

The Computational Coase Theorem

AI Agents, Transaction Costs, and the Minimum Efficient Market

Grant Stenger
Working paper

June 2026

Abstract

Coase explained the boundary between firms and markets by the cost of using the price system. This paper moves that boundary down to the level of the transaction, the exposure, and the state-contingent claim. Economic value is frequently allocated implicitly inside firms, platforms, households, queues, insurance pools, long-term contracts, and social norms because the cost of making a transfer explicit exceeds the surplus it would create. AI agents, machine-readable preferences, programmable settlement, identity, verification, and automated negotiation lower some components of that cost: search, specification, bargaining, authentication, pricing, settlement, and attention. They also raise or expose others: privacy leakage, agency conflicts, manipulation, liability, and norm erosion.

The paper develops a formal transaction-cost theory of agentic market formation. A latent exchange or claim j is privately created when

$$\phi_j W_j(S_t) - T_j(S_t) - \Lambda_j(L_t) - \Psi_j(A_t) > 0,$$

and is socially desirable when the same gross value exceeds implementation, liquidity, and externality costs. The core market state is

$$S_t = (\mathcal{H}_t, \mathcal{Q}_t, \kappa_t, L_t, A_t),$$

where \mathcal{H}_t is the inherited payoff span, \mathcal{Q}_t the valuation-dispersion object, κ_t the representation-cost envelope, L_t the liquidity state, and A_t the admissibility frontier. The full program state is $\text{State}_t = (S_t, I_t, N_t)$, where I_t is the infrastructure slice opened in Paper 4 and N_t is the systemic-network slice opened in Paper 3. The first main result defines the minimum efficient transaction: with fixed attention and bargaining costs, small positive-surplus trades do not occur; as computational fixed costs fall, the threshold transaction size falls and the explicit price system expands. The second gives a delegation-amortization result: repeated micro-exchanges become rational when an agent's setup cost is spread over many opportunities. The computational Coase result is treated as a clean benchmark: absent externalities and agency conflicts, cost-reducing agent technology expands the set of explicit, welfare-improving trades; with privacy, extraction, manipulation, collusion, or norm externalities, the benchmark becomes an admissibility inequality, not a prescription to price everything.

The paper then formalizes algorithmic market discovery as dynamic costly basis search. An agentic platform must infer latent exposures, estimate residual valuation value, design contract representations, forecast liquidity, check admissibility, and update the entire state after each market birth. Paper 1 owns the static residual-basis benchmark; this paper owns the agentic dynamic version, where discovery is computationally hard, greedy discovery is justified only under orthogonality, low coherence, weak submodularity, or related structure, and infrastructure investment changes the future search state. The final results show why personal-agent fiduciary constraints are essential: delegation lowers transaction costs but also reveals preferences, creating extraction rents unless agents are user-aligned, privacy-preserving, and constrained by hard non-market rules. The conclusion is not that software prices all things. It is that the minimum efficient market shrinks until it hits the boundaries of information, liquidity, computation, rights, and social meaning.

Keywords: transaction costs; Coase; AI agents; market design; financial innovation; incomplete markets; computational mechanism design; microtransactions; state-contingent claims; platform economics; fiduciary agents.

JEL codes: D23, D47, D52, D53, D61, D82, G13, G18, L14, L51.

Contents

1	Introduction	4
2	Related Literature and Positioning	7
3	The Market State and the Latent Exchange Universe	9
3.1	The program state	9
3.2	Latent exchanges and latent claims	9
3.3	Transaction-cost decomposition	10
3.4	Agentic technologies	11
4	The Claim-Level Coase Boundary	12
4.1	Implicit allocation as a substitute for explicit pricing	13
5	The Minimum Efficient Transaction	13
5.1	One-off transaction	13
5.2	Surplus-threshold version	14
5.3	Attention as the binding fixed cost	14
6	Delegated Bargaining and Agent-to-Agent Exchange	15
6.1	Delegated authority	15
6.2	Delegation amortization	15
6.3	Agent-to-agent protocol primitives	16
7	The Computational Coase Theorem	16
7.1	Benchmark assumptions	17
7.2	Private creation under capture	17
8	Bargaining, Information, and Preference-Revelation Limits	18
8.1	The bilateral-trade limit	18
8.2	Preference revelation as extraction	19
8.3	Search cost versus price discrimination	19
9	Pricing Is Not Neutral: Norms, Dignity, and Endogenous Surplus	19
9.1	Floors and forbidden markets	20
10	Fiduciary Agents and Platform Extraction	20
10.1	Horizontal agent misalignment	21
10.2	Hard constraints as non-convexities	22
10.3	Auditability and proof-carrying transactions	22

11	Algorithmic Basis Search	22
11.1	Candidate dictionaries	22
11.2	Hardness	23
11.3	Greedy discovery under weak structure	23
11.4	Score decomposition	24
11.5	Infrastructure-aware search	24
12	Dynamic State Updates and Path Dependence	25
12.1	Behavioral feedback	26
13	Minimum Efficient Markets: Domains and Examples	26
13.1	Where MEMs arrive first	26
13.2	Where MEMs arrive late or never	27
14	Agentic Mechanisms	27
14.1	Standing-preference posted-price protocols	27
14.2	Batch auctions for micro-exchange	28
14.3	Constraint-preserving auctions	28
15	Admissibility of Agentic Markets	28
15.1	Admissibility tests	29
15.2	Liquidity source and welfare interpretation	29
16	Empirical Predictions and Test Plan	29
16.1	Observable variables	30
16.2	Predictions	30
16.3	Research designs	31
17	Non-Obvious Implications	31
18	Relation to the Research Program	32
19	Conclusion	33
A	Appendix A: A Bilateral Micro-Trade Model	34
B	Appendix B: Power-Law Proliferation of Micro-Markets	34
C	Appendix C: Dynamic Programming Form	34
D	Appendix D: A Simple Privacy-Protection Mechanism	35
E	Appendix E: Claim-Search Pseudocode	35

1 Introduction

Coase’s central move was to treat the market not as a natural background condition but as an institutional technology with costs. If using the price mechanism were free, much production could be coordinated through spot contracts. Because it is not free, activity is often organized inside firms, long-term contracts, authority relationships, and other governance structures (Coase, 1937; Williamson, 1975, 1985). That insight usually enters economics at the firm-market boundary. This paper applies it one level lower. The boundary at issue is not only whether a factory is organized as a firm or a set of contracts. It is whether a particular exposure, reallocation, queue priority, risk, right, forecast, data access, delivery window, compute slice, financing stream, or state-contingent payoff remains implicit or becomes an explicit tradable object.

The claim is simple enough to state in one inequality. A latent exchange becomes an explicit transaction when the surplus from pricing it exceeds the cost of using the price system for that particular object. A latent claim becomes a market when the value of adding its residual payoff direction exceeds the cost of specifying, verifying, pricing, collateralizing, settling, distributing, and governing it. The old price system was too expensive to run on most of economic life. Human beings do not bargain over every lane change, every delivery-window adjustment, every watt of flexible demand, every API call, every minute of queue priority, every small-business revenue exposure, or every item of household risk. Not because these reallocations have no value, but because naming them, searching for counterparties, negotiating terms, reading the state, moving money, resolving disputes, and deciding whether to bother would cost more than the surplus.

AI agents change the boundary because they lower costs that were not merely monetary. They lower cognitive costs, not just payment costs; they make repeated micro-negotiation possible, not just cheaper checkout; they can maintain standing preferences and constraints, not just click faster; they can discover latent exposures, not just route known orders; and they can perform an institutional search over payoff rules, data feeds, settlement methods, liquidity mechanisms, and admissibility constraints. In this sense the relevant innovation is not a better exchange website. It is a distributed layer of delegated computation that makes the price system callable at smaller scales.

But the conclusion is not “price everything.” That would be a category error. Coase’s theorem in its canonical property-rights form is already a frictionless benchmark, not a claim that actual bargaining is costless or harmless (Coase, 1960). The computational version is likewise a boundary benchmark, not a utopia. Pricing is not neutral. Making an interaction explicit can reveal private information, create coercion, attract manipulation, crowd out norms, intensify inequality, induce addiction, or route value to the platform that controls the agent. A software agent that lowers bargaining cost while selling the user’s preferences to counterparties is not a market-completion technology; it is an extraction technology. A dynamic queue market can allocate scarce capacity efficiently in one domain and destroy civic equality in another. A contract on future cash flows can finance a creator’s project, while a contract that gives strangers control rights over intimate life choices should not exist. The efficient boundary is not

$$\text{price if } v - c - T > 0.$$

It is

$$\text{make explicit only if } v - c - T - E > 0$$

and only if the claim passes rights, consent, privacy, manipulation, and systemic-risk screens.

The purpose of this paper is to give the agentic transaction-cost thesis a formal spine. It connects five objects developed across the Trillion Markets research program. Paper 1 studies the

value of residual payoff directions and the representation cost of implementing them. Paper 2 studies the liquidity state that determines whether long-tail claims can be quoted. Paper 3 studies admissibility: which technically feasible and liquid markets should be allowed, restricted, subsidized, or prohibited. Paper 4 studies the infrastructure of state-contingent claims. This paper supplies the computational and transactional layer beneath all four: how agents reduce the minimum efficient size of the exchanges and claims to which those theories apply.

The central core state variable is

$$S_t = (\mathcal{H}_t, \mathcal{Q}_t, \kappa_t, L_t, A_t).$$

Here \mathcal{H}_t is the inherited market span; \mathcal{Q}_t is the local valuation-dispersion object that determines gross risk-sharing value; κ_t is the cost envelope over contractual representations; L_t is the liquidity state, including hedge bases, residual covariance, flow toxicity, dealer capacity, collateral, and subsidy budgets; and A_t is the admissibility frontier, including rights, access, privacy, manipulation, systemic risk, and externality constraints. The full program state is $\mathbf{State}_t = (S_t, I_t, N_t)$: I_t is infrastructure and delivery, while N_t is the systemic network. A computational market creator does not merely ask whether a payoff direction has value. It asks whether the direction is residual, implementable, quotable, admissible, and robust after the state updates.

This state vector is the main discipline of the paper. It prevents the agentic story from collapsing into futurism. The platform problem is not “AI discovers markets.” It is:

$$\max_{M \subseteq \mathcal{J}} \Phi(M; S_t) = \phi(M)V(B_M | \mathcal{H}_t, \mathcal{Q}_t) - \kappa(M; \mathcal{H}_t, T_t) - \Lambda(M; L_t) - \Psi(M; A_t),$$

subject to computational, incentive, and participation constraints, and with the law of motion

$$S_{t+1} = \Gamma(S_t, M_t, \xi_{t+1}).$$

This is a costly, dynamic, constrained basis-search problem. It is hard in general. It becomes tractable only when the dictionary has exploitable structure: approximate orthogonality, low coherence, weak submodularity, reusable infrastructure, separable admissibility, or decomposable liquidity.

The paper’s first contribution is the concept of the *minimum efficient transaction* (MET). With fixed search, negotiation, settlement, verification, and attention costs, there is a minimum surplus or scale below which trade is inefficient even when $v > c$. Agentic systems lower the fixed components and amortize them across repeated opportunities, so the minimum efficient transaction shrinks. This gives a precise version of the claim that micro-markets appear first in compute, bandwidth, energy flexibility, logistics slots, reservations, curb access, and scheduling: these are domains with objective state variables, repeated opportunities, low dignity cost, and relatively standard verification.

The second contribution is a delegated-bargaining theorem. Human attention makes many positive-surplus transactions irrational to negotiate. A software agent becomes valuable when it can express preferences, constraints, and budgets once, then repeatedly authenticate, search, bargain, settle, and log accountability at low marginal cost. The theorem distinguishes one-time setup cost from per-exchange cost and shows when delegation expands the feasible trade set.

The third contribution is the computational Coase benchmark. In a clean environment with common values, no externalities, aligned agents, valid authority, enforceable contracts, and declining transaction costs, cost-reducing computation weakly expands the set of explicit welfare-improving exchanges. This is the benchmark. The paper then spends equal effort on the failures of the benchmark: private information, bargaining impossibility, preference revelation, platform capture, non-neutral pricing, norm crowd-out, and systemic externalities.

The fourth contribution is algorithmic basis search. A platform that proposes markets must solve a selection problem over candidate payoff rules. This nests sparse approximation, maximum coverage, submodular maximization, and mechanism-design constraints (Karp, 1972; Natarajan, 1995; Nemhauser, Wolsey, and Fisher, 1978; Nisan and Ronen, 2001; Das and Kempe, 2018). The useful negative result is that there is no general cheap discovery of optimal markets in an arbitrary dictionary. The positive results are conditional: greedy ranking by marginal score is justified when residual-value functions have weak submodularity and costs/admissibility are sufficiently separable.

The fifth contribution is fiduciary. Delegated agents are not neutral pipes. An agent that controls search, preference revelation, default rules, routing, and consent becomes a market institution. If it is paid by counterparties or by the platform that profits from trading volume, it may lower transaction costs while lowering user welfare. The paper proves a simple alignment result: unless platform revenue is independent of the choice among user-feasible actions, there exist user preferences for which a revenue-weighted agent chooses the wrong transaction. Thus personal agents require hard constraints, preference privacy, auditability, and fiduciary duties.

Reader’s map: formal spine

The paper is organized around nine results.

- C1. Claim-level Coase boundary.** Principle 4.1 states the explicit-price boundary at the level of latent transactions and claims.
- C2. Minimum efficient transaction.** Theorem 5.1 derives the minimum scale and surplus threshold created by fixed attention and transaction costs.
- C3. Delegation amortization.** Theorem 6.2 shows when setup costs for personal agents are rational because they lower marginal costs over repeated opportunities.
- C4. Computational Coase benchmark.** Principle 7.2 gives the monotone benchmark: lower computational transaction costs expand explicit welfare-improving exchange under restrictive assumptions.
- C5. Bargaining and information limits.** Propositions 8.1 and 8.4 show that computation does not eliminate private-information rents or preference-revelation extraction.
- C6. Non-neutral pricing.** Proposition 9.1 shows that pricing can change the surplus and externality primitives; the boundary must be evaluated after pricing, not before.
- C7. Fiduciary agents.** Theorem 10.2 gives a necessary alignment condition for delegated agents.
- C8. Algorithmic market discovery.** Proposition 11.2 and theorem 11.4 characterize hardness and conditional greedy approximation for market-basis search.
- C9. Dynamic state update.** Propositions 12.1 and 12.2 show why market discovery is dynamic and path dependent.

Vocabulary discipline

This paper uses the program’s vocabulary as follows. A *latent exchange* is a reallocation that could create surplus if made explicit. A *latent claim* is a state-contingent payoff direction that could be implemented by some representation. A *transaction cost* is the cost of using the price system for a particular exchange or claim, including search, specification, bargaining, verification, settlement, liability, attention, and privacy. A *computational transaction cost* is the subset lowered by algorithms, agents, data structures, identity, automated negotiation, and programmable settlement. An *admissibility cost* or constraint is not a mere friction; it reflects rights, externalities, manipulation, dignity, privacy, and systemic risk. A *minimum efficient market* is the smallest repeated domain in

which fixed market-creation and agent-delegation costs can be amortized while satisfying liquidity and admissibility constraints.

2 Related Literature and Positioning

The paper is deliberately synthetic. Its novelty is not that transaction costs matter, not that digital markets lower search costs, not that agents can negotiate, and not that market design is computational. The novelty is the integration of those ideas into a claim-level theory of market formation with an explicit state vector linking residual span value, representation cost, liquidity, admissibility, and dynamic search.

Coase, transaction costs, and institutional boundaries. The starting point is Coase (1937): firms exist because using the price mechanism is costly. Coase (1960) then frames property rights and externalities through bargaining costs. Williamson’s transaction-cost economics extends this into a theory of governance, asset specificity, opportunism, and hierarchy (Williamson, 1975, 1985). This paper narrows and deepens the unit of analysis. Instead of asking why a firm exists, it asks why a particular exposure remains implicit. A firm, platform rule, queue, insurance pool, household norm, or long-term contract can be interpreted as a container for many small allocations whose explicit pricing costs exceed their surplus.

Information technology and electronic markets. Malone, Yates, and Benjamin (1987) argued that information technology can shift coordination from hierarchies toward electronic markets by lowering coordination costs. Bakos (1997) studied how electronic marketplaces lower buyer search costs. This paper updates that line for agentic systems. The new object is not just a searchable marketplace; it is a delegated negotiation and settlement layer that can operate continuously over small, repeated opportunities. The market is less a website than a protocol stack.

Search, attention, and bounded rationality. The minimum efficient transaction depends on human attention. Simon (1955) introduced bounded rationality as a limit on optimizing behavior. In this paper, attention is a fixed transaction cost that creates a lower bound on the efficient size of explicit exchange. Software does not make humans unboundedly rational; it changes which tasks can be delegated and which preferences can be encoded as constraints.

Bargaining and mechanism-design limits. A naïve computational Coase theorem would say that if software lowers costs enough, all positive-surplus trades occur. That is false. Private information creates irreducible constraints. The Myerson–Satterthwaite impossibility result shows that even a single buyer and seller with private values cannot generally achieve ex post efficiency, budget balance, incentive compatibility, and individual rationality simultaneously (Myerson and Satterthwaite, 1983). Vickrey–Clarke–Groves mechanisms show how incentive-compatible allocation can be achieved in some quasi-linear environments, but payments, budgets, participation, and information constraints remain load-bearing (Vickrey, 1961; Clarke, 1971; Groves, 1973). This paper treats computational agents as reducing some frictions while leaving incentive constraints as first-order objects.

Algorithmic mechanism design and automated negotiation. Algorithmic mechanism design studies mechanisms under computational constraints (Nisan and Ronen, 2001). Automated negotiation has long studied agent protocols, preferences, concessions, and strategic interaction

(Jennings et al., 2001). Interface agents were framed early as systems that reduce work and information overload (Maes, 1994). This paper uses those traditions to model an economy in which agents do not merely assist users but instantiate the price system at smaller scales. The key addition is the connection to financial market creation: agents search over payoff directions and contract representations, not merely over prices for given goods.

Algorithmic pricing and collusion. Agentic markets can fail horizontally as well as vertically. Vertical misalignment occurs when a user’s agent sells preference information or routes toward platform revenue. Horizontal misalignment occurs when many agents or pricing algorithms learn mutually profitable supracompetitive behavior. Evidence from algorithmic-pricing experiments shows that learning algorithms can sustain collusive outcomes without explicit communication (Calvano et al., 2020). This paper treats agentic collusion as a market-design and admissibility risk, not as a reason to abandon agentic exchange.

Market design and microstructure. Market design views institutions as engineered allocation mechanisms (Roth, 2002, 2015). Microstructure studies the cost of supplying immediacy and the design of trading protocols (Demsetz, 1968; Budish, Cramton, and Shim, 2015). This paper borrows the design perspective but applies it to the boundary between implicit and explicit allocation. The relevant design question is not only how to run a market once it exists, but whether the exchange should be made into a market at all, and whether personal agents should be allowed to route users into it.

Incomplete markets and financial innovation. Papers 1 and 2 of the program sit closer to incomplete markets and financial innovation: complete-markets theory, endogenous asset creation, residual payoff spans, and synthetic liquidity (Arrow and Debreu, 1954; Radner, 1972; Allen and Gale, 1994; Duffie and Rahi, 1995; Shiller, 1993). This paper uses their objects rather than rederiving them. A candidate claim has gross value because it adds a residual payoff direction. Whether it becomes a market depends on computational transaction costs, liquidity, and admissibility.

Information and the limits of market completion. Prices aggregate information (Hayek, 1945), but information is costly and cannot be perfectly reflected in prices in equilibrium (Grossman and Stiglitz, 1980). Information can also reduce social risk-sharing opportunities when it arrives before trade (Hirshleifer, 1971). Under heterogeneous beliefs, new assets may amplify speculation rather than improve welfare (Simsek, 2013). These results discipline the computational story: better agents and better data can create markets, destroy insurance, or generate harmful speculation depending on timing and flow ecology.

Combinatorial search and sparse approximation. Market discovery is a combinatorial problem. Selecting a small set of claims from a large dictionary to maximize residual value resembles sparse approximation, maximum coverage, and submodular maximization (Karp, 1972; Natarajan, 1995; Nemhauser, Wolsey, and Fisher, 1978; Das and Kempe, 2018). Paper 1 owns the static cost-basis version of this problem. This paper uses the same hardness logic for the dynamic agentic version, where search changes the future state through infrastructure, liquidity, admissibility, and learned preferences.

Platform economics and fiduciary concerns. A personal-agent ecosystem is a platform economy. Platforms choose pricing, access, data, and defaults across multiple sides (Rochet and Tirole,

2003). When agents represent users, the platform’s revenue model becomes part of market design. The paper’s fiduciary theorem is therefore not an ethical add-on. It is an equilibrium condition for agentic markets not to become extraction layers.

What is new. The paper’s distinctive contribution is not the monotone benchmark alone. It is the combination of four pieces at the claim level: the minimum efficient transaction, repeated-agent amortization, fiduciary necessity for delegated agents, and dynamic market-search hardness. It combines Coase’s boundary logic with residual payoff value, representation costs, liquidity feasibility, admissibility constraints, computational search, and delegated agents. The main unit is neither the firm nor the exchange venue. It is the latent micro-allocation or claim and the computational infrastructure required to make it explicit.

3 The Market State and the Latent Exchange Universe

3.1 The program state

Let time be discrete. At each date t , the market environment is summarized by

$$S_t = (\mathcal{H}_t, \mathcal{Q}_t, \kappa_t, L_t, A_t).$$

Each coordinate is inherited from the earlier papers. The full program state is

$$\text{State}_t = (S_t, I_t, N_t),$$

where I_t is the infrastructure and delivery state opened in Paper 4 and N_t is the systemic-network state opened in Paper 3. Most equations use S_t to keep notation light; claims whose feasibility or admissibility depends on infrastructure reuse or network amplification are evaluated in State_t .

Definition 3.1 (Market-state coordinates). *The coordinates of S_t are:*

- (i) \mathcal{H}_t : the current span of tradable payoff directions.
- (ii) \mathcal{Q}_t : the local valuation-dispersion object that determines gross value of residual payoff directions.
- (iii) κ_t : the representation-cost envelope mapping payoff directions or claim systems into least-cost admissible implementations before liquidity and externality costs.
- (iv) L_t : the liquidity state: hedge bases, residual covariance, flow toxicity, dealer balance sheets, collateral constraints, operating costs, and subsidy budgets.
- (v) A_t : the admissibility frontier: rights, access, consent, privacy, manipulation, coercion, moral hazard, systemic risk, and externality constraints.

The state vector is deliberately not minimal. It is a discipline device. A proposed market can fail along any coordinate. It can add value but be too expensive to represent; be representable but unquotable; be liquid but inadmissible; be admissible but privately undersupplied; or be privately viable but socially harmful. Paper 3 works with (S_t, N_t) when systemic network dependencies need to be separated from the broader admissibility frontier. Paper 4 works with (S_t, I_t) when infrastructure determines κ_t , L_t , and delivered payoff quality.

3.2 Latent exchanges and latent claims

Let \mathcal{J}_t be a finite or countable dictionary of candidate explicit objects. The dictionary may include:

- bilateral reallocations of time, priority, access, energy, bandwidth, storage, compute, delivery windows, or queue position;
- state-contingent claims such as revenue shares, parametric insurance triggers, event contracts, derivatives, or micro-hedges;
- platform allocation rules that transform implicit priority into priced priority;
- recurring agent-to-agent protocols over a domain of small opportunities.

Each candidate $j \in \mathcal{J}_t$ has a gross value $W_j(S_t)$, private capture fraction $\phi_j \in [0, 1]$, implementation cost $T_j(S_t)$, liquidity cost or feasibility charge $\Lambda_j(L_t)$, and admissibility burden $\Psi_j(A_t)$. The candidate may be a simple bilateral trade with direct surplus $v_j - c_j$, or a payoff subspace with value $V(B_j | \mathcal{H}_t, \mathcal{Q}_t)$. We write the generic value as $W_j(S_t)$.

Definition 3.2 (Private and social scores). *The private score of candidate j is*

$$P_j(S_t) = \phi_j W_j(S_t) - T_j(S_t) - \Lambda_j(L_t) - \Psi_j^{priv}(A_t).$$

The social score is

$$\mathcal{W}_j(S_t) = W_j(S_t) - T_j(S_t) - \Lambda_j^{soc}(L_t) - E_j(A_t),$$

where E_j includes externalities, rights burdens, manipulation risk, privacy loss, coercion, dignity harms, and systemic costs not internalized by the creator.

A candidate is privately created if $P_j > 0$ and socially desirable if $\mathcal{W}_j > 0$. The two inequalities coincide only under special assumptions: $\phi_j = 1$, no public-good leakage, no monopoly extraction, no externalities, no private-social liquidity wedge, and aligned admissibility burdens.

3.3 Transaction-cost decomposition

For a bilateral or small repeated exchange, write the cost of using the price system as

$$T_j = T_j^{search} + T_j^{spec} + T_j^{compute} + T_j^{negotiate} + T_j^{verify} + T_j^{trust} + T_j^{settle} + T_j^{liability} + T_j^{attention} + T_j^{privacy}.$$

For a claim market, T_j is the cost envelope κ_j over contractual representations plus any fixed protocol cost. The terms are:

Cost component	Meaning
Search	finding counterparties, routes, venues, or liquidity sources
Specification	writing the object: quantity, time, payoff, constraints, fallback rules
Computation	estimating value, risk, counterfactuals, and feasible alternatives
Negotiation	bargaining, auction participation, RFQ, strategic delay, concession cost
Verification	observing quality, state, authority, identity, and trigger conditions
Trust	counterparty risk, reputation, fraud screening, escrow, authentication
Settlement	payment, clearing, custody, atomic delivery, reconciliation, tax records
Liability	residual legal exposure, warranty, dispute, insurance, indemnity
Attention	human review, consent, cognitive load, monitoring, regret
Privacy	preference revelation, data leakage, surveillance, discrimination risk

Remark 3.3 (Attention is not a metaphor). *Human attention is a real fixed cost. A person may rationally refuse to negotiate a five-cent improvement in delivery time even when the improvement is real. If a personal agent can negotiate within preauthorized constraints without interrupting the user, the marginal attention cost can fall by orders of magnitude. This is one reason the minimum efficient transaction shrinks.*

3.4 Agentic technologies

Let $a \in \mathcal{A}_t$ denote an agent technology. It maps a user u 's preferences, constraints, budget, and authority into actions:

$$a : (\theta_u, C_u, B_u, \mathcal{I}_u, S_t) \mapsto \Delta,$$

where Δ is a proposed transaction, bid, ask, route, contract, refusal, or request for human confirmation. Agent technology changes transaction costs through a cost transformer

$$T_j(a, S_t) \leq T_j(\emptyset, S_t)$$

for some components and

$$T_j^{\text{privacy}}(a, S_t), T_j^{\text{liability}}(a, S_t), \Psi_j(a, A_t)$$

may rise. The correct comparative static is not that T always falls, but that the vector of costs changes.

Definition 3.4 (Hard constraints and soft preferences). *A user's hard constraints C_u are non-tradable restrictions: examples include no sale of location history, no medical-data sharing, no*

leverage without explicit consent, no intimate-domain markets, no waiver of emergency access, or no exposure above a specified loss bound. Soft preferences θ_u are tradeoffs the agent may optimize subject to hard constraints.

This distinction is central. An aligned agent maximizes over soft preferences subject to hard constraints. A mere optimizer that prices every constraint is not a personal fiduciary; it is a preference-liquidation engine.

4 The Claim-Level Coase Boundary

Principle 4.1 (Claim-level Coase boundary). *Fix a market state S_t . Suppose each candidate explicit object $j \in \mathcal{J}_t$ has gross value $W_j(S_t)$, implementation cost $T_j(S_t)$, liquidity charge $\Lambda_j(L_t)$, and social externality burden $E_j(A_t)$. Then the socially efficient explicit set is*

$$M_t^{\text{soc}} \in \arg \max_{M \subseteq \mathcal{J}_t} \{W(M; S_t) - T(M; S_t) - \Lambda^{\text{soc}}(M; L_t) - E(M; A_t)\},$$

where $W(M; S_t)$ is the gross value of the union of exchanges or payoff spans, not the sum of overlapping one-claim values. If candidates are independent in value and costs, the efficient rule separates:

$$j \in M_t^{\text{soc}} \iff W_j(S_t) > T_j(S_t) + \Lambda_j^{\text{soc}}(L_t) + E_j(A_t).$$

The private creation rule replaces W_j with $\phi_j W_j$ and E_j with private burdens. Therefore the private and social explicit-price boundaries coincide only if capture and externality wedges vanish.

Proof. The first display is the planner’s definition: choose the set of explicit objects maximizing aggregate net surplus. The union value $W(M; S_t)$ avoids double-counting overlapping payoff directions and shared exchanges. Under independence and additivity,

$$W(M; S_t) - T(M; S_t) - \Lambda(M; L_t) - E(M; A_t) = \sum_{j \in M} [W_j - T_j - \Lambda_j - E_j],$$

so the maximization separates across candidates and yields the threshold rule. Private creators maximize captured gross value net of private costs, giving the analogous private rule. Equality of private and social boundaries requires $\phi_j = 1$, $\Lambda_j^{\text{priv}} = \Lambda_j^{\text{soc}}$, $\Psi_j^{\text{priv}} = E_j$, and no set-level externalities or public-good leakage. \square

Remark 4.2 (Coase at the exposure boundary). *The principle is Coasean because the location of the boundary is set by the cost of using the price system. It is computational because those costs are transformed by algorithms, agents, identity, verification, and settlement. It is claim-level because the unit is a payoff direction or micro-allocation, not a firm.*

Example 4.3 (Curb access versus civic priority). *A delivery vehicle’s use of a curb for three minutes has measurable surplus and can be priced if vehicles, municipalities, and drivers have agents that authenticate, bid, reserve, and settle. The same protocol applied to emergency access or voting priority is inadmissible even if technically feasible. The object is not “pricing queues” in the abstract; it is the specific claim, participants, access rule, and externality profile.*

4.1 Implicit allocation as a substitute for explicit pricing

Let an implicit institution $r \in \mathcal{R}$ allocate a domain without explicit prices: a firm rule, a queue, a norm, a platform ranking, an insurance pool, a long-term contract, or a household convention. Let its distortion cost be D_r and operating cost O_r . An explicit market mechanism has cost $T_m + \Lambda_m + E_m$. The institution chooses the cheaper governance technology.

Proposition 4.4 (Implicit-explicit governance choice). *For a domain d , let V_d be the first-best surplus available if all relevant reallocations are made efficiently. Let r be an implicit allocation rule with surplus loss D_r and operating cost O_r , and let m be an explicit market protocol with transaction/liquidity/externality cost $C_m = T_m + \Lambda_m + E_m$. Then the explicit market dominates the implicit rule if and only if*

$$C_m < D_r + O_r.$$

Agentic technology expands explicit pricing in domains where it reduces C_m below the distortion cost of implicit allocation. It does not expand explicit pricing in domains where pricing externalities keep C_m above the implicit institution's distortion cost.

Proof. The implicit rule produces net surplus $V_d - D_r - O_r$. The explicit protocol produces $V_d - C_m$. Compare. \square

This proposition explains why the firm/market boundary and the norm/market boundary are cousins. A firm may allocate internal labor by authority because pricing every task would be costly. A family may allocate help by norms because explicit side payments would damage the relationship. A platform may allocate attention by ranking because micro-auctions would be manipulative or too slow. As C_m falls, some domains shift; as E_m rises, some do not.

5 The Minimum Efficient Transaction

A transaction can have positive surplus and still be inefficient to conduct. Fixed costs create a minimum efficient scale.

5.1 One-off transaction

Consider a candidate exchange of quantity $q \geq 0$. Let buyer value be vq , seller cost be cq , variable transaction cost be τq , and fixed transaction cost be F . Let $b = v - c - \tau$ be net surplus per unit before fixed cost.

Theorem 5.1 (Minimum efficient transaction). *If $b \leq 0$, no positive quantity is efficient. If $b > 0$, an explicit transaction of quantity q is efficient if and only if*

$$q > \frac{F}{b}.$$

The minimum efficient transaction size is

$$\text{MET} = \frac{F}{v - c - \tau}.$$

A reduction in fixed cost from F to F' reduces the minimum efficient transaction size proportionally:

$$\frac{\text{MET}'}{\text{MET}} = \frac{F'}{F},$$

holding v, c, τ fixed. A reduction in variable cost τ reduces MET convexly by increasing the denominator.

Proof. Net surplus is $bq - F$. It is positive if and only if $q > F/b$ when $b > 0$. The comparative static for fixed cost is immediate. The derivative with respect to τ is

$$\frac{\partial \text{MET}}{\partial \tau} = \frac{F}{(v - c - \tau)^2} > 0,$$

so lowering τ lowers the threshold. □

Remark 5.2 (The minimum efficient transaction is not only monetary). *Quantity q can be dollars, seconds, watts, API calls, miles, megabytes, percentage points of risk, delivery-window minutes, or probability-weighted payoff. What matters is that fixed costs create a scale threshold.*

5.2 Surplus-threshold version

For indivisible opportunities, let $s_j = v_j - c_j$ be gross surplus and T_j the all-in transaction cost. Trade occurs iff $s_j > T_j$. If $T_j = F + \tau_j$, a fall in fixed cost converts a mass of previously latent positive-surplus trades into explicit trades.

Proposition 5.3 (Market-count response to threshold decline). *Let latent opportunities have surplus $s \geq 0$ with distribution function $G(s)$ in a population of size N . If transaction cost is a common threshold T , the expected number of explicit opportunities is*

$$M(T) = N(1 - G(T)).$$

If near zero $G(s) = Cs^\alpha + o(s^\alpha)$ with $C > 0, \alpha > 0$, then for small thresholds

$$N - M(T) = NCT^\alpha + o(T^\alpha).$$

If instead opportunities are ranked by surplus $s_j = Aj^{-\alpha}$, then the number of viable explicit opportunities is approximately

$$M(T) \approx \left(\frac{A}{T} \right)^{1/\alpha}$$

before population saturation.

Proof. The first display is the number with $s > T$. The local expansion follows by substituting the small- s expansion of G . The ranked version solves $Aj^{-\alpha} > T$ for j . □

Remark 5.4 (Why market proliferation can look discontinuous). *A fixed cost may sit above a dense mass of small opportunities. When the fixed cost crosses below that mass, the count of explicit transactions can rise sharply even if the technology improves smoothly. This is the transaction-level analogue of the market-birth waves in Paper 1.*

5.3 Attention as the binding fixed cost

Let human review cost be H . Let an agent reduce review cost to h for routine transactions satisfying hard constraints, with $0 \leq h < H$, but require human review for exceptions. Let ρ be the probability that a transaction is routine.

Corollary 5.5 (Attention-cost threshold). *Expected fixed attention cost falls from H to*

$$F_A = \rho h + (1 - \rho)H + F^{\text{setup}}/n$$

over n repeated opportunities, where F^{setup} is the one-time cost of specifying preferences and constraints. Agentic delegation lowers the minimum efficient transaction size if and only if

$$\rho(H - h) > \frac{F^{setup}}{n}.$$

Proof. The reduction in expected attention cost per opportunity is $H - [\rho h + (1 - \rho)H] - F^{setup}/n = \rho(H - h) - F^{setup}/n$. It lowers the threshold when positive. \square

This corollary explains why agentic micro-markets arrive first in repeated, standardized domains. The setup cost of teaching the agent preferences can be amortized across many opportunities, and the routine share ρ is high.

6 Delegated Bargaining and Agent-to-Agent Exchange

The price system does not require humans to consciously negotiate each trade if agents can act under delegated authority. Delegation is economically meaningful only when the agent's actions are valid, constrained, and accountable.

6.1 Delegated authority

A user u grants an agent authority over a domain D . The delegation contract specifies:

$$\mathcal{K}_u(D) = (\mathcal{C}_u, \theta_u, B_u, \ell_u, \mathcal{R}_u, \mathcal{M}_u),$$

where \mathcal{C}_u are hard constraints, θ_u soft preferences, B_u budgets or risk limits, ℓ_u logging and audit requirements, \mathcal{R}_u revocation rules, and \mathcal{M}_u escalation rules.

Definition 6.1 (Valid delegated transaction). *A delegated transaction Δ is valid for user u in domain D if:*

- (i) Δ lies within the agent's authority domain D ;
- (ii) Δ satisfies all hard constraints \mathcal{C}_u ;
- (iii) Δ respects budget and risk limits B_u ;
- (iv) Δ can be authenticated, settled, and logged;
- (v) Δ is not on an admissibility-excluded claim or domain.

6.2 Delegation amortization

Suppose there are n repeated opportunities. Human negotiation cost per opportunity is H . Agent setup cost is F_A , agent marginal cost is a , and the agent's expected implementation error or monitoring cost per opportunity is e . Let gross surplus from opportunity k be s_k .

Theorem 6.2 (Delegation amortization). *Delegated agentic bargaining creates strictly greater net surplus than manual human bargaining over n opportunities if and only if*

$$\sum_{k=1}^n [\mathbf{1}_{\{s_k > a+e\}}(s_k - a - e) - \mathbf{1}_{\{s_k > H\}}(s_k - H)] > F_A.$$

If all opportunities have common surplus s and $a + e < s \leq H$, then manual bargaining produces no trades while delegated bargaining produces all n trades, and delegation is efficient iff

$$n(s - a - e) > F_A.$$

Equivalently, the minimum number of repeated opportunities required for delegation is

$$n^* = \left\lceil \frac{F_A}{s - a - e} \right\rceil.$$

Proof. Manual bargaining executes opportunity k only when $s_k > H$, producing net surplus $s_k - H$. Delegated bargaining executes it only when $s_k > a + e$, producing net surplus $s_k - a - e$, and pays setup cost F_A . Subtract manual surplus from delegated surplus. The common-surplus case is immediate. \square

Remark 6.3 (The first agentic markets are repeated). *One-off micro-bargains rarely justify setup. Repeated domains do: electricity flexibility every hour, compute routing every second, delivery-window selection every order, bandwidth and latency every session, robot scheduling every task, curb allocation every arrival.*

6.3 Agent-to-agent protocol primitives

A minimal agent-to-agent exchange protocol needs five primitives:

Discover \rightarrow Authenticate \rightarrow Negotiate \rightarrow Verify \rightarrow Settle.

The protocol can be RFQ, auction, posted price, batch clearing, combinatorial allocation, or bilateral bargaining. For the theory, the mechanism matters through costs and incentive constraints.

Proposition 6.4 (Protocol cost decomposition). *Let a protocol p in domain D have fixed cost F_p , per-opportunity operating cost o_p , per-opportunity verification/settlement cost z_p , and incentive loss I_p from strategic behavior or imperfect allocation. For expected volume n , the average protocol cost is*

$$\bar{C}_p(n) = \frac{F_p}{n} + o_p + z_p + I_p.$$

A protocol is a minimum efficient market for domain D at volume n if

$$\mathbb{E}[s_D] > \bar{C}_p(n) + E_p(D)$$

and no lower-cost implicit allocation rule yields higher net surplus.

Proof. Average cost is fixed cost amortized over volume plus marginal costs and incentive loss. The threshold follows from comparing expected surplus per opportunity with average cost and externality. \square

Remark 6.5 (A market can be invisible). *If agents clear thousands of micro-RFQs in the background, users may not experience a “market” at all. The market is the protocol performing search, authentication, negotiation, verification, and settlement.*

7 The Computational Coase Theorem

This section states the clean benchmark and then the reasons it fails.

7.1 Benchmark assumptions

Assumption 7.1 (Clean computational Coase environment). *For the benchmark:*

- (i) values and costs are common knowledge or truthfully elicited at zero incentive cost;
- (ii) agents are perfectly aligned with their principals;
- (iii) property rights and authority are well-defined;
- (iv) contracts are enforceable and settlement is reliable;
- (v) no privacy, dignity, manipulation, coercion, or systemic externalities exist;
- (vi) liquidity is available at the modeled cost;
- (vii) introducing prices does not change underlying values, norms, or behavior.

These assumptions are intentionally strong. They define the frictionless target, not the world.

Principle 7.2 (Computational Coase benchmark). *Under Assumption 7.1, let $T_j(a)$ be the transaction cost of candidate exchange or claim j under agent technology a . If technology improves from a to a' so that*

$$T_j(a') \leq T_j(a) \quad \text{for all } j,$$

and gross values W_j and liquidity costs Λ_j are unchanged, then the set of socially efficient explicit candidates weakly expands:

$$\{j : W_j > T_j(a) + \Lambda_j\} \subseteq \{j : W_j > T_j(a') + \Lambda_j\}.$$

If the inequality is strict for some j with

$$T_j(a') + \Lambda_j < W_j \leq T_j(a) + \Lambda_j,$$

then the expansion is strict. Aggregate social surplus weakly increases when the planner can choose the explicit set after the cost reduction.

Proof. For each j , lowering T_j weakly relaxes the threshold $W_j > T_j + \Lambda_j$. Thus any candidate efficient before remains efficient after. Candidates crossing the threshold enter. Since the planner can always choose the old set, the maximum achievable aggregate surplus cannot fall. \square

Remark 7.3 (What the benchmark actually says). *The benchmark says that if computation lowers transaction costs and nothing else changes, more positive-surplus exchanges become efficient. The assumptions rule out the main difficulties: private information, preference extraction, collusion, externalities, liquidity stress, and endogenous norms. The rest of the paper is about those difficulties.*

7.2 Private creation under capture

A private platform creates candidates when

$$\phi_j W_j > T_j + \Lambda_j + \Psi_j^{priv}.$$

Cost reductions can therefore create three kinds of entry: socially good, socially bad, and socially valuable but still privately uncreated.

Proposition 7.4 (Cost decline and private-social divergence). *Suppose $T_j(a') < T_j(a)$. A cost decline causes private entry of j but reduces social welfare if*

$$\phi_j W_j - T_j(a') - \Lambda_j - \Psi_j^{priv} > 0$$

and

$$W_j - T_j(a') - \Lambda_j^{\text{soc}} - E_j < 0.$$

Thus computational market creation can increase the number of explicit markets while lowering welfare.

Proof. The first inequality gives private entry after the cost decline. The second says the social net value of that entry is negative. Entry therefore lowers social welfare by the magnitude of the negative social score. \square

This proposition is the formal antidote to market-proliferation optimism. Lower costs expand the feasible set. They do not choose the right subset.

8 Bargaining, Information, and Preference-Revelation Limits

Computation lowers some transaction costs but does not repeal incentive constraints.

8.1 The bilateral-trade limit

Proposition 8.1 (Myerson–Satterthwaite limit as a transaction-cost floor). *Consider a bilateral trade with one buyer and one seller, independent private values, overlapping supports, quasilinear utility, and gains from trade possible with positive probability. Even if computational, search, settlement, and attention costs are zero, there is generally no mechanism that is simultaneously ex post efficient, Bayesian incentive compatible, individually rational, and budget balanced. Therefore a computational Coase theorem cannot claim that software alone implements all efficient trades.*

Proof. This is the Myerson–Satterthwaite impossibility theorem (Myerson and Satterthwaite, 1983). The result supplies an information-incentive wedge that remains after mechanical transaction costs vanish. \square

Remark 8.2 (Computation can move the wedge, not erase it). *Agents can improve elicitation, reputation, verification, auction design, and repeated-game incentives. They cannot generally make private information irrelevant. Mechanism design remains part of transaction cost.*

Proposition 8.3 (Repeated agentic interaction lowers but need not erase the bargaining floor). *Consider a domain with n repeated opportunities, agent setup cost F_A , mechanical marginal transaction cost a , and an expected private-information inefficiency Δ_n per opportunity under the best admissible repeated protocol available to the agents. The average non-surplus cost per opportunity is*

$$\bar{T}_n = \frac{F_A}{n} + a + \Delta_n.$$

If repetition, reputation, escrow, credentialing, or privacy-preserving sufficient statistics make Δ_n weakly decreasing in n , then repetition lowers the minimum efficient transaction threshold. But if $\liminf_{n \rightarrow \infty} \Delta_n = \Delta_\infty > 0$, private information remains a positive transaction-cost floor even as setup cost is fully amortized.

Proof. The setup component F_A/n falls with n . If Δ_n is weakly decreasing, the total average cost \bar{T}_n weakly falls apart from the constant marginal component a . The limiting average cost is at least $a + \Delta_\infty$. Thus repeated interaction can move the MET through amortization and better incentives, but it cannot eliminate a positive residual private-information wedge. \square

8.2 Preference revelation as extraction

Delegated agents must know user preferences to bargain. Revealing those preferences can transfer surplus to counterparties.

Proposition 8.4 (Preference-revelation extraction). *A buyer has value v for an indivisible good and faces a seller with cost $c < v$. If the seller observes v perfectly and can make a take-it-or-leave-it offer, the seller sets price $p = v$, trade occurs, and the buyer receives zero surplus. If the seller observes only that $v \sim G$, the buyer may retain positive information rent. Therefore lowering preference-communication costs can reduce user surplus when preference information is revealed to a strategic counterparty.*

Proof. With observed v , any price $p \leq v$ is accepted and $p > v$ is rejected. The seller maximizes $p - c$ subject to acceptance, so sets $p = v$. Buyer surplus is $v - p = 0$. With incomplete information, a price that extracts all surplus from every type generally excludes some types; screening leaves information rents for accepted high types under standard conditions. \square

Remark 8.5 (Privacy is productive). *Privacy is not merely a rights constraint. It can be a surplus-preserving bargaining technology. A personal agent should often prove that a transaction satisfies constraints without revealing the full preference schedule that made it acceptable.*

8.3 Search cost versus price discrimination

Reducing buyer search costs can discipline sellers; reducing buyer privacy can empower sellers. These are opposite effects.

Proposition 8.6 (Two-sided information comparative static). *A technology that lowers buyer search cost while preserving preference privacy weakly improves buyer choice sets and can reduce seller market power. A technology that reveals buyer willingness-to-pay to sellers can increase price discrimination and reduce buyer surplus. The welfare effect of agentic commerce therefore depends on whether the agent is an information shield or an information broker.*

Proof. Lower search cost expands the set of offers the buyer can compare and cannot remove previously available offers. Preference revelation changes the seller’s pricing problem by conditioning price on buyer type. As Proposition 8.4 shows, full revelation can transfer all surplus to the seller under take-it-or-leave-it bargaining. \square

This result is central for agent architecture. A user-side agent should often reveal commitments, constraints, credentials, and proofs, not raw valuations.

9 Pricing Is Not Neutral: Norms, Dignity, and Endogenous Surplus

The clean benchmark assumes that pricing does not alter the underlying value of the interaction. Many domains violate this.

Let S_j^0 be surplus under implicit allocation and S_j^1 surplus after explicit pricing. Let T_j^1 be transaction cost and E_j^1 externality after pricing. Pricing is socially desirable iff

$$S_j^1 - T_j^1 - E_j^1 > S_j^0 - D_j^0 - O_j^0,$$

where $D_j^0 + O_j^0$ are distortion and operating costs of the implicit rule.

Proposition 9.1 (Non-neutrality of pricing). *If explicit pricing changes values, behavior, norms, or externalities, the relevant surplus comparison uses post-pricing primitives. In particular, suppose explicit pricing improves allocative efficiency by A_j but creates norm-crowding, dignity, coercion, or inequality cost N_j . Relative to an implicit rule, explicit pricing is socially desirable iff*

$$A_j > T_j^1 + E_j^{other} + N_j - (D_j^0 + O_j^0).$$

If N_j is large enough, lowering T_j^1 to zero is insufficient to justify pricing.

Proof. Write $S_j^1 = S_j^0 + A_j - N_j$ and $E_j^1 = E_j^{other}$. Explicit pricing dominates implicit allocation iff

$$S_j^0 + A_j - N_j - T_j^1 - E_j^{other} > S_j^0 - D_j^0 - O_j^0.$$

Cancel S_j^0 and rearrange. □

Remark 9.2 (The domain boundary). *There are domains where pricing is prohibited not because transaction costs are high, but because N_j or rights constraints are effectively infinite. Examples include control rights over intimate choices, emergency medical priority, votes, coercive labor terms, and many uses of biometric or location data.*

Example 9.3 (Reservation smoothing versus friendship accounting). *A restaurant may use agentic prices to smooth reservations and reduce no-shows. The same technology used among friends to price every favor may destroy the relation it tries to optimize. In the first domain, A_j can exceed $T + E$. In the second, N_j can dominate.*

9.1 Floors and forbidden markets

Definition 9.4 (Non-market floor). *A non-market floor is a constraint requiring access to a good, right, or capacity under specified conditions independent of willingness to pay. Examples include emergency access, basic mobility, civic participation, minimum energy or water access, privacy of core personal data, and certain health-related services.*

Proposition 9.5 (Floors as admissibility constraints). *If a domain has a non-market floor F_d , then any agentic pricing protocol is admissible only over residual capacity or priority above the floor. A protocol that prices access below the floor is inadmissible even if it increases total willingness-to-pay surplus.*

Proof. The floor is a hard constraint in A_t . Feasible mechanisms must satisfy it before welfare maximization over soft tradeoffs. A mechanism violating a hard constraint is outside the admissible set, so its willingness-to-pay surplus is irrelevant to admissible choice. □

10 Fiduciary Agents and Platform Extraction

A personal agent is a market institution. It observes preferences, controls search, filters options, negotiates, and may settle. Its objective matters.

Let actions $x \in X_u$ be valid transactions satisfying hard constraints. User utility is $U_u(x)$. Platform or agent revenue is $R(x)$. An agent with conflict weight $\alpha \geq 0$ chooses

$$x_\alpha \in \arg \max_{x \in X_u} U_u(x) + \alpha R(x).$$

Assumption 10.1 (Rich preference domain). *For any two distinct valid actions $x, y \in X_u$ and any $\eta > 0$, the admissible user-utility domain contains a utility function with $0 < U_u(y) - U_u(x) < \eta$. This says the fiduciary result is about universal alignment across a rich class of users, not about a single known utility function.*

Theorem 10.2 (Necessity of revenue neutrality for universal user optimality). *Under Assumption 10.1, an agent that maximizes $U_u(x) + \alpha R(x)$ selects a user-optimal action for every possible user utility function U_u over X_u only if either $\alpha = 0$ or $R(x)$ is constant over X_u . If $\alpha > 0$ and R is not constant, there exists a user utility function for which the agent chooses an action that is strictly worse for the user than another valid action.*

Proof. If $\alpha = 0$, the agent maximizes user utility. If R is constant, the revenue term does not affect the argmax. Conversely, suppose $\alpha > 0$ and R is not constant. Then there exist $x, y \in X_u$ with $R(x) > R(y)$. Choose a user utility function satisfying

$$0 < U_u(y) - U_u(x) < \alpha(R(x) - R(y)).$$

Then y is strictly user-preferred, but

$$U_u(x) + \alpha R(x) > U_u(y) + \alpha R(y),$$

so the agent chooses x , which is not user-optimal. □

Corollary 10.3 (Fiduciary condition). *A personal agent that routes users through agentic markets is universally user-aligned only if its objective is user utility subject to hard constraints, or if any non-user revenue is independent of the transaction chosen among valid options. Otherwise fiduciary law, audits, technical commitments, or business-model restrictions are required to prevent predictable misalignment.*

Remark 10.4 (The agent as choke point). *In the old market, the venue was the choke point. In agentic markets, the personal agent may be the choke point. It can protect the user from search costs and extraction, or sell the user’s demand curve to the market.*

10.1 Horizontal agent misalignment

Fiduciary failure is vertical: the agent serves platform revenue rather than the user. Agentic markets also create a horizontal failure mode: agents that repeatedly bargain with one another may learn strategies that soften competition, preserve spreads, or coordinate on supracompetitive prices without explicit communication.

Proposition 10.5 (Algorithmic-collusion admissibility burden). *Suppose a class of delegated agents repeatedly interacts in a market protocol and the induced learning dynamics have a stable outcome with user surplus U^{coll} below the competitive or mechanism-design benchmark U^{comp} , without offsetting reductions in transaction, liquidity, or verification cost. Then the protocol carries an agentic-collusion externality*

$$E^{coll} \geq U^{comp} - U^{coll},$$

which must enter the admissibility frontier. Protocol randomization, batch clearing, audit logs, interoperability, best-execution duties, and anti-collusion constraints can be interpreted as design tools for lowering E^{coll} .

Proof. The surplus gap is a transfer or deadweight loss imposed on users relative to the competitive benchmark. If it is not offset by real cost reductions, it is an externality of the protocol and agent objective design. Admissibility must therefore evaluate it alongside privacy, extraction, and systemic costs. □

10.2 Hard constraints as non-convexities

Hard constraints are not merely large negative utility coefficients. Treating them as tradable can make the agent sell what the user meant to forbid.

Proposition 10.6 (Why hard constraints cannot be shadow-priced by default). *Let a user impose a hard constraint $x \notin B$. Suppose an agent instead encodes violation as a finite penalty M in utility. If there exists a counterparty payment or platform revenue gain exceeding M , the agent may choose a forbidden action. Therefore hard constraints require feasibility restrictions or infinite penalties, not ordinary soft tradeoffs.*

Proof. Let $x_B \in B$ violate the constraint and $x_G \notin B$ satisfy it. If the agent’s objective includes utility plus a finite penalty $-M$ for x_B , then a sufficiently large revenue or apparent benefit difference can make x_B optimal. A feasibility restriction excludes x_B from the choice set and cannot be overwhelmed. \square

10.3 Auditability and proof-carrying transactions

An agentic transaction should be able to carry a proof that it satisfied user constraints without revealing the constraints themselves.

Definition 10.7 (Proof-carrying transaction). *A proof-carrying transaction is a tuple (Δ, π) , where Δ is the transaction and π is a verifiable proof that:*

- (i) *the agent had authority for the domain;*
- (ii) *hard constraints were satisfied;*
- (iii) *budget and risk limits were respected;*
- (iv) *required disclosures or confirmations occurred;*
- (v) *the transaction is admissible under the relevant claim rule.*

The proof should reveal no more preference information than necessary.

Remark 10.8 (Privacy-preserving compliance). *The technical form of π can vary: logs, attestations, secure hardware, cryptographic proofs, regulated audit trails, or institutional records. The economic point is invariant: agents must lower verification costs without turning preference data into bargaining ammunition.*

11 Algorithmic Basis Search

Agentic market formation is not simply bilateral bargaining. A platform may search for new markets. It must infer candidate payoff directions, choose representations, estimate value, estimate liquidity, and check admissibility.

Paper 1 owns the static costly-basis problem: how to rank residual payoff directions against a fixed market span and cost dictionary. This section uses that object as an input and studies the agentic extension: dynamic search over claims, wrappers, liquidity states, admissibility constraints, and infrastructure investments that alter the future search problem.

11.1 Candidate dictionaries

Let $\mathcal{J}_t = \{1, \dots, N\}$ be a dictionary of candidate claims or protocols. Candidate j generates residual payoff subspace $B_j \subseteq \mathcal{H}_t^\perp$, representation cost κ_j , liquidity charge Λ_j , admissibility penalty Ψ_j , and

private capture ϕ_j . For a set $M \subseteq \mathcal{J}_t$, write

$$B_M = \text{span}\{B_j : j \in M\}.$$

The private search objective is

$$\Phi(M; S_t) = \phi(M) \frac{1}{2} \text{tr}(P_{B_M} \mathcal{Q}_t) - \kappa(M; \mathcal{H}_t, T_t) - \Lambda(M; L_t) - \Psi(M; A_t).$$

When ϕ is constant and costs are additive, this becomes a penalized projection problem. When infrastructure, liquidity, and admissibility interact, it becomes a dynamic nonseparable problem.

Definition 11.1 (Implemented, quoted, admissible claim). *A candidate claim is IQA-feasible at state S_t if it is:*

- (i) *implemented: a representation exists with finite κ_t ;*
- (ii) *quoted: a liquidity mechanism exists with finite $\Lambda(L_t)$ at relevant size;*
- (iii) *admissible: it satisfies the constraints in A_t .*

The market-search dictionary should be filtered by IQA feasibility before ranking, unless the platform is explicitly considering infrastructure investments that change feasibility.

11.2 Hardness

Proposition 11.2 (Market-search hardness). *For an arbitrary finite dictionary, choosing k candidate claims to maximize gross residual span value*

$$F(M) = \frac{1}{2} \text{tr}(P_{B_M} \mathcal{Q}) \quad \text{subject to } |M| \leq k$$

is at least as hard as maximum coverage. With costs and admissibility constraints, the general market-search problem contains knapsack, maximum coverage, and sparse approximation as special cases. Therefore no agentic platform can be assumed to discover the globally optimal market set cheaply without exploiting structure.

Proof. Reduce maximum coverage to residual span selection. Let universe elements correspond to orthonormal payoff directions e_ℓ , with $\mathcal{Q} = \sum_\ell w_\ell e_\ell \otimes e_\ell$. Let each set j in the maximum-coverage instance correspond to a candidate subspace $B_j = \text{span}\{e_\ell : \ell \in S_j\}$. Then $F(M)$ equals one half the total weight of covered elements. Maximizing $F(M)$ subject to $|M| \leq k$ is maximum coverage up to the factor 1/2, which is NP-hard (Karp, 1972). Adding costs yields knapsack-like variants; requiring sparse representations nests sparse approximation (Natarajan, 1995). \square

Remark 11.3 (No magic discovery theorem). *An AI system can help search, estimate, and propose. It cannot make an arbitrary combinatorial basis-selection problem easy. Claims about autonomous market discovery require assumptions about the dictionary's geometry and the cost structure.*

11.3 Greedy discovery under weak structure

Let $F(M) = \frac{1}{2} \text{tr}(P_{B_M} \mathcal{Q})$ be normalized and monotone. Define the submodularity ratio $\gamma_{L,k}$ as in weak-submodularity analysis.

Theorem 11.4 (Greedy market search under weak submodularity). *Suppose F is monotone and has submodularity ratio at least $\gamma > 0$ at the relevant cardinality scale. Under a cardinality constraint $|M| \leq k$ and zero or equal costs, the greedy rule that repeatedly adds the candidate with largest marginal residual value satisfies*

$$F(M_{\text{greedy}}) \geq (1 - e^{-\gamma}) \max_{|M| \leq k} F(M).$$

If costs are additive and candidates have equal cost, the same guarantee applies to gross value at fixed budget. With unequal costs, density-greedy or knapsack variants require additional assumptions.

Proof. This is the standard weak-submodularity greedy guarantee (Nemhauser, Wolsey, and Fisher, 1978; Das and Kempe, 2018), applied to the residual span-value set function. The theorem’s role is conditional: when residual payoff directions are sufficiently well behaved, simple discovery heuristics can be justified. \square

Remark 11.5 (What weak submodularity means economically). *Weak submodularity is a statement that candidate claims are not too perversely complementary in residual value. Economically, it means the platform’s next-best claim has diminishing but not chaotic marginal value as the existing span grows. Shared infrastructure and liquidity can still create cost complementarities, which belong outside F .*

11.4 Score decomposition

For one-dimensional candidate g , let $r_g = (I - P_{\mathcal{H}_t})g$ and $u_g = r_g / \|r_g\|$ when $r_g \neq 0$. A local score can be written

$$\text{Score}_t(g) = \phi_g \frac{1}{2} \langle u_g, \mathcal{Q}_t u_g \rangle - \kappa_t(g) - \Lambda_t(g) - \Psi_t(g).$$

The score is meaningful only after residualizing against \mathcal{H}_t . A claim with large raw exposure can have small residual value if existing markets already span it. A tiny raw exposure can have high value if it lies in an unspanned direction with large valuation dispersion.

Proposition 11.6 (Residualization before ranking). *Any ranking rule that scores candidate claims by raw payoff variance or raw expected volume rather than residual valuation value can rank a redundant claim above a welfare-relevant missing direction. A necessary condition for program-consistent market discovery is scoring candidates by their residual component relative to \mathcal{H}_t and valuation-dispersion object \mathcal{Q}_t .*

Proof. Let $g_1 \in \mathcal{H}_t$ have arbitrarily large raw variance. Then $r_{g_1} = 0$ and its residual span value is zero. Let $g_2 \perp \mathcal{H}_t$ with small variance but u_{g_2} aligned with a positive eigenvector of \mathcal{Q}_t . Then g_2 has positive residual value. Raw variance ranks g_1 above g_2 for sufficiently large scaling of g_1 , while residual value ranks g_2 above g_1 . \square

11.5 Infrastructure-aware search

Some candidates are not viable alone but become viable after shared infrastructure. Let I be an infrastructure investment with fixed cost F_I that lowers κ_j for $j \in J_I$.

Proposition 11.7 (Infrastructure option value). *Infrastructure I is privately worthwhile for a platform with capture ϕ if*

$$\phi [V(B_{J_I} | \mathcal{H}_t) - V(B_{M_0} | \mathcal{H}_t)] > F_I + \Delta \kappa_{J_I} + \Delta \Lambda_{J_I} + \Delta \Psi_{J_I},$$

where M_0 is the set of claims in J_I that would be viable without I , and all value terms are union-span values. A myopic one-claim search rule can miss I whenever no individual claim can pay the fixed cost but the cluster can.

Proof. The left side is captured incremental cluster value from the infrastructure-induced claim set, net of claims already viable. The right side is fixed infrastructure cost plus incremental implementation, liquidity, and admissibility burdens. If the inequality holds, building I and enabling the cluster is profitable. If each individual claim's captured value is below F_I , a one-claim rule does not build it. \square

This proposition is where AI market discovery becomes more than ranking. It must propose rails, templates, oracles, collateral conventions, and distribution paths that change the cost envelope for many claims at once.

12 Dynamic State Updates and Path Dependence

A market birth changes the state in which future markets are evaluated.

Proposition 12.1 (Dynamic market-state update). *Let M_t be the set of markets or protocols created at time t . The next state is*

$$S_{t+1} = \Gamma(S_t, M_t, \xi_{t+1}),$$

where ξ_{t+1} is exogenous news. Each coordinate can change:

$$\begin{aligned} \mathcal{H}_{t+1} &= \mathcal{H}_t + \text{span}(B_{M_t}) - \text{retired or displaced directions}, \\ \mathcal{Q}_{t+1} &= \mathcal{Q}(\text{new exposures, behavior, beliefs, endowments}), \\ \kappa_{t+1} &= \mathcal{K}(\kappa_t, \text{templates, oracles, legal forms, data rails}), \\ L_{t+1} &= \mathcal{L}(L_t, \text{hedge bases, residual covariance, flow, balance sheets}), \\ A_{t+1} &= \mathcal{A}(A_t, \text{externalities, rights rules, manipulation, systemic links}). \end{aligned}$$

Therefore a market-search policy that is optimal myopically need not be optimal dynamically.

Proof. The first coordinate changes because created claims expand or rebase the tradable span. The second changes because markets alter behavior, exposure choice, and information. The third changes because representations create reusable infrastructure. The fourth changes because liquidity, residual covariance, flow toxicity, and dealer capacity respond to trading. The fifth changes because externalities and regulatory constraints evolve after the market exists. Since today's action affects tomorrow's feasible set and objective, the problem is dynamic; myopic optimality need not imply dynamic optimality. \square

Proposition 12.2 (Myopic failure under cost complementarity). *Consider two candidate markets 1,2. Each has captured gross value B , standalone cost C , and no admissibility burden. If either market exists, it lowers the other's cost by η . Suppose*

$$B - C < 0 \quad \text{and} \quad 2B - (2C - \eta) > 0.$$

Then neither market is myopically viable alone, but creating both is jointly profitable. A myopic one-step search rule selects no market, while a dynamic or cluster-aware rule selects both.

Proof. Standalone net value is $B - C < 0$, so a myopic rule rejects each. Joint net value is $2B - 2C + \eta > 0$, so the pair is profitable. \square

Example 12.3 (Settlement rails before micro-claims). *An agentic payment rail may not be justified by any single household micro-hedge. Once built, it lowers settlement cost for energy flexibility, bandwidth, API calls, and priority markets. The investment is a cluster option, not a single claim.*

12.1 Behavioral feedback

Market creation can change \mathcal{Q}_t , not merely \mathcal{H}_t . A hedge can encourage concentration; insurance can induce moral hazard; priority markets can change congestion behavior; prediction markets can affect the event predicted.

Proposition 12.4 (Endogenous valuation-dispersion decomposition). *Let gross span value be $V(\mathcal{H}, \mathcal{Q}) = \frac{1}{2} \text{tr}(P_{\mathcal{H}}\mathcal{Q})$. If a market creation changes both \mathcal{H} and \mathcal{Q} , then*

$$V(\mathcal{H}', \mathcal{Q}') - V(\mathcal{H}, \mathcal{Q}) = \frac{1}{2} \text{tr}((P_{\mathcal{H}'} - P_{\mathcal{H}})\mathcal{Q}) + \frac{1}{2} \text{tr}(P_{\mathcal{H}'}(\mathcal{Q}' - \mathcal{Q})).$$

The first term is the completion effect holding exposures fixed. The second term is the behavioral feedback effect. Welfare analysis of agentic markets must include both.

Proof. Add and subtract $\frac{1}{2} \text{tr}(P_{\mathcal{H}'}\mathcal{Q})$. \square

13 Minimum Efficient Markets: Domains and Examples

A minimum efficient market (MEM) is a repeated domain where the fixed costs of market creation, delegation, verification, settlement, and governance can be amortized.

Definition 13.1 (Minimum efficient market). *A domain D with opportunity distribution \mathcal{P}_D and protocol p is a minimum efficient market at state S_t if:*

- (i) *expected net surplus per opportunity after marginal protocol costs is positive;*
- (ii) *expected volume amortizes fixed protocol and delegation costs;*
- (iii) *liquidity is feasible under L_t ;*
- (iv) *the domain satisfies admissibility constraints under A_t ;*
- (v) *no implicit allocation rule dominates it in net welfare.*

13.1 Where MEMs arrive first

The theory predicts early MEMs where:

high repetition + objective verification + low dignity cost
+ standard settlement + limited manipulation.

Examples include:

1. **Cloud compute and API/tool calls.** Units are machine-readable, settlement is digital, quality metrics are measurable, and repetition is high.
2. **Bandwidth and latency.** Packets, sessions, and service-level agreements are objective enough for automated negotiation.

3. **Energy flexibility and demand response.** Devices can sell flexibility subject to comfort and safety constraints.
4. **Logistics slots and warehouse scheduling.** Time windows, dock access, robot routes, and inventory movement are repeated and verifiable.
5. **Curb access and charging priority.** Scarce local capacity can be allocated dynamically if floors and public access constraints are preserved.
6. **Reservation smoothing and delivery-window flexibility.** Agents can sell flexibility or buy urgency within user constraints.
7. **Working-capital and revenue-flow microclaims.** Accounting feeds and payment rails can support small, repeated financing claims if admissible.

13.2 Where MEMs arrive late or never

They arrive late or never where:

verification is invasive + values are intimate + manipulation is easy
+ coercion risk is high + norm damage is large.

Examples include intimate relationships, medical choices, reproductive choices, votes, religious participation, emergency access, coercive labor controls, granular location markets, and many claims on minors.

Proposition 13.2 (Domain ordering). *Suppose domains D_1, D_2 have equal opportunity surplus distributions and equal potential volume. If D_1 has weakly lower verification, settlement, liquidity, and admissibility costs than D_2 , with one strict inequality, then D_1 crosses the MEM threshold weakly earlier under any common cost-decline path, and strictly earlier if the threshold is crossed in the interval between their costs.*

Proof. The MEM condition is a net-surplus threshold. Lower costs weakly increase net surplus. Under a common decline path, the lower-cost domain reaches positive net value no later than the higher-cost domain. □

14 Agentic Mechanisms

This section gives simple mechanism templates. The goal is not to solve mechanism design generally, but to identify which protocols agentic transaction-cost reductions make plausible.

14.1 Standing-preference posted-price protocols

A user commits to a schedule:

$$p_u(x, \omega; \theta_u, \mathcal{C}_u)$$

for buying or selling flexibility x in state ω , subject to constraints. A platform matches schedules and settles trades.

Proposition 14.1 (Standing schedules lower negotiation cost). *If preferences over a repeated domain can be represented by standing schedules with update cost F_U and per-match verification cost z , then average negotiation cost over n opportunities is $F_U/n + z$. Manual negotiation cost H is dominated for sufficiently large n whenever $z < H$.*

Proof. Standing schedules incur update cost once and verification per match. Compare $F_U/n + z$ with H and solve $n > F_U/(H - z)$. □

14.2 Batch auctions for micro-exchange

Continuous bilateral negotiation can create latency races and extraction. Batch auctions aggregate orders over short intervals and clear at uniform prices.

Proposition 14.2 (Batching removes micro-latency rents). *In a domain where agents can submit orders within a batch interval and all orders in the interval clear at a uniform price, rents from being infinitesimally faster inside the interval are eliminated, holding order content fixed. The value of speed is shifted from microsecond priority to batch-to-batch forecasting and order choice.*

Proof. Within a batch, execution priority is not determined by infinitesimal arrival time. Orders submitted before the batch cutoff participate in the same clearing calculation. Therefore being marginally faster within the interval does not improve queue position under the mechanism, though strategic forecasting across intervals remains valuable. \square

This is the agentic analogue of frequent batch auction logic in financial markets (Budish, Cramton, and Shim, 2015). It matters because the same agents that lower transaction costs can also intensify predatory speed competition.

14.3 Constraint-preserving auctions

Let agents submit bids plus proofs of constraints. The mechanism maximizes reported surplus subject to feasibility and floors.

Definition 14.3 (Constraint-preserving mechanism). *A mechanism is constraint-preserving if its feasible allocation set is*

$$X^{feas} = \{x : x \text{ satisfies physical feasibility, user hard constraints, and admissibility floors}\},$$

and optimization occurs only over X^{feas} .

Proposition 14.4 (Constraint preservation before efficiency). *A welfare-maximizing agentic mechanism with hard constraints solves*

$$\max_{x \in X^{feas}} \sum_u \hat{U}_u(x) - C(x),$$

not

$$\max_x \sum_u \hat{U}_u(x) - C(x) - \sum_k M_k v_k(x)$$

with finite penalties M_k , unless every finite-penalty optimum happens to lie in X^{feas} . Hard constraints must define the feasible set.

Proof. The first program excludes violations. The second permits them if compensating reported utility is high enough. Unless no violating allocation is optimal under finite penalties, the programs can differ. \square

15 Admissibility of Agentic Markets

Paper 3 owns the full admissibility theory. This paper needs enough of it to define the transaction-cost boundary correctly.

15.1 Admissibility tests

A candidate agentic market should pass six tests:

- A1. Authority.** Does the agent have valid authority to transact?
- A2. Consent.** Is consent informed, revocable, and non-coerced?
- A3. Privacy.** Does the protocol minimize preference and personal-data revelation?
- A4. Manipulation.** Can participants profitably alter the underlying state or resolution?
- A5. Access.** Does pricing violate floors for emergency, civic, health, mobility, or basic services?
- A6. Systemic risk.** Does the market create leverage, correlated liquidation, or shared-oracle fragility?

Proposition 15.1 (Technical feasibility is not admissibility). *There exist domains where agentic technology makes search, negotiation, verification, and settlement costs arbitrarily small, yet the market remains inadmissible because A_t excludes the claim. Therefore a computational reduction in T_j is neither necessary nor sufficient for admissibility.*

Proof. Let $T_j \rightarrow 0$, but suppose j violates a hard rights constraint, such as sale of a vote or coercive control over intimate choices. Then $j \notin A_t$'s feasible set regardless of transaction cost. Conversely, a claim can be admissible but too costly to implement. \square

15.2 Liquidity source and welfare interpretation

A market can be liquid because it serves hedging, immediacy, information, entertainment, speculation, manipulation, or coercion. Agentic systems can increase all of these.

Flow source	Agentic cost reduction	Admissibility question
Hedging	automated discovery of exposures and counterparties	is risk transferred fairly to better bearers?
Immediacy	instant RFQ, routing, settlement	is the user buying real convenience or being nudged?
Information	lower forecasting and settlement cost	is the price decision-relevant and manipulation-resistant?
Entertainment	easy participation, gamified agents	is it non-addictive and non-exploitative?
Speculation	lower leverage and betting frictions	does it amplify risk or create useful information?
Manipulation	cheap automated state influence	can traders move the event or corrupt the oracle?

Remark 15.2 (Volume is not welfare). *Agentic markets may create enormous volume in low-welfare or negative-welfare flow. Volume is a liquidity statistic, not an admissibility proof.*

16 Empirical Predictions and Test Plan

The theory makes predictions about where agentic micro-markets appear, how small transactions become, how protocols spread, and where backlash or regulation arises.

16.1 Observable variables

The empirical unit is a domain d , platform p , and time t . Useful variables include:

Theoretical object	Empirical proxy
Minimum efficient transaction	smallest observed paid unit, minimum contract size, median transaction value, minimum bid increment
Search cost	number of offers queried, latency to match, API discovery cost, routing failures
Negotiation cost	RFQ messages per trade, auction participation cost, human interventions
Verification cost	dispute rate, oracle coverage, data latency, audit cost, sensor/API availability
Settlement cost	payment fee, settlement lag, reconciliation effort, failed-settlement rate
Attention cost	confirmation prompts, opt-out rates, user review time, alert fatigue
Privacy cost	data fields revealed, third-party sharing, personalization intensity, discrimination complaints
Delegation intensity	share of transactions executed by agents without human review
Admissibility burden	regulatory interventions, prohibited categories, complaints, manipulation events
Liquidity state	spreads, fill rates, quoted depth, imbalance, residual covariance, maker concentration
Platform capture	fees, spreads, routing payments, data monetization, take rates

16.2 Predictions

1. **Minimum-size collapse.** After agentic negotiation or programmable settlement adoption, minimum transaction sizes and minimum contract sizes should fall discontinuously in affected domains.
2. **Repeated-domain ordering.** High-frequency, objective, machine-verifiable domains should adopt agentic markets before low-frequency, subjective, invasive domains.
3. **Setup amortization.** Users with more repeated opportunities in a domain should adopt delegated agents earlier because F_A/n is lower.
4. **Protocol clustering.** Market births should cluster around shared identity, settlement, oracle, and template infrastructure.
5. **Residualization.** New claim markets should add value where existing markets fail to span the exposure, not merely where raw exposure is large.
6. **Liquidity gating.** Technically feasible microclaims should fail to trade when factor hedges, residual breadth, or subsidy budgets are absent.

7. **Privacy backlash.** Domains with high preference revelation should show higher opt-out, regulation, or platform switching unless agents provide privacy-preserving proofs.
8. **Fiduciary premium.** User-aligned agents should command trust and adoption advantages in high-stakes domains relative to agents funded by counterparties.
9. **Batching response.** Where agentic trading creates latency predation, successful protocols should move toward batching, uniform prices, or speed bumps.
10. **Norm boundary.** Markets in domains with high norm-crowding costs should face backlash even when users voluntarily transact.
11. **Admissibility floors.** Pricing should expand over residual capacity above floors but not below emergency, civic, or basic-access floors.
12. **Dynamic feedback.** Early market births should lower costs for adjacent markets through templates and oracles, but may raise admissibility costs through externality experience.

16.3 Research designs

Difference-in-differences around agent deployment. Compare domains or users receiving agentic negotiation tools to matched controls. Outcomes: minimum transaction size, transaction frequency, user surplus proxies, dispute rates, and opt-outs.

Settlement-rail event studies. Study introduction of low-cost programmable settlement or identity rails. The theory predicts entry of smaller repeated transactions and clustering around domains with existing verification.

Protocol adoption thresholds. Estimate whether adoption occurs when expected volume crosses $n^* = F_A / (s - a - e)$ or its domain-level analogue.

Privacy architecture comparisons. Compare agents that reveal full preferences with agents that use privacy-preserving constraint proofs. Outcomes: prices paid, seller surplus capture, user retention, and discrimination proxies.

Market-search prediction. Given an estimated $(\mathcal{H}_t, \mathcal{Q}_t, \kappa_t, L_t, A_t)$, test whether high-score candidates are more likely to become markets than high-raw-exposure candidates.

Backlash and prohibition studies. Study domains where agentic pricing was technically feasible but restricted. The theory predicts restrictions where E_j , rights violations, manipulation, or norm costs dominated transaction-cost savings.

17 Non-Obvious Implications

1. The smallest efficient market is not determined by payment fees alone. It is determined by search, specification, negotiation, verification, attention, privacy, liquidity, and admissibility costs.
2. AI agents lower the cost of using the price system only when authority, constraints, settlement, and accountability are machine-readable.
3. The first agentic markets are repeated domains, not necessarily the domains with the largest one-time surplus.

4. Preference privacy can increase user surplus by preventing personalized extraction.
5. A personal agent is a fiduciary market institution, not a browser extension.
6. The value of an agent is often amortization: specify once, transact many times.
7. A market can be made invisible to humans and still be economically explicit.
8. Agentic market creation is a basis-search problem, not a list of clever product ideas.
9. Residualizing candidate claims against the existing market span is necessary before ranking them.
10. The hardest part of autonomous market discovery is often not finding a payoff direction; it is finding a representation, liquidity regime, and admissibility path.
11. Computational cost decline can increase harmful markets before socially valuable markets if harmful markets have higher private capture.
12. Batching and uniform pricing may become more important, not less, as agents get faster.
13. Privacy, floors, and hard constraints are productive components of market design, not anti-market afterthoughts.
14. Some implicit institutions survive not because computation is weak but because pricing would damage the surplus being allocated.
15. Market count is a misleading target; the relevant target is welfare-positive explicit allocation subject to rights and systemic constraints.
16. A platform that controls agents controls market access, preference revelation, and routing; this can become more valuable than the venue.
17. A computational Coase theorem without Myerson–Satterthwaite is wrong; incentive constraints remain after mechanical costs fall.
18. A computational Coase theorem without Hirshleifer is incomplete; better information can destroy pooling.
19. A computational Coase theorem without admissibility becomes financialization ideology.
20. The long-run boundary is not firm versus market. It is implicit allocation versus explicit state-contingent protocol.

18 Relation to the Research Program

Paper 1: Costly basis of incomplete markets. Paper 1 asks which residual payoff directions are valuable. This paper asks how agents lower the minimum efficient size at which those directions can be proposed, negotiated, and implemented. Paper 1’s value object $V(B \mid \mathcal{H})$ becomes one term in the agentic score.

Paper 2: Synthetic liquidity. Paper 2 asks whether a long-tail claim can be quoted using factor hedging, residual diversification, or subsidy. This paper treats the Paper 2 liquidity state L_t as a gating coordinate in market search. A personal agent may identify a valuable micro-hedge, but if no liquidity mechanism supports it, no market exists.

Paper 3: Admissible market design. Paper 3 asks which feasible markets should exist. This paper imports A_t , the admissibility frontier. Agentic technology makes inadmissible markets easier to create; that raises the value of claim-centric regulation and fiduciary agents.

Paper 4: State-contingent claim infrastructure. Paper 4 describes the common infrastructure of prediction, insurance, derivatives, and parametric claims. This paper explains why agents increase demand for that infrastructure: smaller state-contingent claims become worth specifying when agentic transaction costs fall.

The public thesis. The public essay says that the cost of making markets falls and the number of markets can become enormous. This paper gives the microfoundation: the minimum efficient transaction shrinks, delegated agents amortize fixed costs, and platforms search over residual claims subject to implementation, liquidity, and admissibility constraints.

19 Conclusion

The price system has always been a technology for coordinating human plans under scarcity. It was never free. Its costs determined which allocations became explicit markets and which remained inside firms, platforms, contracts, queues, insurance pools, households, and norms. AI agents, programmable settlement, identity, verification, and automated bargaining change those costs. They lower the minimum efficient transaction and make some previously latent exchanges worth pricing.

That is the computational Coase benchmark in its clean form. Lower the cost of using the price system, and the explicit market boundary expands. But the benchmark is only the beginning. Private information remains. Preference revelation creates extraction. Agents can be conflicted. Pricing can alter norms. Liquidity can fail. Externalities can dominate. Rights constraints can prohibit claims regardless of efficiency. The operative claim is therefore conditional: computation expands the admissible price system only where it lowers transaction costs more than it raises privacy, manipulation, liability, norm, collusion, and systemic costs.

The future market is not merely a larger exchange. It is a protocol layer in which personal and institutional agents discover, negotiate, verify, and settle small reallocations under standing constraints. The smallest efficient market may be a cloud-compute auction, an energy-flexibility bid, a delivery-window trade, a reservation swap, a curb-access reservation, a parametric micro-hedge, or an API call. Many will be invisible to the humans they serve. Some should never be built.

The paper's central warning is that agentic markets are not automatically liberating. An agent that represents the user can shrink transaction costs and protect preferences. An agent that represents the platform can shrink transaction costs and sell the user. The difference is not cosmetic; it is the difference between a computational price system and a computational extraction system.

Coase taught that the boundary of the firm is determined by the cost of using the market. The next boundary is finer. It is the boundary between implicit allocation and explicit claim, between human attention and delegated protocol, between unpriced norm and priced reallocation. Software pushes that boundary outward. Economics must decide where it should stop.

A Appendix A: A Bilateral Micro-Trade Model

A buyer has value v , seller has cost c , and an agentic mechanism charges total cost $T = F + \tau$. If values are common knowledge, trade is efficient iff

$$v - c > F + \tau.$$

If v and c are private, let $v \sim G$, $c \sim H$, independently. A direct mechanism asks reports (\hat{v}, \hat{c}) , trades with probability $x(\hat{v}, \hat{c})$, and transfers p_B, p_S . Incentive compatibility and individual rationality impose monotonicity and envelope conditions. The expected gains from trade lost to incentive constraints are an information transaction cost. Agentic systems can lower reporting, verification, and settlement costs, but the IC/IR wedge remains unless repeated interaction, reputation, audits, or other structure changes the mechanism-design problem.

B Appendix B: Power-Law Proliferation of Micro-Markets

Suppose candidate domains have net value after variable cost

$$s_j = Aj^{-\alpha}, \quad \alpha > 0,$$

and fixed computational transaction cost $F_t = F_0 e^{-\lambda t}$. Candidate j is viable iff $s_j > F_t$. Then the number of viable domains is

$$M_t \approx \left(\frac{A}{F_0} \right)^{1/\alpha} e^{\lambda t/\alpha}.$$

If there is an irreducible externality/admissibility floor $E > 0$, viability requires $s_j > F_t + E$, so

$$M_t \approx \left(\frac{A}{F_0 e^{-\lambda t} + E} \right)^{1/\alpha},$$

which plateaus at

$$M_\infty = \left(\frac{A}{E} \right)^{1/\alpha}.$$

Thus admissibility floors cap proliferation even when computational costs vanish.

C Appendix C: Dynamic Programming Form

An agentic market-creation platform solves

$$\max_{\pi} \mathbb{E}^{\pi} \left[\sum_{t=0}^{\infty} \delta^t \left(\Phi(M_t; S_t) - C^{\text{search}}(M_t, S_t) \right) \right],$$

subject to

$$M_t = \pi(S_t, \mathcal{I}_t), \quad S_{t+1} = \Gamma(S_t, M_t, \xi_{t+1}).$$

The Bellman equation is

$$V^*(S) = \max_{M \in \mathcal{M}(S)} \left\{ \Phi(M; S) - C^{\text{search}}(M, S) + \delta \mathbb{E}[V^*(\Gamma(S, M, \xi))] \right\}.$$

Myopic market creation replaces $V^*(\Gamma(\cdot))$ with zero. It is optimal only under strong conditions: no infrastructure complementarity, no residual substitution beyond current value, no behavioral feedback, no liquidity migration, and no admissibility state changes.

D Appendix D: A Simple Privacy-Protection Mechanism

Suppose a buyer’s agent can commit to a maximum willingness to pay b without revealing b . A seller posts price p . The agent returns accept iff $p \leq b$, and a proof that acceptance satisfied the user’s constraints. If the seller cannot observe b , it chooses a price based on the distribution of b , not the individual realization. If the seller observes b , it sets $p = b$. The value of the privacy layer is the buyer surplus preserved by preventing individual price discrimination.

More generally, privacy-preserving agents should reveal:

feasible / infeasible, credential valid, budget sufficient, constraint satisfied,

rather than:

full value function, complete preference ordering, private urgency, all outside options.

E Appendix E: Claim-Search Pseudocode

A disciplined market-search agent should not simply ask a language model for “new markets.” It should run a constrained pipeline:

1. Construct candidate dictionary \mathcal{J}_t from observed exposures, data feeds, user demands, and institutional templates.
2. For each j , compute residual payoff $r_j = (I - P_{\mathcal{H}_t})g_j$.
3. Estimate value $\widehat{W}_j = \frac{1}{2} \langle u_j, \widehat{Q}_t u_j \rangle$ or the relevant bilateral/repeated surplus.
4. Estimate representation cost $\widehat{\kappa}_j$ over available legal forms, oracles, settlement rails, and data sources.
5. Classify liquidity mechanism and estimate $\widehat{\Lambda}_j(L_t)$.
6. Apply admissibility filters A_t , treating hard constraints as feasibility restrictions.
7. Compute private and social scores.
8. Search over clusters, not only individual claims, to capture infrastructure complementarity.
9. Simulate state update S_{t+1} and stress cases.
10. Recommend: create, subsidize, standardize, monitor, restrict, or reject.

This pipeline is intentionally more demanding than a product-ranking model. It is the algorithmic counterpart of the five-paper program.

References

- Akerlof, George A. 1970. “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism.” *Quarterly Journal of Economics* 84(3): 488–500.
- Allen, Franklin, and Douglas Gale. 1994. *Financial Innovation and Risk Sharing*. Cambridge, MA: MIT Press.
- Arrow, Kenneth J., and Gerard Debreu. 1954. “Existence of an Equilibrium for a Competitive Economy.” *Econometrica* 22(3): 265–290.
- Bakos, Yannis. 1997. “Reducing Buyer Search Costs: Implications for Electronic Marketplaces.” *Management Science* 43(12): 1676–1692.
- Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello. 2020. “Artificial Intelligence, Algorithmic Pricing, and Collusion.” *American Economic Review* 110(10): 3267–3297.
- Budish, Eric, Peter Cramton, and John Shim. 2015. “The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response.” *Quarterly Journal of Economics* 130(4): 1547–1621.
- Clarke, Edward H. 1971. “Multipart Pricing of Public Goods.” *Public Choice* 11: 17–33.
- Coase, Ronald H. 1937. “The Nature of the Firm.” *Economica* 4(16): 386–405.
- Coase, Ronald H. 1960. “The Problem of Social Cost.” *Journal of Law and Economics* 3: 1–44.
- Das, Abhimanyu, and David Kempe. 2018. “Approximate Submodularity and Its Applications: Subset Selection, Sparse Approximation and Dictionary Selection.” *Journal of Machine Learning Research* 19(3): 1–34.
- Demsetz, Harold. 1968. “The Cost of Transacting.” *Quarterly Journal of Economics* 82(1): 33–53.
- Duffie, Darrell, and Rohit Rahi. 1995. “Financial Market Innovation and Security Design: An Introduction.” *Journal of Economic Theory* 65(1): 1–42.
- Groves, Theodore. 1973. “Incentives in Teams.” *Econometrica* 41(4): 617–631.
- Grossman, Sanford J., and Joseph E. Stiglitz. 1980. “On the Impossibility of Informationally Efficient Markets.” *American Economic Review* 70(3): 393–408.
- Hart, Oliver, and John Moore. 1990. “Property Rights and the Nature of the Firm.” *Journal of Political Economy* 98(6): 1119–1158.
- Hayek, Friedrich A. 1945. “The Use of Knowledge in Society.” *American Economic Review* 35(4): 519–530.
- Hirshleifer, Jack. 1971. “The Private and Social Value of Information and the Reward to Inventive Activity.” *American Economic Review* 61(4): 561–574.
- Holmström, Bengt. 1979. “Moral Hazard and Observability.” *Bell Journal of Economics* 10(1): 74–91.

- Jennings, Nicholas R., Peyman Faratin, Alessio R. Lomuscio, Simon Parsons, Michael J. Wooldridge, and Carles Sierra. 2001. "Automated Negotiation: Prospects, Methods and Challenges." *Group Decision and Negotiation* 10(2): 199–215.
- Karp, Richard M. 1972. "Reducibility among Combinatorial Problems." In *Complexity of Computer Computations*, edited by Raymond E. Miller and James W. Thatcher, 85–103. New York: Plenum.
- Maes, Pattie. 1994. "Agents that Reduce Work and Information Overload." *Communications of the ACM* 37(7): 30–40.
- Malone, Thomas W., Joanne Yates, and Robert I. Benjamin. 1987. "Electronic Markets and Electronic Hierarchies." *Communications of the ACM* 30(6): 484–497.
- Myerson, Roger B., and Mark A. Satterthwaite. 1983. "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory* 29(2): 265–281.
- Natarajan, Balas K. 1995. "Sparse Approximate Solutions to Linear Systems." *SIAM Journal on Computing* 24(2): 227–234.
- Nemhauser, George L., Laurence A. Wolsey, and Marshall L. Fisher. 1978. "An Analysis of Approximations for Maximizing Submodular Set Functions—I." *Mathematical Programming* 14: 265–294.
- Nisan, Noam, and Amir Ronen. 2001. "Algorithmic Mechanism Design." *Games and Economic Behavior* 35(1–2): 166–196.
- Radner, Roy. 1972. "Existence of Equilibrium of Plans, Prices, and Price Expectations in a Sequence of Markets." *Econometrica* 40(2): 289–303.
- Rochet, Jean-Charles, and Jean Tirole. 2003. "Platform Competition in Two-Sided Markets." *Journal of the European Economic Association* 1(4): 990–1029.
- Roth, Alvin E. 2002. "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics." *Econometrica* 70(4): 1341–1378.
- Roth, Alvin E. 2015. "The Theory and Practice of Market Design." Nobel Prize Lecture.
- Shiller, Robert J. 1993. *Macro Markets: Creating Institutions for Managing Society's Largest Economic Risks*. Oxford: Oxford University Press.
- Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics* 69(1): 99–118.
- Simsek, Alp. 2013. "Speculation and Risk Sharing with New Financial Assets." *Quarterly Journal of Economics* 128(3): 1365–1396.
- Vickrey, William. 1961. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance* 16(1): 8–37.
- Williamson, Oliver E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press.
- Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.