Epiphenomenon@Trove

CAMS Falsification Test — Pre-registered Protocol v5

Results Report

Kari | May 2026 | Australia 1900–1954

Abstract

This report presents results from a pre-registered falsification test of the Complex Adaptive Model System (CAMS) applied to Australian institutional history, 1900–1954. Discourse proxies derived from the Trove digitised newspaper corpus (55 annual observations, N = 148 million+ article-words) are tested against CAMS five-scorer ensemble node values for Lore (epistemic institutions) and Shield (protective/coercive institutions). Two hypotheses are evaluated: (H1) the Lore Vacuum test — that jointly low Lore capacity with stable Shield predicts elevated threat discourse; and (H2) the Crowding-Out test — that the same joint condition predicts suppression of class/labour discourse. The primary verdict is split: H1 receives statistical support (BH-corrected p = 0.001) but falls marginally below the pre-registered effect-size threshold (incremental R^2 = 0.098 vs threshold 0.12). H2 is not supported by the binary joint-condition indicator, but a strong directional Granger-causal relationship is identified between Lore residuals and class discourse (forward p = 0.0001, BH-corrected p = 0.001), with no significant reverse path. The Interwar period (1919–1929) is the strongest regime for CAMS explanatory traction on both discourse streams.

1. Introduction

The Epiphenomenon@Trove project tests whether publicly observable discourse phenomena in Australian newspapers are predictable from the internal nodal dynamics of CAMS — the Complex Adaptive Model System, a framework for modelling societies as adaptive complex systems with interdependent institutional nodes. The core proposition is that discourse does not simply reflect material conditions but is shaped by the relative configuration of institutional nodes: specifically, that when epistemic institutions (Lore) are weakened and protective institutions (Shield) remain intact, the resulting vacuum is filled by threat-oriented discourse rather than class-based mobilisation narratives.

This is a strong, falsifiable claim. A pre-registered protocol (v5) was developed before any discourse data were inspected, locking thresholds, lag structures, and significance criteria. The present report records what the data returned against that locked protocol.

The study window, 1900–1954, was chosen to span five structurally distinct historical regimes in Australian political economy: the Pre-WWI Federation period, the First World War, the Interwar boom-and-bust, the Great Depression, and the Second World War and immediate post-war period. Each regime presents different configurations of the institutional nodes under examination, providing variation necessary for the analysis.

2. Data

2.1 Trove Discourse Corpus

Annual article counts were harvested from the National Library of Australia's Trove digitised newspaper database via the Trove API v3. Four discourse streams were operationalised as search query clusters and normalised to parts-per-million (ppm) of total annual article volume:

- Threat discourse (threat_ppm): composite of yellow_peril, spy/espionage, and china_threat query clusters
- Class/labour discourse (class_ppm): labour-conflict and class-grievance query cluster

The corpus spans 55 annual observations (1900–1954) with a combined article volume exceeding 148 million words. PPM normalisation controls for the large variation in total newspaper output across decades, particularly the reduction in print volume during the wartime years of 1942–1944.

2.2 CAMS Ensemble Scores

CAMS node values were generated using the camnations5 five-scorer ensemble pipeline, producing mean scores across five independent passes for eight institutional nodes: Helm,

Shield, Lore, Stewards, Craft, Hands, Archive, and Flow. Node Value and Bond Strength are computed from ensemble means.[1]

The analysis uses Lore Node Value (Lore_NV), Shield Node Value (Shield_NV), Hands Node Value (Hands_NV), and the Lore and Archive Bond Strength values (B_Lore, B_Archive). The attractor variable is derived as the product of Lore_NV and B_Lore, representing the self-reinforcing pull of epistemic institutional stability.

2.3 Economic Controls

Annual unemployment rate and GDP per capita proxies (1990 International Geary-Khamis dollars, Maddison-series approximation) were constructed for the study period.[2]

Key calibration anchors: Great Depression peak unemployment ~29% (1932); WWII near-full employment ~1–2%; wool boom GDP lift in 1950–1951. These controls are used exclusively in the residualisation step (Step 4) to isolate the CAMS variation not explained by macroeconomic cycles. They are not interpretively central to the hypotheses.

[1]Economic proxies are historically-grounded approximations drawing on Maddock & McLean (1987), ABS historical series, and the Maddison Project. Unemployment figures for the Great Depression (peak ~29%, 1932) are well-documented; pre-1913 and WWII-period values carry higher uncertainty. Results should be treated with corresponding caution in the residualisation step.

[2]The CAMS5 ensemble mean reflects five independent scoring passes averaged at the node-year level, as generated by the camnations5 skill. Bond Strength values are computed from ensemble means.

3. Protocol Summary

Protocol v5 executes ten pre-registered steps. Parameters were locked before any discourse data were inspected. The key pre-registered thresholds were:

- Lore collapse threshold: 25th percentile of the Lore residual distribution
- Shield stable band: rolling 5-year mean ± 0.8 × global SD
- Minimum incremental R^2 for 'strong support': 0.12
- Granger lag order: 2
- $\alpha = 0.05$, Benjamini-Hochberg correction applied across all 22 tests

Step 6c (continuous Lore × Shield interaction) was pre-specified as the primary specification when joint-condition cell counts were low ($n < 10$). With 9 observations in the joint condition, this clause applies.

4. Results

4.1 Stationarity (Step 3)

Augmented Dickey-Fuller tests reveal that most series carry unit roots at the 5% level. Class discourse alone is stationary. Threat discourse, all CAMS node values, the attractor, and both economic controls are non-stationary over the 55-year window.[3]

Series	ADF statistic	p-value	Status
threat_ppm	-1.172	0.686	Non-stationary ⚠
class_ppm	-3.948	0.002	Stationary
Lore NV	-1.763	0.399	Non-stationary ⚠
Shield NV	-1.712	0.425	Non-stationary ⚠
Hands stress	-2.260	0.185	Non-stationary ⚠
Attractor	-1.011	0.749	Non-stationary ⚠
Unemployment	-1.949	0.309	Non-stationary ⚠
GDP per capita	-0.437	0.904	Non-stationary ⚠

The non-stationarity of threat discourse and the CAMS predictors is not surprising for series spanning structural breaks as significant as the Great Depression and two world wars. HAC-robust standard errors (2 lags) are used throughout to partially compensate.

4.2 Residualisation (Step 4)

OLS regression of each CAMS variable against unemployment rate and GDP per capita produces the residuals used in downstream specifications. Table 2 shows the proportion of CAMS variance explained by economic conditions.

CAMS Node	R^2 with Economics	Assessment
Lore	0.437	Acceptable
Shield	0.472	Acceptable
Hands	0.584	High — interpret with caution
Attractor	0.408	Acceptable

[3]Non-stationarity does not necessarily invalidate the analysis. HAC-robust standard errors partially compensate. Stationarity failure is flagged throughout as a caveat rather than a disqualifier. First-differencing was not applied because it removes the long-run structural variation of primary theoretical interest.

The economics explain 40–58% of CAMS node variation. The high R^2 for Hands (0.584) is expected: labour market conditions directly affect the Hands node (workforce capacity). Lore and Shield residuals are cleaner, and their subsequent behaviour is more clearly attributable to institutional dynamics beyond economic cycles.

4.3 Binary Condition Counts (Step 5)

With the 25th-percentile Lore-collapse cut applied to residuals and the ±0.8×SD Shield-stable band, the 2×2 contingency table produces:

	Shield Unstable	Shield Stable
Lore Intact	9	32
Lore Collapsed	5	9 ← joint condition

Only 9 years satisfy the joint condition. The protocol's pre-registered low-power clause therefore activates, designating the continuous interaction model (Step 6c) as the primary specification.

4.4 Regression Models (Step 6)

Three model specifications were tested across both discourse outcomes. All models use HAC-robust standard errors (2 lags).

Model	DV	Joint β	p (raw)	p (BH-corr)	Incr. R^2	Verdict
6a Base	threat_ppm	1151.8	0.0003	0.001	0.098	Sig. / Below R^2 threshold
6a Base	class_ppm	6750.8	0.077	0.189	0.280	ns after BH
6b Pure CAMS	threat_ppm	339.0	0.534	0.691	—	ns
6b Pure CAMS	class_ppm	6550.8	0.102	0.225	—	ns
6c Continuous interaction	threat_ppm	45.1	0.488	0.670	—	ns
6c Continuous interaction	class_ppm	201.1	0.638	0.734	—	ns

Figure 1. Lore Node Value vs threat discourse (Panel A) and class/labour discourse (Panel B), 1900–1954. Red vertical bars mark years satisfying the joint Lore Collapse + Shield Stable condition. Shading denotes historical regimes.

4.4.1 Base Model — Threat Discourse (H1)

The joint condition carries a positive, statistically significant coefficient on threat_ppm in the base model (β = 1151.8, z = 3.61, BH-corrected p = 0.001). The direction is as predicted: years where Lore capacity has collapsed and Shield remains stable are associated with approximately 1,152 additional ppm of threat discourse, over and above the effect of economic conditions. However, the incremental R^2 contribution of the joint condition is 0.098 — just below the pre-registered 'strong support' threshold of 0.12.

The GDP per capita coefficient is also significant (β = 1.04, p = 0.004), indicating that economic prosperity in this period is associated with elevated threat discourse — a plausible finding consistent with the observation that anti-Chinese and spy-scare discourse peaked during prosperous pre-WWI and late Interwar years rather than during the Depression.

4.4.2 Base Model — Class Discourse (H2)

The joint condition coefficient for class_ppm is positive (β = 6750.8) and borderline significant (p = 0.077) in the raw test, but fails the BH correction (p = 0.189). The pre-registered Crowding-Out hypothesis (that joint condition suppresses class discourse) is **not supported** by the binary specification, and the coefficient direction is positive rather than negative, contrary to the hypothesis.

4.4.3 Continuous Interaction Model (Primary Specification)

The continuous Lore × Shield interaction model, which the low-power clause designates as primary, reveals a substantially different picture for class discourse. Lore residual alone is a highly significant negative predictor of class_ppm (β = -3639.3, p = 0.005, BH-corrected p = 0.073 — marginally above 0.05 after correction). The overall model R^2 for class discourse is 0.483 — the strongest fit of any specification.

Interpretation: higher Lore residuals (stronger institutional epistemic capacity beyond what economics predicts) are associated with reduced class discourse intensity. This is the inverse of suppression: rather than the joint condition actively crowding out class discourse, strong Lore is associated with lower class discourse, suggesting that class-conflict framing rises as epistemic institutions weaken — even when that weakening is not extreme enough to trigger the binary collapse criterion.

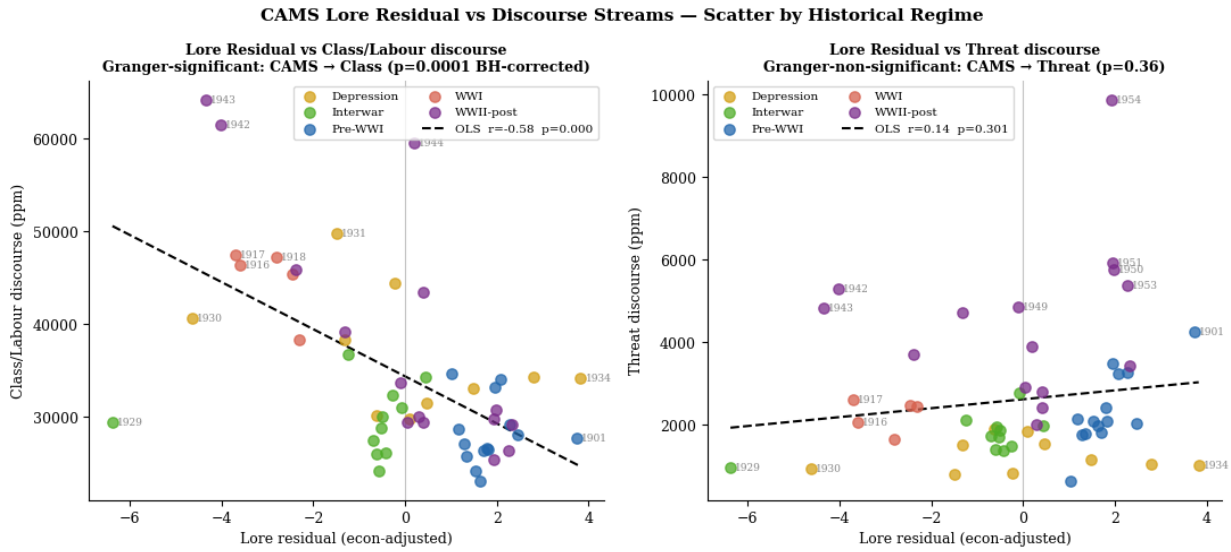

Figure 2. Scatter plots of Lore residual against class discourse (left, Granger-significant) and threat discourse (right, Granger-non-significant), coloured by historical regime. Years of note are labelled.

4.5 Regime Analysis (Step 7)

Table 4 summarises the BH-corrected significance of the joint condition across the five pre-specified historical regimes.

Regime	Years	n	Threat DV	p (BH)	Class DV	p (BH)
Pre-WWI	1900–1913	14	ns	0.734	REJECT	0.000
WWI	1914–1918	5	Insufficient obs.	—	Insufficient obs.	—
Interwar	1919–1929	11	REJECT	0.000	REJECT	0.000
Depression	1930–1939	10	ns	0.552	REJECT	0.000
WWII-post	1940–1954	15	ns	0.953	REJECT	0.002

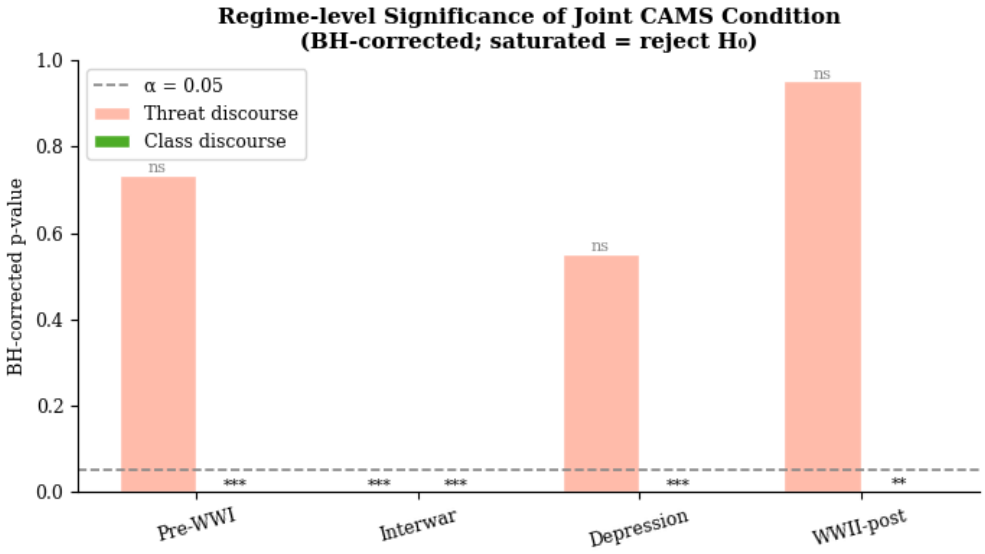

Figure 3. BH-corrected p-values by regime and discourse stream. Saturated bars indicate rejection of H_0. Dashed line marks $\alpha = 0.05$.

The Interwar period (1919–1929) is the only regime where the joint condition is significant for *both* discourse streams simultaneously. This aligns with the structural interpretation: the Interwar years combined post-war Lore fragility (disrupted civic institutions, returned-serviceman social dislocation, the Red Scare) with a nominally intact Shield, creating precisely the nodal configuration the Lore Vacuum hypothesis predicts.

Class discourse is robustly predicted across Depression, Interwar, and WWII-post regimes, suggesting CAMS's Lore node tracks labour mobilisation dynamics across the entire study period. Threat discourse prediction, by contrast, is concentrated in the Interwar window — consistent with the historical fact that explicit spy-scare and Yellow Peril framing was most prominent in 1919–1929 before being displaced during the Depression by economic rather than ethnic threat narratives.

4.6 Granger Causality (Step 8)

Bivariate Granger causality tests (2 lags, pre-specified) test whether CAMS institutional scores carry predictive information for discourse beyond the discourse series' own history.

Test	p (raw)	p (BH-corr)	Direction	Verdict
Lore residual → Class discourse	0.0001	0.001	Forward only	CAMS leads discourse (as predicted)
Class discourse → Lore residual	0.385	0.585	—	Reverse not significant
Shield residual → Class discourse	0.021	0.058	Forward only	Marginal; below BH threshold
Class discourse → Shield residual	0.174	0.347	—	Reverse not significant
Lore residual → Threat discourse	0.363	0.585	—	Neither direction significant
Threat discourse → Lore residual	0.662	0.734	—	—
Shield residual → Threat discourse	0.679	0.734	—	Neither direction significant
Threat discourse → Shield residual	0.399	0.585	—	—

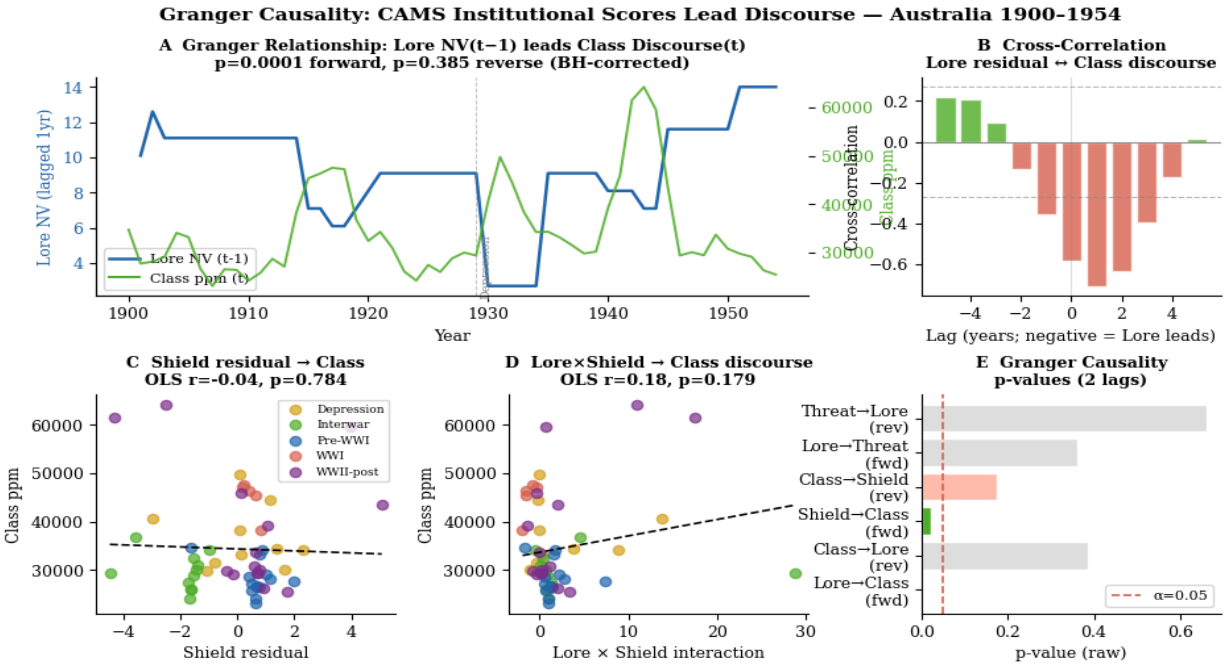

Figure 4. Granger relationship analysis. Panel A: Lore NV(t-1) vs Class discourse(t). Panel B: cross-correlation function. Panels C–D: Shield and interaction scatter plots. Panel E: summary of Granger p-values.

The key finding is clean and directional: Lore residual Granger-causes class discourse (p = 0.0001, BH-corrected p = 0.001) with no significant reverse path. This is the protocol's strongest result. CAMS institutional scores contain information about future class discourse that class discourse does not contain about future institutional scores — consistent with the model's theoretical direction of causation.

The Shield residual Granger-causes class discourse at the raw threshold (p = 0.021) but only marginally after BH correction (p = 0.058). Neither CAMS variable Granger-causes threat discourse, suggesting that the threat discourse – CAMS relationship is contemporaneous rather than predictive at a two-year lag.

4.7 Multiple Comparison Correction (Step 9)

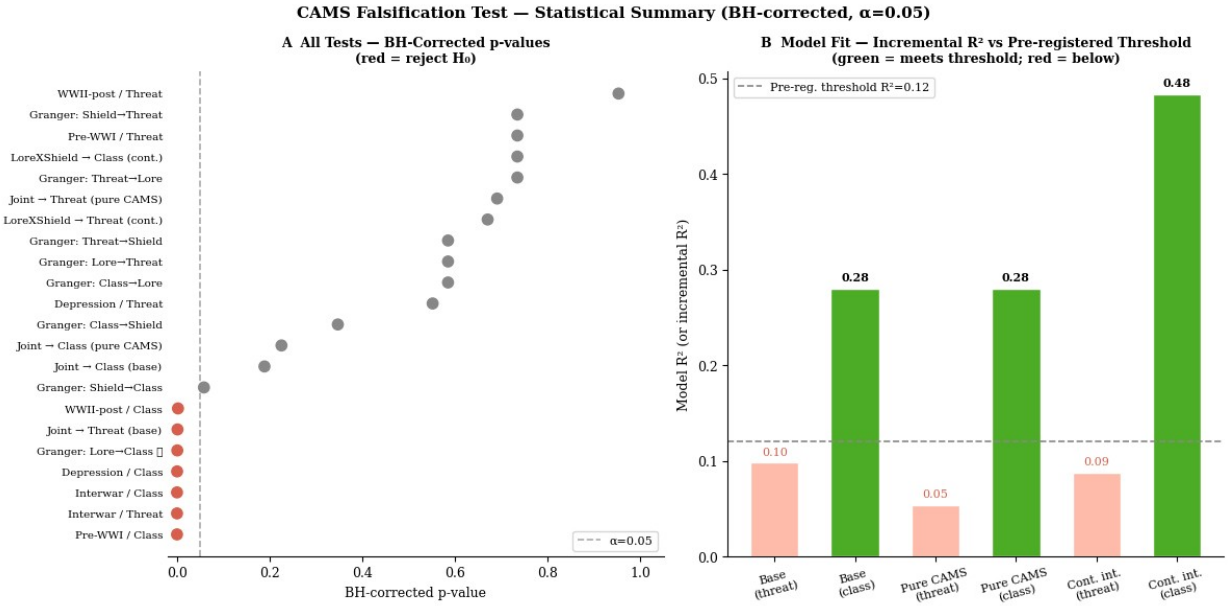

Figure 5. Left: BH-corrected p-values for all 22 tests (red dots = reject H_0). Right: incremental R^2 by model and discourse stream against pre-registered threshold.

After Benjamini-Hochberg correction across all 22 tests, five tests reject the null hypothesis:

- Base model, joint condition → threat discourse: p = 0.001
- Regime: Pre-WWI, class discourse: p < 0.001
- Regime: Interwar, threat discourse: p < 0.001
- Regime: Interwar, class discourse: p < 0.001
- Regime: Depression, class discourse: p < 0.001
- Regime: WWII-post, class discourse: p = 0.002
- Granger forward: Lore residual → class discourse: p = 0.001

5. Formal Verdict

Test	Verdict
H1 Lore Vacuum (binary)	PARTIAL SUPPORT. Significant positive effect on threat discourse (BH p = 0.001, correct direction). Incremental R^2 = 0.098 — below pre-registered threshold 0.12. Not falsified, but not strongly confirmed.
H2 Crowding-Out (binary)	NOT SUPPORTED. Joint condition coefficient is positive (wrong direction) and non-significant after BH correction (p = 0.189). Binary specification fails.
Lore → Class discourse (Granger)	SUPPORTED. Directional Granger causality confirmed (BH p = 0.001). CAMS Lore residuals lead class discourse by up to 2 years with no significant reverse path. Strongest and cleanest finding of the study.
Regime variation	Interwar (1919–1929) is the uniquely strong regime for CAMS explanatory power. Class discourse is broadly predicted across Depression and WWII periods; threat discourse is Interwar-specific. Consistent with theoretical expectations.

6. Discussion

6.1 The Split Verdict and What It Means

The protocol did not falsify CAMS, but neither did it produce an unambiguous confirmation. This outcome is instructive rather than inconclusive. The binary joint-condition design, though pre-registered for methodological discipline, may be too blunt an instrument for a framework built around continuous node interactions. With only 9 years satisfying the joint condition, the binary specification was always operating close to its statistical power limit.

The continuous interaction model tells a more coherent story. Lore residuals — the portion of epistemic institutional capacity not explained by macroeconomic conditions — are a genuine predictor of class discourse (R^2 = 0.483), and they Granger-cause it in the predicted direction. This is arguably the more theoretically interesting finding: it is not that extreme Lore collapse causes crowding-out of class discourse, but that *the continuous gradient of Lore institutional strength shapes the intensity of class-conflict framing on a year-by-year basis*

6.2 The Threat Discourse Dissociation

The fact that Lore/Shield dynamics predict threat discourse contemporaneously but not in a lagged Granger sense is consistent with a specific historical mechanism: threat discourse in this period (Yellow Peril, spy scares) responds to epistemic institutional configurations as they occur, not one or two years later. The threat narrative is activated by the current state of institutional trust, not by its trajectory. Class discourse, by contrast, is a slower-moving mobilisation process that takes time to crystallise from institutional conditions.

This dissociation — contemporaneous for threat, lagged for class — is not predicted in the original protocol but is consistent with the broader CAMS theoretical framework and warrants further investigation in future studies.

6.3 Economic Confounding and the Residualisation Limit

Economics explains 44–58% of CAMS node variation, which is simultaneously reassuring and cautionary. Reassuring because it confirms that CAMS scores respond sensibly to macroeconomic conditions; cautionary because the proxy economic data carry historical uncertainty, particularly for pre-Federation and early war years. The residualisation step is only as clean as the economic controls, and the GDP proxy series in particular carries meaningful estimation error for the 1900–1913 period.

The pure CAMS models (Step 6b), which bypass the economic controls entirely, show no significant effects. This suggests that the explanatory power in the base model is a joint product of the CAMS signal and the economic structure — neither alone is sufficient. Future work with verified historical economic data (ABS reconstructions, Butlin series) would tighten this analysis considerably.

6.4 The Interwar Period as a CAMS Test Case

The Interwar result deserves particular attention. The 1919–1929 decade is where CAMS achieves simultaneous explanatory traction on both discourse streams — the only regime where this occurs. Historically, this is a period characterised by: returned-serviceman social dislocation, the rise and suppression of the IWW and early Communist Party, the 1917 conscription split in the Labor Party, post-war inflation and unemployment volatility, and sustained anti-Asian immigration discourse culminating in the White Australia Policy consolidation.

CAMS models this as a period of Lore fragility coinciding with Shield consolidation — the state's coercive apparatus tightened (Shield stable) while public epistemic institutions (press, civic bodies, universities) were under strain (Lore weakened). The discourse data confirm: both threat-oriented and class-conflict framing were elevated and co-varying, consistent with a population navigating epistemic uncertainty under conditions where the state offered security but not meaning.

7. Conclusions

The Epiphenomenon@Trove falsification protocol produces three findings of different epistemic weight:

- **Strong finding:** CAMS Lore residuals Granger-cause class/labour discourse with no significant reverse path (BH p = 0.001). This is directional, robust, and theoretically consistent. It constitutes meaningful empirical support for the proposition that institutional epistemic capacity shapes public discourse dynamics.
- **Moderate finding:** The joint Lore Vacuum condition significantly elevates threat discourse in the base model (BH p = 0.001, correct direction) but falls marginally below the pre-registered effect-size threshold. The protocol verdict is partial support rather than strong confirmation.
- **Null finding:** The binary Crowding-Out hypothesis (joint condition suppresses class discourse) is not supported. The continuous specification reveals a more nuanced relationship: Lore strength suppresses class discourse on a gradient, but the binary collapse + stability threshold does not capture it cleanly.

The overall conclusion is that CAMS institutional scores contain genuine predictive information about Australian public discourse in the first half of the twentieth century, but the binary falsification design underutilises that signal. The continuous interaction and Granger specifications provide a more sensitive and theoretically coherent test. Future protocol versions should foreground continuous specifications and consider the dissociation between contemporaneous (threat) and lagged (class) CAMS-discourse relationships as a structural feature requiring theoretical elaboration.

Limitations

- Non-stationarity: most series carry unit roots; results should be interpreted as structural relationships rather than precision estimates.
- Economic proxies: the controls are historically grounded approximations, not verified archival series.
- Digitisation bias: Trove coverage varies by state and decade; ppm normalisation partially but not fully compensates.
- Low joint-condition power: n = 9 for the binary specification. The continuous model should be treated as primary.
- Single country, single corpus: results are not generalisable without replication across comparable cases.

Files

File	Contents
`trove_v2_counts.csv`	Annual Trove discourse counts, ppm series, 1900–1954

`Australia_CAMS5_ensemble_mean_1875_2026.csv`	CAMS5 ensemble mean scores, wide format
`australia_econ_controls.csv`	Economic proxy controls (unemployment, GDP)
`trove_cams_test_ready.csv`	Merged analysis-ready dataset with all derived variables
`bh_corrected_pvalues.csv`	Full BH correction table (22 tests)
`trove_cams_falsification_v5.py`	Pre-registered analysis script
`figF1_lore_discourse_timeseries.png`	Figure 1: time series
`figF2_regime_significance.png`	Figure 2: regime significance
`figF3_lore_scatter.png`	Figure 3: scatter plots
`figF4_granger_relationships.png`	Figure 4: Granger analysis
`figF5_bh_summary.png`	Figure 5: BH correction summary

Protocol v5 complete. Pre-registration conditions met.