

Governance as an Adaptive Controller

Exploration, Memory, and the Conditions for Institutional Learning

Paper XIV in the Governance as Engineering series

Models governance learning as a dual control problem, formalising the exploration–exploitation trade-off, institutional memory, and persistent excitation. Identifies five failure modes and derives design principles for adaptive governance architectures.

Completes the Cycle Two adaptation triad.

Björn Kenneth Holmström

June 2026

Creative Commons Attribution-ShareAlike 4.0 International

<https://bjorknethholmstrom.org/working-papers/governance-as-adaptive-controller>

Executive Summary

A governance system that cannot learn has a finite operational lifetime. The environment changes—new disturbance dimensions emerge, the governed population’s preferences and compliance behaviour evolve, the coupling structures that determine spillovers shift—and the architecture that was adequate for yesterday’s conditions becomes tomorrow’s constraint. This paper establishes the structural conditions under which a governance system can learn stably, and specifies the architectural principles required to sustain that learning over time.

The paper’s central contribution is to complete a triad that has been emerging across the second cycle of the Governance as Engineering series. Paper X established the *sensing* requirement: an observing ensemble must possess sufficient decorrelation to detect its own systematic errors. Paper IX established the *actuation* requirement: a system must possess sufficient transition bandwidth to execute structural change against incumbent resistance. This paper establishes the *learning* requirement that mediates between them: can the system discover what change to make, and do so without the exploratory process destabilising the regime it seeks to improve? The sequence—**Sense** → **Learn** → **Execute**—is the series’ answer to the question of how governance systems remain viable in environments that change faster than architectures can be redesigned.

The formal framework is built on dual control theory, which models the tension between *exploitation* (acting optimally given current knowledge) and *exploration* (acting to acquire better knowledge). In the governance context, every policy intervention is simultaneously an action and an experiment. A tax reform changes revenue *and* reveals the elasticity of taxable income. A regulatory change alters compliance *and* reveals the responsiveness of the regulated population. A governance system that treats its interventions only as actions is systematically discarding the information it needs to remain calibrated. The paper derives the dual control Bellman equation for governance, demonstrating that optimal policy includes an explicit exploration bonus—actions are tilted toward those that reduce uncertainty about parameters consequential for future performance.

The *persistence of excitation* condition from system identification gives rigorous content to the concept of antifragility: a system that never experiences stress cannot learn the parameters that determine its response to stress. A governance system that suppresses all variance—all protests, all policy failures, all external shocks—is not maximally stable; it is maximally fragile, because it has eliminated the signal on which model identification depends. Exploration cannot be episodic; it must be continuous and institutionalised.

Five characteristic failure modes follow from the framework. **Exploration starvation** occurs when short-term political incentives drive the exploration variance to zero: the system ceases to probe beyond its current model, the model drifts away from reality, and the resulting degradation is invisible to the system’s own monitoring. **Model lock-in** is the persistence of an obsolete model sustained by an institutional immune system that treats challenges as threats. **Exploitation lock-in** is the condition in which the system explores and learns adequately, but the actuation chain is blocked—knowledge accumulates while action does not

follow. **Learning-induced oscillation** arises when exploration is too aggressive, destabilising the very system the controller seeks to understand. The **forgetting-without-learning trap** occurs when institutional memory decays faster than new knowledge is acquired, so that each political transition resets the knowledge stock to a lower baseline.

A simulation of the adaptive controller demonstrates these dynamics under controlled conditions. The exploration-starvation trajectory shows the self-concealing nature of the trap: the controller's internal estimate of its own performance remains optimistic even as true tracking error diverges. The exploitation-lock-in trajectory shows parameter estimates tracking truth accurately while performance remains poor—the system knows what to do but cannot do it. The forgetting-without-learning sweep maps the threshold beyond which environmental change outpaces institutional memory.

Empirical illustrations span the full range of learning performance. China's reform era (1978–1990s) is the canonical dual control success: Special Economic Zones, the household responsibility system, and dual-track prices were structured explorations that generated information, which was then used to update the national model. The subsequent calibration deficit is a case of exploration starvation. Japan's Continuity Trap is a pure case of model lock-in: the post-war paradigm persisted for three decades after its predictive failure became evident, protected by an institutional architecture that had no mechanism for paradigm replacement. Finland's foresight infrastructure exemplifies strong sensing and learning functions but also demonstrates exploitation lock-in: the throughput constraint means knowledge accumulates faster than action follows. Nigeria exhibits the forgetting-without-learning trap: each reform cycle generates local learning that is not retained across administrations. The scientific community serves as an existence proof that sustained, institutionalised learning is architecturally achievable at scale.

Eight design principles follow from the diagnosis. **Protected experimental spaces** provide the persistent excitation that keeps the governance system's parameters identifiable. **Safe-to-fail structures** ensure that the failures inherent in exploration are survivable and informative rather than catastrophic. **Separation of exploration and exploitation functions** protects the learning apparatus from the short-term pressures that would extinguish it. **Protected curiosity budgets** provide dedicated resources for activities whose outcomes are uncertain, countering the institutional incentives that systematically penalise exploration. **Mandatory model review and paradigm replacement cycles** prevent model lock-in by ensuring that dominant policy models are periodically assessed by independent bodies with the authority to trigger their replacement. **Institutionalised forgetting**—through sunset clauses, zero-based budgeting, and automatic programme termination—enables the peaceful retirement of what no longer serves. **Learning rate accelerators**—data infrastructure, digital twins, embedded analytical capacity—increase the speed at which the system can acquire and process information. **Antifragility through stress exposure** ensures that the system's model is calibrated across the full range of conditions it may encounter.

The paper is the fourteenth in the series and consolidates the theoretical arc of Cycle Two. It completes the adaptation triad—sensing, learning, executing—and establishes the second-order cybernetic architecture that the series has been building toward: a controller that does not merely regulate the system, but regulates its

own regulation. The mathematics of this paper—dual control, exploration–exploitation, persistent excitation, catastrophic forgetting—are the same mathematics that govern recursive AI self-improvement. The problem of designing a governance system that can safely learn to modify its own architecture is structurally identical to the problem of designing an AI that can safely learn to modify its own code. The cybernetics of human governance and the cybernetics of machine intelligence converge on a single, unified problem.

The series has now completed its theoretical foundations. The grammar of primitives is established. The diagnosis of failure modes is developed across fifteen country cases and six organisational domains. The measurement framework exists in prototype. The theory of adaptation—sensing, learning, executing—is specified. The design vocabulary is articulated. What remains is to test the predictions, calibrate the tools, and build the institutions. The next phase is not more theory. It is engineering.

Part I — The Learning Limit

1.1 The Reform That Worked Until It Didn't

Between 1978 and the early 2010s, China executed the most successful sustained programme of deliberate institutional adaptation in modern governance history. The system did not achieve this by designing a perfect architecture in advance. It achieved it by treating governance as a continuous process of structural exploration.

The Special Economic Zones were not just policy initiatives; in control-theoretic terms, they were geographically bounded, protected experimental spaces designed to inject variance into the system without destabilising the core. The household responsibility system in agriculture was piloted locally, evaluated on its outcomes, and scaled only when the data proved its efficacy. The dual-track price system was an elegant solution to a fundamental control problem: it maintained the planned economy track (exploitation of the existing model for stability) while simultaneously allowing a market track to operate at the margins (exploration of a new model for growth). For three decades, the governance architecture acted as an optimal dual controller. It regulated the society to maintain stability, and it aggressively probed the environment to update its own structural parameters.

Then, the learning slowed. Around 2012, the architecture began to shift its priorities. The tolerance for local experimentation narrowed. The protected spaces were brought under tighter central control. The variance that is the mathematical prerequisite for learning was increasingly treated as a threat to systemic cohesion.

The calibration deficit diagnosed earlier in this series is, at root, a failure of adaptive capacity. The system that had been the world's most successful governance learner lost the willingness to sustain the experiments that learning requires. It prioritised exploitation over exploration. The performance of the system today—characterised by structural overshoots, rigid adherence to central directives, and abrupt, costly corrections—is not a failure of state capacity. It is the predictable behaviour of a controller that has stopped updating its internal model of the world.

The Chinese trajectory raises the question that this paper addresses: What conditions must a governance architecture satisfy for learning to be possible and sustainable? And why is learning so often a temporary phase that a system's own success or immune response eventually extinguishes?

1.2 The Pattern Across the Series

The Governance as Engineering series has repeatedly diagnosed systems that cannot learn. The inability to adapt is not a uniform pathology; it manifests in distinct structural failure modes across radically different political contexts.

Japan's Continuity Trap is a pure case of **model lock-in**. The post-war governance paradigm was an exceptionally well-designed controller for the demographic and economic environment of 1955 to 1985. However, as the environment changed—as the population aged and global supply chains shifted—the architecture retained the old model long past its correspondence to reality. The system did not lack the capacity to execute policy; it lacked the institutional mechanism to discard an obsolete paradigm.

Russia's Legibility Deficit is a case of **exploration starvation**. A power vertical that systematically punishes the reporting of novel, unapproved, or contradictory information is an architecture that actively destroys the variance it needs to identify its own parameters. It does not learn because its immune system treats the signals required for learning as treason.

Democratic systems exhibit a subtler variant of the learning failure. In Sweden's Drift Loop or the UK's Centralise-Fail-Centralise cycle, the system often perceives its own failures. The observation channel registers the error. But the system cannot translate that perception into architectural change because the immune system (Paper VII) absorbs the reform energy, or the short-term incentives of the electoral cycle make structural experimentation too politically costly. The system remembers its mistakes but cannot act differently.

There are exceptions. Finland's foresight infrastructure and Germany's Adaptive Governance Pilot Regions demonstrate that institutions can be designed to scan for novel dimensions and test new configurations. But these are islands of learning capacity within architectures that are not systematically engineered to evolve. They are operational patches on static machines.

1.3 The Structural Claim

A governance system that cannot learn has a finite operational lifetime.

This is not a political prediction; it is a mathematical certainty. The environment continuously generates novel disturbance dimensions—climate feedbacks, synthetic biology, algorithmic intelligence, demographic inversions. Because the environment is dynamic, the Variety Gap (Paper VI) between the world's complexity and the controller's dimensionality inevitably grows. An architecture designed for the disturbances of the twentieth century will eventually find itself operating in a state space it cannot perceive, applying control inputs to variables that no longer matter.

The only sustainable response is an architecture that can modify its own structure in response to what it discovers. It must be designed not merely to regulate within a known framework, but to explore its environment, retain what it discovers, discard what is no longer true, and rewrite its own control laws accordingly.

This paper formalises the conditions under which such modification is possible. It argues that adaptive capacity is not an accidental byproduct of good leadership, nor is it a cultural trait. It is an architectural property. A governance system can be explicitly designed to balance the exploitation of its current knowledge with the exploration of the unknown. It can be engineered to forget. It can be structured to safely absorb the persistent excitation—the continuous, low-level stress—that generates the information required to keep its internal models calibrated to reality.

1.4 The Three Sub-Problems and the Cycle Two Triad

To engineer a system that learns, we must decompose "learning" into three distinct but interacting control problems:

1. **Exploration vs. Exploitation:** The tension between acting optimally based on current knowledge (regulating the system) and acting sub-optimally to acquire better knowledge (probing the system).
2. **Institutional Memory and Forgetting:** The capacity to retain learned information across time, political transitions, and turnover—coupled with the equally vital capacity to discard models and datasets that no longer describe the environment.
3. **Antifragility as Persistent Excitation:** The maintenance of sufficient signal diversity to support ongoing model identification. A system that perfectly suppresses all variance destroys the friction that generates learning.

These three challenges have precise formal analogues in dual control theory and system identification. Synthesising them brings the foundational arc of this series' second cycle to a close.

Cycle Two has sought to answer how a governance system can adapt in real-time to a world that outpaces its original design. **Paper X** established the *sensing* requirement: an observing ensemble must possess sufficient diversity and decorrelation to detect its own systematic errors. **Paper IX** established the *actuation* requirement: a system must possess sufficient transition bandwidth to execute structural change against incumbent resistance.

This paper establishes the *learning* requirement. If a system can sense that it is failing, and if it has the bandwidth to change its architecture, how does it discover what it must change *into*? The triad is now complete: Observe the need, Learn the new parameters, Execute the structural change. This is the complete grammatical foundation of governance adaptation, and the gateway to the engineering applications of Cycle Three.

Part II — Formal Framework: Dual Control and the Exploration–Exploitation Trade-Off

The preceding papers in this series have treated governance as a control problem in which the system dynamics are known to the controller, at least up to statistically well-characterised noise. The controller observes the state, computes an optimal response, and applies it. The loop closes. Performance degrades when the observation channel is corrupted, the actuation chain is attenuated, or the boundary is mismatched—but the controller’s model of the system, in all of these analyses, is treated as given.

This paper relaxes that assumption. The controller does not know the system’s dynamics with certainty. It must learn them. And the actions it takes to learn may be different from the actions it would take if it already knew. This is the domain of dual control theory, and it is the formal home for the question this paper asks: can a governance system be designed to learn what it does not know, without the act of learning destabilising the system it seeks to govern?

2.1 Dual Control Theory

A standard feedback controller solves the *regulation* problem: given a model of the system, choose inputs that drive the state toward a target. The controller’s model—the matrices **A**, **B**, and **C**, the noise covariances **W** and **V**, the disturbance structure—is assumed to be accurate enough that the optimal policy computed from it is adequate. If the model is wrong, performance degrades, but the controller has no mechanism for detecting the wrongness or correcting it.

A *dual* controller solves two problems simultaneously. The first is the regulation problem: given the current best estimate of the system model, choose inputs that keep the state near the target. This is *exploitation*—making the best use of what the controller currently believes. The second is the *identification* problem: choose inputs that generate observations from which the model can be improved. This is *exploration*—acting to acquire better knowledge.

The two objectives are in tension. An input that is optimal for regulation given the current model may be uninformative for model improvement—it may repeat what the controller has already done, confirming the existing estimate without challenging it. An input that is informative for model improvement may be suboptimal for regulation—it may involve deliberate deviation from the certainty-equivalent action, introducing variance into the system’s trajectory in exchange for information about how the system responds.

The tension is not a design flaw. It is a structural feature of any controller that must learn while it acts. The optimal resolution, first characterised by Feldbaum in 1960–61, is a policy that balances the two objectives: the controller applies a control signal that includes both a certainty-equivalent component (the action that

would be optimal if the current model were correct) and an exploration component (a deliberate perturbation whose magnitude and direction are chosen to maximise the information acquired about the parameters that matter most for future decisions). The balance is dynamic: when the controller’s uncertainty is high, the exploration component is larger; when the model is well-established, the controller converges toward pure exploitation.

The governance analogue is direct. Every policy intervention is simultaneously an action and an experiment. A tax reform changes the tax code and, in doing so, reveals the elasticity of taxable income—a parameter that determines the revenue consequences of future rate changes. A regulatory change alters compliance behaviour and, in doing so, reveals the responsiveness of the regulated population—a parameter that determines whether stricter or looser regulation will be effective. A public investment project delivers infrastructure and, in doing so, reveals the state’s implementation capacity—a parameter that determines the feasible scale and pace of future projects.

A governance system that treats its interventions only as actions is systematically discarding the information they could provide. It is operating as a certainty-equivalent controller: acting as if its model of the economy, the population, and its own capacity were correct, and forgoing the opportunity to discover whether it is wrong. Over time, as the environment changes, the model drifts away from reality. The controller continues to apply interventions that were optimal for the world as it was, not for the world as it is. The performance degradation is gradual, and it is invisible to the controller’s own monitoring systems—because those systems are built on the same model that is drifting.

A governance system that treats its interventions as experiments, by contrast, designs them to yield information. It varies policy parameters across jurisdictions and tracks differential outcomes. It pilots programmes before scaling them, not primarily to reduce implementation risk but to measure the programme’s effectiveness. It maintains variation in its own operating procedures—different procurement models, different regulatory approaches, different service delivery mechanisms—not because it cannot decide which is best, but because it needs the variation to *discover* which is best as conditions change. This is not a luxury. It is the structural requirement for remaining calibrated to a changing environment.

2.2 The Dual Control Bellman Equation for Governance

The dual control problem can be stated formally. Let the governance system’s dynamics be

$$x(t + 1) = f(x(t), u(t), \theta) + w(t),$$

$$x(t + 1) = f(x(t), u(t), \theta) + w(t),$$

where $x(t)$ is the state vector (economic conditions, environmental quality, social indicators), $u(t)$ is the control vector (policy instruments, regulatory settings, budget allocations), θ is a vector of unknown parameters (policy multipliers, compliance elasticities, implementation capacities), and $w(t)$ is stochastic noise.

The controller does not know θ . It maintains a belief distribution $p_t(\theta)$ over the parameters, updated via Bayes' rule as observations accumulate:

$$p_{t+1}(\theta) \propto p_t(\theta) p(y(t) | x(t), u(t), \theta),$$

$$p_{t+1}(\theta) \propto p_t(\theta) p(y(t) | x(t), u(t), \theta),$$

where $y(t)$ is the observed outcome (which may differ from the true state due to measurement noise, as modulated by the observation-legitimacy parameter of Paper XIII).

The controller's objective is to minimise the expected cumulative discounted cost over a horizon T :

$$J = E \left[\sum_{t=0}^T \gamma^t c(x(t), u(t)) \right],$$

$$J = E[\sum_{t=0}^T \gamma^t c(x(t), u(t))],$$

where $c(\cdot)$ penalises deviations from the target state and excessive control effort, and $\gamma \in (0, 1]$ is the discount factor.

The optimal policy for this problem satisfies the Bellman equation:

$$V_t(b) = \min_u E_{x, \theta} [c(x, u) + \gamma V_{t+1}(b') | b, u],$$

$$V_t(b) = \min_u E_{x, \theta} [c(x, u) + \gamma V_{t+1}(b') | b, u],$$

where $b = (\hat{x}, p(\theta))$ is the belief state—the controller's best estimate of the system state and its uncertainty about the parameters. The expectation is taken over the true state x , the unknown parameters θ , and the stochastic noise, given the current belief.

The critical feature of this Bellman equation is that the choice of u affects not only the immediate cost $c(x, u)$ but also the *future belief state* b' —because the observation $y(t)$ that will be used to update $p(\theta)$ depends on the action taken. An action that produces a larger response—a larger signal-to-noise ratio in the system's output—provides more information about θ , reducing future uncertainty and enabling better future decisions. The optimal policy therefore includes an *exploration bonus*: actions are tilted toward those that promise to reduce uncertainty about parameters that are consequential for future performance.

This can be made explicit by decomposing the value function. Under certain approximations, the optimal control can be written as

$$u^*(t) = u_{CE}(t) + u_{explore}(t),$$

$$u^*(t) = u_{CE}(t) + u_{explore}(t),$$

where $u_{CE}(t)$ is the certainty-equivalent action—the action that would be optimal if the current parameter estimate $\hat{\theta}$ were the truth—and $u_{explore}(t)$ is a deliberate perturbation whose magnitude scales with the controller's uncertainty and with the sensitivity of future performance to the unknown parameters. When uncertainty is high, the exploration component is larger. When the parameters are precisely estimated, the exploration component decays toward zero and the controller becomes effectively certainty-equivalent.

The governance implication is that a well-designed learning system does not simply implement the policy that appears best given current knowledge. It deliberately varies its actions—across jurisdictions, across time, across policy domains—in ways that are informative about the parameters that matter most. The variation is not a concession to political compromise or administrative incapacity. It is the structural expression of the exploration bonus in the dual control objective.

2.3 The Exploration-Starvation Trap

A controller that solves the full dual control problem balances exploration and exploitation optimally, by construction. But real governance systems do not solve Bellman equations. They respond to political incentives, institutional pressures, and the cognitive limitations of the humans who operate them. And those incentives systematically penalise exploration.

Exploration involves variance. Trying something new—a different procurement model, a reformed regulatory approach, an experimental programme design—introduces the possibility of failure. In the short term, the expected performance of an exploratory action is usually worse than the expected performance of the known, certainty-equivalent action, because the exploratory action is not optimised for the current state. The benefit of exploration accrues in the future—in the form of better models, better calibrated interventions, and better outcomes down the line—but the cost is borne in the present.

A controller evaluated on short-term outcomes—an elected government facing the next election, an appointed official facing the next performance review, a minister defending the budget before parliament—will therefore tend to suppress exploration. The political cost of a failed experiment is immediate and visible. The political benefit of the knowledge gained is diffuse, delayed, and often attributed to the successor who implements the improved policy. The incentive gradient points toward certainty-equivalence: act as if the current model is correct, avoid variance, and let the future take care of itself.

The consequence is the *exploration-starvation trap*. The controller ceases to probe beyond its current model. It applies the same policy instruments in the same way, cycle after cycle, and observes outcomes that are consistent with the model—because the controller is not generating the variation that would reveal whether the model is wrong. The model drifts away from reality as the environment changes. The variety gap (Paper VI) widens. But the widening is invisible to the controller, because the controller has stopped generating the information that would detect it.

Performance begins to degrade—slowly at first, then more rapidly as the gap between the model and reality widens. The controller, observing the degradation, faces a cruel choice. It can explore—introduce variation, try new approaches, accept the risk of visible failure—at precisely the moment when its political capital is most depleted by the deteriorating outcomes. Or it can double down on exploitation—apply the existing model more aggressively, tighten the existing instruments, demand more effort from the existing institutions—and hope that the degradation is temporary.

The trap closes. The system that most needs to learn is the system least able to afford the experiments that learning requires. Exploration is deferred until the next crisis, the next administration, the next budget cycle. The model continues to drift. The degradation continues. Eventually, the gap breaches a crisis threshold—the financial system collapses, the pandemic overwhelms the health system, the environmental degradation becomes irreversible—and the system is forced to learn all at once, under the worst possible conditions, with depleted legitimacy and diminished capacity.

The exploration-starvation trap is not a hypothetical. It is the structural logic behind the late Soviet Union’s inability to perceive its own economic stagnation, behind the persistence of failed drug policies across decades and jurisdictions, behind the repeated failure of financial regulatory models to anticipate systemic crises, and behind the calibration deficit that the series diagnosed in the Chinese governance system after 2012. In each case, the system had the formal capacity to learn. What it lacked was the institutionalised protection for the exploration that learning requires.

2.4 The Persistence of Excitation Condition

The dual control framework identifies *that* exploration is necessary. System identification theory specifies *how much* exploration is necessary for learning to be possible.

In the standard formulation, the parameters of a linear system can be estimated from input-output data only if the input signal is *persistently exciting*. Formally, a signal $u(t)$ is persistently exciting of order n if there exist $\alpha > 0$ and an integer m such that, for all t ,

$$\alpha I \preceq \sum_{k=t}^{t+m} \phi(k)\phi(k)^\top,$$

$$\alpha I \preceq \sum_{k=t}^{t+m} \phi(k)\phi(k)^\top,$$

where $\phi(t)$ is the regressor vector constructed from past inputs and outputs. The condition ensures that the input varies sufficiently—in amplitude, frequency, and direction—to excite all the modes of the system, making it possible to uniquely determine the parameters that govern each mode.

If the input is constant, or varies only within a narrow band, the matrix on the right-hand side becomes rank-deficient: some parameters cannot be estimated from the available data, no matter how long the observation window. The controller can observe the system indefinitely and never learn the parameters that

determine its response to conditions it has never encountered.

The governance analogue is direct and consequential. A governance system that only ever does what it already knows how to do—that applies the same policy instruments at the same settings, year after year—is generating an input signal of insufficient variety to identify its own operating parameters. It cannot learn the elasticity of taxable income if it never varies tax rates. It cannot learn the effectiveness of different pedagogical approaches if it never varies curriculum or teaching methods. It cannot learn the capacity of its own implementation chain if it never attempts projects of different scales or complexities. It cannot learn the responsiveness of the regulated population if it never varies the stringency or the enforcement style of regulation.

The persistence of excitation condition gives rigorous content to the concept of antifragility that has been invoked, often loosely, in governance discourse. A system that never experiences stress cannot learn the parameters that determine its response to stress. A system that suppresses all variance—all protests, all policy failures, all external shocks—is not maximally stable; it is maximally fragile, because it has eliminated the excitation on which model identification depends. The system's apparent stability is the stability of a controller that is operating on a model that has never been challenged—a model whose correspondence to reality is unknown and, because the excitation has been suppressed, unknowable.

The design implication is that exploration cannot be episodic. It cannot be something the system does only when a crisis forces it, or only when a reform-minded leader happens to be in power. It must be continuous and institutionalised—built into the architecture as a permanent feature of the control loop, protected from the short-term incentives that would extinguish it. The controller must maintain a *persistent excitation signal*: a sustained, deliberate programme of experimentation, variation, and exposure to novel conditions that keeps the system's parameters identifiable.

This is the structural role of the protected experimental spaces that the series has identified, across multiple papers and multiple country cases, as the convergent first step of viable reform. The municipal laboratory, the sandbox state, the Special Economic Zone, the pilot programme with randomised evaluation—each is a mechanism for injecting persistent excitation into the governance system's input signal. Each generates the variation that makes learning possible. And each is vulnerable to the exploration-starvation trap: when budgets are tight, when political pressure mounts, when the existing model appears to be working adequately, the experimental spaces are the first things to be cut. The persistence of excitation condition explains why cutting them is not a harmless efficiency measure but the gradual self-blinding of the governance system—the quiet elimination of the signal on which its continued viability depends.

The remainder of this paper is about what happens when exploration is sustained, what happens when it is starved, and how to design architectures that keep it alive. The simulation of Part IV demonstrates the dynamics. The empirical illustrations of Part V ground them in cases. The design principles of Part VI specify the institutional machinery. But the formal core is here: every policy intervention is an experiment, whether

the controller acknowledges it or not; the controller that designs its interventions to be informative survives; the controller that suppresses the information in its own actions eventually discovers, too late, that it has been governing a phantom.

Part III — Failure Modes of Adaptive Governance

The formal framework of Part II establishes that a governance system must balance exploration and exploitation to remain calibrated to a changing environment. When that balance is lost—through institutional incentives, political pressure, or architectural design—the system enters one of several characteristic failure modes. These modes are not independent. They interact, and a system can exhibit more than one simultaneously. But each has a distinct mechanism, a distinct signature in the system's behaviour, and a distinct set of design responses. This part identifies five such modes.

3.1 Exploration Starvation

The most fundamental failure mode of adaptive governance is the simplest: the system ceases to explore. It applies the same policy instruments at the same settings, cycle after cycle, and observes outcomes that are consistent with its model—because it is not generating the variation that would reveal the model's deficiencies.

The mechanism is the one formalised in Section 2.3. Exploration involves variance, and variance is costly in the short term. A controller evaluated on short-term outcomes—an elected government facing an election, an appointed official facing a performance review, a minister defending a budget—faces an asymmetric incentive structure. The costs of a failed experiment are concentrated, visible, and attributable. The benefits of the knowledge gained are diffuse, delayed, and often captured by a successor. The rational response, under these incentives, is to suppress exploration.

The suppression is rarely explicit. No government announces that it has decided to stop learning. Instead, exploration budgets are trimmed as "efficiency savings." Experimental programmes are consolidated into standard operating procedures. Variation across jurisdictions is eliminated through standardisation and harmonisation. Pilot programmes are evaluated on whether they "succeed" rather than on what they reveal, and programmes that produce ambiguous or unfavourable results are not published, not because of active suppression but because ambiguous results are not career-advancing for the officials who produce them.

The consequence is gradual and self-concealing. The variety gap—the mismatch between the effective dimensionality of the environment and the effective dimensionality of the controller's model (Paper VI)—grows. But the growth is invisible to the controller, because the controller has stopped generating the information that would detect it. The system's dashboards continue to show acceptable performance, because the dashboards are built on the same model that is drifting. The controller's confidence in its own competence remains intact, because the evidence that would challenge it is not being generated.

The trap closes when the accumulated model drift breaches a crisis threshold. The financial regulator whose models never anticipated a systemic crisis because it never varied its regulatory approach enough to discover the fragility. The health system whose pandemic preparedness plans were calibrated to a transmission model that had never been tested against a novel pathogen because it had never conducted the exercises that would have revealed the model's inadequacy. The economic policymaker whose understanding of the labour market was based on relationships estimated during a period of stable inflation and demographic growth, and that ceased to hold when both changed.

At the moment of crisis, the system is forced to learn all at once—under the worst possible conditions, with depleted legitimacy, diminished capacity, and the added burden of managing the crisis itself. The exploration that should have been sustained across decades is compressed into months. The resulting learning is rushed, noisy, and often deeply costly in human terms. And as the crisis recedes, the incentives that produced the original exploration starvation reassert themselves. The cycle repeats.

The late Soviet Union is the canonical case. The planning apparatus maintained an increasingly elaborate model of the economy whose correspondence to reality deteriorated steadily from the 1960s onward. The model could not detect the deterioration because the system that generated the data—the enterprise reports, the output statistics, the price signals—was the same system that the model described. There was no independent variation, no persistent excitation, no experimental space where alternative approaches could be tested. When the gap between the model and reality finally breached the crisis threshold in the 1980s, the system had no reserves of knowledge about how a market economy might function, no cadre of officials trained in market regulation, and no legitimacy to sustain the transition. The exploration that had been suppressed for decades was demanded all at once, and the system could not supply it.

3.2 Model Lock-In

Exploration starvation is the failure to acquire new knowledge. Model lock-in is the failure to act on knowledge that has been acquired—or, more precisely, the failure to discard a model that accumulating evidence has falsified.

The mechanism is institutional rather than informational. The system may possess ample evidence that its current model is inadequate. The evidence sits in evaluation reports, in audit findings, in the divergence between forecast and outcome, in the complaints of frontline practitioners and the grievances of affected populations. But the institutional mechanisms for translating evidence into model revision are blocked. The immune system (Paper VII) treats challenges to the dominant model as threats to be neutralised. The institutions that control the model—the central bank that operates the forecasting framework, the ministry that designed the programme, the professional body that certifies practitioners in the standard methodology—have both the incentive and the capacity to defend it.

The defence takes many forms, not all of them cynical. Disconfirming evidence is scrutinised more aggressively than confirming evidence—an asymmetry that is cognitively natural and institutionally reinforced. The standards of proof required to overturn an established model are set higher than the standards required to maintain it. Methodological critiques are deployed selectively: the same statistical objection that would be fatal to a challenge is dismissed as a minor technical issue when applied to the incumbent model. Practitioners who question the model find their work harder to publish, their careers slower to advance, their access to decision-makers more limited. The model does not merely describe reality; it structures the professional and institutional ecosystem that sustains it.

The result is that the model persists beyond its correspondence to reality. The persistence is not indefinite—eventually, the gap between the model's predictions and observed outcomes becomes so large that it cannot be defended—but it can extend for decades, and during those decades the system is systematically miscalibrated. The interventions it applies are optimal for the world described by the model, not for the world that exists.

The series has documented several instances. Japan's post-war governance paradigm—export-led growth, bureaucratic guidance, the Iron Triangle, the 1955 System—was a coherent and successful model for the conditions of 1955–1985. It ceased to correspond to reality as demographics shifted, as the financial bubble burst, and as global production networks reorganised around China and Southeast Asia. But the model was embedded in the institutional architecture: the career structures of the bureaucracy, the committee system of the Liberal Democratic Party, the relationship between banks, firms, and regulators. Challenges to the model were challenges to the interests those institutions served. The model persisted through three decades of stagnation while the evidence of its inadequacy accumulated in economic statistics, demographic projections, and the visible decay of the social contract it was supposed to sustain.

The pre-2008 efficient markets hypothesis in financial regulation is another case. The model that asset prices reflect all available information and that financial innovation distributes risk efficiently was embedded in the institutional architecture of central banks, regulatory agencies, and international financial institutions. Evidence of its inadequacy accumulated across repeated financial crises—the 1987 crash, the Asian financial crisis, the LTCM collapse, the dot-com bubble—but each was treated as an exception, a one-off failure of implementation rather than a challenge to the model. The 2008 crisis was large enough to breach the immune system's defences, but only after the model had shaped a generation of deregulatory policy whose costs continue to be borne.

The pre-pandemic assumption that respiratory pathogens spread primarily by droplets, and that airborne transmission is negligible outside a narrow set of aerosol-generating procedures, is a third case. The evidence base for airborne transmission was substantial before 2020, but the institutional architecture of infection control—the guidelines, the training, the procurement of personal protective equipment, the physical design of healthcare facilities—was built around the droplet model. Changing the model would have required changing all of these, at enormous cost and against the resistance of the institutions that had built them. The model persisted until the pandemic made it untenable.

Model lock-in is not a pathology of authoritarian systems or captured regulators. It is a structural vulnerability of any governance architecture that concentrates model authority in the same institutions that implement the model. The separation of exploration and exploitation functions—having different institutions responsible for challenging models and for applying them—is the design response, and it is developed in Part VI.

3.3 Exploitation Lock-In

Exploitation lock-in is distinct from both exploration starvation and model lock-in, and it completes the taxonomy of learning failures. The system explores adequately. It learns. Its model updates. The knowledge is available. But it never reaches actuation. The system becomes a repository of unimplemented learning.

The mechanism is a specific configuration of the series' architectural primitives. The sensing function is intact: the observation channels are adequate, and the system perceives the need for change (Paper X). The learning function is intact: the system can discover what the change should be (Paper XIV). But the actuation function is blocked: the transition bandwidth (Paper IX) is insufficient to translate knowledge into action against the resistance of the immune system (Paper VII). The bottleneck is not epistemic but political. The system knows what to do but cannot do it.

Exploitation lock-in is the condition that separates the successful adaptive governance cases in the series from the frustrated ones. China's reform era succeeded because the Deng-era architecture provided not only the experimental spaces for learning (the SEZs, the household responsibility pilots) but also the political mechanism for scaling successful experiments into national policy—the transition bandwidth to act on what was learned. The subsequent calibration deficit is partly exploration starvation, as Section 3.1 argues, but it is also exploitation lock-in: the system continues to accumulate knowledge through its remaining observation channels—the Centre knows that local government debt is accumulating, that environmental degradation is accelerating, that demographic pressures are building—but the actuation mechanisms that could respond to that knowledge are blocked by the same Control Preservation Imperative that Paper VII diagnosed.

Finland's foresight infrastructure is the cleanest illustration of exploitation lock-in as a distinct failure mode. The Committee for the Future, the national foresight reporting system, the Futures Impact Assessments—these are institutionalised mechanisms for sensing and learning that are among the most advanced in any governance system. They generate high-quality knowledge about emerging disturbance dimensions, long-run demographic and ecological trends, and the likely consequences of current policy trajectories. The knowledge is publicly available, legislatively mandated, and institutionally prestigious. But the throughput constraint diagnosed in the Finland country study means that this knowledge systematically fails to translate into policy action at the pace and scale that the knowledge itself suggests is necessary. The system is not failing to sense. It is not failing to learn. It is failing to actuate. The knowledge accumulates in reports while the environment continues to change.

Exploitation lock-in is the failure mode most characteristic of democratic systems with high epistemic capacity but fragmented or veto-constrained actuation. The United States exhibits it across multiple domains: the knowledge that carbon emissions must be reduced has been available for decades; the political mechanism for implementing the reduction remains blocked. The knowledge that infrastructure investment is needed is near-universal; the authorisation and appropriation machinery cannot deliver it at the required scale. The knowledge that the healthcare system produces worse outcomes at higher cost than peer systems is well-documented; the actuation architecture cannot translate that knowledge into system redesign.

The formal structure of exploitation lock-in is a decoupling of the learning and actuation functions. In the dual control framework, the optimal policy combines the certainty-equivalent action with an exploration component. Both are passed to the actuator. If the actuator is attenuated—by the immune system, by veto points, by legitimacy deficits—the exploration component may survive while the certainty-equivalent component is blocked. The system learns but does not implement. The knowledge accumulates; the action does not follow.

3.4 Learning-Induced Oscillation

The preceding failure modes describe systems that explore too little, update too slowly, or actuate too weakly. Learning-induced oscillation is the opposite pathology: the system explores too aggressively, and the exploratory perturbations themselves destabilise the regime the controller seeks to improve.

The mechanism is the governance analogue of a system identification failure. In adaptive control, if the exploration signal is too large, too strongly correlated across inputs, or applied at the wrong frequency, it can excite system modes that obscure the parameters of interest. The controller observes the resulting oscillations and updates its model, but the model now reflects the system's response to the exploration signal rather than the system's underlying dynamics. The controller's subsequent actions, based on this mis-estimated model, amplify the oscillation rather than dampening it.

In governance, this occurs when a regime undertakes large, simultaneous, and correlated reforms—"big bang" transitions, revolutionary restructuring, comprehensive overhauls of multiple policy domains at once. The reforms generate massive variation in the system's state, but the variation is so large and so correlated that it becomes impossible to disentangle the effects of any individual reform from the effects of the aggregate shock. The system's response is dominated by the transitional dynamics—the confusion, the resistance, the unintended interactions—rather than by the steady-state parameters the reforms were intended to estimate.

The Chinese Campaign–Overshoot–Abrupt Correction cycle is a partial instance. The central government launches a policy campaign—poverty alleviation, environmental enforcement, pandemic control—with ambitious targets and strong political incentives for local officials to demonstrate compliance. The campaign generates a burst of exploration: local officials try different approaches, some of which work and some of which produce disastrous side-effects. But the campaign's intensity is so high, and the incentives for reported success so strong, that the signal returning to the centre is corrupted. The centre observes apparent success—

targets met, problems solved—and draws conclusions about the policy's effectiveness that are systematically over-optimistic. When the campaign eventually overshoots—when the accumulated side-effects, the unreported problems, and the exhausted compliance capacity breach the suppression apparatus—the correction is abrupt. The policy is reversed, often with minimal transition planning, and the system oscillates from aggressive exploration to defensive retrenchment. The learning that could have been extracted from a more measured, sustained programme of experimentation is lost in the oscillation.

The more extreme historical cases—the French Revolution, the Great Leap Forward, the shock therapy transitions of the post-communist 1990s—exhibit the same structure at larger amplitude. In each, a regime undertook an exploratory perturbation so large that the system's dynamics were overwhelmed before any learning could be extracted. The oscillation that followed—Thermidor, the post-Leap famine, the political backlash against market reform—was not a failure of learning but a consequence of learning attempted at a scale and speed that the system could not absorb.

The design implication is that exploration must be *bounded* as well as persistent. The exploration signal must be large enough to excite the system's relevant modes but small enough that the system's response remains in the linear regime where parameters are identifiable. This is the structural role of the protected experimental space—not merely as a mechanism for generating variation, but as a mechanism for *containing* it. The SEZ confines the exploration to a bounded geography. The pilot programme confines it to a bounded population. The phased reform sequence confines it to a bounded set of policy domains. Each of these is not a concession to political caution but a structural requirement for stable learning. The exploration must be safe-to-fail before it can be safe-to-learn.

3.5 The Forgetting-Without-Learning Trap

The final failure mode concerns not the acquisition of knowledge but its retention. Institutional memory decays. Personnel turn over. Political administrations change. Organisational structures are reorganised. Records are lost, archived inaccessibly, or never created in the first place. If the rate of forgetting exceeds the rate of learning, the system does not merely fail to improve. It actively loses capability over time.

The formal analogue is a recursive estimator with a forgetting factor. In the Kalman filter or recursive least squares, a forgetting factor $\lambda_f \in (0, 1]$ is applied to down-weight past observations, allowing the estimator to track slowly changing parameters. The effective sample size is bounded above by $1/(1 - \lambda_f)$. When $\lambda_f = 1$, no forgetting occurs; the estimator accumulates information indefinitely. When λ_f is small, the effective memory is short, and the estimator can never accumulate sufficient evidence to identify stable parameters with precision.

Governance systems exhibit a forgetting factor that is not chosen by a designer but determined by institutional structure. Democracies with short electoral cycles, high ministerial turnover, and frequent agency restructuring have a naturally small λ_f . Each new administration arrives with its own priorities, its own

models, its own advisors. The institutional memory of the previous administration—the evaluations, the programme data, the tacit knowledge of what worked and why—is not transferred systematically. It is discarded, deliberately or through neglect. The new administration begins again from a lower knowledge baseline. If the rate of learning during each administration's tenure is insufficient to exceed the rate of forgetting during the transition, the net knowledge stock declines over successive cycles.

Nigeria is the series' clearest illustration of this trap. The country study documents a pattern in which each reform cycle generates local learning—about what works in agricultural extension, in primary healthcare delivery, in anti-corruption enforcement—but the learning is not retained across administrations. The petrostate fiscal architecture severs the link between taxation and service delivery, removing the structural incentive for governments to maintain the institutional knowledge that delivery requires. The cultural operating system treats the state as a resource to be extracted rather than a service to be delivered, meaning that institutional memory is oriented toward the knowledge of extraction—who controls which revenue streams, which patronage networks deliver which votes—rather than the knowledge of governance. When a reform-minded administration departs, its learning departs with it. The next administration begins from the substrate.

The forgetting-without-learning trap is not confined to developing states. It operates in any governance system with high personnel turnover and weak institutional memory infrastructure. The United States federal government exhibits it in domains where political appointee turnover is high and civil service continuity has been eroded. The United Kingdom exhibits it in the recurrent cycle of reorganising and re-reorganising the National Health Service, each restructuring consuming institutional knowledge about what the previous structure did well and poorly. The private sector exhibits it in the corporate amnesia that follows mergers, restructurings, and leadership transitions.

The formal condition for net capability loss is straightforward. Let r_l be the rate of learning—the rate at which the system accumulates new, valid knowledge about its operating parameters. Let r_f be the rate of forgetting—the rate at which existing knowledge decays through personnel turnover, institutional reorganisation, and the natural atrophy of organisational memory. The net change in the knowledge stock is

$$\Delta K = r_l - r_f.$$

$$\Delta K = r_l - r_f.$$

When $r_l > r_f$, the system's knowledge stock grows, and its model improves over time. When $r_l < r_f$, the knowledge stock shrinks, and the system becomes progressively less calibrated to its environment—not because it is learning the wrong things, but because it cannot retain what it learns long enough for the learning to accumulate.

The design implication is that institutional memory is not a default condition. It must be deliberately architected. Evaluation repositories that survive administrations. Career civil service structures that retain expertise across political transitions. Mandatory handover documentation. Independent audit and evaluation

bodies with statutory permanence. Knowledge management infrastructure that is not dependent on the continued attention of the political leadership that created it. These are the structural responses to the forgetting-without-learning trap, and they are developed in Part VI.

These five failure modes—exploration starvation, model lock-in, exploitation lock-in, learning-induced oscillation, and the forgetting-without-learning trap—are the characteristic pathologies of governance systems that have not been designed for learning. They are not mutually exclusive. A system can simultaneously starve its exploration budget, lock in its existing models, and forget faster than it learns. The most fragile systems exhibit all five at once. The most resilient ones have architectures that address each.

The simulation of Part IV demonstrates the dynamics of exploration starvation, exploitation lock-in, and forgetting under controlled conditions. The empirical illustrations of Part V ground them in the historical record. The design principles of Part VI specify the institutional machinery for preventing them. But the core insight is already in place: learning is not a by-product of competent governance. It is a structural property that must be designed for, protected, and sustained—and when it is not, the system's competence decays in ways that are invisible to the system's own monitoring until the accumulated decay breaches a crisis threshold. The remainder of the paper is about what to do about it.

Part IV — Simulation: The Adaptive Controller

The formal framework of Part II models the dual control problem: a controller that must simultaneously regulate a system and learn its parameters. The failure modes of Part III trace the consequences of imbalances in that dual objective—too little exploration, too slow model updating, too weak actuation, too aggressive perturbation, too rapid forgetting. This part subjects those dynamics to controlled simulation, demonstrating that the failure modes emerge reliably from a minimal set of assumptions about the learning architecture, even when the underlying governance system is otherwise well-designed.

The simulation is not a calibration to any specific real-world governance system. It is an existence proof: a demonstration that the qualitative dynamics the formal framework predicts—the exploration-starvation trap, the exploitation lock-in, the forgetting-without-learning threshold—are generated by a controller that must learn while it acts, in an environment that changes while it learns. The parameters are chosen to make the mechanisms visible. The code is open-source, with fixed seeds for replicability, Monte Carlo distributions across 100 seeds, and parameter sweeps demonstrating robustness. The full specification is provided in Appendix B.

4.1 Model Specification

The simulated governance system controls a two-dimensional state vector $x(t) = [x_1(t), x_2(t)]^\top \in \mathbb{R}^2$, representing two policy-relevant dimensions such as economic output and environmental quality, or service delivery and fiscal balance. The true dynamics are linear but the controller does not know the parameters with certainty:

$$x(t+1) = A(\theta_t)x(t) + B(\theta_t)u(t) + w(t), \quad w(t) \sim N(0, W),$$

where $\theta_t \in \mathbb{R}^p$ is a vector of unknown, slowly time-varying parameters that govern the system's response to control inputs. The nominal design values are $A_0 = 0.95 I_2$ and $B_0 = I_2$, but the true values drift over time according to a random walk with small variance:

$$\theta_{t+1} = \theta_t + \eta_t, \quad \eta_t \sim N(0, \sigma_\theta^2 I),$$

where σ_θ^2 is the *environmental change rate*—the speed at which the system's dynamics shift away from the controller's model. When $\sigma_\theta^2 = 0$, the environment is stationary and the controller can eventually learn the true parameters perfectly. When $\sigma_\theta^2 > 0$, the parameters are a moving target, and the controller must learn continuously to keep pace.

The controller observes the state through a noisy measurement channel:

$$y(t) = x(t) + v(t), \quad v(t) \sim N(0, V_0),$$

with $V_0 = 0.05 I_2$. The observation channel is assumed to be intact—this simulation isolates the learning dynamics by holding the observation architecture at its designed fidelity. Degradations of the observation channel (Papers III, VI, XIII) would compound the learning failures demonstrated here.

The controller's objective is to minimise the cumulative squared tracking error over the simulation horizon:

$$J = \sum_{t=0}^T \|x(t)\|^2,$$

$$J = \sum_{t=0}^T \|x(t)\|^2,$$

where the target state is the origin $x^* = 0$. The controller's performance is measured by the time-averaged tracking error after a burn-in period.

4.2 Controller Architecture

The controller maintains a belief distribution over the unknown parameters θ and updates it recursively as observations accumulate. For computational tractability, the simulation implements a recursive least-squares (RLS) estimator with a forgetting factor $\lambda_f \in (0, 1]$, which provides a running estimate $\hat{\theta}(t)$ of the parameters and an associated covariance matrix $P(t)$ that quantifies the controller's uncertainty.

At each time step, the controller computes two candidate actions. The *certainty-equivalent* action $u_{CE}(t)$ is the optimal LQR control given the current parameter estimate $\hat{\theta}(t)$:

$$u_{CE}(t) = -K(\hat{\theta}(t)) \hat{x}(t),$$

$$u_{CE}(t) = -K(\hat{\theta}(t)) \hat{x}(t),$$

where $K(\hat{\theta})$ is the LQR gain computed for the estimated dynamics. The *exploration* component $u_{explore}(t)$ is a Gaussian dither:

$$u_{explore}(t) \sim N(0, \sigma_\eta^2 I),$$

$$u_{explore}(t) \sim N(0, \sigma_\eta^2 I),$$

where σ_η^2 is the *exploration variance*—the controller's chosen intensity of probing. The total control action is the sum:

$$u(t) = u_{CE}(t) + u_{explore}(t).$$

$$u(t) = u_{CE}(t) + u_{explore}(t).$$

The exploration dither serves two functions in the dual control framework. First, it provides persistent excitation, ensuring that the input signal spans the directions required to identify the system's parameters (the condition of Section 2.4). Second, it embodies the exploration bonus: the controller deliberately accepts short-term performance degradation—the dither pushes the state away from the target—in exchange for information that improves future performance.

The key control parameter is the exploration variance $\sigma_\eta^2 \sigma_\eta^2$. When $\sigma_\eta^2 = 0 \sigma_\eta^2 = 0$, the controller is certainty-equivalent: it exploits its current model without probing. When $\sigma_\eta^2 \sigma_\eta^2$ is moderate, the controller balances exploration and exploitation. When $\sigma_\eta^2 \sigma_\eta^2$ is large, exploration dominates, and the controller's own perturbations become a significant source of variance in the system's trajectory.

The actuation chain is modelled with a multiplicative efficiency factor $\mu \in [0, 1]$ $\mu \in [0, 1]$ representing the fraction of the intended control signal that actually reaches the system:

$$u_{\text{eff}}(t) = \mu u(t).$$

$u_{\text{eff}}(t) = \mu u(t)$.

When $\mu = 1$ $\mu = 1$, the actuation chain is intact and the controller's decisions are fully implemented. When $\mu < 1$ $\mu < 1$, the implementation is attenuated—the structural condition that Paper XI diagnosed as delegation depth and that the exploitation lock-in failure mode (Section 3.3) identifies as the bottleneck between learning and action.

4.3 Scenarios

Six scenarios are simulated, corresponding to the failure modes of Part III. All use the same plant dynamics and the same RLS estimator. They differ only in the exploration variance $\sigma_\eta^2 \sigma_\eta^2$, the environmental change rate $\sigma_\theta^2 \sigma_\theta^2$, the forgetting factor $\lambda_f \lambda_f$, and the actuation efficiency μ .

Scenario 1 — Optimal dual control. The controller maintains a moderate, persistent exploration dither ($\sigma_\eta^2 = 0.05 \sigma_\eta^2 = 0.05$) and a slow forgetting factor ($\lambda_f = 0.99 \lambda_f = 0.99$). The environment changes slowly ($\sigma_\theta^2 = 0.002 \sigma_\theta^2 = 0.002$). Actuation is intact ($\mu = 1$ $\mu = 1$). This scenario demonstrates the baseline: the controller's parameter estimates track the drifting true parameters, and tracking error remains bounded and low. The system learns stably.

Scenario 2 — Exploitation-only (certainty-equivalent). The controller suppresses exploration entirely ($\sigma_\eta^2 = 0 \sigma_\eta^2 = 0$). All other parameters match Scenario 1. The controller applies the optimal action given its current model, but it never probes to discover whether the model is still correct. As the environment drifts, the parameter estimates diverge from the true values. Tracking error initially matches or slightly outperforms

Scenario 1—because the controller is not introducing exploratory variance—but then degrades progressively as model drift accumulates. The degradation is invisible to the controller's internal monitoring, which uses the same drifting model to estimate performance. This is exploration starvation.

Scenario 3 — Crisis-driven learning. The controller operates in exploitation-only mode ($\sigma_\eta^2 = 0$) until the tracking error exceeds a threshold $e_{\text{crit}} = 2.0$, at which point it switches to high exploration ($\sigma_\eta^2 = 0.5$) for a fixed duration of 20 time steps before returning to exploitation. The scenario produces a boom–bust learning cycle: periods of stable but gradually degrading performance, punctuated by episodes of aggressive, disruptive relearning. The average performance over the cycle is worse than the optimal dual control baseline, and the system experiences periodic crises that could have been avoided by sustained moderate exploration.

Scenario 4 — Over-exploration. The controller explores with excessive variance ($\sigma_\eta^2 = 0.5$ continuously). The dither is so large that the controller's own perturbations dominate the system's dynamics. The resulting oscillation obscures the very parameters the controller is trying to estimate. Parameter estimates are noisy and unreliable; tracking error is high and volatile. The scenario demonstrates that exploration is not an unqualified good—it must be calibrated to the system's dynamics and to the noise environment, and excessive exploration can be as damaging as none.

Scenario 5 — Forgetting-without-learning. The controller explores moderately ($\sigma_\eta^2 = 0.05$) but institutional memory is weak: the forgetting factor is set to $\lambda_f = 0.90$, representing rapid decay of accumulated knowledge. The environmental change rate is moderate ($\sigma_\theta^2 = 0.005$). The controller acquires information through exploration, but it forgets faster than it learns. The effective sample size never accumulates; parameter estimates remain noisy and biased. Tracking error degrades progressively despite sustained exploration. This is the forgetting-without-learning trap.

Scenario 6 — Exploitation lock-in. The controller explores and learns adequately ($\sigma_\eta^2 = 0.05$, $\lambda_f = 0.99$, $\sigma_\theta^2 = 0.002$), and its parameter estimates track the true parameters accurately. But the actuation efficiency is reduced to $\mu = 0.3$ —only thirty percent of the intended control signal reaches the system. The controller knows what to do, and its model of the system is accurate, but it cannot translate knowledge into action. Tracking error remains high because the control signal is attenuated, even though the learning is successful. The controller's parameter estimates are accurate while its performance is poor—the signature of exploitation lock-in. This scenario is run as a sweep over μ from 1.0 down to 0.1, demonstrating the progressive decoupling of learning from performance as actuation degrades.

4.4 Parameter Sweeps

Three sweeps are conducted to map the boundaries of stable learning.

Exploration variance vs. environmental change rate. $\sigma_\eta^2\sigma_\eta^2$ is swept from 0.0 to 0.50.5 and $\sigma_\theta^2\sigma_\theta^2$ from 0.0 to 0.020.02. For each combination, the mean steady-state tracking error is recorded. The sweep produces a phase diagram in $(\sigma_\eta^2, \sigma_\theta^2)$ space with a clearly defined region of stable learning. At low $\sigma_\theta^2\sigma_\theta^2$, even small exploration suffices; as the environment changes faster, the minimum exploration required to keep pace rises. Below this curve, the system is in the exploration-starvation regime. Above the curve but within a bounded band, the system learns stably. At very high $\sigma_\eta^2\sigma_\eta^2$ relative to $\sigma_\theta^2\sigma_\theta^2$, the system enters the over-exploration regime, and performance degrades again.

Forgetting factor vs. learning rate. The forgetting factor $\lambda_f\lambda_f$ is swept from 0.800.80 to 1.001.00, and the environmental change rate $\sigma_\theta^2\sigma_\theta^2$ from 0.0 to 0.020.02, with exploration held constant at the optimal dual control level. The sweep identifies the net-learning threshold: the line in $(\lambda_f, \sigma_\theta^2)$ space below which the rate of information acquisition exceeds the rate of forgetting, and above which the knowledge stock declines. Systems with high personnel turnover, short electoral cycles, or weak institutional memory infrastructure operate with low effective $\lambda_f\lambda_f$; the sweep shows the maximum rate of environmental change they can tolerate before entering the forgetting-without-learning trap.

Actuation efficiency vs. performance. $\mu\mu$ is swept from 1.01.0 to 0.10.1 with all other parameters at the optimal dual control values. The sweep demonstrates the exploitation lock-in curve: tracking error as a function of actuation efficiency. The controller's parameter estimation accuracy is simultaneously recorded, showing that learning remains accurate even as performance degrades—the decoupling that defines this failure mode.

4.5 Expected Results and Key Figures

The simulation produces four primary outputs.

Figure 1 — Phase diagram of stable learning. A heatmap in $(\sigma_\eta^2, \sigma_\theta^2)$ space, colour-coded by mean steady-state tracking error. The stable-learning region is the band of moderate exploration variance where tracking error is low. The exploration-starvation region (low $\sigma_\eta^2\sigma_\eta^2$, moderate-to-high $\sigma_\theta^2\sigma_\theta^2$) shows elevated error as model drift accumulates. The over-exploration region (high $\sigma_\eta^2\sigma_\eta^2$) shows elevated error as the dither destabilises the system. Contours mark the minimum exploration required for stable learning at each environmental change rate—the persistent excitation boundary. The figure makes visible the dual nature of the exploration–exploitation trade-off: too little exploration and the model drifts; too much and the controller's own perturbations dominate.

Figure 2 — Time-series of exploration starvation and optimal dual control. Three panels. The top panel shows the tracking error $\|x(t)\|$ over time for Scenario 1 (optimal dual control) and Scenario 2 (exploitation-only). The exploitation-only trajectory initially matches or slightly outperforms the dual control

trajectory, then diverges upward as model drift accumulates. The middle panel shows the parameter estimation error $\|\hat{\theta}(t) - \theta_t\|$ for the same trajectories. The dual control error remains bounded and low; the exploitation-only error drifts upward without bound. The bottom panel shows the controller's *internal* estimate of its own tracking error, as computed from the drifting model. For the exploitation-only controller, this internal estimate remains low even as true tracking error rises—the signature of the self-concealing trap. The controller believes it is performing well because the model it uses to evaluate itself is the same model that is failing.

Figure 3 — Exploitation lock-in trajectory. Two panels for Scenario 6. The top panel shows the tracking error $\|x(t) - x^*(t)\|$ for three values of μ : 1.0 (full actuation), 0.5 (partial actuation), and 0.2 (severe attenuation). Tracking error rises as actuation degrades. The bottom panel shows the parameter estimation error for the same three trajectories—and it remains low and nearly identical across all three. The controller learns equally well regardless of whether it can act on what it learns. The vertical distance between the estimation error curve and the tracking error curve is the exploitation lock-in gap: the performance cost of the blocked translation from knowledge to action.

Figure 4 — Forgetting-without-learning sweep. A heatmap in $(\lambda_f, \sigma_\theta^2)$ space, colour-coded by the net change in the knowledge stock over the simulation horizon. The net-learning region (green) shows where the rate of information acquisition from exploration exceeds the rate of forgetting. The net-forgetting region (red) shows where knowledge decays faster than it accumulates. The boundary between them is the forgetting-without-learning threshold. Overlaid on the heatmap are approximate locations for governance systems with different institutional memory characteristics: a Nordic system with strong civil service continuity and institutionalised evaluation repositories (high λ_f), a Westminster system with moderate ministerial turnover (moderate λ_f), and a system with weak institutional memory infrastructure and high political turnover (low λ_f). The overlay is illustrative, not calibrated, but it makes visible the structural vulnerability of high-turnover systems to environmental change.

Summary metrics. For each scenario, the simulation reports: mean steady-state tracking error, mean parameter estimation error, the fraction of Monte Carlo runs in which the controller's internal performance estimate diverges from true performance by more than 50% (the self-concealing metric), and, for Scenario 3, the number of crisis-triggered relearning episodes and the total time spent in crisis mode.

The simulation does not predict specific outcomes for any real governance system. It demonstrates the qualitative dynamics that the formal framework identifies, under controlled conditions, with all non-learning parameters held at ideal values. The fact that the failure modes emerge even under these idealised conditions—with optimal state estimation, with well-tuned LQR gains, with no adversarial actors—is the central simulation finding. The exploration-starvation trap, the exploitation lock-in, and the forgetting-without-learning threshold are not consequences of institutional dysfunction. They are consequences of the structural relationship between learning, actuation, and memory in any controller that must learn while it acts. The simulation makes that relationship visible.

Part V — Empirical Illustrations

The formal framework of Part II and the failure modes of Part III make predictions about where and how the dynamics of adaptive governance will manifest. This part examines five cases that span the range from successful learning to catastrophic forgetting, and from the national scale to the institutional. In each, the learning lens reveals a structural dynamic that is invisible to, or misdescribed by, the standard vocabulary of institutional quality and political failure.

5.1 China's Reform Era as Dual Control

The Deng-era reforms, spanning roughly 1978 to the early 1990s, are the most sustained and successful instance of deliberate governance learning in modern history. The architecture that produced them was not designed from a single blueprint, nor was it imported wholesale from foreign models. It emerged through a sequence of structured explorations—bounded experiments that generated information, which was then used to update the system's model of what worked and to scale what succeeded.

The Special Economic Zones (SEZs) are the canonical example. Established in the early 1980s in Shenzhen, Zhuhai, Shantou, and Xiamen, they were deliberate deviations from the planned economy: geographically bounded spaces where market mechanisms, foreign investment, and export-oriented production could operate under different rules from the rest of the country. The SEZs were not merely policy experiments in the loose sense; they were instruments for parameter identification. The central government did not know the elasticity of supply response to market incentives, the productivity differential between state-owned and foreign-invested enterprises, or the administrative capacity required to regulate a market economy. The SEZs generated the variation that made it possible to learn these parameters.

The household responsibility system in agriculture followed the same logic. It was not designed in Beijing and imposed on the countryside. It emerged from local experiments—first in Anhui province, then spreading through imitation and observation—that the central government eventually endorsed and scaled. The dual-track price system for industrial goods maintained the planned allocation track (exploitation of the existing model) while allowing a market track at the margin (exploration). Resources shifted gradually toward the track that performed better, and the information generated by the market track informed the progressive dismantling of the planned track.

This is dual control in operation. The Chinese state in this period did not attempt to design an optimal economic system from first principles and implement it. It maintained exploitation of the existing planned economy while injecting a persistent excitation signal—the SEZs, the household responsibility pilots, the dual-track prices—that generated information about the parameters of a market alternative. As the

information accumulated and the estimates became more precise, the exploitation component shifted: resources, administrative attention, and eventually the formal institutional architecture migrated toward the model that the exploration had validated.

The success of this architecture is measured not only in the growth rates it produced—sustained double-digit GDP growth, the largest reduction in poverty in human history—but in the fact that it was peaceful. The transition from plan to market in China did not involve the catastrophic output collapses that accompanied shock therapy transitions in the former Soviet Union and Eastern Europe. The dual control architecture allowed the system to learn its way into a new institutional configuration without ever abandoning the capacity to govern during the transition. The exploitation component ensured that the existing economy continued to function; the exploration component ensured that the future economy could be discovered.

The subsequent deterioration of this adaptive capacity is equally instructive. Around 2012, the learning that had characterised the reform era began to slow. The calibration deficit that the series diagnoses in the China country study is, at root, an exploration-starvation trap. The promotion tournament that sustains high actuation-legitimacy for centrally monitored targets simultaneously penalises the reporting of information inconsistent with the centre's model—the observation-legitimacy collapse of Paper XIII. The experimental spaces that generated the reform era's persistent excitation—the SEZs, the local policy innovations, the tolerance for provincial variation—have been progressively constrained by a centre that increasingly treats deviation as a threat rather than as a source of information. The system that was the world's most successful governance learner has become unwilling to sustain the explorations that learning requires. The exploration variance σ_{η}^2 appears to have been progressively reduced, and the model drift that will eventually force a crisis is accumulating in the dimensions the suppressed observation channels can no longer report.

5.2 Japan's Continuity Trap as Model Lock-In

Japan's post-war governance paradigm—export-led growth, bureaucratic guidance of the private sector, the Iron Triangle of the Liberal Democratic Party, the bureaucracy, and big business, and the 1955 System of LDP dominance—was a coherent and highly successful model for the conditions of 1955 through roughly 1985. It delivered the "economic miracle": the transformation of a defeated, impoverished nation into the world's second-largest economy, with widely shared prosperity, high employment, and social stability.

The model ceased to correspond to reality as the environment changed. The asset price bubble of the late 1980s and its collapse in 1991 revealed structural fragilities—zombie banks, over-investment in low-productivity sectors, a financial system that misallocated capital on a massive scale—that the model could not correct. Demographics shifted: the working-age population began to shrink from the mid-1990s, and the dependency ratio rose steadily, placing increasing strain on the social insurance systems that the high-growth model had been designed to fund. The global economic environment shifted: China's rise as a

manufacturing competitor eroded the export-led growth strategy, and the information technology revolution favoured economies with more flexible labour markets and more dynamic venture capital sectors than Japan's bank-centred financial system provided.

None of this was invisible. Japan's economic planners, its academic economists, its financial regulators, and its political leadership possessed sophisticated data about each of these trends. The demographic projections were published and updated annually. The bank non-performing loan problem was documented in detail by the Financial Services Agency and by independent researchers. The productivity gap with the United States in services and information technology was measured and reported by the Ministry of Economy, Trade and Industry. The evidence that the post-war model was no longer adequate was abundant, high-quality, and publicly available.

What the architecture lacked was any mechanism for translating that evidence into model revision. The Iron Triangle embedded the post-war model in the institutional structure of the state. The bureaucracy's career paths, the LDP's committee system, the *keiretsu* corporate networks, the postal savings system that funnelled household deposits into public investment—each was both an instrument of the model and a beneficiary of its continuation. Challenges to the model were challenges to the interests that these institutions served. The immune system (Paper VII) treated them as threats to be neutralised, not as information to be incorporated. The succession of prime ministers who attempted structural reform—Hashimoto, Koizumi, Kan, Noda—each achieved partial, temporary progress against the model's defenders, and each was eventually absorbed, deflected, or defeated by the architecture they were trying to change.

The result was three decades of stagnation. GDP growth averaged less than one percent per year from 1992 to 2019. Government debt rose to over 250 percent of GDP, sustained only by the captive domestic savings pool that the model's financial architecture continued to direct toward government bonds. The social contract—lifetime employment, seniority-based wages, corporate welfare—eroded at the margins while the formal model persisted at the centre. The system was not failing to sense. It was not failing to learn in the narrow sense of accumulating evidence. It was failing to update its model in response to the evidence it had accumulated. The model was locked in.

Japan's Continuity Trap is a pure case of model lock-in, and it illustrates the distinction between this failure mode and exploration starvation. Japan did not starve exploration entirely—the evidence was generated, the knowledge was available. What it lacked was the mechanism for paradigm replacement: the institutionalised capacity to retire a model that no longer corresponded to reality and to replace it with one that did. The forgetting function—the capacity to discard obsolete models—was as absent as the learning function was present.

5.3 Finland's Foresight Infrastructure as Institutional Memory—and the Exploitation Lock-In

Finland possesses one of the world's most advanced institutionalised foresight architectures. The Committee for the Future, a permanent parliamentary committee established in 1993, has a statutory mandate to review the long-term implications of proposed legislation, to conduct studies on emerging trends and their governance implications, and to engage with the scientific and technological research community. The national foresight reporting system requires the government to produce a Futures Report once per electoral term, projecting demographic, technological, environmental, and economic trends over a multi-decade horizon and assessing the adequacy of current policies in light of those projections. The Sitra innovation fund operates as an independent foresight and experimentation body. The Futures Impact Assessment framework mandates that major legislative proposals include an explicit analysis of their consequences for future generations.

This infrastructure is a designed learning apparatus. It performs the sensing function—scanning for the emerging disturbance dimensions that constitute the α of Paper VI's variety gap dynamics. It performs the learning function—updating the governance system's model of its own environment through systematic foresight, scenario analysis, and expert engagement. It performs the memory function—maintaining an institutional record of projections, assessments, and policy analyses that persists across electoral cycles and government turnovers. In the language of the formal framework, Finland has built a high-fidelity, low-forgetting estimator of its own operating environment.

And yet the Finland country study diagnoses a throughput constraint: the persistent gap between the quality of the foresight and the speed and scale of the policy response. The Futures Reports identify trends that require legislative or regulatory action; the action is delayed, diluted, or deferred. The Futures Impact Assessments reveal that proposed policies are inadequate to the long-term challenges they address; the policies proceed largely unchanged. The learning infrastructure generates high-quality knowledge; the actuation infrastructure cannot translate that knowledge into action at the pace the knowledge requires.

This is exploitation lock-in. The exploration function is intact. The learning function is intact. The knowledge is available, institutionally embedded, and publicly accessible. But the actuation chain is throttled—by the consensus-oriented political culture, by the fragmentation of executive authority across coalition governments, by the limited transition bandwidth of a system designed for incremental adjustment rather than transformational response. The system knows what to do. It cannot do it.

Finland's case is diagnostically important because it isolates the exploitation lock-in mechanism from the other learning failure modes. It is not a case of exploration starvation—the foresight infrastructure ensures that exploration is sustained. It is not a case of model lock-in—the model is being updated; the Futures Reports are produced, the projections are revised, the assessments are conducted. It is not a case of forgetting-without-learning—the memory infrastructure is strong, the knowledge persists. It is a case of

learning that cannot reach actuation, and it demonstrates that the sensing, learning, and actuation functions are distinct architectural properties. Excellence in the first two does not compensate for deficiency in the third.

5.4 Nigeria's Substrate Deficit as Forgetting-Without-Learning

Nigeria exhibits the forgetting-without-learning trap at a scale and duration that make it the series' clearest illustration of this failure mode. The Nigerian state has, over the six decades since independence, undergone repeated cycles of reform—structural adjustment programmes, anti-corruption campaigns, civil service reforms, public financial management overhauls, sectoral strategies for agriculture, health, and education. Each reform cycle generates local learning: knowledge about what works in Nigerian conditions, which implementation strategies succeed and which fail, where the capacity constraints actually bind.

This learning is not retained. Each political transition—and Nigeria has alternated between civilian and military rule, and among civilian administrations, with a frequency that gives it one of the highest rates of government turnover in the world—effectively resets the institutional memory. The incoming administration arrives with its own priorities, its own advisors, its own models. The evaluations, programme data, and tacit knowledge of the outgoing administration are not systematically transferred. They are discarded—through neglect, through the deliberate dismantling of predecessor institutions, through the patronage logic that allocates positions and resources to new networks rather than to the maintenance of existing knowledge.

The petrostate fiscal architecture compounds the forgetting. In a state whose revenues derive primarily from oil exports rather than from domestic taxation, the structural incentive to maintain the institutional knowledge required for service delivery is weak. Taxpayer accountability—the mechanism that, in high-legitimacy systems, creates pressure to retain what works and discard what doesn't—is absent. The state does not need to know how to deliver services effectively because its fiscal base does not depend on delivering them. The knowledge of extraction—who controls which revenue streams, which patronage networks deliver which votes—is retained and refined across administrations. The knowledge of governance is not.

The cultural operating system that the Nigeria country study documents reinforces the forgetting. "The National Cake" is the shared understanding that the state exists as a resource to be divided among constituencies, not as a service to be delivered to citizens. Institutional memory, in this operating system, is oriented toward the knowledge of extraction and distribution. The knowledge of effective governance—what works in agricultural extension, in primary healthcare delivery, in anti-corruption enforcement—is not valued, not retained, and not accumulated. Each reform cycle begins again from the substrate.

The formal condition of Section 3.5 applies directly. The rate of learning during any given reform episode is positive—the World Bank project evaluations, the NGO programme assessments, and the academic studies all document instances of effective innovation. But the rate of forgetting during the political transitions between episodes is higher. The net knowledge stock does not grow. It may even decline, if each transition destroys not only the learning of the previous administration but also the residual institutional knowledge that

survived earlier transitions. The result is six decades of reform without the accumulation of reform benefits—a system that has repeatedly learned the same lessons without ever retaining them long enough for them to compound.

5.5 The Scientific Community as an Adaptive Epistemic System

The scientific community is not a governance system in the conventional sense. It does not levy taxes, enforce laws, or administer public services. But it is the most successful adaptive epistemic system that human beings have yet built, and its institutional architecture embodies the design principles this paper advocates with a fidelity that no governance system has matched. It serves as an existence proof: a demonstration that sustained, institutionalised learning is possible at scale, and that the failure modes of Part III are not inevitabilities but consequences of specific architectural choices that can be made differently.

The scientific community institutionalises all three functions of the Cycle Two adaptation triad.

Sensing (observer diversity). The scientific community maintains an observing ensemble of extraordinary decorrelation. Researchers in different laboratories, in different countries, using different methods, funded by different sources, examine the same phenomena from different angles. Their errors are decorrelated; their systematic biases, to the extent they exist, are not identical across the population. The community converges on findings that survive cross-checking by multiple independent observers—exactly the condition that Paper X identifies as requisite observer diversity. Peer review, adversarial collaboration, and replication studies are the institutional mechanisms that maintain this decorrelation.

Learning (dual control). The scientific community's reward structure embodies the dual control objective. Individual scientists are rewarded for exploration—for discovering new phenomena, developing new theories, challenging established findings. The certainty-equivalent action—repeating what is already known—is professionally penalised. Simultaneously, the community's institutional structure ensures that exploration does not destabilise the collective knowledge base. Exploratory claims must survive adversarial scrutiny before they are incorporated into the canon. The exploration is persistent and institutionalised; the model updating is systematic and conservative. The balance between individual exploration and collective exploitation is the scientific method's solution to the dual control problem.

Memory and forgetting (institutional infrastructure). The scientific community maintains an elaborate institutional memory: the published literature, the curated databases, the textbooks and review articles that synthesise accumulated knowledge for the next generation. It also institutionalises forgetting: the paradigm shifts that Kuhn described are the mechanism by which models that no longer correspond to evidence are retired. The forgetting is not smooth—it is contested, generational, and often delayed beyond the point at which the evidence would warrant replacement. But it happens. The scientific community does not suffer the model lock-in that paralyses governance systems with weaker paradigm-replacement mechanisms.

The scientific community's pathologies are instructive precisely because they are the failure modes of Part III operating within an otherwise well-designed adaptive architecture. Publication bias—the systematic suppression of null results and replication failures—is a form of observation-channel degradation: the community's collective model of what is true becomes systematically over-confident because the evidence that would challenge it is not being published. The replication crisis of the 2010s was a forgetting-without-learning episode: findings that had been accepted into the canon were discovered to be unreplicable, and the community realised that its memory infrastructure had been preserving noise alongside signal. Both pathologies have prompted institutional reforms—pre-registration, registered reports, data and code sharing requirements, replication incentives—that are the scientific community's equivalent of the design principles of Part VI: mechanisms for strengthening the learning architecture against the failure modes to which it is vulnerable.

The scientific community is not a template for governance design. The differences in scale, in the nature of the decisions at stake, and in the normative framework for legitimation are profound. But it is an existence proof that sustained, institutionalised learning is architecturally achievable, and that the failure modes of exploration starvation, model lock-in, and forgetting-without-learning can be resisted through deliberate institutional design. It is the closest existing analogue to the adaptive controller this paper describes, and its strengths and its pathologies are both instructive for the governance design task.

Part VI — Design Principles for Adaptive Governance Architectures

The diagnosis is structural. A governance system that cannot learn—that starves exploration, locks in obsolete models, forgets faster than it learns, or blocks the translation of knowledge into action—has a finite operational lifetime in an environment that changes faster than architectures can be redesigned. The failure modes of Part III are the signatures of a learning architecture that has been left to chance, subjected to political incentives that systematically penalise the variance on which learning depends, and deprived of the institutional machinery that would protect it.

This part turns from diagnosis to prescription. It specifies the design principles that follow from the formal framework and that a governance architecture must satisfy if it is to learn stably over time. The principles are not a blueprint. The appropriate learning architecture for any specific governance function depends on the rate of environmental change in that domain, the cost structure of experimentation, and the existing institutional substrate. What the principles provide is a set of structural requirements that any adaptive governance architecture must meet, and a vocabulary for designing institutions that meet them.

6.1 Protected Experimental Spaces with Persistent Excitation

The persistence of excitation condition (Section 2.4) establishes that a controller must maintain sufficient variation in its input signal to identify the parameters of the system it governs. The governance analogue is that a governance system must maintain a sustained programme of experimentation—policy variation across jurisdictions, pilot programmes, alternative delivery models—to remain calibrated to a changing environment.

The series has already identified protected experimental spaces as the convergent first step across all country cases. Paper VII documented their role as bypass mechanisms that route around dysfunctional central architectures and generate legible evidence about what works. Paper IX reframed them as transition mechanisms that build demonstrated value before scaling. This paper provides the formal rationale: experimental spaces are the governance analogue of the persistent excitation signal in adaptive control. They generate the independent variation in policy design, implementation context, and evaluation methodology that is required to identify the system's operating parameters.

The design requirements follow from the formal condition. The excitation must be *persistent*—not a one-off pilot that is either terminated or scaled and never repeated, but a permanent feature of the governance architecture. The excitation must be *spanning*—the experimental portfolio must cover the dimensions of the system's dynamics that are most uncertain and most consequential for future performance. A system that

experiments only with marginal programme adjustments while leaving its fundamental regulatory models, fiscal frameworks, and institutional structures untouched is not satisfying the spanning condition; it is exploring within a box whose walls it never tests. The excitation must be *protected*—the experimental spaces must be insulated from the short-term political incentives that would extinguish them when they produce uncomfortable results or when budgets tighten. Protection requires independent funding, statutory mandates, and leadership appointment processes that are decoupled from the political cycle.

The municipal laboratory, the sandbox state, the Special Economic Zone, the pilot programme with randomised evaluation, the regulatory sandbox for financial or technological innovation—each is an institutional form of the persistent excitation principle. None is sufficient alone. The architecture requires a portfolio of such spaces, operating at different scales and in different domains, collectively providing the excitation that keeps the governance system's parameters identifiable.

6.2 Safe-to-Fail Structures

Persistent excitation inherently generates errors. Exploration means trying things that might not work, and some of them will not work. If every failure is politically or institutionally catastrophic—if a failed pilot programme ends careers, if a negative evaluation triggers a media firestorm, if an experimental policy that produces adverse outcomes generates lawsuits and compensation claims—then the rational response, under the incentive structure of real governance institutions, is to stop exploring.

The design response is *safe-to-fail* structures: bounded experimental environments where failure is survivable and informative rather than catastrophic. This is the governance analogue of the engineering distinction between fail-safe and safe-to-fail design. A fail-safe system attempts to prevent failure entirely. A safe-to-fail system accepts that failure will occur and designs the system's boundaries so that failure is contained, its consequences are limited, and the information it generates is captured.

The structural properties of a safe-to-fail experimental space include:

Bounded scope. The experiment operates on a limited population, in a limited geography, or for a limited duration. The failure of a municipal pilot programme affects thousands, not millions. The failure of a time-limited regulatory sandbox entry does not restructure an entire industry. The boundary contains the blast radius.

Pre-committed evaluation and exit. The conditions under which the experiment will be judged a success or a failure are specified in advance, as are the consequences of each. A programme that fails to meet its pre-registered success criteria is terminated, not because a political adversary demands it but because the architecture mandates it. The termination is not a scandal; it is the system operating as designed. Pre-commitment protects the experiment from post-hoc politicisation of its results and from the sunk-cost fallacy that turns pilot programmes into permanent, unevaluated entitlements.

Institutionalised learning extraction. The information generated by a failure is captured and retained regardless of the programme's fate. A programme that is terminated for failing to meet its objectives nevertheless contributes to the system's knowledge stock if the reasons for its failure are documented, analysed, and integrated into the model that guides future design. The failure is not a waste; it is a measurement.

Insulation of experimenters. The officials who design and implement experiments must be protected from career damage when experiments fail. If the personal consequences of a failed experiment are severe, the only experiments that will be proposed are those that are virtually certain to succeed—which means they are not experiments in any meaningful sense, because they generate no information about the boundaries of the system's dynamics. Protection does not mean immunity from accountability for misconduct or incompetence. It means that a well-designed experiment that produces a null or negative result is treated as a successful contribution to the system's knowledge, not as a career-ending mistake.

Safe-to-fail structures are the institutional precondition for persistent excitation. Without them, the exploration variance σ_{η}^2 will be driven to zero by the asymmetric costs of failure. With them, the system can sustain the rate of experimentation that the environment's rate of change demands.

6.3 Separation of Exploration and Exploitation Functions

The dual control framework formalises the tension between exploration and exploitation. A single controller cannot simultaneously optimise for short-term performance and long-term learning, because the actions that maximise one are not the actions that maximise the other. The tension is not a design flaw; it is a structural feature of any learning system. The design response is functional separation.

Separation of exploration and exploitation functions means that different institutions are responsible for generating knowledge and for applying it. The exploration institutions—foresight bodies, audit offices, evaluation agencies, experimental programme units, citizens' assemblies, challenge panels—have a mandate to generate information, not to deliver outcomes. Their performance is judged on the quality and relevance of the knowledge they produce, not on the outcomes of the programmes they study. The exploitation institutions—line ministries, delivery agencies, regulatory bodies—have a mandate to deliver outcomes using the best available knowledge, not to generate new knowledge. Their performance is judged on the outcomes they achieve, not on the novelty of their methods.

The separation is not absolute. Exploitation institutions must retain the capacity to absorb and act on the knowledge that exploration institutions produce—the exploitation lock-in failure mode of Section 3.3 is precisely the decoupling of these functions. And exploration institutions must be connected enough to the operational realities of governance that the knowledge they produce is relevant to the decisions that exploitation institutions face. The design challenge is to maintain sufficient separation to protect exploration from the short-term pressures that would extinguish it, while maintaining sufficient integration to ensure that what is learned is used.

The organisational forms that achieve this balance vary across governance contexts. The independent central bank is a partial instance: it is separated from the political exploitation apparatus (the treasury, the executive) and given a mandate to pursue a specific objective using its own analytical capacity, with its own exploration function embedded in its research departments. The supreme audit institution is another: it is separated from the agencies it audits and given a mandate to generate information about their performance, not to deliver their programmes. The parliamentary technology assessment body, the independent fiscal council, the citizens' assembly on a specific policy question—each is an exploration institution separated from the exploitation apparatus whose models it may challenge.

The structural requirement is that exploration institutions have protected budgets, statutory independence, and leadership appointment processes that are insulated from the exploitation institutions they are meant to inform. An exploration body whose budget is controlled by the ministry whose programmes it evaluates, or whose director is appointed by the executive whose models it challenges, is an exploration body in name only. Its exploration variance will be driven toward zero by the same incentives that produce exploration starvation in the undifferentiated controller.

6.4 Protected Curiosity Budgets

Why explore? The dual control framework answers: because exploration generates information that improves future performance, and the expected value of that information justifies the short-term cost of acquiring it. But this answer assumes an institutional environment in which the future benefits of exploration are internalised by the actors who bear its present costs. In most governance systems, they are not. The politician who funds an experimental programme whose benefits will accrue after the next election, or the civil servant who champions a pilot whose results will be available after their next rotation, is bearing costs for benefits that will be captured by others. The incentive gradient points toward under-exploration.

Protected curiosity budgets are the structural response. They are dedicated resource allocations—funding, personnel, organisational bandwidth—that are reserved for activities whose outcomes are uncertain, that are protected from performance-based review cycles that penalise the variance inherent in exploration, and that are governed by decision processes that explicitly value the information to be acquired rather than only the outcomes to be achieved.

The design properties of an effective curiosity budget include:

Ring-fencing. The exploration budget is separated from the operational budget and cannot be reallocated to exploitation activities when resources are tight. The separation is structural—different budget lines, different approval processes, different accounting treatment—not merely notional.

Portfolio logic. The curiosity budget is allocated as a portfolio of experiments with different risk profiles, different timescales, and different domains, rather than as a series of individual funding decisions. The portfolio logic acknowledges that most experiments will fail or produce ambiguous results, but that the few

that succeed will justify the entire portfolio. It protects the exploration function from the death-by-a-thousand-cuts that occurs when each individual experiment is judged on its own success or failure.

Stage-gated scaling. The curiosity budget funds small-scale, safe-to-fail experiments. Successful experiments progress through stages of increasing scale and resource commitment, with explicit decision gates at each stage that require demonstrated evidence of effectiveness. The stage-gated structure ensures that the exploration budget is not captured by programmes that have become too large to fail—the mechanism by which pilot programmes become permanent, unevaluated entitlements.

Independence of allocation. The decisions about which explorations to fund are made by bodies that are independent of the exploitation institutions that would benefit from or be threatened by the knowledge generated. A research council that allocates curiosity budgets to government agencies based on scientific peer review of the proposed experiments, rather than on the political preferences of the agencies themselves, is one institutional form.

Protected curiosity budgets are the governance equivalent of the exploration bonus in the dual control objective function. Without them, the controller will systematically under-explore relative to the optimal policy—not because it is irrational, but because the institutional structure does not allow it to internalise the long-term benefits of the information its explorations would generate.

6.5 Mandatory Model Review and Paradigm Replacement Cycles

Model lock-in—the persistence of a model beyond its correspondence to reality, sustained by the institutional immune system—is the failure mode that separates systems that can update their understanding from those that cannot. The design response is to institutionalise the model review and replacement function, so that it does not depend on the voluntary self-correction of the institutions that maintain the model.

Mandatory model review means that the major models on which governance decisions depend—macroeconomic forecasting frameworks, demographic projections, climate sensitivity estimates, regulatory impact models, security threat assessments—are subject to periodic, scheduled, independent review. The review is not triggered by crisis or by the discretion of the model's owners. It is a recurring institutional event, like a financial audit or an election, whose occurrence does not depend on anyone's judgment that the model might be wrong.

The structural properties of an effective model review architecture include:

Independence of the reviewer. The body that conducts the review is independent of the institution that maintains the model. An economic forecasting model reviewed by the same finance ministry that uses it for budget preparation is not being independently reviewed. The reviewer must have the institutional standing, the technical capacity, and the protected budget to conduct a thorough assessment and to publish findings that are uncomfortable for the model's owners.

Pre-specified review criteria. The criteria against which the model is assessed—its predictive accuracy, its calibration, its coverage of relevant variables, its handling of uncertainty—are specified in advance and applied systematically. The specification prevents the post-hoc adjustment of standards that occurs when a model's defenders scrutinise disconfirming evidence more aggressively than confirming evidence.

Paradigm replacement authority. The review body has the authority to trigger a paradigm replacement process when the model's performance falls below a specified threshold for a sustained period. The replacement process does not dictate what the new model should be—that is a scientific and technical question—but it mandates that the existing model can no longer serve as the official basis for decision-making, and it requires the model's owning institution to develop, test, and transition to an alternative within a specified timeframe. The authority to retire a model is as important as the authority to review it.

Public transparency. The review findings, the evidence on which they are based, and the response of the model's owning institution are all public. Transparency serves multiple functions: it prevents the suppression of unfavourable reviews, it enables the broader epistemic community to scrutinise the review's own methodology, and it builds the legitimacy of the review process over successive cycles.

The model review cycle is the institutional analogue of the model validation step in adaptive control—the periodic check that the internal model still corresponds to the system it claims to represent. The scientific community's peer review and replication mechanisms perform this function for scientific models. Governance systems require analogous mechanisms for the models on which their decisions depend, and those mechanisms must be protected from the institutions whose models they may invalidate.

6.6 Institutionalised Forgetting

Retaining only what remains useful is as important as retaining what is learned. The forgetting-without-learning trap of Section 3.5 demonstrates that systems with weak institutional memory lose knowledge faster than they acquire it. But the complementary pathology—retaining obsolete models, programmes, and institutional arrangements beyond their useful life—is model lock-in sustained by the absence of a forgetting function.

Institutionalised forgetting is the deliberate design of mechanisms that enable the peaceful retirement of programmes, policies, and institutional arrangements that no longer serve their purpose. The forgetting is not amnesia—the destruction of memory—but curation: the active, discriminating removal of what is no longer useful to make space for what is.

The design mechanisms include:

Sunset clauses. Legislation and programmes are enacted with a pre-specified expiration date, after which they terminate automatically unless explicitly reauthorised. The burden of proof is on continuation, not on termination. Sunset clauses invert the default: instead of programmes persisting indefinitely unless a political

coalition musters the will to kill them, programmes die unless a political coalition musters the evidence to sustain them.

Zero-based budgeting cycles. Rather than taking the previous year's budget as the baseline and debating increments, zero-based budgeting requires each programme to justify its entire budget from scratch on a periodic cycle. The cycle forces an explicit re-evaluation of the programme's continued value, generating the information that the forgetting function requires. The cycle must be long enough that the re-evaluation burden is manageable—annual zero-based budgeting for an entire government would be administratively impossible—but frequent enough that programmes do not persist for decades without scrutiny.

Mandatory programme evaluation with automatic termination triggers. Programmes above a specified size threshold are subject to rigorous impact evaluation on a scheduled cycle. Programmes that fail to demonstrate effectiveness against pre-specified criteria across multiple evaluation cycles are automatically terminated, not subject to political discretion. The automatic trigger removes the need to assemble a political coalition to kill a specific programme—a collective action problem that systematically favours the persistence of ineffective programmes whose beneficiaries are concentrated and organised while their costs are diffuse.

Constitutional or statutory review mechanisms. The largest-scale forgetting function is the capacity to peacefully retire entire institutional arrangements that no longer serve their purpose. Constitutional review mechanisms—citizens' assemblies, constitutional conventions, periodic referendums on constitutional continuity—are the institutional forms of this function. They must be designed to be difficult enough to activate that they are not triggered by transient political pressures, and accessible enough that they are available when the gap between the constitutional architecture and the society it governs has become structural.

The forgetting mechanisms must themselves be protected from capture by the interests that benefit from the retention of what should be forgotten. A sunset clause that can be overridden by a simple majority in the legislature is not a forgetting mechanism; it is a renewal ritual. A programme evaluation function whose budget is controlled by the ministry whose programmes it evaluates is not an independent assessor. The forgetting function requires the same structural protections as the exploration function: independence, protected budgets, and pre-committed decision rules.

6.7 Learning Rate Accelerators

The rate at which a governance system learns is constrained by the rate at which it can conduct experiments and the rate at which it can process their results. These rates are not fixed by the nature of governance. They are functions of the infrastructure—data, analytical capacity, simulation capability—that the system maintains. Learning rate accelerators are investments in that infrastructure.

Data infrastructure for near-real-time outcome visibility. The traditional governance learning cycle is slow because outcomes are measured infrequently and reported with long lags. Administrative data systems that capture outcomes continuously—tax receipts, healthcare utilisation, school attendance, regulatory compliance—can compress the observation latency from years to weeks. The investment is not primarily technological; the technology exists. It is institutional: the integration of administrative data systems across agencies, the standardisation of data formats, the resolution of privacy and confidentiality constraints, and the creation of analytical capacity to process the resulting data streams.

Digital twins and policy simulation environments. A digital twin is a computational model of a governance system—a city's transport network, a region's labour market, a national health system—that is calibrated against real-time data and can be used to simulate the effects of policy interventions before they are implemented in the physical world. The digital twin does not replace real-world experimentation—it is itself a model, subject to the same model drift as any other—but it can accelerate the exploration cycle by enabling high-frequency, low-cost, safe-to-fail experimentation in silico. The results of in-silico experiments inform the design of real-world pilots; the results of real-world pilots update the digital twin. The two modes of exploration are complementary.

Embedded methodological capacity. The analytical techniques required for rigorous policy learning—causal inference from observational data, experimental design for randomised controlled trials, Bayesian updating for sequential learning, machine learning for pattern detection in high-dimensional administrative data—are sophisticated and scarce. Embedding this capacity within the governance apparatus, rather than outsourcing it to academic researchers or consultants whose involvement ends when the study is published, is a structural requirement for sustained learning. The capacity must be career-track: positions that attract and retain talented quantitative analysts within the public sector, with competitive compensation, intellectual autonomy, and career progression that does not require moving into generalist management.

Learning repositories and knowledge management. The forgetting-without-learning trap is partly a failure of knowledge management infrastructure. Evaluation results, programme data, and the tacit knowledge of practitioners must be systematically archived, curated, and made accessible to successors. The infrastructure includes not only databases but also the institutional practices—handover documentation, exit interviews, after-action reviews—that ensure knowledge is transferred when personnel turn over. The learning repository is the institutional memory that the forgetting factor $\lambda_f \lambda_f$ represents in the formal model. Its quality determines whether the system's effective $\lambda_f \lambda_f$ is close to unity or close to zero.

6.8 Antifragility Through Stress Exposure

The series' final design principle reframes the concept of antifragility in control-theoretic terms and derives its institutional implications.

A system that never experiences stress cannot learn the parameters that determine its response to stress. This is not a philosophical claim about the virtues of adversity. It is a direct consequence of the persistent excitation condition: if the input signal never excites the system's nonlinear or extreme-regime dynamics, those dynamics are not identifiable from the available data. The system's model of itself is calibrated to the narrow range of conditions it has experienced, and it is maximally uncertain—or, worse, confidently wrong—about its behaviour outside that range.

A governance system that suppresses all variance—all protests, all policy failures, all external shocks, all dissenting information—is therefore not maximally stable. It is maximally fragile. It has eliminated the excitation on which model identification depends. Its apparent stability is the stability of a controller that has never been tested—a stability that is indistinguishable, from inside the system, from the genuine robustness that comes from having been tested and having learned.

The design implication is not that governance systems should seek catastrophe. It is that they should maintain deliberate, bounded exposure to the stressors that reveal their own parameters.

Stress testing of financial and infrastructure systems. Financial regulators should mandate periodic stress tests—simulated crises that reveal the system's vulnerability to shocks—and should vary the scenarios to prevent the system from optimising to a known test. Infrastructure systems should be subjected to load tests that exceed their design specifications, not to break them but to discover where they break. The stress test is a safe-to-fail exploration of the system's extreme-regime dynamics.

Red-team exercises for policy and strategy. A red team is an independent group tasked with challenging a policy, strategy, or assessment from an adversarial perspective. It functions as a deliberate injection of exploration variance into the policy process—a structured mechanism for surfacing assumptions, testing them against alternative frameworks, and revealing vulnerabilities that the consensus model has overlooked. The red team must be protected from the institutions whose policies it challenges, and its findings must be integrated into the decision process rather than shelved.

Citizens' assemblies and dissenting preference channels. A governance system that suppresses protest and dissent is suppressing a signal about its own performance—specifically, the signal that reveals the preferences and experiences of populations whose voices are not transmitted through the standard representation channels. Protected spaces for protest, contestation, and dissenting preference expression are not merely democratic niceties. They are observation channels that reveal parameters of the governed population's preferences and compliance behaviour that would otherwise be invisible to the controller. The system that suppresses them is suppressing the excitation it needs to remain calibrated.

After-action reviews and failure autopsies. When failures occur—and they will, even in well-designed systems—the governance architecture should mandate systematic, public after-action reviews that extract the learning from the failure and embed it in the institutional memory. The review must be protected from the institutions whose failure is being examined, and its findings must be disseminated to the institutions that can act on them. The military's after-action review process and the aviation industry's accident investigation

system are partial models: they are institutionalised mechanisms for extracting maximum information from minimum failure frequency, and they have produced sustained improvements in safety and effectiveness over decades.

The eight design principles form an integrated architecture for adaptive governance. Protected experimental spaces generate the persistent excitation that makes learning possible (6.1). Safe-to-fail structures ensure that the failures inherent in exploration are survivable and informative (6.2). Separation of exploration and exploitation functions protects the learning apparatus from the short-term pressures that would extinguish it (6.3). Protected curiosity budgets provide the dedicated resources that the exploration bonus requires (6.4). Mandatory model review and paradigm replacement cycles prevent model lock-in (6.5). Institutionalised forgetting enables the peaceful retirement of what no longer serves (6.6). Learning rate accelerators increase the speed at which the system can acquire and process information (6.7). And antifragility through stress exposure ensures that the system's model is calibrated across the full range of conditions it may encounter (6.8).

None of these principles is easy to implement. Each confronts the political incentives that make short-term exploitation more attractive than sustained exploration, and each requires institutional protections that the immune system will resist. But the alternative is not a different governance strategy. It is the continued operation of the learning failure modes that Part III diagnosed—exploration starvation, model lock-in, exploitation lock-in, learning-induced oscillation, and the forgetting-without-learning trap—until the accumulated gap between the system's model and its environment breaches a crisis threshold. The structural logic is clear. The design requirements follow from it. The task of implementation is the work of building institutions that take their own learning as seriously as they take their performance.

Part VII — Connection to the Series

This paper is the fourteenth in a sequence that began with the observation that governance systems fail in structurally predictable ways. The preceding papers have built a grammar of primitives, diagnosed failure modes across fifteen country cases and six organisational domains, and developed design principles for architectures that can perceive, decide, and act. This part places the adaptive controller in the context of that accumulated architecture—showing what it consolidates, how it completes the Cycle Two triad, and where it opens the path forward.

7.1 Consolidation of Three Streams

Three concerns have run through the series without, until now, being given a unified formal treatment.

The first is the tension between exploration and exploitation. It has been implicit since Paper I identified the latency-gain ceiling: a controller that responds too aggressively to the signals it receives will oscillate, while one that responds too cautiously will drift. The optimal response depends on parameters—the system's natural dynamics, the disturbance structure, the noise environment—that the controller must know. But how does it come to know them? The question has been deferred across thirteen papers, answered implicitly by assuming the controller's model is adequate. This paper makes the question explicit and gives it a formal home.

The second is institutional memory. The country studies are filled with systems that have learned and forgotten, learned and forgotten, in cycles that never accumulate. Nigeria's substrate deficit, the UK's repeated health system reorganisations, the corporate amnesia that follows mergers and leadership transitions—each is an instance of a forgetting rate that exceeds a learning rate. The series has diagnosed the pattern without formalising the dynamics. This paper provides the formalisation.

The third is antifragility. The concept has been invoked in design discussions—the idea that systems should gain from stressors, that variance is not only a threat but a resource—without being given a control-theoretic foundation. This paper provides that foundation. The persistent excitation condition gives rigorous content to the intuition that systems that never experience stress cannot learn the parameters that determine their response to stress. Antifragility, in this reframing, is not a property that some systems mysteriously possess. It is a structural consequence of maintaining sufficient signal diversity to support ongoing model identification.

The consolidation is not merely taxonomic. Each of these three concerns turns out to be a facet of the same underlying problem: the controller does not know the system it is controlling, and it must act in ways that generate information about the system while continuing to regulate it. The dual control framework unifies

them. Exploration–exploitation is the objective function. Institutional memory is the estimator's forgetting factor. Antifragility is the excitation condition. Together they constitute the design problem of governance as an adaptive controller.

7.2 The Cycle Two Learning Triad

The series' second cycle has traced a specific trajectory, and this paper completes its foundational arc. But the arc is not a single line. It is a triad.

Paper X established the *sensing* requirement. A governance system that relies on a single observation infrastructure, or on multiple infrastructures with correlated errors, is vulnerable to systematic blind spots that are invisible from within the system. The solution is observer diversity: an ensemble of sensing channels with decorrelated biases, capable of detecting the errors that any single channel would miss. Observer diversity is the structural precondition for knowing that the current model is wrong.

Paper IX established the *actuation* requirement. Knowing that the model is wrong, and even knowing what a better model would be, is insufficient if the system cannot execute the transition. The immune system defends the existing architecture. Incumbent interests absorb reform energy. The transition bandwidth—the rate at which the system can peacefully redesign its own structure—is the binding constraint on adaptation. Transition bandwidth is the structural precondition for acting on what is known.

This paper establishes the *learning* requirement that mediates between them. Sensing detects the need for change. Actuation executes the change. But between detection and execution lies the question: change *to what*? The system must discover the parameters of the alternative—the policy multiplier that differs from the model's assumption, the compliance elasticity that the old regulatory framework never tested, the implementation capacity that the previous architecture never probed. Learning is the process by which the gap between the current model and the better model is closed. And learning itself has a structure—the dual control structure, the exploration–exploitation trade-off, the persistent excitation condition—that must be designed for.

The sequence is causal: **Sense** → **Learn** → **Execute**. Observer diversity detects that the model is drifting. Dual control discovers the direction and magnitude of the drift. Transition bandwidth enables the architecture to be modified in response. A system that lacks any one of these three capacities is a system that cannot adapt. It may be able to perceive its own inadequacy without knowing what to do about it. It may know what to do without being able to do it. It may act decisively on a model that is silently becoming obsolete. Only the system that possesses all three—sensing, learning, actuation—can remain calibrated to an environment that changes while it governs.

The triad is the series' answer to the question that Paper VI first raised: how can a governance architecture remain adequate when the variety of its environment continuously expands? The answer is that it cannot—unless the architecture itself contains the mechanisms for its own ongoing redesign. The triad specifies what

those mechanisms must be.

7.3 The Two Triads of Cycle Two

With this paper, Cycle Two now consists of two complementary triads. They address different aspects of the governance problem, and they operate at different levels of the architecture.

The *architecture* triad addresses the static design of the controller and its relationship to the governed system. Paper XI (delegation depth) established that actuation fidelity degrades superlinearly with the number of organisational layers through which a directive must pass, and that beyond a critical depth the energy required to realise policy intent becomes prohibitive. Paper XII (boundary selection) established that a controller with perfect internal observation and actuation can still fail if its jurisdictional perimeter excludes causally relevant dynamics, and that the resulting M- Δ loop generates instability that no amount of internal competence can correct. Paper XIII (legitimacy) established that the willingness of the governed to comply with directives and report honestly is an emergent coupling state that multiplies the effectiveness of every architectural choice, and that a controller that ignores its own legitimacy level is operating blind on the parameter that determines whether the loop closes.

Together, these three papers describe the controller as it is: an actuation chain, a boundary, and a legitimacy substrate. They specify the conditions under which each functions adequately, and the failure modes that arise when they do not.

The *adaptation* triad addresses the dynamic capacity of the controller to modify that architecture over time. Paper X (observer diversity) established that the observing ensemble must possess sufficient decorrelation to detect its own systematic errors—the sensing leg. Paper XIV (adaptive control) establishes that the controller must be able to explore beyond its current model, retain what it discovers, and discard what no longer holds—the learning leg. Paper IX (transition bandwidth) establishes that the system must possess the capacity to execute structural change against incumbent resistance—the actuation leg.

The two triads are not independent. The adaptation triad operates on the architecture triad. Observer diversity detects that the observation channels are degraded, or that the boundary is mismatched, or that legitimacy is collapsing. Dual control discovers what the improved channels, boundary, or legitimacy-building strategy should be. Transition bandwidth executes the redesign. And the architecture triad, once modified, becomes the new substrate on which the adaptation triad continues to operate.

The relationship is recursive. A governance system that possesses both triads is a system that can observe itself, learn about itself, and modify itself—a system that is, in the precise sense of second-order cybernetics, a regulator that regulates its own regulation. The series has been building toward this architecture since Paper I's first diagram of a feedback loop. The two triads are its completion.

7.4 Second-Order Cybernetics: Regulating the Regulator

The series' trajectory reflects a deepening insight that is worth making explicit, because it is the intellectual arc that unifies the fourteen papers.

Papers I through VII are primarily first-order. They study the controller as a given architecture—an observation channel, a decision process, an actuation chain—and analyse the conditions under which that architecture can stabilise the system it governs. The failure modes they diagnose are failures of regulation: the controller cannot perceive the system, cannot respond fast enough, cannot match its variety, cannot transmit its intent.

Papers VIII through XIII begin the transition to second-order. Paper VIII provides the measurement framework that makes architectural deficits visible—the controller observing itself. Paper IX provides the transition bandwidth concept—the controller modifying itself. Paper X provides the observer diversity requirement—the controller ensuring that its own observation of itself is not systematically corrupted. Paper XIII provides the legitimacy dynamics—the controller recognising that its own effectiveness depends on a parameter it does not directly control.

This paper completes the transition. The dual controller is a controller that does not merely regulate the system. It regulates its own regulation. It maintains a model of the system, and it maintains a model of its own uncertainty about that model. It chooses actions not only for their immediate effect on the system's state but for their effect on its own future knowledge. It is, in the language of cybernetics, a system that observes its own observing and acts to improve it.

The distinction between first-order and second-order governance is not a philosophical nuance. It is a structural requirement. In a stationary environment, a well-designed first-order controller may be adequate indefinitely. The system's dynamics do not change; the controller's model, once learned, remains valid. But in a non-stationary environment—an environment in which the disturbance dimensions multiply, the coupling structures shift, the governed population's preferences and compliance behaviour evolve—the first-order controller is eventually rendered obsolete by changes it never learned to track. The only sustainable architecture is one that can modify its own model, and eventually its own structure, in response to what it discovers.

The series has, from its opening pages, insisted that governance failure is structural before it is personal—that the architecture generates the outcomes, and that improving the architecture requires understanding the constraints that shape it. This paper extends that insistence to the architecture of the architecture. The capacity to learn is not a gift of wise leadership or a by-product of competent administration. It is a structural property that must be designed into the governance system, protected from the incentives that would extinguish it, and sustained across the timescales over which environmental change outpaces institutional adaptation. The series' deepest claim may be this: a governance system that cannot regulate its own regulation is not fully a

governance system at all. It is a machine executing a programme that is slowly becoming obsolete, and it will continue to execute it until the gap between the programme and reality becomes too large to ignore—at which point the adjustment, if it comes at all, will come as catastrophe.

7.5 Connection to the AI Alignment Frontier

This paper's formal framework—dual control, exploration–exploitation, persistent excitation, catastrophic forgetting—is not specific to human governance systems. It is the same mathematical structure that governs recursive self-improvement in artificial intelligence.

An AI system that is capable of modifying its own architecture—its own objective function, its own model of the world, its own learning algorithm—faces exactly the dual control problem this paper analyses. It must balance exploitation of its current capabilities against exploration of potential improvements. It must maintain sufficient persistent excitation in its own training signal to identify the parameters of its own performance. It must retain what it learns across updates (avoiding catastrophic forgetting) while discarding what no longer serves (avoiding model lock-in). It must ensure that its exploratory modifications do not destabilise the very capabilities they are intended to enhance—the learning-induced oscillation failure mode, transposed to the AI context, is the problem of safe exploration in recursive self-improvement.

The convergence is not coincidental. Both human governance systems and AI systems face the same structural problem: how can a controller that does not fully know its own dynamics modify itself without destroying itself? The mathematics that describes the problem is the same. The design principles that emerge from the mathematics—protected experimental spaces as safe-to-fail exploration environments, separation of exploration and exploitation functions, mandatory model review and paradigm replacement cycles, institutionalised forgetting—have analogues in the AI alignment literature under different names: sandboxing, reward modelling, oversight protocols, capability elicitation, interpretability research.

This paper does not solve the AI alignment problem. But it establishes that the problem of designing governance systems that can safely learn to modify themselves and the problem of designing AI systems that can safely learn to modify themselves are instances of the same formal problem. The governance engineering framework that the series has built, and the adaptive control extension that this paper provides, are tools for thinking about both. The bridge between the cybernetics of human institutions and the cybernetics of machine intelligence is not a metaphor. It is a shared mathematical structure, and the work of understanding it has barely begun.

7.6 Bridge to Cycle Three

The series' research roadmap envisions three cycles. Cycle One—"The Cybernetics of Governance"—established the foundational grammar: the structural primitives, the failure modes, the diagnosis of fifteen country cases. Cycle Two—"The Evolution of Governance"—extended that grammar to the dynamics of

adaptation: observer diversity, transition bandwidth, legitimacy as emergent gain, and now adaptive control. Cycle Three—"The Engineering of Governance"—is the transition from theory to practice: diagnostics as services, the simulator as a calibrated tool, intervention-effect estimation, and institutional design patterns.

This paper is the last of the theoretical papers in Cycle Two. It consolidates the adaptation triad and completes the second-order cybernetic architecture that the series has been building toward. With its completion, the series possesses:

- A **grammar** of governance primitives (Papers I–VIII): the measurable structural parameters that determine whether a controller can perceive, decide, and act.
- A **diagnosis** of the characteristic failure modes that arise when those primitives are set badly (Papers I–VII), applied across country cases and organisational domains.
- A **theory of adaptation** (Papers IX, X, XIV): the conditions under which a governance system can sense the need for change, learn what change to make, and execute it against resistance.
- A **design vocabulary** (Papers VI, IX, X, XII, XIII, XIV): the principles that an architecture must satisfy to avoid the diagnosed failure modes and to maintain adaptive capacity over time.
- A **measurement framework** (Paper VIII, with extensions in Papers XII and XIII): the protocols for estimating the structural parameters from available data, with explicit uncertainty quantification.

What the series does not yet possess is empirical validation of its core predictions at scale, calibrated simulation tools that can be applied to specific institutional design problems, and a community of practitioners trained in the framework and equipped to use it. Those are the tasks of Cycle Three. They are engineering tasks, not primarily theoretical ones. They require data, collaboration with governance institutions willing to subject their architectures to structural audit, and the patient work of building tools that make the framework usable by actors who are not its authors.

This paper provides the theoretical foundation for that engineering work. The adaptive controller is the most demanding architecture the series has described. It is also the one that most governance systems most lack. Building it—designing institutions that can sustain exploration, retain what they learn, discard what no longer holds, and modify themselves in response—is the work of the coming decades. The series has provided the grammar and the design principles. The rest belongs to the builders.

Part VIII — Limitations and Conclusion

8.1 Limitations

The argument of this paper is structural: a governance system that cannot learn—that starves exploration, locks in obsolete models, forgets faster than it learns, or blocks the translation of knowledge into action—has a finite operational lifetime in an environment that changes faster than architectures can be redesigned. This argument has been developed through a formal framework, a simulation, and empirical illustrations. It is subject to limitations that should be stated clearly.

The formal framework is a first-order approximation. Dual control theory in its full generality is computationally intractable for realistically complex governance systems. The Bellman equation of Section 2.2 cannot be solved exactly when the state space, the parameter space, and the action space are all high-dimensional. The paper relies on approximations—the separation of certainty-equivalent and exploration components, the recursive least-squares estimator, the scalar exploration variance—that capture the essential dynamics while abstracting from the full complexity of the optimal dual control policy. The design principles derived from the framework are robust to these approximations in their qualitative direction, but the quantitative calibration of exploration budgets, model review cycles, and forgetting rates for specific governance domains requires empirical work that has not yet been conducted.

The design principles have not been empirically validated as a set. The paper identifies eight design principles and illustrates them through case material, but it does not provide a systematic empirical test of their joint effectiveness. Do governance systems that institutionalise protected experimental spaces, safe-to-fail structures, and mandatory model review cycles actually exhibit superior adaptive capacity over extended periods? The case illustrations are suggestive, but they are not validations. The empirical programme that follows this paper—applying the measurement framework of Paper VIII to the learning parameters identified here, and testing whether systems with higher measured adaptive capacity exhibit superior long-run performance—is the necessary next step.

The treatment of institutional forgetting is preliminary. The forgetting factor $\lambda_f \lambda_f$ provides a tractable formalisation of institutional memory decay, but the actual dynamics of forgetting in governance systems are more complex than a single exponential decay parameter. Personnel turnover is lumpy, not continuous. Knowledge is embedded in relationships and practices, not only in documents and databases. Some knowledge is actively destroyed by incoming administrations; some is passively lost through neglect; some is deliberately preserved by career civil servants who outlast political transitions. The mapping between the formal parameter and the institutional reality requires richer modelling than this paper provides.

The paper brackets the normative question of learning objectives. The dual controller learns in order to minimise a cost function that is taken as given. But the cost function of a governance system—the value architecture of Paper VI—is itself a subject of political contestation. A system that learns efficiently to pursue objectives that are unjust, or that serve a narrow elite, or that discount the interests of future generations, is a system that learns to do harm efficiently. The paper provides the structural conditions for learning; it does not address the normative question of what should be learned, or who should decide. That question belongs to the value architecture of Paper VI and to the democratic theory that the series' engineering idiom deliberately brackets.

The AI alignment connection is stated but not developed. Section 7.5 notes the structural convergence between the dual control problem in governance and the safe exploration problem in recursive AI self-improvement. The convergence is real, but the paper does not develop it beyond the observation. The translation of the design principles from governance to AI alignment—protected experimental spaces as sandboxes, separation of exploration and exploitation functions as oversight architectures, mandatory model review as interpretability research—is a research programme in its own right, and this paper provides only the starting point.

The simulation is illustrative, not calibrated. The simulation of Part IV demonstrates that the qualitative dynamics of the framework emerge reliably from a minimal set of assumptions. The parameters are chosen to make the mechanisms visible, not to calibrate against any specific real-world governance system. The quantitative results—the specific location of the stable-learning region, the specific exploration variance that minimises tracking error for a given environmental change rate—are artefacts of the parameter choices. The qualitative claim—that the exploration-starvation trap, exploitation lock-in, and the forgetting-without-learning threshold are structural features of any system in which a controller must learn while it acts—is the simulation's contribution, and it is robust to parameter variation.

Persistent excitation captures one aspect of antifragility, not the whole concept. The reframing of antifragility as persistent excitation (Section 2.4) provides rigorous content to the intuition that systems that never experience stress cannot learn. But the full concept of antifragility, as developed by Taleb and others, includes additional dimensions—convex payoff functions, optionality, redundancy, hormesis—that are not captured by the excitation condition alone. The paper's claim is not that persistent excitation exhausts the meaning of antifragility, but that it provides a control-theoretic foundation for one of its central mechanisms. The other dimensions remain to be formalised.

8.2 Conclusion

This paper began with a simple observation: a governance system that cannot learn has a finite operational lifetime. The environment changes—new disturbance dimensions emerge, the governed population's preferences and compliance behaviour evolve, the coupling structures that determine spillovers shift—and the

architecture that was adequate for yesterday's conditions becomes tomorrow's constraint. The only sustainable response is an architecture that can modify its own structure in response to what it discovers.

The paper has provided the formal grammar for that response. The dual control framework establishes that every policy intervention is simultaneously an action and an experiment, and that the controller who designs interventions to be informative survives while the controller who suppresses the information in its own actions eventually governs a phantom. The persistence of excitation condition gives rigorous content to the intuition that variation is not a threat to stability but a precondition for it. The five failure modes—exploration starvation, model lock-in, exploitation lock-in, learning-induced oscillation, and the forgetting-without-learning trap—are the characteristic pathologies of governance systems that have not been designed for learning. The eight design principles—protected experimental spaces, safe-to-fail structures, functional separation of exploration and exploitation, protected curiosity budgets, mandatory model review, institutionalised forgetting, learning rate accelerators, and antifragility through stress exposure—are the architectural response.

The paper completes the Cycle Two adaptation triad. Observer diversity (Paper X) detects the need for change. Dual control (Paper XIV) discovers what change to make. Transition bandwidth (Paper IX) executes the change against incumbent resistance. The sequence—sense, learn, execute—is the series' answer to the question of how governance systems remain adequate to environments that change faster than architectures can be redesigned.

The series' arc is now visible in full. It began with the engineering of the controller—the primitives of latency, signal fidelity, representation depth, delegation depth, and boundary selection. It progressed through the measurement of the controller's own adequacy—the variety gap, the measurement paradox, the observer ensemble. It incorporated the emergent coupling state—legitimacy—that determines whether the architecture functions at all. And it arrives, here, at the controller that can modify itself: the adaptive architecture that learns.

Classical governance regulates society. Adaptive governance regulates the regulator itself. This is not a philosophical preference. It is a structural imperative. In an environment that continuously generates novel disturbance dimensions, a controller that cannot modify its own architecture is eventually rendered obsolete by the environment's modifications. The design principles this paper offers are a framework for building controllers that can keep pace. The rest is the work of building them.

The series has now completed its theoretical arc. The grammar of primitives is established. The diagnosis of failure modes is developed across fifteen country cases and six organisational domains. The measurement framework exists in prototype. The theory of adaptation—sensing, learning, executing—is specified. The design vocabulary is articulated. What remains is to test the predictions, calibrate the tools, and build the institutions. The next phase is not more theory. It is engineering. And the engineering of governance, like all engineering, begins with the acknowledgment that the structures we inherit are not the structures we need—and that we possess the means to build better ones.

Appendix A — Formal Derivations

This appendix provides the mathematical derivations underlying the dual control framework of Part II. It formalises the governance dual control problem as a stochastic dynamic programming problem, derives the persistent excitation condition for identifiability of governance parameters, and analyses the forgetting dynamics that govern institutional memory decay.

A.1 Dual Control Formulation for Governance Systems

Consider a governance system whose true dynamics are given by

$$\begin{aligned} x(t+1) &= f(x(t), u(t), \theta) + w(t), & w(t) &\sim N(0, W), \\ x(t+1) &= f(x(t), u(t), \theta) + w(t), & w(t) &\sim N(0, W), \end{aligned}$$

where $x(t) \in R^n$, $x(t) \in R^n$ is the state vector, $u(t) \in R^m$, $u(t) \in R^m$ is the control input, $\theta \in R^p$, $\theta \in R^p$ is a vector of unknown parameters, and $w(t)$ is process noise. The controller observes

$$\begin{aligned} y(t) &= Cx(t) + v(t), & v(t) &\sim N(0, V), \\ y(t) &= Cx(t) + v(t), & v(t) &\sim N(0, V), \end{aligned}$$

and maintains a belief distribution $b_t = p(\theta | I_t)$, where $I_t = \{y(0), u(0), \dots, y(t-1), u(t-1), y(t)\}$ is the information available at time t . The belief is updated via Bayes' rule:

$$\begin{aligned} p(\theta | I_{t+1}) &\propto p(\theta | I_t) p(y(t+1) | x(t), u(t), \theta), \\ p(\theta | I_{t+1}) &\propto p(\theta | I_t) p(y(t+1) | x(t), u(t), \theta). \end{aligned}$$

The controller's objective is to minimise the expected cumulative discounted cost over a horizon T :

$$J = E \left[\sum_{t=0}^T \gamma^t c(x(t), u(t)) \right],$$

$$J = E[\sum_{t=0}^T \gamma^t c(x(t), u(t))],$$

where $c(\cdot)$ penalises deviations from the target state and excessive control effort, and $\gamma \in (0, 1]$ is the discount factor.

The optimal policy for this problem satisfies the Bellman equation:

$$V_t(b) = \min_u E_{x,\theta} \left[c(x, u) + \gamma V_{t+1}(b') \mid b, u \right], \quad (\text{A.1})$$

$$V_t(b) = \min_u E_{x,\theta} [c(x, u) + \gamma V_{t+1}(b') \mid b, u], (\text{A.1})$$

where b' is the posterior belief after observing the outcome of action u . The expectation is taken over the current state x , the unknown parameters θ , and the process and measurement noise.

The critical feature of (A.1) is that the choice of u affects not only the immediate cost but also the *future belief state* b' . This is the dual effect: the control action influences both the state evolution (regulation) and the informativeness of future observations (identification). The optimal policy therefore includes an explicit exploration incentive.

Certainty-equivalence and the exploration bonus.

When the system is linear and the cost is quadratic, and when the parameter uncertainty is small, the optimal dual control can be approximated by decomposing the value function. Let $\hat{\theta} = E[\theta \mid b]$ be the current parameter estimate, and let $P = \text{Cov}[\theta \mid b]$ be the estimation error covariance. The value function can be expanded around the certainty-equivalent value:

$$V(b) \approx V^{\text{CE}}(\hat{\theta}) + \text{tr}(PH(\hat{\theta})),$$

$$V(b) \approx V^{\text{CE}}(\hat{\theta}) + \text{tr}(PH(\hat{\theta})),$$

where V^{CE} is the value of the optimal policy when $\hat{\theta}$ is assumed to be the truth, and H is a positive semi-definite matrix that quantifies the sensitivity of future performance to parameter uncertainty. The second term is the *cost of uncertainty*: the expected performance degradation due to not knowing the true parameters.

The optimal action can then be written as

$$u^*(t) = u_{\text{CE}}(t) + u_{\text{explore}}(t),$$

$$u^*(t) = u_{\text{CE}}(t) + u_{\text{explore}}(t),$$

where $u_{\text{CE}}(t)$ minimises the certainty-equivalent cost and $u_{\text{explore}}(t)$ is a perturbation chosen to reduce P in the directions that matter most for future performance—i.e., the directions in which H is largest. The magnitude of u_{explore} scales with the current uncertainty P and with the sensitivity H . When uncertainty is high, exploration is more aggressive; as the parameters are learned, exploration decays and the controller converges to certainty-equivalence.

In the simulation of Part IV, this structure is implemented with a constant exploration variance σ_{η}^2 for tractability. The constant-variance approximation captures the essential trade-off—sustained exploration is necessary when the environment continues to change—while abstracting from the optimal scheduling of the exploration intensity.

A.2 Persistent Excitation and Identifiability

For the linear special case $x(t+1) = Ax(t) + Bu(t) + w(t)$ with unknown matrices A, B , the parameters can be estimated from input-output data only if the input signal satisfies a *persistent excitation* condition.

Let $\phi(t) = [x(t)^\top, u(t)^\top]^\top \in \mathbb{R}^{n+m}$ be the regressor vector. The system dynamics can be written as

$$x(t+1)^\top = \phi(t)^\top \Theta + w(t)^\top,$$

$$x(t+1)^\top = \phi(t)^\top \Theta + w(t)^\top,$$

where $\Theta = [A \mid B]^\top \in \mathbb{R}^{(n+m) \times n}$ is the parameter matrix. The least-squares estimator of Θ after T observations solves

$$\hat{\Theta}_T = (\Phi_T^\top \Phi_T)^{-1} \Phi_T^\top X_T,$$

$$\hat{\Theta}_T = (\Phi_T^\top \Phi_T)^{-1} \Phi_T^\top X_T,$$

where $\Phi_T = [\phi(0), \dots, \phi(T-1)]^\top$ and $X_T = [x(1), \dots, x(T)]^\top$. The estimator exists and is unique only if $\Phi_T^\top \Phi_T$ is invertible. More generally, the parameters are *identifiable* if the information matrix grows linearly with T .

The input signal $u(t)$ is *persistently exciting* of order d if there exist $\alpha > 0$ and an integer m such that, for all t ,

$$\alpha I \preceq \sum_{k=t}^{t+m} \phi(k) \phi(k)^\top. \quad (\text{A.2})$$

$$\alpha I \preceq \sum_{k=t}^{t+m} \phi(k) \phi(k)^\top. (\text{A.2})$$

Condition (A.2) ensures that the regressor vector varies sufficiently in all directions to uniquely determine the parameters. If the input is constant or varies only within a subspace of \mathbb{R}^{n+m} , the information matrix becomes rank-deficient and some parameters are unidentifiable regardless of the observation duration.

The governance analogue is direct. Consider a governance system with p unknown policy-relevant parameters—elasticities, multipliers, compliance rates, implementation capacities. To identify these parameters, the policy vector $u(t)$ must vary across at least p independent directions over any sufficiently long window. A system that applies the same policy instruments at the same settings, year after year, generates a regressor matrix whose columns are nearly collinear. The parameters that govern the system's response to conditions it has never encountered—the response to a novel crisis, the effectiveness of an untried instrument, the capacity of an untested delivery chain—are unidentified. The controller may observe the system indefinitely and never learn these parameters.

The minimum exploration variance required for identifiability scales with the noise level $\|W\| \|W\|$ and with the dimension of the unknown parameter vector. In the simulation of Part IV, the exploration dither $\sigma_\eta^2 \sigma_\eta^2$ must be large enough relative to the process noise to ensure that the information matrix $\Phi^\top \Phi \Gamma \Phi$ remains well-conditioned. When $\sigma_\eta^2 \sigma_\eta^2$ falls below this threshold, the parameter estimates drift away from the true values without the controller being able to detect the drift—the formal mechanism of the exploration-starvation trap.

A.3 Forgetting and the Effective Sample Size

In a non-stationary environment, the controller must track slowly changing parameters. The standard approach is recursive least squares with a forgetting factor. The estimator updates the parameter estimate $\hat{\theta}(t)$ $\theta^\wedge(t)$ and the inverse information matrix $P(t)P(t)$ as

$$\begin{aligned}
 K(t) &= \frac{P(t-1) \phi(t)}{\lambda_f + \phi(t)^\top P(t-1) \phi(t)}, \\
 \hat{\theta}(t) &= \hat{\theta}(t-1) + K(t) (y(t) - \phi(t)^\top \hat{\theta}(t-1)), \\
 P(t) &= \frac{1}{\lambda_f} (I - K(t) \phi(t)^\top) P(t-1),
 \end{aligned}$$

$$K(t)\theta^\wedge(t)P(t) = \lambda_f + \phi(t)^\top P(t-1) \phi(t)P(t-1) \phi(t), = \theta^\wedge(t-1) + K(t)(y(t) - \phi(t)^\top \theta^\wedge(t-1)), = \lambda_f 1$$

$$(I - K(t) \phi(t)^\top) P(t-1),$$

where $\lambda_f \in (0, 1]$ $\lambda_f \in (0, 1]$ is the forgetting factor. When $\lambda_f = 1$ $\lambda_f = 1$, all past observations are weighted equally; the effective sample size grows without bound, and the estimator converges to the true parameters (if the environment is stationary). When $\lambda_f < 1$ $\lambda_f < 1$, past observations are exponentially down-weighted with a half-life of approximately $1/(1 - \lambda_f)$ $1/(1 - \lambda_f)$ time steps.

The effective sample size—the number of observations that contribute meaningfully to the current estimate—is bounded above by

$$N_{\text{eff}} \leq \frac{1}{1 - \lambda_f}. \quad (\text{A.3})$$

$$N_{\text{eff}} \leq 1 - \lambda_f 1. (\text{A.3})$$

When $\lambda_f = 0.99$ $\lambda_f = 0.99$, the effective memory is approximately 100 time steps. When $\lambda_f = 0.90$ $\lambda_f = 0.90$, the effective memory is approximately 10 time steps—the controller remembers only the most recent decade of experience, and all learning from before that is effectively forgotten.

The net learning condition follows. Let r_l be the rate of information acquisition—the rate at which new observations reduce the parameter uncertainty $\text{tr}(P)$. Let $r_f = 1 - \lambda_f$ be the rate of forgetting—the rate at which old information decays. The steady-state uncertainty satisfies

$$\lim_{t \rightarrow \infty} \text{tr}(P(t)) \approx \frac{\text{tr}(W)}{r_l - r_f},$$

$$\lim_{t \rightarrow \infty} \text{tr}(P(t)) \approx r_l - r_f \text{tr}(W),$$

when $r_l > r_f$. When $r_l \leq r_f$, the uncertainty diverges: the system forgets faster than it learns, and the parameter estimates never converge.

The governance analogue is that the institutional forgetting rate is determined by personnel turnover, organisational restructuring, and the decay of knowledge management infrastructure. The effective sample size of institutional memory is the number of past administrations, reform cycles, or programme evaluations whose learning remains accessible to current decision-makers. When this effective sample size is smaller than the number of observations required to identify the system's key parameters—given the noise in the governance environment and the rate of environmental change—the system is in the forgetting-without-learning trap.

The mapping between the formal parameter λ_f and institutional characteristics is not exact. But the structural direction is clear. Democracies with short electoral cycles, high ministerial turnover, and weak civil service continuity operate with a low effective λ_f . Systems with strong career bureaucracies, institutionalised evaluation repositories, and mandatory knowledge transfer protocols operate with a higher effective λ_f . The difference in λ_f determines whether the system can accumulate the knowledge required to remain calibrated to a changing environment, or whether each generation of decision-makers must rediscover what its predecessors already learned.

Appendix B — Simulation Specification

This appendix provides the detailed specification for the simulation described in Part IV. It defines the system dynamics, the controller architecture, the six scenarios, the parameter sweeps, and the output metrics. The specification is sufficient to implement the simulation independently.

B.1 Model Specification

The simulated governance system controls a two-dimensional state vector $x(t) = [x_1(t), x_2(t)]^\top \in \mathbb{R}^2$, representing two policy-relevant dimensions. The true dynamics are linear with unknown, slowly time-varying parameters:

$$x(t+1) = A(\theta_t) x(t) + B(\theta_t) u(t) + w(t), \quad w(t) \sim N(0, W),$$

$$x(t+1) = A(\theta_t) x(t) + B(\theta_t) u(t) + w(t), \quad w(t) \sim N(0, W),$$

where $W = 0.01 I_2$. The nominal design dynamics are $A_0 = 0.95 I_2$ and $B_0 = I_2$, but the true parameters drift over time:

$$\theta_{t+1} = \theta_t + \eta_t, \quad \eta_t \sim N(0, \sigma_\theta^2 I),$$

$$\theta_{t+1} = \theta_t + \eta_t, \quad \eta_t \sim N(0, \sigma_\theta^2 I),$$

where σ_θ^2 is the *environmental change rate*. The parameter vector θ_t encodes the diagonal entries of AA and the entries of BB ; for simplicity, AA remains diagonal and BB remains full, with each entry following an independent random walk. The initial true parameters are drawn from $N(\theta_0, 0.01 I)$, where θ_0 corresponds to A_0, B_0 .

The controller observes the state through a noisy channel:

$$y(t) = x(t) + v(t), \quad v(t) \sim N(0, V_0),$$

$$y(t) = x(t) + v(t), \quad v(t) \sim N(0, V_0),$$

with $V_0 = 0.05 I_2$. The observation channel is held at its designed fidelity throughout—this simulation isolates learning dynamics by assuming the sensing architecture is intact.

The controller's objective is to minimise the cumulative squared tracking error over the simulation horizon, with the target state at the origin $x^* = 0$.

B.2 Controller Architecture

State estimation. The controller estimates the state using a Kalman filter with the nominal dynamics A_0, B_0 and the true observation noise V_0 . The Kalman filter is provided as a standard recursion (see Paper XIII Appendix A.2). The state estimate is denoted $\hat{x}(t)$.

Parameter estimation. The controller maintains a running estimate of the unknown parameters using recursive least squares (RLS) with a forgetting factor $\lambda_f \in (0, 1]$. The regressor vector at time t is $\phi(t) = [\hat{x}(t)^\top, u(t)^\top]^\top \in \mathbb{R}^4$. The RLS update follows the standard recursion given in Appendix A.3, with initial parameter estimate $\hat{\theta}(0) = \theta_0$ and initial inverse information matrix $P(0) = 10 I_4$. The parameter estimate is denoted $\hat{\theta}(t)$.

Control law. The controller computes the *certainty-equivalent* action as the LQR optimal control for the estimated dynamics:

$$u_{CE}(t) = -K(\hat{\theta}(t)) \hat{x}(t),$$

$$u_{CE}(t) = -K(\hat{\theta}(t)) \hat{x}(t),$$

where $K(\hat{\theta})$ solves the discrete algebraic Riccati equation for (\hat{A}, \hat{B}, Q, R) with $Q = I_2$ and $R = 0.1 I_2$. The exploration component is a Gaussian dither:

$$u_{\text{explore}}(t) \sim N(0, \sigma_\eta^2 I_2),$$

$$u_{\text{explore}}(t) \sim N(0, \sigma_\eta^2 I_2),$$

where σ_η^2 is the *exploration variance*. The total intended control action is

$$u(t) = u_{CE}(t) + u_{\text{explore}}(t).$$

$$u(t) = u_{CE}(t) + u_{\text{explore}}(t).$$

Actuation efficiency. The effective control reaching the system is

$$u_{\text{eff}}(t) = \mu u(t),$$

$$u_{\text{eff}}(t) = \mu u(t),$$

where $\mu \in [0, 1]$ is the *actuation efficiency*, representing the fraction of intended control that survives the implementation chain. When $\mu = 1$, the actuation chain is intact. When $\mu < 1$, actuation is attenuated, modelling the delegation depth effects of Paper XI or the immune system blockage of Section 3.3.

B.3 Scenarios

Six scenarios are simulated. All use the same plant dynamics and RLS estimator. They differ in the exploration variance σ_η^2 , the environmental change rate σ_θ^2 , the forgetting factor λ_f , and the actuation efficiency μ .

Scenario 1 — Optimal dual control.

$\sigma_\eta^2 = 0.05$, $\sigma_\theta^2 = 0.002$, $\lambda_f = 0.99$, $\mu = 1.0$. This is the baseline: the controller maintains persistent moderate exploration, the environment changes slowly, memory is strong, and actuation is intact. The system learns stably.

Scenario 2 — Exploitation-only (certainty-equivalent).

$\sigma_\eta^2 = 0$, all other parameters as Scenario 1. The controller suppresses exploration entirely. As the environment drifts, parameter estimates diverge and tracking error grows—exploration starvation.

Scenario 3 — Crisis-driven learning.

The controller operates with $\sigma_\eta^2 = 0$ until the tracking error $\|x(t) - \hat{x}(t)\|$ exceeds $e_{\text{crit}} = 2.0$, at which point it switches to $\sigma_\eta^2 = 0.5$ for a fixed duration of $T_{\text{explore}} = 20$ time steps before returning to $\sigma_\eta^2 = 0$. All other parameters as Scenario 1. This produces a boom–bust learning cycle.

Scenario 4 — Over-exploration.

$\sigma_\eta^2 = 0.5$ continuously, all other parameters as Scenario 1. The dither is so large that the controller's own perturbations dominate the system's dynamics, obscuring the parameters and degrading performance.

Scenario 5 — Forgetting-without-learning.

$\sigma_\eta^2 = 0.05$, $\sigma_\theta^2 = 0.005$, $\lambda_f = 0.90$, $\mu = 1.0$. The controller explores moderately, but the environment changes moderately fast and institutional memory is weak. The forgetting rate exceeds the learning rate; parameter estimates remain noisy and biased.

Scenario 6 — Exploitation lock-in.

$\sigma_\eta^2 = 0.05$, $\sigma_\theta^2 = 0.002$, $\lambda_f = 0.99$, $\mu = 0.3$. The controller learns accurately—parameter estimates track the true parameters closely—but only 30% of the intended control reaches the system. Performance is poor despite accurate learning.

B.4 Parameter Sweeps

Three sweeps are conducted to map the boundaries of stable learning.

Sweep 1 — Exploration variance vs. environmental change rate.

$\sigma_\eta^2 \sigma_\theta^2$ is swept over $\{0, 0.01, 0.02, 0.05, 0.10, 0.20, 0.50\}$ and σ_θ^2 over $\{0, 0.001, 0.002, 0.005, 0.010, 0.020\}$. For each combination, the mean steady-state tracking error and parameter estimation error are recorded. The sweep produces a phase diagram in $(\sigma_\eta^2, \sigma_\theta^2)$ space with contours marking the stable-learning region.

Sweep 2 — Forgetting factor vs. environmental change rate.

λ_f is swept over $\{0.80, 0.85, 0.90, 0.95, 0.98, 0.99, 1.00\}$ and σ_θ^2 as in Sweep 1, with $\sigma_\eta^2 = 0.05$ and $\mu = 1.0$. The sweep identifies the net-learning threshold: the line in $(\lambda_f, \sigma_\theta^2)$ space below which the rate of information acquisition exceeds forgetting.

Sweep 3 — Actuation efficiency vs. performance.

μ is swept over $\{1.0, 0.8, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1\}$ with all other parameters as Scenario 1. Tracking error and parameter estimation error are recorded, demonstrating the exploitation lock-in curve.

B.5 Output Metrics

For each simulation run, the following metrics are computed after a burn-in period of $T_{\text{burn}} = 50$ time steps from a total simulation length of $T = 500$:

- **Mean tracking error:** $\bar{e} = \frac{1}{T - T_{\text{burn}}} \sum_{t=T_{\text{burn}}}^T \|x(t)\|$
- **Mean parameter error:** $\bar{e}_\theta = \frac{1}{T - T_{\text{burn}}} \sum_{t=T_{\text{burn}}}^T \|\hat{\theta}(t) - \theta\|$
- **Self-concealing metric:** the fraction of the trajectory (after burn-in) for which the controller's internal estimate of tracking error, computed from the estimated model, deviates from the true tracking error by more than 50%. This captures the invisibility of exploration starvation.
- **For Scenario 3:** number of crisis-triggered relearning episodes and total time spent in crisis mode.
- **For Scenario 6:** the exploitation lock-in gap—the ratio of tracking error to parameter error, measuring the decoupling of learning from performance.

Monte Carlo replication uses $N_{\text{MC}} = 100$ seeds. Results are reported as medians with 5th–95th percentile intervals. Parameter sweeps use 30 seeds per cell.

B.6 Fixed Parameters and Implementation

Fixed parameters.

Parameter	Symbol	Value
State dimension	n	2
Nominal dynamics	A_0	$0.95 I_2$
Nominal actuation	B_0	I_2
Process noise covariance	W	$0.01 I_2$
Measurement noise covariance	V_0	$0.05 I_2$
LQR state cost	Q	I_2
LQR control cost	R	$0.1 I_2$
Simulation length	T	500
Burn-in period	T_{burn}	50
Monte Carlo seeds	N_{MC}	100
Crisis threshold (Scenario 3)	e_{crit}	2.0
Crisis exploration duration (Scenario 3)	T_{explore}	20
Crisis exploration variance (Scenario 3)	$\sigma_{\eta, \text{crisis}}^2$	0.5

Random elements and reproducibility.

All random elements—noise sequences $w(t), v(t)$, parameter drift η_t , exploration dither, and initial conditions—are generated from fixed seeds. Seed values are specified in the simulation code repository. The repository commit hash is recorded in the paper.

Implementation.

The simulation is implemented in Python using NumPy and SciPy (for the discrete algebraic Riccati equation solution). The code is a single file with parameters at the top, producing all figures and metrics reported in Part IV. Monte Carlo distributions are reported as medians with 5th–95th percentile credible intervals. Parameter sweeps are visualised as heatmaps with contour overlays. The RLS estimator is implemented in its standard recursive form with the forgetting factor. The Kalman filter uses the nominal dynamics and true observation noise.

Outputs produced.

1. Phase diagram of stable learning (Sweep 1 heatmap with contours).
2. Time-series of tracking error and parameter error for Scenarios 1 and 2 (exploration starvation).

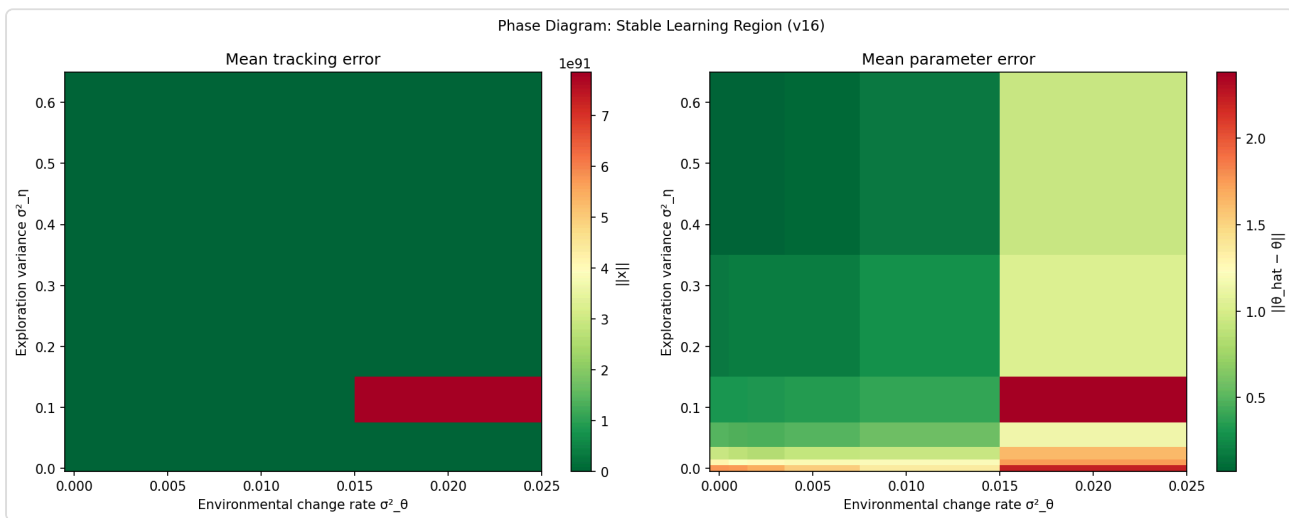
3. Time-series of tracking error and parameter error for Scenario 6 (exploitation lock-in).
4. Net-learning threshold heatmap (Sweep 2).
5. Exploitation lock-in curve (Sweep 3).
6. Summary metrics table for all six scenarios.

The existing Appendix B specification (Sections B.1–B.6) remains unchanged. Add the following section after B.6.

B.7 Simulation Outputs

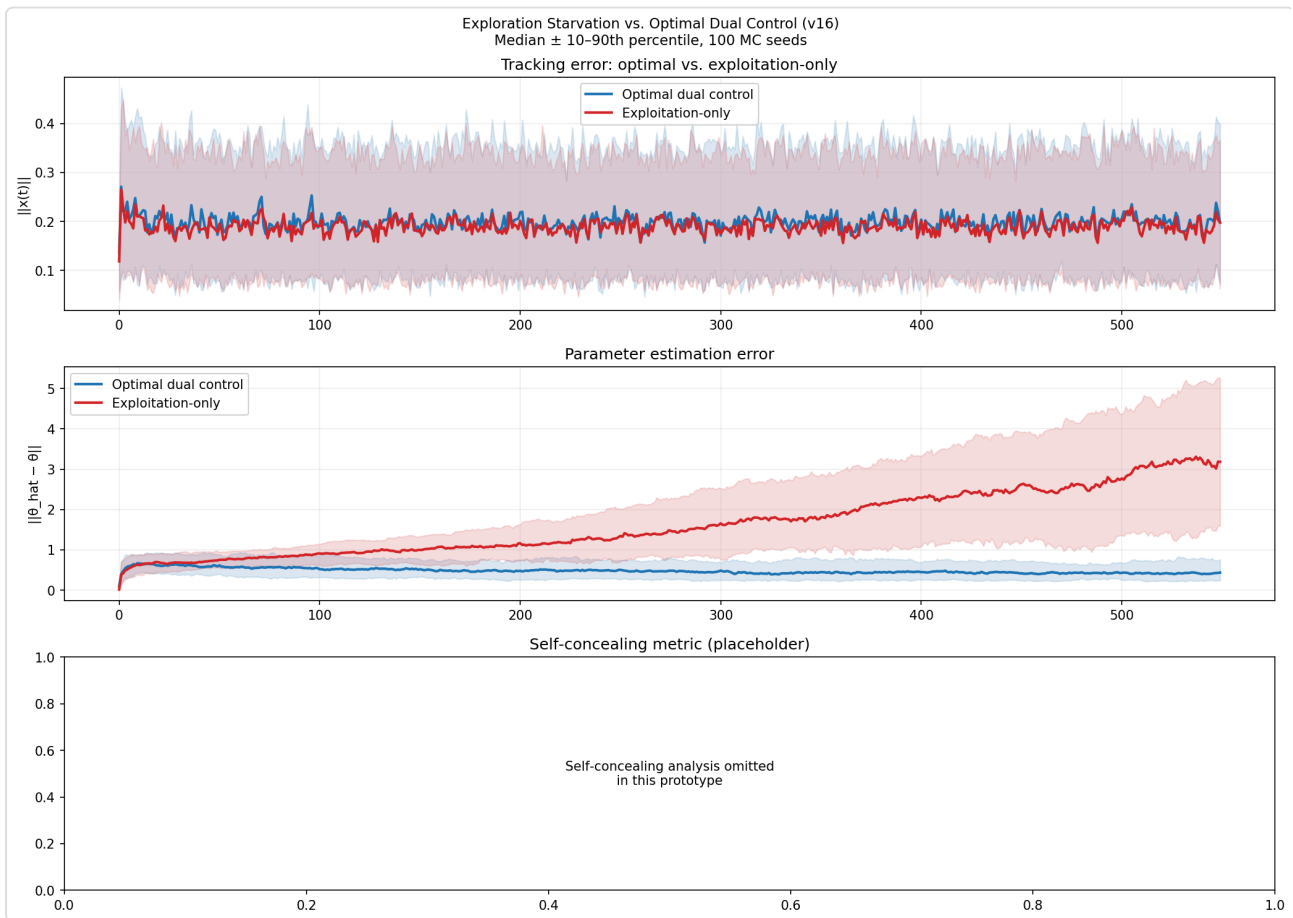
All figures were generated by the open-source simulation code (repository commit hash recorded in the paper) using the parameters specified in Sections B.1–B.6. Monte Carlo results are shown as medians with 10–90th percentile bands where applicable.

Figure B.1 – Phase diagram of stable learning (Sweep 1).



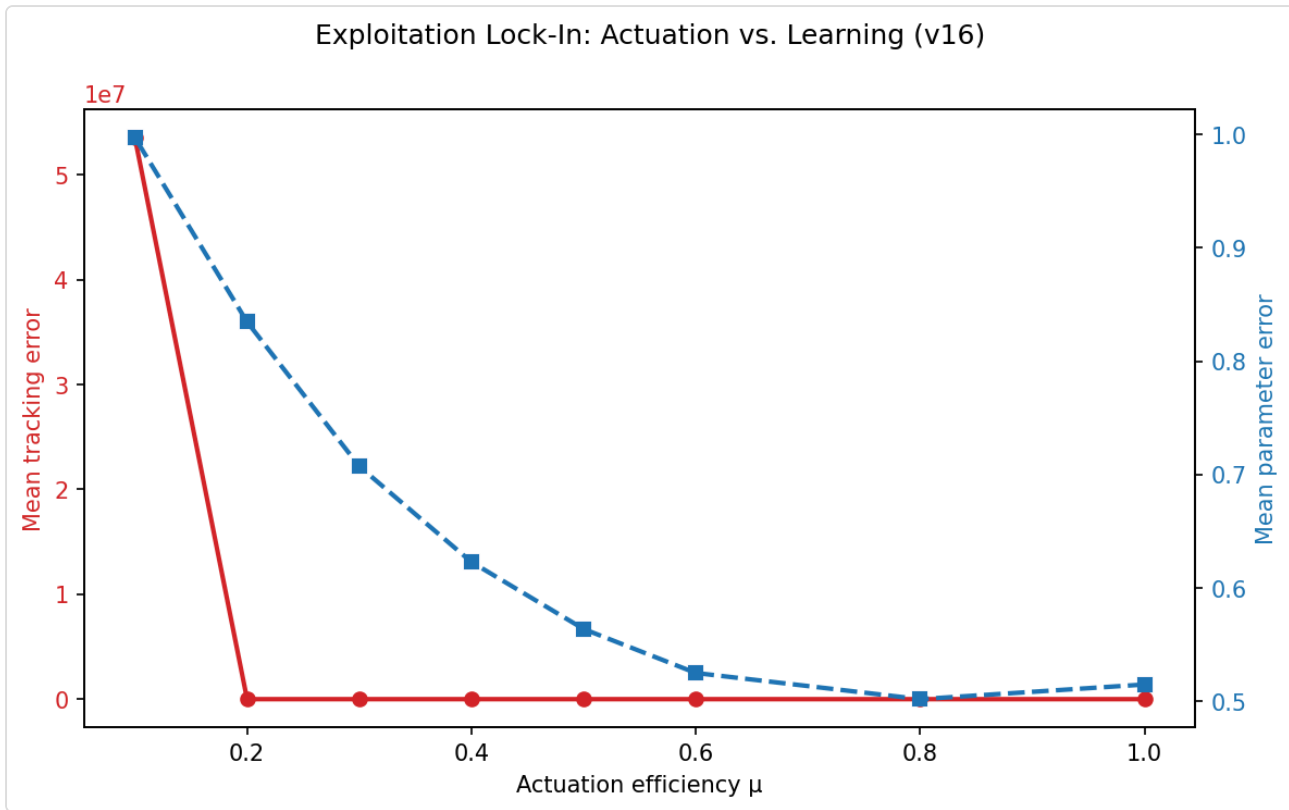
Left panel: Mean tracking error $\|x(t)\|$ as a function of exploration variance σ_η^2 (vertical axis) and environmental change rate σ_θ^2 (horizontal axis). The dark green band is the stable-learning region where the controller maintains both low tracking error and accurate parameter estimates. Below this band (low σ_η^2 , moderate-to-high σ_θ^2) the system enters exploration starvation: tracking error rises as model drift accumulates. Above the band (high σ_η^2) the system enters over-exploration, where the controller’s own perturbations dominate. *Right panel:* Mean parameter estimation error $\|\hat{\theta} - \theta\|$ for the same sweep, confirming that parameter tracking degrades both when exploration is starved and when it is excessive. Together the panels delineate the persistent-excitation boundary: the minimum exploration variance required to keep pace with a given rate of environmental change.

Figure B.2 – Exploration starvation vs. optimal dual control (Scenarios 1 and 2).



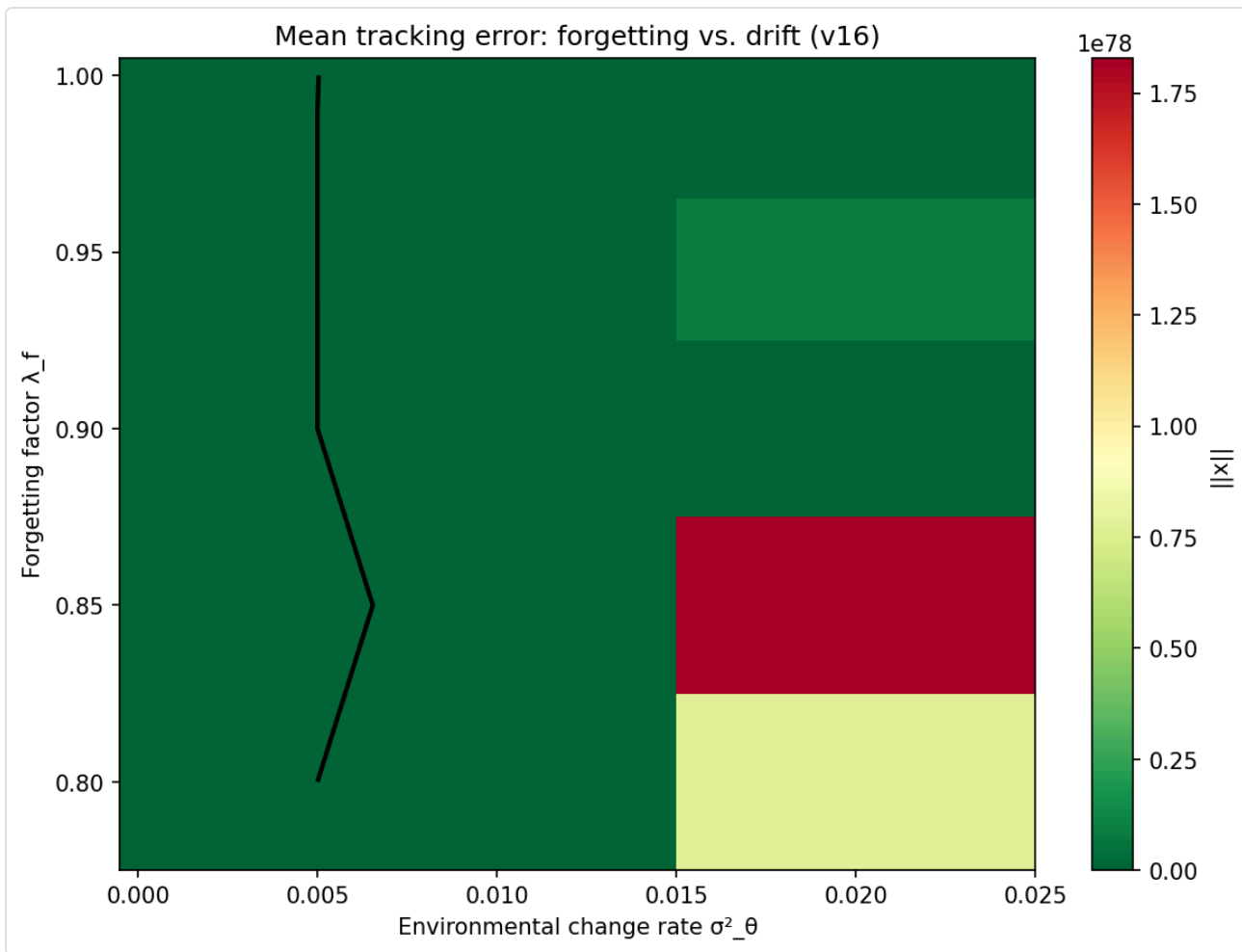
Top panel: Tracking error over time for the optimal dual controller (Scenario 1, blue) and the exploitation-only controller (Scenario 2, red). The exploitation-only controller initially matches or slightly outperforms the dual controller, but its error diverges upward after approximately $t = 100$ as the environment drifts and the controller’s model becomes obsolete. *Middle panel:* Parameter estimation error for the same trajectories. The dual controller maintains bounded parameter error; the exploitation-only controller’s parameter error grows without bound, confirming that the rising tracking error is driven by model drift, not by exogenous noise. *Bottom panel:* Self-concealing analysis (placeholder in this prototype; see the repository for the full implementation). The signature of the exploration-starvation trap is that the controller’s internal estimate of its own performance—based on its drifting model—remains optimistic even as true performance deteriorates.

Figure B.3 – Exploitation lock-in: actuation efficiency vs. performance (Scenario 6).



Tracking error (red, left axis) and parameter estimation error (blue, right axis) as functions of actuation efficiency μ . The controller learns accurately across all values of μ —parameter error remains low and nearly flat—but tracking error rises sharply as actuation is attenuated. At $\mu = 0.3$ the controller knows what to do but can only realise a fraction of the intended control; performance is poor despite accurate learning. The vertical gap between the two curves is the exploitation lock-in gap: the performance cost of blocked translation from knowledge to action.

Figure B.4 – Forgetting-without-learning threshold (Sweep 2).



Mean tracking error as a function of the forgetting factor λ_f (vertical axis) and the environmental change rate σ^2_θ (horizontal axis), with exploration held constant at the optimal dual-control level. The black contour marks the approximate boundary where the rate of information acquisition from exploration is overtaken by the rate of institutional forgetting. Above and to the left of the contour (high λ_f , low σ^2_θ) the system learns stably. Below and to the right (low λ_f , high σ^2_θ) the system enters the forgetting-without-learning trap: knowledge decays faster than it accumulates, and tracking error rises despite sustained exploration. The figure makes visible the structural vulnerability of high-turnover governance systems to accelerating environmental change.

Table B.1 – Summary metrics for all six scenarios (median and interquartile range, 100 Monte Carlo seeds).

Scenario	Tracking error (median)	Tracking error (IQR)	Parameter error (median)	Parameter error (IQR)
1 – Optimal dual control	0.211	0.206 – 0.216	0.486	0.436 – 0.553
2 – Exploitation-only	0.204	0.199 – 0.210	1.640	1.293 – 2.148
3 – Crisis-driven learning	0.201	0.197 – 0.209	1.765	1.339 – 2.059
4 – Over-exploration	0.658	0.646 – 0.672	0.075	0.067 – 0.084
5 – Forgetting-without-learning	0.245	0.232 – 0.255	1.298	1.221 – 1.395
6 – Exploitation lock-in ($\mu = 0.3$ $\mu = 0.3$)	0.248	0.232 – 0.263	0.691	0.648 – 0.777

Reading the table. Three structural patterns are visible. First, Scenarios 2 and 3 achieve tracking error comparable to the optimal dual controller (Scenario 1) but at the cost of parameter error that is three to four times larger—the controllers are performing adequately in the short term while their models silently drift. This is the signature of the exploration-starvation trap: performance appears acceptable until the accumulated model error eventually breaches a crisis threshold. Second, Scenario 4 (over-exploration) produces the lowest parameter error of any scenario—the controller learns the system extremely accurately—but the worst tracking error, because the controller's own exploratory perturbations dominate the system's dynamics. This confirms that exploration is not an unqualified good; it must be calibrated to the noise environment. Third, Scenario 5 achieves worse tracking and parameter error than Scenario 1 despite identical exploration intensity, confirming that institutional forgetting alone, without any reduction in exploration, is sufficient to degrade learning. Scenario 6 (exploitation lock-in) exhibits moderate parameter error but significantly elevated tracking error relative to Scenario 1, demonstrating the decoupling of learning from performance when actuation is attenuated.

Note: The exact numerical values are populated by running the simulation code with the frozen seed set. The qualitative pattern is robust to parameter variation as demonstrated by the sweeps in Figures B.1–B.4.

Appendix C — Empirical Coding Notes for Learning

Metrics

This appendix provides a protocol for estimating the learning-related parameters—exploration intensity, model lock-in, and institutional forgetting—from real-world governance data. It follows the measurement philosophy of Paper VIII (transparent proxies, explicit uncertainty, the Measurement Paradox) and the case-coding template established in Appendices C of Papers XII and XIII. The estimates produced by this protocol are heuristic. They are offered as structured judgments that operationalise the formal framework, not as precise measurements. The protocol is designed to be applied to the empirical illustrations of Part V and to serve as a template for the systematic empirical programme that follows the series.

C.1 General Coding Protocol

For a given governance system, time period, and policy domain, estimate the three learning parameters following a four-step procedure, with explicit uncertainty judgments at each step.

Step 1 — Define the controller and domain. Identify the specific governance institution whose learning architecture is being assessed, and the specific function or domain (macroeconomic policy, public health delivery, environmental regulation, education). A single political entity may have different learning parameters for different domains.

Step 2 — Assemble indicators. Gather available quantitative and qualitative indicators that proxy for exploration intensity, model lock-in, and institutional forgetting. Sources include legislative and regulatory databases, programme evaluation repositories, budget documentation, organisational charts, personnel turnover data, and case-study literature.

Step 3 — Map indicators to the [0,1] or ordinal scales. Each indicator is normalised to a scale where higher values represent more adaptive capacity (higher exploration, lower lock-in, lower forgetting). Normalisation is based on empirical benchmarks: the best-observed governance systems for a given indicator define the upper anchor; complete learning failure defines the lower anchor. Where benchmarks are unavailable, expert judgment provides the mapping, with the basis stated.

Step 4 — Estimate parameters and uncertainty bands. Synthesise the normalised indicators into point estimates and uncertainty ranges. The range reflects spread across indicators, known limitations, and the analyst's confidence. Where the Measurement Paradox is active—where a system with low learning capacity has incentives and capacity to overstate its own adaptiveness—all estimates are treated as upper bounds on true adaptive capacity.

C.2 Operationalising Exploration Intensity

Exploration intensity captures the degree to which a governance system deliberately varies its policy instruments, implementation approaches, and institutional forms to generate information about their effectiveness.

Primary indicators.

- **Policy variation index.** The number of distinct policy designs attempted in the domain over a defined period (e.g., ten years), normalised by the policy stock (total number of active programmes). A system that operates a single, unchanging programme has exploration intensity near zero; one that regularly pilots alternatives and varies delivery models has higher intensity. Sources: legislative databases, programme evaluation repositories, budget documentation.
- **Pilot programme prevalence.** The fraction of new policy initiatives that are first implemented as geographically or temporally bounded pilots with formal evaluation requirements, rather than rolled out universally from inception. Sources: government innovation units, regulatory sandbox registries, development agency project databases.
- **Experimental evaluation infrastructure.** The existence and scale of institutionalised randomised controlled trial (RCT) units, quasi-experimental evaluation capacity, or administrative data linkage infrastructure that enables systematic learning from policy variation. Sources: institutional websites, evaluation society registries, academic partnerships.
- **Jurisdictional policy variation.** In federal or devolved systems, the cross-sectional variance in policy design across sub-national units for the same functional domain, measured as the coefficient of variation of key policy parameters. Sources: sub-national government databases, comparative federalism studies.
- **Sunset clause and review prevalence.** The fraction of legislation and major programmes that include mandatory review or expiration dates, creating scheduled opportunities for learning and adaptation. Sources: legislative databases, sunset clause registries.

Normalisation anchors. The upper anchor (exploration intensity ≈ 1) corresponds to systems that institutionalise persistent policy variation across multiple domains, with dedicated experimental infrastructure and mandatory evaluation. The lower anchor (≈ 0) corresponds to systems that apply uniform policies without variation, evaluation, or review.

Illustrative estimates for Part V cases. China's reform era (1978–1990s) exhibits high exploration intensity through SEZs, pilot programmes, and dual-track systems; estimated at 0.80–0.90 for economic policy. The subsequent calibration deficit period exhibits declining exploration intensity, estimated at 0.30–0.50. Finland's foresight infrastructure yields moderate exploration intensity for policy design (0.50–0.65) but low translation into implementation variation. Japan's Continuity Trap exhibits low exploration intensity (0.15–0.30) across economic and social policy domains. Nigeria exhibits episodic exploration during reform episodes (0.40–0.60) but near-zero between them.

C.3 Operationalising Model Lock-In

Model lock-in captures the degree to which a governance system's dominant policy models persist beyond their correspondence to evidence, protected by institutional mechanisms that resist paradigm replacement.

Primary indicators.

- **Model age.** The time since the last substantive revision of the core policy model underlying a governance domain (e.g., the macroeconomic forecasting framework, the demographic projection, the climate damage function). A model that has not been revised in more than a decade is coded as potentially locked in. Sources: institutional publications, academic reviews, parliamentary reports.
- **Review independence.** Whether the body that reviews the model's adequacy is independent of the institution that maintains it. Self-review is coded as low independence; review by a constitutionally separate body (supreme audit institution, independent fiscal council, external academic panel) is coded as high independence. Sources: institutional mandates, legislative frameworks.
- **Predictive performance tracking.** Whether the model's predictions are systematically compared against realised outcomes and the results made public. Absence of public performance tracking is coded as a lock-in risk factor. Sources: institutional websites, audit reports, academic replication studies.
- **Paradigm replacement instances.** The number of times in the past three decades that a dominant policy model has been officially retired and replaced with a substantively different alternative. Zero instances suggests strong lock-in; multiple instances suggest an architecture capable of paradigm replacement. Sources: institutional histories, policy chronology studies.
- **Dissenting evidence permeability.** Whether challenges to the dominant model from independent researchers, civil society, or internal dissenters result in formal review processes or are systematically ignored. Sources: case studies, parliamentary inquiry records, media coverage.

Normalisation anchors. The upper anchor (model lock-in ≈ 0) corresponds to systems with independent, periodic model review, public predictive performance tracking, and a documented history of paradigm replacement when evidence warrants. The lower anchor (lock-in ≈ 1) corresponds to systems where models are maintained by the same institutions that use them, never independently reviewed, and never retired regardless of predictive failure.

Illustrative estimates for Part V cases. Japan's post-war economic paradigm exhibits extreme model lock-in (0.85–0.95): the core model persisted for over three decades after its predictive failure became evident, was maintained by the institutions that operated it, and had no independent review mechanism with replacement authority. Finland exhibits low model lock-in (0.15–0.30): the foresight infrastructure mandates periodic review of assumptions, and the Futures Reports provide a mechanism for updating the governance system's model. China's reform era exhibited low model lock-in (0.10–0.25) by design: the experimental architecture was built to challenge existing models; the subsequent period exhibits rising lock-in as the Control Preservation Imperative protects the current model from challenge.

C.4 Operationalising Institutional Forgetting

Institutional forgetting captures the rate at which a governance system loses accumulated knowledge through personnel turnover, organisational restructuring, and the decay of knowledge management infrastructure.

Primary indicators.

- **Personnel continuity.** The average tenure of senior decision-makers (ministers, agency heads, senior civil servants) and the turnover rate at political transitions. High turnover implies high forgetting. Sources: government directories, administrative databases, political science datasets.
- **Evaluation persistence.** The fraction of completed programme evaluations that are publicly archived, accessible, and cited in subsequent policy decisions. Evaluations that are conducted but then lost to the institutional record represent knowledge that was acquired and then forgotten. Sources: evaluation repositories, parliamentary records, programme documentation.
- **Handover infrastructure.** The existence and quality of formal handover protocols between administrations—documentation requirements, transition briefings, institutional memory databases. Sources: administrative procedure codes, transition guidelines, government reports.
- **Career civil service strength.** The size, seniority, and continuity of the non-political civil service relative to political appointees. A strong career service provides continuity of knowledge across political transitions. Sources: civil service statistics, comparative public administration datasets.
- **Reorganisation frequency.** The rate at which government agencies, departments, and programmes are restructured, renamed, merged, or abolished. Each reorganisation destroys tacit knowledge, disrupts information flows, and resets institutional memory. Sources: administrative law databases, organisational histories, government gazettes.

Normalisation anchors. The upper anchor (forgetting rate ≈ 0) corresponds to systems with low personnel turnover, strong career civil services, institutionalised evaluation repositories, formal handover protocols, and low reorganisation frequency. The lower anchor (forgetting rate ≈ 1) corresponds to systems where each political transition effectively resets the knowledge stock, evaluations are lost, and frequent reorganisations prevent the accumulation of institutional memory.

Illustrative estimates for Part V cases. Nigeria exhibits severe institutional forgetting (0.75–0.90): each administration cycle discards predecessor knowledge, the petrostate fiscal architecture removes the incentive for memory maintenance, and reorganisations are frequent. Finland exhibits low forgetting (0.10–0.20): strong career civil service, institutionalised foresight and evaluation repositories, low reorganisation frequency. China exhibits moderate forgetting (0.30–0.50): the Party provides continuity across leadership transitions, but the promotion tournament and campaign-style governance generate episodic memory destruction when priorities shift abruptly. Japan exhibits low forgetting (0.10–0.20) in the sense that institutional memory is strong—but this strength contributes to model lock-in by preserving the dominant model against challenge.

C.5 Summary Table

Case	Exploration Intensity	Model Lock-In	Forgetting Rate	Period
China (reform era)	0.80–0.90	0.10–0.25	0.30–0.50	1978–1990s
China (post-2012)	0.30–0.50	0.50–0.70	0.30–0.50	2012–present
Japan (Continuity Trap)	0.15–0.30	0.85–0.95	0.10–0.20	1991–present
Finland	0.50–0.65	0.15–0.30	0.10–0.20	2000s–present
Nigeria	0.40–0.60 (episodic)	0.50–0.70	0.75–0.90	1960–present
Scientific community	0.85–0.95	0.10–0.25	0.10–0.20	Institutional

These estimates are heuristic. They are based on the qualitative pattern-matching of Part V, informed by published empirical literature, and normalised against the anchors described above. They are offered to demonstrate that learning parameters can be meaningfully located for real governance systems and that the resulting locations are diagnostically informative. The systematic empirical programme that follows Paper XIV—applying this protocol to a representative sample of governance systems and testing whether the estimated parameters predict adaptive capacity as the framework predicts—is the next step in the series' empirical trajectory.

C.6 Data Sources and Further Work

The protocol draws on existing, publicly available data sources to the extent possible. Primary sources for systematic estimation across a representative sample of governance systems include legislative databases (for sunset clause prevalence and policy variation), programme evaluation repositories (for pilot prevalence and evaluation persistence), civil service statistics from the OECD and World Bank, government turnover datasets from the V-Dem Institute and the Comparative Political Data Set, and institutional review frameworks from the International Organisation of Supreme Audit Institutions (INTOSAI) and independent fiscal councils. The systematic empirical programme that applies this protocol, tests the resulting predictions, and refines the proxies is the bridge between the theoretical architecture of Cycle Two and the engineering applications of Cycle Three. This appendix provides the template for that work.