Oxford Open Climate Change, 2026, pp. 1–14
10.1093/xxxx/xxxxxx
Published: Date added during production
Paper

An Integrated World Model for Strategic Planning and Scenario Analysis

Nazar S. Sotiriadi,[1,*] Alexander N. Krenke[1] and Andrei A. Osiptsov[2]

[1]Institute of Geography, Russian Academy of Sciences, Staromonetny lane, 29, bld. 4, 119017, Moscow, Russian Federation
[2]Skolkovo Institute of Science and Technology (Skoltech), Bolshoy Boulevard 30, bld. 1, 121205, Moscow, Russian Federation

*Corresponding author. sotiriadi.n.s@gmail.com

Abstract

Motivation. Coupling the economy, climate, society, and geopolitics in one closed simulator is easy to attempt and hard to sustain: such loops generically diverge, stick in absorbing states, or must be damped into triviality. We present a global integrated model (GIM) — 57 country agents, seven coupled domains, an annual deterministic loop — built as an explicit encoding of invariants (conservation laws, accounting identities, slow empirical regularities) with calibrated restoring forces. Methods. Stability is measured, not asserted: numerical linearization of the annual map, twin runs from perturbed initial states, and a 500-member history-matched parametric ensemble; anchoring uses a 2015–2023 retrospective fit, a constraint hold-out frozen at 2019, and an input-vintage-robust conflict ranking. Results. The detrended spectral radius is 1.075–1.094, its above-unity modes localized in slow socio-fiscal directions (trust, tension, debt), while economic and climate perturbations contract (0.97 and 0.94 per year); the parametric, shock-inclusive fan grows near-linearly to $\pm 11.5\,\%$ of median world product at thirty years. In-sample errors are 0.60 trillion USD (country-level output), 0.94 Gt CO_2, $0.145\,^{\circ}C$; scored beyond the 2019 freeze they are 0.77, 0.71, 0.13. Conflict ranking reaches AUC 0.739 (1990–2023), and 0.84 with pre-2000 inputs on 2000–2023. A game layer adversarially samples the tails: $1{,}002$ scripted runs (23.5 % survival) and 200 completed human runs of 401 started (28.5 % of completers) map how a managed world fails, and exposed two structural degeneracies invisible to retrospective validation. Conclusion. GIM is a bounded, anchored, responsive exploration instrument — not a forecast engine.

Keywords integrated assessment model, world simulation, stability, invariants, uncertainty quantification, serious games, weak signals

Introduction: the hard part is staying together

Strategic planning and systemic-risk work need a tool in which climatic, economic, and socio-political processes are considered jointly: the paths that matter — warming to crop yields to prices to social tension to instability — cross domain boundaries that sectoral models do not [Nordhaus, 2017, Riahi et al., 2017]. That gap is well known. What is less discussed is why it persists. Coupling the domains is not conceptually difficult; the difficulty is that a *closed* loop through economy, climate, society, and politics is a high-dimensional nonlinear dynamical system, and such systems do not stay plausible by default [May, 1972, Sterman, 2000]. In our experience they fail in a small number of characteristic ways: variables run to a bound and stick (absorbing states), accounting errors compound into runaway growth, latent variables drift to structural attractors that no policy can move, or — the opposite failure — every feedback is damped until the simulator reproduces a trend line and nothing else. A world model earns its keep in the space between these failures: bounded for decades, yet still able to generate crises, cascades, and lags.

This paper reports a model built and validated to live in that space, and reframes what we consider its main result. The Global Integrated Model (GIM) advances 57 country agents (the 50 largest economies plus seven regional aggregates that conserve world totals) through an annual deterministic loop spanning the economy and prices, climate and the carbon cycle, natural resources, society and politics, interstate relations, and discrete crises with stock-flow-consistent finance. Its headline property is not the breadth of coverage but the combination of three measured behaviours: (i) *convergence* — thirty-year trajectories are bounded, insensitive to small initial-state perturbations, and their ensemble uncertainty grows near-linearly rather than exponentially (Section 4); (ii) *anchor* — the same configuration reproduces 2015–2023 history in-sample, survives a temporal hold-out, and carries out-of-sample skill where structure matters (Section 5); (iii) *responsiveness* — the loop produces structured dynamics (lagged social channels, threshold-stepped sovereign cascades, clustered late failures) whose shapes go beyond the coded terms they build on (Section 6).

The stance behind (i) is methodological and we state it up front: the deterministic core is an explicit hypothesis about *invariants* — the conservation laws, accounting identities, and slowly-varying empirical regularities that we take to govern nature and society (Section 2). Determinism is then a feature, not a naivety: the core computes what follows *if* those invariants hold, and all uncertainty

Received: July 28, 2026 Month 2026. **Revised received:** XXX X, 2026 Month 2026. **Accepted:** XXX X, 2026 Month 2026

is carried by ensembles over the invariants' strengths. On top of this core sits an instrument layer — history-matched ensembles, weak-signal detectors, and a decision game that operates as an adversarial sampler of the tails, including live outcome telemetry from external players (Section 7). The model does not claim predictive foresight; at horizons beyond two–three years, professional growth forecasts lose to a naive mean-growth benchmark in roughly half of countries [Celasun et al., 2021], and we hold ourselves to the same honest bar (Section 5).

Relative to the integrated-assessment and world-modeling traditions — DICE/RICE [Nordhaus, 2017, Nordhaus and Yang, 1996] and FUND [Anthoff and Tol, 2013] (welfare optimisation), GCAM [Calvin et al., 2019] and MESSAGEix [Huppmann et al., 2019] (energy-system equilibrium), E3ME [Mercure et al., 2018] (econometric simulation), T21/iSDG [Millennium Institute, 2002] (single-country system dynamics), World3 [Meadows et al., 1972] (global aggregates) — GIM's distinction is that it closes the loop through society, geopolitics, and discrete crises at country resolution and then *demonstrates* that the closed loop is stable, anchored, and non-trivial. A controlled benchmark quantifying what the closed loop sees and sectoral models structurally cannot is kept in Appendix B; the model specification (functional forms, parameters) is compressed into Appendix A.

A comprehensive review by Keppo et al. [Keppo et al., 2021] identifies six persistent critiques of IAMs: heterogeneous actors, technology diffusion, capital markets, energy–economy feedbacks, policy scenarios, and model interpretation. GIM directly addresses several of these: it embeds heterogeneous country agents with bilateral relations, includes an endogenous R&D-stock mechanism and trade-driven diffusion, implements stock-flow-consistent finance with sovereign debt and crisis dynamics, and couples energy prices, inflation, and social tension in a single loop. The remaining critiques (policy differentiation and result interpretation) are addressed in Sections 6 and 7.

The invariant view

Every long-horizon simulator embodies assumptions; ours are organized deliberately into three classes of invariants, in descending order of confidence, plus a fourth element that keeps the loop near them. We use the term "invariant" in a pragmatic, model-building sense: a relationship that we treat as fixed in the simulation because it is either a conservation law, an accounting identity, or an empirical regularity stable enough across the historical record to serve as an anchor. This is not a philosophical claim about the nature of society; it is a modelling choice, made explicit so that readers can audit its consequences.

Physical conservation. Carbon is conserved: each year's emissions are partitioned into the four IPCC AR6 impulse-response pools and decay on fixed timescales; atmospheric stock, radiative forcing, and the two-reservoir surface–ocean heat balance follow standard reduced-complexity climate physics [Joos et al., 2013, Geoffroy et al., 2013, IPCC, 2021], the same architecture as the FaIR and MAGICC emulators [Smith et al., 2018, Meinshausen et al., 2011], with the emission record anchored to the Global Carbon Budget [Friedlingstein et al., 2023]. The block reproduces both assessed constants that an AR6-consistent emulator must match: equilibrium sensitivity 3.0 °C and transient response 1.79 °C (AR6 best estimate 1.8).

Accounting identities. Money and goods cannot appear from nowhere: the financial side is stock-flow consistent (every loan creates a matching deposit; the money stock is identically deposits and loans) [Godley and Lavoie, 2007]; government budgets close; bilateral trade settles in reserves and world net exports sum to zero by an enforced invariant; excessive borrowing raises the cost of credit — a financial accelerator [Bernanke et al., 1999]. These identities are checked at runtime, not assumed.

Empirical regularities as slow variables. Between hard conservation and free parameters sits a class of relationships that are not laws but have been stable across decades and countries, and we treat them as invariants with priors on their strength: trade follows a gravity law in distance [Head and Mayer, 2014]; unemployment follows Okun's law and inflation a flat post-1990 Phillips curve with a weak quantity-theory term; poorer economies converge toward the frontier at a β-convergence rate validated against the cross-section; the public's response to unemployment and inflation follows the two-to-one misery-index ratio; war sizes follow a power law [Richardson, 1948, Clauset et al., 2009]; institutions are anchored to governance indices, culture to the Hofstede dimensions [Kaufmann et al., 2011, Hofstede et al., 2010]. Monetary policy follows country-level Taylor rules [Taylor, 1993]. Where a regularity failed the data — a classic Richardson arms-race term, rejected at $p = 0.74$ on the 1995–2024 military-expenditure panel — it was left out and the observed state variable used instead. Each of these regularities carries a prior whose range reflects both sampling uncertainty and cross-country heterogeneity; the ensemble analysis of Section 4 explores the consequences of this uncertainty. We do not claim that these relationships are universal laws, only that they are more defensible anchors than arbitrary functional forms.

Equilibrium anchors and restoring forces. Invariants alone do not prevent drift; bounded behaviour comes from explicit restoring terms whose strengths are calibrated, not hidden: resource prices mean-revert in log space toward reference levels; production is normalized so each country's base point is exactly reproduced (a normalized-CES discipline that pins levels while leaving substitution free); trust and tension carry anchor terms; crisis states have exit conditions. Section 4 shows what these forces achieve and Section 7 how two missing anchors were found the hard way. The full functional forms and the 33 literature-grounded priors are in Appendix A.

Determinism, in this view, is the point: the core is a reproducible mapping from an invariant set to trajectories. Following Oreskes et al. [Oreskes et al., 1994], we do not claim the model is verified truth — we claim it is a confirmed, auditable encoding of stated assumptions, evaluated by whether its trajectories stay within the envelope of observed history and whether its failures are informative.

Architecture: one closed loop

One model step is one year; blocks update in fixed order and the year's outputs are the next year's inputs (Fig. 1). The economy is the hub: output comes from capital, labour, and energy through a nested CES function; markets clear by price; inflation, unemployment, budgets, debt, and bilateral trade are tracked per country. Output emits CO_2 (declining structural intensity); emissions accumulate in the carbon pools; warming returns as a quadratic output damage.

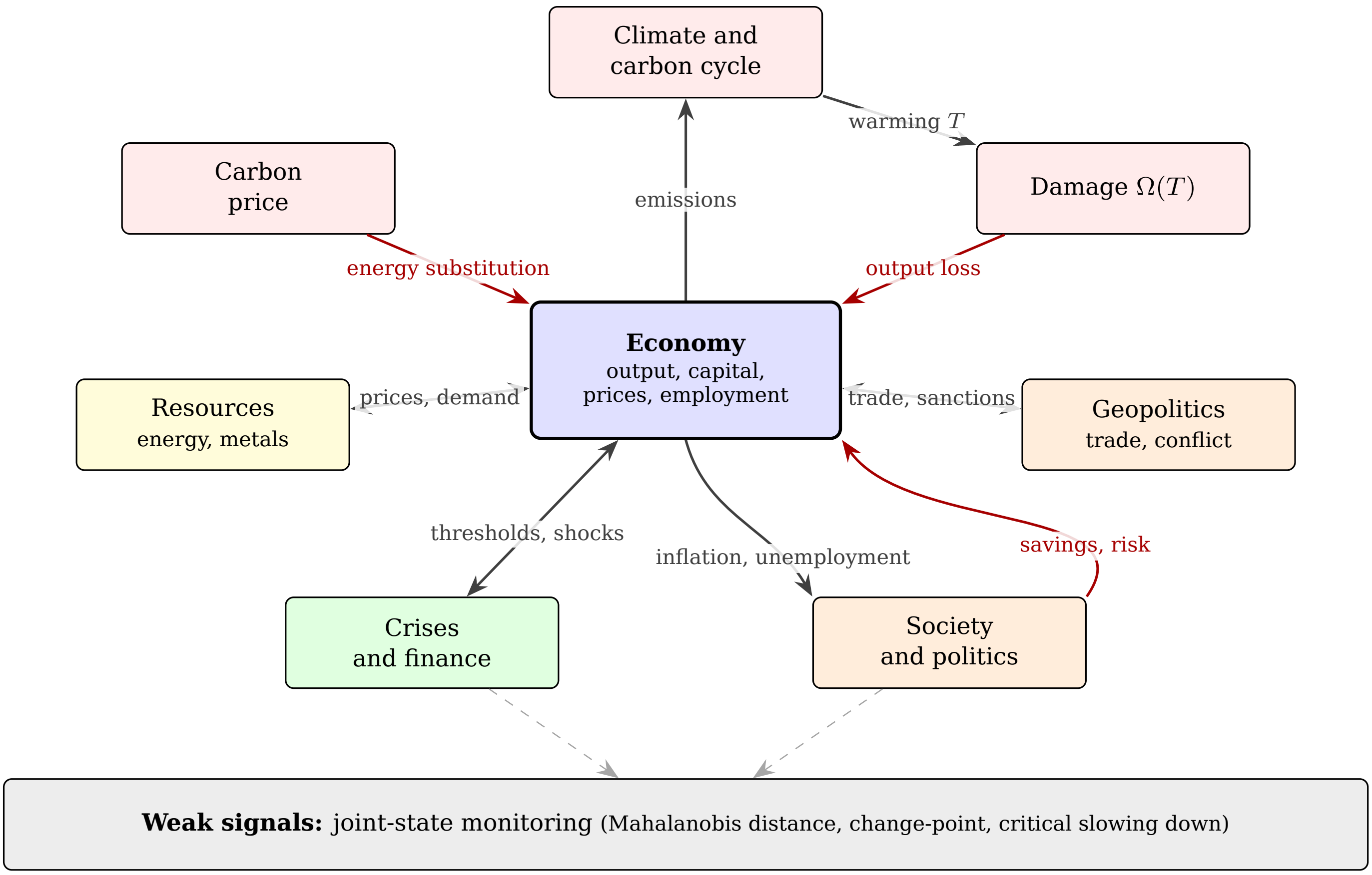

Figure 1 The closed loop. Economy as central hub; black arrows direct effects, red arrows feedbacks reducing output or savings. The climate loop (emissions → warming → damage → output) is complemented by carbon-price, resource, socio-political, geopolitical, and discrete-crisis channels. The bottom layer monitors the joint state for weak signals.

Resource stocks and world prices couple scarcity into inflation. Society tracks trust, tension, inequality, legitimacy, and protest pressure, all responding to economic stress with explicit coefficients; politics converts legitimacy into a policy-space multiplier; geopolitics carries relations, alliances, sanctions, and a CINC-style capability index [Singer et al., 1972] grounded in SIPRI expenditure. The 2023 base-year world product of the 57 agents is ≈107 trillion USD, so the thirty-year baseline median of 241 implies ≈2.7 % median annual growth. Discrete crises — debt, currency, regime collapse — fire on threshold crossings, apply audited state changes (a 40 % debt write-down, output and trust hits), and have exit rules. Conflict, unrest, and climate risk spill over to geographic neighbours; a shock in any block propagates along realistic pathways: carbon price to energy cost to emissions; hardship to tension to savings to crisis risk; crisis to trade partners. The feedback signs matter more than the functional detail (which is deferred to Appendix A): the loop mixes amplifying channels (financial accelerator, crisis contagion, tension–trust spiral) with the restoring forces of Section 2, and the balance between them is what the next section measures.

What keeps it together: stability, measured

Convergence claims are cheap; we support ours with three quantitative diagnostics plus the engineering discipline that maintains them. All runs below use the deterministic core (simple background policy, extreme events off); the harness is in Sotiriadi and Osiptsov [2026].

Spectral radius of the one-year map. We linearize the full annual map $x_{t+1} = F(x_t)$ numerically (central differences) around the baseline trajectory in a 408-dimensional state space — seven variables per agent (output, capital, debt, unemployment, inflation, trust, tension) plus the nine global coordinates (surface and ocean temperature, four carbon pools, three resource prices) — at three points along the run. Eigenvalues are computed on $D^{-1}JD$ with D the diagonal of trajectory-mean coordinate scales — a similarity transform that leaves the spectrum unchanged and only conditions the eigenvector readout. The raw spectral radius is $\rho(J) = 1.075$, 1.076, 1.078 at 2026, 2036, 2046 (Fig. 2a). To separate level growth from genuine amplification we also compute a *detrended* map: coordinates that compound as levels (output, capital, debt, carbon pools, prices) are divided by their own baseline one-year growth factor, so a perturbation growing exactly with the trajectory sits at $|\lambda| = 1$. The detrended radius is $\rho_{\det} = 1.075$, 1.075, 1.094 — almost unchanged — which rules out the comfortable reading that the above-unity modes are mere growth trend. They are not economic at all: summing squared eigenvector weights over the ∼50–60 above-unity modes, trust in government carries 0.53–0.63 of the loading, social tension 0.32–0.39, and public debt most of the remainder.

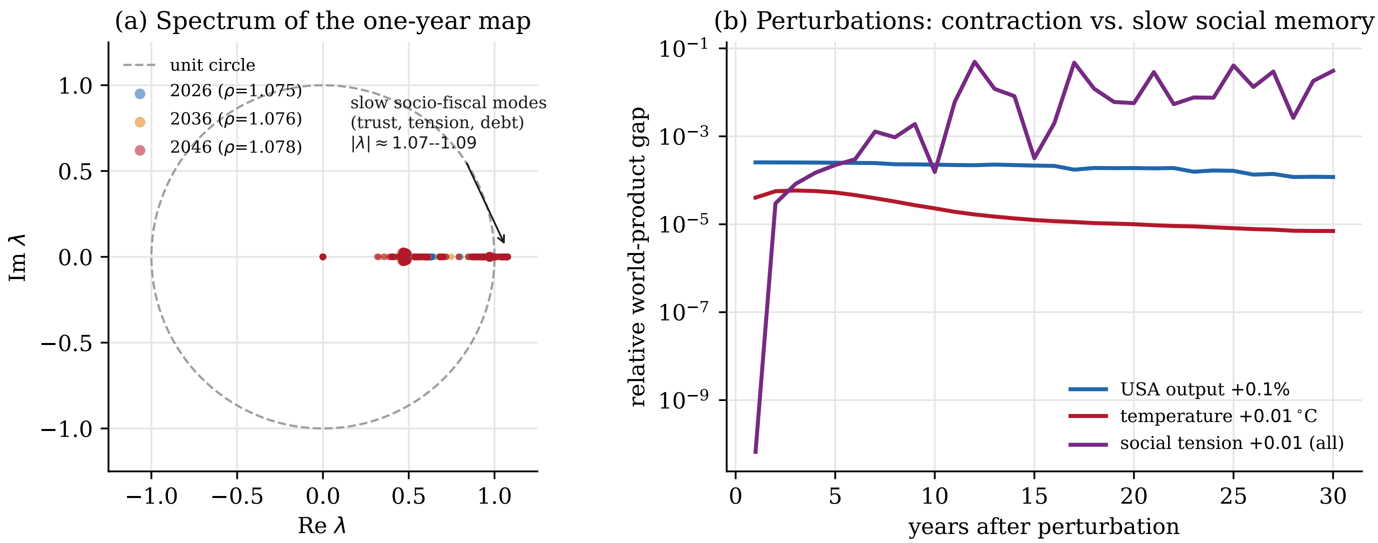

Figure 2 Local stability of the annual map. *(a)* Eigenvalues of the (similarity-) normalized Jacobian (408-dimensional state) at three linearization years: the spectrum lies on and inside the unit circle except a cluster at $|\lambda| \approx 1.07$–1.09 whose eigenvectors load on trust, tension, and debt — the slow socio-fiscal modes; detrending the level-growth coordinates leaves the radius essentially unchanged ($\rho_{\mathrm{det}} = 1.075$–$1.094$). *(b)* Twin-run response (relative world-product gap): output and temperature perturbations contract at 0.97 and 0.94 per year, while a broad social-tension perturbation persists and amplifies — the same long-memory direction.

The super-unit spectrum is the model's *slow socio-fiscal subspace* — the drift of trust (whose equilibrium anchor ships off by default, a conservative choice discussed in Section 7), the accumulation of tension, and debt compounding near the effective interest rate — precisely the directions in which Section 6 finds multi-year lags and the twin runs below find persistence. The economic–climate subspace lies inside the unit circle.

The above-unity spectral radius does not imply global instability; it identifies the local linearised growth rate of perturbations along the socio-fiscal subspace. In the full nonlinear system these directions are bounded by the equilibrium anchors described in Section 2 — the trust anchor (off by default in the headline configuration) and the debt exit rules. The ensemble fan of Fig. 3 demonstrates boundedness over 30 years. The spectral radius is thus the local amplification rate of the *unbounded* components of the linearised dynamics; their nonlinear bounds are what keep the system finite.

Perturbation contraction. The trajectory-level counterpart: twin runs from ε-perturbed initial states, gaps measured as dimensionless relative differences in world product. A $+0.1\,\%$ shock to United States output (a vector with essentially no projection on the socio-fiscal modes above) produces a relative gap of 2.6×10^{-4} after one year that *contracts* to 1.2×10^{-4} by year 30 — an empirical factor of 0.974 per year; a $+0.01\,^{\circ}\mathrm{C}$ temperature perturbation decays at 0.942 per year. Nearby trajectories converge in the economic–climate subspace — the opposite of chaotic divergence (Fig. 2b). The one persistent direction is the one the spectrum names: raising every country's tension by 0.01 leaves output almost untouched for a decade and then amplifies at roughly $1.3\times$ per year between years 10 and 30 as crisis triggers engage, reaching a $3\,\%$ world-product gap by 2053. The spectrum, the twin runs, and the fan below therefore tell one story rather than three: contraction in the physical–economic directions, slow amplification confined to the socio-fiscal subspace. We report the asymmetry as a substantive property, not a defect — and it quantifies the limitation that socio-political initial-state error does not wash out (Section 8).

Bounded, near-linear uncertainty growth. A 500-member ensemble drawing the 33 key priors from their history-matched ranges — with the discrete-crisis machinery *active* (the members average one to three sovereign crises by 2053) — runs the full 30 years without a member diverging or saturating (Fig. 3). This is a *parametric, conditional* fan, not a forecast interval: members share the initial state, so much of its near-linear growth is each member following its own prior-drawn trend, and non-divergence is partly by construction since implausible corners were pruned by history matching. What the fan does bound is the amplification rate: the world-product half-width grows from $\pm 1.7\,\%$ of median at year 5 to $\pm 4.0\,\%$ at 10 and $\pm 11.5\,\%$ at 30 (all 5–95th half-widths), an average exponential rate below $0.08\,\mathrm{yr}^{-1}$ with concave shape after year 10 — any local instability compounding materially faster than the socio-fiscal modes of the spectrum would bend this curve upward, and none does. Two non-trending observables make the point without the trend caveat: the mean-tension fan stays narrow and sub-linear (5–95th width 0.011 at year 5, 0.026 at 30, i.e. 2–5 % of its level), and the temperature fan widens only from 0.72 to $1.10\,^{\circ}\mathrm{C}$ between years 5 and 30 (full widths; the climate-sensitivity prior expresses itself early). Median 2053 world product is 241 trillion USD (5–95th: 215–270) at $+1.94\,^{\circ}\mathrm{C}$ (1.39–2.49); for calibration of scale, a single structural choice — renormalising production to constant returns — moves 2100 world product by 7–9 %, of the same order as this whole parametric fan.

The discipline that preserves this. Stability is a maintained property, not a lucky draw. Every change to the core faces pinned golden backtests (bit-identical or explicitly recalibrated); calibration data are checksum-bound; every optional mechanism is a switch, off by default unless anchored by evidence; and the invariants of Section 2 are runtime-checked. Two failures that this discipline did *not* catch — because they lived outside the historical window — were

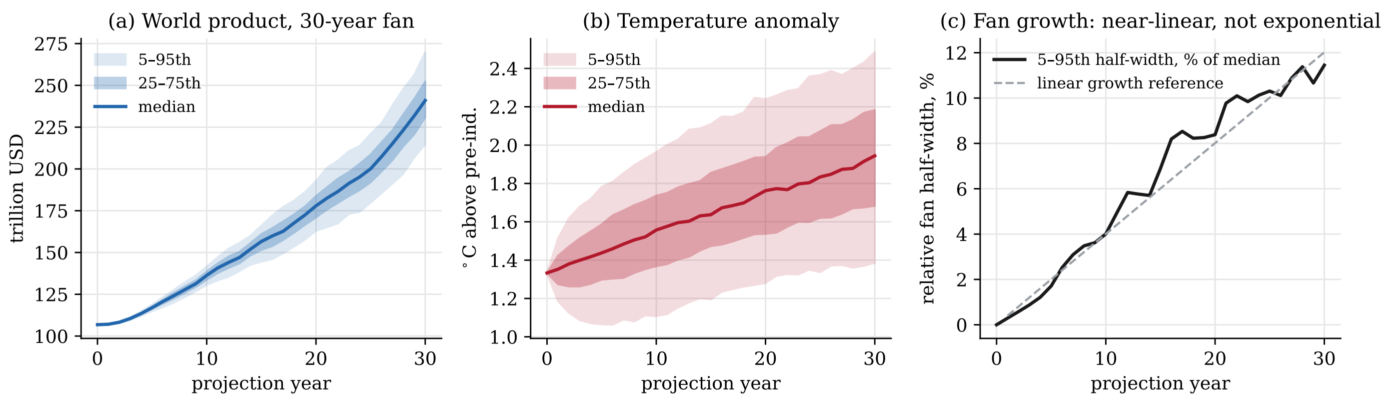

Figure 3 Thirty-year boundedness evidence (500-member parametric ensemble, shared initial state, discrete crises active). *(a)* World product fan: all members bounded, median 241T at 2053. *(b)* Temperature fan: full 5–95th width grows $0.72 \to 1.10\,^\circ$C between years 5 and 30. *(c)* Relative 5–95th half-width of world product tracks a linear reference — parametric uncertainty accumulates; no super-exponential bending appears.

caught by the game layer (Section 7); both fixes shipped as anchors of exactly the Section 2 type, with the historical backtests unchanged to the bit. An honest caveat accompanies all of this: convergence here is engineered and measured, not proven — anchors strong enough to stabilize can in principle mask real-world instabilities, which is why their strengths carry priors and the tails are explored deliberately (Section 7).

Is it anchored? Validation against history

A stable simulator can still be stably wrong; the second requirement is anchor to observation. All retrospective runs are *free-running*: initialized once from the 2015 state and never re-anchored to observations, so errors are multi-year trajectory errors, not one-step-ahead errors. The output metric is the RMSE over country-year pairs (20 largest economies), reported in trillion USD; emissions and temperature are world-series RMSEs. The economic and climate blocks are calibrated by history matching — ruling out parameter regions whose implausibility $I_o(\theta) = \mathrm{RMSE}_o(\theta)/\mathrm{tol}_o$ exceeds the three-sigma cut, a standard approach for expensive simulators [Williamson et al., 2013] — on the 2015–2023 world series; the machinery is specified in Appendix A. Because those metrics are in-sample by construction, predictive adequacy is measured separately by a *constraint hold-out*: history matching is rerun on data through 2019 only and the retained region frozen before 2020–2023 is scored. We name it a constraint hold-out, not a forecast, because only the history-matching constraints move: the SSP2-4.5 forcing path, the SSP2 productivity anchor, and the three registry parameters whose sources name the 2015–2023 window are not re-fit — and since the 2015–2019 window rules out none of the sampled draws at the three-sigma cut, the frozen region coincides with the prior box and the headline configuration is unchanged. What the exercise measures is therefore how a 2019-constrained model degrades when run into 2020–2023, not an information-free forecast.

Results (Table 1; Fig. 4): in-sample RMSE 0.60 trillion USD (country-level output), 0.94 Gt CO_2, $0.145\,^\circ$C; on the 2020–2023 window 0.77, 0.71, and 0.13 (with 2020 excluded: 0.83, 0.10, 0.15). The per-year decomposition is more informative than the window averages: output errors are 0.56, 0.78, 0.80, 0.90 across 2020–2023 — a monotone free-running drift, not a pandemic spike (which is why excluding 2020 *raises* the output figure) — while emissions show the opposite anatomy, a single COVID-sized miss of 1.42 Gt in 2020 collapsing to 0.05–0.12 Gt in 2021–2023, exactly what a model without a pandemic submodel should produce when output is nearly right and lockdown-specific emission suppression is outside its physics; per-year temperature errors are 0.02–$0.17\,^\circ$C. Skill against naive benchmarks built on pre-2020 data divides honestly: temperature beats persistence (+0.05) and linear extrapolation (+0.16), emissions beat both decisively (+0.32, +0.54), output loses to both — short smooth aggregates remain the domain of a ruler, consistent with the full-window comparison and with the forecast-evaluation literature [Celasun et al., 2021]. The climate block additionally tracks the 1990–2023 record ($0.096\,^\circ$C) with best-fit sensitivity $3.0\,^\circ$C. At country level — where the crisis triggers live — the 2015–2023 per-country output RMSE has median 0.10 and interquartile range 0.05–0.13 trillion USD across the 20 anchored economies, with the largest five errors concentrated where levels are largest (China 2.4, United States 1.1, Japan 0.24, India 0.27, United Kingdom 0.17); crisis triggers compare debt and reserves *ratios* within a country, which damps but does not remove the propagation of level errors into trigger distances — a caveat carried by the scenario results.

The socio-political blocks are held to a softer but real bar. The conflict-proneness score — a fixed-weight composite of observed inputs with no coefficient estimated from conflict data — ranks the 57 agents against the UCDP/PRIO registry 1990–2023 [Sundberg and Melander, 2013] at AUC 0.739 (95 % CI $[0.60, 0.86]$, 20,000 bootstrap resamples [Efron and Tibshirani, 1993]; permutation $p \approx 0.001$); its Brier skill against the constant base-rate forecast is +0.123; observed conflict incidence falls monotonically across score quintiles: 6/11, 6/11, 5/11, 3/11, 1/13 countries (55 % at the top to 8 % at the bottom — counts this small support ranking, not calibration, claims). Because regime stability and military burden are themselves affected by conflict, the score is recomputed in a vintage-clean configuration (inputs from pre-2000 data only, evaluated against 2000–2023 conflicts): AUC 0.84 $[0.72, 0.94]$ — evaluated on the 2000–2023 window, hence not directly comparable to the 1990–2023 headline — and identical to the 2023-vintage score on the same subset and window, so ranking skill is not reverse-causation contamination. The social cost of carbon cross-checks against the published range: Ramsey central $\approx$\$90/tCO_2 at 200 years, spanning

Table 1 Validation summary. Economic and climate blocks: retrospective evaluation, in-sample and temporal hold-out (out-of-sample values after semicolons). Other blocks: cross-checks, benchmark improvement, or channel identification.

Block	Source	Criterion	Result
Economy	PWT, WDI	in-sample country-level RMSE 2015–2023; constraint hold-out 2020–2023	output \$0.60T; \$0.77 beyond the 2019 freeze
Climate and carbon	IPCC AR6, GCB	in-sample; hold-out; ECS/TCR anchors	CO_2 0.94 Gt; 0.71 out-of-sample. T $0.145\,^\circ$C; 0.13 out-of-sample
Cost of carbon	EPA-2023, RFF-SP	published cross-check	\$90 Ramsey central; \$90–280 across (ρ, η)
Conflict	UCDP/PRIO	ranking vs. registry; input-vintage test	AUC 0.739 $[0.60, 0.86]$; vintage-clean 0.84 $[0.72, 0.94]$; BSS $+0.123$
Social block	SWIID [Solt, 2020], WGI [Kaufmann et al., 2011]	*internal* channel identification on model output (Sec. 6)	lagged channel; effect sizes $r_s{=}0.59/{-0.52}$, $n{=}120$
Stability	—	detrended spectrum; twins; fan growth (Sec. 4)	$\rho_{\det}{=}1.08$–1.09, confined to socio-fiscal modes; economic/climate directions contract; fan $\pm 11.5\%$ at 30y

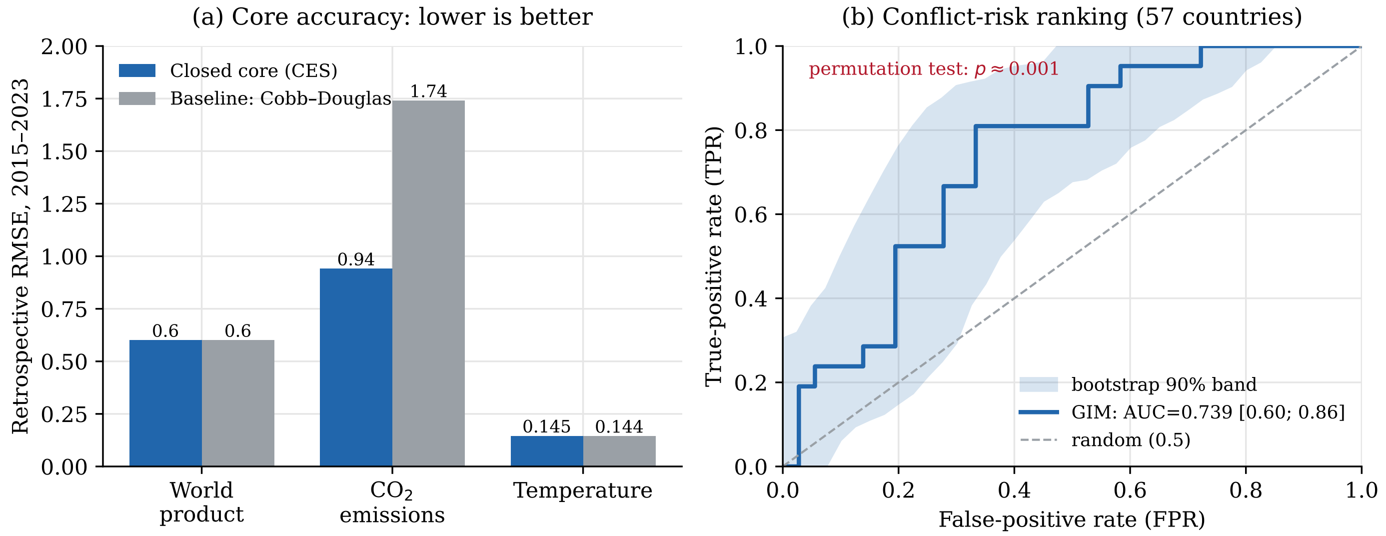

Figure 4 Retrospective anchor. (a) RMSE 2015–2023 of the closed nested-CES core against the former Cobb–Douglas core. (b) Conflict ranking (score inputs at 2023 base-year values; vintage-clean variant in text): AUC 0.739 against 0.5 for chance; the shaded band is the 90 % bootstrap envelope of the ROC curve, while the interval quoted in the text, $[0.60, 0.86]$, is the 95 % CI of the AUC itself.

\$90–280 across (ρ, η) choices, below the EPA-2023 central ≈\$190 [U.S. Environmental Protection Agency, 2023, Rennert et al., 2022]. Fed DICE's own damage function, the engine returns ≈\$15 against DICE's \$31 — confirming the higher native estimate is not produced by an inflated valuation layer; the residual factor of two reflects the 200-year truncation, the different carbon cycle, and the different baseline growth path, and we do not claim a full reconciliation.

The long-horizon baseline deserves its own paragraph rather than a clause. With all headline anchors active, the deterministic baseline reaches $2.69\,^\circ$C in 2100, with the atmospheric stock rising from ≈3,270 to ≈4,380 $GtCO_2$ (≈1,110 $GtCO_2$ of net airborne accumulation). Removing the SSP2 productivity anchor — emissions and decarbonisation are endogenous either way — lowers the no-intervention 2100 outcome to about $2\,^\circ$C, nearest the strong-mitigation SSP1-2.6 envelope: the structural decarbonisation rate fitted to 2015–2023 (1.6 %/yr) is optimistic when extrapolated for decades. We report this as a large, disclosed structural sensitivity — it is the main reason the recommended use mode is scenario *differences*, where anchor uncertainty largely cancels, and it bears on any absolute reading of the game's climate gate in Section 7.

Emergent Behaviour from Coupled Dynamics

A simulator can be stable and anchored by being inert. The third requirement is that the loop produce structured behaviour beyond the coded terms it builds on. We grade this honestly, regularity by regularity: in each case we name the coded ingredient and the consequence that is not coded.

A lagged, mediated economy-to-society channel. The coded ingredients are linear increments to an accumulating tension index; the multi-year timescale itself follows from accumulation. What is *not* coded is the mediation structure: global Morris screening over the 33 priors [Morris, 1991, Campolongo et al., 2007] shows tension insensitive to physico-economic priors at short horizons (elementary effects at the noise floor relative to each output's own scale — μ^* is normalized within each output), while in an ensemble that also varies the social priors, final tension is governed by them (Spearman

$r_s \approx +0.59$ for inequality sensitivity, $r_s \approx -0.52$ for the trust anchor; $n = 120$ members). We report effect sizes rather than the p-values: on model-generated samples p is a function of the run budget, not evidence about the world, and these correlations are stable as the ensemble grows. The prior-induced spread grows by two-thirds between horizons of 5 and 25 years; a three-year stagflation shock raises mean tension persistently, still building a decade later (positive in 78 % of prior draws), with contemporaneous transmission confirmed by distributed-lag regression (Fig. 5). The spectrum of Section 4 sees the same structure from the other side: its slow modes load on trust and tension.

Threshold-stepped sovereign cascades. The staircase shape is coded — thresholds are thresholds. What is not coded is *which* countries tip, in what order, and at what dose: an oil-supply shock propagates through import bills, reserves, and debt issuance until, at a realistic burden (≈3 % of GDP), two additional importers cross, and at a severe one (≈15 %), four more with tripled exposure — an ordering selected by each country's fiscal state, not by hand (Appendix B).

Convex escalation of food-driven instability. The social increments are linear in their inputs (Appendix A); the convex dose–response of protest pressure to yield loss (≈0.11 affordability stress, ≈0.016 protest pressure at a 50 % regional loss) arises from the chain supply gap → price → affordability → protest compounding across blocks — consistent with the nonlinearity reported for food prices and unrest by Lagi et al. [Lagi et al., 2011] (a working paper; we cite it as such).

Late clustering of failures. The cleanest case: no coded term makes the crisis rate time-varying, yet on the deterministic 30-year baseline active crises rise from 1–2 per year in the 2020s to around ten in the 2040s with a dispersion (Fano) index of 2.0 — slow accumulators (debt, temperature, tension) mature together. The game layer shows the same clock from the player's side: 77–90 % of losing runs fail after 2040 (Section 7).

Spatially structured risk. The diffusion itself is coded (neighbour spillover weight 0.05); the non-coded part is that switching it on moves the 2015–2023 retrospective fit by under a quarter of a percent — spatially realistic propagation at no cost in anchor. We do not claim here that the emergent spatial autocorrelation matches the observed record; that comparison lives in the repository's geographic validation materials and is not reproduced as a statistic in this paper.

The instrument: ensembles, weak signals, and a game that hunts the tails

The model is operated as an instrument with three modes, in increasing order of adversarialness.

Ensembles and scenario differences. The core is deterministic, so uncertainty is explicit: 500-member ensembles over the history-matched priors yield outcome distributions, and scenarios are compared as distributions against a common baseline — differences are robust to the anchor uncertainties that dominate absolute levels. The whole workflow runs on a laptop (Apple M4 Pro: a world-year in 30–50 ms, the 500×30-year ensemble in under a minute, the full 527-test suite in two), which is what makes ensemble-first practice a default rather than a special exercise. A worked case: a \$50/tCO_2 carbon tax cuts emissions by ≈5.5 % over ten years while raising tension and lowering trust through the energy-price–inflation channel — the political-economy cost invisible to single-domain models, monotone in the rate, and directly relevant to observed reversals (France 2018, Australia 2014); the full dose–response harness is Appendix B.

Weak signals. Before any threshold fires, the joint state is monitored for early structure: Mahalanobis distance of the stationary increments from their recent baseline [Mahalanobis, 1936, Wilson and Hilferty, 1931], information-criterion change-point detection [Schwarz, 1978], and critical-slowing-down indicators [Scheffer et al., 2009]. On synthetic and model trajectories an injected 15 % output drop is flagged at the correct step while calm runs stay near the false-alarm rate; the working device is comparing a scenario's weak-signal profile against the baseline's.

The game as an adversarial tail sampler. The same 57-actor core, wrapped into a single-player decision game (GODMODE): thirty annual turns from 2023, a deck of 158 dilemma cards compiled into real policy actions, seven catastrophe gates (temperature, biosphere, stability, economic and geopolitical strain, energy and food prices); losing any gate ends the run, reaching 2050 wins. A full game simulates in about a second, so outcome distributions are statistically meaningful; and because players and scripted archetypes actively push the world into corners smooth history never visits, the game functions as an importance sampler of the tails. Three uses justify reporting it in a modeling paper.

A tail census. In the 1,002-game reference batch of scripted archetypes, survival is 23.5 % with losses spread across five causes (35 % climate, 22 % economic, 21 % geopolitical, 13 % stability, 8 % energy price) and 77 % of failures landing in 2040–2045 (Fig. 6). Production telemetry from the public deployment corroborates the shape with humans in the loop: 401 runs were started by external players and 200 completed (50 % attrition). Survival is 57/200 = 28.5 % of completers (95 % CI ±6.3 points, overlapping the best scripted archetype's 28.4 %) but only 57/401 = 14.2 % of starts; since abandonment plausibly correlates with a run going badly, the completer rate is an upper bound. Human losses show a broader spectrum (31 % climate, 29 % economic, 15 % geopolitical, plus biosphere, food, and energy tails the bots rarely hit), 90 % fall after 2040, and scores are bimodal (loss cluster near 26, win cluster 70–81, winners holding warming to 1.6–1.9 °C). The distribution of *ways the world fails under active management* is a model output no backtest can provide. Telemetry is anonymous by design and reported only in aggregate (see Declarations).

The two rates bracket the true human survival rate: the 14.2 % intention-to-start lower bound is likely too pessimistic, since some abandonments reflect player fatigue rather than imminent loss, while the 28.5 % completer rate is an upper bound. The comparison to scripted archetypes is therefore indicative rather than quantitative; what matters is the broader loss spectrum and the late timing of failures, not the point estimate.

Emergent strategy structure. No doctrinaire policy dominates: uniformly random play survives most often (28.4 %), beating consistent decarbonisation (20.1 %) and consistent expansion (21.9 %), and each doctrine dies of its characteristic disease (green of economic strain, 43 % of its losses; growth of climate and stability, 39 % and 34 %); passivity is worst (18.7 %, two-thirds of losses to economic strain — in a stock-flow-consistent world, drifting means compounding debt). Restraint on both strain axes at once wins ≈75 % against ≈5 % for recklessness on both. Single-axis

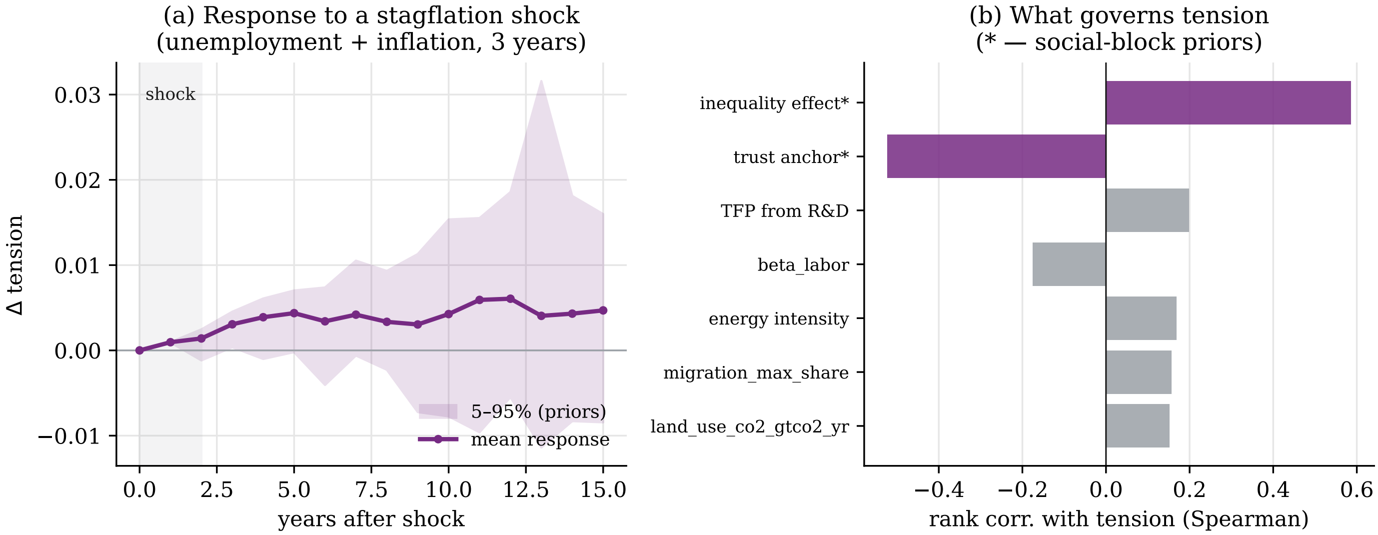

Figure 5 The emergent social channel. *(a)* Impulse response of mean tension to a three-year stagflation shock: slow build, persistence beyond a decade (5–95th band over priors). *(b)* What governs final tension: social-block priors (marked $*$) dominate; physico-economic priors are weak — the channel is real but lagged and mediated.

optimisation shifting the binding constraint to another subsystem is the game-mechanical face of the coupling this paper is about.

A diagnostic that found real defects. Balancing the game exposed two structural degeneracies: resource prices pinned at their clamps (a missing mean-reversion term plus two compounding accounting defects), and government trust declining on hawk and dove paths alike — an attractor from an inequality drag ≈40× the only restoring term, the very socio-fiscal subspace the spectrum of Section 4 flags. Neither was visible in any 2015–2023 backtest; both fixes shipped as equilibrium anchors with the goldens bit-identical (the trust anchor deliberately ships off in the headline configuration, so the paper's stability numbers describe the conservative, weaker-anchored model). Sustained adversarial play explores exactly the regions that validation against smooth history cannot; we regard the game layer as part of the model's quality programme. (Its epistemic status is bounded honestly: the game provides no goodness-of-fit against history, human telemetry reflects self-selected players with completion attrition, and difficulty gates are tuned against the bot batches through threshold-flip analysis.)

Beyond the game, country policies can also be set by language-model agents (intentions compiled into the same typed actions; every consequence computed by the deterministic core; off in all validated configurations, with full prompt/response audit logs) — the strategy space widens without the physics ever leaving the engine. The whole environment — CLI, ensembles, games, a desktop analytical shell — ships as an open, installable codebase with pinned goldens and dated, manifest-bound runs [Sotiriadi and Osiptsov, 2026]; results here correspond to release v18.1.3, archived on Zenodo as v18.1.4 — the same engine with additional supplementary materials.

Limitations

Convergence is engineered, not proven. The anchors that stabilize the loop could mask genuine instabilities of the real system; their strengths carry priors and the tails are probed adversarially, but a formal stability proof for the full nonlinear map does not exist — the spectral and perturbation evidence of Section 4 is local and numerical.

Expectations are adaptive. Agents look at the recent past; a tested model-consistent expectations mechanism is off by default (its marginal effect proved unstable). Anticipatory responses to announced policy are not in the headline runs.

Long-run growth is weakly identified, and forward projections anchor productivity drift to an external scenario; absolute long-horizon levels inherit this, which is why scenario differences are the recommended mode.

The climate block is a fast emulator, not an Earth-system model; the strongest carbon feedbacks are deeply uncertain, switchable, and relegated to tail scenarios. Damage estimates are uncertain across the discipline and carried as an interval (the level function is anchored at the meta-analytic central estimate; a Burke-type growth-damage channel exists as a switch).

The socio-political block mixes anchored and expert-set pieces. Substantial parts are tied to literature (war-size tail, migration gravity, misery-index ratio, governance anchors); coupling thresholds and cross-border tension diffusion remain expert priors, tagged as such in the auditable registry. Conflict is intrinsically hard to predict: the block supports relative risk ranking (AUC ≈0.74; vintage-clean 0.84), not event prediction. The persistent social mode found in Section 4 means socio-political initial-state errors do not wash out — an inherited uncertainty, honestly larger than the economic one. Statistics computed on ensemble output (Section 6) identify internal channels, not external facts, and are reported as effect sizes for that reason.

Aggregation. Resources are aggregate (no sectoral energy technologies, no endogenous renewable learning curves); no nominal exchange rates (a real numeraire with a fundamentals-triggered currency-crisis mechanism); no portfolio capital account; no dedicated pandemic submodel (epidemic shocks are extreme events);

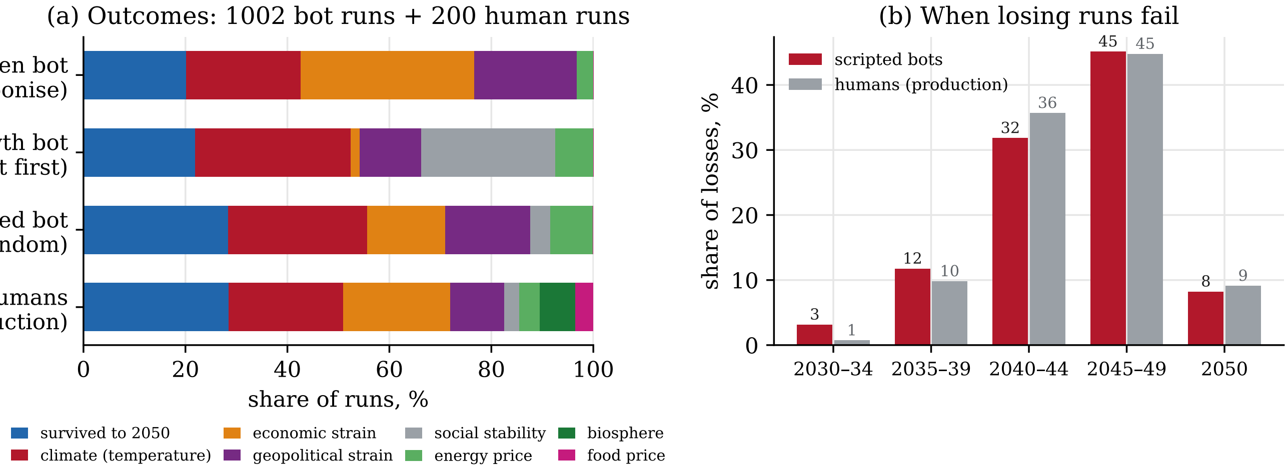

Figure 6 The tail census: 1,002 scripted playthroughs and 200 completed human runs (of 401 started; completer statistics shown, an upper bound under attrition), 2023–2050. *(a)* Outcome composition by archetype and for human players: no doctrine dominates, each dies of its characteristic cause, and the human loss spectrum is broader (biosphere and food-price tails the bots rarely hit). *(b)* Timing of failures, bots vs. humans: both distributions peak in 2040–2049 as slow accumulators come due.

state-dominated political economies are represented by shared functional forms with country-specific states.

Conclusion

The contribution of this work, restated in its proper order: a closed seven-domain, 57-agent world simulator that (i) stays together — a detrended spectral radius whose above-unity modes are confined to the slow socio-fiscal subspace, contracting economic and climate perturbations, and near-linear uncertainty growth over thirty years; (ii) is anchored — in-sample and hold-out retrospective evaluation, external cross-checks, an input-vintage-robust conflict ranking; and (iii) is responsive — lagged social channels, threshold cascades, convex escalations, and clustered late failures whose shapes go beyond the coded terms they build on. Underneath is a stated philosophy: the deterministic core encodes explicit invariants of nature and society — conservation, accounting, slow empirical regularities — with calibrated restoring forces, and all uncertainty lives in ensembles over the invariants' strengths. Around it is an instrument: history-matched ensembles for distributions, weak-signal detectors for early structure, and a game layer that adversarially samples the tails, yields outcome distributions from both scripted archetypes and external human players, and has already paid for itself by exposing two structural degeneracies that no retrospective test could see. The model does not forecast; it maps the consequences of assumptions, the distribution of risks, and the ways a managed world can fail — comparatively, reproducibly, and in an explainable way.

Declarations

Ethics and player telemetry. The game telemetry reported in Section 7 is anonymous by design and analysed only in aggregate. The public deployment records, per run: a randomly generated run identifier, the random seed, the in-game outcome (win/loss, cause, year), the composite score, and final in-game aggregate metrics. There are no accounts; no names, e-mail addresses, IP addresses, demographic attributes, or device identifiers are collected or stored with runs; an optional nickname may be submitted after a run for the public leaderboard and is excluded from the analysis here. On this basis the collection was judged to fall outside the scope of human-subjects review (non-interventional, anonymous gameplay statistics); no ethics approval was therefore sought, and no individual-level data are reported.

Conflict of interest. The authors declare no competing interests.

Funding. This research received no funding of any kind and was carried out entirely on the authors' own time; article processing charges are borne by the authors.

Author contributions (CRediT). N.S.S.: conceptualization, methodology, software (the simulation engine and the game layer), investigation, writing — original draft. A.A.O.: validation (model validation programme), writing — original draft and review & editing. A.N.K.: validation (model and game-simulator testing), formal analysis (assessment of results), writing — review & editing. The archived software is authored by N.S.S. and A.A.O.; the Zenodo deposit's author list reflects the software, the byline reflects the paper.

Generative-AI disclosure. Large-language-model tools were used as coding and manuscript-editing assistants (drafting support, reference formatting, and consistency checks); all analyses, results, and claims were produced by the deterministic model harnesses and verified by the authors, who take full responsibility for the content. LLM use *inside* the model is a documented, optional feature (Section 7) and is off in every validated configuration.

Code and data availability

The complete model, calibration data with checksums, every validation and stability harness cited in the text, the game backend, and the figure-generation scripts are openly available at `https://github.com/Theclimateguy/GIM` [Sotiriadi and Osiptsov, 2026]; the archived Zenodo deposit is version 18.1.4 (`https://doi.org/10.5281/zenodo.21575176`), an archival repackaging of the v18.1.3 engine with no code change — all results in this paper correspond to that engine and are reproducible on a consumer laptop from the pinned release.

References

William D. Nordhaus. Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, 114(7):1518–1523, 2017. doi: 10.1073/pnas.1609244114.

Keywan Riahi et al. The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, 42:153–168, 2017. doi: 10.1016/j.gloenvcha.2016.05.009.

Robert M. May. Will a large complex system be stable? *Nature*, 238: 413–414, 1972. doi: 10.1038/238413a0.

John D. Sterman. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, Boston, 2000.

Oya Celasun, Jungjin Lee, Mico Mrkaic, and Allan Timmermann. An evaluation of World Economic Outlook growth forecasts, 2004–17. IMF Working Paper 2021/216, International Monetary Fund, 2021.

William D. Nordhaus and Zili Yang. A regional dynamic general-equilibrium model of alternative climate-change strategies. *The American Economic Review*, 86(4):741–765, 1996.

David Anthoff and Richard S. J. Tol. The uncertainty about the social cost of carbon: A decomposition analysis using FUND. *Climatic Change*, 117:515–530, 2013. doi: 10.1007/s10584-013-0706-7.

Katherine Calvin, Pralit Patel, Leon Clarke, et al. GCAM v5.1: representing the linkages between energy, water, land, climate, and economic systems. *Geoscientific Model Development*, 12:677–698, 2019. doi: 10.5194/gmd-12-677-2019.

Daniel Huppmann, Matthew Gidden, Oliver Fricko, et al. The MESSAGEix integrated assessment model and the ix modeling platform. *Environmental Modelling & Software*, 112:143–156, 2019. doi: 10.1016/j.envsoft.2018.11.012.

Jean-Francois Mercure, Hector Pollitt, Jorge E. Vinuales, Neil R. Edwards, Philip B. Holden, Unnada Chewpreecha, Pablo Salas, Ida Sognnaes, Aileen Lam, and Florian Knobloch. Macroeconomic impact of stranded fossil fuel assets. *Nature Climate Change*, 8: 588–593, 2018. doi: 10.1038/s41558-018-0182-1.

Millennium Institute. Threshold 21 (T21) overview. Technical report, Millennium Institute, Arlington, VA, 2002. `https://www.millenniuminstitute.net/integrated_planning/tools/t21/`.

Donella H. Meadows, Dennis L. Meadows, Jørgen Randers, and William W. Behrens. *The Limits to Growth*. Universe Books, New York, 1972.

Ilkka Keppo, Isabela Butnar, Nicolas Bauer, Matteo Caspani, Oreane Edelenbosch, Johannes Emmerling, Panagiotis Fragkos, Céline Guivarch, Mathijs Harmsen, Julien Lefevre, et al. Exploring the possibility space: taking stock of the diverse capabilities and gaps in integrated assessment models. *Environmental Research Letters*, 16(5), 2021.

Fortunat Joos, Raphael Roth, Jan S. Fuglestvedt, et al. Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis. *Atmospheric Chemistry and Physics*, 13:2793–2825, 2013. doi: 10.5194/acp-13-2793-2013.

Olivier Geoffroy, David Saint-Martin, Dirk J. L. Olivié, Aurore Voldoire, Gilles Bellon, and Sophie Tytéca. Transient climate response in a two-layer energy-balance model. part I: Analytical solution and parameter calibration using CMIP5 experiments. *Journal of Climate*, 26(6):1841–1857, 2013. doi: 10.1175/JCLI-D-12-00195.1.

IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, 2021.

Christopher J. Smith, Piers M. Forster, Myles Allen, Nicholas Leach, Richard J. Millar, Giovanni A. Passerello, and Leighton A. Regayre. FAIR v1.3: a simple emissions-based impulse response and carbon cycle model. *Geoscientific Model Development*, 11:2273–2297, 2018. doi: 10.5194/gmd-11-2273-2018.

Malte Meinshausen, Sarah C. B. Raper, and Tom M. L. Wigley. Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6. *Atmospheric Chemistry and Physics*, 11:1417–1456, 2011. doi: 10.5194/acp-11-1417-2011.

Pierre Friedlingstein et al. Global carbon budget 2023. *Earth System Science Data*, 15(12):5301–5369, 2023. doi: 10.5194/essd-15-5301-2023.

Wynne Godley and Marc Lavoie. *Monetary Economics: An Integrated Approach to Credit, Money, Income, Production and Wealth*. Palgrave Macmillan, Basingstoke, 2007.

Ben S. Bernanke, Mark Gertler, and Simon Gilchrist. The financial accelerator in a quantitative business cycle framework. In John B. Taylor and Michael Woodford, editors, *Handbook of Macroeconomics*, volume 1, pages 1341–1393. Elsevier, 1999.

Keith Head and Thierry Mayer. Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of International Economics*, volume 4, pages 131–195. Elsevier, 2014. doi: 10.1016/B978-0-444-54314-1.00003-3.

Lewis F. Richardson. Variation of the frequency of fatal quarrels with magnitude. *Journal of the American Statistical Association*, 43 (244):523–546, 1948. doi: 10.1080/01621459.1948.10483278.

Aaron Clauset, Cosma Rohilla Shalizi, and Mark E. J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009. doi: 10.1137/070710111.

Daniel Kaufmann, Aart Kraay, and Massimo Mastruzzi. The worldwide governance indicators: Methodology and analytical issues. *Hague Journal on the Rule of Law*, 3(2):220–246, 2011. doi: 10.1017/S1876404511200046.

Geert Hofstede, Gert Jan Hofstede, and Michael Minkov. *Cultures and Organizations: Software of the Mind*. McGraw-Hill, New York, 3 edition, 2010.

John B. Taylor. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39:195–214, 1993. doi: 10.1016/0167-2231(93)90009-L.

Naomi Oreskes, Kristin Shrader-Frechette, and Kenneth Belitz. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263(5147):641–646, 1994. doi: 10.1126/science.263.5147.641.

J. David Singer, Stuart Bremer, and John Stuckey. Capability distribution, uncertainty, and major power war, 1820–1965. In Bruce M. Russett, editor, *Peace, War, and Numbers*, pages 19–48. Sage, Beverly Hills, 1972.

Nazar S. Sotiriadi and Andrei A. Osiptsov. GIM: Global integrated model. `https://github.com/Theclimateguy/GIM`, 2026. Version 18.1.4 (archival repackaging of the v18.1.3 engine); code, calibration data, and validation harnesses.

Daniel Williamson et al. History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble. *Climate Dynamics*, 41:1703–1729, 2013. doi: 10.1007/s00382-013-1896-4.

Ralph Sundberg and Erik Melander. Introducing the UCDP georeferenced event dataset. *Journal of Peace Research*, 50(4):

523–532, 2013. doi: 10.1177/0022343313484347.

Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall/CRC, New York, 1993.

U.S. Environmental Protection Agency. Report on the social cost of greenhouse gases: Estimates incorporating recent scientific advances. Technical report, U.S. Environmental Protection Agency, 2023.

Kevin Rennert et al. Comprehensive evidence implies a higher social cost of CO2. *Nature*, 610:687–692, 2022. doi: 10.1038/s41586-022-05224-9.

Frederick Solt. Measuring income inequality across countries and over time: The standardized world income inequality database. *Social Science Quarterly*, 101(3):1183–1199, 2020. doi: 10.1111/ssqu.12795.

Max D. Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174, 1991. doi: 10.1080/00401706.1991.10484804.

Francesca Campolongo, Jessica Cariboni, and Andrea Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, 22(10):1509–1518, 2007. doi: 10.1016/j.envsoft.2006.10.004.

Marco Lagi, Karla Z. Bertrand, and Yaneer Bar-Yam. The food crises and political instability in north africa and the middle east. *arXiv preprint arXiv:1108.2455*, 2011.

Prasanta Chandra Mahalanobis. On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, 2(1):49–55, 1936.

Edwin B. Wilson and Margaret M. Hilferty. The distribution of chi-square. *Proceedings of the National Academy of Sciences*, 17(12): 684–688, 1931. doi: 10.1073/pnas.17.12.684.

Gideon Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978. doi: 10.1214/aos/1176344136.

Marten Scheffer et al. Early-warning signals for critical transitions. *Nature*, 461:53–59, 2009. doi: 10.1038/nature08227.

Kenneth J. Arrow, Hollis B. Chenery, Bagicha S. Minhas, and Robert M. Solow. Capital-labor substitution and economic efficiency. *The Review of Economics and Statistics*, 43(3):225–250, 1961. doi: 10.2307/1927286.

Edwin van der Werf. Production functions for climate policy modeling: An empirical analysis. *Energy Economics*, 30(6):2964–2979, 2008. doi: 10.1016/j.eneco.2008.05.008.

Mark J. Koetse, Henri L. F. de Groot, and Raymond J. G. M. Florax. Capital-energy substitution and shifts in factor demand: A meta-analysis. *Energy Economics*, 30(5):2236–2251, 2008. doi: 10.1016/j.eneco.2007.06.006.

Robert C. Feenstra, Robert Inklaar, and Marcel P. Timmer. The next generation of the Penn World Table. *American Economic Review*, 105(10):3150–3182, 2015. doi: 10.1257/aer.20130954.

Douglas Gollin. Getting income shares right. *Journal of Political Economy*, 110(2):458–474, 2002. doi: 10.1086/338747.

Ashoka Mody, Franziska Ohnsorge, and Damiano Sandri. Precautionary savings in the great recession. *IMF Economic Review*, 60(1):114–138, 2012. doi: 10.1057/imfer.2012.5.

Charles I. Jones. R&d-based models of economic growth. *Journal of Political Economy*, 103(4):759–784, 1995. doi: 10.1086/262002.

Gunnar Myhre, Eleanor J. Highwood, Keith P. Shine, and Frode Stordal. New estimates of radiative forcing due to well mixed greenhouse gases. *Geophysical Research Letters*, 25(14):2715–2718, 1998. doi: 10.1029/98GL01908.

Peter H. Howard and Thomas Sterner. Few and not so far between: A meta-analysis of climate damage estimates. *Environmental and Resource Economics*, 68:197–225, 2017. doi: 10.1007/s10640-017-0166-z.

Marshall Burke, Solomon M. Hsiang, and Edward Miguel. Global non-linear effect of temperature on economic production. *Nature*, 527: 235–239, 2015. doi: 10.1038/nature15725.

Marshall Burke, Mustafa Zahid, Noah S. Diffenbaugh, and Solomon Hsiang. Quantifying climate change loss and damage consistent with a social cost of greenhouse gases. NBER Working Paper 31658, National Bureau of Economic Research, 2023.

Alexander Otto, Friederike E. L. Otto, Olivier Boucher, et al. Energy budget constraints on climate response. *Nature Geoscience*, 6:415–416, 2013. doi: 10.1038/ngeo1836.

Nicholas Stern. *The Economics of Climate Change: The Stern Review*. Cambridge University Press, Cambridge, 2007.

Xavier Labandeira, José M. Labeaga, and Xiral López-Otero. A meta-analysis on the price elasticity of energy demand. *Energy Policy*, 102:549–568, 2017. doi: 10.1016/j.enpol.2017.01.002.

Rafael Di Tella, Robert J. MacCulloch, and Andrew J. Oswald. Preferences over inflation and unemployment: Evidence from surveys of happiness. *American Economic Review*, 91(1):335–341, 2001. doi: 10.1257/aer.91.1.335.

Richard Wilkinson and Kate Pickett. *The Spirit Level: Why More Equal Societies Almost Always Do Better*. Allen Lane, London, 2009.

Model specification and key parameters

Production. Output is nested CES in capital, labour, and energy, normalized to Cobb–Douglas at each country's base point [Arrow et al., 1961]:

$$Y = A\,\Theta\,\mathrm{KE}_0^{\alpha+\gamma}\big(a_K(K/K_0)^{\rho_{KE}} + a_E(E/E_0)^{\rho_{KE}}\big)^{(\alpha+\gamma)/\rho_{KE}}L^{\beta}, \qquad \rho_{KE} = \tfrac{\sigma_{KE}-1}{\sigma_{KE}}, \tag{1}$$

with $\alpha = 0.30$, $\beta = 0.60$, $\gamma = 0.042$, $a_K = \alpha/(\alpha+\gamma)$, $a_E = \gamma/(\alpha+\gamma)$, $\mathrm{KE}_0 = K_0^{a_K}E_0^{a_E}$; at the base point the inner aggregator (the parenthesis) equals one because $a_K + a_E = 1$, so levels and first derivatives reproduce the Cobb–Douglas limit for any σ_{KE}, while substitution away from it is governed by $\sigma_{KE} = 0.4$ (gross complementarity [van der Werf, 2008, Koetse et al., 2008]; factor shares from the Penn World Table [Feenstra et al., 2015, Gollin, 2002]). The exponent sum 0.942 encodes mild decreasing returns; per-country scale factors pin base-year levels, so renormalising to constant returns moves 2100 world product by only 7–9 %. Capital: $K_{t+1} = (1-\delta)K_t + sY_t$, $\delta = 0.05$, $s \approx 0.24$ modulated by stability and tension [Mody et al., 2012]. Productivity grows at ≈1 % baseline with R&D (à la Jones [Jones, 1995]), trade diffusion, and β-convergence (slope 0.0093 per log gap, validated on the cross-section); inflation is Phillips-with-pass-through plus a weak quantity-theory term ($\lambda \approx 0.027$); rates follow country Taylor rules capped ±6 pp [Taylor, 1993].

Climate. Four-pool AR6 impulse response [Joos et al., 2013]:

$$C_{t+1}^{(i)} = C_t^{(i)}e^{-1/\tau_i} + a_iE_t, \quad a = (0.2173, 0.2240, 0.2824, 0.2763), \qquad \tau = (\infty, 394, 36.5, 4.3)\,\mathrm{yr}, \tag{2}$$

(shares sum to one exactly; rounded here). Forcing is $F = 5.35\ln(C/C_0) + F_{\text{non-CO}_2}$ — the Myhre et al. [Myhre et al., 1998] coefficient, implying $F_{2\times} = 3.71\ \mathrm{W\,m^{-2}}$ and, at the calibrated equilibrium sensitivity of 3.0 °C, a feedback parameter $\lambda = 1.24\ \mathrm{W\,m^{-2}K^{-1}}$; the AR6-updated $F_{2\times}$ is higher, but since sensitivity is calibrated at doubling the equilibrium anchor is unaffected and the transient response is checked directly (TCR 1.79 °C, within the AR6 likely range) — with the non-CO_2 term on the SSP2-4.5 marker forward. Two-layer heat balance [Geoffroy et al., 2013]:

$$C_s\dot{T} = F - \lambda T - \kappa(T - T_d), \qquad C_d\dot{T}_d = \kappa(T - T_d). \tag{3}$$

Damage: $\Omega(T) = 1 - 0.0078\,T^2$ (3.1/7.0/12.5 % at $+2/+3/+4$ °C), anchored at the Howard–Sterner productivity-corrected central estimate [Howard and Sterner, 2017, Nordhaus, 2017], normalized to the 2023 baseline; Burke-type growth damages switchable [Burke et al., 2015, 2023].

Society and crises. Tension is a $[0,1]$ composite (base year: WGI-derived trust and stability, Gini, water stress, climate risk; increments: unemployment 0.010/pt, inflation 0.005/pt, inequality damped by individualism, trust anchor, neighbour spillover 0.05). Legitimacy $= 0.6\,\mathrm{trust} + 0.4(1 - \mathrm{tension})$ sets the policy-space multiplier. Crisis triggers (expert priors, tagged in the registry): debt — debt > 120 % GDP and effective rate > 12 % (then a 40 % write-down, 7 % output loss, exit below 70 %/8 % within 6 years); currency — external debt > 50 %, current-account deficit > 4 %, reserves < 3 months; regime collapse — trust < 0.20 and tension > 0.80. Severity optionally heavy-tailed (war-size exponent $\alpha_w \approx 1.5$ [Richardson, 1948, Clauset et al., 2009]).

History matching and ensemble formation. Priors are per-parameter triangular (30 of 33), lognormal (2), or normal (1) distributions with the ranges of Tables 2–3 (full registry in the repository). Implausibility uses tolerances tol_o of 6 trillion USD (world product), 2 Gt (emissions), and 0.15 °C (temperature) — observational uncertainty bundled with an explicit model-discrepancy allowance — with the cut $\max_o I_o \leq 3$. The published configuration uses a single wave of 60 Latin-style random draws over the six backtest-driving parameters; at the three-sigma cut the acceptance rate is 100 % (the cut removes only corners more implausible than three combined standard errors, and the priors are already literature-bounded), so the NROY region coincides with the prior box and ensembles draw all 33 parameters independently from their priors. This is disclosed rather than hidden: the tolerances are wide because model discrepancy dominates, and tighter waves are future work. The 33 key parameters carry literature-grounded priors; the full registry of 287 parameters with calibration-status tags, all data-source manifests, and every harness cited in the text are in the repository [Sotiriadi and Osiptsov, 2026].

Table 2 Key priors, part 1: the 19 parameters discussed in the text (units inline; dimensionless unless stated). Central values and 5–95th plausibility ranges.

Parameter	Central	Range	Source
Equilibrium climate sensitivity, $^\circ$C	3.0	1.5–6.0	[IPCC, 2021, Otto et al., 2013]
Surface heat capacity, $\mathrm{W\,yr\,m^{-2}K^{-1}}$	8.0	6.0–18.0	[Geoffroy et al., 2013]
Ocean heat exchange, $\mathrm{W\,m^{-2}K^{-1}}$	1.0	0.4–1.3	[Geoffroy et al., 2013]
Emissions scale	0.98	0.90–1.05	[Friedlingstein et al., 2023]
Structural decarbonization rate, $\mathrm{yr^{-1}}$	0.016	0.010–0.030	GCP/WDI trend [Friedlingstein et al., 2023]; 2015–2023 backtest
Capital share α	0.30	0.22–0.40	[Feenstra et al., 2015, Gollin, 2002]
Labor share β	0.60	0.50–0.72	[Feenstra et al., 2015, Gollin, 2002]
Energy elasticity γ	0.042	0.015–0.080	2015–2023 backtest
Capital–energy substitution σ_{KE}	0.40	0.20–0.60	[van der Werf, 2008, Koetse et al., 2008]
Depreciation rate δ, $\mathrm{yr^{-1}}$	0.05	0.03–0.08	[Feenstra et al., 2015]
TFP sensitivity to R&D	0.30	0.10–0.60	calibration to backtest
Damage coefficient	0.0078	0.0015–0.025	[Howard and Sterner, 2017, Nordhaus, 2017, Burke et al., 2015]
Pure time preference	0.015	0.001–0.030	[Nordhaus, 2017, Stern, 2007]
Elasticity of marginal utility η	1.45	1.0–2.0	[Nordhaus, 2017, Stern, 2007]
Market demand elasticity	0.40	0.20–0.70	[Labandeira et al., 2017]
Social block (validated by sign and channel, Sec. 6)			
Tension stress from unemployment	0.010	0.007–0.014	[Di Tella et al., 2001]; structural prior
Tension stress from inflation	0.005	0.003–0.007	[Di Tella et al., 2001] (aversion ≈2:1); prior
Inequality effect on tension	0.0005	0.0003–0.0007	[Wilkinson and Pickett, 2009]
Trust-anchor strength	0.060	0.039–0.081	calibrated to WGI trust persistence

Table 3 Key priors, part 2: the remaining 14 of the 33 sampled parameters (all switch-off mechanisms sampled only in tail scenarios are excluded from headline runs).

Parameter	Central	Range	Source
Warming-benefit cap	0	0–0.02	prior (CO_2 fertilisation; off headline)
Warming-benefit peak, $^\circ$C	0.5	0–1.0	prior
Damage risk adjustment	0.01	0–0.02	prior
Deep-ocean heat capacity	100	40–200	Geoffroy et al. 2013; DICE-2016
Land-use CO_2, Gt/yr	0.7	0–1.5	Global Carbon Project 2023
Carbon feedback, $GtCO_2/^\circ C$	1.5	0–5.0	AR6 WG1 Ch.5; permafrost literature
CH_4 feedback, $\mathrm{W\,m^{-2}/^\circ C}$	0.03	0–0.1	AR6 WG1 Ch.5
Base birth rate	0.025	0.015–0.035	World Bank WDI
Base death rate	0.012	0.007–0.017	World Bank WDI
Neutral interest rate	0.02	0.005–0.04	IMF WEO; r^* literature
Growth-damage TFP coefficient	0	0–0.001	Burke et al. 2015 (off headline)
War-size tail exponent α_w	1.5	1.3–1.8	[Richardson, 1948, Clauset et al., 2009]
Migration base / max share	0.001/0.003	see registry	UN DESA; UNHCR
Regime-collapse output/capital mult.	0.8/0.7	0.7–0.9 / 0.55–0.85	Barro & Ursúa 2008

Integration benchmark: where sectoral models stop

For three shocks we take a published conclusion from a named sectoral model, feed GIM the same exogenous shock, and ask whether it (a) reproduces the sectoral first-order effect and (b) surfaces a cross-sector consequence the sectoral model cannot represent by construction. Runs are deterministic (57 agents, ten years, extreme events off); each scenario carries a dose–response sweep on which the sectoral model's slope is, by construction, zero.

Carbon tax (anchor: DICE [Nordhaus, 2017]). Fed the same \$50/tCO_2 tax, GIM reproduces the emissions decline (≈5.5 % over ten years) and adds the political-economy channel: tension rises and trust falls monotonically in the rate (the reversal mechanism discussed in Section 7); DICE's response on that axis is zero.

Oil shock (anchor: energy-system models). A 20 % supply cut propagates into sovereign finance: two additional importers enter debt crisis at a realistic burden, four more with triple exposure at a severe one; the response is threshold-stepped (Sri Lanka 2022 is the real-world instance). Energy-system models produce no sovereign crises at any magnitude.

Crop shock (anchor: AgMIP/AR6). A 50 % regional yield loss raises food-affordability stress by ≈0.11 and protest pressure by ≈0.016 on the vulnerable subset, convex in loss depth [Lagi et al., 2011]; the assessment chain stops at the food-security index.

Across the three channels GIM shows a positive, measurable slope where every sectoral model is flat at zero (Fig. 7) — the integration dividend, quantified.

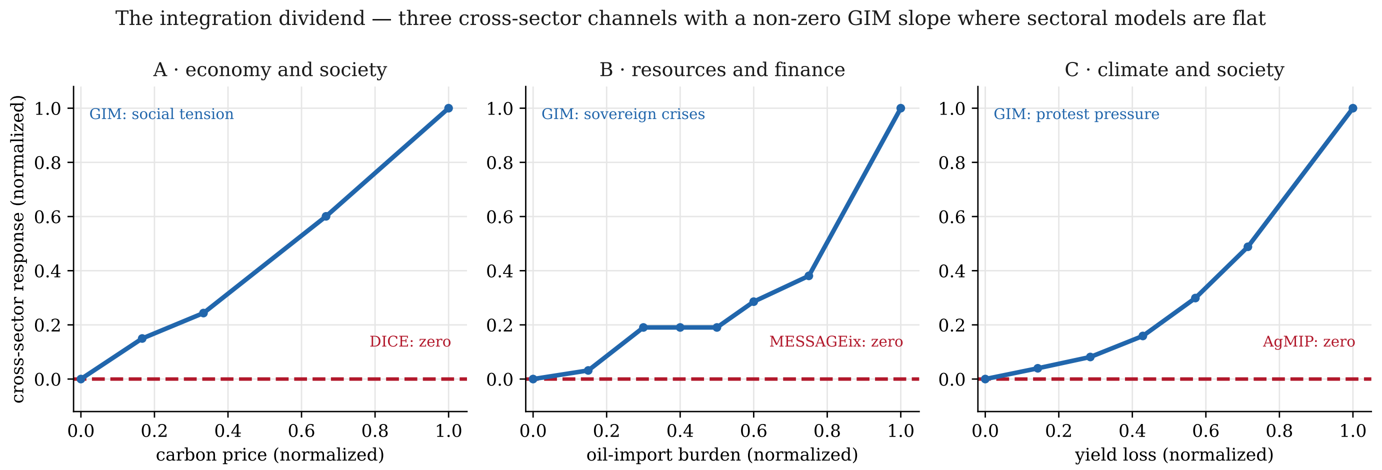

Figure 7 The integration dividend: three cross-sector channels, each normalized to its own maximum; the dashed line is every sectoral model's zero response. Smooth for economy–society, threshold-stepped for resources–finance, convex for climate–society.