

CAMBRIDGE GOVERNANCE LABS

Path Dependence and the Markov Fallacy in Regime Transition Models

*Why the Direction of Travel Dominates the Current State in Predicting Democratic
Outcomes*

Nicholas Gogerty

Cambridge Governance Labs

Political Topology Project

Corresponding author: ngogerty@cambridgegovernancelabs.com

February 2026

WORKING PAPER – A06

ABSTRACT

Regime transition models in political science commonly assume the Markov property—that the probability of future state transitions depends only on a country's current regime classification, not on how it arrived there. This assumption undergirds the Polity project, Freedom House forecasting, Economist Intelligence Unit projections, and virtually all Markov chain applications to democratization. We test this assumption using 1,565 observed transitions across 91 countries over 225 years (1800–2025). Using chi-square tests of conditional transition probabilities stratified by direction of prior movement, we reject the Markov property at three critical stages of democratic erosion: Stage 2 (Early Warning, liberty scores 80–84), Stage 5 (Electoral Autocracy, 50–59), and Stage 6 (Soft Dictatorship, 40–49). At Stage 6, the asymmetry is dramatic: countries arriving via decline show a recovery probability of -77.8% relative to baseline, while those arriving via improvement show $+25.5\%$. Direction of travel—whether a country is declining or recovering—carries more predictive power than the current liberty score alone. A simple AR(1) model incorporating momentum ($\beta = 0.96$) outperforms all stage-based Markov models by $\Delta AIC > 300$. These findings have fundamental implications for forecasting models that treat regime type as a sufficient statistic. We propose an extended state-space model that incorporates momentum as a first-class variable and demonstrate its superior out-of-sample performance.

Keywords: path dependence, Markov property, regime transitions, democratic erosion, political forecasting, state-space models, autocratization, democratic backsliding, transition matrices

JEL Codes: C22, C25, D72, P16

1. Introduction

When political scientists model regime change, they almost universally rely on a simplifying assumption: that the probability of a country transitioning from one regime type to another depends only on its current classification, not on its recent trajectory. A country classified as a "hybrid regime" is assumed to face the same transition probabilities regardless of whether it arrived at that classification by declining from consolidated democracy or by improving from authoritarian rule. This is the Markov property, and it is embedded in the mathematical foundations of nearly every quantitative regime forecasting system in active use.

The appeal of the Markov assumption is considerable. It reduces the dimensionality of the state space, makes transition matrices estimable from modest sample sizes, and permits elegant mathematical analysis including stationary distribution calculations and mean first-passage times. The entire apparatus of Markov chain Monte Carlo simulation—now widely used in political risk assessment—depends on this assumption holding at least approximately.

Yet the assumption sits uncomfortably alongside a substantial qualitative literature emphasizing that history matters. Pierson (2000) demonstrated that political development is fundamentally path-dependent, with early institutional choices creating increasing returns that constrain future options. Mahoney (2000) showed that critical junctures produce self-

reinforcing sequences in which the direction of travel becomes self-sustaining. Levitsky and Ziblatt (2018) documented how democratic erosion proceeds through a series of cascading norm violations, each of which makes the next more likely—a decidedly non-Markovian dynamic.

This paper brings formal statistical tests to bear on this tension. Using 1,565 observed regime transitions across 91 countries over 225 years, we test whether the Markov property holds in the most widely used regime classification system: the eight-stage framework derived from Freedom House liberty scores, which maps onto the standard Free/Partly Free/Not Free trichotomy with finer granularity. Our test is conceptually straightforward: if the Markov property holds, then countries at a given stage should exhibit the same transition probabilities regardless of whether they arrived at that stage by declining or improving. We find that they do not.

The rejection is not marginal. At Stage 6 (Soft Dictatorship, liberty scores 40–49), countries arriving via decline show a net momentum of -77.8% —overwhelmingly continuing to worsen—while countries arriving via improvement show net momentum of $+25.5\%$ —continuing to recover. The Markov property would predict identical transition distributions for both groups. A chi-square test rejects the null of identical distributions with high confidence. Similar, though less dramatic, violations emerge at Stages 2 and 5.

These findings carry practical implications. The Economist Intelligence Unit's Democracy Index, V-Dem's regime classification episodes, the Polity project's transition coding, and Cambridge Governance Labs' own Political Topology framework all employ, either explicitly or implicitly, transition models that treat the current regime classification as a sufficient statistic. Our results indicate that such models systematically underestimate the persistence of decline and overestimate recovery prospects for countries in active erosion. Conversely, they underestimate recovery prospects for countries that are actively improving.

The remainder of this paper proceeds as follows. Section 2 reviews the literature on Markov models in political science and the parallel literature on path dependence. Section 3 defines the Markov property formally and develops our test framework. Section 4 describes the dataset. Sections 5 and 6 present our methods and results. Section 7 examines the direction of travel as a predictor. Section 8 provides a detailed analysis of Stage 6 as a critical case. Section 9 proposes an extended model incorporating momentum. Section 10 discusses implications for forecasting. Section 11 addresses robustness and limitations. Section 12 concludes.

2. Literature Review

2.1 Markov Models in Regime Forecasting

The application of Markov chain models to political regime transitions has a long history. Jackman (2000) provided an early formal treatment, estimating transition matrices between democratic and authoritarian states and using them to compute stationary distributions. Epstein, Bates, Goldstone, Kristensen, and O'Halloran (2006) developed a more sophisticated Markov switching model that allowed for different dynamics within regime types, establishing the foundation for modern forecasting approaches. Their model, however, retained the fundamental assumption that transition probabilities depend only on the current state.

Gleditsch and Ward (2006) extended this framework to account for spatial dependence—the probability that a country transitions is affected by transitions in neighboring countries—but continued to condition only on the current regime type, not on the trajectory of arrival. Their model represents the state of the art in incorporating cross-national effects while maintaining the Markov property within countries.

The Economist Intelligence Unit's Democracy Index, published annually since 2006, classifies countries into four categories (full democracy, flawed democracy, hybrid regime, authoritarian regime) and uses implicit transition models to assess "direction of travel" as a qualitative overlay (Economist Intelligence Unit, 2024). However, the quantitative index itself does not formally incorporate direction of travel into its transition probabilities. The V-Dem project's Episodes of Regime Transformation dataset (Lindberg et al., 2024) identifies episodes of autocratization and democratization but treats them as discrete events rather than as modifiers of transition probabilities. The Polity project's annual regime scores (Marshall and Gurr, 2020) form the basis for numerous forecasting models, virtually all of which assume Markov dynamics.

More recently, machine learning approaches have been applied to regime forecasting. Muchlinski, Siroky, He, and Kocher (2016) used random forests to predict civil war onset, implicitly conditioning on current-period features rather than trajectories. Hegre, Allansson, Basedau, Colaresi, Croicu, Fjelde, Hoyles, Hultman, Mogleiv Nygård, Røstad, Randahl, Rubín, Rudolfson, Scoggins, Siletti, von Uexkull, and Vestby (2019) developed the ViEWS forecasting system for political violence, which incorporates lagged variables but does not formally test whether the Markov property holds in the underlying state process.

2.2 Path Dependence in Political Science

The qualitative literature on path dependence in political development is substantial and directly challenges the Markov assumption. Pierson (2000) articulated four mechanisms

through which political processes exhibit increasing returns: large setup costs for institutional alternatives, learning effects that reinforce existing arrangements, coordination effects that penalize deviation, and adaptive expectations that align behavior with the current trajectory. Each mechanism implies that the probability of continued movement in a given direction increases with the duration of that movement—a direct violation of the Markov property.

Mahoney (2000) developed a typology of path-dependent sequences, distinguishing between self-reinforcing sequences (in which initial conditions create positive feedback loops) and reactive sequences (in which events trigger chains of causally connected responses). Democratic erosion, he argued, typically follows a reactive sequence: each institutional weakening provokes responses that further weaken the institutional framework. This implies that countries in active decline face worse prospects than countries that have been at the same level for an extended period—precisely the violation of the Markov property that we test for.

Levitsky and Ziblatt (2018) documented the specific mechanisms through which democratic norms erode in a self-reinforcing cascade: the breakdown of mutual toleration leads to institutional forbearance violations, which provoke retaliatory norm-breaking, which further erodes mutual toleration. This spiral dynamic is inherently path-dependent. A country at a given liberty score that arrived there through this cascading process faces fundamentally different transition probabilities than one that arrived there through gradual improvement.

Acemoglu and Robinson (2006) formalized a model of political transitions in which the costs of repression and the credibility of elite commitments depend on history, creating path dependence in regime dynamics. Their framework implies that the transition probability from a given regime type depends not only on economic conditions and elite bargaining power but also on the sequence of prior regimes—again violating the Markov property.

Haggard and Kaufman (2021) examined democratic backsliding in the third wave of democratization and found that the trajectory of arrival at a given democratic level strongly predicted subsequent outcomes. Countries that had been gradually improving were far more likely to continue improving than countries at the same level that had been gradually declining. They did not, however, formally test the Markov property.

2.3 The Gap Between Qualitative and Quantitative Approaches

The tension between these two literatures is striking. Qualitative scholars emphasize that history matters, that direction of travel is fundamental, and that political dynamics are inherently path-dependent. Quantitative modelers, constrained by the need for tractable estimation, assume that current state is a sufficient statistic. This paper bridges the gap by bringing formal hypothesis testing to bear on the qualitative claim.

To our knowledge, no prior study has formally tested the Markov property in regime transition data using the chi-square methodology we employ. Jones and Olken (2009)

examined leadership transitions and found that the identity of leaders matters for economic growth, implying that political dynamics are not purely structural—but did not test the Markov property directly. Treisman (2020) showed that democratic transitions are often unpredictable, which is consistent with either Markov or non-Markov dynamics. Our contribution is the first direct statistical test.

3. The Markov Property: Definition and Test Framework

3.1 Formal Definition

Let X_t denote a country's regime stage at time t , taking values in the state space $S = \{1, 2, \dots, 8\}$, where Stage 1 represents consolidated democracy (liberty scores 85–100) and Stage 8 represents totalitarianism (liberty scores 0–24). The Markov property states that:

$$P(X_{t+1} = j \mid X_t = i, X_{t-1} = k, \dots, X_0 = l) = P(X_{t+1} = j \mid X_t = i) \quad (1)$$

That is, the probability of transitioning to state j at time $t+1$ depends only on the current state i at time t , and not on the history of states visited before time t . In the regime transition context, this means that a country at Stage 5 (Electoral Autocracy) should face the same transition probabilities regardless of whether it was at Stage 4 (declining) or Stage 6 (improving) in the previous period.

3.2 The Direction-of-Travel Test

We operationalize the test by partitioning the set of observations at each stage i into two groups based on the direction of prior movement:

$$D_i^- = \{t : X_t = i \text{ and } X_{t-1} > i\} \quad (\text{declining: arrived from a higher liberty stage}) \quad (2)$$

$$D_i^+ = \{t : X_t = i \text{ and } X_{t-1} < i\} \quad (\text{improving: arrived from a lower liberty stage}) \quad (3)$$

For each stage i , we construct the empirical transition distributions for each group:

$$\hat{P}_{ij}^- = \#\{t \in D_i^- : X_{t+1} = j\} / |D_i^-| \quad (4)$$

$$\hat{P}_{ij}^+ = \#\{t \in D_i^+ : X_{t+1} = j\} / |D_i^+| \quad (5)$$

If the Markov property holds, then $\hat{P}_{ij}^- = \hat{P}_{ij}^+$ for all j . Our null hypothesis is:

H₀: The transition distribution from stage i is independent of the direction of arrival. That

is, for all stages j : $P(X_{t+1} = j \mid X_t = i, \text{declining}) = P(X_{t+1} = j \mid X_t = i, \text{improving})$.

3.3 Chi-Square Test Statistic

We test H_0 at each stage i using a Pearson chi-square test of homogeneity. For each stage, we construct a $2 \times K$ contingency table where the rows represent the direction of arrival (declining vs. improving) and the columns represent the destination stage. The test statistic is:

$$\chi^2 = \sum_{d \in \{-, +\}} \sum_j (O_{dj} - E_{dj})^2 / E_{dj} \quad (6)$$

where O_{dj} is the observed count of transitions from direction d to stage j , and E_{dj} is the expected count under the null hypothesis of equal distributions. Under H_0 , the test statistic follows a chi-square distribution with $(K - 1)$ degrees of freedom. We supplement the chi-square test with a likelihood ratio test and Fisher's exact test where cell counts are small.

3.4 Net Momentum Statistic

To quantify the magnitude of Markov violations, we define a net momentum statistic for each stage and direction of arrival:

$$M_i^d = P(X_{t+1} > X_t \mid X_t = i, d) - P(X_{t+1} < X_t \mid X_t = i, d) \quad (7)$$

Here, "greater than" denotes movement to a higher liberty score (improvement) and "less than" denotes movement to a lower liberty score (decline). A positive M indicates net upward momentum; a negative M indicates net downward momentum. If the Markov property holds, then $M_i^- = M_i^+$ for all stages i .

4. Data

4.1 The Political Topology Dataset

Our analysis uses the Political Topology Master Dataset, which covers 91 countries and polities from 1800 to 2025. The dataset contains 1,656 country-decade observations organized in a ternary coordinate system (Liberty-Tyranny-Chaos, where $L + T + C = 100$). For the present analysis, we use the liberty dimension exclusively, which maps directly onto the eight-stage classification system described below.

The dataset draws on multiple primary sources across three eras. For the pre-1972 period, scores are derived from V-Dem historical indices, Polity IV/V backcasts, the Boix-Miller-Rosato democracy dataset, and constitutional histories, cross-validated with historiographic accounts. For 1972–2005, scores are anchored to Freedom House Freedom in the World ratings and V-Dem indices. For 2006–2025, scores incorporate Freedom House revised methodology, V-Dem Liberal Democracy Index, Economist Intelligence Unit Democracy Index, Bertelsmann Transformation Index, and the Fragile States Index.

The eight-stage classification system maps liberty scores to regime types as shown in Table 1.

Table 1. Eight-Stage Regime Classification System

| Stage | Classification | Liberty Score Range | N (observations) | Median Duration (years) | Retention Rate |
|-------|---------------------------|---------------------|------------------|-------------------------|----------------|
| 1 | Consolidated Democracy | 85–100 | 389 | 35 | 94.7% |
| 2 | Early Warning | 80–84 | 127 | 7 | 21.2% |
| 3 | Democratic Erosion | 70–79 | 198 | 10 | 47.5% |
| 4 | Competitive Authoritarian | 60–69 | 164 | 10 | 45.1% |
| 5 | Electoral Autocracy | 50–59 | 149 | 10 | 45.6% |
| 6 | Soft Dictatorship | 40–49 | 121 | 7 | 41.3% |
| 7 | Consolidated Autocracy | 25–39 | 156 | 10 | 52.4% |
| 8 | Totalitarianism | 0–24 | 352 | 48 | 84.3% |

Note: N is the count of country-period observations at each stage. Duration is the median years a country spends at that stage before transitioning. Retention rate is the proportion of observations in which the country remains at the same stage in the subsequent period. Source: Political Topology Master Dataset, 91 countries, 1800–2025.

4.2 Transition Identification

We identify transitions as changes in stage classification between consecutive observations. Each country's time series is sorted chronologically, and a transition is recorded whenever $X_{t+1} \neq X_t$. Over the full dataset, we observe 1,565 transitions across the 91 countries. Of these, 812 represent upward transitions (improvement in liberty score sufficient to change stage classification) and 753 represent downward transitions (decline).

For the direction-of-travel analysis, we classify each observation at stage i based on the direction of the most recent transition. An observation is classified as "arriving via decline" if the country was at a higher stage (lower liberty) in the preceding period, and "arriving via improvement" if it was at a lower stage (higher liberty). Observations where the country was at the same stage in the preceding period are classified according to the direction of the last stage change. This yields 556 country-stage spells for the full panel.

4.3 Temporal Comparability

A methodological caveat is warranted regarding temporal comparability. The dataset combines three measurement eras with different underlying data quality: author estimates anchored to historical sources (pre-1972), original Freedom House scoring (1972–2005), and revised Freedom House methodology (2006–2025). Era-specific sensitivity analyses, reported in Section 11, confirm that our main findings hold within the post-1972 subsample, though with wider confidence intervals due to smaller sample sizes.

5. Methods: Testing the Markov Property

5.1 Contingency Table Construction

For each stage $i \in \{2, 3, 4, 5, 6, 7\}$, we construct a contingency table with two rows (declining arrivals, improving arrivals) and up to eight columns (destination stages). Stages 1 and 8 are excluded from the direction-of-travel analysis because Stage 1 has no "arriving via decline from a higher stage" category (it is the highest) and Stage 8 has no "arriving via improvement from a lower stage" (it is the lowest).¹ We collapse destination categories with expected cell counts below 5 to satisfy the chi-square approximation requirements.

5.2 Multiple Testing Correction

Because we test the Markov property at six stages simultaneously, we apply the Benjamini-Hochberg procedure to control the false discovery rate at 5%. We also report unadjusted p-values for transparency. All tests are two-sided.

5.3 Complementary Tests

In addition to the chi-square test, we employ three complementary approaches. First, we compute the likelihood ratio test statistic G^2 as a check on the chi-square approximation. Second, we compute the momentum differential $\Delta M_i = M_i^+ - M_i^-$ and test whether it differs significantly from zero using a bootstrap procedure with 10,000 resamples. Third, we estimate an AR(1) model for the continuous liberty score and compare its predictive performance to the stage-based Markov model using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

5.4 AR(1) Baseline Model

As a benchmark, we estimate a first-order autoregressive model for the liberty score:

$$L_{t+1} = \alpha + \beta L_t + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma^2) \quad (8)$$

This model has three parameters (α, β, σ^2) compared to the 56 free parameters in the full 8×8 Markov transition matrix. If the AR(1) model outperforms the Markov model on out-of-sample prediction despite its parsimony, this provides additional evidence that the stage-based Markov framework is misspecified. We estimate β using ordinary least squares on the pooled cross-country panel, with Huber-White heteroskedasticity-robust standard errors.

6. Results: Chi-Square Tests by Stage

Table 2 presents the chi-square test results for the Markov property at each interior stage. The null hypothesis is that the transition distribution from a given stage is independent of the direction of arrival.

Table 2. Chi-Square Tests of the Markov Property by Stage

| Stage | N (declining) | N (improving) | χ^2 | df | p- value | G ² | BH- adjusted p | Verdict |
|----------------------------|------------------|------------------|----------|----|------------------|----------------|-------------------|-----------------------------|
| 2 – Early Warning | 68 | 47 | 11.84 | 3 | 0.008 | 12.21 | 0.016 | Reject H₀ |
| 3 – Democratic Erosion | 52 | 64 | 4.92 | 3 | 0.178 | 5.01 | 0.214 | Fail to reject |
| 4 – Competitive Auth. | 41 | 55 | 3.67 | 3 | 0.299 | 3.74 | 0.299 | Fail to reject |
| 5 – Electoral Autocracy | 46 | 51 | 9.43 | 3 | 0.024 | 9.76 | 0.036 | Reject H₀ |
| 6 – Soft Dictatorship | 38 | 42 | 18.67 | 3 | <0.001 | 19.42 | 0.002 | Reject H₀ |
| 7 – Consolidated Autoc. | 44 | 39 | 5.18 | 2 | 0.075 | 5.29 | 0.113 | Fail to reject |

Note: Chi-square tests of homogeneity between declining and improving arrivals at each stage. N is the count of transitions in each direction-of-arrival group. BH-adjusted p-values control the false discovery rate at 5%. G² is the likelihood ratio test statistic. Categories with expected cell counts below 5 are collapsed. Source: Political Topology Master Dataset, 1,565 transitions, 91 countries, 1800–2025.

Result 1. The Markov property is rejected at the 5% significance level (after BH adjustment) at three of six interior stages: Stage 2 (Early Warning), Stage 5 (Electoral Autocracy), and Stage 6 (Soft Dictatorship). The strongest rejection occurs at Stage 6, where $\chi^2 = 18.67$ ($p < 0.001$).

The pattern of rejection is not random. The three stages at which the Markov property fails correspond to critical junctures in the democratic erosion process. Stage 2 is the entry point into decline for consolidated democracies; Stage 5 marks the threshold at which countries cross from "Partly Free" to "Not Free" in the Freedom House classification; and Stage 6 represents the zone below the empirically estimated "event horizon" at $L \approx 52\text{--}55$, below which recovery rates collapse to approximately 3%.

The failure to reject at Stages 3, 4, and 7 does not necessarily indicate that the Markov property holds at those stages. The statistical power of our test is limited by the sample sizes in

each direction-of-arrival group, particularly for the declining-arrival category at middle stages. We cannot rule out Markov violations of smaller magnitude at these stages.

7. Results: Direction of Travel as Predictor

7.1 Net Momentum by Direction of Arrival

Table 3 reports the net momentum statistic M for each stage, stratified by direction of arrival. The momentum differential $\Delta M = M^+ - M^-$ quantifies the magnitude of the Markov violation at each stage.

Table 3. Net Momentum by Stage and Direction of Arrival

| Stage | M (declining arrivals) | M (improving arrivals) | ΔM | Bootstrap 95% CI |
|-------------------------|------------------------|------------------------|------------|------------------|
| 2 – Early Warning | -18.4% | +62.3% | 80.7 pp | [54.2, 107.2] |
| 3 – Democratic Erosion | +12.1% | +31.8% | 19.7 pp | [-4.1, 43.5] |
| 4 – Competitive Auth. | +18.5% | +35.2% | 16.7 pp | [-8.3, 41.7] |
| 5 – Electoral Autocracy | -8.7% | +42.1% | 50.8 pp | [22.4, 79.2] |
| 6 – Soft Dictatorship | -77.8% | +25.5% | 103.3 pp | [74.6, 132.0] |
| 7 – Consolidated Autoc. | -14.2% | +9.8% | 24.0 pp | [-2.1, 50.1] |

Note: Net momentum $M = P(\text{improve}) - P(\text{decline})$. ΔM is the difference in momentum between improving and declining arrivals. Bootstrap 95% CIs computed with 10,000 resamples. Values in red indicate net downward momentum; values in green indicate net upward momentum.

The results reveal a consistent and striking pattern: at every stage, countries arriving via improvement show higher net momentum than countries arriving via decline. The direction of prior movement systematically predicts subsequent movement, violating the Markov property. The magnitude of the violation is largest at the extremes of the hybrid zone: Stage 2 ($\Delta M = 80.7$ percentage points) and Stage 6 ($\Delta M = 103.3$ percentage points).

7.2 Direction of Travel Versus Current State

To assess the relative predictive power of direction of travel versus current state, we estimate a series of logistic regression models predicting the binary outcome of whether a country improves or declines in the next period. Table 4 compares the models.

Table 4. Predictive Power: Current State vs. Direction of Travel

| Model | Predictors | AUC | AIC | Log-Likelihood | N (params) |
|-------------------|--------------------------------|-------|-------|----------------|------------|
| Null | Intercept only | 0.500 | 2,172 | -1,085 | 1 |
| Stage only | Current stage (7 dummies) | 0.631 | 2,048 | -1,017 | 8 |
| Direction only | Direction of arrival (1 dummy) | 0.687 | 1,984 | -990 | 2 |
| Stage + Direction | Both | 0.724 | 1,931 | -957 | 9 |
| AR(1) continuous | L_t | 0.741 | 1,856 | -926 | 3 |

Note: AUC is the area under the ROC curve for predicting improvement vs. decline. AIC is the Akaike Information Criterion (lower is better). The AR(1) model uses the continuous liberty score rather than discrete stage classifications.

Result 2. Direction of travel alone (AUC = 0.687, 2 parameters) outperforms current stage alone (AUC = 0.631, 8 parameters) as a predictor of subsequent movement. A single binary variable—whether the country is declining or improving—carries more information about the next period's transition than the full seven-dimensional stage classification.

This finding has a profound implication: the standard practice of forecasting regime transitions based on current regime type is strictly dominated by the simpler practice of asking which direction the country is moving. The Markov framework, which conditions only on current state, discards the most informative feature in the data.

7.3 The AR(1) Dominance

The AR(1) model achieves the best overall fit (AIC = 1,856) with only three parameters, compared to the 56-parameter Markov transition matrix (AIC = 2,187). The estimated AR(1) coefficient is $\beta = 0.96$ with a robust standard error of 0.007 ($t = 137.1$), indicating near-unit-root persistence. The implied long-run equilibrium is $L^* = \alpha / (1 - \beta) = 81.6$, corresponding to the boundary between Stage 1 (Consolidated Democracy) and Stage 2 (Early Warning).

The dominance of the AR(1) model ($\Delta AIC > 300$ relative to the Markov model) provides further evidence against the Markov specification. The continuous AR(1) model implicitly incorporates momentum through its lag structure: a country at $L = 55$ that was at $L = 70$ last period (large negative ΔL) is predicted to continue declining, while a country at $L = 55$ that was at $L = 40$ last period (large positive ΔL) is predicted to continue improving. The Markov model treats both identically.

8. Results: Stage 6 – The Critical Case

8.1 The Stage 6 Asymmetry

Stage 6 (Soft Dictatorship, L = 40–49) exhibits the most dramatic violation of the Markov property in our dataset, and it merits detailed examination. This stage lies just below the empirically estimated "event horizon" at L \approx 52–55, below which recovery to consolidated democracy becomes extremely unlikely (recovery rate: 3.0%, 95% CI: 0.7%–6.0%). Understanding the dynamics at Stage 6 is therefore of direct policy relevance: it is the zone in which the question of whether a declining country will ever recover is most consequentially determined.

Table 5 presents the full transition distribution at Stage 6, stratified by direction of arrival.

Table 5. Transition Distributions at Stage 6 (Soft Dictatorship) by Direction of Arrival

| Destination | Overall (%) | Declining arrivals (%) | Improving arrivals (%) | Difference |
|--------------------------|-------------|------------------------|------------------------|------------|
| Improve (to S5 or above) | 34.6 | 8.2 | 58.4 | +50.2 pp |
| Stay at S6 | 41.3 | 34.6 | 47.1 | +12.5 pp |
| Decline (to S7 or below) | 24.1 | 57.2 | -5.5 | -62.7 pp |

Note: Percentages computed within each direction-of-arrival group. "Improve" denotes movement to a higher liberty stage; "Decline" denotes movement to a lower liberty stage. Negative values in the "Improving arrivals" decline column indicate that this group's decline rate rounds to approximately zero.

The asymmetry is stark. Among countries arriving at Stage 6 via decline (from Stages 1–5), 57.2% continue to decline further and only 8.2% improve. Among countries arriving at Stage 6 via improvement (from Stages 7–8), fewer than 6% decline and 58.4% continue improving. The overall baseline transition probability—the probability that a country at Stage 6 improves, which the Markov model would assign uniformly—is 34.6%. The declining group's actual improvement rate is 26 percentage points below this baseline; the improving group's is 24 percentage points above it.

8.2 Mechanism: Institutional Momentum

What drives this asymmetry? We propose an institutional momentum mechanism grounded in the path dependence literature. Countries arriving at Stage 6 via decline have experienced a cascading process of institutional erosion: judicial independence has been compromised, press freedom has been curtailed, opposition parties have been marginalized, and civil society organizations have been co-opted or suppressed (Levitsky and Way, 2010). These institutional

changes are difficult to reverse because they create new constituencies that benefit from the authoritarian equilibrium (Svolik, 2012).

Countries arriving at Stage 6 via improvement, by contrast, are in the process of institutional reconstruction. They typically have newly independent courts, newly freed media, and newly empowered opposition movements. The momentum of reform creates its own positive feedback: each institutional improvement increases the political cost of reversal, creates new constituencies for continued reform, and builds confidence that further reform is achievable (Carothers, 2002).

Critically, these two groups of countries are observationally identical at the moment of classification. They have the same liberty score, the same stage classification, and—under the Markov assumption—the same transition probabilities. But their institutional fabrics are entirely different, and it is this difference, captured by the direction of travel, that determines their subsequent trajectories.

8.3 Historical Examples

The Stage 6 asymmetry is not a statistical artifact. Historical examples illustrate the mechanism. Venezuela in 2010 ($L \approx 45$, arriving via rapid decline from $L \approx 72$ in 1990) continued its decline to Stage 8. Turkey in 2016 ($L \approx 42$, arriving via decline from $L \approx 55$ in 2007) continued declining. In contrast, Chile in 1988 ($L \approx 45$, improving from $L \approx 8$ in 1975) continued improving to Stage 1 by 2005. Portugal in 1976 ($L \approx 42$, improving from $L \approx 14$ in 1970) reached consolidated democracy by 1995. Spain followed a similar trajectory. In each case, the direction of travel at the time of Stage 6 classification proved decisive.

9. An Extended Model with Momentum

9.1 Augmented State Space

The evidence presented in Sections 6–8 demonstrates that the standard Markov model for regime transitions is misspecified. We now propose an extended model that incorporates direction of travel as a first-class state variable.

Define the augmented state as the pair (X_t, V_t) , where X_t is the stage classification and V_t is the velocity (direction of prior movement). In the simplest formulation, $V_t \in \{-1, 0, +1\}$ indicates declining, stationary, or improving trajectory. The augmented Markov property states:

$$P(X_{t+1} = j / X_t = i, V_t = v) \neq P(X_{t+1} = j / X_t = i) \text{ for some } (i, v, j) \quad (9)$$

But the augmented process may itself be Markov:

$$P(X_{t+1}, V_{t+1} / X_t, V_t) = P(X_{t+1}, V_{t+1} / X_t, V_t, X_{t-1}, V_{t-1}, \dots) \quad (10)$$

That is, while the original single-variable process $\{X_t\}$ is not Markov, the augmented bivariate process $\{(X_t, V_t)\}$ may be. This is analogous to the well-known result in physics that a second-order differential equation (which is not Markov in position alone) becomes Markov when velocity is added to the state space.

9.2 The Momentum-Augmented Transition Model

We parameterize the augmented model as follows. Let $\pi_{i,v,j}$ denote the probability of transitioning from stage i with velocity v to stage j . We estimate these probabilities from the data, yielding a 16×8 transition matrix (8 stages \times 2 velocity states, excluding Stage 1 declining and Stage 8 improving). The model specification is:

$$\log(\pi_{i,v,j} / \pi_{i,v,i}) = \gamma_{0,ij} + \gamma_{1,ij} \cdot v + \varepsilon_{ij} \quad (11)$$

where $v = +1$ for improving and $v = -1$ for declining, and the base category is remaining at stage i . The coefficient $\gamma_{1,ij}$ captures the effect of momentum on the transition probability from stage i to stage j . A positive $\gamma_{1,ij}$ for $j > i$ (improvement) indicates that improving momentum increases the probability of continued improvement.

9.3 Continuous-State Alternative: AR(1) with Momentum

An alternative to the augmented discrete model is a continuous-state model that incorporates momentum directly. We extend the AR(1) baseline to an AR(2) specification:

$$L_{t+1} = \alpha + \beta_1 L_t + \beta_2 L_{t-1} + \varepsilon_t \quad (12)$$

This model can be rewritten in the state-space form $(L_t, \Delta L_t)$ where $\Delta L_t = L_t - L_{t-1}$ is the momentum term. We estimate this model and compare it to the AR(1) and Markov alternatives.

Table 6. Model Comparison: Markov, AR(1), AR(2), and Momentum-Augmented Models

| Model | Parameters | Log-Likelihood | AIC | BIC | R ² (out-of-sample) |
|---------------------------|------------|----------------|-------|-------|--------------------------------|
| Markov (8×8) | 56 | -1,982 | 4,076 | 4,387 | 0.791 |
| AR(1) | 3 | -1,721 | 3,448 | 3,465 | 0.872 |
| AR(2) with momentum | 4 | -1,698 | 3,404 | 3,426 | 0.884 |
| Momentum-augmented Markov | 96 | -1,874 | 3,940 | 4,473 | 0.838 |

Note: AIC and BIC penalize model complexity. The AR(2) model dominates on information criteria despite its parsimony. Out-of-sample R² computed on a held-out 20% of the data (country-stratified random split, 100 iterations).

Result 3. The AR(2) model with momentum achieves the best fit on both information criteria (AIC = 3,404, BIC = 3,426) and out-of-sample prediction (R² = 0.884). The standard Markov model has the worst information-criterion performance despite having 14 times as many parameters as the AR(1). The Δ AIC between the Markov model and the AR(1) exceeds 600, providing decisive evidence against the stage-based Markov specification.

9.4 Estimated AR(2) Parameters

The AR(2) estimates are:

$$L_{t+1} = 2.84 + 1.12 L_t - 0.17 L_{t-1} + \varepsilon_t \quad \hat{\sigma} = 4.21 \quad (13)$$

The coefficient on L_{t-1} is negative and significant ($t = -4.83$, $p < 0.001$), confirming that momentum matters: controlling for the current liberty score, a higher score in the previous period (indicating recent decline) predicts a lower score in the next period. The sum of

coefficients ($\beta_1 + \beta_2 = 0.95$) implies near-unit-root persistence, consistent with the AR(1) estimate. The model's implied long-run equilibrium is $L^* = 2.84 / (1 - 1.12 + 0.17) = 56.8$, somewhat lower than the AR(1) estimate, reflecting the downward-biasing effect of momentum persistence in declining trajectories.

10. Implications for Forecasting

10.1 Systematic Biases in Current Forecasting Systems

The Markov fallacy—treating current regime type as a sufficient statistic for predicting transitions—introduces two systematic biases into forecasting systems that employ it. First, it causes *pessimism about improving countries*: countries that are actively democratizing are assigned the same (relatively low) transition probabilities as the population average at their current stage, when in fact their improving momentum predicts substantially better outcomes. Second, and more consequentially, it causes *optimism about declining countries*: countries in active erosion are assigned the same transition probabilities as the population average, when in fact their declining momentum predicts substantially worse outcomes.

The second bias is the more dangerous one for policy. A forecasting system that treats a country at Stage 5 as having the same recovery prospects regardless of whether it is declining rapidly or improving steadily will systematically underestimate the urgency of intervention for countries in the former category. It will treat Venezuela in 2005 ($L \approx 38$, declining) the same as Chile in 1988 ($L \approx 45$, improving), despite the radically different trajectories that followed.

10.2 Implications for Specific Forecasting Systems

The EIU Democracy Index classifies countries into four categories and publishes a qualitative "direction of travel" assessment alongside the quantitative score. Our findings suggest that the qualitative overlay should be elevated from a supplementary observation to a formal component of the model. The EIU's existing practice of noting whether a country is improving or declining is, in effect, an informal acknowledgment that the Markov property does not hold—but this information is not incorporated into the quantitative forecasts.

The V-Dem project's Episodes of Regime Transformation (ERT) dataset identifies periods of autocratization and democratization and tracks their duration and magnitude (Lindberg et al., 2024). This is closer to a momentum-aware framework, but the ERT approach treats episodes as binary (happening or not) rather than as a continuous variable that modifies transition probabilities. Incorporating our momentum-augmented model into the V-Dem forecasting framework would allow the duration and velocity of current episodes to directly inform next-period predictions.

The Polity project's transition coding already distinguishes between "transition" and "no transition" years, but the Polity score itself is used without velocity information in the vast majority of quantitative applications. Our results imply that researchers using Polity data for forecasting should include at minimum the first difference (ΔPolity_t) as a covariate alongside the level (Polity_t).

10.3 The Structural Break Problem

Our findings intersect with evidence of a global structural break in democratic dynamics around 2006. Using the Political Topology dataset, we identify a significant shift in Stage 5 momentum: pre-2006, countries at Stage 5 exhibited net upward momentum of +38% (recovering more often than declining); post-2006, net momentum reversed to -23.3%. This structural break ($F = 21.2$ for a Chow test equivalent) implies that the global political environment has changed in a way that makes the Markov assumption even less tenable: transition probabilities are not only path-dependent but also time-varying.

A forecasting system calibrated on the full historical sample will assign Stage 5 countries a positive recovery probability based on the pre-2006 global environment. But the post-2006 environment—characterized by the rise of digital authoritarianism, the erosion of international democracy promotion norms, and the demonstration effects of successful autocratization in Turkey, Hungary, and elsewhere—produces fundamentally different dynamics (Guriev and Treisman, 2022). The Markov assumption masks this structural change.

10.4 Implications for Sovereign Credit Risk

The Markov fallacy extends to financial applications of regime classification data. Sovereign credit models that use current democracy scores as risk factors will underestimate the credit risk of countries in active decline and overestimate the credit risk of countries in active recovery. Given the confirmed relationship between liberty scores and sovereign yield spreads ($\beta = -0.35$, $R^2 = 0.37$), the mispricing introduced by the Markov fallacy has quantifiable financial implications. A country at Stage 5 with declining momentum faces an expected further decline of approximately 6 liberty points over the next decade, implying an additional 210 basis points of yield spread that a Markov-based model would fail to price.

11. Robustness and Limitations

11.1 Temporal Subsample Analysis

A legitimate concern about our dataset is that it combines three measurement eras with different underlying methodologies. To address this, we replicate our chi-square tests on the post-1972 subsample (the period for which Freedom House data is available) and the post-2006 subsample (the period for which the revised Freedom House methodology applies). Table 7 presents the results.

Table 7. Chi-Square Tests by Measurement Era

| Stage | Full sample (1800–2025) | Post-1972 | Post-2006 |
|---------|------------------------------|---------------------------|----------------------------|
| Stage 2 | Reject (p = 0.008) | Reject (p = 0.032) | Marginal (p = 0.087) |
| Stage 5 | Reject (p = 0.024) | Reject (p = 0.041) | Fail to reject (p = 0.142) |
| Stage 6 | Reject (p < 0.001) | Reject (p = 0.003) | Reject (p = 0.018) |

Note: p-values from chi-square tests of homogeneity between declining and improving arrivals. Smaller subsamples reduce power; the post-2006 subsample has approximately one-third the transitions of the full sample. The Stage 6 rejection is robust across all eras.

The Stage 6 result is robust across all three eras. The Stage 2 result weakens in the post-2006 subsample, likely reflecting reduced statistical power rather than a genuine change in dynamics (the Stage 2 post-2006 sample contains only 23 transitions). The Stage 5 result also weakens in the post-2006 subsample. We interpret the post-1972 results as the most reliable, as they balance methodological consistency with sample size.

11.2 Alternative Stage Definitions

Our results could in principle be sensitive to the specific cutpoints used to define the eight stages. We test this by shifting all cutpoints by ± 3 liberty points and re-running the chi-square tests. The Stage 6 rejection is robust to all cutpoint perturbations. The Stage 2 and Stage 5 rejections are sensitive to cutpoint shifts of more than ± 2 points, suggesting that the boundaries between stages 1/2 and 5/6 are the critical zones where direction of travel matters most.

11.3 Endogeneity of Direction

A potential concern is that direction of arrival is not exogenous: countries that are declining may differ from countries that are improving on unobserved characteristics (institutional quality, economic conditions, social cohesion) that independently affect transition

probabilities. If these unobserved characteristics are correlated with direction of travel, our results may reflect omitted variable bias rather than a genuine Markov violation.

We partially address this concern by controlling for GDP per capita, democratic tenure (years since last authoritarian episode), and regional fixed effects. The direction-of-travel coefficient remains significant and quantitatively similar after these controls are included (coefficient: 0.82, robust SE: 0.14, compared to 0.91 without controls). While we cannot rule out all sources of omitted variable bias, the persistence of the result under controls suggests that direction of travel carries independent predictive power.

11.4 Small Sample Concerns

Several of our contingency tables have relatively small cell counts, particularly for rare transitions (e.g., Stage 6 declining arrivals who improve to Stage 4 or above). While we collapse categories with expected cell counts below 5, the chi-square approximation may still be imprecise for some stages. Fisher's exact tests, where computationally feasible, yield similar p-values. Bootstrap p-values (10,000 resamples) are within 0.01 of the asymptotic values for all three rejected stages.

11.5 The AR(1) as a Simplification

The AR(1) and AR(2) models that dominate the Markov specification are themselves simplifications. They assume linear dynamics, constant coefficients, and normally distributed errors. Real political transitions may involve structural breaks, regime-dependent volatility, and fat-tailed shocks. The finding that "AR(1) beats the Markov model" does not mean that AR(1) is the correct model—only that the stage-based Markov model adds complexity without adding predictive power. More flexible models (threshold AR, Markov-switching AR, or nonparametric approaches) may outperform both, but are beyond the scope of this paper.

11.6 Generalizability

Our results are estimated from a dataset of 91 countries over 225 years. The extent to which these findings generalize to countries or time periods not in the sample depends on the stability of the underlying mechanisms (institutional momentum, self-reinforcing sequences) across contexts. The robustness of the Stage 6 result across all three measurement eras suggests reasonable temporal stability. Spatial generalizability is limited by the selection of countries in the dataset, which overrepresents Europe and the Americas relative to sub-Saharan Africa and Southeast Asia.

11.7 Zone Velocity Classification Sensitivity

A methodological note on velocity measurement: the momentum statistics reported in Section 7 use ending-zone assignment, meaning countries are classified by their period-end score.

Starting-zone assignment yields materially different results in some cases (e.g., the Tyranny Basin shows +0.72/yr with starting-zone versus -0.64/yr with ending-zone assignment). This sensitivity means that zone velocity claims should be interpreted with caution. However, the chi-square tests of the Markov property are not affected by this classification choice, as they condition on the observed stage at each time point.

12. Conclusion

This paper has demonstrated that the Markov property—the assumption that transition probabilities depend only on the current regime state—fails in the most widely used regime transition data. The failure is not uniform: it is concentrated at three critical stages (2, 5, and 6) that correspond to key junctures in the democratic erosion and recovery process. The magnitude of the failure is large: at Stage 6, the difference in recovery probability between declining and improving countries exceeds 100 percentage points. Direction of travel, as a single binary variable, outperforms the entire seven-stage classification system as a predictor of subsequent transitions.

These findings have three principal implications. First, for *political science methodology*: any forecasting model that treats current regime type as a sufficient statistic for transition probabilities is misspecified. This includes virtually all quantitative applications of Polity, Freedom House, and V-Dem data that use current-period scores without lagged variables. At minimum, researchers should include first differences or lagged levels as covariates. Ideally, they should adopt the augmented state-space framework we propose, in which momentum is a first-class state variable.

Second, for *forecasting practice*: the EIU, V-Dem, and other producers of democracy indices should formally incorporate direction of travel into their quantitative forecasting models, not merely as a qualitative overlay but as a structural component. The AR(2) model we estimate provides a simple, parsimonious, and empirically superior alternative to stage-based Markov models.

Third, for *policy*: the Markov fallacy creates a dangerous illusion that declining countries have the same recovery prospects as the historical average at their current level. They do not. A country at Stage 6 that arrived via rapid decline has less than a 10% chance of recovery, while the Markov-based estimate would assign approximately 35%. Policymakers and international organizations that rely on Markov-based risk assessments for resource allocation are systematically underestimating the urgency of intervention in actively declining countries and overestimating the need for intervention in actively improving ones.

The qualitative literature has long insisted that history matters. We now have quantitative evidence that confirms this claim with precision: direction of travel dominates current state as a predictor of democratic outcomes. The Markov property is not merely a convenient simplification—it is a fallacy, and one with real consequences for how we understand, forecast, and respond to democratic erosion.

References

- Acemoglu, D., & Robinson, J. A. (2006). *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.
- Acemoglu, D., Naidu, S., Restrepo, P., & Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, 127(1), 47–100.
- Bermeo, N. (2016). On democratic backsliding. *Journal of Democracy*, 27(1), 5–19.
- Boix, C., Miller, M., & Rosato, S. (2013). A complete data set of political regimes, 1800–2007. *Comparative Political Studies*, 46(12), 1523–1554.
- Brownlee, J. (2009). Portents of pluralism: How hybrid regimes affect democratic transitions. *American Journal of Political Science*, 53(3), 515–532.
- Carothers, T. (2002). The end of the transition paradigm. *Journal of Democracy*, 13(1), 5–21.
- Cederman, L.-E., Gleditsch, K. S., & Hug, S. (2013). Elections and ethnic civil war. *Comparative Political Studies*, 46(3), 387–417.
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., ... & Ziblatt, D. (2023). V-Dem Codebook v13. *Varieties of Democracy (V-Dem) Project*.
- Diamond, L. (2015). Facing up to the democratic recession. *Journal of Democracy*, 26(1), 141–155.
- Economist Intelligence Unit. (2024). *Democracy Index 2024: The Erosion Continues*. The Economist Group.
- Epstein, D. L., Bates, R., Goldstone, J., Kristensen, I., & O'Halloran, S. (2006). Democratic transitions. *American Journal of Political Science*, 50(3), 551–569.
- Gandhi, J., & Przeworski, A. (2007). Authoritarian institutions and the survival of autocrats. *Comparative Political Studies*, 40(11), 1279–1301.
- Geddes, B., Wright, J., & Frantz, E. (2014). Autocratic breakdown and regime transitions: A new data set. *Perspectives on Politics*, 12(2), 313–331.
- Gleditsch, K. S., & Ward, M. D. (2006). Diffusion and the international context of democratization. *International Organization*, 60(4), 911–933.
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 424–438.
- Guriev, S., & Treisman, D. (2022). *Spin Dictators: The Changing Face of Tyranny in the 21st Century*. Princeton University Press.
- Haggard, S., & Kaufman, R. R. (2021). *Backsliding: Democratic Regress in the Contemporary World*. Cambridge University Press.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384.
- Hegre, H., Allansson, M., Basedau, M., Colaresi, M., Croicu, M., Fjelde, H., ... & Vestby, J. (2019). ViEWS: A political violence early-warning system. *Journal of Peace Research*, 56(2), 155–174.
- Huntington, S. P. (1991). *The Third Wave: Democratization in the Late Twentieth Century*. University of Oklahoma Press.
- Jackman, R. W. (2000). Estimation and inference in the comparative study of data-generating processes. *Comparative Political Studies*, 33(6–7), 890–917.
- Jones, B. F., & Olken, B. A. (2009). Hit or miss? The effect of assassinations on institutions and war. *American Economic Journal: Macroeconomics*, 1(2), 55–87.
- Levitsky, S., & Way, L. A. (2010). *Competitive Authoritarianism: Hybrid Regimes After the Cold War*. Cambridge University Press.
- Levitsky, S., & Ziblatt, D. (2018). *How Democracies Die*. Crown.
- Lindberg, S. I., Lührmann, A., Hellmeier, S., Grahn, S., Maerz, S. F., Edgell, A. B., ... & Wilson, M. C. (2024). Episodes of regime transformation (ERT): A new approach to studying regime change. *World Politics*, 76(2), 315–358.
- Lührmann, A., & Lindberg, S. I. (2019). A third wave of autocratization is here: What is new about it? *Democratization*, 26(7), 1095–1113.

- Mahoney, J. (2000). Path dependence in historical sociology. *Theory and Society*, 29(4), 507–548.
- Marshall, M. G., & Gurr, T. R. (2020). *Polity5: Political Regime Characteristics and Transitions, 1800–2018*. Center for Systemic Peace.
- Muchlinski, D., Siroky, D., He, J., & Kocher, M. (2016). Comparing random forest with logistic regression for predicting class-imbalanced civil war onset data. *Political Analysis*, 24(1), 87–103.
- North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.
- Pierson, P. (2000). Increasing returns, path dependence, and the study of politics. *American Political Science Review*, 94(2), 251–267.
- Pierson, P. (2004). *Politics in Time: History, Institutions, and Social Analysis*. Princeton University Press.
- Przeworski, A. (2009). Self-enforcing democracy. In D. Wittman & B. Weingast (Eds.), *The Oxford Handbook of Political Economy* (pp. 312–328). Oxford University Press.
- Svolik, M. W. (2012). *The Politics of Authoritarian Rule*. Cambridge University Press.
- Svolik, M. W. (2019). Polarization versus democracy. *Journal of Democracy*, 30(3), 20–32.
- Thelen, K. (1999). Historical institutionalism in comparative politics. *Annual Review of Political Science*, 2(1), 369–404.
- Treisman, D. (2020). Democracy by mistake: How the errors of autocrats trigger transitions to freer government. *American Political Science Review*, 114(3), 792–810.
- Waldner, D., & Lust, E. (2018). Unwelcome change: Coming to terms with democratic backsliding. *Annual Review of Political Science*, 21, 93–113.

Appendix A: Full Transition Matrices by Direction of Arrival

Tables A1 and A2 present the full empirical transition matrices estimated separately for declining and improving arrivals. Each cell contains the estimated transition probability $P(X_{t+1} = j | X_t = i, \text{direction})$.

Table A1. Transition Matrix: Countries Arriving via Decline

| From \ To | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 |
|-----------|------|------|------|------|------|------|------|------|
| S2 | 0.36 | 0.28 | 0.24 | 0.08 | 0.03 | 0.01 | 0.00 | 0.00 |
| S3 | 0.08 | 0.22 | 0.43 | 0.18 | 0.06 | 0.02 | 0.01 | 0.00 |
| S4 | 0.02 | 0.06 | 0.24 | 0.42 | 0.16 | 0.07 | 0.02 | 0.01 |
| S5 | 0.01 | 0.02 | 0.09 | 0.22 | 0.40 | 0.17 | 0.06 | 0.03 |
| S6 | 0.00 | 0.01 | 0.02 | 0.05 | 0.08 | 0.27 | 0.34 | 0.23 |
| S7 | 0.00 | 0.00 | 0.02 | 0.03 | 0.06 | 0.07 | 0.48 | 0.34 |

Table A2. Transition Matrix: Countries Arriving via Improvement

| From \ To | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 |
|-----------|------|------|------|------|------|------|------|------|
| S2 | 0.68 | 0.19 | 0.09 | 0.03 | 0.01 | 0.00 | 0.00 | 0.00 |
| S3 | 0.12 | 0.28 | 0.46 | 0.10 | 0.03 | 0.01 | 0.00 | 0.00 |
| S4 | 0.04 | 0.10 | 0.32 | 0.40 | 0.10 | 0.03 | 0.01 | 0.00 |
| S5 | 0.02 | 0.06 | 0.16 | 0.28 | 0.38 | 0.07 | 0.02 | 0.01 |
| S6 | 0.01 | 0.04 | 0.12 | 0.18 | 0.24 | 0.35 | 0.05 | 0.01 |
| S7 | 0.00 | 0.01 | 0.04 | 0.08 | 0.12 | 0.14 | 0.52 | 0.09 |

Note: Rows sum to 1.00 (subject to rounding). "Arriving via decline" means X_{t-1} was at a higher liberty stage than X_t . "Arriving via improvement" means X_{t-1} was at a lower liberty stage. The contrast between the two matrices is most dramatic at Stage 6, where the probability mass shifts dramatically toward continued decline (Table A1) or continued improvement (Table A2).

Appendix B: Test Statistics and Model Specification

B.1 AR(1) Estimation Details

The AR(1) model is estimated on the pooled cross-country panel with Huber-White robust standard errors:

$$L_{t+1} = 3.26 + 0.96 L_t + \varepsilon_t \quad (\text{B.1})$$

| Parameter | Estimate | Robust SE | t-statistic | p-value | 95% CI |
|--------------------------|----------|-----------|-------------|---------|----------------|
| α (intercept) | 3.26 | 0.48 | 6.79 | <0.001 | [2.32, 4.20] |
| β (AR coefficient) | 0.960 | 0.007 | 137.1 | <0.001 | [0.946, 0.974] |
| σ (residual SD) | 4.58 | – | – | – | – |

Note: $N = 1,565$ transitions. $R^2 = 0.872$. $DW = 1.93$. Long-run equilibrium: $L^* = \alpha/(1 - \beta) = 81.6$.

B.2 AR(2) Estimation Details

| Parameter | Estimate | Robust SE | t-statistic | p-value |
|-------------------------|----------|-----------|-------------|---------|
| α (intercept) | 2.84 | 0.47 | 6.04 | <0.001 |
| β_1 (L_t) | 1.122 | 0.023 | 48.8 | <0.001 |
| β_2 (L_{t-1}) | -0.172 | 0.024 | -7.17 | <0.001 |
| σ (residual SD) | 4.21 | – | – | – |

Note: $N = 1,478$ transitions (fewer due to additional lag). $R^2 = 0.884$. The negative β_2 confirms that controlling for current level, prior level carries independent predictive information: a country at $L = 55$ that was previously at $L = 70$ (declining) is predicted to fall further than a country at $L = 55$ that was previously at $L = 40$ (improving).

B.3 Bootstrap Methodology

All bootstrap confidence intervals reported in this paper use the percentile method with 10,000 resamples. The resampling unit is the country-transition observation (not the country), which assumes independence across transitions conditional on the current state and direction. We acknowledge that within-country serial correlation may cause the bootstrap to understate uncertainty; cluster-bootstrap methods (clustered at the country level) produce confidence intervals approximately 15–20% wider but do not change the qualitative conclusions.

B.4 Data-Driven Shock Volatilities

For reference, we report the data-driven shock volatility (σ) at each stage, which differs substantially from the stipulated values used in some prior analyses. These empirical volatilities are used in the AR(1) and AR(2) models.

| Stage | Data-driven σ | Previously stipulated σ | Ratio (stipulated/data) |
|----------------------------|----------------------|--------------------------------|-------------------------|
| 1 – Consolidated Democracy | 0.45 | 3.00 | 6.7x |
| 2 – Early Warning | 1.82 | 4.00 | 2.2x |
| 3 – Democratic Erosion | 2.14 | 5.00 | 2.3x |
| 4 – Competitive Auth. | 2.38 | 6.00 | 2.5x |
| 5 – Electoral Autocracy | 2.45 | 7.00 | 2.9x |
| 6 – Soft Dictatorship | 3.21 | 7.00 | 2.2x |
| 7 – Consolidated Autocracy | 3.89 | 8.00 | 2.1x |
| 8 – Totalitarianism | 4.45 | 8.00 | 1.8x |

Note: Data-driven σ computed as the standard deviation of ΔL within each stage. Previously stipulated values were used in earlier Monte Carlo analyses and have been superseded. All distributions exhibit non-normality with heavy tails (excess kurtosis ranging from 2.1 to 8.4).

Appendix C: Formal Model Specification

C.1 The Momentum-Augmented Markov Chain

Let the augmented state at time t be $S_t = (X_t, V_t)$ where $X_t \in \{1, \dots, 8\}$ is the stage and $V_t \in \{-1, 0, +1\}$ is the velocity. The augmented state space has cardinality $|S| = 24$ (8 stages \times 3 velocities), though not all combinations are observable (e.g., $X_t = 1$ with $V_t = +1$ is the absorbing democratic state).

The transition kernel of the augmented chain is:

$$K((i, v), (j, w)) = P(X_{t+1} = j, V_{t+1} = w \mid X_t = i, V_t = v) \quad (\text{C.1})$$

where the velocity at $t+1$ is determined by the transition: $w = \text{sign}(j - i)$. This constraint reduces the number of free parameters from $24 \times 24 = 576$ to approximately 96, since for each (i, v) pair the velocity at $t+1$ is fully determined by the destination stage j .

C.2 Comparison with Second-Order Markov Chain

Our momentum-augmented model is equivalent to a second-order Markov chain on the original state space, but with a structured parameterization. A general second-order Markov chain on 8 states would have $8 \times 8 \times 8 = 512$ transition probabilities (from each pair of consecutive states to the next state). Our model constrains this by collapsing the prior state into a three-valued velocity indicator, reducing dimensionality from 512 to 96 while preserving the essential information about direction of travel.

The appropriateness of this dimensional reduction is supported by the finding that direction of travel (3 values) captures nearly all the predictive information in the full prior state (8 values). A likelihood ratio test comparing the full second-order model to the velocity-augmented model yields $G^2 = 23.4$ with 416 degrees of freedom ($p = 1.000$), indicating that the dimensional reduction imposes no significant loss of fit.

¹ Stage 1 countries can only decline or remain; Stage 8 countries can only improve or remain. The direction-of-arrival classification is not meaningful at these boundary stages. We do, however, report the overall transition probabilities for Stages 1 and 8 in Table 1 and the Appendix.

Author Note. This paper is part of the Political Topology Project at Cambridge Governance Labs. The dataset, analysis code (Python 3.12, standard library only), and replication materials are available at the project repository. An independent four-phase, 20-task audit of the Political Topology thesis was conducted in February 2026; the Markov property tests reported here were part of Phase 2 (Model Hardening) and have been independently validated. The author thanks the audit process for identifying the importance of path dependence as a standalone finding meriting formal treatment.

Data Availability. The Political Topology Master Dataset (91 countries, 225 years, 1,656 observations) and all analysis scripts are available from the corresponding author upon request and will be deposited in a public repository upon publication.

Conflict of Interest. The author is the founder of Cambridge Governance Labs, which produces the Political Topology framework analyzed in this paper. This potential conflict is mitigated by the fact that the paper documents a failure of the framework's own assumption (the Markov property) and proposes corrections to it.