The Manufacturing of Bad Faith

Mechanism Design and the Doctrine of the Affected Party

Steven Carroll
Independent Researcher · Founder, KYC.co · Creator, Moral.Money

Working Paper — SSRN Submission Draft
June 2026

Acknowledgment

The author thanks the Centro de Estudos Judiciários for the receipt and library entry of an earlier paper in this sequence, *Network Proximity and the Appearance of Impartiality*,[1] an act of institutional engagement that directly informed the present work and its companion application paper.

[1]Carroll, S. (2026) *Network Proximity and the Appearance of Impartiality: Applying Erdős-Equivalent Distance Measures to the Reform of Court-Appointed Expert Standards in Civil Law Jurisdictions.* SSRN Working Paper No. 6905898.

Abstract

Legal systems routinely manufacture bad faith. They do so not through any failure of individual morality, but as the predictable output of rules that price silence at zero and candour at a cost. The mechanism is general. Wherever a body of law (i) binds two parties together in a relationship of real interdependence, (ii) offers no cheap channel through which one party may disclose a known, relevant condition before harm occurs, and (iii) defaults the allocation of liability or entitlement to the moment of crisis rather than the moment of knowledge, the system settles into a stable and undesirable equilibrium: the party holding private information about a pre-existing condition is rewarded for concealing it, the party who must act is penalised for acting, and the dispute that results is resolved years later by whichever side can sustain litigation longest. The contradiction is not incidental. It is the means by which ordinarily compliant actors are converted into strategic non-disclosers — bad-faith actors made, not born.

This paper names that failure, models it, and proposes a portable correction. Part I describes the manufacturing process: the "missing half" common to administrative, civil, commercial, environmental, and digital law alike. Part II models it formally, showing that the resulting gridlock is a genuine equilibrium — which is precisely why moral exhortation cannot dislodge it, and why only a change in the payoff structure can. Part III specifies that change: the **Affected Party Engine**, a modular mechanism built from two interchangeable variables and one immutable procedural loop — notice, a window, and a narrow evidentiary forfeiture — that requires no new court, no new agency, and no admission of fault from anyone it touches. Part IV demonstrates the Engine across a range of further domains — corporate, environmental, security, medical, consumer, and resource law — including the hard case of resources shared through a common medium, where the mechanism must visibly bend to fit. Part V addresses why a failure this structural has gone unnamed for so long, and why the answer is neither conspiracy nor negligence but something more ordinary and more durable.

The claim is not that bad actors should be punished harder. It is that the rules currently reward the wrong move, and that correcting the payoff matrix — not the character of the players — is both the more tractable problem and the more lasting fix.

A companion paper demonstrates the Engine at full length against a single national legal system.[2]

[2]Carroll, S. (2026) *The Doctrine of the Affected Party: Realigning Standing, Notice, and the Burden of Proof in Interdependent Urban Structures.* SSRN Working Paper. The companion paper works the Engine specified here against Portuguese administrative and civil law, including a documented administrative record and a structurally comparable building-collapse incident. This treatise is written to stand independently of it; the application paper is the proof that the abstract specification survives contact with a real, adversarial doctrinal environment.

Table of Contents

Part I — The Manufacturing of Bad Faith

1.1 The Argument in One Paragraph

Because what follows is detailed, it is worth stating the whole of it once, plainly, before the apparatus arrives. Two parties are bound together by something neither can dissolve — a shared wall, a shared aquifer, a shared codebase, a merged balance sheet. One of them knows something the other does not: a defect, a weakness, a condition that will matter if the relationship is ever stressed. The law, as currently written, gives that party no cheap way to disclose what they know before trouble starts, and every reason to stay quiet, because when trouble does start the law will ask only who was active at the moment of harm, never who knew about the weakness beforehand. So the knowing party stays silent, the acting party proceeds blind, and when the structure fails — physically, financially, operationally — the cost lands on whoever happened to be moving, while the one who could have prevented it cheaply, by saying so in time, pays nothing. This paper changes one thing. It creates a defined moment at which the party in the best position to know must say what they know, gives the other party a fair window and several easy ways to respond, and attaches a single, narrow consequence to silence: the silent party loses the right to blame the other for the harm their own concealment made unmeasurable — but not, and this is the whole of its fairness, for harm the acting party genuinely caused. Everything that follows is an attempt to do that one thing, in any field, without building a single new court, agency, or tax.

1.2 Bad Faith Is Not a Character Trait. It Is a Payoff.

There is a habit, deep in the way law talks about itself, of treating bad faith as a moral fact about a person — a flaw of character that the law's job is to detect and punish. The habit is understandable and, in individual cases, sometimes correct. But it obscures something more important and far more common: most of the bad faith that clogs interdependent-liability disputes is not the product of unusually dishonest people. It is the product of ordinarily honest people responding correctly to a set of rules that has made dishonesty the rational choice.

This is a strong claim, and the paper will earn it formally in Part II. Stated plainly here: when a legal regime arranges its incentives so that disclosure costs the disclosing party and silence costs them nothing, silence is not a moral failure. It is the answer the rules are asking for. A person who stays quiet under such a regime is not betraying the law's expectations; they are meeting them exactly. The fault, if the word still applies, lies not in the player but in the game.

The consequence of this reframing is the entire reason the paper exists. If bad faith in these disputes were primarily a supply problem in human character, the remedy would be exhortation, education, sanction — the familiar toolkit of asking people to be better. But you cannot legislate against a character trait that the law itself is manufacturing. You can only stop manufacturing it. That requires identifying the production process, which is the work of the rest of this Part, and then redesigning it, which is the work of Part III.

1.3 The General Shape of the Failure: The Missing Half

The production process has a recognisable structure, and once named it becomes visible nearly everywhere. Call it the **missing half**.

A missing-half failure arises whenever the duty that would prevent a harm and the remedy that addresses it after the fact are governed by two different bodies of law — two silos — neither of which is required to know what the other knows. The first silo governs the period *before* harm: who is permitted to raise a known risk, who has standing to be heard, who may insist on a precaution. The second silo governs the period *after*: who pays, who is liable, how the cost is allocated. In a well-designed system these two would speak to each other. In a missing-half system they do not, and the gap between them is precisely where the manufactured bad faith is made.

The mechanism works like this. The first silo, governing the *ex ante* period, characteristically narrows the class of people entitled to raise a known risk. It does this by classifying the effect on an interdependent party as something incidental — "reflexive," "indirect," "consequential," a private matter for another forum and another day. The party who can see the risk coming is told, in effect, that the law has no channel for their concern at the stage when raising it would be cheap. The second silo, governing the *ex post* period, then assigns liability by a rule that asks only one question: who performed the act from which the damage flowed? It does not ask whether the damaged thing was already failing. It does not ask who knew. It asks who was active, and it allocates the cost to them.

Put the two silos together and the trap closes. The party holding private knowledge of a pre-existing condition has no doctrinal channel to disclose it cheaply before harm — the first silo will not hear them — and every incentive to stay silent, because the second silo, when harm arrives, will allocate the loss to whoever was active rather than to whoever knew. Silence is not merely permitted. It is rewarded, structurally, at every stage where disclosure would have been cheap and where contesting the matter later will not be.

This shape is not confined to any one field. It recurs, with only the nouns changed, across the whole of the law:

- in **corporate acquisitions**, where a target's undisclosed liabilities surface after closing, and the law asks who now owns the company rather than who knew what was buried in it;

- in **environmental regulation**, where legacy contamination meets a new operator, and liability attaches to present activity rather than to documented prior knowledge;

- in **information security**, where a known but unpatched vulnerability becomes a breach, and the deploying party bears a loss the original vendor could have prevented at a fraction of the cost;

- in **supply chains**, where a latent defect introduced upstream becomes the final assembler's liability;

- in **healthcare**, where an undisclosed prior condition meets a treating provider's exposure to a bad outcome;

- in **digital platforms**, where a user's pre-existing account compromise becomes the platform's liability for the breach that exploits it.

Each of these is developed in Part IV. They are listed here only to make the point that the missing half is not a quirk of one doctrine. It is a structural pattern, and the pattern is the subject of this paper.

1.4 The Law Does Not Always Originate the Distortion. Sometimes It Inherits One.

It would overstate the case — and weaken it — to claim that the law manufactures bad faith from nothing. In an important class of situations, the law's contribution is subtler and, once seen, more damning: the incentive to conceal is already present in the private arrangements of one of the parties *before any legal doctrine touches the dispute at all*, and the missing-half failure consists in the law's leaving that pre-existing distortion uncorrected rather than in its creating the distortion itself.

Consider a contractor paid by the unit of activity — by the metre drilled, the hour billed, the unit produced. Such a fee structure rewards the *volume* of activity, not the *correctness* of the outcome. A driller paid per metre has a built-in reason to keep drilling past the point at which a careful operator, paid for results, would stop and test. This is not dishonesty; it is arithmetic. Now suppose the activity also risks crossing into a shared resource — drilling past a known shallow water strike, for instance, and breaching a deeper stratum whose water is saline, degrading a resource that other, unrelated parties depend on downstream. The fee structure gives the driller no reason to stop and test at the moment of the first strike. The law, for its part, asks no question that would surface the omission: it does not require testing at each material change in condition, and so the cost of an incentive misalignment that began in a private fee schedule is exported, undocumented, onto whoever is downstream. The law did not create the driller's incentive. It simply declined to install the one check that would have caught what that incentive produced.

The same shape appears, transposed, in professional advocacy. A representative compensated by the billable event, or whose standing is better served by a long and visible contest than by the swiftest correct resolution, has a structural reason to prefer the slower path — to advance a fuller, harder-fought defence in place of a single dispositive objection that would end the matter early and cleanly, even where the faster route would better serve the party represented. This observation carries no accusation against any profession or any individual; it is the driller's arithmetic in a different costume, included here because the pattern is more immediately recognisable to most readers in this form than in the language of doctrine.

A third instance extends the same logic from the individual to the institution, and it deserves separate mention because the detection cost involved is lower than in either example above — often close to zero. A utility metering a household's water or energy supply already holds, as a matter of routine billing infrastructure, the data that would reveal an anomalous and ongoing loss on the customer's side: a consumption pattern consistent with an undetected leak or fault, visible the moment a meter is read. Until that anomaly is flagged, the lost resource is, from the utility's own perspective, simply consumption — billed, collected, and revenue-neutral or better for as long as it goes unremarked. The cost of detection here is not the cost of a forensic survey or a professional inspection; it is the cost of a routine reading being read for what it already shows.

Where a utility's practice nonetheless varies — one account flagged for anomalous use as a matter of course, an otherwise identical account left unflagged for years until the customer's own complaint forces the issue — the variance is itself evidence that nothing about the detection is technically difficult. What differs is only whether anyone acts on a signal already in hand. This is, if anything, the starkest illustration in this section of the principle the chapter is making: the harder the detection, the more sympathy a failure to disclose deserves; the cheaper the detection, the less.

In all three cases the lesson is the same: the missing-half failure rarely acts alone. It commonly operates on top of an incentive structure the law never examines, because that structure sits one level below the doctrine the law actually regulates — and the more capable the holder of that structure is of detecting the relevant condition cheaply, whether by professional judgment or by a meter already installed for another purpose, the weaker any excuse for inaction becomes.

This matters for the design that follows, and it bounds the paper's claim honestly. The Engine of Part III does not, and cannot, fix the underlying fee-structure misalignment — that is a contract-level problem, in a different domain, beyond this paper's reach. What the Engine can do, and all it claims to do, is ensure that the *information* the misaligned incentive caused to be withheld becomes visible and actionable before the harm compounds. It corrects what the law can reach. It does not pretend to correct what created the temptation to stay silent in the first place.

1.5 Why "Manufacturing" Is the Right Word

The word is borrowed, deliberately, and it is worth saying what is borrowed and what is not. In its most familiar use, "manufacturing" paired with a thing one would not expect to be manufactured — consent, in the original — describes a process that is systemic rather than conspiratorial: an outcome produced reliably by the structure of a system, serving the interests of those the structure favours, without anyone needing to choose it or even to notice it is happening. That is exactly the sense intended here. The manufacturing of bad faith is not a plot. No legislature sat down to reward concealment; no court intends the gridlock its doctrines produce. The process is automatic, and its automaticity is the point.

This must be stated plainly and kept in view throughout, because the argument is easily and wrongly heard as an accusation. It is not. The institutions implicated in the missing-half failure — the agencies that deny standing, the courts that allocate liability by activity, the bodies that publish notice no one reads — are not acting in bad faith themselves. They are each behaving reasonably within their own silo, applying rules that are, taken one at a time, defensible. The failure is emergent. It lives in the seam between the silos, where no single institution is responsible and therefore none is moved to act. What the paper describes is a reflex of the system, not the malice of any actor inside it — and the corrective it proposes is designed accordingly, asking no institution to confess error and every institution to recognise an improvement in its own position.

1.6 The Cost

The cost of manufactured bad faith is not only, or even primarily, a cost in fairness, though it is that. It is a cost in pure waste. Every dispute the process produces arrives at adjudication stripped of the one thing that would have resolved it cheaply: a record of the relevant condition, made

before the harm, when making it was easy. In its absence, the dispute must be reconstructed from degraded evidence, years after the fact, as a multi-party contest over causation in which each side has every incentive to tell a self-serving story and no contemporaneous document exists to discipline either. The adjudicating system inherits, as a matter of routine, exactly the disputes that a single piece of paper, obtained at the right moment, would have prevented.

This is the legal-system equivalent of a transaction cost that exists for no reason but the absence of a mechanism.[3] And a cost that exists only because a mechanism is missing is, by definition, a cost that the right mechanism removes. If the failure is a payoff problem — and the next Part shows that it is, in the precise sense the term carries in the study of incentives — then it should admit a formal description, and a formal description should point to its own correction.

[3]See Coase, R.H. (1960) "The Problem of Social Cost," *Journal of Law and Economics*, 3, pp. 1–44, on the role of transaction costs in determining whether an externality is resolved by private bargaining or left to fall where it lands.

Part II — The Equilibrium of Legal Gridlock

2.1 Disclosure as a Game

The situations described in Part I share a structure that the study of strategic behaviour was built to analyse, and rendering that structure in its proper terms does more than dress the argument in borrowed rigour. It shows precisely why the gridlock is stable, why it resists the remedies ordinarily aimed at it, and what kind of intervention could actually move it.

Two parties stand in a relationship of interdependence. The first — call it the **Active Party** — is the one whose conduct will eventually trigger scrutiny: the party who builds, acquires, deploys, treats, integrates. The second — the **Passive Party** — holds private information about a pre-existing condition relevant to the interdependence: knowledge of the failing wall, the contaminated soil, the unpatched flaw, the buried liability. Each party has a choice. Each may **Disclose** or **Withhold**. The question is which choice each will rationally make, given how the law currently sets the consequences — and the answer determines everything.

2.2 Why It Is Not a Prisoner's Dilemma

It is tempting to reach for the Prisoner's Dilemma here, because the situation has the surface look of one: two parties, a cooperative outcome that would benefit both, and a slide into mutual defection that benefits neither. The temptation should be resisted, and resisting it is not pedantry — it is the difference between describing the problem loosely and describing it correctly, and the correct description turns out to be the more useful one.

A Prisoner's Dilemma, in its strict form, requires that *both* players hold a dominant strategy to defect — that each is better off defecting no matter what the other does. That is not the structure here. The two parties are not symmetric. The Passive Party does hold a dominant strategy, and it is to withhold. But the Active Party holds nothing of the kind. The Active Party, in the typical case, is not weighing disclosure against silence at all; they are largely *forced to act* — they must build, must acquire, must deploy — and they do so under incomplete information, unable to see the condition the Passive Party is concealing. The Active Party is not a defector. They are a party moving in the dark because the only party holding the light has every reason to keep it covered.

The correct model, then, is not the symmetric dilemma of two prisoners but the asymmetric problem of one-sided concealment under incomplete information — the structure economists know from the analysis of markets in which sellers know the quality of what they sell and buyers do not, and from the analysis of how a party with private information can or cannot credibly signal it.[4] The Passive Party is the holder of the hidden quality; the Active Party is the party who cannot observe it and must transact anyway; and the law, by refusing to ask the one question that would force the hidden quality into the open, ensures it stays hidden.

This correction strengthens the paper's hand rather than complicating it, and it is worth saying why. A true Prisoner's Dilemma is a hard problem to fix, because both parties must be moved at

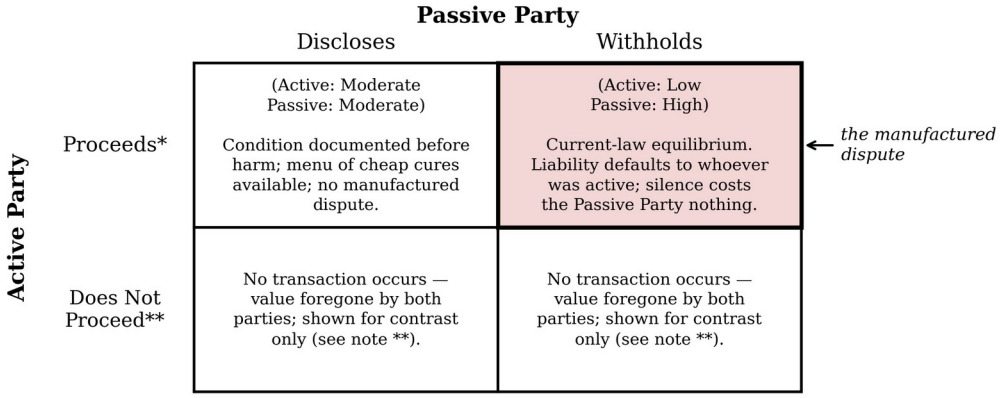
[4]See Akerlof, G.A. (1970) "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84(3), pp. 488–500; and Spence, A.M. (1973) "Job Market Signaling," *Quarterly Journal of Economics*, 87(3), pp. 355–374.

once. A one-sided concealment problem is, by comparison, a far more tractable target: only one party's incentives are misaligned, which means only one party's incentives need to be changed. The asymmetry that makes the current situation unfair is the very thing that makes it fixable.

2.3 The Equilibrium of Silence

With the structure correctly described, the central result follows directly. Under the law as it stands, withholding is the Passive Party's best response to anything the Active Party might do. If the Active Party proceeds — as they generally must — the Passive Party's silence costs them nothing, because the second silo will allocate any resulting loss to the Active Party as the one who was active. If the Active Party were somehow to delay or investigate, the Passive Party's silence still costs them nothing, because no rule compels disclosure and no penalty attaches to its absence. There is no scenario, given current defaults, in which the Passive Party is made worse off by staying silent. Withholding therefore dominates, and the system comes to rest in the outcome that withholding produces.

Figure 1 — Payoff Matrix Under Current Law



Payoffs are ordinal and illustrative, not cardinal measurements.
* Active Party's row reflects observed behaviour under operational necessity, not a freely chosen strategic alternative — see §2.2.
** Shown only to indicate that withdrawal forecloses the interaction entirely; not analysed as a live competing choice.
Within the realised row (Proceeds), Withholds weakly dominates Discloses for the Passive Party — the claim this figure illustrates.

This is a Nash equilibrium in the exact technical sense:[5] a state from which no party can improve its position by changing its own strategy alone. And that technical fact carries the paper's most important practical implication. An equilibrium is not a habit, a culture, or a moral climate. It is a structure of incentives in which the current behaviour is each party's best available response to the others'. You cannot exhort your way out of an equilibrium. Telling the Passive Party to be more forthcoming, urging good faith, appealing to professional conscience — none of it can work, not because people are irredeemable, but because the structure has made silence the correct move, and a correct move is not abandoned in response to being asked nicely. The only thing that dislodges an equilibrium is a change to the payoffs that constitute it.

[5]Nash, J.F. (1950) "Equilibrium Points in N-Person Games," *Proceedings of the National Academy of Sciences*, 36(1), pp. 48–49; Nash, J.F. (1951) "Non-Cooperative Games," *Annals of Mathematics*, 54(2), pp. 286–295.

Aggregate this single-pair equilibrium across an entire docket and the result is a structural account of something usually explained in other terms. The backlog of civil disputes in interdependent-liability matters is routinely attributed to under-resourcing or to a litigious culture. The equilibrium analysis suggests a deeper and less contingent cause: the backlog is the predictable downstream accumulation of an upstream incentive failure. Each manufactured dispute is one pair's equilibrium playing out; the caseload is the sum of them. This is a more durable explanation than under-funding, because it predicts that adding resources will not clear the backlog so long as the upstream incentive continues to produce new disputes at the rate the old ones are cleared — which is, broadly, what observation shows.

2.4 The Cheapest Cost Avoider, and an Inversion

The economic analysis of accidents supplies a principle that names exactly what the missing-half failure violates.[6] The principle holds that the cost of a harm should fall on whichever party could have avoided it most cheaply — not on whoever happened to be nearest the harm when it occurred, but on whoever was best placed, at the lowest cost, to prevent it. Liability assigned this way is not merely fair; it is efficient, because it puts the incentive to prevent harm precisely where prevention is cheapest.

The missing-half doctrine inverts this exactly. It assigns the cost not to the cheapest avoider but to the most causally proximate party — to the Active Party, caught in the act, mid-construction or mid-acquisition or mid-deployment — while the party who could have prevented the whole thing at the lowest cost, by disclosing what they already knew, is asked to bear nothing. The Passive Party is, almost by definition, the cheapest cost avoider: they hold the relevant information already, so their cost of preventing the harm is merely the cost of speaking. The law's allocation runs in the opposite direction from the one efficiency requires.

This sets up precisely what the Engine of Part III is built to do. It is not enough to declare that liability *ought* to track the cheapest avoider, because such a declaration still leaves the question to be litigated after the fact — still requires a court, years later, to reconstruct who could have prevented what and at what cost. The Engine does something more useful. It builds an *ex ante* procedure through which the cheapest avoider is induced to identify themselves, by disclosure, before the question of fault ever arises. It converts a principle that currently operates, if at all, only in retrospect into a mechanism that operates in advance.

2.5 Estoppel, With a Clock

The correction the Engine applies is not, for the most part, a new legal idea. It is an old one, given a feature it has always lacked: a defined moment of operation.

Both great families of law already contain a doctrine for precisely the wrong the missing-half failure embodies — the wrong of allowing another to rely on a state of affairs and then, when it suits, asserting the contrary to their detriment. The civil-law tradition forbids it under the principle that no one may act against their own prior conduct; the common-law tradition reaches

[6]Calabresi, G. (1970) *The Costs of Accidents: A Legal and Economic Analysis.* New Haven: Yale University Press. The principle is conventionally known as the Cheapest Cost Avoider.

the same place through estoppel.[7] Both hold, in their respective idioms, that a party who by silence or representation permits another to proceed on a given understanding may not later turn around and deny it when the denial would harm the party who relied.

The doctrine, in other words, already contains the logic of the Engine's forfeiture. What it lacks is a trigger. As things stand, estoppel-type reasoning — where it enters these disputes at all — enters reactively, mid-litigation, as one contested argument among many, to be established after the fact by whichever party can persuade a court that reliance occurred. It has no fixed moment, no defined event that sets it running. The Engine's contribution is to supply that missing moment: to attach the estoppel to a dated, documented procedural event — the notice and the window described in Part III — so that the doctrine operates determinately and in advance, rather than being reconstructed argumentatively years later.

A second and closer precedent exists for the Notice stage specifically, and it is worth naming because it shows the Engine asking for less than it might appear to. The law of fiduciary and professional duty already imposes, within one narrow class of relationship, exactly the affirmative obligation the Engine proposes to generalise: a professional who knows of a defect curable within a closing window — a filing that can still be corrected, a defence that can still be amended, a deficiency that can still be remedied before a deadline makes it irreversible — already owes their client a duty to say so, and is already liable in negligence for staying silent and letting the window close. This is not the passive logic of estoppel, which merely prevents a party from later denying what they allowed another to rely on. It is an active duty to speak, enforced today, inside the one relationship — advisor and client — where the law has already decided that the cost of silence should fall on whoever held the knowledge. The missing-half failure this paper describes is, in this light, not a gap in legal technology at all. It is a duty the law already knows how to write and already enforces in one corner of itself, never asked why it should stop there.

This framing carries a practical consequence that matters enormously for adoption, and the paper will return to it. Because both substantive doctrines already exist — estoppel for the forfeiture, fiduciary duty for the notice — the Engine is not, in most jurisdictions, a request for new law. It is a request to give two existing doctrines a clock, a form, and a reach beyond the one relationship each currently occupies. That is a far smaller institutional ask than the invention of a new liability theory or the creation of a new tribunal — and the smallness of the ask is, as Part III will show, one of the Engine's principal strengths.

2.6 Reverse Game Theory

There is a name for the discipline of changing a game so that the behaviour one wants emerges as the rational choice of self-interested players. It is the inverse of ordinary game theory: rather than taking the rules as given and asking how rational players will behave, it takes the desired behaviour as given and asks what rules would produce it as the equilibrium outcome. It is, in the most literal sense, the design of mechanisms.[8]

[7]On the comparative civil- and common-law treatment of this principle, see Zimmermann, R. and Whittaker, S. (eds.) (2000) *Good Faith in European Contract Law.* Cambridge: Cambridge University Press.

This names the whole ambition of the paper precisely. Parts I and II have diagnosed an equilibrium — a stable state of mutual silence and downstream gridlock from which no party can unilaterally escape. Part III does not respond to that diagnosis with an appeal for better conduct, because Part II has shown that appeals cannot move an equilibrium. It responds by re-specifying the payoffs, so that disclosure — currently the costly move, the move no rational Passive Party makes — becomes the dominant strategy instead. The Engine is not a code of conduct asking the players to be better. It is a corrected payoff matrix that makes the better move the rational one.

What follows, then, is not an exhortation. It is an engine.

[8]Hurwicz, L. (1973) "The Design of Mechanisms for Resource Allocation," *American Economic Review*, 63(2), pp. 1–30; Myerson, R.B. (1981) "Optimal Auction Design," *Mathematics of Operations Research*, 6(1), pp. 58–73. Hurwicz, Eric Maskin, and Myerson shared the 2007 Nobel Memorial Prize in Economic Sciences for having laid the foundations of mechanism design theory.

Part III — The Affected Party Engine

3.1 Design Constraints

Before the mechanism is specified, the constraints it must satisfy should be stated as constraints, because they are what distinguish a workable correction from a merely well-intentioned one, and because every design choice that follows is answerable to them.

It must require no new institution. The Engine must run on the courts, regulators, and professional bodies that already exist. A correction that depends on a new tribunal or agency to administer it has, in effect, relocated the problem rather than solved it, and has made its own adoption contingent on an act of institution-building that may never come.

It must operate prospectively only. The Engine adjudicates nothing about conduct that predates its adoption. It reopens no settled matter, reallocates no past liability, and asks no institution to answer for what it did before the rule existed. This is not only a matter of fairness; it is a matter of adoptability, since a reform that threatens to unsettle past decisions gives every party with a stake in those decisions a reason to oppose it.

It must be fault-independent at the point of entry. Triggering the Engine must require, and must imply, no admission of wrongdoing by anyone. The conditions that bring it into operation are documentary and objective; they describe a state of affairs, not a verdict on conduct. A mechanism that could be entered only by first conceding fault would be a mechanism no one enters.

It must resolve on documents, not narratives. Wherever the relevant fact already exists in recorded form — a licence, a filing, a report, a registered instrument, a contract term — the Engine must resolve on that record rather than open a fresh contest of competing accounts. The entire purpose is to substitute documentation for adjudication wherever documentation already exists, which in the cases this paper concerns is nearly always.

It must be symmetrically available. Either party must be able to invoke it. An instrument usable by only one side becomes a weapon rather than a mechanism, and a weapon invites the strategic abuse the Engine exists to end.

3.2 Variable A — Interdependence

The Engine takes two inputs. The first is the existence of a qualifying interdependence between the parties — the condition that makes one party's state genuinely the other's concern, and without which the Engine has no business operating at all.

In the abstract, Variable A is satisfied by a documented or readily verifiable linkage — structural, contractual, or systemic — such that the condition of one party materially affects the exposure of the other. The test is deliberately objective and deliberately low-cost: would a competent professional in the relevant field — an engineer, an auditor, a security researcher, an actuary, a physician — recognise the linkage without needing to conduct a fresh investigation? Where the answer is yes, Variable A is met. The interdependence is not something the Engine

asks anyone to adjudicate; it is something the Engine asks only to be recognised, on facts that already exist.

Variable A instantiates differently across domains — a shared foundation in one, a shared dependency in code in another, a merged balance sheet in a third — and Part IV works these out. But one distinction within Variable A is structural rather than merely illustrative, and it must be drawn before the mechanism is specified, because it determines how one stage of the mechanism can operate.

3.3 Two Kinds of Interdependence: Discrete and Diffuse

Most of the cases this paper has so far described involve what may be called **discrete interdependence**: a fixed and identifiable pair, or small named set, of parties joined by something nameable — a party wall, a shared permit, a contract, a single integration between two systems. Discrete interdependence is the easy case for the Engine, because the party who must be notified is already known. There is a specific someone on the other side.

A harder kind deserves naming on its own terms: **diffuse interdependence**, in which the linkage runs not through a discrete connection but through a shared, boundary-crossing medium — an aquifer, an airshed, a watershed, a common network. Here the interdependence is no less real, and is often no less well understood as a class of risk, but the affected party may be unidentified, plural, not yet harmed, or downstream in a way not immediately traceable to any specific actor. The geological fact that two landowners draw from the same body of water is a matter of regional record; the identity of every party who will eventually be affected by a given contamination of it may be unknowable at the moment the contamination begins.

This distinction matters because it bears directly on the notice stage specified below. Where the affected party cannot be individually identified at the time of the triggering conduct, the Engine cannot rely on direct, individually addressed notice, which is the natural form for the discrete case. The honest response — and the paper states it as a genuine limit rather than glossing it — is that diffuse interdependence requires the notice to be routed differently: through the regulatory or licensing body that already aggregates information about the shared resource and is therefore best placed to know who else draws upon it, or, where even that is impractical, by treating the threshold question itself as the trigger for a duty to test and to report to that body, in advance of any specific downstream party being known to exist. The clean bilateral form of the loop is the discrete case’s privilege, not a universal default. The diffuse case is harder, the Engine bends to meet it rather than pretending it does not arise, and Part IV returns to it as the gallery’s hardest instance precisely so that the bending can be shown in operation rather than asserted in principle.

3.4 Variable B — Documented Irregularity

The second input is the condition that makes the interdependence a present concern rather than a latent fact: a documented irregularity in the state of one party, which the other has an interest in knowing before proceeding.

Variable B is satisfied by any one of a defined set of objective, third-party-verifiable conditions — and, importantly, by a residual category anchored to professional judgment on the record, included because the forms an irregularity can take are not exhaustively foreseeable. The

companion paper works out the full catalogue for a single domain; in the abstract, the structure is what matters, and the structure is **a closed list plus a disciplined residual**.

The design reason for this combination is worth stating, because each half answers the other's weakness. A closed list alone goes stale: it captures the forms of irregularity its drafters could foresee and misses the ones they could not, growing less adequate with every year and every novel case. A pure residual standard alone — "any condition a professional would find concerning" — invites abuse and unpredictability, becoming a general-purpose instrument for objection rather than a narrow evidentiary trigger. The combination is durable without being limitless: the closed list handles the ordinary cases predictably, and the residual handles the unforeseen ones, but the residual is anchored to a professional finding on objective evidence, not to suspicion or apprehension, so that it admits the genuinely novel irregularity without opening a door to the speculative grievance the Engine is built to exclude.

One sub-component of Variable B deserves separate mention because it travels unusually well across domains: **the market test.** Wherever a regulated market exists for pricing the relevant risk — liability insurance for a physical structure, warranty markets for goods, cyber-cover for software — a documented refusal by that market to underwrite the *specific disclosed condition* is itself evidence that the condition is real. The refusal must be specific: a blanket unwillingness to write a category of risk at all proves nothing about a particular case, being a fact about the market rather than about the property. But a regulated insurer who, with the condition fully disclosed, declines in writing to cover it has performed, at no cost to the law, exactly the diagnostic the law would otherwise have to perform itself. The market's "no" is a finding the Engine can use.

3.5 The Loop: Notice, Window, Forfeiture

The two variables are the inputs. The procedure that runs on them is a single loop of three stages, and the loop is the heart of the Engine. It is specified here as a protocol — inputs, operation, and the consequence of each path — because that is what it is.

Stage one: Notice. The party in the best position to know that the interdependence exists serves direct notice on the other. In the discrete case this is the party who, in the ordinary course of their work, already knows the interdependence is there — characteristically the professional designing or executing the new activity, who knows it because they are working against it. The notice identifies the interdependence (Variable A) and the specific documented condition (Variable B), in a form the recipient can act upon. It is served directly, individually, and with proof of delivery — and the Engine explicitly rejects constructive or public notice as a substitute in the discrete case. Notice by publication to a party who is known, named, and individually reachable is not a lighter form of notice; it is the appearance of notice without the substance, and the missing-half failure thrives on exactly that substitution. (For the diffuse case, notice routes through the regulator, as Section 3.3 sets out.)

Stage two: Window. The notice opens a defined period within which the notified party may respond — and the menu of responses is arranged deliberately, from the cheapest and least intrusive to the most. The party may **transfer the risk through insurance**: obtain cover or an equivalent regulated instrument that will respond if the condition causes harm, secured before the activity begins, which closes the matter outright because a third party, with capital standing

behind the promise, now bears the risk. They may, in the alternative, **indemnify directly**: rather than routing the risk through a regulated underwriter, the notified party may give the other a personal, direct undertaking to make good any harm that traces to the disclosed condition, leaving the condition itself untouched. This option deserves to be stated on its own terms rather than folded into insurance, because the two differ in exactly the respect that matters: insurance is backed by a regulated balance sheet; direct indemnity is backed only by the indemnifying party's own. It is, for that reason, the cheapest option on the menu to give — no premium, no underwriting delay, no cure cost — and it is also, for the same reason, the option that says the most. A party prepared to stake their own money on the proposition that a disclosed condition will not cause the harm it appears capable of causing is making a costly, self-revealing move of exactly the kind Section 2.2 identified as the missing signal in this class of problem; a party who will not make even that promise, while continuing to insist the condition is benign, has revealed something about their own confidence in the claim. The option must, however, carry a floor: an indemnity from a party with no demonstrable capacity to honour it is no indemnity at all, and the Engine cannot treat a signed promise from an insolvent party as having discharged the Window, on pain of allowing exactly the costless escape the forfeiture exists to close. Whether the indemnifying party can make good on the undertaking — by reference to other assets, a guarantee, or security given for the promise itself — is therefore part of what the Window's instrument must establish, and an indemnity that cannot clear that bar does not count.

They may also **cure the condition** themselves, on their own terms, with their own contractor, at their own cost and schedule. They may **agree to share the cost**, where they cannot fund a cure alone but will cooperate, through a consensual charge registered against their own future proceeds — an ordinary, existing instrument, requiring no new legal machinery, repaid when the asset is eventually realised. Or they may **commission an independent baseline assessment**, establishing the condition's actual state so that any later dispute turns on a documented figure rather than a contest of accounts. The arrangement of this menu is itself the mechanism's central move: every option on it is a way of saying yes, and every option leaves the notified party better off than the alternative the next stage describes. The law's task is only to make saying yes, in any of these forms, more attractive than saying nothing.

Stage three: Forfeiture. Only if the notified party does none of these things — will not insure, will not indemnify, will not cure, will not agree a charge, will not even commission the assessment that would document their position — does any consequence follow. And the consequence is narrow, deliberately and crucially so. The party who lets the window close having done nothing forfeits one thing only: the benefit of the doubt regarding the undocumented pre-existing condition. They lose the right to claim that the other party's activity caused harm that, in truth, traces to their own untreated and now-unmeasurable condition. They do **not** forfeit title. They do **not** forfeit standing in general. And they do **not** — this is the point on which the whole fairness of the Engine turns — forfeit protection against harm the Active Party's own conduct genuinely causes.

This last limit must be read precisely, because its most likely misreading would make the Engine unjust. The forfeiture is not immunity for the Active Party. What it transfers to the silent party is a single, specific risk: the risk of the pre-existing condition that can no longer be measured *because they prevented its measurement by staying silent*. It transfers that risk for one reason only — that the party was given notice, a window, and every means to have the condition

documented, and chose to let the chance pass. It does not transfer, and cannot transfer, the risk of the Active Party's own conduct. If the new activity is performed carelessly and causes harm that a sound structure would equally have suffered, the forfeiture reaches none of it; that harm traces to the conduct, not to the absent baseline, and the ordinary law of liability continues to govern it in full. The forfeiture withdraws the benefit of the doubt on the portion of any failure that traces to untreated decay. It leaves the Active Party carrying the portion that traces to their own action. Each party bears the risk that was always theirs, and neither is permitted to push it onto the other — which is the entire purpose of the rule.

This is estoppel with a clock, exactly as Part II described. The substantive principle — that you may not benefit from your own silence to another's detriment — is unchanged. What is new is only the dated procedural event that makes the principle operate determinately, in advance, rather than being argued after the fact.

3.6 Why the Loop Converts the Equilibrium

It is worth being precise about what the forfeiture accomplishes in the terms of Part II, because this is the mechanism's whole claim. Under current law, the Passive Party's silence carries a cost of zero, which is why silence dominates and the equilibrium holds. The forfeiture changes that single number. Once silence carries a defined, certain cost — the loss of the causation defence on the unmeasurable portion of any future failure — withholding is no longer free, and the five cheap forms of disclosure on the window's menu are each, for a rational party, a better outcome than the forfeiture that follows from doing nothing. Disclosure becomes the dominant strategy. The equilibrium that Part II showed to be immovable by exhortation moves, because the thing that held it in place was a zero that the Engine has replaced with a cost.

Nothing about this requires the Passive Party to become more honest, more cooperative, or more public-spirited. It requires only that they remain rational, and act in their own interest — which, the Engine having corrected the payoff, now points toward disclosure rather than away from it. That is the difference between asking the players to behave better and redesigning the game so that their existing self-interest produces the better behaviour on its own.

3.7 What the Engine Does Not Require

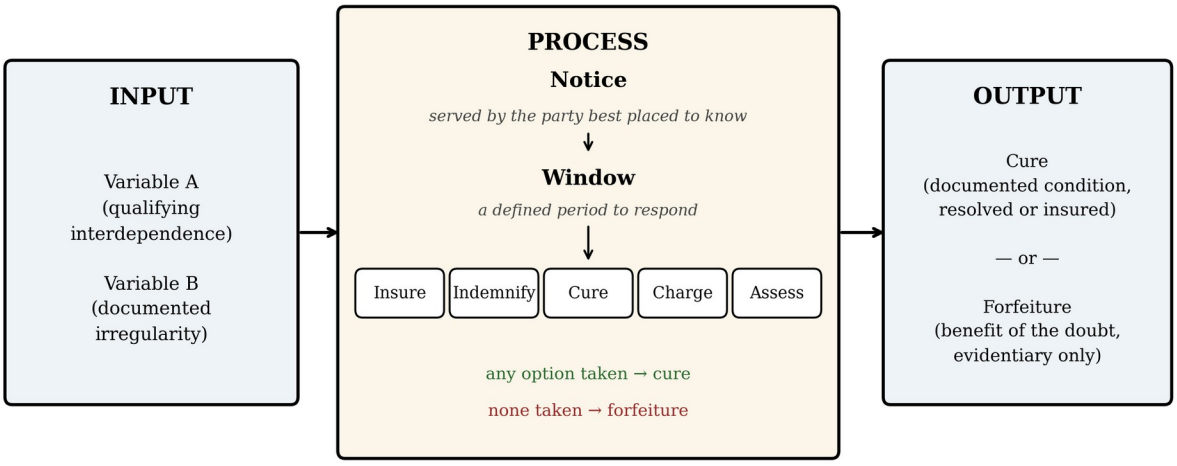
Because the section every adopting institution will read first is the one listing what the Engine will cost it, that list is given here, plainly, as the constraints of Section 3.1 now discharged.

The Engine requires no new tribunal and no new agency. It reopens no past conduct and reallocates no settled liability. It demands no admission of fault as the price of entry, the triggering conditions being documentary and fault-independent by design. And it leaves the underlying substantive law of liability untouched — it operates entirely on the evidentiary plane, changing what a silent party may later claim, not the rule by which liability is ultimately assessed. The substantive allocation of risk that any given field already provides is left exactly as it stands; the Engine adds only a procedure that documents the relevant condition in time, and a consequence for refusing to.

3.8 The Engine as Specification

Stated at its most compact, the Engine is two inputs and one procedure. The inputs are Variable A, a qualifying interdependence, and Variable B, a documented irregularity. The procedure is the loop: notice from the party best placed to know, a window offering risk-transfer or cure or a shared charge or a baseline assessment, and — only on the failure of all of these — a narrow evidentiary forfeiture of the benefit of the doubt on the unmeasurable pre-existing condition.

Figure 2 — The Affected Party Engine as Specification



Only the content of Variable A and Variable B changes between fields; the Notice → Window → Forfeiture loop holds constant. Diffuse interdependence (§3.3) reroutes Notice through a regulator rather than to an individually addressed counterpart; see §4.7.

What makes the Engine portable is that only the content of the two variables changes from one field to the next. The interdependence is a shared wall in one domain and a shared dependency in another; the irregularity is a structural defect here and an unpatched vulnerability there. But the loop — notice, window, forfeiture — holds constant across every field in which the inputs can be defined. That constancy is the claim Part IV now tests.

Part IV — Applications in the Wild

The purpose of this Part is narrow and specific: to show that the Engine's grammar holds across radically different bodies of law without alteration to the loop. Each case below states only what Variable A and Variable B become in that domain, and what the forfeiture would withhold. None is a worked reform; each would require its own application paper, on the model the companion paper provides for one national property-law system. The point of the gallery is not depth in any one field but portability across all of them — and the gallery is arranged to end on its hardest case, because a mechanism shown working only on convenient examples earns less trust than one shown bending, honestly, to meet the case that strains it.

4.1 Corporate Acquisitions and Undisclosed Liability

Variable A is the interdependence a buyer assumes at closing — the contracts inherited, the liabilities assumed, the operations merged into the acquirer's own. Variable B is a known but undisclosed exposure on the target's side: pending litigation not surfaced, a regulatory non-compliance not declared, a contractual breach not flagged in the data room.

The market has already built a partial, informal version of the Engine here, in the representations-and-warranties insurance that now routinely backstops these transactions — evidence, in itself, that the disclosure problem is real and that risk-transfer is a natural response to it. The same recognition appears at the opposite end of the transaction scale, in a register most readers will find more immediately familiar than a data room. Many residential property regimes already impose some duty on a seller to disclose known material defects, qualifying the older rule that a buyer takes a property as found; and insurers, quite separately, routinely exclude cover for a condition shown to have been known to the policyholder and withheld at the time cover was sought. The warranty insurer, the disclosure statute, and the exclusion clause are the same logic, arrived at independently, at three different scales and through three different instruments — and none of them, including the most sophisticated, supplies what the Engine supplies. Each operates only after a dispute has already arisen, deciding in retrospect whether disclosure should have occurred, rather than creating the dated, *ex ante* moment at which it must. The insight recognises itself, apparently, every time a transaction is stressed enough to require it. What has been missing is not the insight. It is the procedure that lets the insight act before the loss rather than after it.

What the Engine would add, concretely, is the formal loop: a defined pre-closing notice-and-cure window, and a specific evidentiary forfeiture — the loss of the "we did not know" defence in a later dispute — attached to the failure to disclose within it. The substantive law of warranty, indemnity, and disclosure is left untouched at every scale; what changes is that the moment of disclosure becomes dated and documented rather than reconstructed afterward, whether the asset in question is a company or a house.

4.2 Information Security and Vulnerability Disclosure

Variable A is the dependency between a software vendor and the parties who deploy what the vendor ships — a supply-chain linkage in code, often invisible until it fails. Variable B is a

known, unpatched vulnerability: a defect the vendor is aware of and has not remediated, the close analogue of a structure whose remediation was authorised and never completed.

The forfeiture would fall on a vendor who, notified of a vulnerability, takes none of the window's options — neither patches, nor discloses to downstream deployers, nor obtains cover for the specific exposure — and thereby forfeits the ability to contest liability on the ground that the breach's root cause was unknown. What makes this case notable is that the security industry has independently arrived at the Engine's timing logic already, in the "responsible disclosure" windows that have become standard practice:[9] a defined period within which a known vulnerability must be addressed before disclosure consequences follow. The Engine does not impose a novel discipline here. It generalises one the field already validated on its own.

4.3 Environmental Liability and Legacy Contamination

Variable A is the interdependence of successive owners of a single contaminated site. Variable B is documented prior contamination — or contamination evidenced through the market test, where no environmental liability cover can be obtained for the disclosed condition.

This case has the most extensive scattered precedent of any in the gallery: several jurisdictions already provide some version of an innocent-landowner or prior-condition defence, recognising that liability should not fall on a new owner for contamination documented before their tenure. The fragmentary existence of these defences is itself an argument for the Engine — not that the logic is novel, but that it has been arrived at piecemeal, field by field and jurisdiction by jurisdiction, and never unified into a single portable procedure. The Engine's contribution is the unification.

4.4 Healthcare and Pre-Existing Conditions

Variable A is the interdependence between a patient's undisclosed condition and a treating provider's exposure to a bad outcome. Variable B is a documented prior diagnosis or treatment history not disclosed at intake.

This case must be handled with more care than any other in the gallery, and the care is instructive about the Engine's design. It is the one domain here in which the Passive Party — the patient — is typically the more vulnerable party, not an institution, and in which an ordinary liability-shifting rule would be plainly inappropriate. The reason the Engine can operate here at all, where a cruder instrument could not, is its two defining restraints: it is fault-independent, and its forfeiture is evidentiary only. It would not, and could not, diminish any obligation of care, nor shift the cost of harm the provider's own conduct causes. It would bear only on the evidentiary weight of a later claim that the provider caused harm in fact attributable to an undisclosed prior condition. That this domain admits the Engine only because of restraints built for entirely different reasons is a point in the design's favour: the narrowness is not a limitation but the thing that makes the mechanism safe to extend into sensitive ground.

[9]International Organization for Standardization / International Electrotechnical Commission (2018) *ISO/IEC 29147:2018 — Information Technology: Security Techniques: Vulnerability Disclosure.* 2nd edn. Geneva: ISO/IEC.

4.5 Digital Platforms and Account Compromise

Variable A is the interdependence between a user's own account-security practices and a platform's liability for unauthorised access. Variable B is a documented prior compromise, credentials flagged in a known breach database, or declined security features made available at no cost — the last a near-exact analogue of the market test's declined-risk-transfer logic.

The forfeiture would fall on a user notified of a specific, documented weakness in their own account security who takes none of the offered remedies within the window, and who thereby forfeits the right to treat a subsequent compromise exploiting that specific weakness as solely the platform's failure. As in every other case, the limit holds: the user forfeits nothing in respect of a breach the platform's own conduct causes. The forfeiture reaches only the harm that traces to the documented weakness the user was notified of and declined to address.

4.6 The Companion Pair: Incentive Misalignment in Plain Sight

Two brief cases return, in concrete form, to the point of Section 1.4 — that the missing-half failure often compounds an incentive distortion the law never examines. They are included not as fresh instances of Variables A and B but because they make the pattern recognisable to any reader, in relationships most people have stood inside.

A contractor paid by the unit of activity — the metre, the hour, the item — has no incentive in their own fee structure to stop and test at the point where continuing risks crossing into a documented irregularity affecting a shared resource. The Engine does not fix the fee structure; it requires testing and disclosure at defined trigger points regardless of what the fee structure rewards, converting an invisible omission into a documented one. And a representative whose compensation or standing is better served by a long contest than by the swiftest correct resolution has a structural reason to prefer the slower path — the same arithmetic, in a different costume, carrying no imputation against anyone but illustrating the same truth: the law's missing half frequently sits atop a private incentive the law does not regulate, and the Engine corrects the half the law can reach while making no pretence about the half it cannot.

4.7 Shared Groundwater: The Hardest Case

The gallery ends, by design, on its most difficult instance — the case of diffuse interdependence introduced in Section 3.3 — because it is here that the Engine's clean bilateral loop visibly bends, and showing the bend is more honest, and finally more persuasive, than hiding it.

Variable A, in its diffuse form, is the interdependence of two or more landowners drawing from a single aquifer system. The linkage runs through a shared medium rather than a discrete connection, and the hydrogeological fact of shared strata is typically a matter of regional geological record rather than case-specific dispute — the interdependence is establishable without fresh investigation, even where the parties affected by a given instance are not yet identified.

Variable B is the pattern that triggers it: a borehole developed in stages — water struck at a shallow depth, drilling continued substantially deeper before a usable yield was obtained, no documented water-quality testing at the first strike or at subsequent material changes in depth —

followed by measurably elevated and inconsistent salinity at the wellhead, readings that fluctuate across samples taken under similar conditions. Inconsistency of this kind is not, by itself, proof of cross-strata mixing between a shallow fresh layer and a deeper saline one. It is precisely the kind of objectively observable pattern the residual limb of Variable B exists to capture at the *threshold* stage: enough to ask the single binary question — does an unresolved condition exist, bearing on an interdependent resource — without yet requiring the full forensic hydrogeological survey that establishing causation or quantifying downstream harm would demand. The distinction the companion paper draws between the threshold question and the baseline question does real work here: the threshold is cheap to reach even where the underlying geology is expensive to resolve fully.

What makes this case the hardest is the notice problem. The affected parties are not a named counterpart but an unknown, possibly plural, possibly not-yet-injured set of downstream users of the same body of water. Direct, individually addressed notice — the form the discrete case relies on — cannot function, because there may be no specific person to address at the moment the irregularity arises. The Engine's honest extension to this case routes notice not to a counterpart but to whichever body already licenses or monitors the shared resource, the entity best placed to know who else draws upon it, converting an unaddressable bilateral notice into a registrable event that a regulator can act on or aggregate over time.

And the forfeiture must be drawn most narrowly of all here. A well-owner who, on documented irregularity, neither tests at the trigger points nor reports to the licensing authority within the window forfeits only the benefit of the doubt as to whether their own continued drilling caused or worsened the cross-strata condition. They do not forfeit title. And the forfeiture must never be read as imposing strict liability for naturally occurring regional salinity that the owner did not cause and cannot be charged with curing. The Engine's evidentiary plane, and not its substantive plane, does all of the work — exactly as the discrete case insists, but with a sharper edge, because the diffuse case is precisely where a careless drafter would be tempted to let the evidentiary forfeiture harden into a substantive liability it was never meant to become.

What this case proves about the Engine generally is the point of placing it last. The loop survives contact with its hardest instance — unidentified counterparties, a continuous rather than discrete medium, evidence that is probabilistic rather than conclusive at the threshold — by rerouting the one component that cannot survive unmodified, the individually addressed notice, while leaving Variable A, Variable B, the window, and the forfeiture's narrow scope all structurally intact. A mechanism that worked only for fully identified bilateral disputes would be a far smaller thing than the one this paper claims to have specified. That it bends to meet the diffuse case without breaking is the strongest available evidence that what has been described is an engine and not merely a single solution wearing general clothes.

4.8 What the Gallery Shows

Across seven cases drawn from corporate, environmental, security, medical, consumer, and resource law, the Engine's core logic held constant: a qualifying interdependence, a documented irregularity, a notice, a window of cheap ways to say yes, and a narrow forfeiture for saying nothing. In one case the notice had to reroute; in another the forfeiture had to be drawn with special care; in every case the loop itself was unchanged. This is what it means to claim that the

Engine is portable — not that every field is identical, but that the differences between fields are absorbed by the content of the two variables, while the procedure that runs on them stays the same. None of these sketches is a finished reform. Each is a demonstration that a finished reform, in that field, would be built from these same parts.

Part V — Correcting the Matrix, Not the Player

5.1 The Central Inversion, Restated

The paper closes where it began, but the opening claim can now be made as a conclusion rather than asserted as a premise. Bad faith, in the disputes this paper concerns, is not a supply problem in human character. It is a demand problem created by a payoff structure that has, until now, gone unexamined as a structure at all. The law has spent its effort trying to detect and punish a disposition it was itself manufacturing — asking players to be better while leaving in place the rules that made the worse move the rational one. The correction is not to ask harder. It is to stop manufacturing, by changing the single number — the cost of silence — that holds the whole equilibrium in place.

5.2 Why Mechanism Design Is the Right Tool

This points to the paper's actual methodological claim, which is larger than the Engine itself. Reform efforts that target conduct directly — heavier penalties, expanded duties of care, exhortations to good faith written into statute — are all attempts to move the players within a fixed game. Part II showed why such attempts fail against an equilibrium: the players are already making their best responses, and asking them to make worse ones, however urgently, does not change what is in their interest. The alternative is to redesign the game, so that the conduct one wants emerges from the players' unchanged self-interest rather than against it. The Engine is one instance of that method. The method is the durable export. Wherever the law produces a stable bad outcome that resists exhortation, the question to ask is not how to make the players better but how the payoffs are arranged to make the bad outcome rational — and then how to rearrange them.

5.3 Why This Has Gone Unaddressed

A reader persuaded this far will be left with a sharp question: if the failure is this structural and this general, why has it gone unnamed for so long? The question deserves an honest answer, and the honest answer is in several parts, none of which is conspiracy and none of which is negligence.

The first part is the division of expert labour. The missing-half failure lives in the seam between two silos, and the two silos are owned by two different sub-disciplines — taught in different courses, practised by different specialists, litigated by different lawyers — who have no occasion to read each other's doctrine against their own. The scholar of standing and procedure and the scholar of liability are not reading one another's cases looking for a shared shape, because nothing in either's training directs them to. The failure is invisible from inside either silo precisely because seeing it requires standing in both at once, which the structure of expertise makes nobody's job. The blindness is not inattention; it is the missing-half problem reproducing itself in the way the profession organises its own knowledge.

The second part is deeper, and it explains not only why the failure went unnamed but why it was never obvious even to someone looking directly at it. The duty to mitigate loss is a well-developed legal idea — but it runs in exactly one direction: inward. A party is expected to act

reasonably to limit damage to themselves. There is no comparably developed doctrine asking a party to act to limit damage that their own inaction is causing someone else — particularly where that someone else is not yet a party to anything, holds no existing claim, and may not even know they are affected. Mitigation was built to discipline a claimant's own recovery, not to police a non-claimant's silence toward a stranger. That gap is not an oversight discovered now in hindsight. It is what gap-filling always looks like at the moment a rule is written: a legislature or a court resolving the dispute actually before it has, by definition, no occasion to ask what its rule will do to a party who is not in the room and may never bring a case. Solving the visible half of the problem is already the entire task. The other half is, structurally, always left for whoever notices it later. That is the honest explanation for the length of the wait — not that anyone failed, but that writing law for the dispute you can see is hard enough on its own.

The third part concerns the kind of attention that finds a failure of this shape, and the paper states it as provenance rather than as any claim of singular insight. Most problem-solving, including very capable problem-solving, proceeds by terrain knowledge: a practitioner learns how a field's existing solutions are shaped, and that shape becomes the frame within which every later fix is built. This is usually the efficient choice and usually the right one. But it carries a specific blind spot — it cannot see a defect built into the shape of the terrain itself, because the terrain is supplying the frame of reference. First-principles design starts elsewhere: not from how a problem has been solved before, but from why the thing in front of you is actually broken, examined on its own structure, independent of inherited solutions. It is slower and less efficient for the ordinary case, and it is the only approach that locates a defect the terrain has normalised. The method of this paper is the latter, applied to law rather than to a machine or a market, and it is available to anyone who has spent real time taking broken systems back to their root cause rather than to the nearest precedent — a discipline of frustration and false starts more than of inspiration, and no one's private possession.

There is a useful and much-misused historical parallel, and the paper takes from it only one thing. When the mathematics of gravitation was first set down, others had already felt the shape of the thing — the sense that something connected the fall of an object to the motion of the moon. What was supplied was not the intuition that a connection existed but the formal description that let the same pattern be recognised on sight across domains that had never been considered together. The contribution of naming a pattern is not the discovery that something is wrong. It is the making of that wrongness visible everywhere at once, to people who had each, in isolation, only ever seen their own corner of it. That is the whole of the claim made here, and nothing larger: the missing-half failure was very likely already half-felt, piecemeal, in the data room and the disclosure window and the contaminated lot. What was missing was the description that lets a single observer recognise it as one failure rather than many unrelated ones.

One honest note on method belongs here, because the temptation it addresses is real and the paper would rather name it than fall to it. Once a pattern is given a name, it becomes conspicuously easy to find — the missing-half shape, once seen in one dispute, has a way of turning up again within days, in a utility bill, in a professional's missed deadline, in a conversation about a neighbour's well. This ease is worth taking seriously rather than dismissing, but it must be taken seriously as evidence of exactly one thing: that the pattern, once named, is **cheap to recognise**. It is not, by itself, evidence of how much of the law's total business the pattern accounts for. Those are different claims, answerable by different kinds of work —

recognisability is established by argument and example, of the kind this paper offers; prevalence would require systematic measurement, of a kind this paper does not undertake and does not claim to have undertaken. The two should not be allowed to blur into one another, however naturally the second seems to follow from the speed of the first.

The fourth part is a lesson from the practice of invention, and it is the reason this paper exists as a standalone treatise rather than as a passage buried inside a single application. An inventor who solves a particular instance of a general problem, and claims only the particular instance, will watch the general engine rebuilt by someone else once it has been seen — sometimes in a form that no longer resembles the original application at all. The discipline of claiming the engine rather than the instance is what separates a transitional fix from a foundational one. A single doctrinal contradiction in a single national legal system is one instance of the missing-half failure — real, fully evidenced, and worked at length in the companion paper. It is not the failure itself. The Engine specified here is offered as the general claim, deliberately, because the alternative — to publish the specific fix and let the general principle be extracted later, by someone else, in a field bearing no relation to the original — is a mistake the author has made once before and does not intend to repeat.

5.4 What Adoption Requires

The Engine has a property worth naming in closing, because it bears directly on whether anything in this paper will ever be built. The reform does not require coordinated adoption. Each institution it touches benefits from adopting it independently of whether any other does. A legislature gains a reduction in the manufactured caseload its courts must carry. A regulator gains a documented record where it now has a publication no one reads. A court gains disputes that arrive with a baseline already established rather than reconstructed from degraded evidence years on. A professional body gains a defined standard of practice in place of a contestable one. An insurer gains a risk it can price accurately because the condition is documented rather than hidden. And two private parties can adopt the loop between themselves, by agreement, with no institutional permission at all, because every instrument it uses already exists in ordinary contract and property law. No adopter needs to be persuaded that the others are right. Each needs only to notice that the alternative it currently tolerates costs it more than the alternative proposed here. A reform adoptable by each party acting alone, in its own interest, is a reform that does not depend on the coordination that defeats most good ideas — which is, fittingly, the same mechanism-design logic that the Engine itself runs on, turned back upon the question of its own adoption.

5.5 The Well

What has been described is narrower than the standing rights it is sometimes mistaken for, faster than the litigation it is built to prevent, and cheaper, for every institution it touches, than the alternative each currently tolerates. It asks for no new court, no new agency, no new tax, and no admission of fault from anyone. It changes one number — the cost of silence — and lets the players' own rationality do the rest.

The failure it corrects is not rare. It recurs wherever interdependence meets a doctrine that asks who acted rather than who knew, and interdependence is everywhere the modern world has built one thing against another. How much of the litigation now pending in any given system traces to

some instance of it is not a thing this paper can measure, and it does not claim to. What it claims is structural and, for now, sufficient: the pattern recurs wherever the conditions named in Part I are met, those conditions are not unusual, and a single procedure resolves every instance of it the same way. The magnitude is for others, with the data, to establish. The shape is what this paper offers — and a shape, once named, is hard to stop seeing.

References

Strategic Behaviour and Mechanism Design

Akerlof, G.A. (1970) “The Market for Lemons: Quality Uncertainty and the Market Mechanism.” *Quarterly Journal of Economics*, 84(3), pp. 488–500.

Hurwicz, L. (1973) “The Design of Mechanisms for Resource Allocation.” *American Economic Review*, 63(2), pp. 1–30.

Hurwicz, L., Maskin, E. and Myerson, R.B. — joint recipients of the 2007 Nobel Memorial Prize in Economic Sciences, awarded “for having laid the foundations of mechanism design theory.”

Myerson, R.B. (1981) “Optimal Auction Design.” *Mathematics of Operations Research*, 6(1), pp. 58–73.

Nash, J.F. (1950) “Equilibrium Points in N-Person Games.” *Proceedings of the National Academy of Sciences*, 36(1), pp. 48–49.

Nash, J.F. (1951) “Non-Cooperative Games.” *Annals of Mathematics*, 54(2), pp. 286–295.

Spence, A.M. (1973) “Job Market Signaling.” *Quarterly Journal of Economics*, 87(3), pp. 355–374.

Law and Economics

Calabresi, G. (1970) *The Costs of Accidents: A Legal and Economic Analysis.* New Haven: Yale University Press.

Coase, R.H. (1960) “The Problem of Social Cost.” *Journal of Law and Economics*, 3, pp. 1–44.

Doctrine and Standards

International Organization for Standardization / International Electrotechnical Commission (2018) *ISO/IEC 29147:2018 — Information Technology: Security Techniques: Vulnerability Disclosure.* 2nd edn. Geneva: ISO/IEC.

Zimmermann, R. and Whittaker, S. (eds.) (2000) *Good Faith in European Contract Law.* Cambridge Studies in International and Comparative Law, vol. 14. Cambridge: Cambridge University Press.

Author’s Prior and Companion Work

Carroll, S. (2026) *Network Proximity and the Appearance of Impartiality: Applying Erdős-Equivalent Distance Measures to the Reform of Court-Appointed Expert Standards in Civil Law Jurisdictions.* SSRN Working Paper No. 6905898. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6905898.

Carroll, S. (2026) *The Doctrine of the Affected Party: Realigning Standing, Notice, and the Burden of Proof in Interdependent Urban Structures.* SSRN Working Paper.