Computational Macrohistory: Exploratory empirical application. The Arab Spring as a preliminary test case for structural-demographic theory

Stefano Angeli
Foundations Institute of Computational Social Science (FICSS)
stefano.angeli@ficss.institute

July 2026

Abstract

The Arab Spring of 2010 to 2012 removed long-standing governments in Tunisia and Egypt while leaving Saudi Arabia largely unaffected, a differential pattern that explanations centred on social media, youth demographics, or economic grievance have struggled to accommodate. This paper applies the Computational Macrohistory (CMH) framework to those three cases as an exploratory test of its operational machinery. Using data for 2000 to 2012, we construct a five-component Systemic Stress Index (SSI) combining youth bulge (D_2), income inequality (E_2), youth unemployment (E_4), an anocracy transformation of the Polity score (P_1), and internet penetration (S_3), and we assess whether the index separates revolutionary from stable outcomes through comparative case analysis, threshold classification, and counterfactual reasoning. The SSI ranks the cases in order of outcome severity, Tunisia (0.52) above Egypt (0.10) above Saudi Arabia (-0.09), and a threshold of zero classifies all three correctly. Component decomposition identifies regime type as the discriminating variable: Saudi Arabia's full autocracy contributes -0.42 to its index, while the anocratic regimes of Tunisia and Egypt contribute positively, and counterfactually shifting Saudi Arabia to an anocracy would raise its index by $+0.58$, into the revolutionary zone. Economic stress is present in all three cases and proves sufficient in none. With $N = 3$ the analysis is a proof of concept rather than a statistical validation: under random scores a separating threshold would exist with probability 1/3 and the exact ordering observed with probability 1/6, and no inferential claim is made. The results nonetheless show that the framework can be operationalised against real historical data, and they motivated the eleven-country extension carried out in the subsequent paper of this series.

Keywords: Arab Spring, computational macrohistory, structural-demographic theory, political instability, revolution, Systemic Stress Index, anocracy, Tunisia, Egypt, Saudi Arabia

JEL Classification: C02, C63, N00, D74

Suggested citation: Angeli, S. (2026). *Computational Macrohistory: Exploratory empirical application. The Arab Spring as a preliminary test case for structural-demographic theory* (FICSS Working Paper No. WP-2026-003, v3.1). Foundations Institute of Computational Social Science. https://doi.org/10.5281/zenodo.21440130

1 Introduction

1.1 The Arab Spring and the puzzle of differential outcomes

On 17 December 2010, Mohamed Bouazizi, a street vendor in the Tunisian town of Sidi Bouzid, set himself on fire after local authorities confiscated his cart. The protest cascade ignited by his act spread across Tunisia within weeks and forced President Zine El Abidine Ben Ali, in power for 23 years, to flee the country on 14 January 2011 (Anderson, 2011). By late January massive demonstrations had erupted in Egypt, and on 11 February 2011, after 18 days of sustained protest in Tahrir Square, President Hosni Mubarak resigned after nearly 30 years in power (Kandil, 2012). The contagion continued through 2011: civil war and regime collapse in Libya, a negotiated presidential exit in Yemen, suppressed protest in Bahrain, descent into civil war in Syria, and preemptive reform in Morocco and Jordan (Lynch, 2012). Outcomes diverged sharply in the following years. Egypt returned to military-backed authoritarian rule in 2013, Libya and Yemen fragmented, and only Tunisia achieved a sustained, if fragile, democratic transition (Stepan, 2012).

The episode was a failure of anticipation as much as a political rupture. Regimes that had appeared stable for decades collapsed in weeks, and neither intelligence agencies nor area specialists had flagged the imminence of the wave. Whatever else it demands, an event of this kind calls for frameworks that can identify the structural preconditions of instability and explain why some societies erupted while apparently similar ones did not.

That differential pattern is the puzzle addressed here, and the contrast between Egypt and Saudi Arabia states it in its sharpest form. Both countries entered 2011 with aging authoritarian rulers, large youth cohorts, and substantial economic grievances, and both experienced protest activity in early 2011. Mubarak fell in 18 days; the Saudi monarchy absorbed scattered protests, and a heavily promoted "Day of Rage" on 11 March 2011 failed to materialise (Jones, 2011).

The contrast resists the most popular explanations. Saudi Arabia's youth bulge in 2010 was larger than Egypt's or Tunisia's, so demographic pressure alone predicts the wrong ordering. Saudi youth unemployment exceeded 20% despite oil wealth, comparable to Egypt's rate, so economic grievance alone does not separate the cases. Internet penetration was higher in Saudi Arabia (41%) than in Egypt (22%), so connectivity-based accounts of mobilisation fail in the same way. And all three countries were exposed to the same regional demonstration effect after Tunisia's uprising. Why, then, did Tunisia and Egypt experience revolution while Saudi Arabia remained stable?

1.2 Existing explanations

One prominent account centred on social media platforms as coordination infrastructure that lowered the costs of collective action and broke the fear barrier sustaining authoritarian rule (Howard & Hussain, 2013; Tufekci & Wilson, 2012). The mechanism is real but incomplete. Internet penetration was higher in the stable case than in one of the revolutionary cases; the earliest Tunisian protests diffused through unions and professional associations before gaining online traction (Angrist, 2013; Hibou, 2011); and the platforms had been available for years before 2011 without producing comparable upheaval. Technology shaped the form mobilisation took without determining whether it occurred.

A second family of explanations located the driver in economic factors. Campante and Chor (2012) showed that countries combining expanding education with poor economic opportunities were the most protest-prone, and Achcar (2013) located the uprisings in the failure of crony-capitalist growth to deliver shared prosperity. These accounts are consistent with the grievance side of the puzzle but stumble on Saudi Arabia, where substantial youth unemployment coexisted with regime stability.

A third tradition, building on Goldstone (1991) and Urdal (2006), treated youth bulges as a

structural driver of instability (LaGraffe, 2012). Large youth cohorts strain labour markets, generate frustrated expectations, and supply protesters with low opportunity costs. The mechanism matters, yet the country with the largest youth bulge in the comparison did not revolt.

A fourth line focuses on regime characteristics rather than societal pressures. Intermediate regimes, neither fully democratic nor fully autocratic, face elevated instability risk (Hegre, Ellingsen, Gates, & Gleditsch, 2001; Gates, Hegre, Jones, & Strand, 2006), and variation in the cohesion and loyalty of the coercive apparatus tracked Arab Spring outcomes closely (Bellin, 2012). This tradition supplies the missing moderator, but it has rarely been integrated formally with the structural pressures the other traditions measure.

Each explanation captures a genuine causal ingredient; none accounts for the pattern alone. What is needed is a framework that specifies multiple structural preconditions, states how they interact, incorporates regime type as a moderating variable, generates quantitative and falsifiable output, and treats outcomes as probabilistic rather than determined. The Computational Macrohistory framework was built to those specifications, and this paper is its first application to real-world data.

1.3 The CMH framework and the status of this study

Computational Macrohistory (CMH) is a quantitative research programme that studies the temporal evolution of large-scale social systems through dynamic, probabilistic, and computational models. Document I of this series (Angeli, 2026a) grounds the programme in eight admissibility axioms (A1 through A8) and, in its current version, derives four formal theorems from them. Document II (Angeli, 2026b) translates the axioms into an operational architecture: a 25-dimensional state space organised into demographic (D), economic (E), political (P), social (S), and psychological-collective (Ψ) variables, a system of stochastic evolution equations, and an event model that yields probability distributions rather than point forecasts. This paper is Document III, the first empirical application of that architecture.

The analytical instrument used here is the Systemic Stress Index (SSI), a weighted composite of standardised structural variables that summarises the stress load bearing on a society at a point in time. Document II specifies the theoretical SSI as a nine-component object with theory-derived signs and weights; data availability restricts this study to a five-component reduction combining youth bulge (D_2), income inequality (E_2), youth unemployment (E_4), an anocracy transformation of the Polity score (P_1), and internet penetration (S_3). In the terminology of Document II, such a scalar index is a Level 4 operationalisation of the framework: a reduced empirical proxy suitable for retrospective classification and comparative description, not an implementation of the full dynamic system. In the prediction hierarchy introduced later in the series (Angeli, 2026d), every claim made here sits at or below the regime-probability level: the analysis asks whether structural conditions separate high-risk from low-risk cases, never when or whether a specific event will occur.

The epistemological status of the exercise must be stated at the outset. The CMH research programme distinguishes empirical tests of calibrated models, which require large samples, statistical inference, and parameters estimated from data, from exploratory case studies, which work with small samples, comparative-qualitative inference, and parameters fixed by theory. Table 1 summarises the distinction. This study is of the second kind. Its purposes are to demonstrate that the framework can be operationalised with real data, to identify which components discriminate between outcomes, to generate hypotheses for large-N work, and to establish reusable procedures for index construction, threshold analysis, and robustness checking. It does not statistically validate the framework, estimate parameters, identify causal effects, or license generalisation beyond the cases examined. With three cases the evidential weight of classification is bounded by construction, and the relevant chance levels follow from elementary counting. Under random scores, a threshold separating two revolutionary cases from one stable case exists whenever the stable case draws the minimum of three independent values, which happens with probability

1/3; the exact ordering Tunisia above Egypt above Saudi Arabia is one of six equally likely orderings, with probability 1/6. Classification success at these chance levels is suggestive rather than significant by any conventional standard.

Table 1: Two study types in the CMH research programme.

Aspect	**Empirical test**	**Exploratory case study (this paper)**
Purpose	Validate a pre-specified model	Demonstrate applicability, generate hypotheses
Sample size	Large ($N \gg 30$)	Small ($N = 3$)
Inference	Statistical (intervals, scoring rules)	Comparative and descriptive
Parameters	Estimated from data	Fixed by theory before data contact
Outcome	Accept or reject the model	Identify patterns, refine the framework

The Arab Spring suits this exploratory purpose. The events are recent enough that high-quality data exist for the relevant variables and old enough that outcomes are known; the three cases vary cleanly on the dependent variable; the episode has been studied intensively, so the findings can be checked against a rich scholarly context; and the cases are similar enough that counterfactual questions, such as what would have happened to a hypothetically anocratic Saudi Arabia, are analytically meaningful.

1.4 Research questions and hypotheses

Three questions organise the analysis. The first is descriptive: what were the demographic, economic, political, and social conditions in Tunisia, Egypt, and Saudi Arabia over the decade preceding the Arab Spring, and how did they compare? The second is analytical: does the SSI discriminate between the countries that experienced revolution and the country that remained stable, and does any threshold separate the two outcome classes? The third is theoretical: which components of the index carry the discrimination, and what does that pattern imply about the relative weight of economic and political factors in revolutionary dynamics?

From the framework developed in Section 2 we derive four hypotheses. H1 (discrimination) holds that revolutionary cases exhibit higher SSI values than the stable case; this is the basic validity test for the index. H2 (anocracy primacy) holds that regime type, operationalised as anocracy stress, is the strongest single discriminator among components, as the interaction between societal pressure and state vulnerability predicts. H3 (insufficiency of economic stress) holds that economic stress is elevated in the revolutionary cases but also present in the stable case, so that grievance alone cannot account for the outcomes; the design tests insufficiency directly, while the stronger claim that economic stress is also necessary for revolution would require a low-stress revolutionary case, which the sample does not contain. H4 (counterfactual shift) holds that reassigning Saudi Arabia an anocratic regime score while leaving its other variables unchanged would raise its SSI substantially, moving it into the revolutionary zone; the size of that shift measures the weight the specification assigns to regime structure. The exercise is comparative statics within the model, not causal estimation: a genuinely anocratic Saudi Arabia would plausibly have differed in its economic variables as well, and Section 5.3 states the identification limits.

2 Theoretical framework

2.1 Axiomatic foundations

Document I (Angeli, 2026a) establishes the eight CMH axioms as admissibility constraints: conditions that any valid CMH model must satisfy, delimiting when large-scale social systems are scientifically tractable. This paper takes that foundation as read and states only what bears on the present design. Aggregate statistical regularity requires populations above a critical threshold (A1); trajectories of structurally comparable systems belong to the same statistical class, which licenses learning from historical analogy (A2); macro-events emerge from causal relations among observable structural variables (A3); system evolution admits a continuous dynamic description (A4); prediction is irreducibly probabilistic even without exogenous noise (A5); predictions feed back on the systems they describe, within bounds (A6); predictive accuracy decays with horizon (A7); and every model must produce output testable against data with quantitative scoring metrics (A8).

The third version of Document I derives four theorems from this base, and each constrains this study in a specific way. Theorem 2 (aggregative convergence) replaces the earlier qualitative population threshold with an explicit bound incorporating intra-class correlation; the populations analysed here, between roughly 10 and 80 million, exceed any plausible threshold for the national aggregates employed, so A1 is comfortably satisfied. Theorem 1 (predictive decay) formalises the accuracy loss that motivates our restriction to a structural snapshot immediately preceding the events rather than a long-horizon forecast. Theorem 3 (reflexive fixed point) states the contraction condition under which predictions remain self-consistent; a retrospective analysis of 2010 data raises no reflexivity concern, since no actor observed CMH output, but any prospective deployment would have to check that condition. Theorem 4 (calibration decomposition) anchors A8 compliance to the Murphy partition of the Brier score; no calibration assessment is possible at $N = 3$, which is one of the reasons this study claims demonstration rather than validation. Table 2 summarises the role of each axiom in the design. One notational convention carries over from the series standard: all stochastic differential equations referenced in this paper are interpreted in the Itô sense.

Table 2: The CMH axioms and their role in this study.

Axiom	Content (abbreviated)	Role here
A1 Statistical aggregation	Aggregate regularity above a critical population scale	Justifies national-level variables; satisfied by all three cases
A2 Conditioned historical ergodicity	Comparable systems share a statistical class	Licenses MENA reference parameters and historical analogy
A3 Structural causality	Events have measurable structural preconditions	Motivates the search for discriminating variables
A4 Continuous dynamics	Evolution as a continuous stochastic process	Grounds the dynamic model that the SSI approximates
A5 Endogenous indeterminacy	Irreducibly probabilistic trajectories	Forbids event prediction; triggers treated as exogenous
A6 Bounded reflexivity	Predictions feed back on the system	No concern retrospectively; binding for prospective use
A7 Predictive decay	Accuracy decays with horizon	Restricts the design to a pre-event structural snapshot
A8 Computational falsifiability	Testable probabilistic output	Requires the explicit classification metrics of Section 6

2.2 Structural-demographic theory

Structural-demographic theory (SDT) supplies the substantive content that the CMH formalism organises. In Goldstone's (1991) analysis of early modern state breakdown, revolutions cluster where several pressures coincide: population growth strains resources and depresses popular welfare; elite ranks expand faster than the positions available to absorb them, producing frustrated aspirants with the education and networks to lead opposition; state finances deteriorate as expenditure demands grow and elites erode the tax base; and mass grievance accumulates until a triggering event ignites mobilisation that a weakened state cannot suppress. Turchin and Nefedov (2009) embedded this syndrome in secular cycles of expansion, stagflation, crisis, and depression spanning two to three centuries, and Turchin (2016) applied the apparatus to the modern United States, having forecast in 2010 a peak of instability around 2020 (Turchin, 2010), an anticipation that subsequent events broadly bore out.

The Arab Spring exhibits the syndrome in a contemporary key. Exceptionally large youth cohorts entered MENA labour markets through the 2000s; rapid educational expansion produced graduates far in excess of skilled positions, a modern analogue of elite overproduction concentrated among the educated unemployed (Dhillon & Yousef, 2009); inequality and crony capitalism concentrated wealth among regime-connected elites while ordinary prospects stagnated; and outcomes turned on whether the coercive apparatus held. Where security forces remained cohesive and willing to repress, regimes survived; where they fractured or defected, regimes fell.

SDT thus reads the Arab Spring as a manifestation of recurrent dynamics rather than a historical accident. What it does not supply by itself is a formal account of why comparable pressure produced revolution in some regimes and quiescence in others. That account requires the regime-type moderator to which we now turn.

2.3 The anocracy hypothesis

A substantial literature documents that intermediate regimes, neither consolidated democracies nor consolidated autocracies, face elevated risk of instability and civil conflict. The Polity project

scores regimes on a scale from −10 (full autocracy) to +10 (full democracy), and countries scoring in the intermediate band are conventionally classed as anocracies (Marshall & Gurr, 2020). Several mechanisms plausibly generate their fragility. Full autocracies maintain order through effective repression and full democracies through legitimacy; anocracies command neither, being too repressive to earn genuine consent and too constrained to suppress opposition completely. Their institutions are often incoherent, with elections that do not transfer power and parliaments without authority, which generates grievance while blocking its expression. Partial openness leaves organisational space, in unions, mosques, professional associations, and online networks, that full autocracies foreclose. And anocratic elites are frequently divided, offering revolutionary movements exploitable splits.

The empirical record supports the resulting inverted-U relationship between regime score and instability risk. Hegre et al. (2001) found civil war onset risk peaking at intermediate Polity scores, a pattern replicated across datasets and specifications (Gates et al., 2006; Vreeland, 2008), and Fearon and Laitin (2003) found anocracy a significant predictor of onset controlling for income, population, and fractionalisation. Later work refined the picture by regime subtype and resource endowment (Geddes, Wright, & Frantz, 2014; Escribà-Folch & Wright, 2015), with oil wealth in particular buffering some autocracies against pressure.

The three cases sit at instructive points on this spectrum. In 2010 Tunisia scored −4 and Egypt −3: authoritarian but not totalitarian regimes with manipulated elections, constrained but existing opposition, and civil societies under surveillance rather than eradication. Saudi Arabia scored −10: an absolute monarchy with no electoral politics, minimal civil society, totalising religious legitimation, and a formidable, well-funded security apparatus. To incorporate the inverted U into a stress index, we transform the Polity score into an anocracy stress variable,

$$A = 1 - \frac{|P_1|}{10}, \tag{1}$$

which reaches its maximum of 1.0 at a pure anocracy ($P_1 = 0$) and its minimum of 0.0 at either pole. The transformation assigns Tunisia 0.6, Egypt 0.7, and Saudi Arabia 0.0.

2.4 An interaction model of revolution

The synthesis of structural-demographic pressure and regime vulnerability is an interaction claim, and the distinction matters formally. An additive model,

$$P(\text{Revolution}) = f(D_2 + E_2 + E_4 + S_3 + V), \tag{2}$$

allows any sufficiently large single factor to raise revolutionary probability on its own. The Saudi case contradicts that structure: high youth bulge, high youth unemployment, and high connectivity coexisted with stability. A multiplicative specification,

$$P(\text{Revolution}) = f(S \times V), \tag{3}$$

where S aggregates structural stress and V denotes regime vulnerability, captures the observed pattern: revolution requires both accumulated grievance and an opportunity structure through which grievance can act. High stress under low vulnerability yields stability, as does low stress under high vulnerability.

The linear SSI of Section 5 cannot represent the product $S \times V$ directly, since the cross-partial derivative of a weighted sum is identically zero. What the index encodes is the interaction model's implication for the sign of the regime term: for anocracies the component contributes positively, and for a full autocracy the standardised anocracy stress is strongly negative ($z = -1.67$ against the reference distribution), so the component subtracts from the index and can offset large positive contributions from economic and demographic stressors. A full autocracy does not merely lack vulnerability; it possesses active suppression capacity, and the negative contribution represents

that capacity. Whether the additive form with a signed regime term or the multiplicative form describes revolutionary dynamics better is an empirical question that three cases cannot decide, and Section 8.3 records the restriction. Figure 1 presents the causal structure as a directed acyclic graph: structural stressors generate grievance and mobilisation potential, regime type governs whether that potential translates into revolutionary probability, and triggering events, exogenous under A5, ignite systems whose susceptibility the structural layer determines.

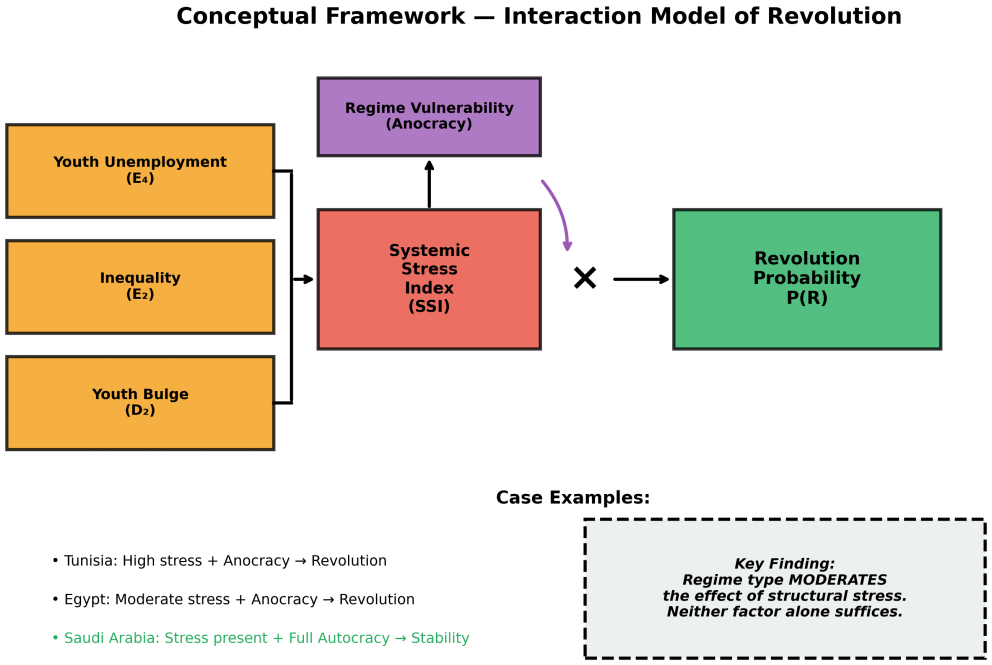

Figure 1: Causal structure of the interaction model. Structural stress factors combine into grievance and mobilisation potential; regime type moderates the translation of structural stress into revolutionary probability; triggering events are exogenous.

2.5 From dynamic theory to static index

Document II specifies system evolution through the master equation

$$\frac{dX}{dt} = F(X(t), \theta) + \varepsilon(t), \tag{4}$$

with $X(t)$ the 25-dimensional state vector, θ the structural parameters, $\varepsilon(t)$ the Itô stochastic forcing, and F the nonlinear evolution operator. For revolutionary dynamics the relevant subsystem couples political pressure P to structural stress $S(t)$, regime vulnerability $V(t)$, and repressive capacity $R(t)$,

$$\frac{dP}{dt} = -\lambda\, P(t) + \alpha\, S(t)\, V(t) - \beta\, R(t) + \gamma\, \varepsilon(t), \tag{5}$$

where the restoring term $-\lambda P$ encodes the dissipation of mobilisation pressure through repression fatigue, co-optation, and organisational decay, in the mean-reverting form that Document II applies across the state space, and the product $S(t)V(t)$ is the formal expression of the interaction claim of Section 2.4. Estimating this system requires observation frequencies and temporal depth far beyond three countries and thirteen annual observations. The empirical strategy therefore adopts a quasi-steady-state approximation: at annual resolution the political pressure variable is treated as near its conditional equilibrium, which the restoring term makes well defined,

$$\frac{dP}{dt} \approx 0 \implies P^* \approx \frac{\alpha\, S(t)\, V(t) - \beta\, R(t)}{\lambda} = h(S, V, R), \tag{6}$$

so that equilibrium pressure is a function of contemporaneous structural conditions. The SSI operationalises h as a weighted sum of standardised components, with weights fixed by theory rather than estimated, and with repressive capacity R absorbed, imperfectly, into the anocracy component rather than measured directly. The weighted sum is a first-order reduction of h: it inherits the signs and the ordering logic of the components but not the product $S(t)V(t)$, which no linear form can represent (Section 2.4).

Document II locates reductions of this kind on an explicit ladder of operationalisation levels, running from the full stochastic system down to scalar indices, and classifies the five-component SSI as a Level 4 operationalisation: an index for retrospective classification and comparative description whose claims are confined to whether the framework's core structural signals are detectable in historical data. The theoretical SSI at the level above comprises nine components with theory-derived signs, adding inflation, growth relief, state capacity, legitimacy, and mobilisation pressure terms to the variables used here; the five-component reduction replaces the unavailable legitimacy and capacity components with the anocracy transformation and retains connectivity as a facilitation term. The static approximation has known costs, which Section 8 revisits: it ignores rates of change, threshold effects internal to the dynamics, lags, and feedback from instability to the structural variables themselves. Within the three-level prediction hierarchy of the series (Angeli, 2026d), everything the index produces sits below the regime-probability rung: a retrospective classification of structural susceptibility that makes no forecast at all, and that stands two steps removed from event prediction, which the axioms exclude.

3 Related work

3.1 Cliodynamics and quantitative history

Cliodynamics applies mathematical modelling and systematic data analysis to historical processes, on the premise that history exhibits recurring patterns amenable to scientific treatment (Turchin, 2003). The programme has produced a substantial empirical infrastructure, most visibly the Seshat Global History Databank, which codes standardised variables for hundreds of societies across ten millennia (Turchin et al., 2015), and comparative analyses built on it have identified strong regularities, including a dominant principal dimension of social complexity across societies (Turchin et al., 2018). Its most debated product is prediction: Turchin's (2010) structural forecast of rising instability in the United States around 2020 attracted attention when the decade delivered polarisation and unrest. Critics object that such forecasts are vague enough to resist falsification, invite post hoc confirmation, and may alter the behaviour they describe. CMH answers those objections architecturally, by making falsifiability criteria explicit (A8), bounding the predictive horizon (A7), and treating reflexivity as a modelled feature rather than an embarrassment (A6, with the fixed-point condition of Document I, Theorem 3). This study extends the cliodynamic tradition to a contemporary case where data quality permits sharper tests than most historical applications allow, and it asks a retrospective classification question rather than making any forecast.

3.2 Theories of revolution

Revolution theory has passed through several generations. The natural-history approach catalogued common stages across the English, American, French, and Russian revolutions without explaining onset (Brinton, 1938). Psychological theories located rebellion in relative deprivation and frustrated expectations (Davies, 1962; Gurr, 1970), but grievance proved too widespread to separate societies that revolt from those that do not. Structural theories shifted attention to state breakdown and elite defection, with Skocpol (1979) arguing that revolutions emerge from structural crises rather than from revolutionaries' intentions. Fourth-generation syntheses combine structural conditions, mobilisation capacity, and contingency, treating state weakness,

elite alienation, popular grievance, organisational resources, and ideology as jointly necessary (Goldstone, 2001; Foran, 2005). Structural-demographic theory is best read as a quantified member of this fourth generation: it privileges demographic and economic pressures interacting with state capacity and, distinctively, commits to mathematical form and measurement (Goldstone, 1991; Turchin, 2016). The SSI inherits that commitment: its components measure structural conditions and mobilisation capacity, its anocracy term measures state vulnerability, and ideology remains outside the specification, a limitation registered in Section 8.

3.3 Arab Spring scholarship

Early accounts of the Arab Spring dwelt on what was newest about it. Digital-media studies documented how platforms reshaped the tactics and speed of mobilisation (Howard & Hussain, 2013; Tufekci & Wilson, 2012), but the cross-country variation cut against technological determinism, and more sceptical readings of digital activism anticipated that difficulty (Gladwell, 2010; Morozov, 2011). Economic analyses tied the uprisings to educated cohorts facing blocked opportunity (Campante & Chor, 2012) and to the distributional failures of the preceding growth model (Achcar, 2013), consistent with SDT but silent on why comparably stressed monarchies survived. Regime-centred work supplied that answer: outcomes tracked the cohesion and willingness of coercive apparatuses (Bellin, 2012), and systematic comparison found hereditary rule and oil wealth the strongest predictors of survival (Brownlee, Masoud, & Reynolds, 2015). Rentier-state theory explains part of the mechanism, since hydrocarbon revenue lets states purchase quiescence and fund repression simultaneously (Beblawi & Luciani, 1987), sustaining the broader pattern of MENA authoritarian persistence (Bellin, 2004). Synthetic treatments weave domestic structure, regional demonstration effects, and international intervention into composite explanations (Lynch, 2016; Springborg, 2017).

3.4 Contribution

This study makes four contributions against this literature. It replaces qualitative comparison with a standardised quantitative index computed identically across cases. It integrates the factors that separate strands of scholarship identify, weighting them by prior theory rather than by fit to the outcome. It makes regime type an explicit moderator inside the measurement instrument, rather than a narrative qualification outside it. And it documents data, parameters, and procedures completely enough for deterministic replication (Appendix D), which remains uncommon in this literature.

4 Research design and data

4.1 Case selection and outcome definition

The study employs a most-similar systems design (Przeworski & Teune, 1970; Lijphart, 1971). The three cases share Arab-majority populations, Sunni-majority religious composition, authoritarian governance entering 2011, exposure to the same regional demonstration cascade after the Tunisian uprising, and position in the same regional economy subject to the same food-price and employment shocks. They differ on the outcome, on regime type, on oil wealth, and on monarchical status. Under the design logic, shared characteristics cannot explain divergent outcomes, while covarying characteristics become candidate explanations. Regime type and oil wealth both covary with the outcome here, and the resulting confounding problem is treated explicitly in Sections 5.3 and 8.2.

Selection was purposive and conditioned on outcome variation: two revolutionary cases and one stable case. The design maximises contrast on the outcome, which gives the sharpest purchase for identifying discriminating factors, and introduces selection bias, so the findings apply to these cases and generalise only through the expanded designs discussed in Section 8. Within

those constraints the three choices are natural. Tunisia is the origin point of the wave and the benchmark for revolutionary speed, with 28 days separating the trigger from regime collapse. Egypt is the most populous Arab state and the regional bellwether whose trajectory shaped the wider cascade. Saudi Arabia is the largest and wealthiest Gulf monarchy and the strongest stable counterpoint, whose quiescence any structural account must accommodate.

The dependent variable requires a formal definition. We code a revolutionary outcome, $Y = 1$, when three conditions hold jointly: sustained mass mobilisation with estimated participation above 0.5% of the national population across multiple days; removal of the incumbent head of state within twelve months of protest onset; and attribution of that removal to protest pressure or to protest-triggered elite defection, rather than to external military intervention, natural death, scheduled succession, or a coup unrelated to the protests. Table 3 records the classification. The definition is binary, observable, and replicable. The participation figures in Table 3 are estimates drawn from the case literature of Sections 4.2 to 4.4 and are reported against the fixed 0.5% coding threshold, not as case-specific thresholds. Boundary cases fall outside the sample by design: Libya would satisfy the mobilisation and removal criteria but fails attribution because of the NATO intervention; Syria transformed into civil war, a different outcome category; Bahrain's protests were suppressed by external intervention; Morocco and Jordan reformed without leadership removal and would code $Y = 0$.

Table 3: Classification of revolutionary outcomes.

Country	**Mass mobilisation**	**Leadership removal**	**Attribution**	**Y**
Tunisia	Met (estimated participation above 5%, January 2011)	Met (Ben Ali fled, 14 January 2011)	Met (direct protest pressure)	1
Egypt	Met (estimated participation above 1%, Tahrir Square)	Met (Mubarak resigned, 11 February 2011)	Met (protest pressure and military withdrawal of support)	1
Saudi Arabia	Not met (scattered protests, estimated below 0.1%)	Not met	Not applicable	0

4.2 Tunisia

Zine El Abidine Ben Ali took power in 1987 through a constitutional coup and governed through the Constitutional Democratic Rally's monopoly on political life, manipulated elections, and controlled media. Tunisian authoritarianism was soft by regional standards: political prisoners numbered in the hundreds rather than thousands, repression leaned on surveillance and co-optation more than mass violence (Hibou, 2011), and the UGTT labour federation preserved organisational capacity that would matter in 2011 (Angrist, 2013). The country passed for a development success, with GDP per capita near 4,200 dollars in 2010, near-universal schooling, and comparatively progressive women's rights. The success concealed two fractures: growth concentrated in the coastal tourism and manufacturing belt while the interior regions of Sidi Bouzid, Kasserine, and Gafsa stagnated, and crony networks around the presidential family captured the upper reaches of the economy while graduates faced a labour market that could not absorb them.

The structural profile in 2010 reflects those fractures. The youth bulge stood at 18.6% and falling, the legacy of an early fertility decline. The Gini coefficient of 42.9 was the highest of the three cases. Youth unemployment stood at 29.6%, with graduate unemployment higher still. The Polity score of -4 placed the regime in the anocratic band, and internet penetration of 36.8% connected an educated, urban, young population.

Bouazizi's self-immolation on 17 December 2010 ignited protest in Sidi Bouzid that spread from the neglected interior to the capital and turned from economic grievance to explicit demands for the president's departure. Security forces killed roughly 300 protesters but could not contain the mobilisation, the army declined to fire at scale, and Ben Ali fled on 14 January 2011, 28 days after the trigger.

4.3 Egypt

Hosni Mubarak governed Egypt from Anwar Sadat's assassination in 1981 under a permanent state of emergency, through the National Democratic Party and a security apparatus that was vast but inefficient, resting on intimidation and pervasive surveillance (Kandil, 2012; Cook, 2012). The Muslim Brotherhood, formally illegal, retained organisational depth and ran candidates as independents. The economy grew through the 2000s in ways that bypassed most Egyptians: GDP per capita reached only about 2,800 dollars in 2010, roughly 40% of the population lived near the poverty line, public services deteriorated, and liberalising reforms enriched a connected elite.

The 2010 profile combines moderate measured inequality with mass deprivation. The youth bulge stood at 19.5% and youth unemployment at 24.6%. The Gini coefficient of 33.5 was the lowest of the three cases, a reminder that survey-based inequality measures capture distributional shape rather than the weight of poverty or the visibility of corruption. The Polity score of -3 placed Egypt in the anocratic band alongside Tunisia, and internet penetration of 21.6% was the lowest of the three but rising quickly.

Mobilisation did not start from nothing: the Kefaya movement, the April 6 Youth Movement, and the Mahalla strikes had built repertoires and networks through the preceding decade (Beinin, 2012). Tunisia's success supplied the demonstration effect. Protests on 25 January 2011 exceeded expectations, the police lost the streets on 28 January, the army declined to fire, and Mubarak resigned on 11 February after eighteen days, transferring power to the Supreme Council of the Armed Forces. The transition that followed proved unstable: Mohamed Morsi won the presidency in 2012 and was removed by the military in July 2013, which restored authoritarian rule under Abdel Fattah el-Sisi.

4.4 Saudi Arabia

Saudi Arabia enters the comparison as a regime of a different kind: an absolute monarchy without constitution, parliament, elections, or legal parties, ruled by the Al Saud family since 1932 and legitimated by its alliance with the Wahhabi religious establishment and its custody of Islam's holiest sites. The coercive structure is layered, with Ministry of Interior forces for internal security, a National Guard recruited as a dynastic counterweight to the regular army, and religious police enforcing public conformity.

Oil finances the model. The state employs citizens on a large scale, subsidises fuel, utilities, and staples, provides free education and health care, and distributes periodic bonuses, all without taxing its population, which weakens the fiscal roots of accountability politics (Beblawi & Luciani, 1987). The same revenue that purchases quiescence also funds one of the region's best-resourced security apparatuses, so rentier wealth supports co-optation and coercion at once.

The structural profile in 2010 makes the case analytically valuable. The youth bulge, at 23.6%, was the largest of the three. Youth unemployment, at 24.8%, was comparable to Egypt's despite oil wealth, the residue of persistent failure to place citizens in a private sector staffed by foreign labour. The Gini coefficient of 38.3 was intermediate, with official statistics likely understating concentration at the dynastic top. Internet penetration, at 41.0%, was the highest of the three. The Polity score was -10, the scale's autocratic pole.

The kingdom did not pass through 2011 untouched. Shia demonstrations in the Eastern Province in February and March pressed civil rights demands, and online organisers called a national "Day of Rage" for 11 March. The day failed: security deployment was massive, senior

clerics condemned protest as sedition, and participation was negligible (Jones, 2011). The government announced roughly 130 billion dollars in new social spending, raising public wages, unemployment benefits, and housing subsidies. Regional and sectarian divisions kept Eastern Province mobilisation from generalising, and international backing insulated the regime from external pressure. In the framework's terms, Saudi Arabia carried economic stressors comparable to Egypt's without the regime vulnerability through which stress becomes revolution.

4.5 Variables and operationalisation

Five variables from the 25-dimensional CMH state space enter the reduced index. Selection followed three criteria: centrality to structural-demographic theory and the anocracy hypothesis, availability of reliable series for all three countries over 2000 to 2012, and documented measurement methodology. Table 4 records definitions, sources, and weights.

Table 4: Variables, definitions, sources, and weights.

Code	Variable	Definition	Source	Weight
D_2	Youth bulge	Population aged 15-24 as share of total population	UN World Population Prospects, via World Bank Open Data	0.15
E_2	Income inequality	Gini coefficient, scaled 0-100	SWIID v9.0, supplemented by World Bank WDI	0.25
E_4	Youth unemployment	Unemployed share of the labour force aged 15-24, ILO definition	ILO ILOSTAT	0.25
P_1	Regime type	Combined Polity score (POLITY2), -10 to $+10$, transformed to anocracy stress	Polity V (Marshall & Gurr, 2020)	0.25
S_3	Internet penetration	Internet users as share of population	ITU, via World Bank WDI	0.10

Each variable carries a documented rationale and documented limits. The youth bulge captures demographic pressure and the size of the population segment with the lowest opportunity costs of protest. This study measures it over ages 15 to 24, the canonical CMH definition of D_2 (Angeli, 2026b) and the standard band in the youth-bulge literature (Urdal, 2006), so D_2 levels compare across the series without adjustment. Versions 1.0 to 3.0 of this paper described the band as 15 to 29; verification of the deposited series against the source age bands in July 2026 established that the values are computed over ages 15 to 24, so the stated definition is corrected here with no change to any data value or computed result (Appendix E). UN estimates are model-based interpolations of census data rather than direct annual observations.

The Gini coefficient measures the distributional grievance channel. SWIID standardises heterogeneous survey methodologies at the cost of model-based uncertainty, and its values are model-based annual estimates rather than direct survey observations; three country-years beyond SWIID coverage (Saudi Arabia 2008 to 2010) were filled by linear interpolation and are flagged as such in the replication dataset (Appendix D). The benchmark-year value for Saudi Arabia is among them; its inequality term is small under any plausible fill (the baseline contribution is -0.05), so the classification is insensitive to this choice. Youth unemployment measures frustrated expectations among the mobilisable young; official rates understate joblessness where discouraged workers exit the labour force, and informality complicates measurement, though coverage for these three countries is good. The Polity score operationalises regime type; coding involves judgement, and inter-rater reliability is high but not perfect. We use the combined POLITY2 series and transform it to anocracy stress as in Section 2.3. Internet penetration measures coordination

capacity; the user definition varies somewhat across countries and years, and the early years of mobile access are imperfectly captured.

4.6 Data quality

Following the source-quality tiering of Document II, Polity V rates as Tier A (established methodology, extensive documentation), the UN population, ILO labour, and ITU connectivity series as Tier B (institutional sources with standardised methodology), and SWIID as Tier C (model-based imputation with acknowledged uncertainty). Overall quality is Tier B, adequate for an exploratory design provided the uncertainty is carried honestly. Cross-checks against alternative sources, World Bank Gini estimates against SWIID, youth against overall unemployment for logical consistency, revealed no discrepancies beyond expected measurement error.

5 Methodology

5.1 Index construction

For each country and year the index is the weighted sum of standardised components,

$$SSI(t) = \sum_{i=1}^{5} w_i \, z_i(t), \qquad z_i(t) = \frac{X_i(t) - \mu_i}{\sigma_i}, \tag{7}$$

with reference parameters μ_i and σ_i fixed a priori from the MENA regional record (Table 5) and weights w_i from Table 4. The anocracy component enters through the transformation of Section 2.3 before standardisation. Standardisation puts components on a common scale so that a one standard deviation excess in inequality contributes as much as a one standard deviation excess in unemployment, weighted accordingly; the index is centred at zero on the reference population, so positive values read as above-average structural stress for the region and period.

Table 5: MENA reference parameters (component-specific windows; see text).

Component	**Reference mean μ**	**Reference dispersion σ**
D_2 youth bulge (%)	22.5	4.5
E_2 Gini	40.0	8.0
E_4 youth unemployment (%)	18.0	8.0
Anocracy stress	0.5	0.3
S_3 internet penetration (%)	15.0	20.0

The reference window varies by component for reasons of data availability: 1960 to 2010 for the youth bulge, 1990 to 2010 for youth unemployment, 2000 to 2010 for internet penetration, and a long-run estimated distribution for the Gini coefficient; the anocracy parameters describe a theoretical reference distribution over the Polity scale rather than an observed sample of country-years. The parameters are conventional reference values fixed before data contact, not moments estimated from a specified country set, and their provenance is documented in the replication materials (Appendix D). Their role in the classification is examined directly in Appendix C: the ranking survives every dispersion perturbation there, while the zero anchor is a property of the stated reference population.

The weights are fixed by theory, not estimated, which removes an estimation step that three cases could not support and protects the design from fitting the outcome. Equal weights of 0.25 on inequality, unemployment, and anocracy stress express the interaction model's claim that economic grievance and regime vulnerability matter comparably; the demographic term receives

0.15 because its pressure operates with a lag through the labour market; connectivity receives 0.10 as a facilitator rather than a driver. Document II specifies the full nine-component theoretical index from which this five-component implementation descends.

A worked example fixes the arithmetic. For Tunisia in 2010, the youth bulge of 18.64% standardises to $z = -0.86$ and contributes -0.13; the Gini of 42.95 gives $z = +0.37$ and $+0.09$; youth unemployment of 29.57% gives $z = +1.45$ and $+0.36$; anocracy stress of 0.6 gives $z = +0.33$ and $+0.08$; internet penetration of 36.8% gives $z = +1.09$ and $+0.11$. The sum is 0.52.

5.2 Analytical procedures

Four procedures generate the results. Group description compares 2000–2010 means across outcome classes; with two outcome groups built from three annual panels whose series are autocorrelated, formal significance tests would be uninterpretable, so the comparison is reported descriptively and no test statistics appear. Threshold classification scans candidate thresholds τ and classifies a country as predicted revolutionary when $SSI(2010) > \tau$, recording accuracy over the three cases. Counterfactual analysis recomputes the index under targeted substitutions, a Saudi Arabia reassigned Egypt's Polity score of -3, and a Tunisia reassigned Saudi Arabia's Gini of 38.3, attributing the change to the altered component. Robustness checking perturbs each weight by $\pm 30\%$ with proportional rescaling of the others, excludes each component in turn with rescaling, applies the alternative weighting schemes documented in Appendix C, applies alternative scaling rules stated in Section 6.7, and repeats the classification at the 2009 benchmark. Document II also specifies a logistic mapping from the index to event probability; its coefficients cannot be estimated from three observations, and it is recorded there without being applied here.

5.3 Causal ordering and identification

The analysis rests on stated causal assumptions rather than on identification. Temporal precedence holds by construction, since predictors are measured in 2010 and outcomes occur in 2011. Regime type is treated as exogenous to short-run economic conditions, which is defensible on path-dependence grounds: the Saudi score of -10 reflects an institutional order in place since 1932, not a response to 2010 unemployment. Reverse causation from outcome to predictors is excluded by the ordering. The vulnerable assumption is conditional independence, the absence of unmeasured variables driving both predictors and outcome, and it is almost certainly violated to some degree. Table 6 lists the salient candidates.

Table 6: Potential confounders.

Confounder	**Affects**	**Concern**
Oil wealth	E_2, E_4, P_1, outcome	Funds patronage and repression simultaneously; absent from the index
Coercive capacity	P_1, outcome	Shapes both regime persistence and protest suppression; not directly measured
Elite cohesion	P_1, outcome	Unified elites sustain autocracy and forestall defection
External support	P_1, outcome	Gulf and Western backing insulated the Saudi regime

Oil wealth is the most serious case. Part of its influence plausibly runs through regime type itself, rentier revenue helping to sustain full autocracy, which would make P_1 a mediator rather than a pure confounder; the two roles cannot be separated at $N = 3$. The study accordingly claims association consistent with the causal model of Figure 1, not causal identification. Triggering events remain exogenous throughout, as Axiom A5 requires: the framework addresses the susceptibility of systems to sparks, not the occurrence of sparks. The graph itself is assumed

rather than tested, and the instruments or discontinuities that would test it are unavailable in a three-case historical design.

6 Results

6.1 Descriptive patterns

Table 7 reports means for 2000 to 2010 alongside 2010 values, and Figure 2 plots the trajectories. Three patterns organise the decade. Tunisia carries the highest inequality and the highest youth unemployment of the three cases throughout. Saudi Arabia carries the largest youth bulge and, by 2010, the highest internet penetration. The Polity scores separate cleanly into the anocratic band (-3, -4) against the autocratic pole (-10). Connectivity rises steeply everywhere over the decade; youth unemployment is persistent everywhere; the youth bulge declines in Tunisia and Egypt while peaking and then declining in Saudi Arabia.

Table 7: Descriptive statistics by country: 2000–2010 means, with 2010 values in parentheses.

Variable	**Tunisia**	**Egypt**	**Saudi Arabia**
D_2 youth bulge (%)	20.40 (18.64)	20.32 (19.51)	22.90 (23.55)
E_2 Gini	45.20 (42.95)	35.15 (33.49)	38.35 (38.28)
E_4 youth unemployment (%)	29.61 (29.57)	27.18 (24.58)	24.32 (24.76)
P_1 Polity	-3.82 (-4)	-4.36 (-3)	-10.00 (-10)
S_3 internet penetration (%)	15.04 (36.8)	11.11 (21.6)	18.97 (41.0)
Outcome	Revolution	Revolution	Stability

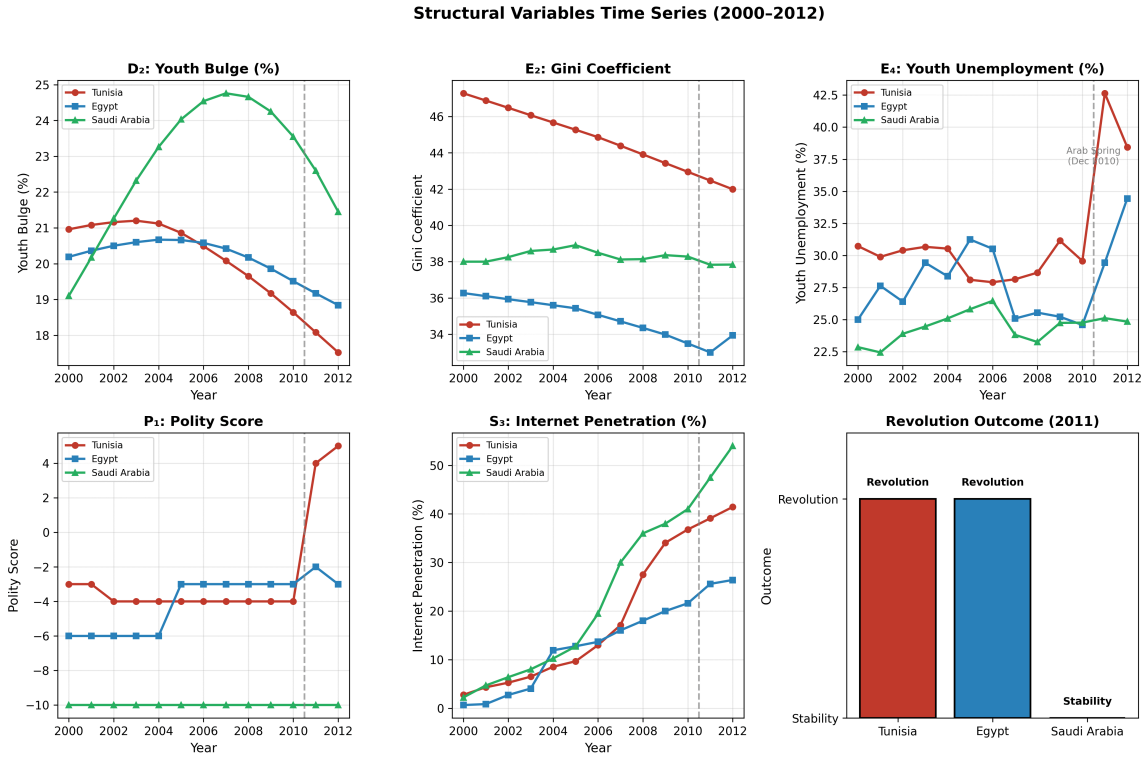

Figure 2: Variable trajectories by country, 2000 to 2012, for the five index components.

6.2 Case profiles and component decomposition

The 2010 index values are 0.52 for Tunisia, 0.10 for Egypt, and −0.09 for Saudi Arabia; Figure 3 shows the decompositions. Tunisia's index is carried by youth unemployment (+0.36), with internet penetration (+0.11), inequality (+0.09), and anocracy stress (+0.08) adding smaller positive terms against a negative youth bulge contribution (−0.13). The country with the highest structural stress is the one where the cascade began.

Egypt's positive but modest index combines youth unemployment (+0.21) and anocracy stress (+0.17) with near-neutral connectivity (+0.03) against negative contributions from the youth bulge (−0.10) and from inequality (−0.20), which sits below the regional reference mean. An index near the regional average sits awkwardly against the outcome, and one reading, offered as hypothesis rather than result, is that Egypt's position in the cascade mattered: a follower case in which Tunisia's demonstration effect lowered the effective mobilisation threshold. The index contains no regional-contagion term, so this reading lies outside what the instrument can show.

Saudi Arabia's decomposition explains the stable outcome. Its positive stress terms are real and substantial, youth unemployment (+0.21), connectivity (+0.13), and youth bulge (+0.04), and its inequality term is mildly negative (−0.05). The anocracy component, at −0.42, overwhelms them: full autocracy enters the index as active suppression capacity, and the net index turns negative. Anticipating Section 6.7, removing that single component would leave Saudi Arabia the highest-stress country in the sample at +0.43.

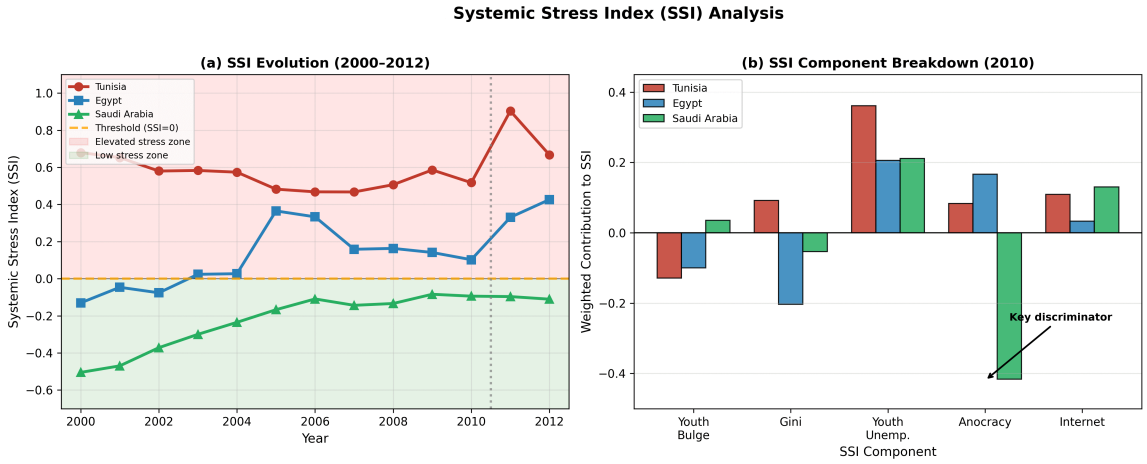

Figure 3: Component decomposition of the 2010 SSI for the three cases. The anocracy term separates Saudi Arabia from the revolutionary cases.

6.3 Group comparison

Table 8 compares 2000–2010 means across outcome classes. As stated in Section 5.2, the comparison is descriptive: no significance tests are reported, because none would be interpretable on two outcome groups built from three annual panels with autocorrelated series.

Table 8: Group means, 2000–2010, by outcome class (descriptive). Differences are computed on unrounded means.

Variable	Revolutionary cases	Stable case	Difference
D_2 youth bulge (%)	20.36	22.90	−2.54
E_2 Gini	40.18	38.35	+1.83
E_4 youth unemployment (%)	28.40	24.32	+4.07
P_1 Polity	−4.09	−10.00	+5.91
S_3 internet penetration (%)	13.08	18.97	−5.89

Only two of the five variables differ in the direction a stress-only account expects: youth unemployment and regime type. The youth bulge and internet penetration run the other way, higher in the stable case, and the inequality difference is modest and driven by Tunisia. The sign pattern already carries the paper's central lesson: single-variable stress stories misorder these cases, and whatever discriminating power exists must come from the political term or from a composite that respects it.

6.4 Threshold classification

We use the term classification assessment rather than predictive validation: with $N = 3$ the question is whether the index correctly classifies known outcomes, not whether it would have predicted them ex ante. Table 9 scans candidate thresholds. The separating region is a continuum: any threshold in the interval $[-0.09, 0.10)$ classifies all three cases correctly. The analysis reports the zero threshold because it is the reference-population mean fixed a priori, not a cutpoint selected from the separating interval; thresholds below miss the stable case, thresholds above miss Egypt. The substantive reading is that both revolutionary cases carried above-average structural stress for the region while the stable case sat below average. The result is suggestive, not significant: under random scores a separating threshold would exist with probability 1/3 and the exact ordering observed with probability 1/6, as derived in Section 1.3.

Table 9: Threshold classification, 2010.

Threshold	Tunisia	Egypt	Saudi Arabia	Accuracy
$SSI > -0.50$	Revolution	Revolution	Revolution	66.7%
$SSI > 0.00$	Revolution	Revolution	Stability	100%
$SSI > 0.25$	Revolution	Stability	Stability	66.7%
$SSI > 0.50$	Revolution	Stability	Stability	66.7%

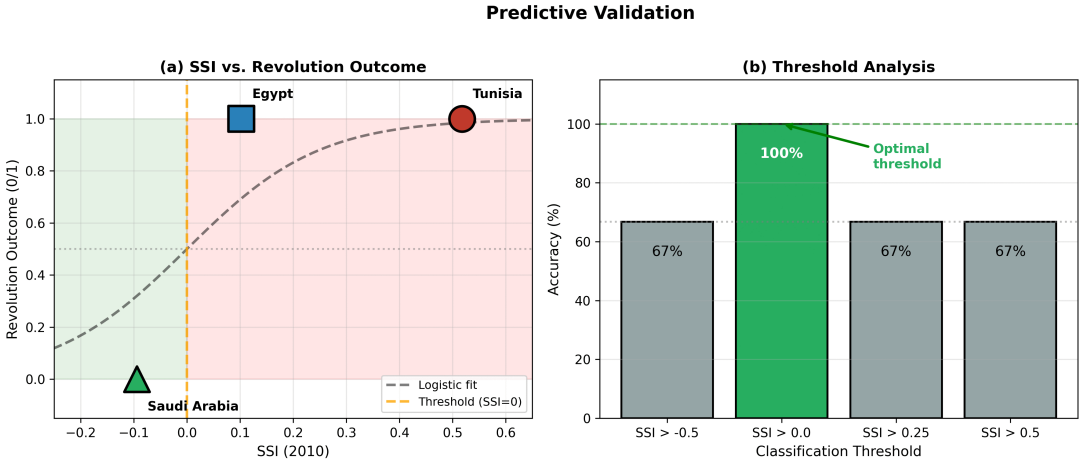

Figure 4: SSI values against outcomes, with the zero threshold separating the outcome classes.

6.5 Counterfactual scenarios

Two substitutions isolate component effects. In the first, Saudi Arabia keeps its 2010 economic, demographic, and connectivity values but takes Egypt's Polity score of -3. The anocracy contribution moves from -0.42 to $+0.17$, and the index rises from -0.09 to $+0.49$, a shift of $+0.58$ that carries the kingdom across the classification threshold into the revolutionary zone. Regime structure alone, with every grievance variable held fixed, separates the stable case from the revolutionary class. The claim is about the index's comparative statics: it shows how the specification trades regime structure against grievance, not that an anocratic Saudi Arabia would in fact have revolted, since a different regime history would have reshaped the economic variables themselves.

In the second, Tunisia keeps everything except inequality, reassigned Saudi Arabia's Gini of 38.3. The inequality contribution moves from $+0.09$ to -0.05 and the index falls from 0.52 to 0.37, still positive and still above Egypt. Tunisia's revolutionary configuration was overdetermined: no plausible single-variable change in the economic block would have moved it out of the risk zone.

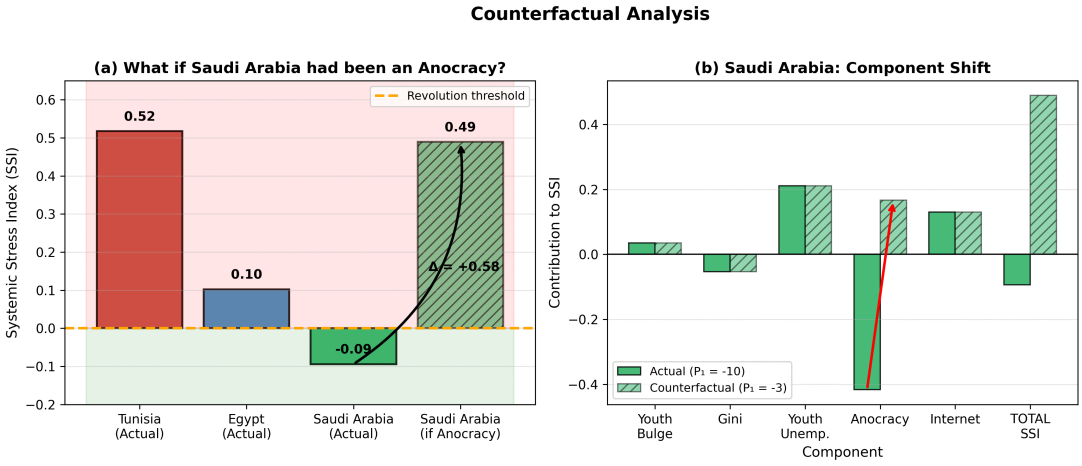

Figure 5: Actual and counterfactual SSI for Saudi Arabia. Reassigning the Polity score from -10 to -3 raises the index by $+0.58$, across the classification threshold.

6.6 Single-variable baselines

A composite index earns its keep only if it beats simpler alternatives. Table 10 assesses each component as a stand-alone classifier of the 2010 outcomes.

Table 10: Single-variable baselines, 2010.

Baseline	**Tunisia / Egypt / Saudi Arabia**	**Separates?**	**Assessment**
Polity score	-4 / -3 / -10	Yes	A threshold near -5 separates these cases, but a linear reading rates full democracies as highest risk; the inverted-U transformation is theoretically required
Anocracy stress	0.60 / 0.70 / 0.00	Yes	Matches the SSI classification in this sample
Youth unemployment (%)	29.6 / 24.6 / 24.8	No	Stable case exceeds one revolutionary case
Youth bulge (%)	18.6 / 19.5 / 23.6	No	Ordered against the outcome
Internet penetration (%)	36.8 / 21.6 / 41.0	No	Highest in the stable case

The economic, demographic, and connectivity baselines fail outright, which is the quantitative face of the puzzle in Section 1.1. The regime-based measures succeed, and at three cases the composite cannot be shown to outperform anocracy stress alone. What the composite adds within this sample is gradient information, since the ordering Tunisia above Egypt matches the cascade's origin and severity where anocracy stress alone would order them the other way, discrimination within the anocratic class by grievance load, and theoretical completeness, since a regime-only classifier is silent about the stress side of the interaction. One caveat from the robustness analysis belongs here: the composite's advantage is conditional on its theory-derived weights. Under equal or economic-emphasis weighting the composite fails where anocracy stress alone would still succeed (Appendix C), so composite measurement is not automatically superior to a single well-chosen variable; the weights carry substantive content.

6.7 Robustness

Weight perturbation comes first. Each weight in turn is moved by $\pm 30\%$ with the others rescaled proportionally to sum to one; Table 11 reports the recomputed 2010 values.

Table 11: Weight perturbation $\pm 30\%$, 2010.

Scenario	**Tunisia**	**Egypt**	**Saudi Arabia**	**Separates?**
Baseline	0.52	0.10	-0.09	Yes
E_2 $+30\%$	0.50	0.01	-0.11	Yes
E_2 -30%	0.53	0.19	-0.08	Yes
E_4 $+30\%$	0.61	0.17	0.00	Boundary
E_4 -30%	0.42	0.03	-0.19	Yes
P_1 $+30\%$	0.50	0.16	-0.25	Yes
P_1 -30%	0.54	0.05	$+0.06$	No

The two failure modes share one cause. Saudi Arabia's economic, demographic, and connectivity z-scores are positive, and the -0.42 anocracy term is the only mechanism holding the

country's index below the threshold. Reducing the anocracy weight by 30% lets the positive terms dominate and the classification fails; raising the unemployment weight by 30% dilutes the anocracy share and pushes the Saudi index to the boundary. The sensitivity is substantive rather than numerical: it restates H2, that a specification which treats regime vulnerability as one stressor among equals will misclassify stable autocracies.

Component exclusion sharpens the point. Each component is removed in turn, with remaining weights rescaled; Table 12 reports the outcome. Only the anocracy exclusion breaks the classification, and it breaks it dramatically, promoting Saudi Arabia to the highest-stress country in the sample (+0.43). The marginal row for youth unemployment reflects range compression rather than reordering: all three countries share elevated unemployment, so removing it shrinks the spread without changing signs.

Table 12: Leave-one-component-out, 2010 (weights rescaled).

Excluded component	**Tunisia**	**Egypt**	**Saudi Arabia**	**Separates?**
None (baseline)	0.52	0.10	−0.09	Yes
D_2 youth bulge	0.76	0.24	−0.15	Yes
E_2 inequality	0.57	0.41	−0.05	Yes
E_4 youth unemployment	0.21	−0.14	−0.41	Marginal (2 of 3 at zero threshold)
S_3 internet penetration	0.45	0.08	−0.25	Yes
Anocracy stress	0.58	−0.09	+0.43	No

Alternative weighting schemes point the same way. Under equal weights of 0.20 the Saudi index turns positive (+0.10) and classification fails; under an economic-emphasis scheme it fails again (+0.08); under a regime-emphasis scheme with anocracy at 0.40 it holds comfortably (−0.35). Appendix C reports the full grid. The baseline anocracy weight of 0.25 is, if anything, conservative.

Alternative scaling rules were recomputed for this version under explicitly stated definitions, replacing a table inherited from the original deposit whose values could not be reproduced from the data panel (Appendix E). Two standard alternatives to reference-population z-scores were applied on the pooled three-country panel for 2000 to 2010. Under min-max normalisation of each component to the unit interval, the ranking survives (Tunisia 0.68, Egypt 0.38, Saudi Arabia 0.37) but the Egypt-Saudi margin nearly vanishes, because bounded non-negative scaling suppresses the negative anocracy contribution that carries the separation under z-scores. Under median and interquartile-range standardisation, with quartiles computed by linear interpolation (the type-7 convention) on the pooled 33-observation panel, classification fails outright (Tunisia 0.19, Egypt −0.30, Saudi Arabia 0.25): the pooled panel is a different reference population from MENA 1960–2010, and the compressed anocracy spread no longer offsets Saudi Arabia's positive terms. The lesson is that the reference population and scaling rule are part of the index design, not neutral conveniences; the MENA-referenced z-score construction is the specification the theory motivates, and departures from it degrade the discrimination for reasons the theory itself predicts.

Temporal stability closes the battery. At the 2009 benchmark the values are 0.59, 0.14, and −0.08, producing identical classifications; the changes into 2010 are −0.07, −0.04, and −0.01. The structural configuration that separates the cases was in place before the trigger, which is what a structural account requires.

Table 13: Robustness summary.

Test	Result	Reading
Weight perturbation ±30%	Holds except P_1 −30% (fails) and E_4 +30% (boundary)	The anocracy weight is load-bearing
Leave-one-component-out	Only anocracy exclusion breaks classification	Anocracy is the indispensable component
Alternative weighting schemes	Regime emphasis holds; equal and economic emphasis fail	Theory-derived weighting is substantive
Alternative scaling rules	Ranking survives min-max; classification fails under pooled median/IQR	The reference population is part of the design
Benchmark year 2009	Identical classifications	The signal precedes the trigger

6.8 Hypothesis assessment

Table 14 gathers the verdicts. The index ranks the cases by outcome severity and classifies all three at the zero threshold (H1). Its discriminating power collapses when the anocracy component is removed or materially down-weighted (H2). The stable case matches the revolutionary cases on economic stress without revolting (H3; the design demonstrates insufficiency, while the necessity half of the claim is untestable at $N = 3$). And the counterfactual reassignment of regime type alone moves Saudi Arabia across the threshold (H4). All four verdicts stand within the exploratory limits of Section 1.3; none is a statistical inference.

Table 14: Hypothesis assessment.

Hypothesis	Finding	Verdict
H1 Discrimination	Ranking Tunisia > Egypt > Saudi Arabia; zero threshold classifies 3 of 3	Supported
H2 Anocracy primacy	Removing or down-weighting anocracy breaks classification	Supported
H3 Insufficiency of economic stress	Stable case matches revolutionary cases on economic stress	Supported (necessity untestable at $N = 3$)
H4 Counterfactual shift	Saudi Arabia under anocracy: $\Delta SSI = +0.58$, crosses threshold	Supported

7 Discussion

7.1 What the results show

Four findings organise the interpretation. First, the index discriminates: it ranks the three cases in order of outcome severity, and the threshold that separates the classes is the reference-population mean rather than a tuned cutpoint. The success is not built in. Four of the five components would misclassify the cases used alone, and two of them, youth bulge and connectivity, run against the outcome. Second, regime type does the discriminating work, and it does so as a moderator rather than as one more addend: the Saudi decomposition shows large positive stress terms neutralised by a single strongly negative political term. Third, economic stress is demonstrably not sufficient: grievance levels comparable to Egypt's coexisted with stability where the regime structure gave mobilisation no purchase. Necessity is consistent with the sample, every revolutionary case here is stressed, but the design contains no low-stress revolutionary case and cannot test it. Fourth, the counterfactuals size the asymmetry: reassigning regime type moves the Saudi index by +0.58,

while reassigning Tunisia's inequality to the Saudi level moves the Tunisian index by -0.15. Within this specification, regime structure moves the index roughly four times as far as a large single-variable economic change.

7.2 Implications for the CMH programme

For the framework itself, the study demonstrates end-to-end applicability: canonical variable definitions survived contact with real sources, the index computed from theory-fixed weights produced interpretable output, and the classification exercise could have failed. Had Saudi Arabia scored high or Tunisia low, the specification would have been contradicted within its own terms, so the exercise was falsifiable in the weak, within-sample sense that A8 requires of every CMH product. The axioms bearing on design held their assigned roles: national aggregates behaved as A1 presumes at these population scales, the MENA reference population gave A2's comparability assumption operational content, and the pattern of component contributions gave A3's structural-precondition claim a concrete instance.

One design insight carries over to the wider programme. Allowing signed contributions, so that consolidated autocracy subtracts stress rather than adding none, is what lets a linear index encode the signed regime effect that the interaction model predicts; it does not let the index express the interaction itself, which no weighted sum can represent. Suppression capacity behaves as an active force, not as the absence of vulnerability, and an index that cannot go negative on the political term cannot represent that fact, as the min-max scaling exercise of Section 6.7 showed concretely. Within the series' hierarchy of claims, everything here remains a Level 4 operationalisation generating regime-level, retrospective statements; graduation to validated regime-probability claims requires the sample sizes, out-of-sample splits, and calibration scoring that Theorem 4 of Document I attaches to A8, none of which three cases can supply.

7.3 The anocracy-autocracy divide and the place of oil

The mechanisms separating scores of -3 and -4 from -10 are visible in the case narratives. Repression in the anocracies was harsh but incomplete: Tunisian security forces killed hundreds without holding the streets, while the Saudi deployment made the cost of assembly prohibitive before assembly occurred. Organisational space differed in kind: the UGTT and the Brotherhood's networks gave Tunisian and Egyptian oppositions infrastructure that Saudi civil society, foreclosed by design, never developed. Legitimacy structures differed: regimes that stage elections while rigging them convert their own institutions into standing evidence of hypocrisy, while the Saudi order claims religious and dynastic authority without democratic pretence, leaving protest to be framed as sedition rather than as the redemption of promised rights. And elite cohesion differed: conscript armies with national identities declined to fire on compatriots, while a dynastic security structure with a dedicated praetorian guard did not fracture.

Oil runs through all of this without appearing in the index. Rentier revenue funded the 130 billion dollar package that met grievance with cash, and it funds the security apparatus that meets assembly with force; hereditary rule and hydrocarbon wealth jointly predicted survival across the region (Brownlee et al., 2015). Whether oil operates alongside regime type or through it, sustaining the autocratic pole of the Polity scale itself, cannot be resolved with three cases, and Section 8.2 records the omission as the specification's most serious.

7.4 Connectivity as facilitator

The connectivity finding is the study's most counterintuitive and its easiest to misread. Internet penetration fails as a discriminator, highest in the stable case, and receives the index's smallest weight, yet no account of how the Tunisian and Egyptian mobilisations unfolded can ignore the platforms. The resolution is the distinction between occurrence and form. Digital infrastructure

shaped the speed, coordination, and visibility of mobilisation where mobilisation happened (Howard & Hussain, 2013; Tufekci & Wilson, 2012); it did not determine whether mobilisation happened, and in the Saudi case the same infrastructure served surveillance and preemption. A conditional reading fits the evidence: in anocracies, where organisational space exists and repression is incomplete, connectivity amplifies mobilisation; in consolidated autocracies it amplifies monitoring. That reading is consistent with the sceptical literature on digital activism (Gladwell, 2010; Morozov, 2011) without dismissing the platforms' tactical importance, and it is the reading the facilitation weight of 0.10 encodes.

7.5 Relation to scholarship and policy

Substantively, the results are consistent with the regime-centred accounts of the Arab Spring (Bellin, 2012; Brownlee et al., 2015) and with structural-demographic theory's insistence on the conjunction of pressure and state weakness (Goldstone, 1991; Turchin, 2016), and they cut against technology-determinist readings. The contribution is not a new causal story but the integration of the standing ones: a single measured instrument in which grievance, demography, connectivity, and regime vulnerability enter with declared weights, applied identically across cases, with every number reproducible from deposited data.

The policy horizon should be stated with restraint. An index of this family could in principle support structural monitoring, tracking country trajectories, flagging threshold approaches, and decomposing which components drive movement. Deployment would be premature. The framework is unvalidated at operational sample sizes; its output is a probability-relevant classification, never an event forecast, by Axioms A5 and A7; published warnings feed back on the systems they describe, so prospective use must check the reflexivity condition that Theorem 3 of Document I formalises; and the underlying data arrive with one to two years of latency. Any eventual operational product would report intervals and drivers, never dates and never point estimates.

8 Limitations and future research

8.1 Sample size and post-hoc character

The controlling limitation is the sample. With three cases there are no degrees of freedom for estimation: regression coefficients, confidence intervals, formal tests, cross-validation, and out-of-sample assessment are all unavailable, and the classification result of Section 6.4 carries the chance probability already stated there. The analysis is also retrospective twice over. It was conducted in knowledge of the outcomes, so variable selection and specification could in principle have been steered by hindsight; fixing weights and definitions from prior theory mitigates that risk without eliminating it, and only out-of-sample application can retire it. And it is a classification of known history, not a demonstration that anyone would have predicted the wave in 2010. Case selection compounds the limits: the three cases were chosen for outcome variation and analytic clarity, exclude the ambiguous middle of the regional distribution, and cannot represent it. What survives these constraints is deliberately modest. The index classified three known cases correctly, ordered them by outcome severity, identified its own load-bearing component, and did so through documented, replicable procedures. What is not claimed, and cannot be claimed from this design, is validation, predictive power, causal identification, or generalisability.

8.2 Specification limits

The five-component implementation is a reduction of the nine-component theoretical index of Document II, and the omissions are not innocent. Of the theoretical components, inflation (E_3), growth relief (E_5), state capacity (P_5), regime legitimacy (Ψ_1), and the mobilisation-pressure terms (Ψ_3, Ψ_5) are absent for data reasons, with the anocracy transformation standing in for

the legitimacy and capacity block. Beyond the theoretical set, three omissions weigh on the present cases. Oil wealth funds patronage and repression at once and differentiates Saudi Arabia from both revolutionary cases; a rents-per-capita term belongs in any expanded specification. Coercive capacity is measured only by proxy through the regime score, though the outcomes turned on whether security forces held. Elite cohesion, the difference between conscript armies that defect and praetorian structures that do not, is consequential and hard to measure before the fact. External intervention, which decided outcomes in Bahrain and Libya, lies outside a domestic-variables framework altogether.

8.3 Theoretical limits

Three limits are theoretical rather than practical. The framework addresses susceptibility, not sparks: triggering events remain exogenous under Axiom A5, so even a fully validated index would speak to conditional structural risk, never to timing. The mechanism behind the anocracy effect is underspecified: repressive capacity, organisational space, legitimacy structure, and elite cohesion all plausibly contribute, the case narratives illustrate each, and this design cannot apportion their shares. And the linear aggregation excludes threshold effects and interactions inside the index itself; the interaction that matters here enters through the signed anocracy component, and Document II defers richer nonlinearities to the event model, where they can eventually be estimated rather than imposed.

8.4 Generalisability

The findings are bounded in space, time, and outcome type: three MENA countries, one decade, one outcome category, classical revolution as defined in Section 4.1. Other regions, other periods, and other forms of instability, civil wars, coups, negotiated transitions, may follow different structural logics, and nothing here licenses extrapolation to them.

8.5 From proof of concept to validation

The path forward was specified by this study's own limits, and its first step has since been taken. The eleven-country extension (Angeli, 2026c) applied the same five-component index across the full spectrum of MENA outcomes at the 2010 benchmark, classifying eight of eleven countries correctly, and converted its misclassifications into specification lessons rather than random error. The later stages stand as planned: a MENA panel for 1990 to 2020 with a few hundred country-years; formal modelling with genuine degrees of freedom, logistic and survival specifications, receiver operating characteristic comparison against baselines, and cross-validation; eventually, pre-registered out-of-sample tests on events postdating the training window, with the fifteen-variable index of Document II conditioned on reaching thirty or more validated country-cases. Only that programme can validate the framework. This study's role was to establish that the programme is worth running.

9 Conclusion

This paper asked whether the operational machinery of Computational Macrohistory can be applied to real historical data and produce interpretable, falsifiable output at the regime level of the framework's prediction hierarchy. Applied to the divergent trajectories of Tunisia, Egypt, and Saudi Arabia through the Arab Spring, the five-component Systemic Stress Index ranked the cases in order of outcome severity (0.52, 0.10, -0.09), classified all three correctly at the reference-mean threshold, and located the discrimination in one theoretically motivated place: the anocracy component, whose removal or material down-weighting collapses the classification, and whose counterfactual manipulation alone moves Saudi Arabia across the threshold by $+0.58$.

Economic stress behaved as the theory requires on the side the design can test, present in all three cases and sufficient in none, and connectivity behaved as a facilitator rather than a driver.

Read as evidence about the Arab Spring, the results join the regime-centred strand of the literature and give it quantitative expression: grievance was regional, vulnerability was not, and the cases separate on vulnerability. Read as evidence about CMH, they establish something narrower and more useful. The canonical definitions survive contact with real sources; the theory-fixed weights carry substantive content, since the specifications that abandon them fail; the procedures replicate end to end from deposited data; and the exercise could have failed in a way that would have counted against the framework.

The claims stop where the design stops. Three cases validate nothing, predict nothing, and identify nothing causally. They demonstrate applicability, and they generate structured expectations, the strongest of which, the primacy of regime vulnerability over accumulated grievance, is the hypothesis the expanded programme was designed to test. That programme is underway: the eleven-country study has since replicated the discrimination pattern at 72.7% descriptive accuracy, and the panel designs that would support genuine inference are specified. The framework's wager is that structural susceptibility is measurable even though sparks are not. Nothing in three cases settles that wager, and nothing in them embarrasses it; the sample that can decide it is the next task of the series.

References

Achcar, G. (2013). *The People Want: A Radical Exploration of the Arab Uprising*. University of California Press. https://doi.org/10.1525/9780520956544

Anderson, L. (2011). Demystifying the Arab Spring: Parsing the differences between Tunisia, Egypt, and Libya. *Foreign Affairs*, 90(3), 2–7. https://www.foreignaffairs.com/articles/libya/2011-04-03/demystifying-arab-spring

Angeli, S. (2026a). *Computational Macrohistory: Axiomatic foundations for a quantitative science of large-scale social systems* (FICSS Working Paper No. WP-2026-001, v3.0). Foundations Institute of Computational Social Science. https://doi.org/10.5281/zenodo.21178550

Angeli, S. (2026b). *Computational Macrohistory: Operational framework. State space, variables, and dynamic equations* (FICSS Working Paper No. WP-2026-002, v2.2). Foundations Institute of Computational Social Science. https://doi.org/10.5281/zenodo.18646832

Angeli, S. (2026c). *Structural stress and political instability in the MENA region: A computational macrohistory analysis of the Arab Spring across eleven countries, 2000–2012* (FICSS Working Paper No. WP-2026-004). Foundations Institute of Computational Social Science. https://doi.org/10.5281/zenodo.19661257

Angeli, S. (2026d). *Beyond the Lyapunov Wall: Extending the predictive horizon of computational macrohistory* (FICSS Working Paper No. WP-2026-005). Foundations Institute of Computational Social Science. https://doi.org/10.5281/zenodo.20111272

Angrist, M. P. (2013). Understanding the success of mass civic protest in Tunisia. *Middle East Journal*, 67(4), 547–564. https://doi.org/10.3751/67.4.13

Beblawi, H., & Luciani, G. (Eds.). (1987). *The Rentier State*. Croom Helm. https://doi.org/10.4324/9781315684864

Beinin, J. (2012). The rise of Egypt's workers. *Carnegie Papers*, Carnegie Endowment for International Peace. https://carnegieendowment.org/research/2012/06/the-rise-of-egypts-workers

Bellin, E. (2004). The robustness of authoritarianism in the Middle East: Exceptionalism in comparative perspective. *Comparative Politics*, 36(2), 139–157. https://doi.org/10.2307/4150140

Bellin, E. (2012). Reconsidering the robustness of authoritarianism in the Middle East: Lessons from the Arab Spring. *Comparative Politics*, 44(2), 127–149. https://doi.org/10.5129/

001041512798838021
Brinton, C. (1938). *The Anatomy of Revolution*. W. W. Norton.
Brownlee, J., Masoud, T., & Reynolds, A. (2015). *The Arab Spring: Pathways of Repression and Reform*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199660063.001.0001
Campante, F. R., & Chor, D. (2012). Why was the Arab world poised for revolution? Schooling, economic opportunities, and the Arab Spring. *Journal of Economic Perspectives*, 26(2), 167–188. https://doi.org/10.1257/jep.26.2.167
Cook, S. A. (2012). *The Struggle for Egypt: From Nasser to Tahrir Square*. Oxford University Press. https://global.oup.com/academic/product/the-struggle-for-egypt-9780199931774
Davies, J. C. (1962). Toward a theory of revolution. *American Sociological Review*, 27(1), 5–19. https://doi.org/10.2307/2089714
Dhillon, N., & Yousef, T. (Eds.). (2009). *Generation in Waiting: The Unfulfilled Promise of Young People in the Middle East*. Brookings Institution Press. https://www.brookings.edu/books/generation-in-waiting/
Escribà-Folch, A., & Wright, J. (2015). Human rights prosecutions and autocratic survival. *International Organization*, 69(2), 343–373. https://doi.org/10.1017/S0020818314000484
Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1), 75–90. https://doi.org/10.1017/S0003055403000534
Foran, J. (Ed.). (2005). *Taking Power: On the Origins of Third World Revolutions*. Cambridge University Press. https://doi.org/10.1017/CBO9780511488962
Gates, S., Hegre, H., Jones, M. P., & Strand, H. (2006). Institutional inconsistency and political instability: Polity duration, 1800–2000. *American Journal of Political Science*, 50(4), 893–908. https://doi.org/10.1111/j.1540-5907.2006.00222.x
Geddes, B., Wright, J., & Frantz, E. (2014). Autocratic breakdown and regime transitions: A new data set. *Perspectives on Politics*, 12(2), 313–331. https://doi.org/10.1017/S1537592714000851
Gladwell, M. (2010, October 4). Small change: Why the revolution will not be tweeted. *The New Yorker*. https://www.newyorker.com/magazine/2010/10/04/small-change-3
Goldstone, J. A. (1991). *Revolution and Rebellion in the Early Modern World*. University of California Press.
Goldstone, J. A. (2001). Toward a fourth generation of revolutionary theory. *Annual Review of Political Science*, 4(1), 139–187. https://doi.org/10.1146/annurev.polisci.4.1.139
Gurr, T. R. (1970). *Why Men Rebel*. Princeton University Press.
Hegre, H., Ellingsen, T., Gates, S., & Gleditsch, N. P. (2001). Toward a democratic civil peace? Democracy, political change, and civil war, 1816–1992. *American Political Science Review*, 95(1), 33–48. https://doi.org/10.1017/S0003055401000119
Hibou, B. (2011). *The Force of Obedience: The Political Economy of Repression in Tunisia*. Polity Press.
Howard, P. N., & Hussain, M. M. (2013). *Democracy's Fourth Wave? Digital Media and the Arab Spring*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199936953.001.0001
Jones, T. C. (2011). Saudi Arabia versus the Arab Spring. *Raritan*, 31(2), 43–59.
Kandil, H. (2012). *Soldiers, Spies, and Statesmen: Egypt's Road to Revolt*. Verso Books.
LaGraffe, D. (2012). The youth bulge in Egypt: An intersection of demographics, security, and the Arab Spring. *Journal of Strategic Security*, 5(2), 65–80. http://dx.doi.org/10.5038/1944-0472.5.2.4
Lijphart, A. (1971). Comparative politics and the comparative method. *American Political Science Review*, 65(3), 682–693. https://doi.org/10.2307/1955513

Lynch, M. (2012). *The Arab Uprising: The Unfinished Revolutions of the New Middle East*. PublicAffairs.

Lynch, M. (2016). *The New Arab Wars: Uprisings and Anarchy in the Middle East*. PublicAffairs.

Marshall, M. G., & Gurr, T. R. (2020). *Polity5: Political Regime Characteristics and Transitions, 1800–2018*. Center for Systemic Peace.

Morozov, E. (2011). *The Net Delusion: The Dark Side of Internet Freedom*. PublicAffairs.

Przeworski, A., & Teune, H. (1970). *The Logic of Comparative Social Inquiry*. Wiley-Interscience.

Skocpol, T. (1979). *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815805

Springborg, R. (2017). *Egypt*. Polity Press. https://www.politybooks.com/bookdetail?book_slug=egypt--9781509520480

Stepan, A. (2012). Tunisia's transition and the twin tolerations. *Journal of Democracy*, 23(2), 89–103. https://doi.org/10.1353/jod.2012.0031

Tufekci, Z., & Wilson, C. (2012). Social media and the decision to participate in political protest: Observations from Tahrir Square. *Journal of Communication*, 62(2), 363–379. https://doi.org/10.1111/j.1460-2466.2012.01629.x

Turchin, P. (2003). *Historical Dynamics: Why States Rise and Fall*. Princeton University Press.

Turchin, P. (2010). Political instability may be a contributor in the coming decade. *Nature*, 463(7281), 608. https://doi.org/10.1038/463608a

Turchin, P. (2016). *Ages of Discord: A Structural-Demographic Analysis of American History*. Beresta Books.

Turchin, P., Currie, T. E., Turner, E. A. L., & Gavrilets, S. (2015). Seshat: The Global History Databank. *Cliodynamics*, 6(1), 77–107. https://doi.org/10.21237/C7clio6127917

Turchin, P., Currie, T. E., Whitehouse, H., François, P., Feeney, K., Mullins, D., et al. (2018). Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization. *Proceedings of the National Academy of Sciences*, 115(2), E144–E151. https://doi.org/10.1073/pnas.1708800115

Turchin, P., & Nefedov, S. A. (2009). *Secular Cycles*. Princeton University Press.

Urdal, H. (2006). A clash of generations? Youth bulges and political violence. *International Studies Quarterly*, 50(3), 607–629. https://doi.org/10.1111/j.1468-2478.2006.00416.x

Vreeland, J. R. (2008). The effect of political regime on civil war: Unpacking anocracy. *Journal of Conflict Resolution*, 52(3), 401–425. https://doi.org/10.1177/0022002708315594

A Complete data tables

Table 15: Tunisia, 2000–2012.

Year	D_2 (%)	E_2 **(Gini)**	E_4 (%)	P_1	S_3 (%)	**SSI**
2000	20.96	47.28	30.73	−3	2.75	0.68
2001	21.08	46.88	29.89	−3	4.30	0.65
2002	21.16	46.48	30.40	−4	5.25	0.58
2003	21.20	46.07	30.67	−4	6.49	0.58
2004	21.12	45.67	30.53	−4	8.53	0.57
2005	20.86	45.27	28.10	−4	9.66	0.48
2006	20.49	44.86	27.91	−4	12.99	0.47
2007	20.08	44.39	28.14	−4	17.10	0.47
2008	19.65	43.91	28.66	−4	27.53	0.51
2009	19.17	43.43	31.15	−4	34.07	0.59
2010	18.64	42.95	29.57	−4	36.80	0.52
2011	18.08	42.47	42.64	4	39.10	0.90
2012	17.52	41.99	38.44	5	41.44	0.67

Table 16: Egypt, 2000–2012.

Year	D_2 (%)	E_2 **(Gini)**	E_4 (%)	P_1	S_3 (%)	**SSI**
2000	20.19	36.27	24.99	−6	0.64	−0.13
2001	20.36	36.10	27.63	−6	0.84	−0.05
2002	20.50	35.93	26.40	−6	2.72	−0.08
2003	20.60	35.77	29.43	−6	4.04	0.02
2004	20.67	35.60	28.37	−6	11.92	0.03
2005	20.66	35.43	31.24	−3	12.75	0.37
2006	20.58	35.07	30.52	−3	13.66	0.33
2007	20.42	34.71	25.08	−3	16.03	0.16
2008	20.17	34.35	25.54	−3	18.01	0.16
2009	19.86	33.98	25.22	−3	20.00	0.14
2010	19.51	33.49	24.58	−3	21.60	0.10
2011	19.17	33.00	29.43	−2	25.60	0.33
2012	18.84	33.94	34.44	−3	26.40	0.43

Table 17: Saudi Arabia, 2000–2012.

Year	D_2 (%)	E_2 **(Gini)**	E_4 (%)	P_1	S_3 (%)	**SSI**
2000	19.10	38.00	22.85	−10	2.21	−0.50
2001	20.17	38.00	22.44	−10	4.68	−0.47
2002	21.26	38.25	23.89	−10	6.38	−0.37
2003	22.32	38.59	24.46	−10	8.00	−0.30
2004	23.26	38.67	25.09	−10	10.23	−0.24
2005	24.03	38.91	25.81	−10	12.71	−0.17
2006	24.54	38.49	26.46	−10	19.46	−0.11
2007	24.76	38.12	23.81	−10	30.00	−0.14
2008	24.66	38.14	23.25	−10	36.00	−0.13
2009	24.25	38.36	24.74	−10	38.00	−0.08
2010	23.55	38.28	24.76	−10	41.00	−0.09
2011	22.60	37.83	25.11	−10	47.50	−0.10
2012	21.45	37.84	24.84	−10	54.00	−0.11

B Calculation procedure

The index for a country-year is computed in three steps. The Polity score is transformed to anocracy stress, $A = 1 - |P_1|/10$. Each of the five components is standardised as $z = (X - \mu)/\sigma$ against the reference parameters of Table 5. The standardised components are combined as

$$SSI = 0.15\, z(D_2) + 0.25\, z(E_2) + 0.25\, z(E_4) + 0.25\, z(A) + 0.10\, z(S_3). \tag{8}$$

Every SSI value reported in this paper, including all robustness variants, is reproducible to the reported two decimal places from the data of Appendix A, the parameters of Table 5, and the weight specifications of Sections 5.1 and 6.7 and Appendix C, using any spreadsheet or statistical software. All values were re-verified by independent recomputation for this version.

C Alternative weighting schemes and parameter sensitivity

Table 18 reports the 2010 index under the alternative weighting schemes referenced in Section 6.7. All schemes sum to 1.00. Versions 1.0 to 3.0 specified the economic and regime schemes with weights summing to 1.01; version 3.1 normalises them by trimming the connectivity weight, the component with the smallest theoretical weight, by one percentage point (S_3 from 0.08 to 0.07 in the economic scheme and from 0.15 to 0.14 in the regime scheme). The adjustment subtracts $0.01 \times z(S_3)$ from each country's index (a shift of −0.013 for Saudi Arabia in 2010), so the values differ from earlier releases by that component-specific amount rather than by a uniform rescaling; rankings, thresholds, and classifications are unaffected.

Table 18: Alternative weighting schemes, 2010.

Scheme	D_2	E_2	E_4	**Anocracy**	S_3	**Tunisia**	**Egypt**	**Saudi Arabia**	**Separates?**
Baseline	0.15	0.25	0.25	0.25	0.10	0.52	0.10	−0.09	Yes
Equal	0.20	0.20	0.20	0.20	0.20	0.48	0.07	+0.10	No
Economic emphasis	0.08	0.35	0.35	0.15	0.07	0.69	0.07	+0.08	No
Regime emphasis	0.10	0.18	0.18	0.40	0.14	0.53	0.25	−0.35	Yes

Sensitivity to the reference parameters of Table 5 follows two regimes. Perturbing a reference mean shifts every country's index by the same constant for that component, so the ranking and all pairwise gaps are unchanged and the optimal threshold relocates by the same constant; the zero threshold is therefore a property of the stated reference population rather than of the index structure. Perturbing the reference dispersions is not neutral: rescaling each σ in turn by factors between 0.5 and 1.5 preserves the full ranking Tunisia above Egypt above Saudi Arabia in every case, and a separating threshold continues to exist in every case, but the anchor at zero is crossed by one country in four of the ten perturbations. Classification at a fixed zero threshold accordingly presumes the reference parameters as stated; separability of the outcome classes does not.

D Replication materials

Data and code for replication are available in the FICSS umbrella repository at `https://github.com/FICSS-Institute/computational-macrohistory/tree/main/CMH-Arab-Spring`. The folder contains the raw data in CSV format, the SSI calculation results, a variable codebook with sources and definitions, and the figures of this paper. Replication is deterministic: the procedure of Appendix B applied to the data of Appendix A reproduces every reported value. Figures 1 to 5 of this version correspond to the files `fig1_causal_diagram.png`, `fig3_time_series.png`, `fig4_ssi_analysis.png`, `fig5_validation.png`, and `fig6_counterfactual.png`; the files `fig2_probability_model.png` and `fig7_causal_flowchart.png` belong to earlier versions and are retained for continuity.

E Version history

Version 1.0 (March 2026) is the original deposit under DOI 10.5281/zenodo.18848734. Version 2.0 (May 2026 - never published) was an errata release: numerical inconsistencies between the interpretive text of the robustness section and the values of the robustness and appendix tables were resolved by recomputing those tables exactly from the Appendix A data panel and the stated reference parameters. The recomputation surfaced one previously unreported boundary case (the E_4 +30% perturbation), reconciled the paper's two leave-one-out tables, and corrected the arithmetic of the temporal-stability comparison. No substantive finding changed.

Version 3.0 (July 2026) is a revised and shortened edition. The main text was brought within the FICSS working paper length standard, from roughly 20,000 to roughly 12,500 words, with no changes to the data panel, the reference parameters, or any computed value; all values were re-verified by independent recomputation for this release. The author name was updated to Stefano Angeli. Cross-references were reconciled with Document I version 3.0 (Angeli, 2026a), whose theorems now formalise conditions this paper invokes, and with Document II version 2.2 (Angeli, 2026b), including the nine-component theoretical index and the ladder of operationalisation levels. Two analytical presentations changed. The group comparison of Section 6.3 is reported without test statistics, which are not interpretable on this design; the underlying means and differences are unchanged. The alternative-scaling table of versions 1.0 and 2.0, whose values could not be reproduced from the deposited data panel, was replaced by a recomputed analysis under explicitly stated definitions (Section 6.7), and its conclusion changes from scaling-insensitivity to reference-population dependence, a revision that reinforces the paper's central finding on the anocracy component. The duplicate leave-one-out table of the former Appendix C.2 was removed, the former one-line sensitivity appendix was merged into Appendix C, and the appendices were relettered. Rankings, threshold classification, counterfactuals, and all hypothesis verdicts are unchanged.

Version 3.1 (July 2026, this document) is a technical correction and compliance release; no data value, reference parameter, or weight or computed result under the baseline specification

changes, and all rankings, classifications, counterfactuals, and hypothesis verdicts are unchanged. The corrections are six. First, the stated definition of the youth bulge D_2 is corrected from ages 15 to 29 to ages 15 to 24: verification of the deposited series against the source age bands (World Bank Open Data, UN population estimates) matched the deposited values to the reported two decimals for ages 15 to 24 (Tunisia 2010: 18.64; Egypt 2010: 19.51; Saudi Arabia 2010: consistent within sex-ratio uncertainty), establishing that the series was always computed over the canonical band and that the earlier band description was a documentation error. Second, the revolutionary-dynamics subsystem of Section 2.5 gains the pressure-dissipation term that its steady-state argument requires, bringing it into the mean-reverting form of Document II; the SSI and every downstream number are unaffected. Third, the chance-classification probabilities are derived and matched to the claims they support, 1/3 for the existence of a separating threshold and 1/6 for the exact ordering, replacing the previous 12.5% to 16.7% interval. Fourth, the interaction language is tightened throughout: the linear index encodes the signed regime main effect implied by the interaction model of Section 2.4, not the product term itself, and the additive-versus-multiplicative question is explicitly reserved for the large-N programme. Fifth, the reference-parameter table gains explicit component-wise provenance, the alternative-scaling analysis states its quartile convention, the alternative weighting schemes are renormalised to sum to one by trimming the connectivity weight (Appendix C), which shifts their 2010 values by $0.01 \times z(S_3)$ relative to earlier releases without affecting rankings or classifications, Table 7 reports decade means for all five components, and Table 12 clarifies the marginal row. Sixth, hypothesis H3 is relabelled to reflect that the design tests insufficiency only, the counterfactual H4 is framed as within-model comparative statics, a citation mismatch in Section 1.2 is corrected (Beinin, 2012 now supports the Egyptian labour history of Section 4.3; the Tunisian diffusion claim cites Angrist, 2013 and Hibou, 2011), Table 3 clarifies that participation figures are observed estimates against a fixed coding threshold, the Gini interpolation note identifies the three filled country-years (Saudi Arabia 2008 to 2010), and the prose is swept for residual banned-form constructions.

Document III of the Computational Macrohistory Series

Foundations Institute of Computational Social Science
`https://ficss.institute`

Contact: stefano.angeli@ficss.institute