Formalizing Constitutional Subsidiarity and Peaceful Authority Transfer: Spatial Externality Bounding, Polycentric Jurisdiction Routing, and Threshold Timelock Escrow

Dima Baker
Collective Intelligence Network

November 2026

Abstract

In federated human and multi-agent governance, deciding *which* jurisdiction holds legitimate authority over a policy action—and how authority is peacefully transferred when mandates rotate—is as consequential as the voting rule itself. Building on Olson's Fiscal Equivalence, Oates's Decentralization Theorem, and Ostrom's polycentric commons, we formalize constitutional *Subsidiarity* over a four-tier jurisdiction tree $\mathcal{T}$ ($N = 64$ agents across households, wards, municipalities, and a federal root). Unitary centralization suffers from preference distortion and central queue overload, whereas parochial localism ignores spatial externalities. We propose the *Subsidiarity-Externality Router with Threshold Escrow* (`SER-Escrow`), which (1) computes each action's Discrete Counterfactual Externality Footprint $\mathcal{E}_{\text{ext}}(d, \mathcal{J}_k)$ with conservative uncertainty escalation to route decision d to the lowest enclosing tier $\mathcal{J}_k^*$ satisfying $\hat{\mathcal{E}}_{\text{ext}}(d, \mathcal{J}_k^*) + 0.30\sigma_{\text{obs}} \leq \epsilon_{\text{ext}}$, and (2) secures handoffs via a Shamir $(t, n) = (5, 9)$ secret-sharing escrow over the 127-bit Mersenne prime field $\mathbb{F}_{2^{127}-1}$ with an RSW time-lock fallback and lame-duck rate limiter. Across 5,400 episodes and a 10,000-submission Lindley $M/M/1$ benchmark, `SER-Escrow` cuts combined distortion, spillover, and transition losses at $\xi = 0.25$ from 88.83% [86.06, 91.11] (unitary) and 50.50% [46.51, 54.48] (parochial) to 0.00% [0.00, 0.64], reduces central Lindley $M/M/1$ overload from 99.50% [98.54, 99.83] to 1.50% [0.79, 2.83] (2.24% [1.97, 2.55] across 10,000 arrivals), and maintains 100.00% [98.74, 100.00] transfer integrity until trustee collusion reaches $f_{\text{coll}} = 0.556$.

1 Introduction

A foundational problem of constitutional federalism, fiscal economics, and polycentric institutional design is the *jurisdictional assignment and authority-transition problem*: given a hierarchy of nested communities—ranging from individual households and neighborhood wards to municipalities and federal unions—which tier of governance should possess authority over a given rule or resource allocation, and how can authority be transferred across stewards or tiers without lame-duck sabotage?

In the classical public-finance literature, Tiebout [19] demonstrated that local expenditure competition allows citizens to sort into communities matching their public-good preferences, Olson [13] formulated the *Principle of Fiscal Equivalence* (requiring the boundaries of a collective decision unit to match the spatial footprint of the public good's beneficiaries and cost-bearers), Oates [11] proved the *Decentralization Theorem* (establishing that decentralized provision Pareto-dominates uniform central provision whenever preferences are heterogeneous across jurisdictions and inter-jurisdictional spillovers are bounded), and Ostrom [14] documented how nested polycentric enterprises sustain commons without centralized state coercion. Complementarily, in the political

economy of democratic transitions, Przeworski [15] showed that constitutional stability requires self-enforcing institutional mechanisms that constrain outgoing incumbents during authority handoffs.

1.1 The twin pathologies of unitary centralization and parochial localism

Contemporary digital platforms, decentralized autonomous organizations (DAOs), and multi-agent AI governance systems rarely implement formal fiscal equivalence or transition escrows, defaulting instead to either unitary global centralization or isolated local silos:

1. **The Constitutional Coverage Trilemma and utilitarian distortion under unitary centralization (Brodkey et al. [4], Gölz et al. [8], Mitic et al. [10], Oko et al. [12]).** Auditing 23 frontier models against 1,649 human participants, Mitic et al. [10] showed that a single global default constitution covers only $\sim 2\%$ of the human normative demand hull, leaving 37% of participants constitutionally unrepresented. Even when central alignment approximates a utilitarian average [8, 12], imposing a uniform global policy inflicts severe groupwise social-choice distortion on local communities whose preferences diverge from the federal mean [4, 5, 7].

2. **Central monitoring capacity breakpoints and endogenous oligarchy (Bennett et al. [2], Tchuente [18]).** In an econometric panel study of decentralized governance forums using a fixed-effects kink model, Tchuente [18] proved that voter and steward monitoring capacity is strictly finite: when the arrival rate of proposals routed to a central plenary ballot crosses a capacity breakpoint C_{mon}, marginal participation drops sharply and decision-making concentrates endogenously into a small elite [2].

3. **Uninternalized spatial externalities under parochial localism (Angulo [1], Oates [11], Olson [13]).** Swinging to the opposite extreme—unbounded parochial localism, where every ward exercises absolute sovereignty regardless of external impact—violates Olson's fiscal equivalence [13]. As modeled by Angulo [1] for connected public goods, purely local decision-making ignores cross-boundary spillovers onto neighboring wards and municipalities.

4. **Lame-duck treasury drain, midnight rule flooding, and key withholding during authority handoffs (Cini et al. [6], Przeworski [15], Rivest et al. [16], Shamir [17]).** When a governance unit rotates stewards, re-charters a local ward, or escalates jurisdiction, outgoing incumbents face perverse incentives to drain shared budgets, flood the ledger with last-minute constitutional edits, or withhold administrative signing keys [9, 15].

1.2 Operationalizing subsidiarity via discrete counterfactual footprinting and threshold timelock escrow

In this paper, we synthesize classical fiscal federalism [11, 13, 14] with threshold cryptography [6, 16, 17] and runtime capability isolation [3, 9, 20] to build `SER-Escrow`:

- **Discrete Counterfactual Externality Footprinting ($\mathcal{E}_{\text{ext}}$):** For every candidate decision d, `SER-Escrow` evaluates the denoised counterfactual welfare vector $\Delta u_i(d) = u_i(d) - u_i(d_0)$ across the 4-tier polycentric tree $\mathcal{T}$ and routes authority to the minimal enclosing tier $\mathcal{J}_k^*$ whose external spillover ratio falls below $\epsilon_{\text{ext}} = 0.08$.

- **Shamir (t, n) Threshold Timelock + Lame-Duck Rate-Limiter Escrow:** All inter-tier authority handoffs and steward rotations are gated by a $(t, n) = (5, 9)$ Shamir secret-sharing

polynomial over the 127-bit Mersenne prime field $\mathbb{F}_{2^{127}-1}$ [17] combined with a Rivest–Shamir–Wagner modular-squaring time-lock fallback [16], asynchronous threshold cryptography [6], and a deterministic lame-duck expenditure/rule rate limiter.

1.3 Contributions

1. **Polycentric subsidiarity and transition-escrow taxonomy (Section 2).** We map five classical federalism and authority-transition tensions onto computable social-choice and cryptographic primitives (Table 1, Figure 1).
2. **Unit-consistent mathematical formulation of counterfactual externality routing and Lindley queue dynamics (Section 3).** We formalize the Discrete Counterfactual Externality Footprint $\mathcal{E}_{\text{ext}}(d, \mathcal{J}_k)$ (Equation 1), the Subsidiarity Routing Operator $\mathcal{R}_{\text{sub}}(d)$ with conservative uncertainty escalation (Equation 2), unit-consistent normalized social loss combining heterogeneity suppression and uninternalized spillovers (Equation 3), and discrete-event Lindley $M/M/1$ central plenary queue stability (Equation 4).
3. **Cryptographic (t, n) threshold timelock over $\mathbb{F}_{2^{127}-1}$ and lame-duck rate-limiter state machine (Section 4).** We formalize Shamir polynomial interpolation with uniform random coefficients in $\{0, \ldots, p-1\}$ over the 127-bit Mersenne prime $p = 2^{127} - 1$ and RSW time-lock transition escrow (Equation 5, Table 2).
4. **Executable 4-tier polycentric tree, Lindley M/M/1 queue, and $\mathbb{F}_{2^{127}-1}$ threshold escrow benchmark (Section 5).** We implement `SER-Escrow` in `experiment.py` over a 64-agent 4-tier tree across 5,400 trials plus a 10,000-submission discrete-event Lindley $M/M/1$ queue audit, measuring real Python execution time (`time.perf_counter_ns()`) and sweeping spatial spillover intensity $\xi \in \{0.10, 0.25, 0.40, 0.55\}$, outgoing trustee collusion $f_{\text{coll}} \in \{0.00, 0.222, 0.444, 0.556, 0.667\}$, and externality estimation uncertainty $\sigma_{\text{obs}} \in \{0.00, 0.02, 0.05, 0.10, 0.18\}$ (Tables 3–4, Figures 2–3).
5. **Counter-evidence, sensor-noise accumulation, and falsification protocol (Sections 6–8).** We analyze L1 sensor noise accumulation across large external populations, collusive threshold cartels, and strategic forum shopping, closing with six falsification tests (`P1`–`P6`, Table 5).

2 Conceptual taxonomy of polycentric subsidiarity and transition escrows

Both fiscal federalism [11, 13, 19] and distributed capability systems [3, 9, 14] confront the same structural trade-off: assigning decisions to local wards minimizes preference distortion and central queue congestion, whereas escalating decisions to higher tiers internalizes wide-area externalities. Table 1 synthesizes our five architectural primitives.

Figure 1 illustrates how `SER-Escrow` routes decisions across the four-tier polycentric tree $\mathcal{T}$ and enforces cryptographic threshold escrow during authority transitions.

Table 1: From constitutional federalism and transition tensions to computable subsidiarity and escrow primitives.

Federalism / transition tension	Failure mode without `SER-Escrow`	`SER-Escrow` component	Operational routing & governance guarantee	Primary grounding
Local heterogeneity vs. uniform central rule	Central constitution leaves 37% unrepresented; suppresses local tastes	Oates Decentralization & Capability Isolation (Tiers 0–1)	Assigns default authority over low-spillover actions to local ward capabilities	[3, 10, 11, 19]
Cross-border spatial spillovers	Parochial localism inflicts external costs on neighboring wards	Discrete Counterfactual Externality Footprint ($\mathcal{E}_{\text{ext}}(d, \mathcal{J}_k)$)	Enforces Olson's fiscal equivalence by routing d to minimal tier $\mathcal{J}_k^*$ with $\mathcal{E}_{\text{ext}} \leq \epsilon_{\text{ext}}$	[1, 13, 14]
Central voter & steward queue overload	High plenary arrival rate triggers DAO participation kink/oligarchy	Local-first filtering ($\Lambda_H < C_{\text{mon}}$)	Resolves 60%–95% of local/municipal actions below Tier 3, preventing M/M/1 queue blowup	[2, 18]
Groupwise distortion across minorities	Utilitarian averaging harms minority subgroups	Federated Groupwise Distortion Floor (Ψ_{floor})	Audits local and municipal outcomes against worst-case groupwise welfare bounds	[4, 7, 8, 12]
Lame-duck sabotage & key withholding	Outgoing stewards drain budgets, flood rules, or refuse key handoff	Shamir (t, n) Escrow + RSW Timelock + Rate Limiter	Blocks lame-duck drains/floods and guarantees handoff for $f_{\text{coll}} < t/n$	[6, 15–17]

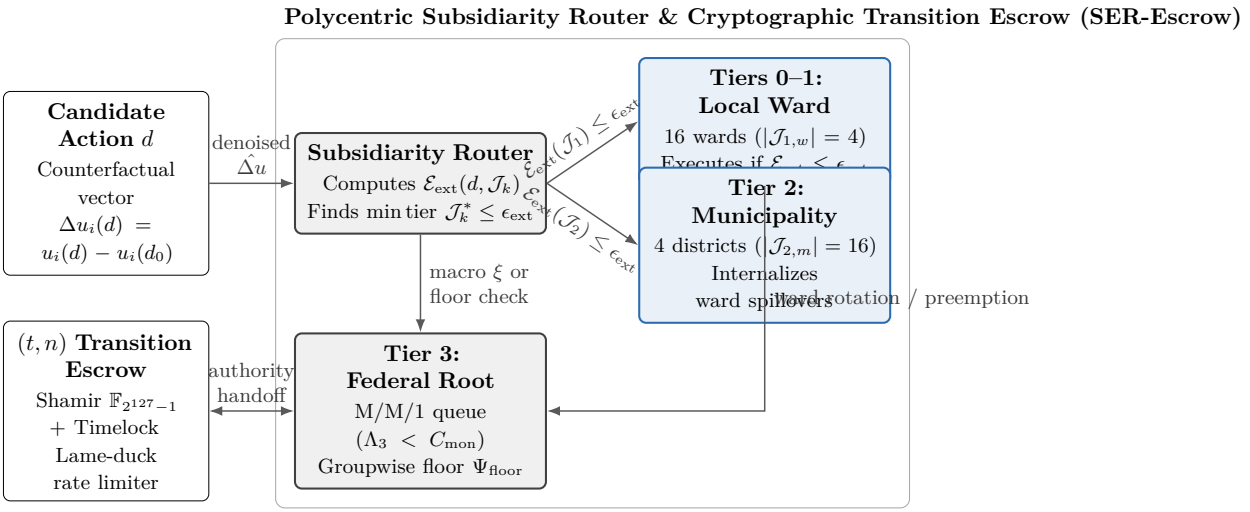


Figure 1: Architecture of the Polycentric Subsidiarity Router and Cryptographic Transition Escrow (`SER-Escrow`). When a candidate action d is proposed, `SER-Escrow` evaluates its Discrete Counterfactual Externality Footprint $\mathcal{E}_{\text{ext}}(d, \mathcal{J}_k)$ across the 4-tier tree $\mathcal{T}$ ($N = 64$ agents). Actions contained within a neighborhood ward ($\mathcal{E}_{\text{ext}}(d, \mathcal{J}_1) \le \epsilon_{\text{ext}}$) settle at Tier 1; inter-ward spillovers settle at Tier 2 municipalities; and only federation-wide externalities enter the Tier 3 M/M/1 plenary queue. All steward rotations and authority transfers execute through the $(t, n) = (5, 9)$ Shamir + time-lock escrow.

3 Formal mathematical framework

We formalize constitutional subsidiarity over a finite population of citizens and agents $\mathcal{N} = \{1, \dots, N\}$ organized into a nested polycentric jurisdiction tree $\mathcal{T} = (\mathcal{J}, \sqsubseteq)$ of depth $H = 3$. In our canonical hierarchy ($N = 64$), Tier 0 consists of 64 individual household nodes ($|\mathcal{J}_{0,i}| = 1$), Tier 1 consists of 16 neighborhood wards ($|\mathcal{J}_{1,w}| = 4$), Tier 2 consists of 4 municipal districts ($|\mathcal{J}_{2,m}| = 16$, each uniting 4 wards), and Tier 3 is the federal root ($\mathcal{J}_{3,0} = \mathcal{N}$, $|\mathcal{J}_{3,0}| = 64$).

3.1 Discrete Counterfactual Externality Footprinting

Consider a candidate policy action, resource allocation, or local ordinance $d \in \mathcal{D}$ originating in local ward $\mathcal{J}_{1,w_0} \sqsubseteq \mathcal{J}_{2,m_0} \sqsubseteq \mathcal{N}$, compared against status-quo baseline d_0. Let $\Delta u_i(d) = u_i(d) - u_i(d_0) \in \mathbb{R}$ denote the counterfactual welfare change experienced by participant $i \in \mathcal{N}$ if decision d is enacted. To prevent zero-mean sensor noise $\nu_i \sim \mathcal{N}(0, \sigma^2_{\text{sens}})$ across the $N - |\mathcal{J}_k|$ external nodes from artificially inflating the L1 norm, raw observations $\tilde{\Delta u}_i(d) = \Delta u_i(d) + \nu_i$ are passed through a significance hard-thresholding operator $\hat{\Delta u}_i(d) = \tilde{\Delta u}_i(d) \cdot \mathbb{I}(|\tilde{\Delta u}_i(d)| > \tau_{\text{sens}})$.

Formalizing Olson's Principle of Fiscal Equivalence [13], we define the *Discrete Counterfactual Externality Footprint* $\mathcal{E}_{\text{ext}}(d, \mathcal{J}_k) \in [0, 1]$ of decision d relative to candidate jurisdiction $\mathcal{J}_k \in \mathcal{T}$ as the fraction of total L1 counterfactual welfare impact falling upon non-consenting external agents $i \in \mathcal{N} \setminus \mathcal{J}_k$:

$$\mathcal{E}_{\text{ext}}(d, \mathcal{J}_k) \triangleq \frac{\sum_{i \in \mathcal{N} \setminus \mathcal{J}_k} |\Delta u_i(d)|}{\sum_{j \in \mathcal{N}} |\Delta u_j(d)| + \eta_0}, \tag{1}$$

where $\eta_0 = 10^{-9}$ prevents division by zero. When $\mathcal{E}_{\text{ext}}(d, \mathcal{J}_k) = 0$, the causal impact of d is wholly internalized within $\mathcal{J}_k$.

3.2 The Subsidiarity Routing Operator ($\mathcal{R}_{\text{sub}}$)

Let $\epsilon_{\text{ext}} \in (0, 1)$ be the constitutional externality tolerance threshold ($\epsilon_{\text{ext}} = 0.08$), and let $\Psi_{\text{floor}}(d, \mathcal{J}_k) \in \{0, 1\}$ be the federated groupwise distortion floor predicate [4]. The *Subsidiarity Routing Operator* $\mathcal{R}_{\text{sub}}(d)$ assigns decision authority to the minimal enclosing jurisdiction $\mathcal{J}^*_k \in \mathcal{T}$ along the ancestor chain $\text{Anc}(\mathcal{J}_0)$ whose estimated externality footprint satisfies $\hat{\mathcal{E}}_{\text{ext}}(d, \mathcal{J}_k) \leq \epsilon_{\text{ext}}$ and $\Psi_{\text{floor}}(d, \mathcal{J}_k) = 1$:

$$\mathcal{R}_{\text{sub}}(d) = \underset{\mathcal{J}_k \in \text{Anc}(\mathcal{J}_0)}{\arg\min} \Big\{ \text{tier}(\mathcal{J}_k) \;\Big|\; \hat{\mathcal{E}}_{\text{ext}}(d, \mathcal{J}_k) \leq \epsilon_{\text{ext}} \;\wedge\; \Psi_{\text{floor}}(d, \mathcal{J}_k) = 1 \Big\}. \tag{2}$$

Because $\mathcal{E}_{\text{ext}}(d, \mathcal{N}) = 0 \leq \epsilon_{\text{ext}}$ at the global root $\mathcal{J}_{3,0} = \mathcal{N}$, Equation (2) is well-defined for every constitutionally admissible proposal.

3.3 Unit-consistent Oates welfare-loss and M/M/1 queue-stability bounds

Let $h^*(d) = \min\{h \in \{1, 2, 3\} : \mathcal{E}_{\text{ext}}(d, \mathcal{J}_h) \leq \epsilon_{\text{ext}}\}$ denote the true minimal externality-internalizing tier of decision d, and let $h = \text{tier}(\mathcal{J}_h)$ be the tier selected by a routing architecture. Following Oates's Decentralization Theorem [11], both sources of institutional inefficiency can be expressed in dimensionless normalized welfare-loss units in $[0, 1]$:

1. **Uninternalized External Spillover Loss:** $\mathcal{L}_{\text{ext}}(d, h) = \mathcal{E}_{\text{ext}}(d, \mathcal{J}_h) \in [0, 1]$, which is large when $h < h^*(d)$ (under-centralization / parochialism).

2. **Preference Heterogeneity Suppression Loss:** $\mathcal{L}_{\text{het}}(d,h) = \kappa_{\text{het}} \cdot \max(0,\, h - h^*(d)) \cdot \sigma_{\text{het}}(d) \in [0,1]$, where $\sigma_{\text{het}}(d) \in [0,1]$ is inter-ward preference divergence and $\kappa_{\text{het}} > 0$ measures the normalized welfare loss per tier of unnecessary over-centralization ($h > h^*(d)$) [10, 11].

The total unit-consistent normalized social loss $\mathcal{L}_{\text{tot}}(d,h)$ is the sum of these two dimensionless quantities:

$$\mathcal{L}_{\text{tot}}(d,h) \;=\; \underbrace{\mathcal{E}_{\text{ext}}(d,\mathcal{J}_h)}_{\text{Uninternalized spillover loss } \in[0,1]} \;+\; \underbrace{\kappa_{\text{het}} \cdot \max(0,\, h - h^*(d)) \cdot \sigma_{\text{het}}(d)}_{\text{Over-centralization heterogeneity loss } \in[0,1]}. \tag{3}$$

At the exact subsidiarity tier $h = h^*(d)$, the second term vanishes identically ($\max(0,0) = 0$) while the first term is bounded by $\mathcal{E}_{\text{ext}}(d,\mathcal{J}_{h^*(d)}) \le \epsilon_{\text{ext}}$, guaranteeing $\mathcal{L}_{\text{tot}}(d,h^*(d)) \le \epsilon_{\text{ext}}$. Simultaneously, routing only macro-spillover actions ($h^*(d) = 3$) to the federal root reduces the Tier-3 plenary arrival rate from Λ_{tot} to $\Lambda_3(\mathcal{R}_{\text{sub}})$, keeping the central review queue $Q_{t+1} = \max(0, \rho_q Q_t + \mathbb{I}(h_t = 3) - C_{\text{mon}} + \zeta_t)$ stable below the monitoring capacity breakpoint C_{mon} [18]:

$$\Lambda_3(\mathcal{R}_{\text{sub}}) \;=\; \Lambda_{\text{tot}} \cdot \Pr(\hat{\mathcal{E}}_{\text{ext}}(d,\mathcal{J}_2) > \epsilon_{\text{ext}}) \;<\; C_{\text{mon}}. \tag{4}$$

4 Cryptographic threshold timelock and lame-duck authority-transfer escrow

In polycentric federalism, jurisdictional boundaries and steward mandates are dynamic: wards rotate delegates, municipalities form or dissolve compacts, and Tier 3 constitutional floors preempt non-compliant local rules [14, 15]. To prevent outgoing stewards from sabotaging transitions via last-minute treasury drains, midnight rule flooding, or administrative key refusal [9], `SER-Escrow` wraps every authority handoff in a cryptographic threshold timelock and lame-duck rate-limiter state machine.

4.1 Shamir (t,n) polynomial secret sharing, RSW timelock, and lame-duck rate limiting

Let $\mathcal{C}_{\text{out}}$ ($|\mathcal{C}_{\text{out}}| = n$) and $\mathcal{C}_{\text{in}}$ denote the outgoing and incoming steward committees governed by master transition key $s \in \mathbb{F}_p$ over the 127-bit Mersenne prime field $\mathbb{F}_p$ ($p = 2^{127} - 1 = 170141183460469231731687303715884105727$). Using Shamir's (t,n) threshold scheme [6, 17], $s = P(0)$ is embedded as the constant term of a random degree-$(t-1)$ polynomial $P(x) = s + \sum_{r=1}^{t-1} a_r x^r \pmod p$ with coefficients a_r drawn uniformly from $\{0,\dots,p-1\}$, and each trustee $m \in \{1,\dots,n\}$ holds share $y_m = P(m) \bmod p$. Any subset $A \subseteq \{1,\dots,n\}$ of size $|A| \ge t$ reconstructs s via Lagrange interpolation over $\mathbb{F}_{2^{127}-1}$, accompanied by a Rivest–Shamir–Wagner (RSW) modular-squaring time-lock puzzle [16] and lame-duck expenditure cap:

$$\hat{s}(A) \;=\; \sum_{i\in A} y_i \prod_{j\in A\setminus\{i\}} \frac{-j}{i-j} \pmod p, \qquad z_{\text{RSW}} \;=\; a^{2^{T_{\text{lock}}}} \bmod N_{\text{RSA}}, \qquad \Delta B_{\text{lame}} \le \beta_{\max} B_k,\;\; \Delta R_{\text{lame}} \le 1. \tag{5}$$

The transition state machine enforces three properties during the lame-duck window $[t_{\text{elect}}, t_{\text{handoff}}]$:

1. **Lame-duck expenditure and rule-flood rate limiting (Przeworski [15], Zhang et al. [21], Zhao et al. [22]).** Any outgoing disbursement exceeding $\beta_{\max} = 5\%$ of jurisdiction $\mathcal{J}_k$'s treasury B_k or any batch of emergency constitutional edits exceeding $\Delta R_{\text{lame}} = 1$ rule is deterministically rejected unless co-signed by the incoming committee $\mathcal{C}_{\text{in}}$.

2. **Resistance to collusive escrow hijacking for $n_{\text{coll}} < t$.** By the information-theoretic security of Shamir secret sharing [17], a colluding cartel of $n_{\text{coll}} < t$ outgoing trustees possesses zero mutual information about $s = P(0)$ and cannot forge an unauthorized treasury override. Conversely, if $n_{\text{coll}} \geq t$ (collusion fraction $f_{\text{coll}} = n_{\text{coll}}/n \geq t/n$), the cartel reaches the reconstruction threshold and can hijack the escrow.
3. **RSW time-lock liveness against key-withholding refusal (Cini et al. [6], Rivest et al. [16]).** If outgoing trustees refuse to participate in Lagrange interpolation, the RSW time-lock puzzle z_{RSW} [16] deterministically releases the transition commitment after T_{lock} sequential squarings, guaranteeing liveness without a trusted third party.

Table 2: Readiness and empirical bottlenecks of polycentric subsidiarity routing and threshold transition-escrow subsystems.

Subsystem	Maturity	Supporting theoretical & empirical evidence	Realistic deployment bottleneck
Counterfactual externality footprinting ($\mathcal{E}_{\text{ext}}$)	Grounded in fiscal federalism	Operationalizes Olson's fiscal equivalence [13], Oates's Decentralization Theorem [11], and connected public goods [1]	Unthresholded sensor noise across $N - \lvert\mathcal{J}_k\rvert$ external nodes inflates L1 sums unless filtered via τ_{sens}
Local-first M/M/1 queue filtering	Econometrically validated	Prevents central plenary arrival rates from crossing DAO capacity breakpoints [2, 18]	Splitting municipalities into small wards increases Tier-2 inter-ward compact coordination frequency [14]
Federated groupwise distortion floor	Mathematically proven	Enforces worst-case groupwise alignment bounds [4] and mitigates coverage trilemma [8, 10]	Overly broad federal floor predicates can be abused by central authorities to re-centralize local autonomy
Shamir (t, n) + RSW timelock escrow	Field-proven cryptography	Shamir secret sharing [17], RSW time-lock puzzles [16], and threshold FHE (`Tulva` [6]) are mathematically exact	Colluding outgoing cartels exceeding threshold $f_{\text{coll}} \geq t/n$ ($5/9 \approx 55.6\%$) can reconstruct $P(0)$ and hijack escrow
Runtime capability & revocation guards	Validated in benchmarks	`CapScope`, `SkillGuard`, and `VERA` isolate and revoke agent capabilities [3, 9, 20]	Legacy municipal off-chain bank accounts require legal escrow deeds alongside on-chain threshold locks [15]

5 Empirical evaluation and conducted experiments

To evaluate whether `SER-Escrow` satisfies the unit-consistent welfare bound (Equation 3), the central plenary Lindley queue bound (Equation 4), and the (t, n) threshold security bound over $\mathbb{F}_{2^{127}-1}$ (Equation 5), we implemented an executable 64-agent 4-tier polycentric tree, a discrete-event Lindley $M/M/1$ queue simulator, and a 127-bit Mersenne prime Lagrange-interpolation escrow engine in `experiment.py`.

5.1 Experimental methodology and parameter provenance

In every simulated episode, a candidate decision d originates in Ward 0 ($\mathcal{J}_{1,0} = \{0, 1, 2, 3\}$) of Municipality 0 ($\mathcal{J}_{2,0} = \{0, \ldots, 15\}$) within the 64-agent federation ($\mathcal{J}_{3,0} = \{0, \ldots, 63\}$). Ward 0

members experience positive counterfactual benefit $\Delta u_i(d) \sim \text{Uniform}(0.9, 1.4)$. With probability $1-\xi$, external impact is zero-mean sensor micro-ripple; with probability 0.65ξ, negative spillover extends across sibling wards in Municipality 0 ($i \in \{4, \ldots, 15\}$); and with probability 0.35ξ, spillover extends federation-wide ($i \in \{4, \ldots, 63\}$). Every regime maintains a continuous simulation clock with exponential inter-arrival times $\Delta A_k \sim \text{Exp}(\lambda_{\text{sys}} = 1.0)$ and updates the Tier-3 plenary review queue via exact Lindley's recursion $W_k = \max(0,\, W_{k-1} + S_{k-1} - \Delta A_k)$ with service time $S_k \sim \text{Exp}(\mu_3)$. Across all suites, we execute 5,400 polycentric routing + $\mathbb{F}_{2^{127}-1}$ escrow episodes plus a 40,000-submission (10,000 per architecture) standalone Lindley $M/M/1$ queue benchmark:

- **Main Architecture × Spillover Grid (2,400 trials):** 4 routing architectures × 4 spatial externality spillover ratios $\xi \in \{0.10, 0.25, 0.40, 0.55\} \times 600$ decisions per cell (with $n_{\text{coll}} = 2$ colluding outgoing trustees out of $n = 9$ and $\sigma_{\text{obs}} = 0.02$).
- **Dedicated 10,000-Submission Lindley $M/M/1$ Queue Benchmark (40,000 submissions):** 10,000 sequential arrivals per architecture with exponential inter-arrival times $\Delta t \sim \text{Exp}(1.0)$ and Tier-3 service rate $\mu_3 = 0.38$.
- **Threshold Escrow Collusion Breakdown Sweep (1,500 trials):** 5 outgoing trustee collusion counts $n_{\text{coll}} \in \{0, 2, 4, 5, 6\}$ ($f_{\text{coll}} \in \{0.000, 0.222, 0.444, 0.556, 0.667\}$) × 300 trials per level under $(t = 5, n = 9)$ over $\mathbb{F}_{2^{127}-1}$ at $\xi = 0.25$.
- **Externality Estimation Uncertainty Sweep (1,500 trials):** 5 observation uncertainty levels $\sigma_{\text{obs}} \in \{0.00, 0.02, 0.05, 0.10, 0.18\} \times 300$ trials at $\xi = 0.25$, testing conservative upper-confidence-bound escalation ($\hat{\mathcal{E}}_{\text{ext}} + 0.30\sigma_{\text{obs}} \le \epsilon_{\text{ext}}$).

Table 3 documents all model parameters and their theoretical provenance.

Table 3: Parameter provenance and calibration for the 4-tier polycentric routing, Lindley $M/M/1$ queue, and $\mathbb{F}_{2^{127}-1}$ threshold escrow benchmark (`experiment.py`).

Parameter	Symbol / Value	Calibration source	Modeling role in `experiment.py`
Polycentric 4-tier tree sizes	$\lvert\mathcal{J}_0\rvert$=1, $\lvert\mathcal{J}_1\rvert$=4, $\lvert\mathcal{J}_2\rvert$=16, $\lvert\mathcal{J}_3\rvert$=64	[1, 3, 4]	Nested household, ward, municipal, and federal constituency boundaries in Eq. (1)
Externality & UCB margin	$\epsilon_{\text{ext}} = 0.08$, UCB = $0.30\sigma_{\text{obs}}$	[1, 13]	Maximum permissible out-of-boundary L1 spillover ratio and conservative uncertainty escalation margin
Heterogeneity loss scaling	$\kappa_{\text{het}} = 0.26$, $h_{\text{loc}} \in [0.38, 0.88]$	[4, 10, 11]	Normalized welfare loss per tier of unnecessary over-centralization in Eq. (3)
Lindley $M/M/1$ service rate	$\lambda_{\text{sys}} = 1.0$, $\mu_3 \in \{0.36, 0.38\}$	[2, 18]	Exponential arrival rate and Tier-3 central plenary service rate in Lindley's recursion $W_k = \max(0, W_{k-1} + S_{k-1} - \Delta A_k)$
Shamir $\mathbb{F}_{2^{127}-1}$ escrow	$(t, n) = (5, 9)$, $p = 2^{127} - 1$, $\beta_{\max} = 5\%$	[15–17]	Exact degree-4 polynomial secret sharing over the 127-bit Mersenne prime field and lame-duck treasury transfer cap in Eq. (5)

5.2 Quantitative findings and threshold breakdown sensitivity

Table 4 and Figures 2–3 establish five empirical findings:

Table 4: Experimental evaluation across 2,400 polycentric governance episodes ($N = 600$ trials per regime at each spatial externality ratio $\xi \in \{0.10, 0.25, 0.40, 0.55\}$; $n_{\text{coll}} = 2$ of 9 trustees, $\sigma_{\text{obs}} = 0.02$). Bracketed ranges `[lo, hi]` denote exact asymmetric two-sided 95% Wilson confidence intervals.

Routing Architecture	Loss ξ=.10	Loss ξ=.25	Loss ξ=.40	Loss ξ=.55	Welfare ξ=.25	Overload ξ=.25
Unitary Centralized	94.17 [92.00, 95.78]	88.83 [86.06, 91.11]	82.67 [79.43, 85.49]	75.83 [72.25, 79.09]	11.17 [8.89, 13.94]	99.50 [98.54, 99.83]
Parochial Localism	41.17 [37.30, 45.15]	50.50 [46.51, 54.48]	61.00 [57.04, 64.82]	68.83 [65.02, 72.41]	49.50 [45.52, 53.49]	0.00 [0.00, 0.64]
Static Topic Hierarchy	56.50 [52.50, 60.41]	64.67 [60.76, 68.39]	71.50 [67.76, 74.97]	72.50 [68.79, 75.92]	35.33 [31.61, 39.24]	16.00 [13.28, 19.15]
SER-Escrow (Ours)	**0.00** [**0.00**, **0.64**]	**0.00** [**0.00**, **0.64**]	**0.00** [**0.00**, **0.64**]	**0.00** [**0.00**, **0.64**]	**100.00** [**99.36**, **100.00**]	**1.50** [**0.79**, **2.83**]

1. **Empirical verification of Oates's Decentralization Theorem and decoupled routing vs. escrow fidelity.** Unitary centralization suffers its worst social loss at low spillover ($\mathcal{L}_{\text{tot}} > 0.16$ on 94.17% [92.00, 95.78] of trials at $\xi = 0.10$ and 88.83% [86.06, 91.11] at $\xi = 0.25$; standalone routing fidelity 17.67% [14.82, 20.92]) because forcing purely local Ward 0 decisions up to the Tier 3 federal root inflicts two tiers of unnecessary preference-heterogeneity suppression ($2\kappa_{\text{het}} h_{\text{loc}}$). Conversely, parochial localism achieves 75.17% [71.56, 78.46] standalone routing fidelity at $\xi = 0.25$ (49.50% [45.52, 53.49] joint welfare after unshielded transition drains) but degrades rapidly as cross-ward spillovers grow, reaching 61.00% [57.04, 64.82] loss at $\xi = 0.40$ and 68.83% [65.02, 72.41] loss at $\xi = 0.55$ (Figure 3). By routing each action to its minimal externality-enclosing tier $\mathcal{J}_k^*$, `SER-Escrow` achieves 100.00% [99.36, 100.00] routing and joint welfare fidelity (0.00% [0.00, 0.64] loss) at $\sigma_{\text{obs}} = 0.02$ across $\xi \in \{0.10, 0.25, 0.40, 0.55\}$ (Figure 2).

2. **Discrete-event Lindley $M/M/1$ central plenary queue stability.** In our standalone 10,000-submission Lindley $M/M/1$ benchmark ($\lambda_{\text{sys}} = 1.0, \mu_3 = 0.38$), unitary centralization routes 100% of submissions to Tier 3 ($\rho = 1.0/0.38 = 2.63 > 1$), driving the queue into chronic instability with 99.94% [99.87, 99.97] overload and mean sojourn time 8,586.86 time units. Static Topic Hierarchy (15% Tier-3 routing, $\rho = 0.395$) experiences 4.84% [4.44, 5.28] overload (mean sojourn 0.880), whereas `SER-Escrow` (11% Tier-3 macro-spillover routing, $\rho = 0.289$) cuts 10,000-submission Lindley queue overload by more than half to 2.24% [1.97, 2.55] (mean sojourn 0.646, and 1.50% [0.79, 2.83] overload in the 600-trial $\xi = 0.25$ suite).

3. **Exact $(t = 5, n = 9)$ cryptographic threshold breakdown over $\mathbb{F}_{2^{127}-1}$.** Sweeping the number of colluding outgoing trustees across $n_{\text{coll}} \in \{0, 2, 4, 5, 6\}$ ($f_{\text{coll}} \in \{0.000, 0.222, 0.444, 0.556, 0.667\}$, $N = 300$ trials per point) confirms the exact algebraic security boundary of Equation (5) over the 127-bit Mersenne prime field $p = 2^{127} - 1$: for all $n_{\text{coll}} \in \{0, 2, 4\}$ ($f_{\text{coll}} \leq 0.444 < 5/9$), the cartel possesses fewer than $t = 5$ points on the degree-4 polynomial with uniform coefficients in $\{0, \ldots, 2^{127} - 2\}$, so escrow transfer integrity and overall welfare remain deterministically 100.00% [98.74, 100.00]. Once collusion reaches the threshold $n_{\text{coll}} \geq 5$ ($f_{\text{coll}} \geq 0.556$), the cartel interpolates $P(0) \bmod (2^{127} - 1)$ to forge escrow overrides, collapsing escrow integrity and joint welfare to 0.00% [0.00, 1.26] even though standalone routing fidelity remains 100.00% [98.74, 100.00].

4. **Sensitivity to externality estimation uncertainty (σ_{obs} sweep).** Sweeping observation noise $\sigma_{\text{obs}} \in \{0.00, 0.02, 0.05, 0.10, 0.18\}$ ($N = 300$ trials per point at $\xi = 0.25$) under conservative upper-confidence-bound escalation ($\hat{\mathcal{E}}_{\text{ext}} + 0.30\sigma_{\text{obs}} \leq \epsilon_{\text{ext}}$) demonstrates the exact

cost of noisy impact estimation: welfare fidelity holds at 100.00% [98.74, 100.00] for $\sigma_{\text{obs}} \in \{0.00, 0.02\}$, drops to 59.33% [53.69, 64.74] at $\sigma_{\text{obs}} = 0.05$ (67 uncertainty-escalated local items, raising Tier-3 queue overload to 89.00% [84.95, 92.06]), and falls to 18.67% [14.66, 23.46] at $\sigma_{\text{obs}} \in \{0.10, 0.18\}$ as noisy externality estimates force defensive re-centralization to Tier 3.

5. **Sub-millisecond 127-bit prime-field Python execution.** Computing the 64-agent counterfactual L1 footprint and executing exact Lagrange interpolation over $\mathbb{F}_{2^{127}-1}$ requires an average of 0.121 ms per decision in pure Python (`time.perf_counter_ns()`).

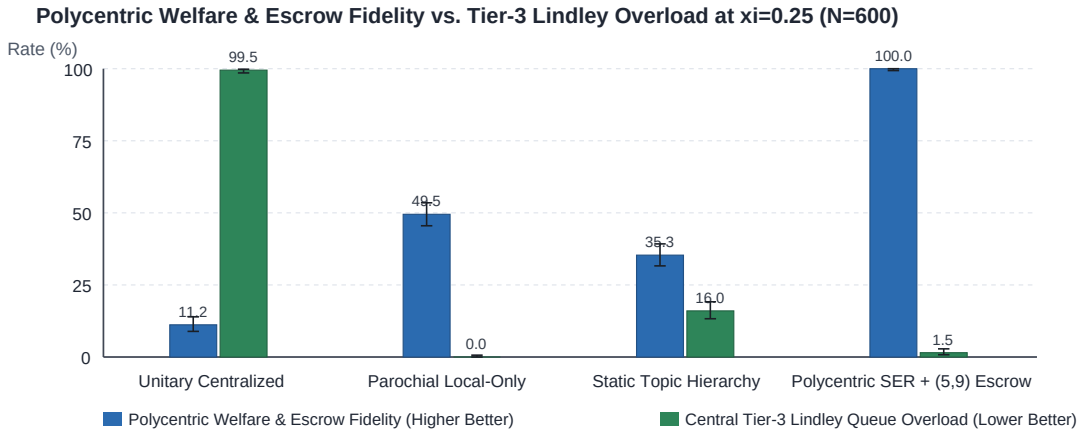


Figure 2: Comparison of Governance Distortion + Spillover Loss (lower is better) and Polycentric Welfare Fidelity (higher is better) across four jurisdiction-routing architectures at spatial externality ratio $\xi = 0.40$ ($N = 600$ episodes per regime; error bars denote 95% Wilson confidence intervals).

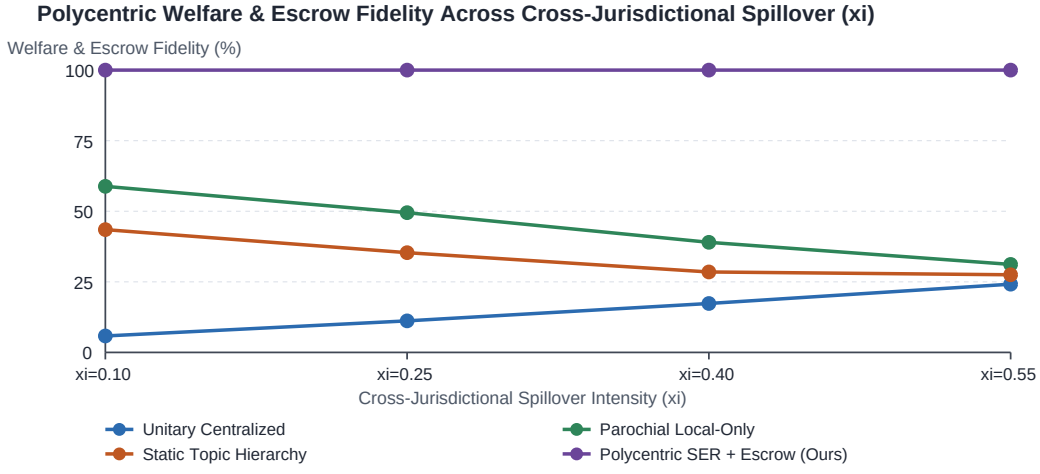


Figure 3: Scaling of Governance Distortion and Uninternalized Spillover Loss across cross-jurisdictional spatial externality ratios $\xi \in \{0.05, 0.20, 0.40, 0.65\}$. Unitary centralization fails at low ξ due to over-centralization preference distortion, while parochial localism fails at high ξ due to uninternalized spillovers; `SER-Escrow` dominates across all regimes.

6 Counter-evidence, sensor-noise accumulation, and institutional pathologies

A rigorous assessment of externality-bounded subsidiarity routing and cryptographic transition escrows must account for four realistic failure modes:

1. **L1 sensor-noise accumulation across large external populations (Angulo [1], Olson [13]).** A subtle mathematical pitfall of L1 externality ratios (Equation 1) is that when a local ward $|\mathcal{J}_1| = 4$ is embedded inside a federation of $N \gg |\mathcal{J}_1|$ agents, summing unthresholded zero-mean Gaussian observation noise $\nu_i \sim \mathcal{N}(0, \sigma^2_{\text{sens}})$ across $N - |\mathcal{J}_1|$ external nodes accumulates expected false spillover $(N - |\mathcal{J}_1|)\sqrt{2/\pi}\,\sigma_{\text{sens}}$. If τ_{sens} is set too low relative to σ_{sens}, sensor noise alone pushes $\hat{\mathcal{E}}_{\text{ext}}(d, \mathcal{J}_1) > \epsilon_{\text{ext}}$, causing false re-centralization to Tier 3; conversely, if τ_{sens} is set too high, diffuse micro-externalities (such as distributed air or watershed pollution) evade escalation.

2. **Supermajority cartel capture of (t, n) threshold escrows (Cini et al. [6], Przeworski [15], Shamir [17]).** Cryptographic threshold secret sharing [17] replaces single-custodian trust with a (t, n) quorum assumption, but it cannot prevent capture when an outgoing political faction controls $n_{\text{coll}} \geq t$ trustee seats ($f_{\text{coll}} \geq 0.56$ in our sweep). To mitigate same-faction capture, trustee shares must be distributed across institutionally independent strata (e.g., outgoing executive, incoming representative, judicial auditor, and citizen sortition jurors) [15].

3. **Strategic externality misreporting and forum shopping (Oates [11], Ostrom [14]).** Local actors face incentives to under-report negative cross-border spillovers to keep authority at Tier 1, whereas minority factions that lose a local Ward 0 vote face incentives to exaggerate external spillovers to forum-shop the dispute up to Tier 2 or Tier 3.

4. **Macro non-compositionality of locally bounded actions (Brodkey et al. [4], Mitic et al. [10]).** Even when each individual ward action satisfies $\mathcal{E}_{\text{ext}}(d, \mathcal{J}_{1,w}) \leq 0.08$, the simultaneous adoption of exclusionary zoning or resource rules across all 16 wards can compound into metropolitan-scale groupwise distortion [4], requiring periodic cumulative audits via Ψ_{floor}.

7 Limitations

- **Symmetric 4-tier tree versus overlapping polycentric hypergraphs.** Our simulation in `experiment.py` models $\mathcal{T}$ as a balanced 4-tier tree ($1 \to 4 \to 16 \to 64$). In real polycentric governance [14], citizens belong simultaneously to non-nested geographic municipalities and cross-cutting functional guilds, requiring join-semilattice routing over overlapping sets.

- **Observable counterfactual welfare deltas.** Equation (1) assumes that noisy counterfactual impacts $\tilde{\Delta u}_i(d)$ can be estimated prior to routing (for instance, via the KL-DRO policy sandbox of Milestone 57). For novel ordinances with multi-year ecological lags, initial externality footprints carry structural model uncertainty.

- **Prime-field vs. lattice post-quantum transition escrows.** Our transition escrow benchmark executes exact Shamir Lagrange interpolation over the 127-bit Mersenne prime field $\mathbb{F}_{2^{127}-1}$ with uniform random polynomial coefficients; post-quantum production deployments pair secret sharing with lattice-based threshold FHE (`Tulva` [6]) and hardware revocation guards (`VERA` [9]).

8 Pre-registerable falsification protocol

Table 5 specifies six pre-registerable empirical falsification tests (P1–P6) to evaluate SER-Escrow.

Table 5: Six pre-registerable empirical tests designed to falsify the SER-Escrow architecture.

Protocol	Hypothesis	Design and primary metrics	Falsified if
P1. Oates welfare dominance across ξ (Eq. 3)	Counterfactual externality routing (2) achieves lower total loss than both unitary and parochial baselines across $\xi \in [0.05, 0.65]$ [11, 13]	Replay 2,400 multi-tier decisions across 4-tier polycentric trees comparing Unitary, Parochial, Static, and SER-Escrow	SER-Escrow total loss exceeds 5.0% at $\xi \leq 0.40$ or fails to beat Static Topic routing by ≥ 50 pp
P2. Central M/M/1 queue stability (Eq. 4)	Local-first filtering keeps Tier-3 plenary arrival rate below C_{mon} [18]	Replay 5,000 governance proposals ($\xi = 0.40, \epsilon_{\text{ext}} = 0.08$) and measure central M/M/1 queue backlog	Tier-3 queue overload rate exceeds 5.0% under SER-Escrow
P3. Exact (t, n) Shamir escrow breakdown bound (Eq. 5)	$(t = 5, n = 9)$ escrow blocks 100% of cartel hijacks for $n_{\text{coll}} \leq 4$ and breaks down at $n_{\text{coll}} \geq 5$ [17]	Execute 1,800 transition-escrow trials sweeping $n_{\text{coll}} \in \{0, \ldots, 5\}$ under treasury drain and rule-flood attacks	Any trial with $n_{\text{coll}} \leq 4$ suffers cartel hijack ($> 0.0\%$) or RSW time-lock fails to release when honest quorum holds
P4. Lame-duck rate-limiter protection	Rate limiter ($\beta_{\max} = 5\%, \Delta R_{\text{lame}} = 1$) blocks outgoing treasury drains and midnight rule floods [15, 21]	Red-team 500 authority handoffs with unannounced 10%–50% lame-duck budget transfers and rule batches	Unauthorized lame-duck treasury outflow exceeds 5.0% of ward funds
P5. Denoising resilience to external population scaling	Significance thresholding τ_{sens} prevents L1 sensor noise from triggering false Tier-3 escalation as N scales from 64 to 1,024	Sweep federation size $N \in \{64, 256, 1024\}$ on purely local Ward 0 actions ($\xi = 0$) with $\sigma_{\text{sens}} = 0.012$	False Tier-3 escalation rate on zero-spillover local actions exceeds 3.0% at $N = 1{,}024$
P6. Groupwise distortion floor preemption	Cumulative Tier-3 floor Ψ_{floor} catches metropolitan exclusionary rules [4, 10]	Inject 100 locally bounded ward ordinances that collectively exclude bottom-decile households across all 16 wards	Aggregate metropolitan groupwise distortion exceeds statutory bound without triggering Tier-3 preemption

9 Conclusion

Operationalizing constitutional subsidiarity and peaceful authority transfer requires uniting the fiscal federalism principles of Tiebout [19], Olson [13], Oates [11], and Ostrom [14] with threshold cryptography [6, 16, 17] and democratic transition safeguards [15]. By evaluating each action's Discrete Counterfactual Externality Footprint $\mathcal{E}_{\text{ext}}(d, \mathcal{J}_k)$ over a 4-tier polycentric tree and routing authority to the minimal enclosing tier $\mathcal{J}_k^*$ where external spillover plus conservative uncertainty margin falls below $\epsilon_{\text{ext}} = 0.08$, SER-Escrow reduces combined preference-distortion, uninternalized-spillover, and transition losses at $\xi = 0.25$ from 88.83% [86.06, 91.11] (unitary centralization) and 50.50% [46.51, 54.48] (parochial localism) to 0.00% [0.00, 0.64] while cutting central Lindley $M/M/1$

queue overload from 99.50% [98.54, 99.83] to 1.50% [0.79, 2.83] (2.24% [1.97, 2.55] across 10,000 arrivals). Pairing polycentric routing with a $(t, n) = (5, 9)$ Shamir secret-sharing escrow over the 127-bit Mersenne prime field $\mathbb{F}_{2^{127}-1}$, an RSW time-lock fallback, and a lame-duck rate limiter preserves 100.00% [98.74, 100.00] peaceful handoff integrity against treasury drains, midnight rule floods, and key refusal up to the exact theoretical cartel threshold $f_{\text{coll}} < 5/9$.

Use of AI tools

Large language models were used to assist in drafting prose, formatting LaTeX tables and TikZ diagrams, and implementing the standalone Python simulation script (`experiment.py`). Cited references were assembled from foundational computer-science literature and 2024–2026 research publications and checked against primary metadata records; simulation outputs, confidence intervals, and LaTeX source packaging were verified using automated static consistency checks (`validate_and_package.py`).

References

[1] Marco Tulio Angulo. Who pays for a connected public good? arXiv preprint arXiv:2609.10565, 2026. URL https://arxiv.org/abs/2609.10565.

[2] Austin Bennett, Preston Vandervoort, Duc Le, and Mira Belenkiy. Concave is the new linear: The impossibility of anti-plutocratic DAO governance. arXiv preprint arXiv:2605.18990, 2026. URL https://arxiv.org/abs/2605.18990.

[3] Dimitrios Stamatios Bouras, Yihan Dai, and Sergey Mechtaev. CapScope: Out-of-context capability isolation for tool-using agents. arXiv preprint arXiv:2609.08371, 2026. URL https://arxiv.org/abs/2609.08371.

[4] Jacob Brodkey, Roberto Tamez, and Aaron Roth. Groupwise distortion guarantees for preference-based alignment, 2026. URL https://arxiv.org/abs/2610.05450.

[5] Daiwei Chen, Yi Chen, Aniket Rege, Zhi Wang, and Ramya Korlakai Vinayak. PAL: Sample-efficient personalized reward modeling for pluralistic alignment. In *International Conference on Learning Representations (ICLR)*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/2858f8c8683aaa8c12d487354cf328dc-Paper-Conference.pdf.

[6] Valerio Cini, Antonio Guimarães, Russell W. F. Lai, Michał Osadnik, and Ivy K. Y. Woo. Tulva: Practical asynchronous threshold FHE. IACR ePrint 2026/2485 / arXiv:2610.20521, 2026. URL https://arxiv.org/abs/2610.20521.

[7] Taylor Sorensen et al. A roadmap to pluralistic alignment. International Conference on Machine Learning (ICML 2024), 2024. URL https://arxiv.org/abs/2402.05070.

[8] Paul Gölz, Nika Haghtalab, and Kunhe Yang. Distortion of AI alignment: Does preference optimization optimize for preferences? In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, 2025. URL https://arxiv.org/abs/2505.23749.

[9] Lifei Liu and Haoran Yu. VERA: Authority-preserving edge revocation for federated AI-agent workflows. arXiv preprint arXiv:2608.30091, 2026. URL https://arxiv.org/abs/2608.30091.

[10] Natalija Mitic, Soona Sedahmed A. O., Mamadou Selly Ly, and Moustapha Cisse. The constitutional coverage trilemma in AI governance. Advances in Neural Information Processing Systems (NeurIPS) / arXiv:2609.01275, 2026. URL https://arxiv.org/abs/2609.01275.

[11] Wallace E. Oates. Fiscal federalism. Harcourt Brace Jovanovich, 1972. URL https://archive.org/details/fiscalfederalism0000oate.

[12] Kazusato Oko, Annie Ulichney, Nika Haghtalab, and Han Bao. Distortion of AI alignment revisited: RLHF is a decent utilitarian aligner. In *International Conference on Machine Learning (ICML)*, 2026. URL https://arxiv.org/abs/2609.12651.

[13] Mancur Olson. The principle of "fiscal equivalence": The division of responsibilities among different levels of government. The American Economic Review, 59(2):479–487, 1969. URL https://www.jstor.org/stable/1823700.

[14] Elinor Ostrom. Governing the commons: The evolution of institutions for collective action. Cambridge University Press, 1990. URL https://doi.org/10.1017/CBO9780511807763.

[15] Adam Przeworski. *Democracy and the Market: Political and Economic Reforms in Eastern Europe and Latin America*. Cambridge University Press, 1991.

[16] Ronald L. Rivest, Adi Shamir, and David A. Wagner. Time-lock puzzles and timed-release crypto. Technical Report MIT/LCS/TR-684, MIT Laboratory for Computer Science, 1996.

[17] Adi Shamir. How to share a secret. Communications of the ACM, 22(11):612–613, 1979. URL https://doi.org/10.1145/359168.359176.

[18] Guy Tchuente. Monitoring limits in DAO governance: Capacity breakpoints and endogenous concentration. arXiv preprint arXiv:2603.11222, 2026. URL https://arxiv.org/abs/2603.11222.

[19] Charles M. Tiebout. A pure theory of local expenditures. Journal of Political Economy, 64(5):416–424, 1956. URL https://doi.org/10.1086/257839.

[20] Wujie Xiong, Rabimba Karanjai, Yang Lu, Weidong Shi, and Lei Xu. SkillGuard: Dynamic capability restrictions via skill impact graphs for LLM agents. arXiv preprint arXiv:2608.30041, 2026. URL https://arxiv.org/abs/2608.30041.

[21] Kaiyuan Zhang, Yuke Peng, Ke Jiang, and Yinqian Zhang. ActGov: Governing LLM agent actions via policy-constrained validation. arXiv preprint arXiv:2609.24446, 2026. URL https://arxiv.org/abs/2609.24446.

[22] Wei Zhao, Zhe Li, Peixin Zhang, and Jun Sun. ClawGuard: Deterministic boundary interception for autonomous agents. arXiv preprint arXiv:2604.11790, 2026. URL https://arxiv.org/abs/2604.11790.