

Institutional Failure Modes: A Configurational Theory of Institutional Reproduction and Fragility

Nicolò Bellanca

Associate Professor

Department of Economics and Management

University of Florence

Via delle Pandette 9

50127 Florence, Italy

Email: nicolo.bellanca@unifi.it

ORCID: <https://orcid.org/0000-0002-3809-3455>

Working Paper

Version: March 2026

Abstract

Why do some institutional changes become durable while others remain fragile and reversible? Existing theories explain institutional persistence primarily through single mechanisms — path dependence, legitimacy, or incentive alignment. This article argues that such accounts misidentify the unit of analysis: durability is not a property of mechanisms individually but of their configuration. An institutional arrangement becomes self-reproducing when three mechanisms sustain one another in a relation of constitutive interdependence — incentive alignment, normative stabilization, and organizational embedding.

The article's central contribution is not the identification of these mechanisms, which are each well established in the literature, but the demonstration that configurations lacking any one of them generate a specific and predictable failure mode: technocratic stability (vulnerable to legitimacy crises), normative persistence (vulnerable to resource erosion), or fragile legitimacy

(vulnerable to organizational drift). Identifying these failure modes and their structural logic is the analytical core of the paper.

The framework is grounded in the sociological tradition of reproduction theory — Bourdieu, Giddens, and Sewell — and positions the concept of constitutive interdependence explicitly in relation to Luhmann's autopoietic account, from which it departs in treating self-reproduction as a continuous and configurational variable and in retaining a constitutive role for actors, agency, and conflict. The argument is developed through a conceptual framework, a dedicated section on the politics of institutional reproduction, three analytical sketches, and a stylized formal model confined to an appendix.

Keywords: institutional reproduction; institutional durability; institutional fragility; path dependence; legitimacy; organizational embedding

The manuscript is currently under review at a peer-reviewed journal – comments welcome.

Author Note

No external funding was received for this research.

The author declares no competing interests.

In preparing this manuscript, the author used a large language model as an editorial aid for structural revision, language editing, and consistency of argumentation. All theoretical arguments, analytical decisions, and interpretations remain the responsibility of the author.

1. Introduction

Theories of institutional stability confront an uncomfortable asymmetry. They are adept at explaining why institutions persist in general, but they struggle to explain why some institutional arrangements remain stable even under severe external pressure while others, superficially similar, collapse rapidly or degrade quietly from within. Path dependence, legitimacy, and incentive alignment each illuminate part of the phenomenon. Yet applied individually, each also generates a characteristic blind spot: it accounts for institutional durability under favorable conditions while failing to predict the specific ways in which stability breaks down when conditions change. This article argues that the source of this asymmetry is a shared methodological commitment to single-mechanism explanation. Institutions do not reproduce themselves through any one mechanism; they do so through the structured interaction of several mechanisms whose relation of constitutive interdependence is the generative condition of genuine durability.

Social theory has long understood reproduction as a relational, not a mechanical, phenomenon. For Bourdieu (1977), social fields reproduce themselves through the reciprocal constitution of *habitus* and objective social structures — a dynamic in which neither element is reducible to the other. For Giddens (1984), the reproduction of institutions is the recursive effect of structuration: agents draw on rules and resources that, in the very act of being drawn upon, are reconstituted. For Sewell (1992), structures are composed of mutually dependent schemas and resources, such that the reproduction of either requires the reproduction of both. What unites these accounts is the insistence that reproduction is a *configurational* achievement: it requires the alignment of heterogeneous elements whose interaction is non-additive and non-substitutable. The present article transposes this insight into the analysis of specific institutional arrangements.

Institutional analysis has developed three major accounts of persistence: path-dependent increasing returns (Pierson 2000; Mahoney 2000), normative and cultural stabilization (Meyer and Rowan 1977; DiMaggio and Powell 1983), and incentive-based political equilibria (North 1990; Acemoglu and Robinson 2006). Each account identifies real mechanisms. Yet each, used in isolation, misrepresents institutional durability as a property of a single mechanism rather than as the emergent property of their interaction. The consequence is a systematic failure to predict institutional fragility: arrangements that look stable by one account — well-incentivized, culturally legitimate, or organizationally entrenched — collapse in ways that the single-mechanism account cannot anticipate.

The concept of self-reproduction adopted here requires a word of clarification, because it connects to a major tradition in social theory that the present framework both engages and departs from. Niklas Luhmann's theory of autopoietic social systems (1984, 1995) defines self-reproduction as the operational closure of a system: a social system reproduces itself by producing the elements of which it consists through those same elements, with no constitutive inputs from the environment. On this account, self-reproduction is an all-or-nothing property of system-types, not a variable property of particular institutional configurations. The present framework departs from Luhmann on two connected points. First, it treats institutional self-reproduction as a *continuous* and *configurational* variable — a property that admits of degrees and depends on the joint strength of multiple mechanisms — rather than as a systemic threshold property. Second, it retains a central role for actors, agency, and conflict in both the generation and the disruption of self-reproducing dynamics, whereas Luhmann's framework brackets agency as a product of systemic communication rather than its source. The present account is closer to Bourdieu and Sewell than to Luhmann in its assumptions about the relationship between structure and action.

The central argument of the article is that institutional arrangements acquire genuine self-reproducing properties when three mechanisms operate jointly: *incentive alignment*, *normative stabilization*, and *organizational embedding*. The distinctive contribution is not the identification of these mechanisms, which are each well established, but the *constitutive interdependence* that connects them. Each mechanism, absent the others, generates a specific and predictable failure mode. Identifying these failure modes — their characteristic dynamics and their conditions of occurrence — is the analytical core of the article.

The argument proceeds as follows. Section 2 reviews existing single-mechanism accounts and identifies the shared limitation that motivates the present framework. Section 3 develops the concept of institutional self-reproduction, positioning it in relation to Luhmann and to the sociological tradition of reproduction theory. Section 4 presents the triadic architecture: the three mechanisms, their constitutive interdependence, and the failure modes generated by each partial configuration. Section 5 analyzes the role of agency and conflict in the production and disruption of self-reproducing institutions. Section 6 offers three analytical sketches illustrating the framework's explanatory reach. Section 7 discusses implications for institutional theory. Section 8 concludes. Appendix A presents a stylized formal model.

2. The Limits of Single-Mechanism Accounts

Three traditions in institutional analysis have generated the most influential explanations of why institutions persist. The purpose of this section is not to offer yet another literature review but to identify the specific structural limitation that each tradition shares — a limitation that motivates the configurational approach developed in the remainder of the article.

2.1 Path dependence and the limits of increasing returns

Historical institutionalism explains persistence through increasing returns: early institutional choices generate coordination effects, adaptive expectations, and sunk costs that make deviation progressively more costly (Pierson 2000; Mahoney 2000). The mechanism is self-reinforcing: the more an institutional path is followed, the more following it becomes advantageous. The macro-level result is lock-in.

This account is powerful for explaining trajectories that are already stable. Its limitation is that it cannot predict *which* institutions will develop increasing returns and which will not — or, equivalently, why some apparently locked-in institutional arrangements unravel rapidly when external conditions shift. Lock-in due to coordination constraints or sunk costs is not the same as lock-in due to endogenous self-reproduction. The former depends on the continuation of the conditions that generate switching costs; the latter generates its own conditions of continuation. Path dependence can be the macro-level *result* of institutional self-reproduction, but it is not identical to it, and treating them as equivalent leads to systematically overpredicting institutional durability.

2.2 Normative legitimacy and the limits of cultural stabilization

Sociological institutionalism argues that institutions persist because they become embedded in systems of shared meanings and taken-for-granted practices (Meyer and Rowan 1977). Once an institution acquires *social facticity* — it appears natural, inevitable, beyond question — it can sustain itself even in the absence of continuous material incentives (DiMaggio and Powell 1983). Legitimacy provides compliance without enforcement and confers a kind of resilience that purely incentive-based accounts cannot explain.

The limitation of this account is symmetrical. Cultural legitimacy is a real stabilizing force, but it is insufficient to sustain institutional practices in the absence of material resources and organizational carriers. The sociology of organizations documents numerous cases of institutions that retain symbolic legitimacy while their operational infrastructure decays — a pattern Meyer and Rowan (1977) themselves anticipated in their analysis of the decoupling between institutional myth and organizational practice. Legitimate forms can survive the

hollowing-out of the substance they are supposed to embody; the resulting arrangements are symbolically durable but substantively fragile in characteristic ways.

2.3 Incentive alignment and the limits of distributional equilibria

Political economy approaches explain persistence through the incentives of relevant actors (North 1990; Acemoglu and Robinson 2006). Institutions endure because they generate benefits for actors who have the capacity to defend them. Research on gradual institutional change has refined this picture: Mahoney and Thelen (2010) show that actors exploit ambiguities in institutional rules to introduce incremental changes whose cumulative effect may be transformative. Thelen (2004) demonstrates that institutions can be adapted to new purposes by new coalitions, so that formal persistence coexists with substantial functional change.

This tradition is essential for understanding the politics of institutional maintenance and change. Its limitation, for present purposes, is that it conceptualizes stability as an equilibrium among actors rather than as a property of institutional configurations. Coalitional equilibria can be disrupted by shifts in the distribution of power, by changes in economic conditions, or by the mobilization of new actors. An institution sustained primarily by the political power of a winning coalition is stable only as long as that coalition remains intact. Its fragility, when the coalition fragments, is not predicted by the account that explained its stability.

2.4 The shared limitation: misidentifying the unit of analysis

The three traditions just surveyed are not wrong about the mechanisms they identify. They are, individually, incomplete. Each privileges a single mechanism — increasing returns, legitimacy, or incentive alignment — and treats institutional durability as a function of that mechanism's strength. This is a misidentification of the unit of analysis. Durability is not a property of mechanisms taken singly but of their *configuration*: of whether and to what degree they sustain one another in a relation of constitutive interdependence. The consequence of the misidentification is not only theoretical incompleteness but also empirical misprediction — specifically, a systematic tendency to underestimate institutional fragility by ignoring the failure modes that partial configurations generate.

The following section develops a positive alternative by specifying the conditions under which institutional arrangements acquire genuine self-reproducing properties and, equally, the conditions under which they do not.

3. Institutional Self-Reproduction: Beyond Persistence and Autopoiesis

An institutional arrangement is *self-reproducing* when it generates endogenously the mechanisms required for its own persistence. Three clarifications of this definition are necessary.

First, the definition specifies an *endogenous* relationship between the institution and its conditions of continuation. Many institutions persist for exogenous reasons: because political coalitions happen to remain favorable, because no viable alternative has emerged, or because the costs of disruption exceed what any single actor is willing to bear. Such institutions are *persistent* but not *self-reproducing*. The distinction is empirically tractable: one can ask the counterfactual whether the institution would survive a substantial alteration of the external conditions that currently support it. A self-reproducing institution generates its own incentives, normative frameworks, and organizational carriers. A merely persistent institution borrows them from its environment.

Second, the definition treats self-reproduction as a *continuous* rather than a categorical property. No real institution achieves perfect endogenous self-reproduction; all are to some degree dependent on external conditions. The concept identifies a continuum along which institutional arrangements can be located, and the analytical interest lies in specifying the mechanisms through which arrangements move toward or away from the self-reproducing end of the continuum.

Third, the definition differs from Luhmann's autopoietic concept of self-reproduction in important respects. For Luhmann (1984, 1995), a social system reproduces itself autopoietically when its operations produce the elements of which it is composed, establishing the operational closure that defines the system boundary. Self-reproduction, on this account, is a systemic property that either holds or does not: a system is either autopoietically closed or it is not a system in the relevant sense. The present framework departs from this in two ways. First, it treats self-reproduction as a *variable* property — a function of the strength and configuration of specific mechanisms — rather than as a threshold property of system-types. Second, and more fundamentally, it retains the constitutive role of actors, coalitions, and conflict in the production and disruption of self-reproducing dynamics. Luhmann dissolves agency into communication: actors are not the authors of institutional reproduction but its effects. The present account, following Bourdieu (1977), Giddens (1984), and Sewell (1992), treats the recursive relationship between actors and structures as the generative process through which self-reproducing institutions are both built and, sometimes, unraveled.

The concept is also distinct from, though related to, path dependence. Path dependence, in the standard account (Pierson 2000), is a macro-level trajectory property: a historical sequence in which prior choices constrain subsequent ones through mechanisms of increasing returns. Institutional self-reproduction provides a micro-level account of why and how some trajectories exhibit this property. When an institutional arrangement is self-reproducing — when it generates the incentives, norms, and organizations that sustain it — the macro-level pattern of path dependence is a predictable emergent consequence. But institutional arrangements can exhibit path dependence without being genuinely self-reproducing, if the lock-in results primarily from switching costs or coordination constraints rather than from endogenous reproduction of the supporting conditions.

The sociological tradition of reproduction theory illuminates a further dimension. Bourdieu's concept of *habitus* describes how dispositions internalized through socialization reproduce the practices and perceptions that sustain social fields — without deliberate coordination and often without conscious awareness. Sewell's (1992) account of structural reproduction shows how schemas and resources are co-constitutive: actors enact schemas by mobilizing resources, and in doing so both reproduce the schema and confirm the existence of the resource. These accounts suggest that institutional self-reproduction operates through multiple channels simultaneously — cognitive and normative, material and organizational — and that its durability is a function of how tightly these channels are constitutively coupled. This is precisely the claim that the triadic architecture makes explicit.

4. The Triadic Architecture: Mechanisms, Constitutive Interdependence, and Failure Modes

The central theoretical argument of this article is that institutional self-reproduction is a *configurational* property. It emerges when three mechanisms sustain one another in a relation of constitutive interdependence, and it fails in three characteristic ways when any one of them is absent. This section presents the three mechanisms, the logic of their constitutive interdependence, and the failure modes that each partial configuration generates.

4.1 Incentive alignment

Incentive alignment refers to the process through which institutional arrangements generate material or strategic benefits for relevant actors, creating a stake in the institution's continuation. When institutional rules produce advantages for those who operate within them, actors develop incentives to reproduce the framework through their own behavior. Over time, investments accumulate around the institutional structure, adaptive

expectations form, and the prospect of institutional disruption comes to be experienced as a material threat (North 1990; Pierson 2000).

Incentive alignment is not, however, limited to direct material benefit. It includes the generation of *political* and *strategic* stakes: actors who have built careers, coalitions, or reputations within an institutional framework have incentives to defend it that go beyond the immediate material calculus. The distributional dimension is also essential: incentive alignment means that the institutional arrangement produces a specific distribution of costs and benefits that gives its beneficiaries more to lose from institutional disruption than they have to gain from it (Acemoglu and Robinson 2006).

Incentive alignment alone, however, is insufficient for self-reproduction. Distributional coalitions can fragment as economic conditions change, as new actors with different interests are mobilized, or as the institutional arrangement itself generates distributional dynamics that gradually erode the coalition that sustains it. Without normative legitimacy, an institution sustained by incentive alignment alone is vulnerable to contestation whenever its distributional effects become salient. Without organizational embedding, its continued operation depends on the continuous renewal of the political will that animates it — a condition that cannot be guaranteed.

4.2 Normative stabilization

Normative stabilization refers to the process through which institutional practices come to be supported by shared expectations regarding appropriate and legitimate behavior. Through socialization, professionalization, and cultural diffusion, institutional arrangements may come to appear not merely advantageous but *right*, *natural*, or *inevitable*. Once an institution achieves normative stabilization, compliance no longer requires continuous material incentives: violation carries cognitive dissonance and social sanction. Legitimacy functions as a form of institutional self-insurance (Meyer and Rowan 1977; DiMaggio and Powell 1983).

Normative stabilization connects directly to the sociological tradition of reproduction theory. For Bourdieu (1977), the *doxa* of a field — the set of presuppositions taken for granted by all participants — is both the product and the precondition of field reproduction. For Sewell (1992), schemas are reproduced when actors mobilize resources in ways that confirm their validity. These processes are not mechanical: they require active social work by the actors who internalize, enact, and transmit institutional norms across generations. Professional

associations, educational institutions, and the specialized vocabularies of expert communities are central vehicles of normative stabilization.

Normative stabilization alone is insufficient for the same reason as incentive alignment: it does not guarantee the material and organizational preconditions for institutional practice. An institution may retain symbolic legitimacy long after its operational substance has eroded — the decoupling phenomenon that Meyer and Rowan (1977) identified. More broadly, normative frameworks that are disconnected from material incentive structures and organizational carriers become increasingly formal and decreasingly action-guiding, until a sufficiently severe disruption reveals the gap between normative appearance and operational reality.

4.3 Organizational embedding

Organizational embedding refers to the process through which institutional rules and practices become incorporated into durable organizational structures capable of reproducing routines across time and across changes in the individual actors who enact them. Organizations are institutional carriers: they sustain rules through procedures and formal hierarchies, enforce compliance through monitoring and sanctioning, and transmit institutional knowledge through training, socialization, and accumulated records (Mahoney and Thelen 2010; Thelen 2004).

The centrality of organizations to institutional reproduction reflects a general feature of social structures that Sewell (1992) identified in his analysis of resource accumulation. Resources — material, organizational, cognitive — tend to concentrate in organizations that then become structural anchors for the practices they house. Administrative agencies, regulatory bodies, professional associations, educational institutions, and legal courts are paradigmatic organizational carriers. Their procedures continue to reproduce institutional practices even when the original rationale for those procedures has been forgotten, when the individuals who enacted them have departed, and when the normative frameworks that once justified them have been contested.

Organizational embedding alone is also insufficient. Organizations without incentive support decay through resource erosion and defection: staff leave for more attractive opportunities, funding dries up, and the organization loses the operational capacity to perform its functions. Organizations without normative legitimacy are vulnerable to challenges that they lack the cultural authority to deflect; when their formal authority is contested and their functional legitimacy is questioned, the procedural routines that embody the institution can be disrupted or redirected.

4.4 Constitutive interdependence and the logic of failure modes

The relation of *constitutive interdependence* among the three mechanisms is the analytical core of the framework. The concept draws directly on Sewell's (1992) account of the co-constitution of schemas and resources, extended here to the interaction among incentive structures, normative frameworks, and organizational carriers. Three mechanisms are constitutively interdependent when each reproduces the conditions that make the others operative — when none can sustain its own function without the simultaneous functioning of the other two. This is a stronger claim than mutual reinforcement or complementarity: it asserts that the mechanisms do not merely *enhance* one another's effectiveness but *co-produce* the conditions of each other's reproduction.

The constitutive character of these interdependencies can be specified concretely. Incentive alignment is not self-sustaining: the material resources it distributes require organizational infrastructure to be reliably delivered, and the legitimacy of that distribution requires a normative framework to render it socially acceptable. Remove normative stabilization, and incentive alignment becomes nakedly coercive — dependent on continuous enforcement whose costs rise as legitimacy declines. Remove organizational embedding, and incentive alignment becomes politically contingent — renewed only as long as the coalition that sustains it retains power. Normative stabilization, equally, is not self-sustaining: the norms that legitimate an institutional arrangement require organizational carriers to be transmitted and enacted across generations, and they require material incentive structures that give actors reasons to comply with them in practice. Organizational embedding, for its part, requires both the material resources that incentive alignment provides and the legitimacy that normative stabilization confers: organizations that are resource-starved decay, and organizations that are normatively illegitimate face challenges that their procedural authority alone cannot deflect.

This mutual constitution — in which each mechanism produces the conditions without which the others cannot function — is what generates genuinely self-reproducing institutional dynamics. As Giddens (1984) argued for structural reproduction more generally, what is reproduced is not simply the institution itself but the very conditions that make institutional action possible. When all three mechanisms operate jointly, the institution continuously regenerates the material stakes, the normative expectations, and the organizational forms through which it is enacted. Durability, in this account, is not a property that institutions *have* but a process they *perform* — one that requires all three mechanisms to be simultaneously active.

The logic of constitutive interdependence also implies that the absence of any one mechanism generates a *specific* failure mode, not merely a weaker version of full self-reproduction. This is the article's central empirical

claim, and it distinguishes the configurational approach from additive or threshold models of institutional stability. Consider the three partial configurations:

Technocratic stability (Incentives + Organizations, no Normative Stabilization). An institution with strong incentive alignment and organizational embedding but weak normative legitimacy functions effectively as long as its organizational capacity remains intact and its beneficiaries retain political power. It is vulnerable, however, to legitimacy crises: challenges to its formal authority that it lacks the cultural resources to deflect. Because no shared normative framework defends the institution against direct contestation, its stability depends entirely on the correlation of forces between its beneficiaries and its opponents. When political conditions shift — when the coalition that supports the institution loses power, or when public mobilization against technocratic governance intensifies — the institution may collapse rapidly, leaving no normative residue that would make reconstruction straightforward.

Normative persistence (Norms + Organizations, no Incentive Alignment). An institution with strong normative stabilization and organizational embedding but weak incentive alignment retains legitimacy and formal organizational structure while losing the material resources and coalitional support needed to operate effectively. The characteristic failure mode is gradual *resource erosion*: the institution continues to exist and to be regarded as legitimate, but its operational capacity hollows out. Staff are underpaid and unmotivated; resources are insufficient for effective functioning; the organizational forms that embody institutional rules persist as shells while their operational substance deteriorates. This is a pattern familiar from many nominally legitimate institutions in contexts of fiscal austerity or political abandonment.

Fragile legitimacy (Incentives + Norms, no Organizational Embedding). An institution with strong incentive alignment and normative stabilization but weak organizational embedding commands broad material and symbolic support but fails to reproduce its practices reliably across time and across actor turnover. The characteristic failure mode is *organizational drift*: practices diverge progressively from institutional rules, because no durable organizational carrier exists to implement and transmit them. Each generation of actors must reinvent or re-interpret what the institution requires, producing cumulative departures from the founding intent. The institution retains legitimacy and coalitional support but loses substantive coherence.

These four configurations — full triad, technocratic stability, normative persistence, and fragile legitimacy — and the single-mechanism case are summarized in Table 1 and Figure 1 below.

Table 1. Configurations, dynamics, and characteristic failure modes

Configuration	Mechanisms Present	Institutional Dynamic	Characteristic Vulnerability
Full triad	I + N + O	Self-reproducing	External shocks only; highly resistant to endogenous challenge
Technocratic stability	I + O (no N)	Functionally stable	Legitimacy crisis; collapse when political authority is contested
Normative persistence	N + O (no I)	Symbolically stable	Resource erosion and defection as material incentives weaken
Fragile legitimacy	I + N (no O)	Coalitionally stable	Organizational drift; practices diverge from institutional rules over time
Single mechanism	I, N, or O alone	Unstable equilibrium	Rapid breakdown under even moderate external pressure

Note: I = incentive alignment; N = normative stabilization; O = organizational embedding. The characteristic vulnerability identifies the specific failure mode generated when one mechanism is absent, not a general measure of institutional weakness.

Fig. 1 The Triadic Architecture: Mechanisms and Failure Modes

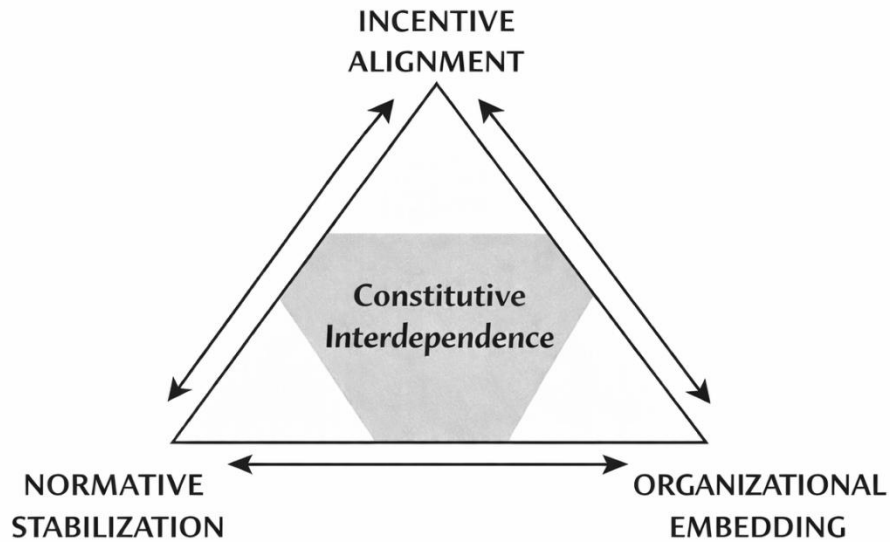


Fig 1 *The Triadic Architecture: Mechanisms and Failure Modes. Each side represents a relation of constitutive interdependence: the reproductive capacity of each mechanism depends on the simultaneous presence of the other two. Constitutive interdependence is the zone in which all three mechanisms co-produce the conditions of institutional self-reproduction. Each absent vertex generates a characteristic failure mode (see Table 1).*

5. Agency, Conflict, and the Politics of Institutional Reproduction

A configurational framework risks the appearance of structural determinism: institutions either have the right configuration or they do not, and actors are merely the bearers of mechanisms whose logic unfolds above their heads. This is not the intention, and the present section takes up explicitly the role of agency, conflict, and power in the production, maintenance, and disruption of self-reproducing institutional arrangements.

The triadic architecture is not given in advance. It is *built* through the deliberate and often conflictual action of coalitions with interests in institutional durability. Incentive alignment does not appear spontaneously: it is the product of distributional politics in which some actors succeed in anchoring their material interests in institutional rules while others fail. Normative stabilization is not simply a consequence of institutional longevity: it requires active legitimation work by professionals, educators, journalists, and public intellectuals who invest in making a particular institutional arrangement appear natural or necessary. Organizational embedding requires

resources, political will, and the defeat of competing organizational models. In all three dimensions, the consolidation of institutional arrangements involves winners and losers — actors whose interests are served by the emerging configuration and actors who resist it (Streeck and Thelen 2005).

The Bourdieuan tradition is illuminating here. Field reproduction, for Bourdieu (1977), is not a frictionless process: it is an ongoing struggle in which agents compete for the forms of capital — economic, cultural, social, symbolic — that define the stakes of the field. Dominant actors work to reproduce the field's rules of the game (*nomos*) because the existing configuration of capital advantageously reflects their own accumulated resources; dominated actors either comply — and thereby also reproduce the field — or challenge the rules in ways that risk transforming the field's structure. Institutional reproduction, in this framework, is always also institutional politics.

Conflict does not cease once a self-reproducing configuration is established. It is, rather, *transformed*. When all three mechanisms of the triadic architecture are functioning, challenges to the institution face a formidable combination of material, normative, and organizational resistance. Challengers must simultaneously undermine the material interests of the institution's beneficiaries, delegitimize its normative framework, and disrupt or capture its organizational carriers. This is why genuinely self-reproducing institutions — mature welfare states, established professional systems, entrenched regulatory regimes — tend to change, when they change at all, through the gradual and incremental mechanisms that Mahoney and Thelen (2010) identify: layering, drift, conversion, and displacement, each of which involves exploiting one of the partial vulnerabilities that even self-reproducing institutions exhibit rather than directly challenging the configuration as a whole.

Partial configurations, by contrast, transform conflict dynamics in characteristic ways. Technocratic stability concentrates conflict around questions of legitimacy and democratic accountability: the institution's organizational and material resources are intact, but its lack of normative embedding makes it vulnerable to political entrepreneurs who can frame it as unaccountable, elitist, or self-serving. This is the conflict dynamics that Streeck (2014) identifies in the crisis of European technocratic governance. Normative persistence, by contrast, displaces conflict toward questions of resources and effectiveness: legitimacy is taken for granted, but the material deficit generates pressures from within the institution — from underpaid professionals, understaffed agencies, under-resourced programs — as well as from users who experience deteriorating service quality. Fragile legitimacy concentrates conflict around interpretation and control of institutional practice: when no organizational

carrier reliably transmits institutional rules, the institution becomes a terrain of contestation over what, precisely, it requires.

Power, in this account, is not simply the resource that explains which institutional configurations get built. It is also differentially affected by which configuration exists. Self-reproducing institutions generate specific distributions of power among their beneficiaries, legitimators, and organizational managers, and these distributions shape the political terrain on which both institutional maintenance and institutional change take place.

6. Analytical Sketches

The following three illustrations are offered as *analytical sketches*: theoretically organized accounts of empirical cases whose purpose is to demonstrate the framework's capacity to organize and clarify familiar institutional phenomena. They do not constitute case studies in the methodological sense and make no claim to causal identification. Their function is to show that the configurational categories and failure-mode predictions of the triadic architecture are recognizable in well-documented institutional trajectories.

6.1 *The postwar welfare state: from precarious reform to self-reproducing institution*

The consolidation of welfare states in Western Europe after 1945 offers a paradigmatic instance of the transition from a fragile to a self-reproducing institutional arrangement. In the immediate postwar period, welfare institutions were politically precarious: dependent on favorable electoral conditions, on the organizational capacity of labor movements, and on a specific geopolitical configuration. The puzzle is not how welfare states were created — the politics of their origins is well understood — but how they acquired the durability that made subsequent retrenchment so politically costly (Pierson 1994).

The answer, in configurational terms, is that welfare institutions gradually built all three components of the triadic architecture into a relation of constitutive interdependence. Incentive alignment developed through the creation of large, politically organized constituencies of beneficiaries — pensioners, health care users, workers insured against unemployment — whose material interests became directly tied to the continuation of provision. Normative stabilization developed through the diffusion of social citizenship as a legitimate moral framework: social protection came to be understood as a right of membership in a political community, not as charitable provision (Marshall 1950). This normative framework made direct frontal attacks on welfare institutions

politically costly, even when ideological preferences shifted rightward. Organizational embedding was provided by the extensive administrative apparatus of welfare delivery — the social security agencies, public health systems, and professional corps of welfare workers who developed both material stakes and normative commitments to the institutions they staffed.

The transition to full self-reproduction was not linear and was itself contested. In the early years, many welfare institutions exhibited the fragile legitimacy configuration: broad normative support and growing incentive alignment, but organizational embedding that was incomplete or uneven. The characteristic failure mode — organizational drift — was visible in the divergence between formal entitlement rules and actual access to provision in the early postwar period. Full constitutive interdependence was achieved only once organizational capacity was built to the point where it could reliably reproduce the incentive structures and normative frameworks of social citizenship. This sequencing helps explain why Mahoney and Thelen's (2010) analysis of welfare state change focuses on incremental mechanisms — layering, conversion, drift — rather than direct dismantlement.

6.2 Central bank independence: technocratic stability and its crisis

The institutionalization of central bank independence in advanced economies from the 1980s onward offers an illustration of the *technocratic stability* configuration — and of the distinctive failure mode it implies. The initial institutionalization was primarily an exercise in incentive alignment and organizational embedding: financial markets rewarded governments that credibly committed to anti-inflationary monetary policy, and the organizational form of central bank independence — statutory autonomy, professional expertise, operational insulation — was rapidly diffused through international policy networks.

The normative dimension was, initially, relatively weak. Central bank independence was technically legitimate among economists and market participants, but it lacked deep popular legitimacy as a mode of democratic governance. This is the characteristic profile of *technocratic stability*: strong on incentive alignment and organizational embedding, but weak on normative stabilization in the broader public sphere. The constitutive interdependence between the three mechanisms was, accordingly, incomplete: the organizational and material pillars were in place, but the normative pillar had not yet been built to the point where it could reproduce the conditions of the institution's cultural authority.

The financial crisis of 2008 revealed the limits of this configuration. The crisis disrupted incentive alignment — the credibility benefits of inflation targeting came into question — and simultaneously triggered a

legitimacy challenge: the technocratic authority of unelected central bankers became politically contested. The framework predicts this as the characteristic failure mode of technocratic stability: when the correlation of political forces shifts and the normative vacuum becomes visible, the institution's organizational and material resources are insufficient to defend it against direct contestation of its legitimacy.

6.3 Post-communist administrative reform: fragile legitimacy and organizational drift

A contrasting illustration of *fragile legitimacy* is provided by administrative reform in Central and Eastern Europe during the 1990s. These reforms were typically enacted rapidly, with strong formal normative commitments: legislation, administrative restructuring, professionalization requirements. Yet many proved fragile and were partially reversed or quietly hollowed out within years of their introduction.

The triadic architecture framework identifies the source of this fragility. In most cases, normative stabilization and organizational embedding were attempted through legislative and administrative fiat, without the prior development of the incentive structures that would give relevant actors a stake in the institution's continuation. Political elites who had supported the reforms when they served immediate electoral or geopolitical purposes had limited intrinsic interests in maintaining them once those external incentives weakened. This is the profile of fragile legitimacy: strong normative commitment at the moment of enactment, some degree of organizational form, but weak incentive alignment — and therefore incomplete constitutive interdependence.

The characteristic failure mode — *organizational drift* — is precisely what followed. Civil servants trained under the previous system reproduced informal practices that undercut formal organizational rules. Each change of government brought reinterpretations of what the formal rules required. The institution retained its normative framing while its operational content drifted progressively from it. Without the incentive structures that would have stabilized compliance, the constitutive coupling between normative frameworks and organizational practices could not be maintained.

7. Implications for Institutional Theory

7.1 Path dependence as emergent property

The triadic architecture framework reframes the interpretation of path dependence. Standard accounts treat it as an independent causal mechanism — a property of historical sequences (Pierson 2000). The present

framework treats it as the *emergent macro-level manifestation* of micro-level self-reproducing processes. When an institutional arrangement develops constitutive interdependence among all three mechanisms, the macro-level trajectory becomes increasingly resistant to reversal — producing the pattern that historical institutionalism describes as path dependence. But the reverse is not true: institutional trajectories can exhibit apparent lock-in without genuine constitutive interdependence, if the lock-in results from exogenous constraints rather than endogenous reproduction. The distinction predicts that apparently locked-in but not genuinely self-reproducing institutions will prove more fragile than path-dependence accounts expect, once the exogenous constraints are relaxed.

7.2 Institutional isomorphism and the risk of normative decoupling

The framework also qualifies the sociology of institutional isomorphism. DiMaggio and Powell (1983) showed that organizational fields converge on legitimate institutional forms through coercive, normative, and mimetic pressures. The triadic architecture framework adds a critical qualification: normative convergence that is not accompanied by the development of corresponding incentive structures and organizational capacity produces the *normative persistence* configuration — arrangements that are symbolically isomorphic but operationally divergent. This is the configurational basis of the decoupling that Meyer and Rowan (1977) identified: when normative stabilization advances ahead of incentive alignment and organizational embedding, the gap between institutional form and practice widens, and the failure mode of resource erosion and defection follows.

7.3 The temporality of institutional consolidation

The three mechanisms do not develop simultaneously. The typical sequence in cases of successful institutional consolidation involves an initial period in which incentive alignment is established — a coalition of beneficiaries is created — followed by a period in which normative stabilization accumulates, followed by the progressive deepening of organizational embedding as administrative capacity is built and professional cultures form. Different sequencing patterns produce different vulnerabilities at different points in the consolidation process. Constitutive interdependence is achieved only when all three mechanisms reach levels sufficient to sustain each other's reproduction — a condition that may require decades and that is never guaranteed.

The temporality of consolidation also implies that institutional fragility is not a stable property but a moving target. An institution that exhibits the fragile legitimacy profile at one point may transition to full constitutive interdependence as organizational capacity is built and normative frameworks deepen. Conversely, a

self-reproducing institution can be degraded toward a partial configuration by policies that systematically undermine one of the three mechanisms. The triadic architecture framework offers a vocabulary for analyzing these transitions that is more specific, and therefore more empirically useful, than general accounts of institutional stability and change.

7.4 Implications for empirical research

The framework generates several concrete research implications. Studies of institutional durability should attend not only to the presence of the three mechanisms individually but to their configuration — whether they sustain each other in constitutive interdependence. Comparative studies could examine how different institutional domains build the components of the triadic architecture in different sequences and analyze whether different sequencing patterns produce the predicted vulnerability profiles. Studies of institutional failure could test the prediction that failures take configuration-specific forms: legitimacy crises should be disproportionately concentrated in technocratic stability configurations; resource erosion in normative persistence configurations; organizational drift in fragile legitimacy configurations.

8. Conclusion

This article has argued that institutional durability is a configurational property, not a property of any single mechanism. Institutions become genuinely self-reproducing when incentive alignment, normative stabilization, and organizational embedding sustain one another in a relation of constitutive interdependence — the triadic architecture. The article's central contribution is the demonstration that partial configurations — each missing one of the three mechanisms — do not merely produce weaker versions of full self-reproduction. They produce qualitatively distinct failure modes: legitimacy crises, resource erosion, and organizational drift. Identifying these failure modes, their logic, and their conditions of occurrence is the analytical core of the configurational approach.

The article has engaged with three traditions in institutional analysis — historical institutionalism, sociological institutionalism, and political economy — and argued that each, used in isolation, misidentifies the unit of analysis. It has also positioned the concept of institutional self-reproduction in relation to Luhmann's autopoietic framework, departing from it by treating self-reproduction as a continuous and configurational variable and retaining a constitutive role for actors, coalitions, and conflict. The sociological tradition of

reproduction theory — Bourdieu, Giddens, Sewell — provides the conceptual background for this alternative account.

The concept of constitutive interdependence shifts the analytical focus of institutional analysis in a specific direction: from asking why institutions persist or change to asking how the conditions of their own reproduction are generated, sustained, and disrupted. When all three mechanisms co-produce each other's operative conditions, institutional arrangements acquire a genuine capacity for self-reproduction that is qualitatively different from the contingent stability of partial configurations. Understanding this distinction — and understanding the specific failure modes it implies — is essential for explaining why some institutional transformations become durable while others, however well designed and broadly supported, remain fragile. That, ultimately, is the practical question that this theoretical framework serves.

Appendix A. A Stylized Formal Model

The formal model is offered as an illustrative device that makes the logic of constitutive interdependence mathematically transparent. It is not offered as an empirical model, and its function is to generate clear structural implications rather than to represent institutional processes in their full complexity.

A.1 Setup. Let $I, N, O \in [0,1]$ represent the strength of incentive alignment, normative stabilization, and organizational embedding respectively. Institutional reproduction is represented as $R = I \cdot N \cdot O$.

A.2 The multiplicative specification and constitutive interdependence. The choice of a multiplicative rather than additive specification reflects the sociological claim of constitutive interdependence. An additive specification ($R = I + N + O$) would imply that the mechanisms are *substitutable*: a high value of one could compensate for the complete absence of another. The theoretical argument denies this. Constitutive interdependence, as developed in the tradition of Sewell (1992) and Bourdieu (1977), implies *non-substitutability*: each mechanism reproduces the conditions that make the others operative, so the absence of any one removes a constitutive precondition of the others' functioning, not merely a contributing factor. The multiplicative specification is the most parsimonious formal representation of this claim: when any mechanism is absent (approaches zero), R collapses to zero regardless of the strength of the others. A generalization $R = I^\alpha N^\beta O^\gamma$ preserves this non-substitutability for any positive parameters and allows for asymmetric weights; the baseline specification is the symmetric special case, chosen for parsimony.

A.3 Constitutive interdependence in partial derivatives. The cross-partial derivatives are $\partial^2 R / \partial I \partial N = O > 0$, $\partial^2 R / \partial I \partial O = N > 0$, and $\partial^2 R / \partial N \partial O = I > 0$. All cross-partial derivatives are strictly positive at any interior point: each mechanism becomes more productive of institutional reproduction as the others are present. This is the formal counterpart of the sociological claim that the mechanisms co-produce each other's operative conditions.

A.4 Dynamics. Let $S_{t+1} = \alpha S_t + \beta INO$, where $\alpha \in (0,1)$ captures institutional persistence and $\beta > 0$ captures the strength of the reproductive feedback generated by the triadic architecture. At steady state $S^* = \beta INO / (1 - \alpha)$. The equilibrium is globally stable for $\alpha < 1$. If any mechanism approaches zero, $S^* \rightarrow 0$, irrespective of the strength of the other two — confirming the failure-mode logic of Section 4.4.

A.5 Threshold dynamics. Let $\theta > 0$ denote the minimum institutional support required for persistence. The arrangement survives if and only if $\beta INO / (1 - \alpha) \geq \theta$, i.e. $INO \geq \theta(1 - \alpha) / \beta$. This threshold condition is not satisfied if any mechanism is zero. Near the threshold, small shocks to any mechanism can push S below θ , generating the non-linear collapse associated empirically with institutional tipping points. This formalizes the prediction that partial configurations are more vulnerable to abrupt breakdown than the full triadic architecture.

References

- Abdelal, R. (2007). *Capital rules: The construction of global finance*. Harvard University Press.
- Acemoglu, D., & Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. Cambridge University Press.
- Bourdieu, P. (1977). *Outline of a theory of practice*. Cambridge University Press.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160.
<https://doi.org/10.2307/2095101>
- Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. Polity Press.
- Luhmann, N. (1984). *Soziale Systeme: Grundriss einer allgemeinen Theorie*. Suhrkamp.
- Luhmann, N. (1995). *Social systems*. Stanford University Press.
- Mahoney, J. (2000). Path dependence in historical sociology. *Theory and Society*, 29(4), 507–548.
<https://doi.org/10.1023/A:1007113830879>

- Mahoney, J., & Thelen, K. (2010). *Explaining institutional change: Ambiguity, agency, and power*. Cambridge University Press.
- Marshall, T. H. (1950). *Citizenship and social class*. Cambridge University Press.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363. <https://doi.org/10.1086/226550>
- North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.
- Pierson, P. (1994). *Dismantling the welfare state? Reagan, Thatcher, and the politics of retrenchment*. Cambridge University Press.
- Pierson, P. (2000). Increasing returns, path dependence, and the study of politics. *American Political Science Review*, 94(2), 251–267. <https://doi.org/10.1017/S0003055400261780>
- Scott, W. R. (2008). *Institutions and organizations: Ideas and interests* (3rd ed.). Sage.
- Sewell, W. H., Jr. (1992). A theory of structure: Duality, agency, and transformation. *American Journal of Sociology*, 98(1), 1–29. <https://doi.org/10.1086/229967>
- Streeck, W. (2014). *Buying time: The delayed crisis of democratic capitalism*. Verso.
- Streeck, W., & Thelen, K. (2005). *Beyond continuity: Institutional change in advanced political economies*. Oxford University Press.
- Thelen, K. (2004). *How institutions evolve: The political economy of skills in Germany, Britain, the United States, and Japan*. Cambridge University Press.