Machine Verification and the Institutions of Trust: Reputation, Certification, and Contract when Checking Becomes Cheap*

Amrit Gill

Independent Researcher

amritbir1@gmail.com

August 2026

Abstract

Many institutions of commerce exist because verification is expensive: brands bond unobservable quality, certifiers amortize the fixed cost of checking, and intermediaries are paid, in part, a verification wage. This paper asks what happens to those institutions when machine intelligence lowers the cost of two historically expensive tasks: establishing whether existing obligations were satisfied (*conformity*), and whether a better contractual arrangement exists (*optimality*). In a model where trading relationships choose among opacity, reputation bonds, third-party certification, and direct machine verification, five results follow. First, verification de-institutionalizes: amortization vanishes with construction costs, and because shared-certifier adoption is strategically complementary, certification markets can end by tipping rather than decay. Second, brands migrate rather than die: the bonding share of brand value compresses toward the cost of credibly attested verification while search, taste, and status shares persist. Third, optimality rents compress through repricing and exit: as individualized counterfactuals become cheap, the minimum mismatch worth acting on falls, incumbents reprice before customers switch, and compression stops at switching friction, not zero. Fourth, accusing gets cheap faster than proving innocence, expanding nuisance challenge unless rebuttal costs fall symmetrically or fee-shifting intervenes. Fifth, opacity relocates toward unverifiable attributes. An implementation in commercial energy contracting operationalizes both layers, with the verifier construction-cost series $K(t)$ as a novel measured object.

Keywords: verification cost; artificial intelligence; reputation; brands; certification; intermediation; contract design; nuisance claims.

JEL codes: D82, D83, D86, L14, L15, K41, M42, O33.

*Working paper, August 2026. Comments welcome. This paper builds on the verification frontier developed in Gill (2026); the two papers are complements. Large language models were used extensively in drafting, revising, and stress-testing this paper; all claims, judgments, and errors are the author's own. Disclosure: the author holds commercial interests in automated verification of commercial energy billing.

1 Introduction

A companion paper (Gill 2026) asks when checking an economic claim becomes worth its cost, and derives the frontier that answer traces in claim value and claim volume. This paper asks the question that frontier makes urgent: *what happens to an economy whose institutions were designed around checking not being worth it?*

The premise is a dual cost collapse. Machine intelligence lowers the marginal cost c of checking a claim—parsing, retrieval, comparison, recomputation. Less noticed but institutionally more consequential, it lowers the fixed cost K of *constructing* the checker: the bespoke parser, the custom reconciliation pipeline, the single-counterparty audit tool that once required a bespoke engineering project and can increasingly be produced through model-assisted configuration at a small fraction of that cost. High K is why verification historically lived inside institutions built to amortize it—audit firms, certification bodies, ratings agencies, inspection regimes—and why relationships too small, too idiosyncratic, or too bespoke to amortize anything ran instead on brands, reputations, and tolerated opacity. When K collapses, the economic logic that centralized verification collapses with it.

The paper models this directly. Trading relationships differ in transaction volume, in the gains from misconduct, and in how standardized their claims are. Each relationship chooses the cheapest available credibility technology: remain opaque and bear the informational penalty; pay the reputation premium that makes a seller's promise self-enforcing (Klein and Leffler 1981); hire a certifier who amortizes the fixed cost of verification across many clients; or verify directly by machine. The pre-shock equilibrium assignment reproduces the institutional landscape we observe: standardized, high-volume claims are certified and audited; valuable bespoke relationships are governed by brands and reputations; the small and idiosyncratic remain opaque. The paper then lowers c and K and solves for where the assignment moves.

Five results organize the answer.

Machine verification, throughout, performs two tasks that were both historically expensive, and the paper carries the distinction from theory to measurement. *Conformity* verification asks whether a realized claim complies with an existing obligation: bill against contract, delivery against order, settlement against policy. *Optimality* verification asks whether the existing arrangement is the right one for the relationship—contract against counterfactual—where nothing need be misstated, nothing is illegal, and the loss survives purely because constructing and acting on the comparison was historically expensive analysis. The institutions this paper studies were adapted to the scarcity of both.

Result 1 — verification de-institutionalizes. A certifier's economic advantage over in-house checking is fixed-cost amortization: spreading K across M clients. That advantage is proportional to K and vanishes as K collapses, so any positive intermediation margin eventually makes direct verification strictly cheaper. The migration has a predicted order: it enters the bespoke, reputation-governed relationships first, extends to arbitrarily low volumes as construction costs collapse—territory no institution previously served—and reaches the certifiers' standardized core last. Endogenizing the certifier's clientele sharpens the prediction into the paper's least obvious result: adoption is strategically complementary, so certification markets under a falling construction cost do not shrink smoothly—they tip, discontinuously and with hysteresis. Intermediaries do not simply disappear: where credibility requires recognized standing and someone to bear liability—where the evidence stage of the pipeline depends on *who* attests—the intermediary survives on that margin, and intermediation rents migrate from amortization to recognition and liability.

Result 2 — brands migrate rather than die. Decompose brand value into a bonding component (the capitalized premium that makes cheating irrational when buyers cannot check), a search component, and a taste-and-status component. Machine verification attacks only the first: where quality becomes cheaply and credibly verifiable, the sustainable assurance premium falls toward the resource cost of verification, and high-quality entrants can substitute evidence for accumulated reputation at a minimum scale that falls with K. The other components are untouched, sorting improves because only truthful quality benefits from verification, and there is a genuine counterforce: evidence overload raises the value of brand-as-heuristic. The prediction is a migration of brand capital toward intrinsically unverifiable attributes—not a collapse of brands.

Result 3 — optimality rents compress through repricing and exit. A contract can be perfectly performed and still wrong for the relationship. Establishing the better arrangement—repricing the customer's realized data against the obtainable menu, or against contracts assemblable from already-priced components—was itself expensive analysis, and that cost sustained inertia rents no court could touch, because nothing was breached. When the counterfactual becomes cheap, the minimum mismatch worth acting on falls with it. Enforcement needs no adjudication: a credible alternative disciplines the incumbent through exit and, mostly, through the threat of it—repricing before switching, deterrence's twin. Compression stops at switching friction, not at zero; and because an optimality loss is a flow that no restitution reaches backward, the value of checking lies in its frequency—the layer's limit is continuous verification.

Result 4 — the rebuttal asymmetry. The technology that lowers the cost of checking lowers the cost of *accusing*. The cost of proving innocence—human, legal, procedural—falls much more slowly. In a simple challenge-and-settlement game, nuisance challenge against honest parties is profitable exactly when accusation cost is below the target's rebuttal cost, so an asymmetric collapse expands the nuisance region even as it improves true deterrence. Cheap verification is welfare-improving where rebuttal costs fall symmetrically and can be welfare-destroying where they do not; fee-shifting and pleading standards are the instruments that restore symmetry. The first automated private enforcement wave and the judicial pushback it triggered illustrate both halves of the result.

Result 5 — opacity relocates. Sellers allocate effort between verifiable and unverifiable attributes. Verification disciplines the first and, when efforts are substitutes, diverts distortion toward the second (Holmström and Milgrom 1991). The transition does not eliminate opacity; it moves opacity toward judgment, ambiguity, and whatever the machines cannot yet measure—the honest boundary on Results 1–3.

The contribution is deliberately positioned. That reputation substitutes for verification is Klein and Leffler (1981); that certifiers are information intermediaries with market power is Lizzeri (1999); that monitoring technology re-selects governance is documented in trucking by Baker and Hubbard (2004); that nuisance suits extract settlements from innocent defendants is Rosenberg and Shavell (1985) and Bebchuk (1988); that verification can be modeled as a technology is Townsend (1979) through Ball and Kattwinkel (2025); that centralizing verification in one delegated monitor economizes on duplicated checking is Diamond (1984), the classical ancestor of the certifier here, with Biglaiser (1993) supplying the middleman-as-expert version; and the critical-mass and tipping mechanics of Proposition 2 descend from the network-effects lineage of Katz and Shapiro (1985) and Farrell and Saloner (1985), applied to a shared cost-recovery base rather than a demand-side externality. Closest in time and subject, Zhang (2026) models the make-or-buy of checking AI output, where the verifier's diligence inherits the difficulty of the verified task and liability capacity bounds the verification market: that paper supplies deep structure for why the recognition margin μ exists, where this paper takes μ in reduced form and studies the institutional assignment around it. To my knowledge, however, prior work has not modeled the joint collapse of the marginal and construction costs of verification as a shock to the *menu* of credibility institutions, derived which institutions lose their function in which order, and traced the consequences for brand value, intermediation rents, and the welfare economics of accusation. That combined architecture is the contribution; every ingredient's ancestry is cited where it is used. Stated as one sentence: treat machine intelligence as a shock not only to the marginal cost of verification but to the fixed cost of constructing bespoke verification systems, and study how that shock re-selects the institutions society previously used to manufacture trust when direct verification was uneconomic—where verification covers both checking realized claims against agreed rules and checking agreed arrangements against attainable counterfactuals. The paper is offered in the spirit that sentence implies—an organizing framework and research agenda whose models are deliberately transparent and whose weight rests on the predictions they generate—rather than as a claim to deep theorems; where a result follows quickly from its assumptions, the text says so.

The paper proceeds as follows. Section 2 positions the paper in the literature. Section 3 states the verification shock, importing the frontier from the companion paper in compressed form. Section 4 builds the model of credibility institutions and characterizes the pre-shock assignment. Sections 5 through 9 derive the five results. Section 10 states the empirical program: an implementation in commercial energy contracting now in progress, operationalizing both verification layers, with the construction-cost series $K(t)$ as the novel measured object and the EU Digital Product Passport as a prospective quasi-experimental setting for Result 2. Section 11 concludes.

2 Related Literature

Reputation as a bond. Klein and Leffler (1981) show that price premia can assure unobservable quality: a cheating seller forfeits the capitalized premium, so the premium is the return on sunk brand capital. The mechanism inherits Klein–Leffler's information assumption, stated here explicitly: cheating is detected by buyers with certainty one period after it occurs, without any verification technology—experience reveals it. Everything the shock changes runs through making detection contemporaneous and pre-contractual rather than experiential; the premium's equilibrium production (entry dissipating the quasi-rent into sunk brand investment) is likewise Klein–Leffler's, imported not re-derived. Shapiro (1983) develops the premium as the return to reputation investment. This paper treats that mechanism as one technology on a menu and asks what happens to its market share—and to the premia it sustains—when a substitute technology's costs collapse.

Certification and intermediation. Lizzeri (1999) shows certification intermediaries can strategically coarsen information and capture surplus; Dranove and Jin (2010) survey quality disclosure and certification; Pérignon et al. (2019) describe operating third-party certification of computational reproducibility. The economics of why certifiers exist at all—fixed-cost amortization across clients—is the margin this paper endogenizes and then shocks.

Monitoring technology and governance. Baker and Hubbard (2004) provide the canonical evidence that a monitoring technology re-selects institutions: on-board computers changed contractibility in trucking and with it asset ownership. Grossman and Hart (1986) and Hart and Moore (1990) supply the deeper logic—what cannot be verified must be governed by ownership—so a technology that enlarges the verifiable set re-draws that boundary. Holmström and Milgrom (1991) supply the multitask logic behind Result 5.

Nuisance claims. Rosenberg and Shavell (1985) and Bebchuk (1988) model suits brought for settlement value against defendants for whom defense costs exceed demands. Result 4 is the machine-age version: the asymmetry between collapsing accusation costs and sticky rebuttal costs is new, the extraction logic is theirs. The optimality mechanism of Result 3 likewise claims no new ancestors: search and switching frictions, inertia, and shrouding are Gabaix and Laibson (2006) and, in electricity, Hortaçsu, Madanizadeh, and Puller (2017); what is new is the falling cost of *constructing the individualized counterfactual* as the driving technology, and the institutional consequences of that fall.

Verification theory and the AI shock. Costly state verification begins with Townsend (1979); mechanism design with costly or probabilistic verification runs through Ben-Porath, Dekel, and Lipman (2014), Halac and Yared (2020), and Ball and Kattwinkel (2025). Catalini, Hui, and Wu (2026) model human verification as the automation era's bottleneck; Merane and Stremitzer (2026) document the first credibly identified field study of automated private enforcement; Jin, Sokol, and Wagman (2026) show AI-augmented monitoring below institutional capacity produces workload without harm reduction; Fedyk et al. (2022), Law and Shen (2025), and Commerford et al. (2022) document AI reorganizing audit work and the human discounting of machine evidence. The companion paper (Gill 2026) develops the verification frontier, pipeline, and deterrence results this paper builds on; the present paper takes the institutional question those results make tractable.

3 The Verification Shock

The technological environment is imported from the companion paper in compressed form; full development, including the multi-stage pipeline and its comparative statics, is in Gill (2026).

A verification technology for a claim class is a fixed construction cost K, a marginal cost c per claim checked, and an end-to-end reliability $q \in (0, 1]$: the probability that a checked misstatement produces an enforceable consequence F for the misstating party. Reliability is a pipeline—detection, substantiation, evidentiary acceptance, enforcement—and if its institutional stages fail, $q \approx 0$ and no fall in c or K matters: verification without consequence is a dashboard (Gill 2026, Corollary 1). Throughout, the results below are conditional on the pipeline clearing; where it does not, the pre-shock institutions persist by default.

The shock is a joint decline in (c, K), driven respectively by machine intelligence performing verification and by machine intelligence *constructing* verifiers—code generation turning bespoke parsers, reconciliation pipelines, and single-counterparty audit tools from engineering projects into configuration. The paper does not model the engineering process; it conditions on the cost movement and studies the institutional response. Behavior within each institution is endogenous throughout: deterrence conditions, not exogenous compliance rates, determine outcomes.

4 A Model of Credibility Institutions

4.1 Environment

A continuum of trading relationships. A relationship generates N transactions per period. In each transaction the seller either performs or shirks: shirking yields the seller a private gain $g > 0$ and imposes a loss $\ell > g$ on the buyer, so shirking destroys $\ell - g > 0$ of surplus. Sellers and buyers are risk-neutral; the common discount factor is $\delta \in (0, 1)$; prices are set competitively, so buyers pay expected value and the cost of credibility is ultimately borne within the relationship. Relationships differ in (N, g, β), where $\beta \in [0, 1]$ is the *bespokeness* of the relationship's claims: the share of its verification architecture that cannot be shared with other relationships. A statutory financial statement has β near zero; a customized supply contract with idiosyncratic pricing formulas has β near one. A units convention governs every cost comparison in the paper: K denotes the *per-period equivalent* of owning a verifier for the relationship's claim class—annuitized construction plus maintenance, where maintenance is substantial in practice because tariffs, taxes, formats, and regulation drift—so K, Nc, the reputation premium, and certification fees are all flows and are added legitimately. This matters twice below: comparisons across technologies are level comparisons of flows, and the certifier dynamics of Section 5 require exactly the recurring component of K—a certifier's historical construction outlay is sunk, but its maintenance and cost-recovery obligation is not, and it is the recurring part that a shrinking clientele must carry.

Without credibility, rational buyers price in shirking. Trade then either proceeds at lemons terms or collapses; either way the relationship forgoes a surplus that we summarize as an informational penalty Δ per transaction, the value of credible performance. A relationship adopts a credibility technology when its per-period cost is below $N\Delta$, and adopts the cheapest one.

4.2 The menu

Reputation (the Klein–Leffler bond). The seller charges a premium P per period above cost. Compliance is self-enforcing when the capitalized premium exceeds the one-period gain from milking the reputation:

$$\frac{\delta}{1-\delta}P \;\geq\; Ng \qquad \Longleftrightarrow \qquad P \;\geq\; P_{KL} \equiv \frac{(1-\delta)\,Ng}{\delta}. \tag{1}$$

The premium is a transfer, but it requires sunk brand capital—the below-cost introductory period, advertising, the years of track record—whose return it constitutes; entrants must sink that capital before the bond binds. If direct verification of intensity q operates alongside reputation, the per-transaction gain from shirking net of expected consequence falls to $(g-qF)_+$, and the required premium falls to

$$P^*(q) \;=\; \frac{(1-\delta)\,N\,(g-qF)_+}{\delta}. \tag{2}$$

Direct machine verification. The relationship (in practice, the buyer or a seller-installed, provenance-carrying pipeline) verifies its own claims: per-period cost $K+Nc$, deterrence condition $qF \geq g$ for full discipline; below that threshold verification is partial and the residual premium (2) applies.

Certification. A certifier builds the verification architecture once and serves M clients whose claims share the standardizable component: the certifier amortizes $(1-\beta)K$ across clients but each client's bespoke share βK cannot be shared. Under competitive certification the fee per client is

$$f_C \;=\; \frac{(1-\beta)K}{M} \;+\; \beta K \;+\; Nc \;+\; \mu, \tag{3}$$

where $\mu \geq 0$ is the intermediation margin: the price of recognized standing, liability-bearing, and the credibility that makes the certifier's attestation acceptable evidence. The margin is not decoration; it is the pipeline's evidence stage priced. Certification's deterrence operates like direct verification's.

Opacity. Cost $N\Delta$ per period: the default when every technology above costs more.

4.3 The pre-shock assignment

Lemma 1 (The institutional landscape: an organizing taxonomy)**.** *Fix (μ, M, δ, c, F) common across relationships, let sanctions satisfy the feasibility condition $qF \geq g$ where verification is used, and compare technologies pairwise at the costs (1)–(4). With high construction cost K, the cost-minimizing assignment partitions relationships as follows: (i) standardized,* *high-volume relationships (low β, high N) certify, because amortization makes f_C the cheapest credible technology; (ii) bespoke, valuable relationships (high β, high g) use reputation, because their bespoke share of K is unamortizable and P_{KL}, which scales with Ng but requires no architecture, undercuts $K + Nc$ at high K; (iii) small and idiosyncratic relationships (low $N\Delta$) remain opaque. The partition is offered as an organizing taxonomy under these sufficient conditions: it is the pairwise cost comparison made exhaustive, not a full characterization over the parameter space, and heterogeneity in the fixed parameters can redraw local boundaries.*

The lemma is verified by comparing the three per-period costs and is stated because its content is the observed world: audited financial statements and graded commodities on one side; branded suppliers, relationship banking, and trusted contractors on the other; and a long tail of unverified claims tolerated because nothing cheaper than tolerance existed. The model's discipline is that the same comparison, re-run at post-shock costs, must predict where each region moves.

4.4 Private choice versus social cost

The institutional comparison above is private: each relationship minimizes the price of credibility, and to the buyer the reputation premium is a price like any other. Welfare accounting differs, because the premium is a transfer while $K + Nc$ is a resource cost. Reputation's social cost is not the premium but what sustaining it burns—sunk brand-building expenditure, delayed and deterred entry, and the competitive barriers that protect the quasi-rent—while verification's social cost is the engineering it consumes. The private substitution from reputation to verification is therefore not automatically efficient in either direction: it saves brand-building waste and entry barriers while spending real verification resources, and when the displaced premium was mostly transfer over already-sunk capital, privately rational verification spending can be socially excessive—the institutional-choice face of the private–social divergence derived in the companion paper (Gill 2026). The propositions below are statements about private adoption and market structure; welfare claims are made only where this distinction is carried explicitly.

5 Result 1: Verification De-Institutionalizes

Compare certification with direct verification for the same relationship. Both pay Nc; both bear the bespoke share βK in construction (the certifier passes it through). Certification's advantage is amortization of the sharable component; its disadvantage is the margin:

$$f_C - (K + Nc) \;=\; -(1-\beta)\,K\Big(1 - \tfrac{1}{M}\Big) \;+\; \mu. \tag{4}$$

Proposition 1 (De-institutionalization)**.** *(i) Certification dominates direct verification if and only if the amortization saving exceeds the margin,* $(1-\beta)K(1-1/M) > \mu$*. The saving is proportional to* K*; hence there exists* $\bar{K} = \mu M/[(1-\beta)(M-1)]$ *such that for* $K < \bar{K}$ *direct verification strictly dominates for any* $\mu > 0$*: as construction costs collapse, the fixed-cost rationale for centralized verification vanishes. (ii) Reputation loses to direct verification when* $K + Nc \le P_{KL}$*, i.e. when* $g \ge \frac{\delta}{1-\delta}\Big(c + K/N\Big)$*; as* $K \to 0$ *the condition becomes independent of* N*, so verification reaches relationships of arbitrarily small volume. (iii) The migration is ordered: it begins in the bespoke, high-stakes region (*β *and* g *high) that certifiers could never serve and reputation governed at the highest premia, proceeds down the volume scale as* K *falls—the switching gain* $P_{KL} - (K + Nc)$ *turns positive first for high-*g*, high-*N *cells and reaches arbitrarily small* N *in the limit—and arrives at the certifiers' standardized low-*β *core last, where amortization retains value longest; the ordering is stated under fixed* (μ, M) *and common* (c, F, δ) *across regions, as sufficient conditions, and heterogeneity in those parameters can reorder the transition locally. (iv) Intermediaries survive on the recognition margin: relationships for which acceptable evidence requires an attester with standing and liability continue to pay* μ*, and intermediation rents migrate from amortization to recognition.*

The menu is mutually exclusive by construction; real governance mixes these technologies, and formal and relational governance are often complements rather than substitutes (Poppo and Zenger 2002)—the exclusive assignment is a simplifying restriction whose predictions concern the dominant technology in a relationship, not its full portfolio. The proof is the comparison of (1)–(4) and is completed in the appendix. Three readings. First, the historical, stated at the strength the model supports: one historically important source of comparative advantage for verification institutions—audit firms, certification bodies, ratings agencies, inspection regimes—has been fixed-cost amortization. The shock removes that particular advantage and forces these institutions to compete increasingly on what μ bundles. Three components of that bundle should be kept distinct even in reduced form: resource costs of recognized attestation (professional labor, insurance, regulatory capital), expected liability, and any market-power markup—the first two survive competition because they are costs, while the third requires entry barriers; and recognition properly operates on the benefit side as well, raising the evidence's pipeline reliability (q_E) rather than only adding to price. The prediction is not their disappearance but the re-pricing of their function, with survival contingent on the margins other than amortization that they actually supply. Second, the extensive margin: part (ii) says the collapse of K carries verification into the region where no institution could previously afford it—one-off claims, idiosyncratic counterparties, the long bespoke tail—which is where reputation premia and tolerated opacity were largest. Third, the ordering in (iii), which is two-dimensional because the shock is: a fall in the *marginal* cost c converts claims where volume is largest—standardized, high-N classes whose verifiers already exist—while a fall in the *construction* cost K converts claims that previously could not carry a verifier at all: bespoke formats, idiosyncratic counterparties, small claim classes. Which territory moves first in any market is a statement about which component fell there. The companion paper's field episode and formula-governed billing are c-shock conversions; the bespoke tail opening is the distinctive signature of the K-shock; and intermediaries whose function is fixed-cost amortization lose their bespoke fringes to the K-shock before their standardized core, which converts only when in-house ownership beats a shared fee outright. One reordering is predictable rather than local: the bespoke-first path presumes private-ordering enforcement—settlement, repricing, exit—which is the operative mode in commercial billing and most B2B disputes. Where remedies run through courts, evidentiary quality is typically lowest precisely where bespokeness is highest, and the ordering can invert: standardized claims migrate first because only they survive the evidentiary stage at scale. A re-concentration qualification belongs here rather than in a footnote: if provenance, proprietary data access, or liability capacity are themselves scale goods, the verification layer that replaces the amortizers can be more concentrated than what it displaced. The de-institutionalization of the checking function and the market structure of the checking industry are separate questions, and the model answers only the first.

5.1 Endogenous clientele: certifier tipping

Treating the clientele M as fixed understates the transition's sharpness, because M is an equilibrium object with a feedback: the shared component of the fee falls in clientele, and clientele rises as fees fall. Let a mass $\bar{M}$ of potential clients differ in their next-best credibility cost, and let $D(f)$ denote the mass for whom certification at fee f is cheapest—nonincreasing in f, with the fee schedule $f_C(M)$ of equation (3) defined on $M \geq 1$ and understood as the recurring cost-recovery obligation of the units convention of Section 3. Equilibrium clientele solves $M = D\Big(f_C(M)\Big)$, a fixed point of a nondecreasing map: certification adoption is strategically complementary. What follows from complementarity alone is existence and a lattice of equilibria; multiplicity, discontinuous collapse, and hysteresis are *not* implied by it and require conditions on demand and adjustment, which the proposition states and a worked example in the appendix exhibits in closed form. The direction of the comparative static itself is conditional: for clients whose binding outside option is stand-alone verification, a falling K improves the outside option faster than it lowers the fee; for clients whose binding alternative is reputation or opacity, a falling K lowers the fee against an unchanged alternative and can *raise* certification demand. Tipping is a prediction about clienteles of the first kind.

Proposition 2 (Certifier tipping, conditional)**.** *Suppose clients' binding outside option is stand-alone verification, demand D is continuous, and adoption follows myopic best-response dynamics. Then: (i) adoption is strategically complementary and equilibria form a lattice; where the equilibrium map crosses the diagonal more than once, the largest fixed point $M^H(K)$ is locally stable under a slope condition ($\Phi' < 1$ at M^H). (ii) A fall in K lowers each such client's outside option one-for-one while lowering the fee by only $\beta + (1-\beta)/M < 1$ per unit, so demand shifts down pointwise and $M^H(K)$ declines; if the shifted map becomes tangent to the diagonal at some $K_{crit} > 0$, the high-clientele equilibrium vanishes by a saddle-node and adoption collapses discontinuously—marginal exit raising the per-client cost-recovery burden on inframarginal clients and inducing further exit. (iii) Where collapse occurs, it is hysteretic under the same dynamics: re-entry at small clientele faces the high small-M fee, so restoring K marginally above K_{crit} does not restore the market. (iv) Whether a niche survives is a parameter statement: in the worked example, a market whose certification fee premium over stand-alone costs is below the top of clients' distribution of outsourcing value shrinks smoothly to a positive niche as $K \to 0$, while a market priced above it collapses to zero at an interior K_{crit} with closed-form threshold—tipping is the signature of certifiers whose fee exceeds their non-amortization value to clients.*

This is the paper's sharpest structural prediction, stated at the strength the mathematics supports: the discontinuity is not implied by complementarity alone, but it emerges—with an explicit interior threshold—in the fully solved example of Appendix A.2, under interpretable conditions rather than an assumed cliff. Certification markets under a construction-cost collapse should exhibit long apparent stability followed by rapid unraveling, segment by segment—and the segments that tip first are those whose members' outside options improve fastest, which by Proposition 1(iii) are the high-bespokeness fringes of each certifier's book. The empirical signature follows equation (3): per-client fees *escalate* as the book shrinks—the shared cost-recovery burden concentrating on remaining clients—so the observable path is rising fees, shrinking client counts, and then abrupt exit, not gradual fee erosion.

Remark 1 (Availability, not uptake)*.* The assignment characterizes which technology each relationship would choose with the full menu in hand. Direct verification, however, requires buyer-side deployment, and adoption is heterogeneous in ways the model does not carry—firm size, tooling, attention. The transition equilibrium is therefore two-tier: verified relationships coexist with opaque ones inside the same market, sellers price-discriminate between the tiers, and the inertia rents of the optimality layer concentrate on non-adopters, in the manner of shrouded-attribute equilibria (Gabaix and Laibson 2006). The propositions are statements about availability—where the frontier moves—and the lag between availability and uptake is itself a measurable object: the implementation of Section 10 observes it directly as the gap between achievable and realized savings across the client base.

6 Result 2: Brands Migrate Rather than Die

Decompose the value of an established brand to its owner as

$$V \ = \ V_{bond} \ + \ V_{search} \ + \ V_{taste}, \tag{5}$$

where V_{bond} is the capitalized assurance premium—the Klein–Leffler component—V_{search} the value of reducing buyers' search and attention costs, and V_{taste} the value of identity, status, and preference that would survive full information. Only the first is a substitute for verification. The decomposition is definitional—verification affects the other components only through the channels modeled—so the proposition's content lies not in the decomposition but in the cross-sectional and entry predictions it generates; and the sorting result below holds under the maintained assumption that verification is truthful, an assumption governed by the recognition margin of Section 5 and the pipeline of the companion paper.

Proposition 3 (Brand migration)**.** *(i) The sustainable assurance premium in a relationship with verification intensity q is $P^*(q)$ of equation (2): weakly decreasing in q, strictly where $qF < g$, and zero once $qF \geq g$. Where the shock makes direct verification cheaper than the pre-shock premium ($K + Nc < P_{KL}$), competition from verified sellers compresses the premium toward the resource cost of* credibly attested *verification—$K + Nc$ plus the recognition margin μ wherever standing is required for evidence to move a counterparty—and V_{bond} erodes not to zero but to the price of that recognition. The bonding function does not vanish; it migrates to whoever bears liability for the attestation, which is where Proposition 1(iv) sends surviving intermediary rents. Zero is the limit only where self-produced evidence is credible without standing. (ii) The minimum scale at which a high-quality entrant can substitute evidence for accumulated reputation is $N_E = K/(\pi - c)$—abstracting from reliability, recognition, buyer adoption, and non-verification entry costs, each of which raises the threshold—, where π is the per-transaction premium the incumbent's assurance commanded; N_E falls with K, so the construction-cost collapse specifically lowers the scale of credible entry. (iii) Verification improves sorting: a low-quality entrant gains nothing from truthful verification, so the entrants* *the technology helps are disproportionately those whose quality is real. (iv)* V_{search} *and* V_{taste} *are unaffected by q, and* V_{search} *can rise in the volume of verified evidence: when buyers face many verified alternatives, the brand's value as a default that economizes on attention increases. Total brand value therefore moves ambiguously; its composition migrates toward the unverifiable components.*

The economic statement is the one worth remembering: machine verification strips brands of their role as generalized bonds for verifiable quality, and brand capital migrates toward attributes that remain intrinsically difficult to verify. The result is emphatically not that brands collapse. It is a re-selection *within* the brand: assurance-heavy premia—safety, provenance, specification compliance, durability, billing accuracy—are exposed; taste, identity, status, and attention economics are not, and the last of these can strengthen. The prediction is cross-sectional and testable: premium erosion should be largest in categories where the premium was mostly assurance and the attributes are objectively verifiable, smallest where brand value is mostly identity—and the entry effect should appear as high-quality small entrants gaining share in assurance-heavy categories as K falls, marketing evidence where incumbents market names.

One accounting clarification closes a loose end: equation (2) sends the *necessary* assurance premium to zero as verification substitutes for bonding, while the discussion says the bonding function migrates to the attester—both hold, because the buyer's payment does not vanish but changes payee and character, from a quasi-rent paid to the brand into a fee paid for attestation; what reaches zero is the rent, not the expenditure. The mechanism has, moreover, already run once at partial scale. Review platforms are a partial verification technology—crowd-sourced, noisy, limited to experienced attributes—and their diffusion is a completed natural experiment on this proposition. Luca (2016) finds that a one-star Yelp increase raises independent-restaurant revenue five to nine percent while chain-affiliated restaurants are unaffected, and that chain market share declined as review penetration rose: reviews substituted for the chain's bonding function exactly where that function was the binding constraint on entry. Hollenbeck (2018) finds the hotel chain-affiliation premium fell by more than half between 2000 and 2015, that the decline is attributable to online reputation mechanisms, and that it concentrates among low-quality and small-market properties—the assurance-heavy cells where the model locates the largest V_{bond}—while the premium among high-end properties, where taste and status dominate, largely persists, and revenue's correlation with brand-wide reputation is replaced by correlation with the individual property's own record. Every element of the proposition appears: compression concentrated where bonding was the premium's substance, persistence of the taste and search residual, entry by small independents, and the migration of credibility from the umbrella name to the unit-level evidence trail. Reviews are crowd-sourced reputation aggregation rather than machine verification, and Hollenbeck states his evidence with corresponding caution, so the episode is the closest historical analogue rather than a completed test; machine verification is a shock of the same sign with higher evidentiary quality and a far wider reach of attributes, which is why Result 2 is offered as disciplined extrapolation rather than speculation.

7 Result 3: Optimality Rents Compress through Repricing and Exit

The results so far concern conformity: claims checked against obligations. The second task defined in the introduction has its own economics, and it is where an important class of fully legal rents lives, precisely because no obligation is violated; whether it is the larger class is an empirical question Section 10 is designed to answer. Let P_0 be a relationship's current contract cost over the relevant horizon, and define three counterfactuals of increasing ambition. P_M^* is the cost of the best *genuinely obtainable* equivalent contract—matched on term, volume profile, risk allocation, service, and attributes—so $G_M = P_0 - P_M^*$ is the market-optimality gap: value available from moving to an existing alternative. P_R^* is the cost of the best contract *assemblable from already-priced components*—existing wholesale products, published tariff building blocks, standard risk terms—without material additional true cost, whether or not any supplier currently offers it; $G_R = P_0 - P_R^*$ is the recombination gap, and it repairs the market gap's blind spot: where all incumbents coordinate on the same ill-fitting contract structures, G_M is small while the true mismatch is large, and G_R measures against the *contestable* menu rather than the offered one. P_E^* is the lowest economically feasible cost including all legitimate costs and a normal return; $G_E = P_0 - P_E^*$ is a welfare benchmark, and this paper rests no result on it. Since the obtainable set is contained in the recombination-feasible set, which is contained in the economically feasible set, $P_E^* \leq P_R^* \leq P_M^*$ and $G_M \leq G_R \leq G_E$.

Establishing and validating an actionable counterfactual is the optimality layer's verification task: the machine may *generate* the alternative by optimization, but what carries economic weight is verifying that the alternative is genuinely equivalent, feasible, and actionable. That task has the companion paper's cost structure: a fixed cost K_O of building the counterfactual engine for a contract-type class and a marginal cost c_O of running it against one relationship's realized data, summing to an all-in cost C_O per evaluation; and acting on the result carries a friction s—search, renegotiation, notice periods, exit fees, implementation. Historically C_O was consulting, brokerage, and benchmarking labor: a EUR 300 annual mismatch rationally survived a EUR 1,000 analysis. The comparison never happened, so the rent never appeared as one.

The enforcement mapping is the layer's distinctive economics. A conformity claim runs the pipeline: detection, substantiation, evidence, enforcement—its floor is legal. An optimality claim runs: counterfactual, credible alternative, repricing or exit—its floor is frictional. The customer proves no wrongdoing; the customer says *an equivalent contract exists at lower cost; match it or I leave.* For G_M the outside option exists; for G_R it is at most one procurement away—a machine-generated specification invites some competing supplier to create the contract and win the account. That step is not automatic: entry costs, capacity and risk limits, credit constraints, minimum portfolio sizes, regulatory restrictions, market power, or strategic refusal to quote can all leave a creatable contract uncreated, which is why G_R's enforceability sits strictly between G_M's and G_E's rather than matching the former. The fill-or-no-fill protocol of Section 10 is the empirical resolution: each specification sent to procurement tests, case by case, whether *could exist* becomes an outside option, and the no-fill outcomes map exactly where these frictions bind.

Proposition 4 (Optimality frontier, repricing before switching)**.** *(i) A relationship acts on an optimality gap G iff $G > C_O + s$; the action threshold $G^* = C_O + s$ falls one-for-one with the counterfactual cost, so a collapse in C_O makes progressively smaller mismatches actionable, and with C_O composed of (K_O, c_O) the threshold inherits the companion frontier's comparative statics, claim class by claim class. (ii) An incumbent facing a customer with a credible counterfactual, whose own cost of serving at the matched terms does not exceed the alternative, weakly prefers repricing to losing the relationship; realized switching is therefore selected on incumbent disadvantage, and the equilibrium response to systematic optimality verification is repricing first and switching second—gaps can compress without switching volumes rising in proportion. (iii) Enforceability orders with the counterfactual ladder: G_M is disciplined by exit to an existing offer, G_R by procurement against a generated specification, G_E by neither, immediately. (iv) Compression has a floor: as $C_O \to 0$, actionable gaps compress toward the action friction s, which does not fall with the verification technology; optimality rents are eliminated only where s is. (v) The two layers differ in recoverability: a conformity violation is a stock—restitution reaches backward—while an optimality mismatch is a flow no remedy recovers, so expected optimality loss is proportional to detection delay, the value of verification is increasing in its frequency, and continuous rather than episodic checking is the limit policy as $C_O \to 0$.*

Every part follows almost immediately from the definitions and the stated participation comparisons, and the text says so: the proposition's contribution is the mapping—

counterfactual ladder to enforcement mechanism to compression floor—not the mathematics. Its content is in what it organizes and predicts. Three remarks carry the institutional weight.

First, part (ii) is the deterrence result of the conformity layer, transposed. Cheap conformity verification changes counterparty behavior before violations are challenged (the companion's threat channel); cheap optimality verification changes incumbent pricing before customers switch. Both predict that the visible activity metric—recoveries there, switching here—understates the mechanism, and both imply the same evaluation warning: a successful optimality-verification system may show modest switching precisely because it works.

Remark 2 (Incidence: who captures the identified gap). Part (ii) says the incumbent reprices; it does not say by how much. When a customer holding a credible counterfactual renegotiates rather than exits, the concession is a bargained share of the identified gap: with bargaining weights $(\alpha, 1-\alpha)$ over the surplus at stake, the customer recovers αG net of frictions and the incumbent retains the rest, so verification's *incidence*—the pass-through of identified gaps to customers—is an equilibrium object between zero and one, not an assumption. It should vary with the customer's outside options (G_M versus only G_R), with whether the counterfactual is disclosed or merely held, and with who runs the audit, since an intermediary may capture part of α as its own margin. The pass-through share is, to my knowledge, unmeasured anywhere, and Section 10 logs it as a primary outcome: the ratio of realized concession to identified gap, client by client.

The compression-floor parallel completes the correspondence, and unifies the paper: brand premia compress to the cost of credibly attested verification including μ (Result 2); conformity rents compress to what the evidence and enforcement stages can carry (the companion's pipeline); optimality rents compress to the switching friction s. In each case machine intelligence collapses the computation, and an institution sets the floor—which is also where policy lives, since s is partly regulatory (notice periods, exit fees, portability) while C_O is not a policy variable at all.

Second, the intermediaries of this layer are brokers, benchmarking consultants, and procurement advisors, and Result 1's logic transposes to them without modification: their margin has been, in substantial part, a verification wage—amortizing across a client book the cost of knowing the market, modeling suitability, and computing counterfactuals. The falling (K_O, c_O) attacks exactly that component, their books are shared-cost bases subject to the tipping mechanics of Proposition 2, and the migration prediction is the certifier prediction re-instantiated: surviving broker value moves to negotiation, execution, aggregation, supplier access, trusted representation, and liability—the margins machine counterfactuals do not supply. No disappearance claim is made, and none follows.

Third, suppliers and contract form. Systematic optimality verification compresses the portion of margins supported by costly comparison—inertia premia, mismatch tolerated because discovering it was expensive—without implying perfect competition or zero margins: differentiation, risk-bearing, and service margins are not this object. The contract itself becomes subject to a standing question the conformity layer never asks: not *was this formula applied correctly* but *is this still the right formula for this relationship's realized data*—extending institutional re-selection from contract enforcement to contract design and renegotiation, the Grossman–Hart boundary question in operating form. And the seller's counter-move is Section 9's: complexity and differentiation on dimensions the counterfactual engine cannot compare, which is where the relocation result and this one meet. One honesty note on scope: the repricing logic is partial equilibrium—one informed customer against one incumbent. If verification becomes universal, suppliers' best response may not be client-by-client concessions at all but menu simplification and commitments to uniform pricing, ending discrimination rather than negotiating it away—a different and arguably better equilibrium, in the spirit of the obfuscation literature run in reverse; which response prevails is an open question the implementation can observe but this model does not settle.

8 Result 4: The Rebuttal Asymmetry

The same technology that checks claims cheaply generates *challenges* cheaply. Let a challenger issue a challenge at cost a (detection plus claim generation, collapsing with the shock). A challenged party who is actually compliant can establish innocence only at rebuttal cost r—document assembly, legal process, management time—which is human, procedural, and falls far more slowly than a; or the party can settle at demand σ. A compliant target pays any demand $\sigma \leq r$ rather than rebut.

Proposition 5 (Nuisance expansion under asymmetric collapse)**.** *(i) Challenging a compliant party is profitable if and only if $a < \sigma \leq r$: the nuisance region is $a < r$, and its measure expands when accusation costs fall while rebuttal costs do not. (ii) Let challengers bear a fixed participation cost $\phi > 0$ (filing, standing, attention), so challenging a compliant party is viable only if $r - a > \phi$. A symmetric proportional collapse $a' = \lambda a$, $r' = \lambda r$ with $\lambda < 1$ preserves the sign condition $a < r$ but contracts every viable gap to $\lambda(r - a)$, so the set of claim types supporting profitable nuisance strictly shrinks against the fixed ϕ—symmetric collapse shrinks nuisance in stakes and, with any participation friction, in extent, while retaining all true-deterrence gains. (iii) Fee-shifting that charges a failed challenger the target's rebuttal cost makes challenging a compliant party who rebuts strictly unprofitable ($-a - r < 0$), so nuisance survives only where targets settle regardless—where fee recovery is slow, partial, or uncertain; pleading and evidence standards operate on the same margin by raising the effective* *accusation cost of low-merit challenges. (iv) The welfare effect of an accusation-cost collapse is therefore signed by its symmetry: positive where the enabled marginal challenges are true and rebuttal is cheap, ambiguous to negative where compliant parties are extortable at scale.*

The appendix states the game and completes the argument; the mechanism is Rosenberg and Shavell (1985) and Bebchuk (1988) with a technological asymmetry those papers had no reason to consider. The result reframes what the transition's welfare economics turn on. It is not whether machine verification finds true violations—it does—but whether the cost of *demonstrating innocence* falls in step with the cost of alleging guilt. The first automated private enforcement wave displays both halves within one episode: automated detection produced a large, genuine compliance response, and the same mechanized claim generation was met by courts with abuse-of-rights pushback—the enforcement stage responding endogenously to a challenge flood (Merane and Stremitzer 2026). That endogeneity deserves stating as its own remark.

Remark 3 (Courts respond). The enforcement reliability q and the legal cost parameters are not invariant to the volume of machine-generated claims: filing fees, standing doctrines, fee-shifting, and abuse-of-rights rulings are the system's endogenous response to challenge floods. A full dynamic treatment is beyond this paper; the static results should be read as short-run, with institutional counter-adaptation as the medium run—which is itself an instance of Result 1's logic applied to the legal system.

9 Result 5: Opacity Relocates

Let the seller allocate effort between a verifiable attribute v and an unverifiable attribute u, with buyer value increasing in both and the seller's cost convex in total effort (Holmström and Milgrom 1991). Pre-shock, neither attribute is priced accurately and distortion is spread. Post-shock, v is verified and priced; the seller's first-order conditions shift effort toward v and, when efforts are substitutes in cost, divert shirking and misrepresentation toward u.

Proposition 6 (Relocation)**.** *Verification of v weakly raises delivered quality on v and weakly lowers it on u when efforts are substitutes; total surplus rises if and only if the verifiable attribute's marginal value is high relative to the diverted distortion. Opacity is not eliminated; it migrates to the unverifiable margin.*

This is the honest boundary on Results 1–3, and the referee's natural objection stated as a result. The transition's reach is the reach of T, the set of technically verifiable attributes: judgment, taste, and radically incomplete states stay outside, and strategic effort flows toward them. The companion paper's obfuscation analysis (Gill 2026) is the within-attribute version; this is the across-attribute version, and both predict the same empirical signature—substitution, not disappearance, documented at national scale in tax settings by Carrillo, Pomeranz, and Singhal (2017).

9.1 Endogenous bespokeness: complexity as a choice

The model so far holds β fixed, and the natural objection is that sellers choose it: contractual complexity is an investment, and raising β to keep one's claims unamortizable is the institutional-scale version of the obfuscation just modeled. Two observations discipline the objection. First, complexity protected opacity under the *old* technology because construction costs rose steeply in format novelty: a certifier's template could not follow a bespoke contract, so bespokeness bought unverifiability. The defining property of the shock is that K falls across formats—machine verifiers generalize where templates could not—so the protective return to complexity on formula-governed dimensions collapses with it. A seller who buys β after the shock buys delay, not shelter, and rational obfuscation reallocates toward the genuinely unverifiable attributes of this section: complexity migrates from syntax to substance. Second, the mirror-image investment—*lowering* β by standardizing, in Barzel's (1982) sense of designing transactions to economize on measurement cost—also loses its return, since amortization no longer requires it; the framework predicts declining pressure for *syntactic* standardization—uniform formats and templates, whose function was to make claims checkable by template verifiers—precisely where machine verification is adopted. *Semantic* standards move the other way: shared identifiers, definitions, and provenance conventions are complements to automated verification at scale, and gain value as it spreads. The companion paper's machine-readable-terms prediction and this section's are therefore the two halves of one statement, not a contradiction: machines relax the format constraint while raising the return to common meaning. A full treatment would put β on the menu as a priced choice and re-derive the assignment; this paper flags the margin, states the direction of both effects, and leaves the formalization—including the redrawing of ownership boundaries that Grossman and Hart (1986) and Hart and Moore (1990) tie to the verifiable set—to the research program.

10 Evidence in Progress and Predictions

This is a theory paper, and it says so: no result above is tested here. The empirical program is nonetheless specified, because the theory's value is that it tells the coming data what to measure.

An implementation is in progress in commercial energy invoicing—a setting chosen because bills are generated from explicit contracts, tariff formulas, metered consumption, and regulated network components, so claims are formula-governed, high-volume, and individually small: the region the companion paper's verification-frontier results, and Result 1's marginal-cost ordering, predict moves first. The author holds commercial interests in this implementation, disclosed on the title page; the design follows the companion paper's measurement framework (Gill 2026, Section 8), with pre-registered claim classes, disclosed false-positive and dispute outcomes, and reproducible evidence trails. Beyond the verification ratio and coverage objects defined there, this paper adds the measured object its own Result 1 turns on: the *construction-cost series* $K(t)$—the cost and time to stand up verification for each new counterparty format or claim class, quarter by quarter. A falling $K(t)$ is the shock; the model says where its consequences appear.

The implementation operationalizes the two layers defined in the introduction and developed as Result 3, measuring them as separate objects from the first audit onward. The conformity layer verifies bills against their contracts; its telemetry is error incidence, recovery, dispute outcomes, false positives, and the counterparty's behavioral response. The optimality layer measures the counterfactual ladder of Section 7 in ascending order of assumptions. G_M is the headline series: current annualized contract cost against the best genuinely obtainable equivalent, repriced against metered consumption, matched on term, volume profile, hedging, and attributes—observable, replicable, and conservative. G_R has its own protocol, which converts the recombination counterfactual from definition into data: the engine emits the specification of the contract that could exist without material additional cost, the specification goes to procurement, and the fill-or-no-fill outcome is logged—each filled specification is a realized G_R, each unfilled one is evidence about where the contestable menu ends. G_E is reported, where at all, only as a bounded, assumption-flagged extension. Alongside the gaps, the implementation logs what closes them: incumbent repricing versus realized switching, the pass-through share (realized concession over identified gap, the incidence object of Section 7's remark), quote acquisition, time-to-resolution, and the action friction s itself—notice periods, exit fees, renegotiation time—which Proposition 4(iv) identifies as the compression floor. Three disciplines govern the optimality object, because an ex-post gap between what a customer paid and the cheapest retrospective alternative is not itself a rent. The gap decomposes into a *computation gap* (the contract was dominated at signing, given the consumption forecast and offers available then—the component sustained by verification cost, and the only component that is this framework's object); a *forecasting gap* (the contract was ex-ante optimal and consumption realized differently—risk, not opacity, though cheap re-verification converts contract choice from a multi-year commitment into continuous re-optimization and so shortens exposure to it); and *compensated unmeasured attributes* (risk protection, green attributes, service quality, exit flexibility—priced dimensions a spot comparison omits, no rent at all, and by Section 9 the predicted destination of relocated opacity). Measurement must therefore be timestamped to information available at signing and reported as an *optimization gap under stated assumptions*; the word rent is earned only by the causal design—exogenous variation in C_O compressing the gaps—not by the size of any gap. Search frictions and inertia in retail electricity are independently documented (Hortaçsu, Madanizadeh, and Puller 2017), and public offer comparators already exist for standardized profiles, so the computation itself is not the novelty—the novelty is systematic, continuous, contract-level auditing at collapsed construction cost, with the cost series $K(t)$ measured alongside the gaps it opens. The two layers carry separate telemetry: error incidence, recovery, and dispute outcomes for conformity; the distribution of achievable savings per audit, its decay over time as counterparties learn they are checked, tariff-complexity counts as the obfuscation response, and switch-versus-reprice outcomes for optimality. Which layer carries the larger rent is an empirical question the implementation is designed to answer rather than assume. The optimality layer's decay under systematic auditing is nonetheless the cleaner test of rent compression, because it requires no enforcement pipeline at all: the threat of exit disciplines pricing before anyone switches.

The two-layer structure is not specific to energy. Any market in which a standing contract meets an evolving feasible set has both objects: insurance (claims settlement against policy terms is conformity; premium and coverage structure against the risk profile and available policies is optimality), telecommunications and software licensing (invoiced usage against the rate card; the rate card against the market), freight, banking fees, and procurement generally. The existing institutional landscape already reflects the two layers priced at high K: recovery-audit firms sell conformity verification, benchmarking consultants sell optimality verification, and both amortize construction costs across clients in exactly the certifier pattern of Section 5—which makes both, on this paper's account, candidates for the tipping dynamics of Proposition 2. Brokers in retail and commercial energy are the optimality layer's incumbent intermediaries, and the implementation's own trajectory—tooling adopted first by brokers, who amortize it across a book of clients, before direct-to-customer deployment—is the migration sequence of Result 1 observed from inside.

The predictions, stated so that future work must test against them (1–4 from the institutional results, 5–8 from Result 3):

1. *De-institutionalization ordering.* Verification adoption grows fastest in bespoke, previously reputation-governed relationships, proceeding down the volume scale; intermediaries whose function is fixed-cost amortization lose share before intermediaries who bear liability or

supply recognized standing; certifier margins re-price from amortization toward recognition.

2. *Brand migration.* Brand premia erode most in categories where the premium is assurance on objectively verifiable attributes (safety, provenance, specification, billing accuracy) and least where it is taste, identity, or status; high-quality small entrants gain share in assurance-heavy categories as K falls. The EU Digital Product Passport—standardized, machine-readable product information arriving category by category—is a prospective staggered rollout that may support quasi-experimental designs rather than an automatic natural experiment: product groups and timing are policy-selected, concurrent ecodesign rules touch the same outcomes, and machine-readable disclosure does not itself verify the upstream assertion. A credible design needs treatment dates, comparable non-passported controls within category, pre-trend tests, and separation from accompanying regulation; granted those, the prediction is that premium compression is larger for passported attributes than for experiential ones within the same product.

3. *Rebuttal asymmetry* (Result 4). Where automated challenge arrives without symmetric rebuttal technology or fee-shifting, nuisance-challenge volume and settlement extraction against compliant parties rise; jurisdictions adopting fee-shifting or heightened pleading show the deterrence gains without the nuisance growth.

4. *Relocation* (Result 5). Post-verification, misconduct and marketing effort shift measurably toward unverifiable attributes within the same relationships—the across-attribute analogue of documented within-tax substitution (Carrillo, Pomeranz, and Singhal 2017).

5. *Optimality frontier and compression.* As counterfactual-construction cost falls, progressively smaller customer-specific mismatches become worth identifying and acting on; among relationships under systematic optimality verification, the distribution of G_M shifts down over time, and the gap between achievable and realized savings narrows.

6. *Repricing before switching.* Incumbents facing systematically informed customers reprice before customers exit, so gap compression outruns realized switching volumes; anticipating verification, suppliers change offered contract structures before switching statistics move.

7. *Ladder and intermediary ordering.* Market gaps (G_M) compress before recombination gaps (G_R), which compress before any movement in economic gaps (G_E), because enforcement runs exit, then procurement, then nothing; correspondingly, brokers and benchmarking intermediaries lose the information-processing component of their margin before the negotiation, execution, aggregation, and liability components.

8. *Adoption heterogeneity and the complexity response.* Non-adopters retain measurably larger optimality gaps than adopters within the same market; and suppliers respond by differentiating on dimensions the counterfactual engine cannot compare—the optimality-layer face of Result 5's relocation.

11 Conclusion

The companion paper asked when checking becomes worth its cost. This paper asked the larger question that answer opens: what happens to an economy whose institutions—brands, certifiers, auditors, contracts, tolerated opacity—were selected precisely because checking was not worth it. The model's answer is a re-selection with structure. The fixed-cost collapse removes the amortization logic that centralized verification, sending it down the volume scale and into the bespoke tail, while intermediation rents migrate from amortization to recognition and liability. Brands lose their bonding function attribute by attribute and keep their taste, identity, and attention functions—a migration of brand capital, not a death. Between conformity and accusation sits the layer where nothing is breached: as individualized counterfactuals become cheap to construct, the mismatches they reveal are disciplined by repricing and exit, compressing toward switching friction—potentially large, fully legal rents, ended without a courtroom, and only stoppable forward, which is why the layer's limit is continuous verification. The welfare economics turn on an asymmetry nobody chose: accusing became cheap faster than proving innocence, and whether the transition improves or degrades welfare in a given domain depends on whether rebuttal costs and legal design restore the symmetry. And opacity does not disappear; it relocates to what machines cannot yet measure, which is where the next version of every result above will have to be fought. The institutions of trust were built to economize on verification. The economizing is ending; the institutions are next; and the order in which they move is now written down precisely enough to be wrong.

A Proofs

Proof of Lemma 1 and Proposition 1. Per-period credibility costs: reputation $P_{KL} = (1-\delta)Ng/\delta$; direct verification $K + Nc$; certification f_C of equation (3); opacity $N\Delta$. (i) $f_C - (K+Nc) = -(1-\beta)K(1-1/M)+\mu$; certification is preferred iff $(1-\beta)K(1-1/M) > \mu$, i.e. iff $K > \bar{K} \equiv \mu M/[(1-\beta)(M-1)]$, which is finite for $\beta < 1$, $M \geq 2$ and diverges as $\beta \to 1$: bespoke relationships never certify. This delivers Lemma 1(i) at high K and Proposition 1(i) as K falls. (ii) Direct verification beats reputation iff $K + Nc \leq (1-\delta)Ng/\delta$, i.e. $g \geq \frac{\delta}{1-\delta}(c + K/N)$; at high K this fails for small N (Lemma 1(ii)); as $K \to 0$ it becomes $g \geq \delta c/(1-\delta)$, independent of N. (iii) The switching gain from reputation to verification, $G(N, g; K) = N[(1-\delta)g/\delta - c] - K$, is increasing in g and, when the bracket is positive, in N; hence along a declining K path the gain turns positive first in high-g, high-N reputation cells and last at small N, while the certified region converts only when K crosses $\bar{K}$. Since the reputation region is the high-β region (part i), the migration enters bespoke relationships before standardized ones. (iv) At $K < \bar{K}$ the only remaining certification advantage is whatever μ prices; relationships whose evidentiary acceptance requires an attester with standing continue to pay it, which is the stated rent migration. □

Proof of Proposition 2. (i) $\Phi(M) = D(f_C(M))$ is nondecreasing (D nonincreasing, f_C decreasing), so by Tarski's theorem its fixed points form a nonempty complete lattice on any invariant interval; under myopic best-response dynamics $M_{t+1} = \Phi(M_t)$, a fixed point M^* with $\Phi'(M^*) < 1$ and Φ crossing from above is locally stable, which at the largest crossing is the stated slope condition. (ii) For clients whose binding outside option is stand-alone verification, a unit fall in K lowers the outside option by one and the fee by $\beta + (1-\beta)/M < 1$, so certification's relative attractiveness falls for every such client and Φ shifts down pointwise; $M^H(K)$ is then nonincreasing, and if the family $\Phi(\cdot\,; K)$ becomes tangent to the diagonal at K_{crit}, the upper crossing and its companion interior crossing merge and vanish—a saddle-node—so the dynamics from any M below the vanished pair converge to the next fixed point below, which is the discontinuous drop. Tangency is a property of the demand distribution, not of complementarity alone; Appendix A.2 exhibits a distribution where it occurs at an interior K_{crit} with closed form, and one where it does not. (iii) After collapse, re-entry from small M faces $f_C(M)$ with the $1/M$ term large, so $\Phi(M) < M$ in a neighborhood of the collapsed state even at K slightly above K_{crit}: the high equilibrium's basin is not re-entered by marginal reversal, which is hysteresis under the maintained dynamics. (iv) Both niche and no-niche limits occur in Appendix A.2's example depending on whether the fee premium over stand-alone costs, μ, lies below or above the top of the distribution of clients' outsourcing value; the general statement is that a niche as $K \to 0$ exists iff a positive mass of clients values certification at $K = 0$ net of μ. □

Proof of Proposition 3. (i) Equation (2) gives $\partial P^*/\partial q = -\frac{(1-\delta)NF}{\delta} < 0$ on $qF < g$ and $P^* = 0$ beyond; where $K + Nc < P_{KL}$, a verified entrant can profitably undercut any price carrying the full pre-shock premium, so the sustainable assurance premium is bounded above by the cost of credible verification—$K + Nc$, plus μ where acceptable evidence requires recognized standing—compressing V_{bond} to the price of that recognition and to zero only where self-produced evidence suffices. (ii) An entrant capturing per-transaction premium π by verifying adopts iff $N\pi \geq K + Nc$, giving $N_E = K/(\pi - c)$ for $\pi > c$, with $\partial N_E/\partial K > 0$. (iii) Verification is truthful by assumption of the technology: a low-quality entrant verified reveals low quality and earns no premium, so the entry benefit accrues only to genuine quality; average entering quality on the verified channel rises. (iv) Let a buyer with per-option attention cost α choose between the default (branded) option and evaluating k verified alternatives at cost αk; the default's attention value is nondecreasing in k, so V_{search} can rise with the volume of verified evidence while V_{taste} is unaffected by q by construction. $\square$

Proof of Proposition 4. (i) is the action condition's definition, with $\partial G^*/\partial C_O = 1$; writing $C_O = K_O/n + c_O$ for a contract-type class supporting n evaluations reproduces the companion frontier's comparative statics in (K_O, c_O, n). (ii) Retaining the relationship at matched terms yields the incumbent a nonnegative margin iff its cost of serving does not exceed the alternative's price; losing the relationship yields zero; repricing is weakly preferred exactly there, so realized switching selects on incumbent disadvantage. (iii) The obtainable set is contained in the recombination-feasible set, which is contained in the economically feasible set; minima over nested sets are ordered, giving $P_E^* \leq P_R^* \leq P_M^*$ and hence the gap ordering; the enforcement ordering restates which counterfactual corresponds to an existing outside option, a creatable one, or none. (iv) is the limit of (i) as $C_O \to 0$ with s fixed. (v) is accounting: conformity remedies operate on realized past deviations; an optimality mismatch involves no deviation, so foregone surplus accrues until detection, expected loss scales with detection delay, and the value of an audit schedule is decreasing in its interval. Each step is immediate from the definitions, as stated in the text. $\square$

Proof of Proposition 5. A challenger who challenges a compliant target at cost a and demands σ receives $\sigma - a$ if the target settles; the target settles iff $\sigma \leq r$; the challenger optimally demands $\sigma = r$, so nuisance profit is $r - a$, positive iff $a < r$ (i). Under a proportional symmetric collapse $a' = \lambda a$, $r' = \lambda r$ with $\lambda < 1$, the condition $a' < r'$ holds exactly when $a < r$ did—the region is scale-invariant in ratio terms—but the extractable surplus contracts to $\lambda(r - a)$, so with participation cost $\phi > 0$ the viable set $\{(a, r) : \lambda(r - a) > \phi\}$ is strictly contained in the pre-collapse viable set (ii). Under fee-shifting, a compliant target litigates (net cost zero when made whole), the challenge loses, and the challenger's payoff is $-a - r < 0$, eliminating nuisance against litigating targets; raising effective accusation cost for low-merit claims operates identically on the condition (iii). Welfare: the challenge sets enabled by a fall in a decompose into true challenges (deterrence and remediation gains net of process costs) and nuisance challenges (pure extraction plus process waste); the sign of the total follows the composition, which part (i) ties to the evolution of r (iv). $\square$

Proof of Proposition 6. Let delivered values be increasing in efforts e_v, e_u with seller cost $C(e_v, e_u)$, C convex with $C_{vu} > 0$ (substitutes). Verification prices e_v: the seller's return to e_v rises, and the first-order conditions give $de_v > 0$ and, by $C_{vu} > 0$, downward pressure on e_u; where misrepresentation is the shirking margin, it migrates to u correspondingly (Holmström and Milgrom 1991). Total surplus rises iff the value gain on v exceeds the diverted loss on u, which is a parameter statement, not a theorem—hence the proposition's conditional form. $\square$

A.2 Worked example: tipping and its absence in closed form

Let a mass $\bar{M}$ of potential clients share (N, c, β) and differ only in the value σ they place on outsourcing verification—integration, attention, and staffing costs avoided by buying rather than making—with $\sigma \sim U[0, b]$. A client's stand-alone cost is $K + Nc + \sigma$; the certifier's fee at clientele $M \geq 1$ is $f_C(M) = (1-\beta)K/M + \beta K + Nc + \mu$. Certification is chosen iff $\sigma \geq \sigma^*(M) = \mu - (1-\beta)K\,(1 - 1/M)$: the client weighs the fee premium μ against the amortization saving. Demand is $D = \bar{M}\,(1 - \sigma^*/b)$ truncated to $[0, \bar{M}]$. Writing $\rho \equiv (1-\beta)K/b$ and $m_0 \equiv 1 - \mu/b$, interior equilibria solve

$$M^2 - \bar{M}(m_0+\rho)M + \bar{M}\rho = 0, \qquad M^{\pm} = \tfrac{1}{2}\Big[\bar{M}(m_0+\rho) \pm \sqrt{\Delta}\Big], \quad \Delta = \bar{M}^2(m_0+\rho)^2 - 4\bar{M}\rho,$$

with M^+ stable ($\Phi'(M^+) = \bar{M}\rho/(M^+)^2 < 1$ whenever $\Delta > 0$) and M^- unstable; for K large enough the corner $M = \bar{M}$ binds and the interior analysis applies below it. Two regimes:

Smooth shrinkage to a niche ($\mu < b$, so $m_0 > 0$). As $K \to 0$, $\rho \to 0$ and $M^+ \to \bar{M}m_0 = \bar{M}(1-\mu/b) > 0$: the clientele declines continuously to the mass of clients whose outsourcing value exceeds the fee premium. No tipping; the certifier survives as an outsourcing service priced on μ.

Discontinuous collapse ($\mu > b$, so $m_0 < 0$). The market exists only where $m_0 + \rho > 0$ and $\Delta \geq 0$; since $\Delta < 0$ at $\rho = -m_0$ and $\Delta > 0$ for ρ large, there is an interior threshold $\rho_{crit} > -m_0$ solving $\bar{M}(m_0+\rho)^2 = 4\rho$, i.e.

$$\rho_{crit} = \frac{2}{\bar{M}}\left(1 - m_0\tfrac{\bar{M}}{2} + \sqrt{1 - m_0\bar{M}}\right) \text{ (larger root)},$$

at which M^+ and M^- merge and vanish: as K falls through $b\rho_{crit}/(1-\beta)$, the clientele drops discontinuously from $M^+ = \bar{M}(m_0 + \rho_{crit})/2$ to zero, and for ρ marginally above ρ_{crit} the collapsed state $M = 0$ is outside the high equilibrium's basin (which requires $M > M^-$), so the market does not re-form: hysteresis. No niche survives, because at $K = 0$ the fee premium μ exceeds every client's outsourcing value.

The dichotomy is the economics of Proposition 2(iv): certifiers whose fee, net of amortization, is covered by the independent value clients place on outsourced attestation shrink smoothly to that base; certifiers priced above it are held together by amortization alone, and amortization-held markets end by tipping, not by decay.

References

Baker, George P., and Thomas N. Hubbard. 2004. "Contractibility and Asset Ownership: On-Board Computers and Governance in U.S. Trucking." *Quarterly Journal of Economics* 119(4): 1443–1479.

Barzel, Yoram. 1982. "Measurement Cost and the Organization of Markets." *Journal of Law and Economics* 25(1): 27–48.

Ball, Ian, and Deniz Kattwinkel. 2025. "Probabilistic Verification in Mechanism Design." *Theoretical Economics* 20(4): 1247–1284.

Bebchuk, Lucian Arye. 1988. "Suing Solely to Extract a Settlement Offer." *Journal of Legal Studies* 17(2): 437–450.

Ben-Porath, Elchanan, Eddie Dekel, and Barton L. Lipman. 2014. "Optimal Allocation with Costly Verification." *American Economic Review* 104(12): 3779–3813.

Biglaiser, Gary. 1993. "Middlemen as Experts." *RAND Journal of Economics* 24(2): 212–223.

Carrillo, Paul, Dina Pomeranz, and Monica Singhal. 2017. "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement." *American Economic Journal: Applied Economics* 9(2): 144–164.

Catalini, Christian, Xiang Hui, and Jane Wu. 2026. "Some Simple Economics of AGI." MIT Sloan Research Paper, February. arXiv:2602.20946.

Commerford, Benjamin P., Sean A. Dennis, Jennifer R. Joe, and Jenny W. Ulla. 2022. "Man Versus Machine: Complex Estimates and Auditor Reliance on Artificial Intelligence." *Journal of Accounting Research* 60(1): 171–201.

Diamond, Douglas W. 1984. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies* 51(3): 393–414.

Dranove, David, and Ginger Zhe Jin. 2010. "Quality Disclosure and Certification: Theory and Practice." *Journal of Economic Literature* 48(4): 935–963.

Fedyk, Anastassia, James Hodson, Natalya Khimich, and Tatiana Fedyk. 2022. "Is Artificial Intelligence Improving the Audit Process?" *Review of Accounting Studies* 27(3): 938–985.

Farrell, Joseph, and Garth Saloner. 1985. "Standardization, Compatibility, and Innovation." *RAND Journal of Economics* 16(1): 70–83.

Gabaix, Xavier, and David Laibson. 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics* 121(2): 505–540.

Gill, Amrit. 2026. "The Economics of Machine Verification: Verification-Cost Shocks and the Extensive Margin of Monitoring." SSRN Working Paper 7307578.

Grossman, Sanford J., and Oliver D. Hart. 1986. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *Journal of Political Economy* 94(4): 691–719.

Halac, Marina, and Pierre Yared. 2020. "Commitment versus Flexibility with Costly Verification." *Journal of Political Economy* 128(12).

Hart, Oliver, and John Moore. 1990. "Property Rights and the Nature of the Firm." *Journal of Political Economy* 98(6): 1119–1158.

Hollenbeck, Brett. 2018. "Online Reputation Mechanisms and the Decreasing Value of Chain Affiliation." *Journal of Marketing Research* 55(5): 636–654.

Holmström, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7(special issue): 24–52.

Jin, Ginger Zhe, D. Daniel Sokol, and Liad Wagman. 2026. "Adaptive Enforcement with AI-Augmented Monitoring." NBER Working Paper 35010.

Hortaçsu, Ali, Seyed Ali Madanizadeh, and Steven L. Puller. 2017. "Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market." *American Economic Journal: Economic Policy* 9(4): 192–226.

Katz, Michael L., and Carl Shapiro. 1985. "Network Externalities, Competition, and Compatibility." *American Economic Review* 75(3): 424–440.

Klein, Benjamin, and Keith B. Leffler. 1981. "The Role of Market Forces in Assuring Contractual Performance." *Journal of Political Economy* 89(4): 615–641.

Law, Kelvin K. F., and Michael Shen. 2025. "How Does Artificial Intelligence Shape Audit Firms?" *Management Science*.

Lizzeri, Alessandro. 1999. "Information Revelation and Certification Intermediaries." *RAND Journal of Economics* 30(2): 214–231.

Luca, Michael. 2016. "Reviews, Reputation, and Revenue: The Case of Yelp.com." Harvard Business School NOM Unit Working Paper 12-016.

Merane, Jakob, and Alexander Stremitzer. 2026. "Automated Private Enforcement: Evidence from the Google Fonts Case." *Journal of Legal Analysis* 18(1): 182–206.

Pérignon, Christophe, Kamel Gadouche, Christophe Hurlin, Roxane Silberman, and Eric Debonnel. 2019. "Certify Reproducibility with Confidential Data." *Science* 365(6449):

127–128.

Poppo, Laura, and Todd Zenger. 2002. "Do Formal Contracts and Relational Governance Function as Substitutes or Complements?" *Strategic Management Journal* 23(8): 707–725.

Rosenberg, David, and Steven Shavell. 1985. "A Model in Which Suits Are Brought for Their Nuisance Value." *International Review of Law and Economics* 5(1): 3–13.

Shapiro, Carl. 1983. "Premiums for High Quality Products as Returns to Reputations." *Quarterly Journal of Economics* 98(4): 659–679.

Townsend, Robert M. 1979. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory* 21(2): 265–293.

Zhang, Hong. 2026. "The Division of Verification: Adjudication, Liability, and the Boundary of the Firm in Checking AI." SSRN Working Paper 7086002.