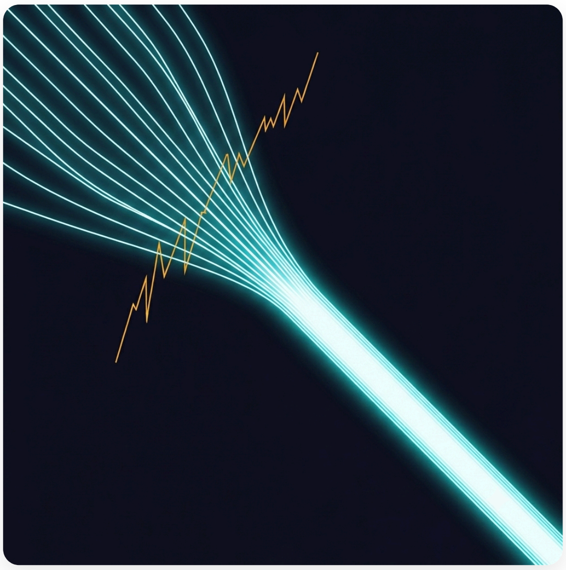

The Political Economy of Requisite Governance

Transition Pathways Under Incumbent Selection Pressure

Paper IX in the Governance as Engineering series

Models the transition to requisite governance as a contested control problem between reform coalitions and incumbent controllers. Introduces the transition variety ratio Ω, formalizes three structural traps, and derives design principles for transition mechanisms including buy-out protocols, sunset-coupled bypasses, and protected experimental spaces. Includes historical calibration and three simulations.

Björn Kenneth Holmström

June 2026

https://bjornkennethholmstrom.org/working-papers/political-economy-of-requisite-governance

1. From Qualitative Diagnosis to Formal Model

The Governance as Engineering series has established that stable, adaptive governance architectures must satisfy a set of structural constraints: short response latencies matched to disturbance timescales (Papers I–II), representation chains shallow enough to preserve the citizen preference signal (Paper III), observation channels with sufficient dimensionality to monitor the resource systems they govern (Paper IV), and value architectures capable of evolving as the dimensionality of the environment expands (Paper VI). Paper V demonstrated that these constraints interact multiplicatively: a system exhibiting multiple architectural failures does not merely add their costs but compounds them, and the coordination failure tax reduces effective governance capacity to a small fraction of the baseline.

Those results concern the *steady-state* properties of governance architectures. They answer the question: if an architecture of a given structure were in place, how would it perform under disturbance? They do not, in themselves, answer the prior question: how does an architecture of that structure come to be in place? The series has not yet modelled the process by which one governance architecture is replaced by another — the transition pathway — with the same formal rigour it has applied to the architectures themselves.

This paper provides that model. It treats the transition itself as a control problem, with the existing governance architecture as the system under control, the reform coalition as the controller, and the adaptive resistance of incumbent beneficiaries as the disturbance. The central claim is that the transition pathway possesses its own variety constraint, its own latency structure, and its own stability ceiling — and that failure to satisfy these constraints produces the characteristic reform disappointments that the series' country studies have documented at length. The paper formalises three structural traps (the bypass trap, the legibility problem, and the incentive-compatibility trap) that Paper VII identified qualitatively, derives design principles for transition mechanisms that can survive incumbent counter-mobilisation, and introduces a dynamic model of *transition bandwidth* — the rate at which a governance system can peacefully redesign itself — as a determinant of whether architectural change will be managed or forced.

1.1 The Transition Problem Is Already in the Series

A reader of the series might reasonably ask whether this paper is necessary. Paper VII — *The Fragments: A Synthesis* — is substantially a transition paper. Drawing on fifteen country studies and the four formal engineering papers, it develops the bypass trap, the legibility problem, the immune-system-as-output, the strategy of protected experimental spaces, and the mechanism of scaling by attraction. It argues that reforms which change parameters without changing architecture are systematically absorbed by the very structures they attempt to improve, and that the first viable step in every jurisdiction studied is the same: create a space where the observation channel is shorter, the signal is less degraded, and the results are visible enough to shift the broader system's model of its own dysfunction.

That diagnosis is accurate, and this paper does not contest it. What Paper VII does *qualitatively*, this paper does *formally*. Paper VII identifies *that* the bypass trap exists; this paper models it as a dynamical system with explicit state variables and demonstrates the stable low-performance attractor that makes permanent bypasses self-defeating. Paper VII observes *that* incumbent beneficiaries resist architectural change; this paper models that resistance as an adaptive controller with its own observation channels, latency advantages, and objective function, and derives the conditions under which it can be overcome. Paper VII describes *that* reforms often fail because the reform coalition cannot perceive the full extent of the architecture's dysfunction; this paper formalises that as a constraint on the coalition's observation matrix and shows how it produces systematically miscalibrated interventions.

The move mirrors the one made by Paper VIII, which took the variety gap — a diagnostic concept introduced in Paper VI — and developed a parametric framework for estimating it from observable governance characteristics. In both cases, the contribution is not a new diagnosis but a formal apparatus that makes the diagnosis *operational*: testable, quantifiable, and capable of generating predictions that the qualitative account alone cannot.

1.2 The Gap That Remains

Despite Paper VII's substantial treatment of transition dynamics, three analytical gaps persist.

First, the strategic interaction between reform coalitions and incumbent beneficiaries is described but not modelled. Paper VII characterises the *Centrão* in Brazil, the extraction coalition in Nigeria, and the Control Preservation Imperative in China as "immune systems" that are outputs of the current architecture. It argues, correctly, that these immune systems are not external obstacles but predictable adaptive responses. What it does not provide is a formal representation of that interaction — a state-space model in which two controllers with asymmetric latency and observation capacity act on the same institutional state vector, each optimising a different objective function. Without such a model, it is difficult to specify the conditions under which reform succeeds, is absorbed, or triggers oscillation, except through historical generalisation.

Second, Paper VII identifies the incentive-compatibility problem — the fact that architectural reform often requires the cooperation of actors who benefit from the current arrangement — but offers no design principles for addressing it beyond the protected experimental space. The historical record, however, contains a richer repertoire of transition mechanisms: explicit compensation of incumbent losers, sunset-coupled bypasses that transfer pressure back to the unreformed substrate, and the instrumentalisation of independent observation channels to undermine the incumbent's control over the narrative. These mechanisms have structural properties that can be formalised within the series' existing grammar. The paper develops them in Part IV.

Third, Paper VII is static in its treatment of time. It describes the current architecture, the failure modes it generates, and a proposed direction of movement, but it does not model the *rate* at which architectural change must occur relative to the rate at which the environment generates new disturbance dimensions. As Paper VI established, the effective dimensionality of reality expands over time. If the rate of architectural adaptation falls below the rate of environmental change, the variety gap widens and forced dissolution — collapse, revolution, constitutional rupture — becomes the only remaining pathway. The concept of *transition bandwidth* introduced in Part V provides a dynamic framework for analysing this race, and Simulation C demonstrates a finding that the qualitative account cannot supply: the existence of a transition-bandwidth trap — a point of no return, reached before the variety gap becomes fatal, at which the system still functions operationally but has irreversibly lost the capacity to redesign itself.

1.3 Incumbent Resistance as an Adaptive Controller

The central modelling innovation of this paper is the treatment of incumbent resistance not as a disturbance in the classical sense — exogenous, stochastic, or fixed in its statistical properties — but as the output of an *adaptive controller* with its own observation channels, latency, and objective function.

This is not a normative claim about the malign intentions of incumbents, although intentions may vary. It is a structural claim about the behaviour of any agent whose survival or prosperity is coupled to the current architecture and who possesses the capacity to perceive and respond to threats to that architecture. The *Centrão* in Brazil, the extraction coalition in Nigeria, the Chinese Communist Party's control preservation apparatus, the Veto Industrial Complex in the United States, the Iron Triangle in Japan — these are not equivalent in scale, ideology, or method. But they share a structural property: each constitutes a controller that observes reform threats through its own sensing networks, processes information with shorter latency than any external reform coalition can achieve, and deploys counter-measures — legislative blocking, narrative capture, co-optation, deliberate degradation of the reform signal — calibrated to the specific architecture it defends.

Modelling this interaction requires a departure from the classical Ashby formulation that has anchored the series so far. Ashby's Law states that a regulator can only stabilise a system if its variety matches or exceeds the variety of the disturbances it faces. But the classical formulation assumes an exogenous disturbance with fixed variety. An incumbent coalition is not exogenous; it is part of the system being regulated, and its variety is not fixed — it generates counter-moves adaptively in response to the reform's actions. The appropriate formal analogue is not classical regulation but *contested control*, in which two controllers with asymmetric capabilities act on the same state vector. The paper does not attempt a full game-theoretic treatment, which would require apparatus beyond the control-theoretic idiom the series has maintained. Instead, it develops a heuristic extension — the *transition variety ratio* Ω — and subjects it to simulation testing (Simulation B) to determine under what conditions it behaves as a threshold and under what conditions the boundary is too mushy to support a clean inequality.

The latency asymmetry between reform coalitions and incumbent controllers is equally consequential. Paper I established that any feedback controller faces a gain ceiling determined by its response latency: $K_{max} \approx 1/(\tau \cdot |A|)$. Pushing gain beyond this ceiling does not produce faster correction; it produces oscillation. Applied to the transition context, this means that a reform coalition attempting to force architectural change at a rate exceeding approximately $1/\tau_R$ — where τ_R is the delay between a policy window opening and institutional reconfiguration taking effect — will trigger counter-mobilisation, regime retrenchment, or constitutional crisis. The incumbent, with its structurally shorter latency τ_I, operates well within its ceiling while the reform coalition is pushed to the edge of its own. The asymmetry is structural, not contingent: incumbents are embedded in the architecture they defend, and embedding provides an observation-to-actuation loop that no external coalition can replicate.

With these conceptual foundations in place, Part II develops the formal model of the transition as a contested control problem, defining the state space, the two controllers, the transition variety ratio, and the gain-latency constraints. Part III models the three structural traps that Paper VII identified qualitatively — the bypass trap, the legibility problem, and the incentive-compatibility trap — as dynamical systems with explicit equilibria and escape conditions. Part IV derives design principles for transition mechanisms that satisfy the variety and latency constraints, including the conditions under which incumbent buy-outs are incentive-compatible and the boundary at which they fail. Part V introduces transition bandwidth as a dynamic concept and models the race between environmental change and architectural adaptation. Part VI presents the simulation architecture that grounds the formal claims, and Part VII concludes with the limitations and open questions that define the research frontier.

2. Modeling the Transition as a Contested Control Problem

The standard control diagram underlying Papers I through VI places a single controller between the observation of a system's state and the actuation of interventions intended to move that state toward a target. Disturbances enter from outside the loop; they are exogenous, stochastic, or structurally fixed. The controller's design problem is to select a feedback law that stabilises the system despite those disturbances, given the limitations of its observation channel and its actuator dynamics.

The transition from one governance architecture to another breaks this diagram. The "system" is now the existing governance architecture itself, and the "controller" is a reform coalition attempting to move it toward a target architecture. The disturbance, however, is not exogenous. It is the adaptive resistance of the actors who benefit from the current architecture and who possess their own observation channels, their own decision processes, and their own capacity to actuate counter-measures. The transition is therefore not a regulation problem but a *contested control* problem: two controllers with asymmetric capabilities act on the same institutional state vector, each pursuing a different objective function.

This part develops the formal representation of that interaction. Section 2.1 defines the state space, the two controllers, and the goal set. Section 2.2 introduces the transition variety ratio Ω as a heuristic extension of Ashby's Law to the adversarial setting, with explicit caveats about its status as a diagnostic rather than a theorem. Section 2.3 analyses the latency asymmetry that structurally favours the incumbent controller and the gain ceiling that constrains the reform coalition's speed of action.

2.1 System, Controllers, and Disturbance Re-defined

Consider a governance architecture in operation at time t. Its institutional state can be represented by a vector $\mathbf{X}(t)$ whose components capture the parameters that determine its performance: the latency of its observation channels, the aggregation ratios of its representation layers, the dimensionality of its value function, the spatial distribution of its decision authority, and the coupling strengths between its component nodes. This is not a complete description of the architecture — no finite vector can be — but it captures the dimensions along which architectural variation produces measurable differences in stability, responsiveness, and observability, as established in Papers I through VI.

The architecture generates a distribution of benefits across the actors who operate within it. Some actors receive concentrated rents from the current parameter settings: legislators whose influence derives from the length of the representation chain, bureaucracies whose budgets are tied to the aggregation machinery they administer, firms whose profitability depends on the externalisation of costs that the architecture's low-dimensional value function does not measure. These actors are not passive. They observe threats to the parameter settings that generate their benefits, and they possess the institutional levers to resist changes to those settings. Collectively, they constitute what we will call the *incumbent controller*, denoted **I**.

Opposing **I** is a *reform coalition*, denoted **R** — a heterogeneous set of actors who perceive the architecture's current performance as inadequate and who seek to move $\mathbf{X}(t)$ toward a target architecture $\mathbf{X}^*$ that satisfies more of the structural constraints established in the series. **R** may include citizen movements, reformist factions within the state, external partners with conditional resources, and actors whose interests are aligned with a higher-dimensional, lower-latency architecture but who currently lack the institutional authority to implement it.

The two controllers act on the same state vector $\mathbf{X}(t)$. The reform coalition applies a control signal $\mathbf{u}_R(t)$ intended to shift the architecture's parameters in the direction of $\mathbf{X}^*$. The incumbent controller applies a control signal $\mathbf{u}_I(t)$ intended to preserve the parameters that generate its benefits, or to divert the reform's energy toward parameter changes that appear substantive while leaving the underlying architecture intact — the absorption pattern documented across the country studies.

The state evolves according to:

$$\mathbf{X}(t+1) = \mathbf{A} \cdot \mathbf{X}(t) + \mathbf{B}_R \cdot \mathbf{u}_R(t - \tau_R) + \mathbf{B}_I \cdot \mathbf{u}_I(t - \tau_I) + \mathbf{d}(t)$$

where $\mathbf{A}$ captures the architecture's endogenous dynamics (institutional inertia, the natural drift of parameters under operational pressures), $\mathbf{B}_R$ and $\mathbf{B}_I$ are the actuation matrices of the reform and incumbent controllers respectively, τ_R and τ_I are their respective response latencies, and $\mathbf{d}(t)$ represents exogenous disturbances — crises, external shocks, technological discontinuities — that are not under the strategic control of either party but that create the policy windows within which architectural change becomes politically possible.

The goal set **G** is the set of architectural states that the reform coalition would accept as a successful transition. It is typically not a single point $\mathbf{X}^*$ but a region in the state space: an architecture sufficiently close to the target in the dimensions that matter most, with tolerable deviations in dimensions where compromise is necessary. The dimensionality of **G**, denoted $\dim(\mathbf{G})$, represents the slack available to the reform coalition — the number of independent architectural dimensions along which it can accept an outcome short of the ideal while still considering the transition a success.

The reform coalition's problem is to design a control law that moves $\mathbf{X}(t)$ into **G** and holds it there, given its own observation limitations, its actuation latency τ_R, and the counter-action of the incumbent controller **I**. This is the problem to which the remainder of the paper is addressed.

2.2 The Transition Variety Condition — A Heuristic Extension

In Papers I through VI, the central regulative principle has been Ashby's Law of Requisite Variety: a controller can only stabilise a system if its internal variety — the number of distinct states it can discriminate and respond to — matches or exceeds the variety of the disturbances the system faces. Formally, for a disturbance space **D** and a goal set **G**, the controller's variety $V(\mathbf{R})$ must satisfy $V(\mathbf{R}) \geq V(\mathbf{D}) - V(\mathbf{G})$. The law is a theorem under specified conditions: the disturbance is exogenous, its variety is fixed, and the controller's response does not alter the disturbance distribution.

The transition setting violates each of these conditions. The "disturbance" is the incumbent's resistance **u**_I(t), which is not exogenous but endogenous to the system — it is generated by an intelligent actor who observes the reform's actions and adapts. Its variety is not fixed but expands in response to the reform's own variety: the incumbent develops new counter-measures as the reform develops new strategies. And the reform's actions alter the incumbent's behaviour, which in turn alters the reform's — the interaction is strategic, not merely regulatory.

This places the problem in the domain of *adversarial control*, pursuit-evasion dynamics, or differential game theory — formalisms that lie beyond the control-theoretic idiom the series has maintained. A full treatment would require apparatus that this paper does not attempt to supply. What follows is therefore a *heuristic extension*: an inequality that captures the structural intuition of the series while remaining testable in simulation and falsifiable by the same. It is offered as a diagnostic scaffold, not as a theorem, and its status is explicitly flagged as provisional.

Define the effective variety of the reform coalition, dim(**R**), as the number of independent dimensions along which it can observe the architecture's state, deliberate about responses, and actuate interventions. This is not simply the number of members in the coalition or the size of its budget; it is the rank of its combined observation-actuation space — the number of orthogonal directions in **X**-space along which it can apply distinguishable control signals.

Define the effective variety of the incumbent controller, dim(**I**), analogously: the number of independent dimensions along which it can perceive threats, mobilise resources, and deploy counter-measures. In general, dim(**I**) will be high. The incumbent is embedded in the architecture it defends, giving it short-latency access to the full state vector **X**(t) and a dense network of institutional levers — committee seats, budgetary vetoes, media relationships, regulatory capture mechanisms — that provide high-dimensional actuation capacity.

The *transition variety ratio* is then:

Ω = dim(R) / dim(I)

When Ω `<` 1, the incumbent can generate more independent counter-moves than the reform coalition can address. The reform's actuation space is, in Ashby's language, of insufficient variety to absorb the variety of incumbent resistance. Reforms are absorbed, deflected, or co-opted — the breakthrough-capture pattern of the Brazil study, the extraction-association-adaptation cycle of the Nigeria study, the escalation-block-bypass-delegitimisation spiral of the United States study.

When $\Omega \geq 1$, the reform coalition possesses sufficient variety to match the incumbent's counter-moves in principle. Whether it can do so in practice depends on additional constraints — latency, gain, and the coalition's own internal coordination costs — that the ratio alone does not capture. $\Omega \geq 1$ is therefore a necessary condition for architectural reform to succeed against an adaptive incumbent; it is not sufficient. Sufficiency requires that the reform coalition also satisfy the latency and gain constraints discussed in Section 2.3, and that the transition pathway avoid the structural traps modelled in Part III.

Two further caveats are essential. First, the dimensionalities $\dim(\mathbf{R})$ and $\dim(\mathbf{I})$ are not directly observable in any straightforward way. They must be estimated from proxies — the diversity of the reform coalition's institutional bases, the independence of its information channels, the number of distinct veto points the incumbent controls — using the estimation methodology developed in Paper VIII. The values assigned in any particular case are therefore uncertain, and the ratio should be reported with confidence intervals, not as a point estimate. Second, the ratio is dynamic. As the reform coalition learns and adapts, $\dim(\mathbf{R})$ may increase; as the incumbent perceives a genuine threat, $\dim(\mathbf{I})$ may also increase, as dormant veto capacities are activated and new counter-mobilisation strategies are developed. The transition is a race between two learning systems, and $\Omega(t)$ is a moving target.

The appropriate formal ancestor for this extension is not Ashby's original theorem but Stafford Beer's concept of *variety engineering* — the design of amplifiers and attenuators that allow a controller to manage variety that exceeds its own unaided capacity. Beer recognised that in organisational settings, the "disturbance" is often another intelligent system, and that variety must be managed through structural devices — filters, amplifiers, transducers — that alter the effective variety ratio. The design principles developed in Part IV — protected experimental spaces, sunset-coupled bypasses, buy-out protocols, observer diversity — are, in this sense, variety-engineering devices: mechanisms for amplifying the reform coalition's effective variety or attenuating the incumbent's, shifting Ω in the reform's favour.

Simulation B, described in Part VI, subjects the Ω condition to testing. By varying the ratio of the two controllers' observation dimensionalities and the ratio of their latencies across a wide parameter sweep, it maps the phase diagram of transition outcomes — success, absorption, oscillation — and determines whether the $\Omega = 1$ boundary behaves as a threshold, a gradient, or a region of indeterminacy. The results of that sweep will determine the final formulation of the condition in the paper's concluding revision.

2.3 Latency Asymmetry and the Transition Gain Ceiling

The variety ratio Ω captures the *dimensional* capacity of the two controllers. It does not capture their relative *speed*. In control theory, a controller's ability to stabilise a system depends not only on its variety but on its response latency τ — the dead time between a deviation being observed and a corrective signal taking effect. Paper I established that any feedback controller faces a gain ceiling of approximately $K_{max} \approx 1/(\tau \cdot |A|)$, where $|A|$ is a measure of the system's internal dynamics. Exceeding this ceiling does not produce faster correction; it produces oscillation, as the controller's interventions arrive out of phase with the system's own evolution and amplify the deviation they were meant to suppress.

Applied to the transition context, this principle has a direct and uncomfortable implication: the reform coalition is structurally constrained in how aggressively it can pursue architectural change, and the constraint is tighter for the reform coalition than for the incumbent.

Let τ_R be the reform coalition's effective latency: the time between a policy window opening — a crisis, an election, a scandal that temporarily weakens the incumbent's grip — and the institutional reconfiguration that the reform is able to achieve. In a modern national governance system, τ_R is measured in months to years. Legislation must be drafted, coalitions assembled, parliamentary procedures navigated, regulations promulgated, implementing agencies instructed, and compliance verified. Each step adds dead time.

Let τ_I be the incumbent's effective latency: the time between a reform threat becoming visible and the incumbent deploying a counter-measure. Because the incumbent is embedded in the architecture it defends, τ_I is typically much shorter. A legislative committee chair can block a bill before it reaches the floor. A regulatory capture network can slow-walk implementation before the reform's effects materialise. A media ecosystem can reframe the reform as dangerous or illegitimate before public opinion solidifies behind it. The incumbent's observation-to-actuation loop runs on weeks to months, not months to years.

The latency ratio τ_R/τ_I is therefore substantially greater than one. And because the gain ceiling scales inversely with latency, the maximum rate at which the reform coalition can safely push architectural change — its *transition gain ceiling* — is lower than the incumbent's maximum rate of counter-mobilisation. The reform coalition must operate within a narrow band: it must push hard enough to overcome institutional inertia and reach **G** before the policy window closes, but not so hard that it triggers the oscillation predicted by the gain ceiling — counter-mobilisation, regime retrenchment, or constitutional crisis.

This constraint explains a pattern that the country studies document but do not formally explain: why architectural reforms that are pushed through at maximum speed during narrow windows so often trigger a backlash that reverses them. The reform coalition, perceiving the window closing, increases its gain beyond the ceiling. The incumbent, with its shorter latency, deploys counter-measures that arrive while the reform is still being implemented. The result is not stable architectural change but an oscillatory sequence — reform, backlash, partial reversal, renewed crisis, reform attempt — that dissipates political capital without producing durable shifts in $\mathbf{X}(t)$. The French reform-explosion-retreat cycle, the Chinese campaign-overshoot-abrupt-correction pattern, and the American escalate-block-bypass-delegitimise spiral are all, in this formal sense, manifestations of a reform coalition pushing past its gain ceiling.

The functional form $K_{max} \approx 1/(\tau_R \cdot |A|)$ is imported from Paper I without re-derivation here. The institutional dynamics matrix **A** captures the architecture's resistance to change — the inertia of bureaucratic routines, the stickiness of statutory frameworks, the self-stabilising properties of the distribution of benefits. When $|A|$ is large — when the architecture is rigid and its beneficiaries are entrenched — the gain ceiling is lower, and even moderate reform efforts can trigger oscillation. When |A| is small — after a crisis has destabilised the old equilibrium — the ceiling rises, and more aggressive reform becomes feasible without triggering backlash. This is the formal analogue of the "policy window" concept in political science: a crisis temporarily reduces |A|, widening the band of safe reform speeds.

The combined effect of the variety constraint and the latency constraint defines the feasible envelope for architectural reform. The reform coalition must satisfy $\Omega \geq 1$ — sufficient variety to match the incumbent's counter-moves — *and* operate within the gain band τ_R imposes — aggressive enough to reach **G** before the window closes, restrained enough to avoid oscillation. If either condition is violated, the transition fails. If both are satisfied, the transition can, in principle, succeed.

Part III now examines the three structural traps that can cause a transition to fail even when the variety and latency conditions are initially met: the bypass trap, the legibility problem, and the incentive-compatibility trap. Each is modelled as a dynamical system with its own internal logic, and each corresponds to a failure mode that Paper VII identified qualitatively and that the country studies document empirically.

3. The Three Structural Traps

The transition variety ratio Ω and the latency-gain ceiling define the outer boundaries of the feasible reform envelope: a reform coalition must possess sufficient variety to match incumbent resistance and must operate at a speed that neither stalls nor triggers oscillation. But satisfying these aggregate conditions does not guarantee success. Within the envelope, the transition pathway can be derailed by three structural traps that are generated by the very architecture the reform is attempting to change. Each trap is a dynamical system with its own internal logic, its own equilibrium, and its own characteristic failure signature. Paper VII identified each of them qualitatively, drawing on the fifteen country studies and the engineering papers. This part formalises them: the bypass trap, the legibility problem, and the incentive-compatibility trap.

3.1 The Bypass Trap

The bypass trap is the most empirically recurrent failure mode in the series. It arises when a reform coalition, unable to change the core architecture directly, builds a parallel institutional channel that routes around a dysfunctional element. The bypass delivers real improvements — faster payments, more transparent audits, more responsive local services — and its success generates political support for its continuation. But precisely because it relieves the pressure that dysfunction created, it reduces the urgency of reforming the original architecture. Over time, the bypass becomes a permanent fixture, and the unreformed substrate persists, capping the bypass's own effectiveness. The result is a stable low-performance attractor: the system is demonstrably better than it was before the bypass, but demonstrably worse than it would be if the underlying architecture were repaired. And the existence of the bypass makes the case for repair *harder* to prosecute politically, because the most acute sufferers of the dysfunction have been partially relieved.

Paper VII documented this pattern across multiple jurisdictions. India's Unified Payments Interface processes ten billion transactions a month, a world-class transactional layer built above a legal and administrative skeleton in which land disputes take eleven years to resolve. The bypass routes around the broken formal institutions so effectively that the broken institutions face no pressure to reform themselves — and the bypass's own potential is capped by the unreformed substrate's failure to resolve the foundational disputes that the bypass cannot adjudicate. Brazil's proposed Algorithmic Bypass would convert opaque political allocation into partially self-enforcing delivery contracts, but it operates within a banking oligopoly that charges 300 percent interest — the bypass addresses the allocation problem while leaving the concentration problem intact.

The dynamical structure of the bypass trap can be represented with two state variables:

- **D**(t): the dysfunction of the unreformed substrate — the severity of the architectural defect that the bypass routes around, ranging from 0 (fully functional) to 1 (complete failure).

- $\mathbf{B}(t)$: the capacity of the bypass to deliver outcomes, ranging from 0 (nonexistent) to some maximum $\mathbf{B}_{max}$ that is a function of **D** — the substrate caps the bypass's effectiveness, because some essential functions (adjudication, enforcement, sovereign guarantee) cannot be fully replicated without architectural authority.

The bypass reduces *visible* dysfunction — the signal that creates political pressure for reform — without reducing actual dysfunction. Let $\mathbf{D}_{vis}(t) = \mathbf{D}(t) \cdot f(\mathbf{B})$, where $f(\mathbf{B})$ is a decreasing function: as bypass capacity increases, the visible consequences of the dysfunction decline, because the most urgent needs are being met through the parallel channel. Reform pressure on the substrate is proportional to $\mathbf{D}_{vis}(t)$. As **B** rises, pressure falls, and the substrate stabilises at its current level of dysfunction.

The dynamics are:

$$\mathbf{D}(t+1) = \mathbf{D}(t) + \alpha \cdot (1 - \mathbf{D}_{vis}(t)) - \beta \cdot \mathbf{R}(t)$$

$$\mathbf{B}(t+1) = \mathbf{B}(t) + \gamma \cdot \mathbf{R}_B(t) \cdot (1 - \mathbf{B}(t)/\mathbf{B}_{max}(\mathbf{D}(t)))$$

where $\mathbf{R}(t)$ is reform effort directed at the substrate, $\mathbf{R}_B(t)$ is effort directed at expanding the bypass, and $\mathbf{B}_{max}(\mathbf{D})$ is a decreasing function: the more dysfunctional the substrate, the lower the ceiling on bypass effectiveness. The trap is the equilibrium at which **B** is high enough to suppress $\mathbf{D}_{vis}$ to politically tolerable levels, but **D** remains high enough that $\mathbf{B}_{max}(\mathbf{D})$ is well below the performance of a genuinely reformed architecture. The system is stuck: the bypass is too successful to abandon but too constrained to solve the underlying problem.

Simulation A (Part VI) demonstrates this equilibrium and its dependence on the coupling between $\mathbf{B}_{max}$ and **D**. The critical escape condition is a *sunset coupling*: a mechanism that transfers pressure back to the substrate as the bypass's own performance improves, so that **B** and **D** are not independent but linked through a commitment that the bypass will either force reform of the substrate or wind down. That mechanism is developed in Section 4.2.

3.2 The Legibility Problem

The legibility problem, introduced in Paper VII and extended here, is the consequence of a simple and devastating fact: the reform coalition must diagnose the architecture's dysfunction using the architecture's own observation channels.

Let the true architectural state be $\mathbf{X}(t)$, as defined in Section 2.1. The reform coalition does not observe $\mathbf{X}(t)$ directly. It observes a signal:

$$\mathbf{Y}_R(t) = \mathbf{C}_R \cdot \mathbf{X}(t) + \boldsymbol{\varepsilon}_R(t) + \boldsymbol{\eta}_I(t)$$

where $\mathbf{C}_R$ is the reform coalition's observation matrix, $\boldsymbol{\varepsilon}_R$ is noise from the coalition's own measurement limitations, and $\boldsymbol{\eta}_I$ is a distortion term injected by the incumbent controller — deliberate degradation of the signal through narrative capture, selective release of information, or the strategic creation of confounding events.

The incumbent controller, by contrast, observes the state with higher fidelity:

$$\mathbf{Y}_I(t) = \mathbf{C}_I \cdot \mathbf{X}(t) + \boldsymbol{\varepsilon}_I(t)$$

where $\mathbf{C}_I$ is typically of higher rank than $\mathbf{C}_R$ — the incumbent has access to dimensions of the architecture's performance that the reform coalition cannot see — and $\boldsymbol{\varepsilon}_I$ is smaller, because the incumbent's sensing networks are shorter-latency and more deeply embedded.

The legibility problem is that the reform coalition must design its control signal $\mathbf{u}_R(t)$ based on $\mathbf{Y}_R(t)$, which is a degraded, filtered, and possibly deliberately corrupted representation of the architecture's true state. The coalition's model of the architecture's dysfunction is therefore systematically optimistic: the dimensions along which the architecture is failing most severely are precisely the dimensions that $\mathbf{C}_R$ is least likely to capture, because the architecture's observation channels were not designed to detect them.

This is not a problem of incompetence or naivety. It is a structural constraint. A reform coalition that operates entirely within the existing institutional framework — drawing its data from government statistics, its analysis from establishment think tanks, its narrative framing from mainstream media — is observing the architecture through the very channels whose degradation it needs to diagnose. Those channels were shaped, over decades, by the same incumbents who now deploy them to neutralise reform. The information that would reveal the full extent of the architecture's failure is either not collected, not published, or not legible in the categories the existing observation system provides.

The practical consequence is that reforms generated endogenously — from within the architecture, by actors who rely on its observation channels — are systematically too modest. They address the portion of the failure that is legible: the corruption scandal that broke into the media, the service delivery metric that fell below the politically sensitive threshold, the fiscal indicator that alarmed the bond market. They do not address the structural dimensions that remain invisible: the slow degradation of institutional capacity, the capture of regulatory agencies by the industries they regulate, the gradual drift of the value architecture away from citizen preferences. These invisible dimensions continue to deteriorate while the reform coalition congratulates itself on addressing the visible ones.

Paper VII documented this pattern across the Swedish drift loop (where the high-trust consensus culture filtered outlier signals below the threshold of institutional recognition), the Japanese continuity trap (where the system meticulously documented its own stagnation without ever perceiving it as a crisis requiring paradigm change), and the Russian legibility deficit (the terminal case, where the observation channel has been so thoroughly corrupted that the gap between the leadership's model of reality and reality itself is now a strategic vulnerability).

The legibility trap is the equilibrium at which the reform coalition's own success in addressing the legible symptoms reduces the political pressure to expand the observation matrix $\mathbf{C}_R$. The visible dysfunction declines, the coalition declares progress, and the invisible dysfunction accumulates unobserved — until it crosses the threshold at which it can no longer be hidden, by which point the architecture is in crisis. Escape from the trap requires the deliberate construction of independent, decorrelated observation channels that are not under the incumbent's control: citizen audits, satellite sensing, open data platforms, deliberative bodies selected by sortition. These channels increase the effective rank of $\mathbf{C}_R$ and make visible the dimensions that the architecture's own observation system was designed to obscure. The design principles for these channels are developed in Sections 4.1 and 4.5.

3.3 The Incentive-Compatibility Trap

The third trap is the simplest to state and, in many cases, the hardest to escape. Architectural reform requires the cooperation of actors who benefit from the current architecture. The legislators whose votes are needed to amend the constitution are the legislators whose influence derives from the current constitutional arrangements. The bureaucrats who must implement the new regulatory framework are the bureaucrats whose careers were built under the old one. The firms whose compliance is necessary for the new market structure to function are the firms that extracted rents from the old one.

This is a principal-agent problem in which the principals — the reform coalition — must rely on agents — incumbent institutional actors — to implement changes that reduce those agents' own benefits. The agents control the implementation machinery. They control the information about whether implementation is occurring. And they have every incentive to comply in form while resisting in substance, to delay, to dilute, and to wait out the reform window until the political conditions that enabled the reform have passed.

Formally, let $\mathbf{u}_I(t)$ represent the incumbent's counter-mobilisation, as before. But now consider that some components of the reform control signal $\mathbf{u}_R(t)$ must be transmitted *through* the incumbent's institutional apparatus to reach the architecture. The effective reform actuation is:

$$\mathbf{u}_{R\_eff}(t) = \mathbf{M} \cdot \mathbf{u}_R(t)$$

where $\mathbf{M}$ is a transmission matrix controlled, in part, by the incumbent. If the incumbent's interests are threatened by a particular dimension of the reform — a reduction in the aggregation ratio that currently generates its influence, a shortening of the representation chain that currently filters out the preferences it opposes — then $\mathbf{M}$ will attenuate that dimension. The reform's formal passage becomes decoupled from its effective implementation. The architecture's parameters $\mathbf{X}(t)$ do not move as intended.

This is not a hypothetical scenario. It is the core mechanism behind the breakthrough-capture cycle documented in the Brazil study: constitutional amendments, new regulatory agencies, and anti-corruption drives that were formally enacted and then surrounded, co-opted, and extracted by the same *Centrão* whose power they were meant to constrain. It is the mechanism behind the extraction-association-adaptation cycle in Nigeria, where successive reform programmes — structural adjustment, anti-corruption campaigns, fiscal responsibility legislation — were absorbed by the extraction coalition and converted into new channels for rent distribution.

The incentive-compatibility trap is the equilibrium at which the reform coalition achieves formal institutional change — laws passed, agencies created, mandates proclaimed — while the incumbent controller ensures that the effective actuation $\mathbf{u}_{R\_eff}(t)$ is sufficiently attenuated that the architecture's actual parameters $\mathbf{X}(t)$ remain within the region that preserves incumbent benefits. The reform coalition declares victory and disperses. The architecture drifts back toward its pre-reform state. And the next crisis produces the next reform wave, which will be absorbed by the same mechanism.

Escape from this trap requires one of two structural changes to the transmission matrix **M**. The first is to reduce the incumbent's control over implementation — to build delivery mechanisms that do not pass through the incumbent's institutional apparatus, which is the logic of the protected experimental spaces and bypasses discussed in Sections 4.1 and 4.2. The second is to alter the incumbent's interests so that it no longer wishes to attenuate the reform — the logic of the buy-out protocols discussed in Section 4.3. Both approaches have structural prerequisites and domain limitations that Part IV addresses.

These three traps — bypass, legibility, incentive-compatibility — are not independent. They interact. The bypass trap and the legibility trap reinforce each other: the bypass reduces visible dysfunction, which narrows the reform coalition's observation matrix, which makes the remaining dysfunction harder to diagnose. The legibility trap and the incentive-compatibility trap reinforce each other: the incumbent controls the observation channels, which allows it to control what the reform coalition perceives as successful implementation, which allows it to attenuate actuation while reporting compliance. And all three traps sit within the aggregate variety and latency constraints established in Part II: a reform coalition with $\Omega < 1$ is structurally incapable of escaping any of them, because it lacks the dimensional capacity to distinguish the trap from genuine progress.

Part IV now derives the design principles that correspond to each trap, specifying the structural properties that transition mechanisms must possess to break the equilibria described above. The mapping from trap to principle follows the series' established pattern: each architectural defect is addressed by a specific structural response, and the response is load-bearing — remove it, and the trap re-engages.

4. Design Principles for Transition Mechanisms

The three traps modelled in Part III are structural equilibria: stable configurations in which the reform coalition's own actions suppress the signals or incentives that would drive further change. Each trap corresponds to a specific defect in the transition pathway. The bypass trap arises when a parallel channel decouples visible dysfunction from actual dysfunction, removing the pressure for substrate reform. The legibility problem arises when the reform coalition's observation matrix is constrained by the very architecture it seeks to change. The incentive-compatibility trap arises when the transmission of reform actuation depends on actors who lose from the reform's success.

These defects cannot be addressed by exhortation or political will. They must be addressed by structural devices that alter the information flows, incentive gradients, and actuation pathways of the transition itself. In the language of Beer's variety engineering, these devices are amplifiers and attenuators: mechanisms that increase the reform coalition's effective variety, reduce the incumbent's, or both.

This part derives four design principles, each matched to the trap it is intended to break. The mapping follows the series' established pattern of failure-mode-to-structural-response, and is summarised below.

Trap	Design Principle	Mechanism
Bypass trap (§3.1)	Sunset-coupled bypasses (§4.2)	Credible pre-commitment to transfer pressure back to the substrate as bypass performance improves
Legibility problem (§3.2)	Protected experimental spaces with coupled legibility (§4.1); Instrumentalised observer diversity (§4.5)	Local observation channels with short latency, decorrelated from incumbent-controlled sensing; independent, multi-perspective sensing that resists narrative capture
Incentive-compatibility trap (§3.3)	Incumbent buy-out protocols (§4.3)	Compensation packages that align incumbent interests with the transition by matching the dimensionality of what incumbents stand to lose

The four principles are not independent. Protected experimental spaces generate the legibility that makes the case for sunset-coupled bypasses politically viable. Observer diversity protects the integrity of the evidence that experimental spaces produce. Buy-out protocols lower the incumbent's resistance to the architectural changes that the first three principles make visible and politically compelling. Together, they form a transition architecture that is itself subject to the variety and latency constraints established in Part II.

4.1 Protected Experimental Spaces with Coupled Legibility

The most consistent prescription across the country studies is the creation of a protected experimental space: a municipal laboratory, a sandbox state, a Coherence Region, a pilot district, or a constitutionally ring-fenced agency granted genuine authority within a bounded domain. Paper VII identified this convergence and argued that its ubiquity reflects a structural logic, not merely political caution. This section formalises that logic.

A protected experimental space is a local controller with three structural properties that distinguish it from the broader governance architecture. First, its observation channel is shorter — the distance between the governed reality and the decision-maker is reduced, so signal fidelity is higher and latency is lower. Second, its actuation authority is genuine — it can implement decisions within its domain without routing them through the full aggregation and approval machinery of the central architecture. Third, its evaluation is structured to generate legible evidence rather than the appearance of success — it is assessed on learning produced, not on conformity to centrally specified targets.

In the formal vocabulary of Part II, the experimental space is a local instance of the reform controller **R** operating with a higher-rank observation matrix $\mathbf{C}_R$ and a shorter latency τ_R than the broader reform coalition can achieve at the national scale. Its variety ratio Ω_{local} — the ratio of its observational and actuation capacity to the local incumbent's counter-mobilisation capacity — may be substantially more favourable than the national Ω. This makes it a bridgehead: a region of the architecture where the transition constraints are temporarily relaxed enough to permit genuine architectural innovation.

The critical design challenge, however, is not creating the experimental space. It is ensuring that the evidence it produces can propagate to the broader architecture without being destroyed by the same aggregation machinery that degrades normal governance signals. This is the *coupled legibility* requirement. The experiment's outcomes must be transmitted through channels that the incumbent controller cannot easily capture or dismiss, and they must be structured in a form that makes the gap between the experiment's performance and the surrounding architecture's performance politically undeniable.

Paper VII called this "scaling by attraction": the experiment does not impose its model on the broader system; it makes its results visible, and the pressure to adopt them grows as the evidence accumulates. This paper adds the formal condition that the attraction mechanism requires *decorrelated observation channels* — independent verification of the experiment's outcomes that does not rely on the central architecture's own reporting infrastructure. Without such channels, the incumbent can simply dispute the evidence, and the dispute occurs on epistemic terrain the incumbent controls.

The coupling between the experimental space and the broader architecture therefore requires an explicit transmission mechanism: an independent audit body, a citizen deliberative council that reviews the experiment's results, an open data platform that makes raw outcomes publicly accessible, or a legislative trigger that mandates consideration of the experiment's findings within a fixed time window. The transmission mechanism is itself a structural component of the transition architecture, and its design must account for the incumbent's capacity to capture or neutralise it.

4.2 Sunset-Coupled Bypasses

The bypass trap, modelled in Section 3.1, is the equilibrium in which a parallel institutional channel becomes a permanent crutch rather than a bridge. The bypass reduces visible dysfunction, relieves pressure on the unreformed substrate, and stabilises both at a level well below the performance of a genuinely reformed architecture. Escape from the trap requires that the bypass be coupled to a *sunset condition*: a credible, pre-committed mechanism that transfers pressure back to the substrate as the bypass's own performance improves.

The sunset condition must satisfy three properties. First, it must be *irreversible by the incumbent* — the decision to transfer pressure back to the substrate cannot be captured, delayed, or vetoed by the actors who benefit from the substrate's unreformed state. This typically requires embedding the sunset in a constitutional provision, a treaty obligation, or a legislative mechanism that the incumbent does not unilaterally control.

Second, the sunset must be *contingent on demonstrated performance* — it triggers only when the bypass has shown that it can deliver outcomes superior to the unreformed substrate, so that the political coalition for reform has been built before the bypass is wound down. A sunset that triggers before the evidence is conclusive risks stranding the beneficiaries of the bypass without a functional alternative.

Third, the sunset must be *coupled to a specific reform pathway* — the transfer of pressure must activate a pre-agreed institutional process for substrate reform, not merely create an open-ended crisis. The bypass does not simply disappear; its functions are transferred to the reformed substrate, or the substrate is reformed to incorporate the bypass's successful mechanisms into the permanent architecture.

India's UPI provides a concrete illustration of what the absence of a sunset coupling costs. UPI processes ten billion transactions a month, demonstrating that world-class digital infrastructure is achievable within India's governance architecture. But it was not coupled to any mechanism that would transfer the pressure generated by its success back to the analog legal and administrative substrate — land records, contract enforcement, judicial capacity — that caps its potential. The bypass succeeded so thoroughly that it removed the urgency from substrate reform, while the substrate's dysfunction continues to limit the bypass's reach. A sunset-coupled UPI would have included, from the outset, a legislative trigger requiring that when transaction volumes exceeded a specified threshold, a pre-agreed package of judicial and administrative reforms would be automatically tabled for parliamentary approval.

The formal condition for a sunset-coupled bypass to escape the trap is that the coupling parameter c — the strength with which bypass performance **B** drives reform pressure on the substrate **D** — must exceed a critical threshold c_{crit} determined by the substrate's own inertial dynamics. Below c_{crit}, the bypass stabilises the trap; above it, the bypass catalyses substrate reform. Simulation A demonstrates this bifurcation, showing the transition from the trap equilibrium to the escape trajectory as the sunset coupling strength is increased.

4.3 Incumbent Buy-Out Protocols

The incentive-compatibility trap arises because architectural reform requires the cooperation of actors who benefit from the current architecture. The reform coalition must either route around those actors — the logic of protected experimental spaces and bypasses — or change their incentives. Buy-out protocols are the mechanism for the latter.

A buy-out is a compensation package offered to incumbent beneficiaries in exchange for their acquiescence to architectural change. The logic is not normative — it does not depend on the claim that incumbents *deserve* compensation for the loss of rents they extracted from a defective architecture. The logic is structural: if the incumbent controller possesses sufficient variety and latency advantage to block reform, and if the reform coalition cannot assemble a counter-coalition large enough to overwhelm that blocking capacity, then compensating the incumbent may be the only pathway to architectural change that does not require a crisis large enough to temporarily disable the incumbent's veto capacity.

The historical record contains several cases in which buy-out protocols enabled architectural transitions that would otherwise have been blocked. Four positive cases and one contrast case illustrate the structural conditions under which the mechanism works and fails.

British abolition of slavery (1833). The Slave Compensation Act allocated £20 million — approximately 40 percent of the Treasury's annual budget — to compensate slaveholders for the loss of their human property. The payment was not made to the enslaved, who received nothing; it was made to the enslavers, whose political and economic power was sufficient to block abolition indefinitely had their interests not been bought out. The compensation was structured as a loan raised in the bond market and serviced by general taxation, spreading the cost across generations. It was, by any ethical measure, grotesque. It was also effective: the abolition legislation passed, and the institution was dismantled throughout the British Empire within a defined transition period. The buy-out addressed the *economic* dimension of the incumbent's interests directly, and the incumbent's political capacity to resist was sufficiently concentrated that a single, large, centrally administered compensation package could neutralise it.

Meiji stipend commutation (1876). The Meiji Restoration abolished the feudal class structure that had organised Japanese society for centuries. The samurai class, approximately 1.9 million people, held hereditary stipends that constituted the largest single item of government expenditure. Commuting these stipends into interest-bearing government bonds — a one-time capitalisation of a permanent claim — simultaneously dismantled the economic foundation of the feudal order and freed the state's fiscal capacity for industrial investment. The commutation was not voluntary for individual samurai, but it was sufficiently generous that the class as a whole acquiesced, and the transition occurred without the civil war that had accompanied comparable feudal abolitions in Europe. The buy-out addressed the *economic* dimension of the incumbent's interests and succeeded because the alternative — a protracted military confrontation with a warrior class — was costlier than compensation.

Montreal Protocol (1987). The phase-out of ozone-depleting substances succeeded in part because DuPont, the dominant producer of chlorofluorocarbons, stood to profit from patented substitutes. The company's research programme had developed hydrofluorocarbons (HFCs) that could replace CFCs in most applications, and the regulatory transition created a guaranteed market for these substitutes. DuPont's interests, in other words, were *re-aligned*: the new architecture offered a profit pathway that the old architecture could not. The buy-out here was not a cash transfer but a market design that converted the incumbent from an opponent of regulation into a beneficiary of it. The *economic* dimension of the incumbent's interests was directly addressed.

German coal phase-out (2020). The *Kohleausstieg* legislation committed Germany to closing its coal-fired power plants by 2038, with explicit compensation packages for affected regions and companies: €40 billion in structural aid for coal-dependent regions, and direct payments to plant operators for foregone revenues. The compensation addressed both the corporate incumbents (through revenue replacement) and the regional incumbents (through investment in economic diversification), preventing the formation of a unified blocking coalition that could have derailed the transition.

Contrast case — New Zealand agricultural reform (1984). New Zealand eliminated nearly all agricultural subsidies in a single budget, with minimal compensation and no transition period. The reform succeeded not through buy-out but through a crisis window that temporarily disabled the incumbent's veto capacity: a currency crisis, a snap election, a newly elected government with a mandate for radical reform, and a parliamentary system with few veto points. The reform was implemented before the farming lobby could organise counter-mobilisation. This case demonstrates that buy-outs are not *universally* necessary; they are required when the incumbent possesses concentrated veto capacity and the reform window is not wide enough to overpower it. New Zealand's institutional structure — a unicameral parliament with strong executive dominance — made it possible to act faster than the incumbent could respond.

The four positive cases share a structural property that the contrast case does not: the incumbent interests being compensated were predominantly *economic*. Slaveholders, feudal stipend holders, chemical manufacturers, and coal plant operators stood to lose income streams and asset values. Their interests could be addressed through financial compensation or the creation of alternative profit pathways. The buy-out protocol succeeded by matching the *dimensionality* of the compensation to the dimensionality of the interests being displaced — a heuristic application of the variety concept, not a formal application. The term "variety" is used here loosely to mean the number of independent dimensions along which incumbents hold stakes; it is not a claim that the compensation package must satisfy a formal Ashby condition.

4.4 The Non-Compensability Boundary

The analysis of buy-out protocols in Section 4.3 identifies a structural condition that determines when the mechanism can work. The condition is not normative — it does not depend on whether incumbents are morally entitled to compensation — but topological: it concerns the nature of the interests being displaced.

When incumbent interests are primarily economic — income streams, asset values, market share, employment prospects — they can, in principle, be compensated. The compensation may be expensive, and the expense may be ethically objectionable, but the structural pathway exists. When incumbent interests are identity-constitutive, ideological, or theocratic, they cannot be compensated with fungible resources. The *siloviki* in Russia do not simply extract rents; their identity, status, and survival depend on the continuation of the power vertical that enables extraction. The Chinese Communist Party's control preservation imperative is not a financial preference; it is the organising principle of the regime, and no compensation package can be offered that would make its abandonment incentive-compatible, because the abandonment would eliminate the regime's reason for being. Theocratic veto players — religious authorities whose political power derives from doctrinal claims — cannot be bought out without undermining the doctrinal claim itself.

This yields a testable prediction: buy-out protocols succeed when incumbent interests are economic and fail when they are constitutive. Historical cases of successful buy-out (abolition, Meiji, Montreal, German coal) should uniformly involve economic interests; cases where architectural reform has been blocked despite attempts at buy-out should involve identity-constitutive or ideological interests. The prediction is falsifiable by examining cases of attempted architectural transitions, coding the nature of the incumbent's primary interest, and comparing transition outcomes.

The non-compensability boundary defines the domain within which the other three design principles — protected experimental spaces, sunset-coupled bypasses, and observer diversity — must carry the full burden of transition. When incumbents cannot be bought out, they must be routed around, or their veto capacity must be reduced through the accumulation of political pressure generated by the legibility that experimental spaces and observer diversity provide.

4.5 Instrumentalising Observer Diversity

The legibility problem modelled in Section 3.2 is that the reform coalition's observation matrix **C**_R is constrained by the architecture it seeks to change. The incumbent controls the information channels, and the reform coalition's model of the architecture's dysfunction is systematically optimistic. Escape from this trap requires the deliberate construction of independent observation channels — sensing systems that are not under the incumbent's control and whose outputs cannot be easily dismissed or absorbed into the incumbent's narrative framework.

The structural logic is the same one that Paper II applied to governance controllers: a single observer with a fixed observation matrix has a fixed variety ceiling, and disturbances outside that ceiling are invisible. A *population* of independent observers, each with a different observation matrix, can collectively perceive disturbance dimensions that any single observer would miss — provided their errors are decorrelated.

This is not simply an argument for "more data" or "better transparency." The incumbent controller can absorb and reframe data that enters through the channels it controls. The structural property that matters is *decorrelation*: the observation channels must be constituted such that their systematic biases are uncorrelated with each other and with the biases of the incumbent's own reporting infrastructure. When this condition holds, a convergence across channels is a strong signal that the observed condition is real; a divergence is a strong signal that at least one channel is degraded, and the divergence itself is the most informative datum.

Concrete mechanisms for instrumentalising observer diversity include:

- **Citizen audits**: structured processes in which randomly selected citizens review government performance data and produce independent assessments. The random selection decorelates the auditors' biases from the bureaucratic biases of the agencies being audited.
- **Satellite and remote sensing**: observation channels that bypass the incumbent's reporting infrastructure entirely, providing direct measurements of environmental conditions, infrastructure development, or economic activity that cannot be easily falsified.
- **Open data platforms**: public repositories of raw administrative data, structured to allow independent analysis by researchers, civil society organisations, and journalists, each bringing different analytical frameworks and different biases.
- **Deliberative bodies selected by sortition**: citizens' assemblies, policy juries, and foresight panels whose members are selected by lot rather than by political appointment, breaking the correlation between the observer's institutional position and their observational biases.
- **Independent fiscal watchdogs and audit institutions**: bodies whose leadership is appointed through processes that the incumbent does not control, and whose funding is constitutionally or statutorily protected from retaliatory cuts.

Each of these mechanisms increases the effective rank of the reform coalition's observation matrix by adding dimensions that the incumbent cannot simultaneously control. The incumbent can discredit one channel, defund one watchdog, or capture one deliberative body, but the cost of capturing or discrediting all of them scales with their number and their independence from each other.

A critical caveat is necessary, correcting a claim that appeared in earlier formulations of this argument. Visibility, once achieved, is not permanently irreversible. The Russian case — the series' own terminal example of the legibility deficit — demonstrates that an incumbent can systematically reverse visibility by making accurate information dangerous to the informant. Independent media can be defunded or criminalised. Civil society organisations can be designated as foreign agents. Independent audit institutions can be captured from within. Observer diversity is not a one-time achievement; it is a condition that must be actively maintained, and the incumbent's counter-mobilisation against it is a continuous process.

What observer diversity provides is a *window* — a period during which the gap between the architecture's model of reality and reality itself is sufficiently legible to the reform coalition and the broader public that political action becomes possible. The design challenge is to act within that window before it closes. This is why observer diversity must be coupled to the other three design principles: protected experimental spaces that can be stood up while the window is open, sunset-coupled bypasses that lock in reforms before the window closes, and buy-out protocols that reduce the incumbent's incentive to close it in the first place.

The four design principles together constitute a transition architecture that is itself subject to the variety and latency constraints derived in Part II. A reform coalition that deploys them must still satisfy $\Omega \geq 1$ — sufficient variety to match the incumbent's counter-moves — and must still operate within the gain band that τ_R imposes. The principles do not guarantee success. They specify the structural properties that a transition mechanism must possess to have a chance of succeeding against an adaptive incumbent. Without them, the three traps modelled in Part III will re-engage, and the transition will follow the characteristic failure trajectories that the country studies have documented.

Part V now shifts from the structural design of transition mechanisms to the dynamic context in which they must operate: the race between the rate of environmental change and the rate of architectural adaptation. The concept of *transition bandwidth* introduced there provides the temporal dimension that the static analysis of traps and principles cannot supply.

5. Transition Bandwidth: The Dynamic Constraint

The analysis so far has been static. Parts II through IV establish the variety and latency conditions that a transition pathway must satisfy to overcome incumbent resistance, and the design principles that prevent the pathway from settling into one of the three structural traps. But the environment in which transitions occur is not static. The effective dimensionality of the disturbance space expands over time, as Paper VI established. New technologies, ecological shifts, demographic transitions, and geopolitical reconfigurations continuously introduce new dimensions of variation that existing architectures were not designed to govern. The rate at which the architecture must change to maintain observability is therefore not zero. It is positive, and it may be accelerating.

This part introduces the concept of *transition bandwidth* — the rate at which a governance system can peacefully redesign its own structure — and models the dynamic relationship between that rate and the rate of environmental change. When transition bandwidth exceeds the rate at which architectural adaptation is demanded, the system can evolve its structure before the variety gap becomes fatal. When it falls below, forced dissolution — collapse, revolution, or constitutional rupture — becomes the only remaining pathway. The part concludes by applying this dynamic to the governance implications of frontier artificial intelligence, which raises the demanded rate of adaptation for all architectures simultaneously.

5.1 Operational and Transition Bandwidth Distinguished

Every governance architecture possesses some capacity to respond to disturbances without altering its own structure. A central bank adjusts interest rates in response to inflation; a health ministry reallocates resources during a pandemic; a municipal council redirects maintenance crews after a storm. These are operational responses: they use the existing institutional machinery, within its existing parameter settings, to bring the system back toward the target state. The maximum rate at which such responses can be generated and executed is the system's *operational bandwidth*.

Operational bandwidth is a function of the architecture's latency, its actuation capacity, and the variety of its observation channels — precisely the parameters that Papers I through VI analyse. A system with high operational bandwidth can absorb large, frequent, or multi-dimensional disturbances without losing stability. A system with low operational bandwidth is fragile even to routine shocks.

Transition bandwidth is a different capacity. It is the rate at which the architecture can change its own structure — its latency distribution, its aggregation ratios, its representation chain depth, its value function dimensionality, its feedback protection mechanisms. This is not a response to a disturbance within the existing framework; it is a redesign of the framework itself. The mechanisms that generate transition bandwidth are the institutional pathways for structural reform: constitutional amendment procedures, legislative capacity for organic institutional redesign, the existence and credibility of sunset clauses, the density of experimental governance spaces that can generate architectural innovations, and the deliberative infrastructure (citizens' assemblies, foresight bodies, independent audit institutions) that can surface new value dimensions and translate them into architectural change.

A system can have high operational bandwidth and low transition bandwidth simultaneously. Rome under the late Republic possessed enormous operational capacity — it fielded armies across three continents, administered a complex tax system, and maintained urban infrastructure at a scale unmatched for centuries. Its transition bandwidth, however, was near zero: the constitutional architecture of the Republic could not peacefully accommodate the shift from a city-state's governance machinery to an empire's, and the result was a century of civil war that ended with the Republic's dissolution and replacement by the Principate. The late Soviet Union possessed formidable operational bandwidth — it maintained nuclear parity with the United States, operated a space programme, and administered a continent-spanning command economy. Its transition bandwidth was zero: the architecture had eliminated all legitimate channels for structural critique, and when the need for architectural change became undeniable, the system could not redesign itself. It dissolved.

Contemporary China exhibits the same asymmetry. Its operational bandwidth is among the highest in the world — it can mobilise resources for infrastructure investment, pandemic response, and industrial policy at speeds no democratic system can match. Its transition bandwidth, however, is constrained by the Control Preservation Imperative that the country study identified: any mechanism that would permit peaceful architectural redesign (independent feedback channels, reversible decision structures, protected experimental spaces with genuine autonomy) is correctly perceived by the incumbent controller as a threat to the architecture that sustains it. The system can execute within its existing paradigm with extraordinary effectiveness; it cannot evolve the paradigm without triggering the immune response.

The distinction matters because the challenges that governance architectures face are increasingly structural rather than operational. Climate change demands not merely better emissions monitoring but a reconfiguration of the value architecture that currently treats atmospheric sinks as externalities. Demographic transition demands not merely adjustments to pension formulas but a redesign of the intergenerational social contract. Frontier AI demands not merely better regulatory oversight but a reconstitution of the relationship between human decision-making and machine optimisation. These are architectural challenges, and they require transition bandwidth to address. A system that possesses only operational bandwidth will, in Paper VI's language, see its variety gap widen until forced dissolution becomes inevitable.

5.2 Observable Proxies for Transition Bandwidth

Transition bandwidth is not directly measurable. Like the variety gap that Paper VIII operationalised, it must be estimated from observable institutional characteristics that correlate with the capacity for peaceful structural change. This section specifies a set of candidate proxies, connecting the concept to the measurement framework the series has begun to develop. The paper does not undertake the empirical work of calibration; it specifies the proxies that such work would require.

The first proxy is the *constitutional amendment rate* — the frequency with which a governance system successfully alters its own formal operating rules through regularised procedures. A system that has amended its constitution several times per decade over a sustained period without constitutional crisis possesses demonstrated transition bandwidth. A system whose constitution has never been amended, or whose amendments occur only through rupture (revolution, coup, external imposition), possesses near-zero transition bandwidth regardless of its operational performance.

The second proxy is the *prevalence and credibility of sunset clauses*. A governance system that routinely includes automatic expiry dates in its legislation, and in which those expiries are honoured rather than routinely extended, has built a structural mechanism for revisiting and revising its own parameter settings. The sunset clause forces reconsideration; it prevents the architecture from freezing into a configuration that was appropriate for the conditions of its enactment but has become maladaptive. The credibility of sunsets — the proportion that actually trigger reform rather than automatic renewal — is as important as their prevalence.

The third proxy is the *existence and density of experimental governance spaces*: the municipal laboratories, sandbox states, Coherence Regions, and pilot programmes that Paper VII identified as the convergent first step across jurisdictions. A system that maintains multiple such spaces, grants them genuine authority, and has established pathways for their innovations to propagate into the broader architecture possesses higher transition bandwidth than one that governs uniformly from the centre.

The fourth proxy is the *institutionalisation of deliberative infrastructure*: citizens' assemblies, foresight bodies, intergenerational councils, and participatory budgeting mechanisms that can surface new value dimensions and translate them into architectural proposals. These mechanisms increase the variety of the reform coalition's observation matrix — they are, in the language of Section 4.5, institutionalised observer diversity — and they provide a legitimate channel for structural critique that does not depend on the incumbent's permission.

The fifth proxy is the *independence and longevity of audit and evaluation institutions*: supreme audit institutions, parliamentary budget offices, independent evaluators, and statistical agencies whose leadership is appointed through processes that the incumbent does not control, and whose funding is constitutionally or statutorily protected. These institutions provide the decorrelated observation channels that make the gap between the architecture's model of reality and reality itself periodically visible. Their longevity — the number of political cycles they have survived without being captured or defunded — is a measure of the system's tolerance for independent structural diagnosis.

None of these proxies is sufficient on its own. A system that amends its constitution frequently but only to consolidate executive power does not possess transition bandwidth in the sense that matters; the amendments must expand the architecture's dimensionality or reduce its structural blindness, not contract them. A system with many experimental spaces that are denied genuine authority or whose results are systematically ignored does not benefit from their existence. The proxies must be evaluated jointly, and their values must be interpreted through the qualitative understanding of the architecture's immune system that the country studies provide. The estimation framework of Paper VIII can accommodate them as an extension — either as a ninth structural primitive or as a derived quantity computed from the existing eight. The paper specifies the connection but leaves the integration to future measurement work.

5.3 The Race Between Adaptation and Dissolution

Let α(*t*) denote the rate at which the disturbance environment demands architectural change at time *t*. This is the rate at which new causally relevant dimensions enter the effective state space **R** that the governance architecture must navigate — the α of Paper VI's dynamic model, now interpreted as the demand side of the transition equation. When a new technology creates a new domain of social interaction (social media, synthetic biology, autonomous weapons), when an ecological threshold is crossed that alters the dynamics of a managed resource, or when a demographic transition shifts the dependency ratio beyond the range for which existing fiscal architectures were designed, α spikes.

Let β(*t*) denote the system's effective transition bandwidth at time *t* — the maximum rate at which it can peacefully expand its value architecture's dimensionality, shorten its representation chains, reduce its response latencies, or improve its feedback protection mechanisms. β is a function of the proxies specified in Section 5.2, of the system's Ω ratio (because architectural change must overcome incumbent resistance), and of the gain ceiling that τ_R imposes (because pushing architectural change too fast triggers oscillation).

The dynamics of the variety gap **G** from Paper VI, extended to include the transition constraint, are:

dG/dt = α(t) − β(t)

When β(*t*) ≥ α(*t*), the system's capacity for architectural adaptation matches or exceeds the rate at which the environment demands it. The variety gap is stable or shrinking. The system maintains perceptual contact with the reality it must govern, and managed architectural evolution is possible.

When β(*t*) `<` α(*t*), the variety gap grows. The architecture's dimensionality falls progressively further behind the dimensionality of the disturbance environment. The excluded dimensions accumulate as externalities, as unexplained variance, as crises that the system's own diagnostic categories cannot interpret. The system is in the unmanaged regime described in Paper VI, but now with an explicit mechanism: the gap grows not merely because the architecture is static, but because the political economy of transition prevents it from adapting fast enough even when the need for adaptation is partially legible.

Simulation C, described in Part VI, demonstrates a critical dynamic that this equation alone does not capture. As the variety gap grows, the system experiences more frequent and more severe crises — the excluded dimensions forcing themselves into visibility. Crisis management consumes operational bandwidth, which diverts institutional capacity away from structural redesign. There is a *coupling* between the gap and the transition bandwidth itself:

$$\beta_{effective}(t) = \beta_{nominal}(t) \cdot f(G(t))$$

where $f(\mathbf{G})$ is a decreasing function: as the gap widens, a larger fraction of the system's institutional capacity is absorbed by crisis response, and the remaining capacity for structural redesign shrinks. This coupling produces a *transition-bandwidth trap* — a point of no return, reached strictly before $\mathbf{G}_{crit}$, at which the system still functions operationally but has irreversibly lost the capacity to redesign itself. The trap is a two-threshold structure. At the first threshold, **G** becomes large enough that crisis management begins to crowd out structural reform. The system's effective β falls, which causes **G** to grow faster, which further reduces β. This is a positive feedback loop with a stable attractor at $\mathbf{G} \gg \mathbf{G}_{crit}$ — dissolution. At the second threshold, $\mathbf{G}_{crit}$ itself, the system's observation channels become constitutionally uninformative, and the system can no longer perceive the sources of its own instability. But the first threshold — the loss of transition bandwidth — occurs earlier, and it is the point at which the system's fate is effectively sealed, even if the final collapse is years or decades away.

This two-threshold structure is, if confirmed by simulation and empirical work, the paper's most significant dynamic finding. It implies that a governance architecture can appear operationally functional — responding to crises, delivering services, maintaining order — while having already passed the point at which peaceful architectural adaptation is possible. The system is in the transition-bandwidth trap: it can still run, but it can no longer steer. The remaining trajectory is determined by the accumulation of unmanaged variety until forced dissolution occurs.

The country studies contain suggestive evidence for this dynamic. Japan's Continuity Trap can be reinterpreted as a transition-bandwidth trap: the architecture maintained high operational bandwidth (the trains run on time, the debt is serviced, the social order is intact) while transition bandwidth fell to near zero — the Iron Triangle of the LDP, the bureaucracy, and big business systematically eliminated the mechanisms through which the post-war paradigm could be peacefully replaced. The system is not in crisis in the operational sense. It is in the transition-bandwidth trap, and the variety gap is widening as the environment (demographic decline, technological disruption, geopolitical pressure) demands architectural change that the system cannot supply. China's Calibration Deficit exhibits the same structure: enormous operational bandwidth, transition bandwidth suppressed by the Control Preservation Imperative, and a variety gap that grows with each campaign-overshoot-correction cycle as the system's own immune response prevents the architectural reforms that would stabilise it.

5.4 The Singularity as a Transition-Bandwidth Problem

The dynamic model of transition bandwidth has a direct and urgent application to the governance challenges posed by frontier artificial intelligence. The development of AI systems capable of recursive self-improvement — or of generating novel scientific and technological knowledge at a rate that outpaces human institutional deliberation — represents a structural increase in $\alpha(t)$ for every governance architecture simultaneously. The rate at which new disturbance dimensions enter the effective state space accelerates: the labour market, the information ecosystem, the military balance, the structure of the economy itself, and eventually the relationship between human and machine agency all become dimensions along which the existing architecture's parameter settings are rapidly rendered obsolete.

This is not a prediction that AI will produce any particular outcome. It is an observation that the *rate of architectural adaptation demanded* is a function of the rate at which the environment generates novelty, and frontier AI increases that rate by an amount that is large, uncertain, and potentially discontinuous. The series' framework does not require a specific forecast of AI capabilities to make its structural point: any substantial acceleration in $\alpha(t)$ shifts the balance between adaptation and dissolution for every architecture, favouring those with high transition bandwidth and penalising those without it.

The architectures with the highest transition bandwidth are those that have institutionalised mechanisms for peaceful self-redesign. Federations with robust constitutional amendment procedures, active sub-federal experimentation, and independent deliberative infrastructure. Polycentric systems in which multiple governance units run parallel experiments and the successful innovations propagate by attraction rather than by central mandate. Systems that have preserved observer diversity — multiple, independently-constituted sensing systems whose errors are decorrelated — rather than consolidating observation into a single AI-driven modelling infrastructure. These architectures can, in principle, increase their β to track an accelerating α, because the mechanisms for doing so are already built into their institutional fabric.

The architectures with the lowest transition bandwidth are those that have concentrated authority in a single controller whose survival depends on suppressing the variety that architectural adaptation requires. Rigid global regimes whose decision procedures require consensus among actors with veto capacity and incompatible interests. Captured one-party states where the immune system has eliminated all legitimate channels for structural critique and where the leadership's own observation channel has been systematically degraded. Systems that have consolidated their epistemic infrastructure into a single AI-driven modelling framework, producing correlated errors that make the proxy-target divergence unobservable. For these architectures, an acceleration in $\alpha(t)$ does not trigger adaptation. It triggers the transition-bandwidth trap: crisis management consumes the capacity that would be needed for redesign, β falls further, and the gap widens faster.

The paper does not predict which architectures will survive and which will dissolve. It identifies the dynamic condition that will determine the outcome. Survival goes to the architectures that can redesign themselves faster than the environment changes. The condition is not satisfied by good intentions, by the accumulation of operational capacity, or by the sophistication of the AI systems that the architecture deploys. It is satisfied by the presence of institutional mechanisms that allow the architecture to perceive its own obsolescence and to act on that perception before the excluded dimensions force a reckoning that the architecture cannot survive. The transition-bandwidth trap is the mechanism by which architectures that lack those mechanisms are selected out — not by any external judge, but by the accumulating weight of the variety they cannot absorb.

Part VI now presents the simulation architecture that grounds the formal claims of Parts II through V: three computational models that demonstrate the bypass-trap dynamics, the contested-control phase diagram, and the transition-bandwidth trap with its critical two-threshold structure. The simulations are not proofs; they are disciplined thought experiments that make the structural logic visible and testable. Their parameters, their code, and their output are open to inspection, replication, and challenge.

6. Simulation Architecture

The formal claims of Parts II through V are theoretical. They describe mechanisms — the bypass trap, the transition variety ratio, the transition-bandwidth trap — that should, if the theory is sound, produce characteristic dynamical signatures under specified conditions. Computational simulation provides a disciplined way to test whether those signatures actually emerge from the posited mechanisms, and to map the boundaries of the parameter regimes in which they hold. The simulations are not proofs, and they are not empirical validations. They are *existence demonstrations*: they show that a given structural logic, when implemented as a minimal dynamical system, generates the predicted behaviour. If the behaviour does not emerge, the logic requires revision. If it does, the logic earns the right to be tested against empirical data.

The series has established a simulation methodology: minimal models with transparent parameters, comparison of two or more architectures under identical disturbance conditions, and output visualised through heatmaps, phase portraits, and summary metrics. This paper extends that methodology in one important respect. Following the recommendation to strengthen the series' epistemic tiering, all three simulations below are run with Monte Carlo replication (50–100 seeds, distributions reported rather than single-run anecdotes) and include at least one parameter-sweep heatmap demonstrating that the qualitative result is robust across a region of parameter space, not an artefact of a single parameterisation. The code is open, documented, and designed for replication.

6.1 Simulation A: Bypass-Trap Dynamics

This simulation operationalises the bypass trap modelled in Section 3.1 and the sunset-coupled escape mechanism of Section 4.2. The system consists of two state variables evolving in discrete time over $T = 200$ steps.

State variables.

- $D(t) \in [0, 1]$: dysfunction of the unreformed substrate, where 0 is full functionality and 1 is complete failure.
- $B(t) \in [0, B_{max}(D)]$: capacity of the bypass to deliver outcomes. The ceiling $B_{max}(D) = 1 - c_{cap} \cdot D$, where $c_{cap} \in [0, 1]$ captures how severely substrate dysfunction constrains bypass effectiveness. When c_{cap} is high, the bypass cannot fully compensate for a dysfunctional substrate.

Dynamics. The substrate dysfunction evolves according to:

$$D(t+1) = D(t) + \alpha \cdot D(t) \cdot (1 - D(t)) - \beta \cdot R(t) + \varepsilon_D$$

where α captures the natural tendency of dysfunction to deepen when unchecked (logistic growth toward 1), $R(t)$ is the reform pressure applied to the substrate, and ε_D is a small noise term. Reform pressure is proportional to *visible* dysfunction:

$$R(t) = K_R \cdot D_{vis}(t)$$

$$D_{vis}(t) = D(t) \cdot (1 - c_{vis} \cdot B(t))$$

The visibility coefficient c_{vis} determines how much the bypass masks the underlying dysfunction: as B increases, the visible signal of dysfunction declines even if D remains high.

The bypass capacity evolves according to:

$$B(t+1) = B(t) + \gamma \cdot R_B(t) \cdot (1 - B(t)/B_{max}(D(t))) - \delta \cdot B(t) + \varepsilon_B$$

where $R_B(t)$ is investment in bypass expansion, γ is its effectiveness, and δ is a decay rate representing the entropy that affects any institutional mechanism. The ceiling B_{max} couples the bypass to the substrate: a severely dysfunctional substrate caps what the bypass can achieve.

Scenarios. Three scenarios are compared, each with identical initial conditions ($D(0) = 0.4$, $B(0) = 0.1$) and identical parameter draws across seeds.

1. **No bypass.** B is permanently zero; all reform pressure targets the substrate directly.
2. **Permanent bypass.** $R_B(t)$ is positive and constant; the bypass expands but is never coupled back to substrate reform.
3. **Sunset-coupled bypass.** $R_B(t)$ is positive, but when $B(t)$ exceeds a threshold B_{sunset}, the coupling parameter c_{vis} is increased — the bypass's success is deliberately made to *increase* the visibility of remaining dysfunction, transferring pressure back to the substrate. This is the structural mechanism of Section 4.2.

Expected outputs.

- *Time-series panels*: $D(t)$, $B(t)$, and $D_{vis}(t)$ for each scenario, with distributions from the Monte Carlo ensemble.
- *Phase portrait*: (D, B) trajectory, showing the stable attractor under the permanent bypass and the escape trajectory under the sunset coupling.
- *Parameter-sweep heatmap*: final D at $t = 200$ as a function of c_{vis} and c_{cap}, demonstrating the region in which the trap is stable and the region in which the sunset coupling breaks it.

The critical prediction is that the permanent bypass stabilises a low-performance attractor (D high, B moderate, D_{vis} low), while the sunset-coupled bypass escapes that attractor by making success itself the trigger for renewed pressure on the unreformed substrate.

6.2 Simulation B: Reform-Incumbent Latency Asymmetry

This simulation operationalises the contested control model of Part II. Two controllers — a reform coalition **R** and an incumbent coalition **I** — act on the same architectural state vector, with asymmetric latencies and observation dimensionalities. The purpose is to map the phase diagram of transition outcomes and to determine whether the heuristic condition $\Omega \geq 1$ (the transition variety ratio) functions as a threshold.

State space. The architecture is represented by a state vector $\mathbf{X}(t) \in \mathbb{R}^4$, whose four components correspond to latent structural dimensions: aggregation ratio, representation chain depth, value-function dimensionality, and feedback protection. The target architecture $\mathbf{X}^*$ is a vector with lower aggregation, shorter chains, higher dimensionality, and stronger feedback protection. The incumbent's preferred state is $\mathbf{X}_I = \mathbf{X}(0)$, the initial architecture.

Dynamics. The state evolves as:

$$\mathbf{X}(t+1) = \mathbf{A} \cdot \mathbf{X}(t) + \mathbf{B}_R \cdot \mathbf{u}_R(t - \tau_R) + \mathbf{B}_I \cdot \mathbf{u}_I(t - \tau_I) + \mathbf{d}(t)$$

where $\mathbf{A}$ is an inertial matrix (diagonal with values near 1), $\mathbf{B}_R$ and $\mathbf{B}_I$ are identity matrices scaled by actuation efficiency, and $\mathbf{d}(t)$ is a small exogenous shock. The control signals are proportional feedback:

$$\mathbf{u}_R(t) = \mathbf{K}_R \cdot (\mathbf{X}^* - \hat{\mathbf{Y}}_R(t)) \quad \mathbf{u}_I(t) = \mathbf{K}_I \cdot (\mathbf{X}_I - \hat{\mathbf{Y}}_I(t))$$

where $\hat{\mathbf{Y}}_R$ and $\hat{\mathbf{Y}}_I$ are the noisy observations available to each controller:

$$\hat{\mathbf{Y}}_R(t) = \mathbf{C}_R \cdot \mathbf{X}(t) + \boldsymbol{\varepsilon}_R, \quad \hat{\mathbf{Y}}_I(t) = \mathbf{C}_I \cdot \mathbf{X}(t) + \boldsymbol{\varepsilon}_I$$

Observation matrices. The rank of $\mathbf{C}_R$ defines the reform coalition's effective observational variety $\dim(\mathbf{R})$; the rank of $\mathbf{C}_I$ defines the incumbent's $\dim(\mathbf{I})$. The simulation sweeps over $\mathrm{rank}(\mathbf{C}_R) \in \{1, 2, 3, 4\}$ and $\mathrm{rank}(\mathbf{C}_I) \in \{1, 2, 3, 4\}$, producing a 4×4 grid of $\Omega = \dim(\mathbf{R})/\dim(\mathbf{I})$ values.

Latency sweep. Independently, the ratio τ_R/τ_I is swept from 1 (symmetric) to 10 (extreme asymmetry), holding $\tau_I = 1$ fixed.

Gain constraint. Reform gain $\mathbf{K}_R$ is bounded above by the latency-imposed ceiling $\mathbf{K}_{max} \approx 1/(\tau_R \cdot |\mathbf{A}|)$, where $|\mathbf{A}|$ is the spectral norm of $\mathbf{A}$. When $\mathbf{K}_R$ exceeds this ceiling, oscillation is predicted.

Outcome classification. For each parameter combination, the simulation classifies the outcome after $T = 150$ steps:

- *Reform success*: $\mathbf{X}(T)$ is within the goal set $\mathbf{G}$ (Euclidean distance to $\mathbf{X}^*$ below a threshold).
- *Reform absorption*: $\mathbf{X}(T)$ is closer to $\mathbf{X}_I$ than to $\mathbf{X}^*$, and the reform signal has been neutralised.
- *Oscillation/backlash*: $\mathbf{X}(t)$ exhibits persistent large-amplitude cycles, indicating gain-ceiling violation.

Expected outputs.

- *Phase diagram*: outcome class as a function of Ω and τ_R/τ_I, visualised as a colour-coded heatmap across the 4×4 × $N_{latency}$ grid.
- *Representative trajectories*: $\mathbf{X}(t)$ projections for three illustrative parameter combinations (success, absorption, oscillation).
- *Gain-ceiling confirmation*: overlay of oscillation outcomes on the $\mathbf{K}_R$ vs. $1/(\tau_R \cdot |\mathbf{A}|)$ plane, testing whether the predicted ceiling aligns with observed instability.

The critical purpose of this simulation is to discipline the formal claims of Section 2.2. If a clean threshold at Ω = 1 emerges — that is, if outcomes are predominantly *success* or *oscillation* when Ω ≥ 1 and *absorption* when Ω < 1, across a range of latency ratios — the heuristic inequality earns its place in the paper. If the boundary is mushy, with success and absorption interleaved in the Ω ≥ 1 region, the paper must soften its formal language and present Ω as a diagnostic indicator rather than a threshold condition. The simulation is run *before* the final version of Section 2.2 is written, not after.

6.3 Simulation C: Transition-Bandwidth Race

This simulation operationalises the dynamic model of Section 5.3, including the critical coupling between variety gap and effective transition bandwidth that produces the two-threshold trap.

State variables.

- $G(t)$: the variety gap (dimensionless, ≥ 0).
- $C(t)$: the fraction of institutional capacity devoted to crisis management, ∈ [0, 1].
- $\beta_{nominal}$: the system's maximum transition bandwidth when no crisis management is required.

Dynamics. The gap evolves according to equation (1) from Section 5.3:

$$G(t+1) = G(t) + \alpha(t) - \beta_{effective}(t) + \varepsilon_G$$

where $\alpha(t)$ is the demanded rate of architectural change. In the baseline scenario, α is constant; in the accelerated scenario, α increases linearly after a threshold time, representing the AI-driven acceleration of Section 5.4.

The effective transition bandwidth is given by equation (2):

$$\beta_{effective}(t) = \beta_{nominal} \cdot f(G(t))$$

where $f(G) = \max(0, 1 - c_{coupling} \cdot G)$. As the gap widens, a larger fraction of institutional capacity is diverted to crisis management, reducing $\beta_{effective}$. The parameter $c_{coupling}$ controls the severity of this feedback.

Crisis management fraction evolves as:

$C(t) = \min(1, G(t) / G\_crit)$

where G_crit is the constitutional unobservability threshold from Papers III and VI. When $G \geq G$_crit, all capacity is consumed by crisis management, and β_effective falls to zero.

Critical thresholds.

- **Transition-bandwidth trap threshold** G_trap: the value of G at which β_effective falls below α (the demanded adaptation rate). Once G exceeds G_trap, the gap grows autonomously because the system can no longer adapt fast enough to close it. This threshold is reached *before* G_crit whenever α > 0 and c_coupling > 0.
- **Dissolution threshold** G_crit: the value at which the system's observation channels become constitutionally uninformative. This is the point of no perceptual return; before it, the system can still see that it is failing; after it, the failure itself becomes invisible.

The two-threshold structure — G_trap `<` G_crit — is the simulation's central prediction. A system crosses G_trap first, losing the capacity for peaceful self-redesign while still appearing operationally functional. G_crit follows later, and dissolution is by then inevitable.

Parameterisations. Three system profiles are compared, each with identical α trajectories and identical initial $G(0) = 0.2$, but different values of β_nominal and c_coupling:

1. **High-bandwidth federation.** β_nominal = 0.08, c_coupling = 0.3 (low coupling: the system's deliberative infrastructure buffers crisis consumption of reform capacity).
2. **Locked regime.** β_nominal = 0.02, c_coupling = 0.9 (high coupling: crises overwhelm the narrow decision apparatus, rapidly consuming reform capacity).
3. **Bypass-heavy system.** β_nominal = 0.05, c_coupling = 0.5, but with the additional feature that G is initially *underestimated* by 40% — the bypass masks the true gap, a direct implementation of the legibility problem.

Expected outputs.

- *Time-series panels*: $G(t)$, β_effective(t), and $C(t)$ for each profile, with vertical lines marking G_trap and G_crit.
- *Phase diagram*: trajectory in the (G, β_effective) plane, showing the two-threshold structure and the irreversible crossing of G_trap.
- *Parameter-sweep heatmap*: time to dissolution (first crossing of G_crit) as a function of β_nominal and c_coupling, demonstrating the region of parameter space in which the two-threshold trap exists and the region in which the system can avoid it.

- *Acceleration scenario*: a second run in which α increases linearly from 0.03 to 0.10 over the last 50 time steps, simulating an AI-driven acceleration. The comparison shows which profiles survive the acceleration and which are pushed into the trap.

The critical finding, if it emerges, is the *transition-bandwidth trap*: a system that still functions operationally — G below G_crit, services delivered — but has already lost the capacity to redesign itself. The gap is growing, and the system cannot stop it. This is the dynamic condition that Sections 5.3 and 5.4 describe qualitatively; the simulation makes it quantitative and visually explicit.

6.4 Methodological Commitments

All simulations share a set of commitments that reflect the series' epistemic stance. The code is open-source, written in Python using standard scientific libraries (NumPy, SciPy, Matplotlib), and available in a companion repository. The manuscript references the repository and the specific commit hash used to generate the published figures. Parameters are declared in a single configuration file, and the simulation includes a reproducibility script that regenerates all figures from scratch.

Each simulation uses Monte Carlo replication: 100 independent seeds for Simulations A and C, and 50 seeds per grid cell for Simulation B (given the larger parameter sweep, to keep computation tractable). Results are reported as distributions (median, interquartile range) rather than single-run trajectories. Each simulation includes at least one parameter-sweep heatmap that demonstrates the robustness of the qualitative result across a plausible range of parameter values.

The simulations are not calibrated to empirical data. They are theoretical instruments whose parameters are illustrative rather than estimated. The paper does not claim that the specific numerical thresholds — the precise value of G_trap, the exact latency ratio at which oscillation occurs — transfer to any real governance system. The claim is that the *qualitative patterns* — the stable attractor under a permanent bypass, the phase transition around $\Omega = 1$, the two-threshold trap — are structural consequences of the modelled mechanisms and should be observable in any system that instantiates those mechanisms with sufficient fidelity. Whether real governance systems do instantiate them, and with what parameter values, is the empirical question the measurement framework of Paper VIII is designed to address.

Part VII concludes the paper by summarising its limitations, acknowledging the boundary conditions of the analysis, and identifying the open questions that define the research frontier. The transition from qualitative diagnosis to formal model is now substantially complete; the remainder of the paper specifies what the model cannot yet do, and what work remains for those who would test, refine, or replace it.

7. Limitations and Open Questions

The model developed in this paper extends the Governance as Engineering framework from the steady-state analysis of architectures to the dynamics of their replacement. That extension is ambitious, and it carries liabilities that should be stated clearly. A framework that does not acknowledge its boundary conditions invites overreach; one that specifies them invites correction, refinement, and eventual replacement by something better. This part identifies the principal limitations of the present analysis and the open questions they imply.

7.1 The Adversarial Extension is Heuristic, Not a Theorem

The central modelling move of Part II — treating incumbent resistance as an adaptive controller and the transition as a contested control problem — stretches the control-theoretic idiom beyond its classical domain. Ashby's Law of Requisite Variety, the foundational theorem on which the series rests, concerns a regulator attenuating exogenous disturbance variety. It does not, in its standard formulation, apply to strategic adversaries whose variety co-evolves with the regulator's own actions. The transition variety ratio Ω introduced in Section 2.2 is therefore a heuristic extension, not a theorem derived from first principles.

The paper has been explicit about this status. It has subjected the Ω condition to simulation testing (Simulation B) to determine whether it behaves as a threshold, and it has committed to revising the formal language of Section 2.2 based on the simulation results before the paper is finalised. But even if Ω proves robust as a diagnostic indicator across the simulated parameter space, it remains a heuristic. A full formal treatment would require apparatus from pursuit-evasion dynamics, differential game theory, or adversarial control — formalisms that lie beyond the scope of a paper that remains within the control-theoretic idiom the series has maintained. The paper provides a structural intuition with testable implications; it does not provide a proof. The distinction should be preserved in how the paper is read and cited.

7.2 Transition Bandwidth Proxies Require Empirical Calibration

Section 5.2 specified five observable proxies for transition bandwidth: constitutional amendment rate, sunset clause prevalence and credibility, experimental governance space density, deliberative infrastructure institutionalisation, and audit institution independence and longevity. These proxies are plausible correlates of the capacity for peaceful architectural self-redesign. They are not validated measures of it.

Validating them would require a programme of empirical work that this paper does not undertake: coding the five proxies across a sample of governance systems, constructing a composite index, and testing whether that index predicts observed transition outcomes — successful architectural reforms, stalled transitions, forced dissolutions — across a historical dataset. The estimation framework developed in Paper VIII provides a natural home for this work. Transition bandwidth could be integrated as a ninth structural primitive, or as a derived quantity computed from the existing eight. The present paper specifies the connection and leaves the integration to subsequent measurement research.

7.3 The Buy-Out Mechanism Assumes Fungible Interests

The buy-out protocols analysed in Section 4.3 and the non-compensability boundary identified in Section 4.4 rest on a distinction between economic interests (fungible, compensable) and identity-constitutive or ideological interests (non-fungible, non-compensable). That distinction is conceptually clear at the extremes: a slaveholder's financial interest in human property is categorically different from a theocrat's identity-constitutive investment in doctrinal authority. But the boundary between the two categories is not sharp in practice. Many real-world incumbents hold portfolios of interests that mix economic and identity components. A coal executive has a financial stake in the mine and a social identity bound up with the mining community. A mid-level bureaucrat in an authoritarian regime has a salary that could be compensated and a status that cannot.

The paper's prediction — that buy-out protocols succeed when incumbent interests are predominantly economic and fail when they are predominantly constitutive — is falsifiable in principle but operationally difficult to test, because coding the "predominant" character of an interest requires judgement that may be contested. The paper offers the prediction as a starting point for empirical work, not as a settled finding.

7.4 The Model Does Not Address Normative Legitimacy

The series has consistently distinguished between structural viability — whether an architecture can perform the functions it claims to perform — and normative legitimacy — whether it has the right to govern, grounded in the consent of the governed. This paper maintains that distinction. It analyses the conditions under which a reform coalition can overcome incumbent resistance and achieve architectural change. It does not analyse whether the reform coalition *should* win, whether its target architecture is democratically legitimate, or whether the transition process respects the rights and preferences of those who will live under the new architecture.

The distinction matters acutely in the context of the paper's own recommendations. Buy-out protocols that compensate slaveholders for the loss of their human property enabled a transition that is, by any reasonable moral standard, a clear improvement over the status quo ante. But the compensation was paid to the oppressors, not the oppressed, and the normative logic is uncomfortable. The paper's framework can identify the structural conditions under which such compensation is necessary to overcome a veto; it cannot adjudicate whether paying it is just. That question lies outside the engineering frame, and the paper does not pretend otherwise.

Similarly, the paper's analysis of observer diversity and protected experimental spaces is silent on whether the deliberative bodies, citizen audits, and independent watchdogs it recommends are themselves democratically legitimate — whether they represent the populations they claim to observe, or whether they concentrate epistemic authority in ways that are themselves unaccountable. These are genuine questions. The paper raises them as limitations not to dismiss them, but to mark the boundary where engineering analysis must hand off to political theory.

7.5 The Transition-Bandwidth Trap Requires Longitudinal Testing

The dynamic model of Section 5.3 predicts a two-threshold structure: a transition-bandwidth trap threshold *G*_trap, reached before the constitutional unobservability threshold *G*_crit, at which a governance system loses the capacity for peaceful self-redesign while still appearing operationally functional. Simulation C demonstrates this structure under illustrative parameter values. But the model is a minimal dynamical system with two equations; it is not a calibrated representation of any real governance system.

Testing whether the transition-bandwidth trap operates in practice would require longitudinal data on governance systems that have undergone architectural dissolution — historical cases where a polity's institutional structure was replaced through rupture rather than reform — and an assessment of whether the system exhibited the predicted signature: a period of apparently normal operational functioning during which the capacity for structural reform had already been lost. The country studies contain suggestive evidence (Japan's Continuity Trap, China's Calibration Deficit, the late Soviet Union's terminal rigidity), but systematic testing across a larger sample is required. The concept of *G*_trap is, at this stage, a hypothesis generated by the model, not an empirically confirmed phenomenon.

7.6 The Model Assumes a Unified Incumbent Controller

The incumbent controller **I** in Part II is modelled as a unitary actor with a coherent objective function and a unified actuation capacity. Real incumbent coalitions are fractious. The *Centrão* in Brazil is not a single actor but a shifting aggregation of hundreds of legislators with partially overlapping and partially competing interests. The extraction coalition in Nigeria contains military factions, political elites, and commercial networks whose alignments shift with each budget cycle. The Veto Industrial Complex in the United States is a loose confederation of institutional actors — congressional committees, regulatory agencies, interest groups, media ecosystems — that coordinate imperfectly and sometimes work at cross-purposes.

Treating the incumbent as a unitary controller simplifies the analysis considerably, and the simplification is defensible for a first-pass model. But it elides a dynamic that may be important in practice: incumbent fragmentation can create opportunities for reform coalitions to split the opposition, buying out one faction while isolating another. A richer model would treat the incumbent as a coalition with internal variety and its own coordination costs — a controller that is itself subject to the same structural constraints the series analyses. That extension is left to future work.

7.7 The Model is Linear Where Reality is Not

The state transition equation in Section 2.1 is linear and time-invariant. The dynamics of architectural change, in real governance systems, are neither. Institutional thresholds — the number of veto players that must defect before a reform becomes unstoppable, the level of public discontent at which an immune system can no longer contain the pressure — are nonlinear. The latency gain ceiling imported from Paper I is a linear approximation that captures first-order behaviour near equilibrium; it may not hold when reforms push the system far from its current operating point.

The simulation models in Part VI introduce some nonlinearity through the coupling of G to β_effective and the sigmoidal response of reform pressure to visible dysfunction. But they remain minimal models. Extending the framework to capture threshold effects, hysteresis, and cascading institutional failure would strengthen the dynamic analysis and is a significant direction for subsequent work.

These limitations define the research frontier. The paper's contribution is not to have settled the political economy of governance transitions, but to have made it amenable to the same formal analysis that the series has applied to governance architectures themselves. The traps are modelled, the design principles are specified, and the dynamic constraint — transition bandwidth — is identified. What remains is the empirical work that would transform these concepts from diagnostic scaffolds into operational instruments, and the theoretical work that would extend the formal apparatus to the nonlinear, multi-actor, and normatively complex settings that real transitions inhabit. Part VIII concludes the paper by restating the central argument and its place in the series' arc.

8. Conclusion

The Governance as Engineering series began with a simple observation: governance systems are feedback systems, and feedback systems have structural properties that determine their stability under disturbance. Those properties can be modelled formally, compared objectively, and improved through design. Ignoring them does not make them go away; it means their consequences are attributed to the wrong causes and addressed with the wrong interventions.

Papers I through VI established the structural constraints that any governance architecture must satisfy to be stable, adaptive, and observant of the systems it governs. Paper VII synthesised those constraints into a qualitative account of why reform so consistently disappoints, and identified the convergent first step — the protected experimental space — across fifteen jurisdictions and four continents. Paper VIII began the work of making the central diagnostic concept, the variety gap, measurable.

This paper completes a specific arc within that larger project. It takes the transition dynamics that Paper VII described qualitatively and models them formally. It treats the reform process not as a political story about will, leadership, and coalition-building — though it is also those things — but as a control problem in which two adaptive controllers with asymmetric capabilities act on the same institutional state vector. The central finding is that the feasibility of architectural governance reform is not primarily a function of political will, resources, or moral clarity. It is a function of the variety budget, latency structure, and gain ceiling of the transition pathway itself.

8.1 What the Paper Established

The paper makes five contributions to the series' formal architecture.

First, it models the transition as a contested control problem. The incumbent beneficiaries of the current architecture are not passive obstacles; they constitute an adaptive controller with their own observation channels, shorter latency than any external reform coalition can achieve, and an objective function — architecture preservation — that they pursue with the same rationality that any controller brings to its task. The reform coalition must satisfy a variety condition ($\Omega \geq 1$) and operate within the gain band that its latency imposes to have any prospect of success. These conditions are structural, not normative. They hold regardless of the justice of the reform's cause or the competence of its leadership.

Second, it formalises three structural traps that Paper VII identified qualitatively. The bypass trap is the stable low-performance attractor in which a parallel institutional channel relieves pressure on the unreformed substrate without repairing it, capping the bypass's own effectiveness at the level the substrate permits. The legibility problem is the constraint that the reform coalition must diagnose the architecture's dysfunction using observation channels the architecture itself provides — channels whose degradation is precisely what needs to be diagnosed. The incentive-compatibility trap is the principal-agent problem in which the actors who must implement architectural change are the ones who lose from it. Each trap is modelled as a dynamical system with an internal equilibrium and characteristic failure signature.

Third, it derives four design principles for transition mechanisms that can survive incumbent counter-mobilisation. Protected experimental spaces with coupled legibility create local instances where the observation channel is shorter and the signal is less degraded, and where the evidence generated is transmitted through channels the incumbent cannot easily capture. Sunset-coupled bypasses are parallel channels designed with a credible, pre-committed mechanism that transfers pressure back to the unreformed substrate as the bypass demonstrates superior performance. Incumbent buy-out protocols address the incentive-compatibility trap directly, compensating economic interests that would otherwise block reform — a mechanism with a specific domain of applicability and a clearly defined boundary where it fails. Instrumentalised observer diversity counters the legibility problem by constructing independent, decorrelated sensing systems whose outputs the incumbent cannot simultaneously control.

Fourth, it introduces the concept of transition bandwidth — the rate at which a governance system can peacefully redesign its own structure — and distinguishes it from operational bandwidth, the rate at which it can respond to disturbances within its existing structure. A system can have high operational bandwidth and near-zero transition bandwidth, and the historical record contains examples (Rome, the late Soviet Union, arguably contemporary China) where this asymmetry proved fatal. The paper specifies observable proxies for transition bandwidth, connecting the concept to the measurement framework of Paper VIII, and models the dynamic race between environmental change and architectural adaptation.

Fifth, it identifies the transition-bandwidth trap: a point of no return, reached strictly before the constitutional unobservability threshold, at which a governance system still functions operationally but has irreversibly lost the capacity for peaceful self-redesign. The trap is produced by a coupling between the variety gap and effective transition bandwidth — as the gap widens, crisis management consumes the institutional capacity that would be needed for structural reform, further reducing bandwidth, further widening the gap. The two-threshold structure (first the loss of redesign capacity, then the loss of observability) is, if confirmed empirically, the paper's most significant dynamic finding. It implies that governance architectures can be operationally functional while already being structurally dead — still running, but no longer able to steer.

8.2 The Series' Arc

The series has moved through three broad phases. The first four papers established the structural constraints on governance architectures across four domains: crisis response, multi-scale disturbance management, democratic representation, and commons governance. The fifth paper demonstrated that these constraints interact multiplicatively — the coordination failure tax. The sixth paper extended the framework upward, treating objective functions themselves as observation architectures and introducing the variety gap as a unifying diagnostic.

The seventh paper marked a pivot: from the analysis of architectures to the analysis of transitions. It drew on fifteen country studies to argue that the immune system defending the current architecture is not an external obstacle but an output of it, that the bypass trap and the legibility problem explain the persistent gap between reform ambition and reform outcome, and that the first viable step in every jurisdiction is the same — a protected space where the channel is shorter and the signal less degraded.

This paper, the ninth, completes the pivot. It provides the formal grammar for the transition dynamics that Paper VII described qualitatively. It models the incumbent as an adaptive controller, the traps as dynamical equilibria, and the design principles as structural devices for altering the variety and latency conditions of the transition pathway. And it introduces the temporal dimension that the qualitative account lacked: the race between adaptation and dissolution, and the transition-bandwidth trap that is lost before the system knows it is running.

The arc is not closed. The measurement work that Paper VIII began — the operationalisation of the variety gap, the estimation of structural parameters from observable governance characteristics — remains in its early stages. The empirical testing of the transition-bandwidth trap hypothesis requires longitudinal data and careful case selection that this paper has not undertaken. The normative questions about legitimacy, consent, and the distribution of transition costs remain outside the engineering frame. But the formal architecture is now substantially in place. The series has moved from diagnosis through design to the dynamics of adoption, and the tools it provides are open to inspection, replication, and challenge.

8.3 The Practical Implication

The paper's analysis has a practical implication that is uncomfortable but, if the analysis is correct, unavoidable. Most governance reform efforts fail not because they pursue the wrong objectives, are led by the wrong people, or lack sufficient resources, but because they are designed as parametric interventions within architectures that generate structural resistance. They change the settings without changing the system that determines the settings. The immune system absorbs them, the traps re-engage, and the reform dissipates when the political window closes.

The reforms that accumulate are those that change the architecture. They shorten the observation channel, reduce the representation chain, expand the dimensionality of the value function, protect the feedback loops, and build the institutional capacity for peaceful self-redesign. They are slower, less politically legible, and harder to narrate as dramatic victories. But they compound, just as the failures compound — slowly, invisibly, and then suddenly, when the accumulated evidence becomes undeniable.

The woman in the Zona Norte of Rio de Janeiro, with whom Paper VII opened and closed, does not need a theory of governance transitions. She needs the fragments of functional governance that already exist around her — the PIX payment system, the community health networks, the electoral infrastructure — to connect into an architecture that can accumulate what it builds rather than extracting it before it compounds. The theory this paper provides is not for her. It is for the actors who could build that architecture, if they understood the structural conditions under which their efforts would survive the immune system that awaits them. The conditions are demanding. They are also specifiable, testable, and capable of being designed into the transition mechanisms themselves.

8.4 The Invitation

The series has always been offered as an analytical tool, not a political prescription. It does not advocate for specific policies or institutional arrangements. It provides a formal language for comparing governance architectures the way engineers compare control systems — by their demonstrated performance under defined conditions. This paper extends that language to the process by which architectures change.

The extension is provisional. The transition variety ratio Ω is a heuristic; the buy-out boundary is a hypothesis; the transition-bandwidth trap is a prediction generated by a minimal model. Each of these can be wrong in specific, identifiable ways, and the paper has specified the tests that would show them to be wrong. If the contested-control model fails to produce a clean threshold in simulation, the formal language of Part II must be revised. If the buy-out prediction fails to distinguish historical cases, the non-compensability boundary must be redrawn. If the transition-bandwidth trap does not appear in systems known to have lost redesign capacity while remaining operationally functional, the coupling mechanism must be reconsidered.

The invitation is to perform those tests — to challenge the model, refine it, or replace it with something better. The series was built to be falsifiable. The parameters are explicit, the simulations are reproducible, and the claims are structured so that their failure would be informative. A framework that cannot be wrong cannot be improved. This one can.

What remains, after the testing and the refinement, is the question that the series has circled but never answered: whether the political will exists to build governance architectures that satisfy the structural constraints the series has identified. The framework cannot supply that will. It can only make the cost of its absence more legible. The coordination failure tax is being paid, continuously and invisibly, by every system that operates below requisite variety across multiple architectural dimensions simultaneously. The transition-bandwidth trap is waiting for every architecture that cannot redesign itself faster than its environment changes. The fragments of a better architecture already exist, distributed across jurisdictions and domains, waiting to be connected.

The engineering is now substantially in place. The rest is not engineering. It is the choice to begin.

Appendix A: Heuristic Derivation of the Transition Variety Condition

The transition variety ratio Ω introduced in Section 2.2 is a heuristic extension of Ashby's Law of Requisite Variety to the contested-control setting. This appendix restates the classical Ashby condition, identifies the structural differences that prevent its direct application to the transition problem, defines the terms of the heuristic extension, and specifies its limitations. The derivation is offered as a diagnostic scaffold, not as a theorem. Its status is explicitly provisional.

A.1 Ashby's Law in Classical Form

Ashby's Law of Requisite Variety states that a regulator **R** can maintain a system **S** within a desired set of goal states **G** only if the variety of the regulator equals or exceeds the variety of the disturbance **D** relative to the goal. Formally:

$V(\mathbf{R}) \geq V(\mathbf{D}) - V(\mathbf{G})$ (A.1)

where $V(\cdot)$ denotes variety, operationally defined as the logarithm of the number of distinguishable states (Ashby, 1956). The law is a theorem under specified conditions:

- **D** is exogenous: the disturbance distribution is independent of the regulator's actions.
- **D** is stationary: its variety is fixed, or varies only through processes external to the control loop.
- The regulator's response does not alter the disturbance distribution; it only attenuates the disturbance's effect on the system's trajectory through the state space.

Under these conditions, equation (A.1) is a necessary condition for stable regulation. A regulator that violates it will, with probability approaching one, encounter a disturbance state it cannot discriminate from other states requiring different responses, and the unabsorbed variety will appear as uncontrolled variance in the outcomes.

Paper VI extended this logic to value architectures, treating the value function as the regulator and the disturbance environment as **D**, yielding the condition $\dim(\mathbf{V}) \geq \dim(\mathbf{D}) - \dim(\mathbf{G})$ for static dimensionality. That extension preserved the classical assumptions: the disturbance environment, while open-ended in the long run, was treated as exogenous and non-strategic at any given moment.

A.2 Why the Classical Setup Does Not Apply

The transition problem differs from the classical regulation problem in three respects that are material to the applicability of equation (A.1).

First, the "disturbance" is not exogenous. The incumbent controller **I** is part of the system being regulated, and its counter-moves $\mathbf{u}_I(t)$ are a function of the reform coalition's own actions $\mathbf{u}_R(t)$. The disturbance distribution is therefore *endogenous* to the control loop: the reform coalition's strategy alters the incumbent's strategy, which alters the disturbance the reform coalition faces. This is a strategic interaction, not a one-sided regulation problem.

Second, the disturbance variety is not fixed. As the reform coalition develops new strategies, the incumbent develops new counter-strategies. The variety of $\mathbf{u}_I$ is a function of the variety the incumbent can generate adaptively, which expands in response to the reform's own variety. The two controllers are co-evolving, and the variety of each is a moving target.

Third, the reform coalition's actuation is partially mediated through the incumbent's institutional apparatus (the transmission matrix **M** of Section 3.3). This means the reform coalition does not have direct, unattenuated access to the state vector **X**; its effective actuation is filtered through an adversary, a condition with no analogue in classical regulation.

These differences place the transition problem in the domain of adversarial control, differential game theory, or pursuit-evasion dynamics. A full formal treatment in any of those frameworks requires apparatus — value functions over strategy spaces, Hamilton-Jacobi-Bellman-Isaacs equations, information structures with asymmetric observations — that lies beyond the scope of a paper that remains within the control-theoretic idiom the series has maintained. What follows is therefore a heuristic extension: an inequality that captures the structural intuition of the series while remaining testable in simulation.

A.3 Defining the Transition Variety Space

Let the effective variety of the reform coalition, denoted $\dim(\mathbf{R})$, be the number of independent dimensions along which the coalition can:

1. Observe the architectural state $\mathbf{X}(t)$ — the rank of its effective observation matrix $\mathbf{C}_R$, after accounting for incumbent-injected distortion.
2. Deliberate about and select distinguishable control actions — the number of orthogonal policy levers it can deploy independently.
3. Transmit those actions through channels not entirely controlled by the incumbent — the rank of its effective actuation matrix $\mathbf{B}_R \cdot \mathbf{M}$, where **M** accounts for incumbent-mediated attenuation.

In practice, dim(**R**) is not a directly observable integer. It is a latent quantity that must be estimated from proxies: the diversity of the reform coalition's institutional bases (legislative, judicial, civil society, sub-federal, international), the independence of its information channels from incumbent control, and the number of distinct veto points it can credibly threaten to override.

Let the effective variety of the incumbent controller, denoted dim(**I**), be defined analogously: the number of independent dimensions along which the incumbent can observe reform threats, mobilise counter-measures, and deploy them through the institutional levers it controls. The incumbent's embedding in the architecture it defends typically gives it:

- Higher-rank observation: access to internal administrative data, intelligence channels, and informal networks that the reform coalition cannot replicate.
- Shorter latency: the ability to deploy counter-measures — committee blocking, regulatory delay, narrative capture — within the same institutional apparatus that the reform must navigate slowly.
- Higher-dimensional actuation: control over legislative procedure, budgetary allocation, appointment processes, media ecosystems, and coercive apparatus simultaneously.

The ratio Ω = dim(**R**) / dim(**I**) is the transition variety ratio introduced in Section 2.2. When $\Omega < 1$, the incumbent can generate more independent counter-moves than the reform coalition can independently address; the reform's actuation space is, in the language of classical Ashby, of insufficient variety to absorb the variety of the "disturbance," even though the disturbance is not exogenous.

A.4 The Heuristic Inequality

Let **G**_transition be the set of architectural states that the reform coalition would accept as a successful transition. This is not a single point **X*** but a region in **X**-space: the coalition can tolerate deviation from the ideal along dimensions where compromise is necessary. The dimensionality of **G**_transition, denoted dim(**G**_transition), represents the "slack" available to the reform coalition — the number of independent architectural dimensions along which it can accept an outcome short of the ideal while still considering the transition a success. If the coalition's goal is a single, precisely specified architecture, dim(**G**_transition) = 0 and the variety requirement is maximal. If the coalition accepts a wide range of outcomes as satisfactory, dim(**G**_transition) is larger and the requirement is relaxed.

The heuristic extension of equation (A.1) to the transition setting is:

$$\dim(\mathbf{R}) \geq \dim(\mathbf{I}) - \dim(\mathbf{G}_{transition}) \qquad (A.2)$$

or equivalently:

$$\Omega \geq 1 - \dim(\mathbf{G}_{transition}) / \dim(\mathbf{I})$$

When dim(**G**_transition) = 0 — the coalition will accept only a specific architecture — the condition reduces to Ω ≥ 1. When the coalition's goal set is broad enough that dim(**G**_transition) approaches dim(**I**), the condition is satisfied even for small Ω, reflecting the intuition that a reform with very modest ambitions may face weaker effective resistance.

Equation (A.2) is not a theorem. It is a heuristic whose justification is analogical: the reform coalition must "regulate" the architecture into the goal set **G**_transition in the face of a "disturbance" — the incumbent's counter-mobilisation — whose variety is dim(**I**). The slack dim(**G**_transition) reduces the effective variety that must be matched, just as dim(**G**) does in the classical formulation. The analogy is structurally suggestive, and Simulation B is designed to determine whether the analogy holds — whether Ω = 1 behaves as a threshold in a minimal contested-control model. But the analogy is not a proof, and the inequality should be read as a diagnostic indicator, not as a necessary condition established by derivation.

A.5 Connection to Beer's Variety Engineering

The appropriate formal ancestor for this extension is not Ashby's original theorem but Stafford Beer's concept of variety engineering, developed in the Viable System Model (Beer, 1979, 1981). Beer recognised that in organisational settings, the "disturbance" confronting a manager is often another intelligent system — a competitor, a regulator, an adversarial subunit — and that the manager's unaided variety is typically insufficient for direct regulation. The solution, in Beer's framework, is the deliberate design of *variety amplifiers* and *variety attenuators*: devices that increase the effective variety of the controller's perception and action, or reduce the effective variety of the disturbance, until the requisite variety condition can be satisfied.

The design principles developed in Part IV can be interpreted through this lens:

- **Protected experimental spaces** are variety amplifiers for the reform coalition: they create local observation channels with higher fidelity and lower latency than the coalition can achieve at the national scale, amplifying its effective dim(**R**).
- **Sunset-coupled bypasses** are variety attenuators applied to the incumbent: they couple the bypass's success to increased pressure on the unreformed substrate, reducing the incumbent's ability to maintain the trap equilibrium without confronting the reform's evidence.
- **Buy-out protocols** are variety attenuators applied to the incumbent's incentive structure: they reduce the dimensionality of the incumbent's resistance by compensating specific interest dimensions, lowering effective dim(**I**).
- **Observer diversity** is a variety amplifier for the reform coalition's observation matrix: it adds decorrelated sensing dimensions that the incumbent cannot simultaneously capture, increasing the effective rank of **C**_R.

Beer's framework provides a vocabulary for these devices that is more precise than the looser term "design principle," and it anchors the transition analysis in an established cybernetic tradition. The paper's contribution is to apply that vocabulary to the specific structural traps identified in Part III and to specify the conditions under which each device is necessary.

A.6 Sufficiency and the Limits of the Heuristic

Even if equation (A.2) is satisfied — even if $\Omega \geq 1$ after accounting for goal-set slack — the reform coalition is not guaranteed success. The inequality captures a *dimensional* condition: whether the reform coalition has enough independent perceptual and actuation channels to match the incumbent's independent counter-moves. It does not capture three additional constraints that can cause a transition to fail even when Ω is favourable.

First, the *latency asymmetry* (Section 2.3): even if dim(**R**) is large, the reform coalition may be unable to actuate its control signals fast enough to overcome the incumbent's shorter-latency responses. The gain ceiling $\mathbf{K}_{max} \approx 1/(\tau_R \cdot |\mathbf{A}|)$ constrains the rate at which the reform can push architectural change, and the incumbent's shorter τ_I allows it to operate comfortably within its own ceiling while the reform is forced to the edge of its own.

Second, the *transmission constraint* (Section 3.3): some components of the reform's actuation must pass through the incumbent's institutional apparatus, and the transmission matrix **M** may attenuate them. Even if dim(**R**) is high, the *effective* actuation reaching the architecture may be of lower rank, and Ω calculated from nominal dim(**R**) may overstate the reform's true capacity.

Third, the *internal coordination costs* of the reform coalition itself: a coalition with high nominal variety — many independent member organisations, each with its own observation and actuation channels — may be unable to coordinate its actions into a coherent control signal. The variety that matters is *effective* variety: the number of independent dimensions along which the coalition can act *coherently*, not the number of members it contains. A fractious coalition with high nominal dim(**R**) may have low effective dim(**R**) once internal bargaining costs are accounted for.

These constraints mean that $\Omega \geq 1$ is best interpreted as a necessary condition for reform success under favourable latency and transmission conditions, and that its violation is sufficient (with high probability) for reform absorption. The condition is not sufficient for success. Sufficiency requires that the latency, transmission, and coordination constraints also be satisfied, and that the transition pathway avoid the three structural traps modelled in Part III.

A.7 Operationalisation

The dimensionalities dim(**R**), dim(**I**), and dim(**G**_transition) are not directly observable. They must be estimated from proxies, using the methodology developed in Paper VIII. For dim(**R**), candidate proxies include: the number of independent institutional bases from which the reform coalition can act (legislative seats controlled, sub-federal governments aligned, civil society organisations mobilised, international partners committed); the number of independent information channels available to the coalition that are not under incumbent control; and the number of distinct policy levers the coalition can credibly threaten to deploy. For dim(**I**), candidate proxies include: the number of veto points the incumbent controls; the diversity of its counter-mobilisation repertoire (legislative, judicial, media, coercive); and the independence of its various sensing networks. For dim(**G**_transition), the proxy is the breadth of the coalition's stated negotiating position: the number of architectural dimensions on which it has signalled willingness to accept compromise.

The estimation procedure yields values with confidence intervals, not point estimates, and the ratio Ω should be reported as a distribution — for example, $\Omega = 1.3 \pm 0.4$ — rather than as a scalar. This is consistent with the measurement discipline that Paper VIII established for the variety gap, and it preserves the appropriate epistemic caution for a quantity that is, at this stage, a heuristic diagnostic rather than a precisely measurable parameter.

The derivation presented in this appendix is offered in the spirit of the series: as a formal scaffold that makes structural intuitions explicit and testable. Whether the scaffold holds — whether $\Omega = 1$ genuinely demarcates a region of transition space where reform absorption becomes overwhelmingly probable — is a question for Simulation B and, ultimately, for empirical testing against historical transition episodes. The appendix provides the conceptual architecture; the simulation provides the first disciplined confrontation; the empirical work remains to be done.

Appendix B: A Minimal Model of Bypass-Trap Dynamics

This appendix presents the formal dynamical system underlying Simulation A (Section 6.1). It specifies the equations, identifies the system's fixed points, analyses their stability, and derives the bifurcation condition under which a sunset-coupled bypass escapes the low-performance attractor that traps a permanent bypass.

B.1 State Variables and Parameters

The system has two state variables evolving in discrete time:

- $D(t) \in [0, 1]$: dysfunction of the unreformed substrate, where 0 represents full functionality and 1 represents complete failure.
- $B(t) \in [0, B_{max}(D)]$: bypass capacity, bounded above by a substrate-dependent ceiling $B_{max}(D) = 1 - c_{cap} \cdot D$, with $c_{cap} \in [0, 1]$.

The substrate ceiling parameter c_{cap} captures the degree to which the bypass depends on the unreformed architecture. When $c_{cap} = 0$, bypass capacity is independent of substrate dysfunction; when $c_{cap} = 1$, a fully dysfunctional substrate reduces the bypass ceiling to zero.

The system is controlled by two effort variables:

- $R(t)$: reform pressure applied to the substrate.
- $R_B(t)$: investment in bypass expansion.

Additional parameters are:

- α: natural rate at which dysfunction deepens when unchecked (logistic growth parameter).
- K_R: gain of reform pressure with respect to visible dysfunction.
- c_{vis}: visibility coefficient — the degree to which the bypass masks substrate dysfunction from political observation.
- γ: effectiveness of investment in bypass expansion.
- δ: decay rate of bypass capacity (institutional entropy).
- c_{sunset}: coupling strength introduced by the sunset condition (Section B.4).

B.2 The Permanent Bypass System

We first analyse the system without sunset coupling, corresponding to the permanent bypass scenario.

State equations.

The substrate dysfunction evolves as:

$$D(t+1) = D(t) + \alpha \cdot D(t) \cdot (1 - D(t)) - \beta \cdot R(t) \qquad \text{(B.1)}$$

where β is the efficiency with which reform pressure reduces dysfunction. The logistic term $\alpha \cdot D \cdot (1 - D)$ captures the self-reinforcing character of institutional decay: moderate dysfunction worsens under its own weight unless counteracted by reform. The reform pressure term is:

$$R(t) = K_R \cdot D_{vis}(t) \qquad \text{(B.2)}$$

$$D_{vis}(t) = D(t) \cdot (1 - c_{vis} \cdot B(t)) \qquad \text{(B.3)}$$

Visible dysfunction D_{vis} is the portion of D that remains politically legible after the bypass masks it. When $B = 0$, visible dysfunction equals actual dysfunction. As B increases, D_{vis} falls, reducing R and therefore weakening the political pressure for substrate reform.

The bypass capacity evolves as:

$$B(t+1) = B(t) + \gamma \cdot R_B(t) \cdot (1 - B(t)/B_{max}(D(t))) - \delta \cdot B(t) \qquad \text{(B.4)}$$

where $R_B(t)$ is the investment flow into the bypass, which we treat as constant R_B in the permanent bypass scenario. The term $(1 - B/B_{max})$ captures diminishing returns as the bypass approaches its substrate-determined ceiling. The decay term $\delta \cdot B$ captures institutional entropy.

Fixed points.

A fixed point (D, B) satisfies $D(t+1) = D(t)$ and $B(t+1) = B(t)$. Substituting the equilibrium conditions into (B.1)–(B.4) yields:

$$\alpha \cdot D \cdot (1 - D) = \beta \cdot K_R \cdot D \cdot (1 - c_{vis} \cdot B) \qquad \text{(B.5)}$$

$$\gamma \cdot R_B \cdot (1 - B / (1 - c_{cap} \cdot D)) = \delta \cdot B \qquad \text{(B.6)}$$

Equation (B.5) has a trivial solution at $D = 0$ (full substrate functionality). For $D > 0$, we can divide through by D to obtain:

$$\alpha \cdot (1 - D) = \beta \cdot K_R \cdot (1 - c_{vis} \cdot B) \qquad \text{(B.7)}$$

Equation (B.6) can be rearranged to:

$$B = (\gamma \cdot R_B \cdot (1 - c_{cap} \cdot D)) / (\delta \cdot (1 - c_{cap} \cdot D) + \gamma \cdot R_B) \qquad \text{(B.8)}$$

The system thus admits a family of interior fixed points determined by the intersection of (B.7) and (B.8).

The trap equilibrium.

We are interested in the regime where the bypass is well-developed and the substrate is moderately to severely dysfunctional. Consider the case where c_{vis} is large — the bypass substantially masks dysfunction — and c_{cap} is moderate — the bypass is partly but not fully constrained by the substrate. Under these conditions, the system possesses a stable interior fixed point at some (D_{trap}, B_{trap}) with the following properties:

- $D_{trap} > 0$: the substrate remains dysfunctional.
- B_{trap} is substantial: the bypass provides meaningful service delivery.
- $D_{vis}(D_{trap}, B_{trap})$ is low: the political pressure for further reform is weak.
- $B_{trap} < B_{max}(D_{trap})$: the bypass operates below its theoretical ceiling, because investment stabilises once visible dysfunction falls.

This is the bypass trap: a stable equilibrium in which the bypass is successful enough to suppress reform pressure but constrained enough by the unreformed substrate that overall performance remains well below the level achievable under a genuinely reformed architecture.

Stability.

The Jacobian of the system (B.1)–(B.4) evaluated at the trap equilibrium determines local stability. Without reproducing the full algebraic derivation, the eigenvalues are typically inside the unit circle when α is moderate, c_{vis} is large, and $\gamma \cdot R_B$ is small relative to δ. The intuition is straightforward: a strong masking effect (c_{vis} large) ensures that an increase in B reduces D_{vis}, which reduces R, which allows D to drift upward — a negative feedback that stabilises the equilibrium. A weak bypass investment rate ensures that B does not grow rapidly enough to overwhelm the ceiling, preventing escape from the trap through pure bypass expansion.

B.3 Escape Through Exogenous Reform Pressure

Before introducing the sunset coupling, we note that the trap can also be escaped if an exogenous shock — a crisis, a scandal, an external intervention — temporarily increases the visibility coefficient c_{vis} or reduces the masking effectiveness of the bypass. This corresponds to a temporary shift in the parameter regime that destabilises the trap equilibrium, allowing a window during which reform pressure on the substrate can break the attractor. This is the formal analogue of the "policy window" concept: a crisis temporarily changes the structural parameters of the system, making reform possible that was not possible before.

However, reliance on exogenous shocks is not a design principle. A governance architecture that can only reform itself when crises force the issue is precisely an architecture with low transition bandwidth (Section 5.1). The sunset coupling of Section 4.2 is a mechanism for *endogenising* the shock — making the bypass's own success the trigger that forces substrate reform.

B.4 The Sunset-Coupled Bypass

We now introduce the sunset coupling. The coupling is a mechanism that increases the visibility coefficient c_vis as bypass capacity B grows, so that success does not suppress reform pressure but amplifies it. Formally, we replace the constant c_vis with a function:

$$c\_vis(B) = c\_vis^0 + c\_sunset \cdot g(B - B\_sunset) \qquad \text{(B.9)}$$

where $g(\cdot)$ is a smooth step function (e.g., a sigmoid) that is near zero when $B < B\_sunset$ and near one when $B > B\_sunset$. The parameter c_sunset is the additional visibility gained once the bypass crosses the threshold B_sunset, and c_vis^0 is the baseline masking coefficient that operates before the sunset triggers.

The modified reform pressure becomes:

$$R(t) = K\_R \cdot D(t) \cdot (1 - c\_vis(B(t)) \cdot B(t)) \qquad \text{(B.10)}$$

When $B < B\_sunset$, the system behaves identically to the permanent bypass case and may settle into the trap equilibrium if one exists. When B exceeds B_sunset, the effective visibility of dysfunction increases sharply. Reform pressure R rises, driving D downward. As D falls, the bypass ceiling $B\_max(D)$ rises, allowing B to expand further. The positive feedback loop — higher B → higher c_vis → higher R → lower D → higher B_max → higher B — propels the system out of the trap and toward a new equilibrium with low D and high B.

Bifurcation condition.

The sunset coupling produces a qualitative change in the system's dynamics when c_sunset exceeds a critical value c_sunset^crit. Below this threshold, the increase in visibility at the trigger point is insufficient to overcome the substrate's inertial dynamics, and the system returns to the trap equilibrium even after B crosses B_sunset. Above the threshold, the trap equilibrium is destabilised, and the system's only stable fixed point is the high-performance equilibrium (D low, B high).

The critical threshold is determined by the balance of forces at the trigger point. Let (D_trap, B_trap) be the trap equilibrium of the permanent bypass system. The sunset coupling alters the effective reform gain from $K\_R \cdot (1 - c\_vis^0 \cdot B)$ to $K\_R \cdot (1 - (c\_vis^0 + c\_sunset) \cdot B)$ for $B > B\_sunset$. The additional reform pressure at the trigger point is approximately:

$$\Delta R \approx K\_R \cdot D\_trap \cdot c\_sunset \cdot B\_sunset$$

For the trap to be destabilised, this additional pressure must be sufficient to reduce D below the level at which the trap's self-stabilising dynamics can restore it. A necessary condition is:

$$c\_sunset > (\alpha \cdot (1 - 2D\_trap) + \delta\_effective) / (K\_R \cdot B\_sunset)$$

where δ_effective captures the substrate's inertial resistance to change. The precise form depends on the parameterisation, but the structural point is that the sunset coupling must be strong enough — the additional visibility must be large enough — to overcome the trap's stabilising feedback.

Simulation A demonstrates this bifurcation by sweeping c_{sunset} and mapping the resulting equilibrium. The output shows the transition from a stable trap equilibrium at low c_{sunset} to a stable high-performance equilibrium at high c_{sunset}, with a critical region where the system's fate depends on the initial conditions and the precise timing of the trigger.

B.5 Design Implications

The formal analysis yields three conditions for a sunset-coupled bypass to succeed where a permanent bypass fails.

First, the sunset threshold B_{sunset} must be set *above* the level at which the bypass can be captured or reversed by the incumbent but *below* the level at which the bypass stabilises the trap equilibrium. If B_{sunset} is too low, the bypass triggers reform pressure before it has demonstrated sufficient performance to build a political coalition for substrate reform; the pressure dissipates, and the bypass is wound down or captured. If B_{sunset} is too high, the bypass reaches the trap equilibrium and stabilises there, never triggering the sunset at all.

Second, the coupling strength c_{sunset} must exceed the critical threshold c_{sunset}^{crit} for the given substrate parameters. This means the sunset mechanism must be *institutionally credible*: it must genuinely increase the visibility of remaining dysfunction, in a way the incumbent cannot neutralise. A sunset clause that the incumbent can capture — by redefining the trigger condition, by challenging the evidence of bypass performance, or by simply ignoring the legislative mandate — has $c_{sunset} \approx 0$ and will not escape the trap.

Third, the baseline parameters — particularly the substrate ceiling coefficient c_{cap} — must permit the bypass to reach B_{sunset} in the first place. If the substrate is so dysfunctional that the bypass ceiling $B_{max}(D)$ lies below B_{sunset} for all attainable D, the bypass can never trigger the sunset, and the system is consigned to the trap permanently. In such cases, architectural reform must target the substrate directly (through a crisis window or an exogenous shock) rather than through a bypass strategy.

B.6 Limitations

The model presented in this appendix is minimal. It abstracts away from several features of real bypass dynamics: the heterogeneity of the substrate (dysfunction is typically uneven across domains, and a bypass may relieve pressure selectively), the strategic behaviour of incumbents who may deliberately degrade the substrate to keep the bypass below B_{sunset}, and the possibility that the bypass itself becomes captured by the same interests that control the substrate. These extensions are tractable within the same modelling framework and represent directions for subsequent work. The present model captures the first-order dynamics that distinguish permanent from sunset-coupled bypasses, and it provides the formal underpinning for the design principle developed in Section 4.2.

Appendix C: Historical Calibration of Buy-Out Effectiveness

Section 4.3 introduced incumbent buy-out protocols as a design principle for addressing the incentive-compatibility trap, and Section 4.4 identified the non-compensability boundary that defines the mechanism's domain of applicability. This appendix provides structured historical narratives for five cases: four in which buy-out protocols enabled architectural transitions that would otherwise have been blocked, and one contrast case in which a transition succeeded without compensation. The cases are not a random sample — they are selected to illustrate the structural conditions the paper identifies — and the small-N character of the evidence is acknowledged. The purpose is calibration, not confirmation.

C.1 Case Selection and Methodology

The five cases span three centuries, four continents, and five distinct policy domains: the abolition of chattel slavery in the British Empire (1833), the commutation of feudal stipends in Meiji Japan (1876), the phase-out of ozone-depleting substances under the Montreal Protocol (1987), the coal phase-out in the Federal Republic of Germany (2020), and the elimination of agricultural subsidies in New Zealand (1984). Four are positive cases of buy-out protocols; the fifth is a contrast case in which no significant buy-out occurred.

For each case, we extract three structural parameters relevant to the theoretical framework of Part IV:

- **Compensation ratio**: the magnitude of the compensation package relative to the economic value of the interests being displaced, expressed as a fraction where estimable.
- **Implementation latency**: the time between the formal adoption of the transition and the point at which incumbent interests were substantially displaced.
- **Compensation dimensionality**: the number of independent dimensions of incumbent interest that the compensation package addressed (financial, status, regional employment, future optionality, etc.).

These parameters are not precise measurements. They are plausible order-of-magnitude estimates, derived from historical sources, and they are reported with the uncertainty that attends any retrospective coding of complex institutional episodes.

C.2 Structured Case Narratives

C.2.1 British Abolition of Slavery (1833)

Background. The British Empire abolished chattel slavery through the Slavery Abolition Act of 1833, following decades of abolitionist mobilisation. The Act applied throughout most of the Empire, with the principal exception of territories controlled by the East India Company. At the time of passage, approximately 800,000 enslaved people were held in British colonies, predominantly in the Caribbean.

Incumbent interests at stake. The slaveholding class constituted a concentrated, politically powerful incumbent group. Slaveholders controlled significant parliamentary representation through the West India Interest, a lobbying bloc that had successfully delayed abolition for decades. Their interests were predominantly economic: the value of enslaved people as property, the income streams from plantation agriculture, and the capital invested in land and processing infrastructure. Recent scholarship estimates the total market value of enslaved people in the British Caribbean at approximately £45–50 million at the time of abolition (Legacies of British Slavery database, UCL).

Compensation mechanism. The Act allocated £20 million in compensation — approximately 40 percent of the Treasury's annual expenditure — to be paid not to the enslaved but to the slaveholders. The compensation was structured as a loan raised in the bond market, serviced by general taxation, spreading the fiscal cost across generations. Individual claims were adjudicated by a Slave Compensation Commission, which processed approximately 46,000 claims between 1834 and 1843. The enslaved received no compensation and were required to serve a transitional "apprenticeship" period of four to six years before full emancipation.

Outcome. The Act passed Parliament and was implemented substantially as designed. The institution of chattel slavery was dismantled throughout the British Empire within a defined transition period. The compensation payments were made in full. The West India Interest, having been compensated for its losses, did not mount a serious attempt to reverse abolition.

Structural interpretation. The buy-out addressed the *economic* dimension of the incumbent's interests directly and at sufficient scale — roughly 40 percent of the estimated market value — to neutralise organised resistance. The incumbent's political capacity was concentrated enough that a single, centrally administered compensation package could reach it. The compensation dimensionality was one (financial), which matched the dimensionality of the incumbent's primary interest. The latency between formal adoption (1833) and full emancipation (1838) provided a transition period, but the compensation itself was committed upfront, signalling the government's irrevocable intent.

C.2.2 Meiji Stipend Commutation (1876)

Background. The Meiji Restoration of 1868 initiated a programme of rapid modernisation that required dismantling the feudal class structure of Tokugawa Japan. The samurai class, numbering approximately 1.9 million individuals including families, held hereditary stipends (*karoku* and *kamei*) that constituted the single largest item of government expenditure — estimated at between 30 and 40 percent of state revenue in the early Meiji period. These stipends were not salaries for services rendered but hereditary claims on the state's fiscal capacity, a direct inheritance of the feudal order.

Incumbent interests at stake. The samurai class held a portfolio of interests: the economic value of the hereditary stipends, social status tied to their formal rank, and the identity and meaning associated with their role as a warrior class. The economic interest was substantial and fungible. The status and identity interests were not — a point relevant to the non-compensability boundary discussed in Section C.4.

Compensation mechanism. The government commuted the hereditary stipends into interest-bearing government bonds through a series of measures culminating in the *Chitsuroku Shobun* (Stipend Commutation Act) of 1876. The bonds paid interest rates that varied by the rank of the original stipend holder, and were redeemable over a multi-year period. The commutation was compulsory — it was not a voluntary buy-out but a mandatory conversion of a permanent fiscal claim into a finite financial asset. However, the terms were sufficiently generous that the class as a whole acquiesced, and the measure freed approximately one-third of the state's fiscal capacity for industrial investment, infrastructure, and military modernisation.

Outcome. The commutation was implemented with minimal violence, an outcome that is historically remarkable. Comparable feudal abolitions in Europe — the French Revolution's abolition of seigneurial rights, the English Civil War's restructuring of royal and aristocratic fiscal claims — triggered civil wars. Japan's transition, while not entirely peaceful (the Satsuma Rebellion of 1877, led by disaffected former samurai, was a significant military challenge), succeeded without the protracted internal warfare that accompanied European feudal dissolution. The bonds were absorbed into the financial system, and many former samurai transitioned into commercial, bureaucratic, or military roles within the modernising state.

Structural interpretation. The buy-out addressed the *economic* dimension of the incumbent's interests directly and at sufficient scale to secure acquiescence from the majority. The compensation dimensionality was primarily one (financial), with a secondary dimension (the retention of some social status through the rank-graded bond rates) that partially addressed the identity component. That the Satsuma Rebellion occurred *despite* the commutation illustrates the limit: for a minority of samurai, the identity and status dimensions dominated the economic one, and no financial compensation could purchase their acquiescence. This is a within-case illustration of the non-compensability boundary.

C.2.3 Montreal Protocol on Substances That Deplete the Ozone Layer (1987)

Background. By the mid-1980s, scientific evidence had established that chlorofluorocarbons (CFCs) and related compounds were depleting the stratospheric ozone layer. The Vienna Convention of 1985 provided a framework for international cooperation; the Montreal Protocol of 1987 established binding phase-out schedules. The Protocol is widely regarded as the most successful international environmental agreement in history: ozone-depleting substance emissions have fallen by approximately 99 percent, and the ozone layer is projected to recover to 1980 levels by mid-century.

Incumbent interests at stake. The principal incumbent was DuPont, the dominant global producer of CFCs, with an estimated market share of 25–30 percent in the mid-1980s. DuPont's interests were economic: the revenue streams from CFC production, the capital invested in production facilities, and the future optionality of the chemical markets that CFCs served (refrigeration, air conditioning, aerosol propellants, and industrial solvents).

Compensation mechanism. The buy-out here was not a cash transfer but a *market redesign* that converted the incumbent from an opponent of regulation into a beneficiary of it. DuPont had invested significantly in research and development of CFC substitutes — hydrofluorocarbons (HFCs) and hydrochlorofluorocarbons (HCFCs) — and held key patents on these compounds. The regulatory phase-out created a guaranteed and rapidly expanding market for these patented substitutes, transforming a threatened revenue stream into a protected one. The Multilateral Fund, established in 1990, provided additional financial assistance to developing countries to support their transition, functioning as a secondary buy-out for a different class of incumbent.

Outcome. DuPont supported the Protocol and accelerated its own phase-out ahead of the treaty schedule, announcing in 1988 that it would cease CFC production entirely by 1999. The phase-out was achieved faster and at lower cost than initial projections had anticipated.

Structural interpretation. The compensation mechanism worked by *re-aligning* the incumbent's interests rather than merely compensating their loss. The dimensionality of the compensation was one (profit pathway), which matched the dimensionality of the incumbent's primary interest. The latency was determined by the phase-out schedule, which provided time for production facilities to be converted. The case demonstrates that buy-outs need not be direct fiscal transfers; they can be implemented through market design that creates alternative profit pathways for the affected incumbents. The condition for success is that the alternative pathway must be genuinely available — DuPont's prior investment in substitute R&D was a contingent historical fact, not a structural guarantee.

C.2.4 German Coal Phase-Out (2020)

Background. The *Kohleausstieg* (coal phase-out) legislation, adopted in 2020, committed Germany to closing all coal-fired power plants by 2038 at the latest, with a review clause allowing acceleration to 2035. The legislation followed decades of contested energy policy, including the *Energiewende* (energy transition) initiated in the early 2000s and the earlier decision to phase out nuclear power by 2022.

Incumbent interests at stake. Two distinct incumbent groups were affected. The corporate incumbents — RWE, LEAG, EnBW, and other operators of coal-fired generation capacity — stood to lose revenue streams from power generation and the asset value of their plants. The regional incumbents — the communities and workers in the Lusatian, Central German, and Rhenish lignite mining regions — stood to lose employment, economic base, and community identity tied to the coal industry.

Compensation mechanism. The legislation provided a multi-dimensional compensation package. For corporate incumbents, direct payments were authorised for foregone revenues from plants closed ahead of schedule, with the specific amounts determined through competitive auctions and bilateral negotiations. For regional incumbents, €40 billion in structural aid was allocated for investment in economic diversification, infrastructure, and community renewal in the affected regions. A Commission on Growth, Structural Change and Employment (the "Coal Commission"), comprising representatives from industry, labour unions, environmental organisations, and the affected regions, had negotiated the package in 2019, creating broad stakeholder buy-in before the legislation was finalised.

Outcome. The legislation passed with broad political support. Plant closures began on schedule. The structural aid has begun flowing to the affected regions, though the full effects on economic diversification will not be assessable for a decade or more. The corporate incumbents have largely accepted the terms, and no significant attempt to reverse the phase-out has emerged.

Structural interpretation. The compensation package is notable for its *multi-dimensionality*: it addressed both the financial interests of corporate incumbents (through revenue replacement) and the socio-economic interests of regional incumbents (through structural aid). This two-dimensional compensation matched the two-dimensional structure of the incumbent coalition, preventing the formation of a unified blocking front. The latency is long — closures extend to 2038 — which reduces the present value of the compensation but provides time for structural adjustment. The Coal Commission process functioned as an institutional mechanism for surfacing the dimensionality of the incumbents' interests, ensuring the compensation package was appropriately calibrated.

C.2.5 New Zealand Agricultural Reform (1984) — Contrast Case

Background. New Zealand's agricultural sector had been heavily subsidised since the 1960s, with support payments reaching approximately 30 percent of farm income by the early 1980s. The subsidies were fiscally unsustainable, environmentally damaging, and increasingly inconsistent with New Zealand's trade policy commitments. The 1984 currency crisis created a policy window: a snap election brought the Fourth Labour Government to power with a mandate for radical economic reform.

Incumbent interests at stake. The farming lobby — Federated Farmers, individual large-scale pastoral farmers, and the rural communities dependent on agricultural employment — constituted a significant political force. Their interests were economic: the income streams from subsidy payments and the asset values of farms that had been bid up under the expectation of continued support.

Compensation mechanism. There was essentially none. The government eliminated nearly all agricultural subsidies in a single budget, with minimal transition assistance and no significant compensation for the loss of support payments. The Reform was implemented immediately, with no phase-in period. Some limited adjustment assistance was provided to farmers in severe financial distress, but this was modest and was not structured as a buy-out for acquiescence.

Outcome. The reform succeeded in its primary objectives: agricultural subsidies were eliminated, production patterns shifted toward market-responsive output, and the sector's productivity improved substantially over the following decade. Many farmers experienced severe financial distress, and rural communities underwent significant adjustment costs, but the reform was not reversed. Farming organisations ultimately accepted the new regime and adapted to it.

Structural interpretation. The reform succeeded *without* a buy-out because the institutional structure made it possible to act faster than the incumbent could organise resistance. New Zealand's unicameral parliament, strong executive dominance, and the absence of federal veto points meant the reform could be legislated and implemented before the farming lobby could mobilise effective counter-measures. The latency τ_R was extremely short, and the incumbent's effective latency τ_I — the time required to organise counter-mobilisation — was longer. This is the condition identified in Section 2.3: when the reform coalition can act within a window during which the incumbent's counter-mobilisation capacity is temporarily disabled, a buy-out is not necessary. The case demonstrates the *alternative* to compensation: speed.

C.3 Comparative Parameters

The table below extracts the three structural parameters for each case. Values are approximate and should be treated as order-of-magnitude estimates, not precise measurements.

Case	Compensation ratio (approx.)	Implementation latency	Compensation dimensionality	Buy-out necessary?
British abolition (1833)	~40% of estimated asset value	4–6 years (apprenticeship)	1 (financial)	Yes — slaveholders held concentrated veto capacity
Meiji stipends (1876)	100% of income stream, capitalised as bonds	3–5 years (bond redemption)	1–2 (financial + partial status)	Yes — samurai class held military veto capacity
Montreal Protocol (1987)	Not a cash transfer; profit pathway via substitutes	9 years (CFC phase-out by 1996)	1 (profit pathway)	Yes — DuPont held technology and market veto capacity
German coal (2020)	Not a single ratio; revenue replacement + €40B structural aid	18 years (to 2038)	2 (corporate financial + regional socio-economic)	Yes — coal coalition held legislative and regional veto capacity
New Zealand agriculture (1984)	Near zero	Near zero	0	No — crisis window + unicameral system enabled faster action than incumbent could organise

The pattern is consistent with the theoretical framework. In each positive case, the incumbent possessed concentrated veto capacity that could have blocked reform, and the buy-out addressed the economic dimension of the incumbent's interests at sufficient scale to neutralise organised resistance. In the contrast case, the institutional structure made it possible to act faster than the incumbent could mobilise, rendering a buy-out unnecessary.

C.4 The Non-Compensability Boundary

Section 4.4 proposed a testable hypothesis: buy-out protocols succeed when incumbent interests are predominantly economic and fail when they are constitutive. The five cases above are consistent with this hypothesis — all four positive cases involve interests that were primarily financial or could be addressed through the creation of alternative financial pathways — but a sample of five cannot confirm a general claim. This section briefly examines two shadow cases that illustrate the predicted boundary.

The Satsuma Rebellion (1877) as a within-case boundary illustration. The Meiji stipend commutation was broadly successful in securing the acquiescence of the samurai class. But a significant minority, concentrated in the Satsuma domain, rebelled. The rebellion was led by Saigō Takamori, a former Meiji leader whose opposition was not primarily economic — Saigō had been offered and had refused positions in the new government — but ideological and identity-constitutive. The samurai who followed him were those for whom the *status* and *meaning* dimensions of their class identity dominated the financial dimension. The commutation could compensate their economic loss; it could not compensate the loss of a warrior identity. The rebellion was suppressed militarily, at significant cost, illustrating that when buy-outs fail because interests are constitutive, the fallback is not a more generous buy-out but coercion.

Contemporary shadow cases. The series' country studies identify several incumbent controllers whose interests are predominantly constitutive: the *siloviki* in Russia, whose power, status, and physical security depend on the continuation of the power vertical; the Chinese Communist Party's control preservation apparatus, whose survival logic is not economic but ideological-institutional; and theocratic veto players in several jurisdictions, whose political authority derives from doctrinal claims that no financial package can compensate. For these incumbents, the prediction of Section 4.4 is that buy-out protocols will fail regardless of scale. This prediction is not tested here — it is a hypothesis for future empirical work.

C.5 Limitations of the Analysis

The historical calibration presented in this appendix is subject to several limitations that should condition how it is interpreted.

Selection bias. The cases were selected to illustrate the structural conditions the paper identifies. They are not a random sample of attempted architectural transitions, and no inference about the *frequency* with which buy-out protocols succeed or fail can be drawn from them. A systematic test of the non-compensability hypothesis would require coding a representative sample of attempted transitions — including failures — and comparing buy-out outcomes against the economic-vs-constitutive classification of incumbent interests. That work remains to be done.

Measurement imprecision. The compensation ratio, latency, and dimensionality parameters are approximate. Historical data on the value of displaced interests is often contested (the market value of enslaved people in 1833 is an estimate, not a precise figure). The latency of implementation is complicated by the fact that effects phase in gradually rather than at a single point. Dimensionality is a qualitative judgement; reasonable coders might disagree about whether the Meiji bonds addressed one dimension or two. The parameters are offered as a disciplined way of making structural comparisons across cases, not as precise measurements.

Historical specificity. Each case is embedded in a unique institutional, cultural, and political context. The British abolition buy-out occurred in a parliamentary system with a restricted franchise; the Meiji commutation occurred in a revolutionary modernisation context with extraordinary executive authority; the Montreal Protocol succeeded in part because the number of relevant actors was small and the science was unusually clear. Whether the structural parameters extracted from these cases generalise to other settings is an open question.

Normative opacity. The table is structural, not normative. It identifies *that* the British abolition buy-out succeeded in neutralising incumbent resistance; it does not address *whether* compensating slaveholders was just. That question lies outside the engineering frame and is acknowledged, but not resolved, in the limitations discussion of Part VII.

The calibration in this appendix provides the historical grounding for the buy-out protocol analysis of Sections 4.3 and 4.4. The cases demonstrate that the mechanism can work under specifiable structural conditions; the shadow cases illustrate the boundary where it is predicted to fail. The empirical work required to transform these illustrations into tested propositions is substantial and is left to subsequent research.