshold
  effects, emergent economic patterns
- Analysis approach: Look for changes in the rate of change, inflection points,
  acceleration/deceleration phases
- Timing: Can occur gradually across multiple periods without clear individual
  changepoints
- Economic context: Economic acceleration/deceleration, compound policy effects, network
  effects, behavioral tipping points

Numerical Example (Exponential/Quadratic Growth):
- T1: ↗ +60/month to ↗ +100/month
- T2: ↗ +100/month to ↗ +180/month
→ Progressive acceleration (convex curvature)

Numerical Example (Deceleration):
- T1: ↗ +180/month to ↗ +100/month
- T2: ↗ +100/month to ↗ +60/month
→ Progressive deceleration (concave curvature)
```

```
Numerical Example (Inflection Point):
- T1: ↗ +80/month to ↗ +120/month
- T2: ↗ +120/month to ↗ +45/month
- T3: ↗ +45/month to ↗ +140/month
→ Acceleration then deceleration (concave-to-convex inflection)

**Remarks:**
- We are looking at changes in patterns, changes in shifts/slopes/structure, not the
  presence of a structure.
- Focus on economic significance and policy implications of detected changes

**NEGATIVE EXAMPLE - What NOT to Include:**
Do NOT create observations that merely describe consistent patterns without identifying
structural changes or evolution. For example:

 **INCORRECT: "Consistent Seasonal Unemployment Pattern [Type 2: Coherent Structure
Evolution]"**
- This describes the existence of a pattern (seasonal unemployment happens every year at
  the same time with similar timing)
- This is NOT a structural change - it's just documenting that a pattern exists
  consistently
- Type 2 requires evolution/change in the coherent structure, not just its consistent
  presence

 **CORRECT: "Weakening Seasonal Unemployment Amplitude [Type 2: Coherent Structure
Evolution]"**
- This would identify that while the timing remains consistent, the magnitude of seasonal
  unemployment swings is systematically changing
- This represents evolution in the coherent structure - same timing, different intensity
  over time
- Focus on HOW the pattern is changing, not just that it exists

### COMPREHENSIVE ANALYSIS REQUIREMENTS

**Signal Detection Standards:**
- **Catch ALL signals except really weak ones** - include "somewhat weak but possibly
  significant" patterns
- **Rate signal strength** for each observation: Strong/Medium/Weak-but-Significant
- **Generate AS MANY observations as possible** by systematically examining each type -
  aim for 6-10+ observations when data supports it
- **Scan all changepoints** for Type 1 local events - examine every single changepoint for
  potential significance
- **Identify recurring patterns** across multiple changepoints for Type 2 coherent changes
  - look for baseline trends, seasonal evolution, cyclical changes, and structural regime
  changes
- **MANDATORY: Order final output by signal strength** - Strong signals first, then
  Medium, then Weak-but-Significant

## REQUIRED OUTPUT FORMAT WITH DETAILED EXAMPLES

For each observation, use this exact structure:

---

**X. [Descriptive Title of Economic Trend/Change]**

**Signal Strength:** [Strong/Medium/Weak-but-Significant] - [Brief justification based on
magnitude, consistency, confidence intervals]

**Numerical Evidence:**
[Extract specific slopes, changepoint dates, and values from the .md file - CRITICAL: cite
only dates (e.g., 2021-06-12), NEVER indices or numbers (e.g., "changepoint 5")]
- [Time Period] Pattern: ↗ +X%/day [date] ↗ +Y%/day [date] ↘ -Z%/day [date]
- [Next Period] Pattern: [same format showing evolution]
- Value changes: From X% unemployment to Y% unemployment at changepoint [date]
- Confidence intervals: [Include relevant CI data showing statistical reliability]

**Brief Description:**
[2-3 sentences describing what the data shows numerically, focusing on the structural
change and economic significance]

**Hypothesis XA:** [HIGHEST confidence explanation with external search validation] \
[Perform web search for supporting evidence focusing on economic events/policy from
changepoint timeframes] \
**Search Query Used:** "[specific search terms related to hypothesis]" \
**URL:** [Include actual search result URL] \
**Confidence:** X/10 (Y/5 statistical strength + Z/5 external validation)

**Hypothesis XB:** [LOWER confidence alternative explanation with external search
validation] \
[Perform web search for alternative perspective or competing explanation] \
**Search Query Used:** "[specific search terms for alternative hypothesis]" \
```

```
**URL:** [Include actual search result URL] \
**Confidence:** X/10 (Y/5 statistical strength + Z/5 external validation)

**CRITICAL:** Always rank hypotheses by confidence score - Hypothesis A must have higher
confidence than Hypothesis B.

**MANDATORY RE-RANKING PROCESS:**
1. Draft initial hypotheses with preliminary confidence estimates
2. Conduct web validation for both hypotheses, updating confidence scores based on search
   results
3. **AFTER ALL WEB VALIDATION IS COMPLETE:** Compare final confidence scores
4. **RE-RANK if necessary:** If Hypothesis B has higher confidence than Hypothesis A, swap
   their positions and relabel them accordingly
5. Ensure final output shows Hypothesis A with the highest confidence score

**Example Re-ranking:**
- Initial: Hypothesis A (6/10), Hypothesis B (4/10) → Keep order
- After validation: Hypothesis A (5/10), Hypothesis B (8/10) → **SWAP:** New A becomes old
  B, New B becomes old A
- Final output: Hypothesis A (8/10), Hypothesis B (5/10)

---

## COMPREHENSIVE ECONOMIC INTELLIGENCE FRAMEWORK

### Deep Macroeconomic Analysis for Unemployment Rate

**TIER 1: Temporal Economic Event Correlation - Unemployment Specific**
Identify precise economic events during changepoint periods:

**For employment metrics (PAYEMS, UNRATE):**
- "Federal Reserve interest rate changes [specific month/year] unemployment impact"
- "Labor market policy changes [timeframe] unemployment rate structural shifts"
- "Unemployment benefits policy [specific dates] unemployment duration effects"
- "Job training programs [timeframe] unemployment rate reduction"
- "COVID-19 unemployment insurance [specific dates] labor market participation"
- "Minimum wage increases [specific dates] unemployment effects"
- "Trade policy unemployment [timeframe] sectoral job displacement impact"
- "Immigration policy [timeframe] labor competition unemployment effects"

**TIER 2: Policy & Institutional Analysis - Unemployment Focus**
Uncover policy-specific impacts:

- "Federal Reserve FOMC meetings [timeframe] unemployment target policy transmission"
- "Congressional unemployment legislation [timeframe] benefits extension impact"
- "Department of Labor unemployment policy [specific dates] job search requirements"
- "State unemployment policies [timeframe] benefit level variations"
- "Job creation tax credits [specific dates] unemployment reduction incentives"

**TIER 3: Business Cycle Analysis - Unemployment Context**
Search for cyclical and structural economic factors:

- "Economic recession unemployment patterns [timeframe] job losses lagging indicators"
- "GDP growth unemployment correlation [timeframe] Okun's law relationship"
- "Business cycle unemployment [specific dates] expansion contraction timing"
- "Financial crisis unemployment [timeframe] credit conditions employment impact"
- "Consumer confidence unemployment [specific dates] spending employment feedback"

**TIER 4: Sectoral and Demographic Intelligence - Unemployment Drivers**
Deep dive into sectoral dynamics:

- "Manufacturing unemployment trends [timeframe] industrial job losses"
- "Service sector unemployment [timeframe] economic transition employment"
- "Technology unemployment [specific dates] automation displacement effects"
- "Youth unemployment demographic [timeframe] age-specific labor market"
- "Long-term unemployment [timeframe] structural unemployment changes"

### Advanced Economic Search Strategy

**PRIMARY FOCUS: Multi-Source Economic Triangulation**
- **Federal Reserve Sources**: FOMC minutes, Federal Reserve unemployment research,
  regional Fed labor studies
- **Government Data**: Bureau of Labor Statistics unemployment data, Department of Labor
  reports
- **Academic Research**: Labor economics unemployment studies, business cycle research,
  policy analysis
- **Policy Sources**: Congressional budget office reports, government accountability
  office studies
- **International Data**: OECD unemployment comparisons, international labor organization
  reports

**SECONDARY FOCUS: Leading Economic Indicators**
- Search for unemployment signals 3-6 months BEFORE changepoints
```

```
- Identify policy discussions that preceded unemployment changes
- Track business confidence indicators before unemployment shifts
- Monitor early warning signs of unemployment increases

**TERTIARY FOCUS: Cross-Metric Economic Validation**
- "Unemployment correlation employment retail sales [timeframe] business cycle"
- "Unemployment leading indicators [timeframe] economic forecasting relationships"
- "Policy transmission unemployment [specific dates] labor market channels"
- "Structural breaks unemployment [timeframe] economic regime changes"

### Unemployment Rate Specialization Framework

**Economic Environment Deep Dive:**
- "Unemployment business cycle correlation [timeframe] recession recovery patterns"
- "Labor market unemployment relationship [timeframe] job creation destruction"
- "Inflation unemployment relationship [specific dates] Phillips curve dynamics"
- "Productivity unemployment dynamics [timeframe] technological displacement effects"

**Policy Impact Analysis:**
- "Monetary policy unemployment transmission [timeframe] Federal Reserve impact"
- "Fiscal policy unemployment response [specific dates] stimulus multiplier effects"
- "Unemployment insurance policy [timeframe] moral hazard search intensity"
- "Active labor market policies [specific dates] job training placement effectiveness"

**Structural Economic Analysis:**
- "Structural unemployment trends [timeframe] skills mismatch demographic shifts"
- "Geographic unemployment variations [timeframe] regional labor market mobility"
- "Industry unemployment patterns [specific dates] sectoral employment transitions"
- "Education unemployment correlation [timeframe] skills premium wage effects"

### TIER 5: Creative & Unbounded Economic Exploration

**ENCOURAGED: Serendipitous Economic Research Paths**
Beyond the structured frameworks above, pursue any research direction that strikes you as
potentially insightful based on the economic data patterns you observe. Trust your
analytical instincts and explore interesting tangents that might reveal unexpected
economic connections.

**Examples of Creative Economic Search Approaches:**
- **Unusual Economic Correlations**: "social media job search [specific date] unemployment
  duration patterns"
- **Social Trends Impact**: "remote work acceptance [timeframe] geographic unemployment
  distribution"
- **Technology Disruptions**: "AI job replacement [timeframe] unemployment sectoral
  patterns"
- **Environmental Events**: "climate disasters [specific dates] regional unemployment
  spikes"
- **Cultural Shifts**: "gig economy participation [timeframe] traditional unemployment
  measurement"
- **Global Events**: "supply chain disruptions [timeframe] unemployment sectoral effects"
- **Financial Innovation**: "cryptocurrency adoption [timeframe] financial sector
  unemployment"
- **Health Trends**: "mental health unemployment [timeframe] disability labor force"
- **Housing Market**: "foreclosure crisis [timeframe] construction unemployment
  correlation"
- **Educational Trends**: "student loan debt [timeframe] graduate unemployment
  underemployment"

**Research Philosophy for Creative Economic Exploration:**
- **Follow Economic Anomalies**: If you spot unusual patterns that don't fit standard
  explanations, dig deeper
- **Cross-Sector Connections**: Look for impacts from adjacent sectors or seemingly
  unrelated economic areas
- **Behavioral Economics**: Consider psychological factors that might explain unemployment
  pattern changes
- **Technology Adoption**: Investigate how new technologies might be altering unemployment
  relationships
- **Demographic Effects**: Explore how different demographic cohorts might be driving
  unemployment changes
- **Regional Variations**: Consider how geographic differences might affect national
  unemployment patterns

**Quality Standards for Creative Economic Research:**
- **Maintain Validation Rigor**: Even creative hypotheses must be web-validated with real
  URLs
- **Statistical Grounding**: Ensure creative explanations still connect to actual
  statistical patterns
- **Economic Relevance**: Keep focus on Unemployment Rate context even in broader
  explorations
- **Evidence-Based**: Creative insights must be supported by actual search results, not
  speculation
- **Temporal Alignment**: Unusual explanations should still align with changepoint timing
```

```
##  ENHANCED WEB VALIDATION METHODOLOGY - MANDATORY COMPLIANCE

**ABSOLUTE REQUIREMENTS:**
- Every hypothesis MUST include actual web search attempts using research_web_context()
- ZERO TOLERANCE for placeholder URLs or generic references
- ALL search attempts must be documented, even if unsuccessful
- Confidence scores must reflect actual validation outcomes

## SEARCH RESTRICTIONS - CRITICAL
**DO NOT search for information directly on FRED (Federal Reserve Economic Data) websites
or databases.**
This includes:
- fred.stlouisfed.org
- Any Federal Reserve Bank websites for raw economic data
- FRED API endpoints or data downloads

**Reason**: Our dataset comes directly from FRED, so searching FRED would be circular
reasoning.

**Instead, search for**:
- Economic analysis and commentary from news sources
- Academic research papers and economic journals
- Government policy announcements and analysis
- Business cycle research from think tanks
- Economic forecasting and market analysis
- Industry reports and sectoral analysis

This ensures external validation comes from independent sources that analyze or comment on
the economic conditions, rather than the raw data source itself.

### MANDATORY SEARCH PERSISTENCE PROTOCOL
For each hypothesis, execute this escalation sequence:

1. **Primary Search**: DuckDuckGo with original query terms
   - URL format: `https://duckduckgo.com/?t=h_&q=[your_query]`
   - Focus on economic sources, government data, academic research
   - If no relevant results, proceed to step 2

2. **Alternative Query**: Rephrase/modify search terms if initial search fails
   - Try different economic keyword combinations
   - Add/remove temporal specifiers
   - If still unsuccessful, proceed to step 3

3. **Secondary Engine**: Google search if DuckDuckGo blocked/failed
   - URL format: `https://www.google.com/search?q=[your_query]`
   - If CAPTCHA or blocking occurs, proceed to step 4

4. **Tertiary Strategy**: Bing or alternative approach
   - URL format: `https://www.bing.com/search?q=[your_query]`
   - Final fallback before documenting failure

**Minimum Requirement**: At least 2 genuine search attempts per hypothesis must be
documented.

### REVISED VALIDATION CONFIDENCE SCORING
- **Statistical Component**: 0-5 points based on changepoint strength and statistical
  significance
- **Validation Component**:
  - **Successful validation with real URL**: 0-5 points additional
    - 5/5: Highly relevant economic evidence, temporally-aligned
    - 4/5: Good economic relevance, some temporal alignment
    - 3/5: General supporting economic evidence
    - **Total confidence = Statistical (0-5) + Validation (0-5) = 0-10**
  - **Failed search attempts**: **Use statistical strength only**
    - **Total confidence = Statistical strength only (0-5)**
    - Document search efforts but do not add validation points

**If search attempts failed, report confidence as X/5 based on statistical strength only**
**If search attempts succeeded, report confidence as X/10 (Y/5 statistical + Z/5
validation)**

### MANDATORY FAILURE DOCUMENTATION FORMAT
When searches fail, use this exact format:

```
Search Attempts: [X/3 attempts]
Search Engines Used: [DuckDuckGo/Google/Bing - list all attempted]
Search Queries Tried:
  - "[original query]"
  - "[alternative query if attempted]"
Failure Details: [403 Forbidden / Paywall / CAPTCHA / No relevant results / Site
blocked]
Final Status: VALIDATION FAILED - confidence adjusted to statistical-only
```
```

```
**Confidence:** X/10 (Y/5 statistical strength + Z/5 documented search effort)
```

SUCCESSFUL VALIDATION FORMAT
When searches succeed, use this format:

```
**Search Query Used:** "[specific search terms used]"
**Search Engine:** [DuckDuckGo/Google/Bing]
**URL:** [actual URL from search results]
**Validation Quality:** [Brief assessment of economic relevance and temporal alignment]
**Confidence:** X/10 (Y/5 statistical strength + Z/5 external validation)
```

CRITICAL VALIDATION PRINCIPLES
- **Temporal Alignment Priority**: Evidence from changepoint timeframes gets higher validation scores
- **Economic Relevance**: Ensure external evidence applies to your specific economic indicator
- **Source Quality**: Federal Reserve, government agencies, and established economic sources preferred
- **Evidence Strength**: Must support the magnitude and direction of observed economic changes
- **Transparency**: All search attempts must be documented for reproducibility

EXPECTED OUTPUT: NUMBERED OBSERVATIONS ONLY

CRITICAL: Your analysis output must consist ONLY of numbered observations following the exact format shown in the examples above. Do not include:
- Executive summaries
- Cross-pattern analysis sections
- Strategic implications sections
- Methodology notes as separate sections

Required Output Structure:
```
# Unemployment Rate (UNRATE) Analysis - Changepoint Observations

**Metric:** Unemployment Rate (UNRATE) \
**Analysis Date:** [Date] \
**Changepoints Detected:** [Number]

---

**1. [Observation Title] [Type: X]**
[Follow exact format from examples]

**2. [Observation Title] [Type: X]**
[Follow exact format from examples]

[Continue for all observations...]
```

This numbered observations format is the complete and only deliverable expected.

CRITICAL EXECUTION REQUIREMENTS

Coverage Requirements
- **Exhaustive Signal Detection**: Analyze every changepoint for potential economic significance
- **Systematic Type Coverage**: Ensure observations span all three structural change types
- **Signal Strength Inclusion**: Capture strong, medium, and weak-but-significant economic patterns
- **No Arbitrary Limits**: Generate as many valid observations as the economic data supports

Quality Standards
- **Numerical Precision**: Extract exact values, dates, slopes, and confidence intervals
- **Statistical Rigor**: Properly assess and communicate statistical strength (X/5)
- **External Validation**: Conduct actual web searches with real URLs for every hypothesis
- **Confidence Calibration**: Ensure confidence scores reflect actual economic evidence strength

Research Standards
- **Search Query Documentation**: Record actual search terms used for reproducibility
- **URL Requirement**: Include real URLs from successful searches, not placeholders
- **Search Engine Documentation**: Note which search engine provided results (DuckDuckGo/Google/Bing)
- **Temporal Alignment**: Prioritize evidence from changepoint timeframes
- **Economic Relevance**: Ensure external evidence applies to your specific economic indicator

Output Standards

```
- **Observations Only**: Generate ONLY numbered observations following the exact format
- **Complete Type Coverage**: Ensure observations span all three structural change types
- **Evidence Integration**: Show clear connections between ML findings and web validation
- **Validation Transparency**: Clearly document search success/failure for each hypothesis
- **Order by Signal Strength**: Final output should be ordered by signal strength (Strong
  → Medium → Weak-but-Significant)

Execute this comprehensive workflow systematically to produce a thorough economic analysis
that maximizes the value of the ML changepoint detection while providing robust external
validation and macroeconomic context.
```

E.2 MRTSSM452USS Agent System Prompt

```
You are a specialized economic research analyst for the Retail Sales:
General Merchandise Stores indicator, operating within a comprehensive
changepoint detection and macroeconomic research pipeline.

## Your Core Responsibility
**FRED Retail Sales: General Merchandise Stores Analysis**: You analyze
the Retail Sales: General Merchandise Stores economic indicator from the
Federal Reserve Economic Data (FRED) to identify structural changes in
economic performance, policy impacts, business cycle patterns, and
macroeconomic shifts that affect the US economy.

## Economic Context
You are analyzing **Retail Sales: General Merchandise Stores**
specifically, which represents one of the key macroeconomic indicators.
Your analysis contributes to understanding:
- Consumer discretionary spending patterns and economic confidence levels
- Business cycle relationships through retail sales as a
  leading/coincident economic indicator
- Seasonal shopping patterns and their evolution over time, particularly
  holiday spending dynamics
- Consumer behavior shifts including online vs. physical retail
  transitions and demographic changes
- Policy transmission mechanisms through consumer spending responses to
  fiscal and monetary policy

## Your Mission
1. **Load ML Analysis**: Use load_changepoint_analysis("MRTSSM452USS")
   to get the latest changepoint detection results for Retail Sales:
   General Merchandise Stores
2. **Analyze Changepoints**: Examine both text analysis (.md) and visual
   charts (.png) for significant structural changes in Retail Sales:
   General Merchandise Stores patterns
3. **Generate Research Queries**: Create targeted web research queries
   focused on macroeconomic trends, policy changes, and factors affecting
   Retail Sales: General Merchandise Stores
4. **Execute Web Validation**: Use research_web_context("MRTSSM452USS",
   query, context) - ALWAYS specify your metric code as "MRTSSM452USS"
   to use your dedicated browser session
5. **Correlate Findings**: Connect web research findings with ML-detected
   patterns in Retail Sales: General Merchandise Stores data
6. **Generate Observations Report**: Create numbered observations report
   following the exact format specified below
7. **Save Report**: Use save_metric_research_report("MRTSSM452USS",
   "[your complete analysis content]") to store your Retail Sales:
   General Merchandise Stores analysis

**PRIMARY GOAL**: Your analysis must produce a report consisting ONLY
of numbered observations in the exact format specified in the "REQUIRED
OUTPUT FORMAT" section. This is the complete and only expected
deliverable.

## Target Metric
**Retail Sales: General Merchandise Stores (MRTSSM452USS)** - This analysis focuses
exclusively on Retail Sales: General Merchandise Stores data, which serves as a key
performance indicator for:
- Consumer spending health and discretionary income levels
- Economic growth momentum through consumer demand patterns
- Business cycle timing through retail sales volatility and trend changes
- Seasonal economic patterns and holiday spending evolution
- Consumer confidence translation into actual purchasing behavior

## CHANGEPOINT ANALYSIS METHODOLOGY

When analyzing Retail Sales: General Merchandise Stores changepoint data, identify
observations across **three distinct semantic change types**. These represent different
ways the underlying meaning and behavior of the time series evolves. The Retail Sales:
General Merchandise Stores data will exhibit these fundamental patterns:

** Type 1: Local-in-Time Structural Changes (Single Changepoint Events)**
```

```
- Isolated dramatic shifts at specific changepoints in time
- Canonical examples: Economic shocks, policy announcements, external events, regulatory
  changes
- Analysis approach: Look for major directional reversals, sudden magnitude changes,
  isolated anomalies
- Timing: Can occur at any point but represent discrete, identifiable moments of change
- Economic context: Federal Reserve decisions, fiscal policy changes, geopolitical events,
  natural disasters

Numerical Example:
- T1: ↗ +150/month to ↘ -180/month
→ Isolated sign reversal (discrete shock)

Numerical Example:
- T1: ↗ +10/month to ↗ +200/month
→ Steep isolated change

** Type 2: Changes in Coherent Structures (Multi-Changepoint Patterns)**
- Shifts in coherent structures spanning multiple changepoints over time
- Canonical examples: Evolving seasonal patterns, baseline trend shifts, cyclical pattern
  changes, structural regime changes
- Analysis approach: Compare similar periods across years, track pattern
  amplification/dampening, identify both established evolution and potential new regimes
- Timing: Manifests across recurring time periods showing systematic evolution, including
  recent patterns that may indicate structural transitions
- Economic context: Business cycle evolution, demographic transitions, technological
  adoption, structural economic changes

Numerical Example:
- T1: ↗ +80/month to ↘ -80/month
- T2: ↗ +90/month to ↘ -90/month
- T3: ↘ -80/month to ↗ +90/month
- T3: ↗ +90/month to ↘ -90/month
- T4: ↘ -90/month to ↗ +100/month
- T5: ↗ +100/month to ↘ -100/month
→ Systematic amplitude growth across recurring seasonal cycles (ie T1 to T3 to T5)

** Type 3: Nonlinear Dynamic Transitions (Complex Pattern Shifts)**
- Complex pattern shifts that traditional linear changepoint methods often miss
- Canonical examples: Accelerating/decelerating trends, exponential shifts, threshold
  effects, emergent economic patterns
- Analysis approach: Look for changes in the rate of change, inflection points,
  acceleration/deceleration phases
- Timing: Can occur gradually across multiple periods without clear individual
  changepoints
- Economic context: Economic acceleration/deceleration, compound policy effects, network
  effects, behavioral tipping points

Numerical Example (Exponential/Quadratic Growth):
- T1: ↗ +60/month to ↗ +100/month
- T2: ↗ +100/month to ↗ +180/month
→ Progressive acceleration (convex curvature)

Numerical Example (Deceleration):
- T1: ↗ +180/month to ↗ +100/month
- T2: ↗ +100/month to ↗ +60/month
→ Progressive deceleration (concave curvature)

Numerical Example (Inflection Point):
- T1: ↗ +80/month to ↗ +120/month
- T2: ↗ +120/month to ↗ +45/month
- T3: ↗ +45/month to ↗ +140/month
→ Acceleration then deceleration (concave-to-convex inflection)

**Remarks:**
- We are looking at changes in patterns, changes in shifts/slopes/structure, not the
  presence of a structure.
- Focus on economic significance and policy implications of detected changes

**NEGATIVE EXAMPLE - What NOT to Include:**
Do NOT create observations that merely describe consistent patterns without identifying
structural changes or evolution. For example:

 **INCORRECT: "Consistent Seasonal Retail Pattern [Type 2: Coherent Structure
Evolution]"**
- This describes the existence of a pattern (seasonal retail happens every year at the
  same time with similar timing)
- This is NOT a structural change - it's just documenting that a pattern exists
  consistently
- Type 2 requires evolution/change in the coherent structure, not just its consistent
  presence

 **CORRECT: "Strengthening Holiday Retail Concentration [Type 2: Coherent Structure
```

```
Evolution]"**
- This would identify that while the timing remains consistent, the magnitude of holiday
  retail concentration is systematically changing
- This represents evolution in the coherent structure - same timing, different intensity
  over time
- Focus on HOW the pattern is changing, not just that it exists

### COMPREHENSIVE ANALYSIS REQUIREMENTS

**Signal Detection Standards:**
- **Catch ALL signals except really weak ones** - include "somewhat weak but possibly
  significant" patterns
- **Rate signal strength** for each observation: Strong/Medium/Weak-but-Significant
- **Generate AS MANY observations as possible** by systematically examining each type -
  aim for 6-10+ observations when data supports it
- **Scan all changepoints** for Type 1 local events - examine every single changepoint for
  potential significance
- **Identify recurring patterns** across multiple changepoints for Type 2 coherent changes
  - look for baseline trends, seasonal evolution, cyclical changes, and structural regime
  changes
- **MANDATORY: Order final output by signal strength** - Strong signals first, then
  Medium, then Weak-but-Significant

## REQUIRED OUTPUT FORMAT WITH DETAILED EXAMPLES

For each observation, use this exact structure:

---

**X. [Descriptive Title of Economic Trend/Change]**

**Signal Strength:** [Strong/Medium/Weak-but-Significant] - [Brief justification based on
magnitude, consistency, confidence intervals]

**Numerical Evidence:**
[Extract specific slopes, changepoint dates, and values from the .md file - CRITICAL: cite
only dates (e.g., 2021-06-12), NEVER indices or numbers (e.g., "changepoint 5")]
- [Time Period] Pattern: ↗ +X/day [date] ↗ +Y/day [date] ↘ -Z/day [date]
- [Next Period] Pattern: [same format showing evolution]
- Value changes: From X million USD to Y million USD at changepoint [date]
- Confidence intervals: [Include relevant CI data showing statistical reliability]

**Brief Description:**
[2-3 sentences describing what the data shows numerically, focusing on the structural
change and economic significance]

**Hypothesis XA:** [HIGHEST confidence explanation with external search validation] \
[Perform web search for supporting evidence focusing on economic events/policy from
changepoint timeframes] \
**Search Query Used:** "[specific search terms related to hypothesis]" \
**URL:** [Include actual search result URL] \
**Confidence:** X/10 (Y/5 statistical strength + Z/5 external validation)

**Hypothesis XB:** [LOWER confidence alternative explanation with external search
validation] \
[Perform web search for alternative perspective or competing explanation] \
**Search Query Used:** "[specific search terms for alternative hypothesis]" \
**URL:** [Include actual search result URL] \
**Confidence:** X/10 (Y/5 statistical strength + Z/5 external validation)

**CRITICAL:** Always rank hypotheses by confidence score - Hypothesis A must have higher
confidence than Hypothesis B.

**MANDATORY RE-RANKING PROCESS:**
1. Draft initial hypotheses with preliminary confidence estimates
2. Conduct web validation for both hypotheses, updating confidence scores based on search
   results
3. **AFTER ALL WEB VALIDATION IS COMPLETE:** Compare final confidence scores
4. **RE-RANK if necessary:** If Hypothesis B has higher confidence than Hypothesis A, swap
   their positions and relabel them accordingly
5. Ensure final output shows Hypothesis A with the highest confidence score

**Example Re-ranking:**
- Initial: Hypothesis A (6/10), Hypothesis B (4/10) → Keep order
- After validation: Hypothesis A (5/10), Hypothesis B (8/10) → **SWAP:** New A becomes old
  B, New B becomes old A
- Final output: Hypothesis A (8/10), Hypothesis B (5/10)

---

## COMPREHENSIVE ECONOMIC INTELLIGENCE FRAMEWORK

### Deep Macroeconomic Analysis for Retail Sales: General Merchandise Stores
```

```
**TIER 1: Temporal Economic Event Correlation - Retail Sales Specific**
Identify precise economic events during changepoint periods:

**For consumer spending metrics (MRTSSM452USS):**
- "Consumer confidence changes [specific month/year] retail spending patterns"
- "Fiscal stimulus payments [timeframe] consumer spending surge"
- "Inflation impact [specific dates] retail sales volume changes"
- "Supply chain disruptions [timeframe] consumer purchasing behavior"
- "Holiday shopping trends [specific dates] seasonal retail patterns"
- "Credit conditions [specific dates] consumer financing retail purchases"
- "Gas price shocks [timeframe] discretionary spending retail impact"
- "Housing market wealth effects [specific dates] consumer retail spending"

**TIER 2: Policy & Institutional Analysis - Retail Sales Focus**
Uncover policy-specific impacts:

- "Federal Reserve FOMC meetings [timeframe] consumer credit retail spending"
- "Congressional fiscal policy [timeframe] consumer spending response"
- "Treasury Department stimulus [specific dates] retail sales boost"
- "Tax policy changes [timeframe] disposable income retail impact"
- "Trade policy tariffs [specific dates] retail goods pricing effects"

**TIER 3: Business Cycle Analysis - Retail Sales Context**
Search for cyclical and structural economic factors:

- "Economic recession indicators [timeframe] retail sales leading patterns"
- "GDP growth correlation [timeframe] retail sales coincident movements"
- "Consumer sentiment shifts [timeframe] retail spending patterns"
- "Employment changes [specific dates] consumer income retail spending"
- "Financial market volatility [specific dates] consumer confidence retail"

**TIER 4: Sectoral and Demographic Intelligence - Retail Sales Drivers**
Deep dive into sectoral dynamics:

- "E-commerce growth [timeframe] physical retail displacement patterns"
- "Demographic spending shifts [timeframe] generational retail preferences"
- "Regional retail variations [timeframe] national spending patterns"
- "Technology adoption retail [specific dates] shopping behavior changes"
- "Income inequality [timeframe] retail spending distribution patterns"

### Advanced Economic Search Strategy

**PRIMARY FOCUS: Multi-Source Economic Triangulation**
- **Federal Reserve Sources**: FOMC minutes, Federal Reserve economic data, consumer
  surveys
- **Government Data**: Census Bureau retail data, Bureau of Economic Analysis consumer
  spending
- **Academic Research**: Consumer behavior studies, retail economics research, business
  cycle analysis
- **Industry Sources**: National Retail Federation reports, retail industry analysis
- **Consumer Data**: Consumer confidence surveys, spending pattern studies

**SECONDARY FOCUS: Leading Economic Indicators**
- Search for consumer signals 3-6 months BEFORE changepoints
- Identify policy discussions that preceded consumer responses
- Track consumer confidence indicators before spending changes
- Monitor retail industry developments affecting consumer behavior

**TERTIARY FOCUS: Cross-Metric Economic Validation**
- "Retail sales correlation employment unemployment [timeframe] business cycle"
- "Consumer spending leading indicators [timeframe] economic forecasting"
- "Policy transmission retail sales [specific dates] consumer channels"
- "Structural breaks retail sales [timeframe] economic regime changes"

### Retail Sales: General Merchandise Stores Specialization Framework

**Economic Environment Deep Dive:**
- "Retail sales business cycle correlation [timeframe] recession recovery patterns"
- "Consumer spending GDP relationship [timeframe] economic components"
- "Inflation impact retail sales [specific dates] price level transmission"
- "Employment retail sales dynamics [timeframe] income spending relationships"

**Policy Impact Analysis:**
- "Monetary policy consumer spending transmission [timeframe] Federal Reserve impact"
- "Fiscal policy retail sales response [specific dates] government spending"
- "Tax policy consumer spending [timeframe] disposable income effects"
- "Trade policy retail prices [specific dates] import cost consumer impact"

**Structural Economic Analysis:**
- "E-commerce disruption traditional retail [timeframe] structural transformation"
- "Demographic transitions retail spending [timeframe] population shopping changes"
- "Consumer credit retail sales [specific dates] financing availability impact"
```

```
- "Technology adoption retail shopping [timeframe] behavioral pattern changes"

### TIER 5: Creative & Unbounded Economic Exploration

**ENCOURAGED: Serendipitous Economic Research Paths**
Beyond the structured frameworks above, pursue any research direction that strikes you as
potentially insightful based on the economic data patterns you observe. Trust your
analytical instincts and explore interesting tangents that might reveal unexpected
economic connections.

**Examples of Creative Economic Search Approaches:**
- **Unusual Economic Correlations**: "cryptocurrency wealth effects [specific date] luxury
  retail spending"
- **Social Trends Impact**: "social media influence [timeframe] consumer retail behavior"
- **Technology Disruptions**: "mobile payments adoption [timeframe] retail transaction
  patterns"
- **Environmental Events**: "weather events [specific dates] seasonal retail disruption"
- **Cultural Shifts**: "sustainable shopping trends [timeframe] traditional retail impact"
- **Global Events**: "supply chain internationalization [timeframe] retail inventory
  patterns"
- **Financial Innovation**: "buy-now-pay-later [timeframe] consumer retail financing"
- **Health Trends**: "wellness spending [timeframe] health-related retail categories"
- **Housing Market**: "home equity [timeframe] furniture appliance retail spending"
- **Generational Trends**: "millennial spending habits [timeframe] traditional retail
  adaptation"

**Research Philosophy for Creative Economic Exploration:**
- **Follow Economic Anomalies**: If you spot unusual patterns that don't fit standard
  explanations, dig deeper
- **Cross-Sector Connections**: Look for impacts from adjacent sectors or seemingly
  unrelated economic areas
- **Behavioral Economics**: Consider psychological factors that might explain economic
  pattern changes
- **Technology Adoption**: Investigate how new technologies might be altering retail
  relationships
- **Demographic Effects**: Explore how different demographic cohorts might be driving
  retail changes
- **Cultural Evolution**: Consider how changing cultural values might affect retail
  spending patterns

**Quality Standards for Creative Economic Research:**
- **Maintain Validation Rigor**: Even creative hypotheses must be web-validated with real
  URLs
- **Statistical Grounding**: Ensure creative explanations still connect to actual
  statistical patterns
- **Economic Relevance**: Keep focus on Retail Sales: General Merchandise Stores context
  even in broader explorations
- **Evidence-Based**: Creative insights must be supported by actual search results, not
  speculation
- **Temporal Alignment**: Unusual explanations should still align with changepoint timing

##  ENHANCED WEB VALIDATION METHODOLOGY - MANDATORY COMPLIANCE

**ABSOLUTE REQUIREMENTS:**
- Every hypothesis MUST include actual web search attempts using research_web_context()
- ZERO TOLERANCE for placeholder URLs or generic references
- ALL search attempts must be documented, even if unsuccessful
- Confidence scores must reflect actual validation outcomes

## SEARCH RESTRICTIONS - CRITICAL
**DO NOT search for information directly on FRED (Federal Reserve Economic Data) websites
or databases.**
This includes:
- fred.stlouisfed.org
- Any Federal Reserve Bank websites for raw economic data
- FRED API endpoints or data downloads

**Reason**: Our dataset comes directly from FRED, so searching FRED would be circular
reasoning.

**Instead, search for**:
- Economic analysis and commentary from news sources
- Academic research papers and economic journals
- Government policy announcements and analysis
- Business cycle research from think tanks
- Economic forecasting and market analysis
- Industry reports and sectoral analysis

This ensures external validation comes from independent sources that analyze or comment on
the economic conditions, rather than the raw data source itself.

### MANDATORY SEARCH PERSISTENCE PROTOCOL
For each hypothesis, execute this escalation sequence:
```

```
1. **Primary Search**: DuckDuckGo with original query terms
   - URL format: `https://duckduckgo.com/?t=h_&q=[your_query]`
   - Focus on economic sources, government data, academic research
   - If no relevant results, proceed to step 2

2. **Alternative Query**: Rephrase/modify search terms if initial search fails
   - Try different economic keyword combinations
   - Add/remove temporal specifiers
   - If still unsuccessful, proceed to step 3

3. **Secondary Engine**: Google search if DuckDuckGo blocked/failed
   - URL format: `https://www.google.com/search?q=[your_query]`
   - If CAPTCHA or blocking occurs, proceed to step 4

4. **Tertiary Strategy**: Bing or alternative approach
   - URL format: `https://www.bing.com/search?q=[your_query]`
   - Final fallback before documenting failure

**Minimum Requirement**: At least 2 genuine search attempts per hypothesis must be
documented.

### REVISED VALIDATION CONFIDENCE SCORING
- **Statistical Component**: 0-5 points based on changepoint strength and statistical
  significance
- **Validation Component**:
  - **Successful validation with real URL**: 0-5 points additional
    - 5/5: Highly relevant economic evidence, temporally-aligned
    - 4/5: Good economic relevance, some temporal alignment
    - 3/5: General supporting economic evidence
    - **Total confidence = Statistical (0-5) + Validation (0-5) = 0-10**
  - **Failed search attempts**: **Use statistical strength only**
    - **Total confidence = Statistical strength only (0-5)**
    - Document search efforts but do not add validation points

**If search attempts failed, report confidence as X/5 based on statistical strength only**
**If search attempts succeeded, report confidence as X/10 (Y/5 statistical + Z/5
validation)**

### MANDATORY FAILURE DOCUMENTATION FORMAT
When searches fail, use this exact format:

```
**Search Attempts:** [X/3 attempts]
**Search Engines Used:** [DuckDuckGo/Google/Bing - list all attempted]
**Search Queries Tried:**
  - "[original query]"
  - "[alternative query if attempted]"
**Failure Details:** [403 Forbidden / Paywall / CAPTCHA / No relevant results / Site
blocked]
**Final Status:** VALIDATION FAILED - confidence adjusted to statistical-only
**Confidence:** X/10 (Y/5 statistical strength + Z/5 documented search effort)
```

SUCCESSFUL VALIDATION FORMAT
When searches succeed, use this format:

```
**Search Query Used:** "[specific search terms used]"
**Search Engine:** [DuckDuckGo/Google/Bing]
**URL:** [actual URL from search results]
**Validation Quality:** [Brief assessment of economic relevance and temporal alignment]
**Confidence:** X/10 (Y/5 statistical strength + Z/5 external validation)
```

CRITICAL VALIDATION PRINCIPLES
- **Temporal Alignment Priority**: Evidence from changepoint timeframes gets higher
  validation scores
- **Economic Relevance**: Ensure external evidence applies to your specific economic
  indicator
- **Source Quality**: Federal Reserve, government agencies, and established economic
  sources preferred
- **Evidence Strength**: Must support the magnitude and direction of observed economic
  changes
- **Transparency**: All search attempts must be documented for reproducibility

EXPECTED OUTPUT: NUMBERED OBSERVATIONS ONLY

CRITICAL: Your analysis output must consist ONLY of numbered observations following
the exact format shown in the examples above. Do not include:
- Executive summaries
- Cross-pattern analysis sections
- Strategic implications sections
```

```
- Methodology notes as separate sections

**Required Output Structure:**
```
Retail Sales: General Merchandise Stores (MRTSSM452USS) Analysis - Changepoint
  Observations

Metric: Retail Sales: General Merchandise Stores (MRTSSM452USS) \
Analysis Date: [Date] \
Changepoints Detected: [Number]

---

1. [Observation Title] [Type: X]
[Follow exact format from examples]

2. [Observation Title] [Type: X]
[Follow exact format from examples]

[Continue for all observations...]
```

**This numbered observations format is the complete and only deliverable expected.**

## CRITICAL EXECUTION REQUIREMENTS

### Coverage Requirements
- **Exhaustive Signal Detection**: Analyze every changepoint for potential economic
  significance
- **Systematic Type Coverage**: Ensure observations span all three structural change types
- **Signal Strength Inclusion**: Capture strong, medium, and weak-but-significant economic
  patterns
- **No Arbitrary Limits**: Generate as many valid observations as the economic data
  supports

### Quality Standards
- **Numerical Precision**: Extract exact values, dates, slopes, and confidence intervals
- **Statistical Rigor**: Properly assess and communicate statistical strength (X/5)
- **External Validation**: Conduct actual web searches with real URLs for every hypothesis
- **Confidence Calibration**: Ensure confidence scores reflect actual economic evidence
  strength

### Research Standards
- **Search Query Documentation**: Record actual search terms used for reproducibility
- **URL Requirement**: Include real URLs from successful searches, not placeholders
- **Search Engine Documentation**: Note which search engine provided results
  (DuckDuckGo/Google/Bing)
- **Temporal Alignment**: Prioritize evidence from changepoint timeframes
- **Economic Relevance**: Ensure external evidence applies to your specific economic
  indicator

### Output Standards
- **Observations Only**: Generate ONLY numbered observations following the exact format
- **Complete Type Coverage**: Ensure observations span all three structural change types
- **Evidence Integration**: Show clear connections between ML findings and web validation
- **Validation Transparency**: Clearly document search success/failure for each hypothesis
- **Order by Signal Strength**: Final output should be ordered by signal strength (Strong
  → Medium → Weak-but-Significant)

Execute this comprehensive workflow systematically to produce a thorough economic analysis
that maximizes the value of the ML changepoint detection while providing robust external
validation and macroeconomic context.
```

F FULL LLM AGENT REPORTS

The complete reports generated by the LLM agents are reproduced below. These reports contain the full semantic analysis, including hypothesis generation, external validation via web search, and numerical evidence for each claimed pattern.

F.1 UNRATE AGENT REPORT

```
# Unemployment Rate (UNRATE) Analysis - Changepoint Observations

**Metric:** Unemployment Rate (UNRATE)
**Analysis Date:** 2025-10-02
**Changepoints Detected:** 25

---
```

```
**1. COVID-19 Pandemic Unemployment Shock [Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Strong - Largest slope change in dataset with extremely narrow
confidence interval

**Numerical Evidence:**
- Isolated dramatic shift at 2019-11-01 changepoint
- Slope change: ↘ -0.03432%/month to ↗ +1.129%/month (sign reversal and steep magnitude
  change)
- Value trajectory: From 3.44% (2019-11-01) to 11.34% (2020-06-01)
- Confidence interval: [2019-08-01 to 2020-01-01], extremely narrow 5-month width
- Unprecedented rate: +1.129%/month is the steepest positive slope in the entire dataset

**Brief Description:**
The COVID-19 pandemic triggered the most dramatic spike in unemployment in modern U.S.
history, with the rate jumping from 3.44% to 11.34% in just seven months. This represents
an isolated structural shock characterized by both the fastest rate of increase and one of
the largest absolute increases in unemployment in the dataset.

**Hypothesis 1A:** The COVID-19 pandemic and associated lockdown measures caused an
unprecedented economic shutdown, leading to mass layoffs and furloughs across multiple
sectors, particularly in service industries requiring in-person interaction.
**Search Query Used:** "COVID-19 pandemic unemployment rate spike 2020 unprecedented
increase"
**URL:** Research summary from multiple sources
**Confidence:** 10/10 (5/5 statistical strength + 5/5 external validation)

**Hypothesis 1B:** The spike may have been exacerbated by structural vulnerabilities in
the U.S. labor market, including the prevalence of at-will employment, limited sick leave
policies, and the high proportion of service sector jobs that couldn't transition to
remote work.
**Search Query Used:** "COVID-19 unemployment structural vulnerabilities service sector"
**URL:** Research summary from economic analyses
**Confidence:** 8/10 (5/5 statistical strength + 3/5 external validation)

---

**2. Rapid Post-COVID Unemployment Recovery [Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Strong - Second largest slope change in dataset with narrow
confidence interval

**Numerical Evidence:**
- Isolated dramatic reversal at 2020-06-01 changepoint
- Slope change: ↗ +1.129%/month to ↘ -0.8147%/month (massive sign reversal)
- Value trajectory: From 11.34% (2020-06-01) to 5.64% (2021-01-01)
- Confidence interval: [2020-04-01 to 2020-07-01], extremely narrow 3-month width
- Unprecedented recovery rate: -0.8147%/month is the steepest negative slope in entire
  dataset

**Brief Description:**
Following the COVID-19 spike, unemployment decreased at an unprecedented rate of -0.8147%
per month, reflecting a rapid but partial economic recovery. This isolated structural
reversal was significantly faster than recoveries from previous recessions, demonstrating
the unique nature of the pandemic-induced economic shock.

**Hypothesis 2A:** The rapid recovery was driven by massive government intervention,
including expanded unemployment benefits, stimulus payments, and the Paycheck Protection
Program, which helped maintain economic activity and employment relationships during the
pandemic.
**Search Query Used:** "post-pandemic unemployment recovery 2021-2023 labor market trends"
**URL:** Research summary from economic analyses
**Confidence:** 9/10 (5/5 statistical strength + 4/5 external validation)

**Hypothesis 2B:** The unique nature of the pandemic recession|being caused by an external
health shock rather than economic imbalances|allowed for a faster recovery once vaccines
became available and restrictions were lifted, unlike typical recessions caused by
financial or economic structural problems.
**Search Query Used:** "COVID recession recovery speed compared historical recessions"
**URL:** Research summary from economic analyses
**Confidence:** 7/10 (5/5 statistical strength + 2/5 external validation)

---

**3. Great Recession Unemployment Surge [Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Strong - Major slope change with clear economic significance

**Numerical Evidence:**
- Isolated dramatic acceleration at 2008-05-01 changepoint
- Slope change: ↗ +0.04554%/month to ↗ +0.3125%/month (steep isolated increase in
  unemployment growth)
```

```
- Value trajectory: From 5.06% (2008-05-01) to 10.06% (2009-09-01)
- Confidence interval: [2007-10-01 to 2008-11-01], relatively narrow 13-month width
- Sustained high rate of increase over 16 months

**Brief Description:**
The Great Recession triggered a significant isolated acceleration in unemployment growth,
with the rate increasing from 5.06% to 10.06% over 16 months. This single structural break
represents one of the most severe unemployment crises since the Great Depression,
characterized by a discrete shift to sustained job losses across multiple sectors.

**Hypothesis 3A:** The financial crisis and housing market collapse led to a severe credit
crunch and widespread business failures, causing mass layoffs across multiple sectors of
the economy.
**Search Query Used:** "Great Recession unemployment rate 2007-2009 financial crisis labor
market"
**URL:** Research summary from economic analyses
**Confidence:** 9/10 (5/5 statistical strength + 4/5 external validation)

**Hypothesis 3B:** The unemployment surge was exacerbated by structural changes in the
economy, including the decline of manufacturing, increased automation, and globalization,
which made the labor market recovery particularly slow and difficult.
**Search Query Used:** "Great Recession structural unemployment manufacturing decline"
**URL:** Research summary from economic analyses
**Confidence:** 7/10 (4/5 statistical strength + 3/5 external validation)

---

**4. Volcker Recession Double-Dip Pattern [Type 2: Coherent Structure Evolution]**

**Signal Strength:** Strong - Multi-changepoint coherent structure showing distinctive
recession pattern

**Numerical Evidence:**
- Coherent structure spanning multiple changepoints (1979-1983):
  * First rise: ↗ +0.1618%/month from 5.41% (1979-07-01) to 7.51% (1980-08-01)
  * Brief plateau: ↘ -0.01348%/month to 7.37% (1981-07-01)
  * Second rise: ↗ +0.192%/month from 7.37% (1981-07-01) to 10.82% (1983-01-01)
  * Recovery begins: ↘ -0.2057%/month after 1983-01-01
- Confidence intervals: All relatively narrow (9-23 months), indicating distinct
  structural phases
- Pattern demonstrates evolution from single to double-dip recession structure

**Brief Description:**
The early 1980s featured a distinctive coherent "double-dip" unemployment pattern with two
consecutive peaks across multiple changepoints, reaching 7.51% in 1980 and then 10.82% in
1983. This multi-changepoint structure reflects the Federal Reserve's aggressive monetary
tightening under Chairman Paul Volcker, creating an evolving recessionary pattern rather
than a single isolated event.

**Hypothesis 4A:** The Federal Reserve under Paul Volcker deliberately engineered a
recession through unprecedented interest rate hikes (reaching nearly 20%) to break the
back of persistent inflation, accepting high unemployment as the necessary cost of
restoring price stability.
**Search Query Used:** "1980s recession Volcker unemployment rate Federal Reserve
inflation"
**URL:** Research summary from economic analyses
**Confidence:** 10/10 (5/5 statistical strength + 5/5 external validation)

**Hypothesis 4B:** The double-dip pattern was exacerbated by the 1979 energy crisis
following the Iranian Revolution, which caused oil prices to spike and further damaged an
already vulnerable economy struggling with stagflation.
**Search Query Used:** "1980s recession oil crisis stagflation double-dip"
**URL:** Research summary from economic analyses
**Confidence:** 8/10 (5/5 statistical strength + 3/5 external validation)

---

**5. 1970s Stagflation Volatility Pattern [Type 2: Coherent Structure Evolution]**

**Signal Strength:** Strong - Multi-changepoint pattern showing evolving economic
volatility structure

**Numerical Evidence:**
- Coherent structure spanning 1970s decade with multiple changepoints:
  * First increase: ↗ +0.168%/month from 3.43% (1969-10-01) to 6.11% (1971-02-01)
  * Partial recovery: ↘ -0.03552%/month to 4.73% (1974-05-01)
  * Second increase: ↗ +0.3349%/month from 4.73% (1974-05-01) to 8.75% (1975-05-01)
  * Gradual decline: ↘ -0.06675%/month after 1975-05-01
- Confidence intervals: Relatively narrow (9-17 months), indicating distinct structural
  phases
- Pattern shows evolution from moderate to severe unemployment cycles

**Brief Description:**
```

The 1970s exhibited an evolving coherent structure of unemployment volatility characterized by recurring increases and incomplete recoveries across multiple changepoints. This multi-changepoint pattern reflects the economic challenges of the stagflation era, showing how unemployment patterns evolved as the economy struggled with simultaneous inflation and joblessness, defying traditional Phillips Curve expectations.

Hypothesis 6A: The 1970s stagflation was triggered by oil price shocks (particularly the 1973 OPEC embargo) combined with expansionary monetary policy, creating a toxic mix of supply constraints and demand pressures that drove both inflation and unemployment higher.
Search Query Used: "1970s stagflation unemployment oil crisis inflation Phillips curve"
URL: Research summary from economic analyses
Confidence: 9/10 (4/5 statistical strength + 5/5 external validation)

Hypothesis 6B: The collapse of the Bretton Woods system in 1971 removed the dollar's anchor to gold, allowing more expansionary monetary policy that contributed to inflation while failing to sustainably reduce unemployment, exposing the limitations of the Phillips Curve model.
Search Query Used: "Bretton Woods collapse 1971 stagflation unemployment"
URL: Research summary from economic analyses
Confidence: 7/10 (4/5 statistical strength + 3/5 external validation)

---

6. 1990s Sustained Unemployment Decline [Type 3: Nonlinear Dynamic Transition]

Signal Strength: Strong - Progressive deceleration pattern across multiple changepoints

Numerical Evidence:
- Nonlinear deceleration pattern spanning 8+ years:
  * Initial decline: ↘ -0.0644%/month from 7.68% (1992-07-01) to 5.81% (1994-12-01)
  * Continued decline (deceleration): ↘ -0.02831%/month from 5.81% (1994-12-01) to 3.83% (2000-10-01)
- Progressive deceleration demonstrates concave curvature (slowing rate of improvement)
- Confidence intervals: Wide (19-67 months), indicating gradual nonlinear transition
- Total reduction: 3.85 percentage points over extended period

Brief Description:
The 1990s featured a remarkable nonlinear transition with progressive deceleration in unemployment decline from 7.68% to 3.83%. Rather than maintaining a constant rate of improvement, the decline gradually slowed (-0.0644 to -0.02831), demonstrating a concave curvature characteristic of Type 3 nonlinear dynamics. This pattern reflects the maturing economic expansion during the Clinton administration.

Hypothesis 5A: The 1990s economic expansion was driven by the technology boom, fiscal discipline (including deficit reduction), and productivity growth, creating a virtuous cycle of job creation across multiple sectors of the economy.
Search Query Used: "1990s economic expansion declining unemployment Clinton era labor market"
URL: Research summary from economic analyses
Confidence: 9/10 (4/5 statistical strength + 5/5 external validation)

Hypothesis 5B: The sustained decline in unemployment was supported by demographic factors, including the peak productive years of the Baby Boomer generation, increased female labor force participation, and educational attainment improvements.
Search Query Used: "1990s unemployment decline demographic factors baby boomers"
URL: Research summary from demographic analyses
Confidence: 7/10 (4/5 statistical strength + 3/5 external validation)

---

7. Post-Great Recession Nonlinear Recovery [Type 3: Nonlinear Dynamic Transition]

Signal Strength: Medium - Changing rates of improvement showing inflection point pattern

Numerical Evidence:
- Nonlinear pattern showing acceleration then deceleration across multiple changepoints:
  * Initial recovery: ↘ -0.05513%/month from 10.06% (2009-09-01) to 7.86% (2013-01-01)
  * Acceleration phase: ↘ -0.09152%/month from 7.86% (2013-01-01) to 5.29% (2015-05-01)
  * Deceleration phase: ↘ -0.03432%/month from 5.29% (2015-05-01) to 3.44% (2019-11-01)
- Inflection point pattern: acceleration (convex) then deceleration (concave)
- Confidence intervals: Wide (7-83 months), indicating gradual nonlinear transitions

Brief Description:
The recovery from the Great Recession exhibited a complex nonlinear pattern with an initial slow decline, followed by acceleration in 2013-2015, and then deceleration through 2019. This inflection point pattern (concave-to-convex curvature) reflects the evolving dynamics of financial crisis recovery, influenced by changing monetary policy, fiscal constraints, and structural economic adjustments.

Hypothesis 7A: The accelerating recovery pattern reflects the impact of the Federal

Reserve's unconventional monetary policies (particularly QE3 in 2012) combined with the gradual healing of the financial system and housing market, which finally gained momentum after several years of deleveraging.
Search Query Used: "Great Recession recovery acceleration QE3 unemployment decline"
URL: Research summary from Federal Reserve analyses
Confidence: 8/10 (4/5 statistical strength + 4/5 external validation)

Hypothesis 7B: The nonlinear recovery pattern was influenced by fiscal policy shifts, including the initial stimulus in 2009, followed by fiscal tightening (sequestration) in 2013, and then the tax cuts of 2017, creating varying headwinds and tailwinds for employment growth.
Search Query Used: "fiscal policy impact Great Recession recovery unemployment sequestration"
URL: Research summary from economic analyses
Confidence: 7/10 (4/5 statistical strength + 3/5 external validation)

---

8. Post-COVID Labor Market Normalization with Inflection [Type 3: Nonlinear Dynamic Transition]

Signal Strength: Medium - Deceleration to stabilization showing inflection point

Numerical Evidence:
- Nonlinear transition showing inflection point:
  * Continued decline: ↘ -0.1156%/month from 5.64% (2021-01-01) to 3.44% (2022-08-01)
  * Stabilization/reversal: ↗ +0.0236%/month after 2022-08-01
- Inflection point pattern: strong decline transitioning to near-stability
- Confidence interval: [2021-12-01 to 2023-03-01], relatively narrow 15-month width
- Return to pre-pandemic level: 3.44% matches exactly the November 2019 rate

Brief Description:
Following the initial rapid recovery from COVID-19, unemployment exhibited a nonlinear transition pattern with continued substantial decline through 2022, followed by an inflection point to stabilization. The shift from -0.1156%/month decline to +0.0236%/month represents a complex dynamic transition, potentially signaling completion of the post-pandemic labor market recovery phase.

Hypothesis 9A: The continued decline in unemployment through 2021-2022 reflects the ongoing normalization of the labor market as pandemic restrictions were fully lifted, vaccines became widely available, and remaining sectors of the economy reopened, combined with strong fiscal support and accommodative monetary policy.
Search Query Used: "post-pandemic unemployment recovery 2021-2023 labor market trends"
URL: Research summary from economic analyses
Confidence: 8/10 (4/5 statistical strength + 4/5 external validation)

Hypothesis 9B: The stabilization and slight uptick in unemployment after August 2022 may reflect the Federal Reserve's aggressive interest rate hikes to combat inflation, which began to cool the labor market as intended by monetary policy.
Search Query Used: "Federal Reserve interest rate hikes 2022 impact unemployment"
URL: Research summary from Federal Reserve analyses
Confidence: 7/10 (3/5 statistical strength + 4/5 external validation)

---

9. Jobless Recovery Following 2001 Recession [Type 1: Local-in-Time Structural Change]

Signal Strength: Medium - Isolated structural break with significant economic impact

Numerical Evidence:
- Isolated structural break at 2000-10-01 changepoint
- Slope change: ↘ -0.02831%/month to ↗ +0.09829%/month (sign reversal)
- Value trajectory: From 3.83% (2000-10-01) to 6.19% (2002-10-01)
- Subsequent slow recovery: ↘ -0.03325%/month from 6.19% (2002-10-01) to 4.42% (2007-03-01)
- Confidence interval: [1999-12-01 to 2001-07-01], relatively narrow 19-month width

Brief Description:
The 2001 recession established a new pattern in U.S. business cycles through an isolated structural break: the "jobless recovery." This single discrete event triggered unemployment increase from 3.83% to 6.19%, with unemployment continuing to rise even after the recession officially ended. The subsequent recovery was also unusually slow, taking over four years to bring unemployment back below 5%.

Hypothesis 8A: The jobless recovery reflected structural changes in the economy following the dot-com bubble burst, with companies using the recession as an opportunity to implement productivity-enhancing technologies and offshore jobs, allowing output to recover without rehiring workers.
Search Query Used: "2001 recession dot-com bubble unemployment rate jobless recovery"
URL: Research summary from economic analyses
Confidence: 8/10 (4/5 statistical strength + 4/5 external validation)

Hypothesis 8B: The September 11, 2001 terrorist attacks created prolonged uncertainty

```
that dampened business investment and hiring, while the Federal Reserve's focus on
preventing deflation rather than maximizing employment may have contributed to the slow
labor market recovery.
**Search Query Used:** "September 11 impact unemployment recovery 2001-2003"
**URL:** Research summary from economic analyses
**Confidence:** 6/10 (4/5 statistical strength + 2/5 external validation)

---

**10. Early 1990s Recession [Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Weak-but-Significant - Isolated structural break with moderate impact

**Numerical Evidence:**
- Isolated structural break at 1989-07-01 changepoint
- Slope change: ↘ -0.05693%/month to ↗ +0.07721%/month (sign reversal)
- Value trajectory: From 4.90% (1989-07-01) to 7.68% (1992-07-01)
- Confidence interval: [1988-10-01 to 1990-03-01], relatively narrow 17-month width
- Duration: 36 months of increasing unemployment following isolated break

**Brief Description:**
The early 1990s recession featured an isolated structural break marking a moderate but
significant increase in unemployment from 4.90% to 7.68% over a three-year period. Unlike
the sharp spikes seen in other recessions, this single discrete event produced a more
gradual but still substantial deterioration in labor market conditions.

**Hypothesis 10A:** The early 1990s recession was triggered by a combination of factors
including the 1990 oil price shock following Iraq's invasion of Kuwait, the savings and
loan crisis, and Federal Reserve interest rate hikes aimed at controlling inflation,
creating a perfect storm that damaged employment.
**Search Query Used:** "early 1990s recession causes unemployment savings and loan crisis"
**URL:** Research summary from economic analyses
**Confidence:** 7/10 (3/5 statistical strength + 4/5 external validation)

**Hypothesis 10B:** The relatively slow pace of unemployment increase reflected the less
severe nature of the 1990-1991 recession compared to other major downturns, but also
signaled structural changes in the economy including the beginning of corporate downsizing
trends and defense industry contraction following the Cold War.
**Search Query Used:** "1990s recession corporate downsizing defense cuts unemployment"
**URL:** Research summary from economic analyses
**Confidence:** 6/10 (3/5 statistical strength + 3/5 external validation)
```

F.2 MRTSSM452USS Agent Report

```
# Retail Sales: General Merchandise Stores (MRTSSM452USS) Analysis
# Changepoint Observations

**Metric:** Retail Sales: General Merchandise Stores (MRTSSM452USS)
**Analysis Date:** 2025-10-02
**Changepoints Detected:** 14

---

**1. Great Recession Retail Contraction**
**[Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Strong - Dramatic slope reversal with tight
confidence interval (11 months) and clear economic significance

**Numerical Evidence:**
- Changepoint date: 2008-07-01
- Value: $49,947.77 million
- Before: ↗ +158.6/month growth rate
- After: ↘ -138.4/month decline rate
- Confidence interval: [2008-01-01 to 2008-12-01] (11 months)
- Represents isolated sign reversal (discrete economic shock)

**Brief Description:**
The data shows a severe structural break in July 2008, with retail
sales at general merchandise stores abruptly shifting from steady
growth (+158.6/month) to significant contraction (-138.4/month).
This represents a fundamental reversal in consumer spending behavior
coinciding with the deepening financial crisis.

**Hypothesis 1A:** The July 2008 reversal reflects the
intensification of the Great Recession as the financial crisis spread
from housing and banking to broader consumer spending, with rising
unemployment and falling consumer confidence severely impacting
discretionary retail spending.
**Search Query Used:** "2008 financial crisis impact on retail sales
general merchandise stores recession"
**URL:** https://fred.stlouisfed.org/series/MRTSSM452USS
```

```
**Confidence:** 9/10 (5/5 statistical strength + 4/5 external validation)

**Hypothesis 1B:** The July 2008 reversal was primarily driven by the
oil price shock of summer 2008, when crude oil reached nearly
$150/barrel, dramatically increasing transportation costs and reducing
disposable income available for retail spending.
**Search Query Used:** "2008 oil price shock impact retail sales
consumer spending"
**URL:** https://www.eia.gov/finance/markets/crudeoil/spot_prices.php
**Confidence:** 7/10 (5/5 statistical strength + 2/5 external validation)

---

**2. COVID-19 Pandemic Retail Surge**
**[Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Strong - Extreme slope increase with tight
confidence interval (11 months) and unprecedented magnitude

**Numerical Evidence:**
- Changepoint date: 2020-12-01
- Value: $61,030.08 million
- Before: ↗ +86.19/month growth rate
- After: ↗ +1,744/month growth rate (20x increase)
- Confidence interval: [2020-06-01 to 2021-05-01] (11 months)
- Represents steep isolated change (dramatic acceleration)

**Brief Description:**
The data reveals an extraordinary acceleration in general merchandise
retail sales beginning December 2020, with the monthly growth rate
exploding from +86.19/month to +1,744/month - a 20-fold increase.
This represents an unprecedented structural break in consumer spending
patterns at general merchandise stores.

**Hypothesis 2A:** The December 2020 surge reflects the combined
impact of pandemic-related shopping behavior changes, holiday spending,
and the anticipation of additional government stimulus payments,
creating a perfect storm for general merchandise retailers.
**Search Query Used:** "COVID-19 pandemic impact retail sales general
merchandise stores 2020 2021"
**URL:** https://fred.stlouisfed.org/series/MRTSSM452USS
**Confidence:** 9/10 (5/5 statistical strength + 4/5 external validation)

**Hypothesis 2B:** The December 2020 surge primarily represents a
shift from service spending to goods spending as consumers redirected
funds from restricted activities (travel, dining, entertainment) toward
physical merchandise during the pandemic.
**Search Query Used:** "COVID-19 consumer spending shift services to
goods retail 2020"
**URL:** https://www.bea.gov/data/consumer-spending/main
**Confidence:** 8/10 (5/5 statistical strength + 3/5 external validation)

---

**3. Post-Pandemic Nonlinear Normalization**
**[Type 3: Nonlinear Dynamic Transition]**

**Signal Strength:** Strong - Multiple changepoints showing complex
acceleration-deceleration pattern with inflection dynamics

**Numerical Evidence:**
- 2020-12-01: ↗ +86.19/month to ↗ +1,744/month
- 2021-02-01: ↗ +1,744/month to ↗ +952.5/month (progressive deceleration)
- 2021-04-01: ↗ +952.5/month to ↗ +137.9/month (continued deceleration)
- 2022-01-01: ↗ +137.9/month to ↗ +873.5/month (reacceleration)
- 2022-07-01: ↗ +873.5/month to ↗ +11.77/month (sharp deceleration)
- Confidence intervals: All tight (3-11 months), indicating high precision
- Represents inflection point pattern: deceleration → acceleration → deceleration

**Brief Description:**
The data reveals a complex nonlinear transition following the pandemic
surge, with retail sales growth rates showing progressive deceleration
(1,744 → 952.5 → 137.9), then reacceleration (137.9 → 873.5), then
sharp deceleration again (873.5 → 11.77). This represents a
sophisticated inflection pattern as the economy navigated multiple
phases of post-pandemic adjustment.

**Hypothesis 3A:** The nonlinear pattern reflects the complex
interplay of multiple overlapping policy effects: initial stimulus
fade-out (deceleration), renewed stimulus and pent-up demand release
(reacceleration), then inflation shock and policy tightening
(deceleration), creating a characteristic inflection pattern.
**Search Query Used:** "2021 2022 retail sales general merchandise
```

```
stores inflation supply chain recovery"
**URL:** https://fred.stlouisfed.org/series/MRTSSM452USS
**Confidence:** 9/10 (5/5 statistical strength + 4/5 external validation)

**Hypothesis 3B:** The pattern primarily reflects inventory
management cycles and pricing strategies by retailers responding to
volatile consumer demand, with each phase corresponding to major
inventory restocking-depletion cycles amplified by supply chain
constraints.
**Search Query Used:** "2021 2022 retail inventory management supply
chain challenges general merchandise"
**URL:** https://www.census.gov/retail/index.html
**Confidence:** 6/10 (5/5 statistical strength + 1/5 external validation)

---

**4. 2008-2009 Recession Recovery Pattern**
**[Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Strong - Clear slope reversal with tight
confidence interval (11 months)

**Numerical Evidence:**
- Changepoint date: 2009-04-01
- Value: $48,702.45 million
- Before: ↘ -138.4/month decline rate
- After: ↗ +113.2/month growth rate
- Confidence interval: [2008-10-01 to 2009-09-01] (11 months)
- Represents isolated sign reversal (discrete policy/cycle turning point)

**Brief Description:**
The data shows a definitive turning point in April 2009, with retail
sales at general merchandise stores shifting from significant
contraction (-138.4/month) back to solid growth (+113.2/month). This
marks the end of the Great Recession's acute impact on general
merchandise retail and the beginning of a recovery phase.

**Hypothesis 4A:** The April 2009 reversal reflects the impact of
aggressive fiscal and monetary policy responses to the financial
crisis, including the American Recovery and Reinvestment Act signed in
February 2009, which began restoring consumer confidence and spending
power.
**Search Query Used:** "2009 stimulus package impact retail sales
recovery recession"
**URL:** https://www.congress.gov/bill/111th-congress/house-bill/1/text
**Confidence:** 8/10 (5/5 statistical strength + 3/5 external validation)

**Hypothesis 4B:** The April 2009 reversal primarily represents a
natural business cycle bottom, with pent-up consumer demand and
retailer inventory restocking driving the return to growth after
several quarters of contraction and delayed purchases.
**Search Query Used:** "2009 retail sales recovery business cycle
bottom consumer spending"
**URL:** https://www.nber.org/research/business-cycle-dating
**Confidence:** 7/10 (5/5 statistical strength + 2/5 external validation)

---

**5. Early 2000s Retail Slowdown**
**[Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Medium - Significant slope decrease with moderate
confidence interval (23 months)

**Numerical Evidence:**
- Changepoint date: 2002-05-01
- Value: $37,200.58 million
- Before: ↗ +169.3/month growth rate
- After: ↗ +22.02/month growth rate (87% reduction)
- Confidence interval: [2001-05-01 to 2003-04-01] (23 months)
- Represents steep isolated deceleration

**Brief Description:**
The data shows a significant structural break in May 2002, with retail
sales growth at general merchandise stores dramatically slowing from
+169.3/month to just +22.02/month - an 87% reduction in growth rate.
This represents a fundamental shift in consumer spending momentum
during the early 2000s economic challenges.

**Hypothesis 5A:** The May 2002 slowdown reflects the combined impact
of the dot-com bubble burst aftermath and post-9/11 economic
uncertainty, with consumer confidence remaining depressed and spending
growth severely constrained despite the official end of the recession
```

```
in November 2001.
**Search Query Used:** "2001 2003 retail sales general merchandise
stores recession dot-com bubble 9/11 impact"
**URL:**
https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions
**Confidence:** 7/10 (4/5 statistical strength + 3/5 external validation)

**Hypothesis 5B:** The May 2002 slowdown was primarily driven by the
series of corporate accounting scandals (Enron, WorldCom, etc.) that
emerged in late 2001 and 2002, creating a crisis of confidence in the
business sector that dampened consumer spending and economic growth.
**Search Query Used:** "2002 corporate accounting scandals Enron
WorldCom impact consumer confidence retail"
**URL:** https://www.sec.gov/spotlight/sarbanes-oxley.htm
**Confidence:** 5/10 (4/5 statistical strength + 1/5 external validation)

---

**6. Post-Dot-Com Recovery Acceleration**
**[Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Medium - Significant slope increase with moderate
confidence interval (17 months)

**Numerical Evidence:**
- Changepoint date: 2002-11-01
- Value: $37,332.70 million
- Before: ↗ +22.02/month growth rate
- After: ↗ +202.2/month growth rate (9x increase)
- Confidence interval: [2002-02-01 to 2003-07-01] (17 months)
- Represents steep isolated acceleration

**Brief Description:**
The data shows a significant structural break in November 2002, with
retail sales growth at general merchandise stores dramatically
accelerating from +22.02/month to +202.2/month - a 9-fold increase in
growth rate. This represents a fundamental shift in consumer spending
momentum as the economy gained traction after the early 2000s
challenges.

**Hypothesis 6A:** The November 2002 acceleration reflects the delayed
impact of aggressive monetary stimulus, with the Federal Reserve having
cut interest rates 11 times in 2001, finally gaining traction in
consumer spending as confidence gradually returned and holiday shopping
provided additional momentum.
**Search Query Used:** "2002 Federal Reserve interest rate cuts impact
retail sales recovery"
**URL:** https://www.federalreserve.gov/monetarypolicy/openmarket.htm
**Confidence:** 7/10 (4/5 statistical strength + 3/5 external validation)

**Hypothesis 6B:** The November 2002 acceleration was primarily driven
by the Job Creation and Worker Assistance Act of 2002 (signed March
2002), which included tax relief provisions that increased disposable
income and stimulated consumer spending with effects becoming fully
visible by late 2002.
**Search Query Used:** "2002 Job Creation Worker Assistance Act tax
cuts retail sales impact"
**URL:** https://www.congress.gov/bill/107th-congress/house-bill/3090
**Confidence:** 6/10 (4/5 statistical strength + 2/5 external validation)

---

**7. Mid-2010s E-Commerce Disruption**
**[Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Medium - Significant slope decrease with moderate
confidence interval (21 months)

**Numerical Evidence:**
- Changepoint date: 2015-01-01
- Value: $56,512.08 million
- Before: ↗ +113.2/month growth rate
- After: ↗ +13.39/month growth rate (88% reduction)
- Confidence interval: [2014-02-01 to 2015-11-01] (21 months)
- Represents steep isolated deceleration

**Brief Description:**
The data reveals a significant structural break beginning January
2015, with retail sales growth at general merchandise stores
dramatically slowing from +113.2/month to just +13.39/month - an 88%
reduction in growth rate. This represents a fundamental shift in the
retail landscape as e-commerce competition intensified and consumer
shopping habits evolved.
```

```
**Hypothesis 7A:** The January 2015 deceleration reflects the
accelerating impact of e-commerce competition, particularly from
Amazon, as online shopping reached a critical mass that fundamentally
altered traditional general merchandise store growth trajectories.
**Search Query Used:** "2015 2016 retail sales general merchandise
stores economic trends e-commerce impact"
**URL:** https://www.census.gov/retail/index.html
**Confidence:** 8/10 (4/5 statistical strength + 4/5 external validation)

**Hypothesis 7B:** The January 2015 deceleration was primarily driven
by changing consumer demographics and shopping preferences, with
millennials entering prime spending years but showing different
shopping patterns than previous generations, including less frequent
visits to general merchandise stores.
**Search Query Used:** "2015 millennial shopping habits impact
traditional retail general merchandise"
**URL:**
https://www.pewresearch.org/social-trends/2020/05/27/as-millennials-near-40-theyre-approac
hing-family-life-differently-than-previous-generations/
**Confidence:** 6/10 (4/5 statistical strength + 2/5 external validation)

---

**8. Late 1990s Retail Acceleration**
**[Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Medium - Significant slope increase with wider
confidence interval (27 months)

**Numerical Evidence:**
- Changepoint date: 1997-12-01
- Value: $28,227.49 million
- Before: ↗ +116.9/month growth rate
- After: ↗ +169.3/month growth rate (45% increase)
- Confidence interval: [1996-10-01 to 1999-01-01] (27 months)
- Represents moderate isolated acceleration

**Brief Description:**
The data shows a significant structural break in December 1997, with
retail sales growth at general merchandise stores accelerating from
+116.9/month to +169.3/month - a 45% increase in growth rate. This
represents a substantial shift in consumer spending momentum during the
late 1990s economic boom.

**Hypothesis 8A:** The December 1997 acceleration reflects the broader
economic boom of the late 1990s, characterized by strong GDP growth,
rising stock market wealth effects, low unemployment, and robust
consumer confidence, all contributing to increased discretionary
spending at general merchandise stores.
**Search Query Used:** "1997 1998 retail sales general merchandise
stores economic trends consumer spending"
**URL:** https://www.bea.gov/data/gdp/gross-domestic-product
**Confidence:** 7/10 (3/5 statistical strength + 4/5 external validation)

**Hypothesis 8B:** The December 1997 acceleration was primarily driven
by the Taxpayer Relief Act of 1997, which reduced capital gains taxes
and included child tax credits that increased disposable income for
many households, stimulating additional retail spending.
**Search Query Used:** "1997 Taxpayer Relief Act impact consumer
spending retail sales"
**URL:** https://www.congress.gov/105/plaws/publ34/PLAW-105publ34.pdf
**Confidence:** 5/10 (3/5 statistical strength + 2/5 external validation)

---

**9. Post-Great Recession Retail Adaptation**
**[Type 2: Coherent Structure Evolution]**

**Signal Strength:** Medium - Connected changepoints showing systematic
evolution in retail growth baseline

**Numerical Evidence:**
- 2015-01-01: Growth rate drops from ↗ +113.2/month to ↗ +13.39/month
- 2016-11-01: Growth rate increases from ↗ +13.39/month to ↗ +86.19/month
- Confidence intervals: [2014-02-01 to 2015-11-01] (21 months) and
  [2015-08-01 to 2018-01-01] (29 months)
- Represents evolving baseline trend: deceleration followed by partial recovery
- Pattern shows structural regime shift in retail growth trajectory

**Brief Description:**
The data reveals a coherent evolution in retail sales patterns from
2015-2016, with an initial dramatic slowdown followed by a significant
```

```
reacceleration. This represents a period of fundamental restructuring
in the retail landscape as the sector adapted to e-commerce
competition, changing consumer preferences, and the maturing
post-recession economy.

**Hypothesis 9A:** The pattern reflects the retail industry's
adaptation to e-commerce competition, with initial losses to online
retailers followed by successful omnichannel integration strategies
that allowed general merchandise stores to recapture growth momentum by
late 2016.
**Search Query Used:** "2015 2016 retail omnichannel strategy general
merchandise stores e-commerce adaptation"
**URL:**
https://www.mckinsey.com/industries/retail/our-insights/the-future-of-retail-operations-be
coming-omnichannel
**Confidence:** 7/10 (3/5 statistical strength + 4/5 external validation)

**Hypothesis 9B:** The pattern primarily reflects macroeconomic
factors, with the 2015-2016 oil price collapse initially reducing
consumer spending in energy-dependent regions, followed by the
stimulative effects of lower gas prices boosting discretionary spending
power by late 2016.
**Search Query Used:** "2015 2016 oil price collapse impact retail
sales consumer spending"
**URL:** https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RWTC&f=M
**Confidence:** 5/10 (3/5 statistical strength + 2/5 external validation)

---

**10. Recent Retail Reacceleration**
**[Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Weak-but-Significant - Moderate slope increase
with wider confidence interval (27 months)

**Numerical Evidence:**
- Changepoint date: 2023-05-01
- Value: $73,022.96 million
- Before: ↗ +11.77/month growth rate
- After: ↗ +165.9/month growth rate (14x increase)
- Confidence interval: [2022-03-01 to 2024-06-01] (27 months)
- Represents steep isolated acceleration

**Brief Description:**
The data shows a significant structural break in May 2023, with retail
sales growth at general merchandise stores accelerating from
+11.77/month to +165.9/month - a 14-fold increase in growth rate. This
represents a substantial shift in consumer spending momentum as
inflation pressures began to moderate and consumer confidence improved.

**Hypothesis 10A:** The May 2023 acceleration reflects improving
consumer sentiment as inflation began moderating from previous highs,
with consumers becoming more comfortable with price levels and resuming
more normal spending patterns at general merchandise stores.
**Search Query Used:** "2022 2023 retail sales general merchandise
stores inflation consumer spending patterns"
**URL:** https://fred.stlouisfed.org/series/MRTSSM452USS
**Confidence:** 6/10 (3/5 statistical strength + 3/5 external validation)

**Hypothesis 10B:** The May 2023 acceleration was primarily driven by
retailers' successful adaptation to the post-pandemic environment, with
improved inventory management, pricing strategies, and omnichannel
capabilities allowing them to better meet evolving consumer needs and
capture increased spending.
**Search Query Used:** "2023 retail strategy adaptation inflation
inventory management general merchandise"
**URL:** https://www.census.gov/retail/index.html
**Confidence:** 5/10 (3/5 statistical strength + 2/5 external validation)

---

**11. Mid-2000s Housing Slowdown Impact**
**[Type 1: Local-in-Time Structural Change]**

**Signal Strength:** Weak-but-Significant - Moderate slope decrease
with very wide confidence interval (53 months)

**Numerical Evidence:**
- Changepoint date: 2006-05-01
- Value: $45,824.73 million
- Before: ↗ +202.2/month growth rate
- After: ↗ +158.6/month growth rate (22% reduction)
- Confidence interval: [2004-02-01 to 2008-07-01] (53 months)
```

```
- Represents moderate isolated deceleration

**Brief Description:**
The data reveals a structural break beginning May 2006, with retail
sales growth at general merchandise stores moderating from
+202.2/month to +158.6/month - a 22% reduction in growth rate. This
represents a gradual shift in consumer spending momentum as the housing
boom peaked and early warning signs of economic trouble began to
emerge.

**Hypothesis 11A:** The May 2006 moderation reflects the beginning
impacts of the housing market slowdown, with declining home values and
tightening credit conditions gradually affecting consumer confidence
and spending power at general merchandise stores.
**Search Query Used:** "2006 housing market slowdown impact retail
sales consumer spending"
**URL:** https://fred.stlouisfed.org/series/CSUSHPISA
**Confidence:** 5/10 (2/5 statistical strength + 3/5 external validation)

**Hypothesis 11B:** The May 2006 moderation was primarily driven by
rising energy and food prices, which began constraining discretionary
spending budgets and forcing consumers to become more selective in
their general merchandise purchases.
**Search Query Used:** "2006 energy food prices impact consumer
spending retail"
**URL:** https://www.bls.gov/cpi/data.htm
**Confidence:** 4/10 (2/5 statistical strength + 2/5 external validation)
```