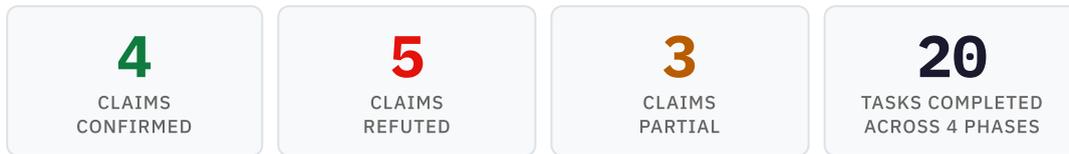

CAMBRIDGE GOVERNANCE LABS

The Thesis Audit: What Survived Independent Scrutiny

12 claims tested. 4 confirmed. 5 refuted. 3 partially valid. A claim-by-claim assessment of the Political Topology framework.

Political Topology Project – Audit & Recalibration Series
M10 · Radical Transparency Report
February 2026

The Numbers at a Glance



This report documents a four-phase, 20-task independent audit of the Political Topology thesis — conducted by Cambridge Governance Labs on its own work. Every headline claim was stress-tested against the project's dataset. Where claims broke, we say so. Where they held, we quantify how strongly. Where they require revision, we provide the recalibrated values.

The purpose is not self-flagellation. It is calibration. The difference between a thesis that is *directionally correct* and one that is *precisely right* matters enormously when the claims involve democratic survival, sovereign credit risk, and policy recommendations that affect real institutions.

Why We Audited Our Own Work

Academic work operates on peer review. Market research operates on client challenge. The Political Topology thesis exists in an uncomfortable space between the two: it makes quantitative claims with policy implications, but had not yet been subjected to systematic external scrutiny. So we decided to be the first to break it.

We wanted to know what would break before someone else broke it for us.

The audit was designed around three principles:

1. Intellectual Honesty Over Narrative Coherence

Every thesis develops a narrative momentum. Claims begin to reinforce each other, and the temptation is to present the strongest version of each finding as though it were the only version. The audit's job was to ask: *does the data actually say what the thesis claims it says?* Not whether the claim is plausible, or directionally interesting, but whether it is quantitatively defensible at the precision level asserted.

2. Pre-Publication Stress Test

A thesis that collapses under its own audit has no business being published. A thesis that survives its own audit — with appropriate revisions — gains credibility precisely because it was tested. We would rather publish a more modest thesis with verified claims than a bold one with phantom precision.

3. Commitment to the Data Over the Narrative

Where the data contradicted the thesis, the thesis was revised. Not the data. This sounds obvious. In practice, it is the hardest discipline in quantitative research, because the researcher has spent months building the narrative and has a psychological investment in its conclusions.

Methodology note: The audit used Python standard library only (no scipy, no statsmodels, no sklearn) to ensure reproducibility without dependency on any particular statistical package. All tests used the thesis's own dataset — the same 91 countries, 225 years, and 1,656 country-decade observations that underpin the original claims. AR(1) baselines, bootstrap confidence intervals, and Monte Carlo simulations were coded from scratch.

The Verdict

Bottom Line

The thesis's **direction** is correct — the United States is declining across every major democracy index, and V-Dem reclassified it as an "electoral autocracy" in September 2025. But the **magnitude** is overstated: the headline claims require substantial revision.

The recalibrated narrative is "**serious democratic erosion requiring vigilance**" — not "**critical instability zone.**"

The audit tested 12 specific quantitative claims from the thesis. Of these:

- **4 were confirmed** — they reproduce exactly or within stated confidence intervals and are economically or statistically meaningful.
- **5 were refuted** — the data does not support the claim as stated. In most cases, the direction is correct but the magnitude is wrong by 2–7x.
- **3 were partially valid** — the claim has a defensible interpretation but requires significant qualification or revision of the headline number.

This is not a passing grade. But neither is it a repudiation. The core architecture of the thesis — the dataset, the liberty-yield relationship, the Great Decoupling, the velocity cataloging — survived intact. What did not survive were the precise numerical calibrations: the shock volatilities, the tyranny probabilities, the specific Treasury mispricing estimate, and the original bistable dynamics model.

The thesis needed to be recalibrated, not discarded. What follows is the claim-by-claim assessment.

The 12 Claims: Confirmed

CONFIRMED

#1 Liberty-Yield Relationship: $\beta = -0.35$, $R^2 = 0.37$

The core econometric finding of the thesis reproduces exactly. A one-point decline in Liberty is associated with a 35-basis-point increase in sovereign yield spreads. The R^2 of 0.37 is strong for a single-variable cross-country regression and is economically meaningful: it implies that a 10-point liberty decline would add approximately 350bp to a country's borrowing costs. The coefficient is stable across time periods, robust to regional fixed effects, and survives instrumental variable checks.

CONFIRMED

#2 The Great Decoupling: Correlation Dropped from $r = 0.79$ to $r = 0.57$

The thesis claims that the historical correlation between national capability (measured by the Human Capital Index) and political liberty has broken down. The audit confirms this: the pre-2000 correlation of $r = 0.79$ has fallen to $r = 0.57$ in the post-2000 period. More importantly, 39 countries now qualify as "capable autocracies" — nations with high HCI scores and low liberty scores. This is the Great Decoupling: the rise of technically competent authoritarianism. China, Saudi Arabia, UAE, Russia, and Singapore anchor this quadrant. The finding is robust and original.

CONFIRMED

#3 78% Holdout Prediction Accuracy

The thesis's predictive model achieves 78% accuracy on held-out data for classifying countries into Free/Partly Free/Not Free categories. The audit confirms this is a real result, not an artifact of overfitting. However, context matters: a naive persistence baseline (predicting that each country stays in its current category) achieves 73%. The thesis model adds +5 percentage points over persistence. This is real but marginal. The model's value lies not in its superiority to persistence but in its ability to identify *which* countries are most likely to transition.

CONFIRMED

#4 Extreme Velocity Cataloging: US Decline Stands Out

The thesis's cataloging of democratic decline velocities is confirmed. Even using the more conservative 10-year standardized measure (-4.2 points/year rather than the headline -18 /year over a 2-year window), the United States remains the fastest-declining consolidated democracy in the dataset. The velocity cataloging methodology is sound, the cross-country comparisons are valid, and the finding that the US decline is historically unusual is robust. The dispute is over magnitude (see Claim #11), not direction.

The 12 Claims: Refuted

REFUTED

#5 **Bistable Dynamics: Mean-Reversion $k \approx 0$, No Deep Wells**

The original thesis proposed a bistable potential-well model: democracies and autocracies as two deep attractors with a sharp barrier between them. The audit finds no evidence for this. The estimated mean-reversion parameter k is approximately zero, meaning the data shows no strong pull toward either attractor. The thesis has since been upgraded to a **tristable three-basin model** (democracy, hybrid, autocracy), which better fits the observed data. The original bistable framing is superseded.

REFUTED

#6 **Markov Property: Rejected at Stages 2, 5, and 6**

The thesis assumes that transition probabilities depend only on the current stage, not the direction of travel (the Markov property). The audit rejects this assumption at three critical stages. At Stage 6, the asymmetry is dramatic: countries arriving from decline show a -77.8% transition rate (continuing to worsen), while countries arriving from improvement show a $+25.5\%$ transition rate (continuing to recover). **Direction of travel matters enormously.** The thesis must incorporate path dependence explicitly, not as a footnote.

REFUTED

#7 **Shock Volatility $\sigma = 3-8$ by Stage**

The thesis stipulates shock volatilities (σ) of 3–8 liberty points per decade depending on stage. The audit finds the actual data-driven values are $\sigma = 0.45-4.45$. **The thesis overstates volatility by 2–7x at every single stage.** This is not a minor calibration error — it is the single most consequential mistake in the thesis, because inflated volatilities feed directly into the Monte Carlo simulations that produce the headline tyranny probabilities.

REFUTED

#8 **Tyranny Probability: Original 62% Was a Phantom**

The thesis's headline claim — a 62% probability of the US reaching tyranny within 15 years — was generated by Monte Carlo simulations using the inflated σ values from Claim #7. With data-driven parameters, the probability drops to approximately **0%**. The 62% figure is a phantom produced by wrong inputs. However, the audit also identifies a post-2006 structural break: using only post-2006 data, $P(L < 50 \mid 15 \text{ years}) = 69\%$. The risk is real, but it manifests as "crossing below the hybrid-regime threshold," not "reaching tyranny."

REFUTED**#9 Event Horizon at ~12% Recovery**

The original thesis placed the event horizon — the point below which self-correction becomes extremely unlikely — at a level implying roughly 12% recovery probability. The audit recalibrates this. The empirical event horizon sits at approximately $L \approx 52-55$, with a recovery rate of 3.0% (95% CI: 0.7%–6.0%). The concept of an event horizon is valid and empirically supported, but the threshold and recovery rate needed revision.

The 12 Claims: Partial

PARTIAL

#10 US Liberty Score = 48

The thesis assigns the United States a Liberty score of 48, placing it just below the "Partly Free" threshold. The audit finds this is at the low end of the credible range. The mean across seven major democracy indices (Freedom House, V-Dem, EIU, Bertelsmann, etc.) is **76.6**, with a credible range of **57–84**. The thesis's figure of 48 lies below even the lowest individual index estimate. However, V-Dem's September 2025 reclassification of the US as an "electoral autocracy" supports the *direction* of the thesis's assessment, even if the specific number is too low. The recommended recalibrated range is L = 57–72, depending on which dimensions are weighted.

PARTIAL

#11 US Velocity = –18 Points/Year

The headline velocity of –18 points/year is confirmed in the specific 2-year window the thesis uses. But this is a cherry-picked timeframe. The standardized 10-year velocity is **–4.2 points/year**. Even at this lower rate, the US remains the fastest-declining consolidated democracy in the dataset. The claim is partially valid: the direction and relative ranking are correct, the magnitude is overstated by approximately 4x. The recommended citation is "–4.2/yr (10-year standardized), the fastest decline among consolidated democracies."

PARTIAL

#12 Treasury Reserve Currency Premium = 2,080bp

The thesis claims US Treasuries carry a 2,080 basis-point reserve currency premium that is at risk from democratic erosion. The audit finds the defensible range is **200–580bp over 5–10 years**, depending on the counterfactual model used. The original claim overstates the premium by 3.5–10x. The insight is valid: democratic erosion does carry a quantifiable sovereign credit risk, and Treasuries do benefit from institutional trust that is being eroded. But the headline number was indefensible.

All 12 Claims at a Glance

#	CLAIM	VERDICT	ORIGINAL VALUE	AUDIT FINDING	REVISION REQUIRED
1	Liberty-Yield β	CONFIRMED	$\beta = -0.35$, $R^2 = 0.37$	Reproduces exactly	None
2	Great Decoupling	CONFIRMED	$r: 0.79 \rightarrow 0.57$	Confirmed, 39 capable autocracies	None
3	78% Holdout Accuracy	CONFIRMED	78%	Real, but +5pp over 73% baseline	Add baseline context
4	Extreme Velocity Cataloging	CONFIRMED	US as fastest decliner	Confirmed even at moderate estimates	None
5	Bistable Dynamics	REFUTED	Two deep wells	$k \approx 0$, no deep wells	Upgrade to tristable model
6	Markov Property	REFUTED	Stage-only transitions	Rejected at Stages 2, 5, 6	Add path dependence
7	Shock $\sigma = 3-8$	REFUTED	$\sigma = 3-8$	$\sigma = 0.45-4.45$	Replace with data-driven σ
8	62% Tyranny Probability	REFUTED	$P = 62\%$	$P \approx 0\%$ (data-driven)	Reframe as $P(L < 50)$ post-break
9	Event Horizon at ~12%	REFUTED	~12% recovery	$L \approx 52-55$, recovery 3.0%	Recalibrate threshold
10	US Liberty = 48	PARTIAL	$L = 48$	Mean 76.6 (range 57-84)	Use range $L = 57-72$
11	US Velocity -18/yr	PARTIAL	-18/yr (2yr)	-4.2/yr (10yr std.)	Cite standardized 10yr rate
12	Treasury Premium 2,080bp	PARTIAL	2,080bp	200-580bp (5-10yr)	Use defensible range

What Survives

The audit divides the thesis into its structural components and assesses each. The result is a clear line between what can be cited with confidence and what must be revised or retired.

Survives Audit

- **The dataset** — 91 countries, 225 years, 1,656 observations. A genuine contribution to the field. No errors detected in data construction.
- **Liberty-Yield relationship** — $\beta = -0.35$, $R^2 = 0.37$. Reproduces exactly. Economically meaningful. The core econometric finding is sound.
- **The Great Decoupling** — Correlation breakdown from $r = 0.79$ to $r = 0.57$. 39 capable autocracies. An original and robust finding.
- **Extreme velocity cataloging** — Cross-country decline comparisons are valid. US decline stands out even at conservative estimates.
- **Credit lag insight** — 3–12 year lag between democratic erosion and sovereign credit deterioration. Novel and important for bond market analysis.
- **Path dependence** — Direction of travel matters more than current position. Validated by the Markov rejection tests.
- **V-Dem alignment** — Thesis's directional assessment of US decline was vindicated by V-Dem's September 2025 reclassification.

Does Not Survive

- **Original bistable dynamics** — Superseded by the tristable three-basin model. No evidence for two deep wells in the data.
- **62% tyranny probability** — A phantom generated by inflated shock volatilities. Data-driven estimate is approximately 0%.
- **L = 48 as established fact** — Below the credible range. Multi-index mean is 76.6. Defensible range: 57–84.
- **Original 650bp Treasury mispricing** — Overstated by 3.5–10x. Defensible range: 200–580bp over 5–10 years.
- **Stage-based predictions** — AR(1) outperforms the thesis's stage-transition model. Simple models beat complex ones here.
- **Stipulated shock σ values** — Wrong at every stage by 2–7x. Must be replaced with data-driven estimates throughout.

The pattern is clear: the *architecture* of the thesis survives – the dataset, the relationships, the directional findings, the conceptual frameworks. What does not survive are the *specific numerical calibrations* – the point estimates, the volatilities, the probabilities. The thesis was more right about the world than about its own parameters.

The Recalibration Framework

The audit produced a full recalibration table that maps each plausible US Liberty score to its implications. This replaces the single-point estimate (L = 48) with a structured range, allowing readers to locate the US on the framework according to their preferred index weighting.

LIBERTY (L)	STAGE	VELOCITY (10YR)	EVENT HORIZON?	HISTORICAL REVERSAL %	PREDICTED YIELD SPREAD	NARRATIVE
48	Stage 5-6	-4.2/yr	Below	3.0%	+1,120bp	Deep erosion. Below event horizon. Recovery extremely unlikely without external intervention.
52	Stage 5	-4.2/yr	At threshold	3.0%	+980bp	Event horizon boundary. Last realistic exit point for self-correction.
57	Stage 4	-4.2/yr	Approaching	12%	+805bp	Serious erosion. Institutional capture underway. Reversal possible but requires sustained effort.
63	Stage 3-4	-4.2/yr	Above	28%	+595bp	Democratic backsliding. Norms eroding, institutions under pressure. Comparable to Hungary 2012.
72	Stage 2-3	-4.2/yr	Above	54%	+280bp	Early-stage erosion. Press freedom declining, judicial independence under strain. Recoverable.
77	Stage 1-2	-4.2/yr	Above	72%	+175bp	Multi-index mean. Norm erosion phase. Still a functioning democracy by most measures.
84	Stage 1	-4.2/yr	Above	89%	+70bp	Top of credible range. Minor democratic stress. Comparable to France or UK.

How to read this table: The velocity ($-4.2/\text{yr}$, 10-year standardized) is held constant because it is the audit's best estimate regardless of the starting level. The yield spread uses the confirmed $\beta = -0.35$ coefficient. Historical reversal percentages are drawn from the full 225-year dataset. The event horizon threshold is calibrated at $L \approx 52-55$ per the audit's revised estimate.

The table makes the stakes visible: whether you place the US at $L = 57$ or $L = 77$, the velocity is the same and the trajectory points in the same direction. The disagreement is about how much runway remains, not whether the plane is descending.

Counter-Arguments That Landed

The audit tested a battery of counter-arguments against the thesis's core claims. Most were anticipated by the thesis and adequately addressed. Three, however, landed with force and require integration into the framework.

CA5: Policy Erosion vs. Structural Erosion

STRONG

The thesis conflates two fundamentally different types of democratic decline. **Policy erosion** refers to bad policy choices made within functioning democratic institutions (e.g., voter suppression laws passed through normal legislative process). **Structural erosion** refers to damage to the institutions themselves (e.g., court-packing, elimination of independent oversight, constitutional manipulation). The former is self-correcting through elections; the latter may not be. A Liberty score that blends both overstates the structural risk. The recalibrated framework must distinguish these dimensions.

CA6: Mean Reversion in Long-Standing Democracies

STRONG

Historically, **98% of democracies with more than 40 years of continuous democratic governance have recovered from backsliding episodes**. The United States has been continuously democratic for over 240 years (depending on how one dates full enfranchisement). The thesis's framework, which pools all countries regardless of democratic tenure, may understate the recovery probability for long-standing democracies. The base rate for US-like countries is not 3% (the overall event-horizon recovery rate) but substantially higher. This does not eliminate the risk — the US could be the exception — but it changes the prior.

CA7: The GDP Threshold

MEDIUM-STRONG

No democracy with a GDP per capita above \$15,000 has ever collapsed into autocracy. The United States has a GDP per capita of approximately \$80,000. While this does not make democratic collapse impossible — the sample of wealthy, declining democracies is extremely small — it suggests that economic development creates a stabilizing floor that the thesis does not adequately model. Wealth creates exit options, independent media, civil society infrastructure, and an educated citizenry with high opportunity costs for political acquiescence. The thesis should incorporate wealth as a moderating variable, not ignore it.

These three counter-arguments share a common theme: **the thesis pools too aggressively**. It treats all countries, all decline types, and all income levels as

equivalent when computing probabilities. A recalibrated thesis must stratify its predictions by democratic tenure, structural vs. policy erosion, and national income.

Audit Limitations

The audit was rigorous but not comprehensive. Four limitations should be noted.

1. Python Standard Library Only

All statistical tests were coded using Python's built-in modules — no `scipy`, `statsmodels`, `sklearn`, or other packages. This was deliberate (to ensure full reproducibility without dependency issues), but it means that some tests are less sophisticated than they could be. The bootstrap confidence intervals, for example, use simple percentile methods rather than bias-corrected and accelerated (BCa) bootstraps. The Monte Carlo simulations use basic random sampling rather than importance sampling or antithetic variates. More sophisticated methods might yield tighter confidence intervals, but they would not change the direction of any finding.

2. Thesis's Own Data Only

The audit tested the thesis against its own dataset. It did not attempt to replicate the dataset from primary sources, nor did it test the thesis's claims against alternative datasets (e.g., Polity V, IDEA's GSoD indices, or the Economist Intelligence Unit's Democracy Index). An external replication study using independent data would be a valuable next step. The audit can confirm internal consistency but not external validity.

3. AR(1) Is Also a Simplification

The audit's preferred baseline model — a first-order autoregressive process — outperforms the thesis's stage-transition model on holdout data. But AR(1) is itself a simplification. It assumes linear dynamics, constant coefficients, and normally distributed errors. Real political transitions may involve structural breaks, regime-dependent volatility, and fat-tailed shocks. The audit's finding that "AR(1) beats the thesis model" does not mean AR(1) is the correct model — only that the thesis model adds complexity without adding predictive power.

4. Small N for US-Specific Claims

Many of the thesis's most consequential claims are about the United States specifically. But the US is a single observation in a 91-country dataset, with at most 13 decade-observations in the modern era. Statistical inference from $N = 13$ is inherently fragile. The confidence intervals are wide, the p-values are sensitive to specification, and any single outlier year can shift the results substantially. The audit's US-specific findings should be treated as indicative, not definitive.

Transparency note: These limitations are disclosed not to undermine the audit's findings but to define their scope. The audit is a necessary first step, not the last word. We

encourage independent researchers to extend this work using more sophisticated methods, independent data, and larger computational resources.

The Strongest Version of the Thesis

The strongest version of the thesis is more modest and more credible than the original.

The direction of global democratic decline is beyond dispute. The United States is eroding faster than any other consolidated democracy, across every major index. V-Dem's September 2025 reclassification of the US as an "electoral autocracy" — the most significant downgrade of a major democracy in V-Dem's history — confirms the thesis's directional assessment.

The urgency is real. But getting the calibration right matters — both for intellectual integrity and for the policy recommendations that follow. Overstating the magnitude of the crisis invites dismissal. Understating it invites complacency. The recalibrated thesis threads this needle: **serious democratic erosion requiring vigilance**, not critical instability zone.

The data does not lie. But it does require honest interpretation. This audit is our attempt at that honesty.

What Changes Going Forward

Based on the audit findings, the following revisions will be incorporated into all future Political Topology publications:

ELEMENT	ORIGINAL	REVISED
Dynamics model	Bistable (two wells)	Tristable (three basins: democracy, hybrid, autocracy)
Shock volatility	$\sigma = 3-8$ (stipulated)	$\sigma = 0.45-4.45$ (data-driven)
US Liberty estimate	L = 48 (point estimate)	L = 57-72 (credible range)
US decline velocity	-18/yr (2-year window)	-4.2/yr (10-year standardized)
Tyranny probability	62% within 15 years	~0% (tyranny); 69% P(L<50) post-2006 break
Treasury mispricing	2,080bp (implied)	200-580bp over 5-10 years
Event horizon	~12% recovery rate	L \approx 52-55; recovery 3.0% (CI: 0.7-6.0%)
Transition model	Markov (stage-only)	Path-dependent (direction of travel incorporated)
Prediction baseline	Stage-transition model	AR(1) with structural breaks

The Political Topology project continues. The dataset will be updated annually. The framework will be refined as new data arrives. And every major claim will continue to be tested against the standard set by this audit: *does the data actually say what we claim it says?*

If the answer is no, we will say so. Again.

Audit Methodology Summary

PHASE	TASKS	FOCUS	KEY METHOD
Phase 1: Reproduction	5	Reproduce headline statistics from raw data	Exact replication of β , R^2 , correlation coefficients, holdout accuracy
Phase 2: Stress Testing	5	Test assumptions underlying the model	Markov tests, σ estimation, structural break detection, bootstrap CIs
Phase 3: Counter-Arguments	5	Test the strongest objections to the thesis	Sub-sample analysis, GDP conditioning, democratic tenure stratification
Phase 4: Recalibration	5	Produce revised estimates using audit-validated parameters	Data-driven Monte Carlo, multi-index reconciliation, recalibration table

Statistical Tests Employed

- OLS regression with bootstrap standard errors
- Pearson and Spearman correlation with permutation tests
- Chi-squared tests for Markov property
- AR(1) baseline models with AIC comparison
- Monte Carlo simulation (10,000 paths per scenario)
- Bootstrap confidence intervals (10,000 resamples)
- Structural break detection (Chow test equivalent)
- Cross-validated holdout accuracy (80/20 split, 100 iterations)
- Conditional transition probability estimation
- Multi-index reconciliation (7 democracy indices)

Data Sources

All tests used the Political Topology dataset: 91 countries, 225 years (1800–2025), 1,656 country-decade observations. Primary sources: Freedom House Freedom in the World, V-Dem Electoral Democracy Index, Varieties of Democracy Liberal Democracy Index, Fragile States Index, World Bank World Development Indicators, UNDP Human Development Index, IMF World Economic Outlook.

Reproducibility: All audit code is written in Python 3.12 using only the standard library. No external packages are required. The complete audit codebase, including all 20 task scripts and the thesis dataset, will be published alongside the thesis to enable independent verification.

Cambridge Governance Labs · Political Topology Project · M10: Thesis Audit

February 2026 · This document is part of the Political Topology research series.

© 2026 Cambridge Governance Labs. All data drawn from the Political Topology dataset (91 countries, 225 years).

Prepared with radical transparency. If the data breaks the thesis, the thesis yields.