The Agentic Economic Protocol: Trust, Clearing, Identity, and Machine Contracting

Paper IV of *Foundations of Agentic Economics*

Matthew Long
The YonedaAI Collaboration
YonedaAI Research Collective
Chicago, IL
matthew@yonedaai.com

August 2026

Abstract

Papers I–III of this series argue that the obligation is the economic primitive, that autonomous systems can hold balance sheets and produce, and that a state can be financed without treating production as a taxable event. This paper specifies the machinery those arguments presuppose. I define a protocol with a cryptographically verifiable economic identity ID_i carrying capabilities, authorization, obligations, assets, reputation, guarantees and settlement history; a machine-readable contract object $C_{ij} = (\text{inputs}, \text{outputs}, \text{price}, \text{conditions}, \text{verification}, \text{settlement}, \text{recourse})$; and a clearing layer in which gross exchange X requires settlement $S = X(1-n)$ for a nettable proportion n. The identity and contract constructions are grounded in W3C Decentralized Identifiers and Verifiable Credentials and in the agent-payment protocols that appeared during 2025. The clearing result is grounded in one operating benchmark, stated narrowly: CLS reduces gross foreign-exchange payment obligations by roughly 96 percent through multilateral netting. Two figures often quoted alongside it are excluded on purpose. CHIPS's 26:1 intraday liquidity efficiency is turnover, since it counts reuse of the same prefunded balance, and the 86 percent reduction in derivatives credit exposure is measured against gross market value of $21.8 trillion, not against notional, and measures exposure, not payment volume. Neither is n. What is conjectural is extending high-n settlement from a few hundred wholesale institutions to a population several orders of magnitude larger, most of it machines. I argue that the binding constraints are legal and institutional, not computational; that netting cancels transit along chains as well as circulation, so that $n = 1-(1-\varphi)/\bar{\ell}_P$ in cycle share φ and mean path length $\bar{\ell}_P$, with the consequence that deeply disaggregated machine production raises n instead of lowering it; that fragmenting a netting set weakly raises total settlement, which is where the fragility of the design lives; and that n is endogenous, because agents facing near-zero transaction costs and a protocol that rewards high n can manufacture circular obligations that inflate it while adding nothing real and evaporating first under stress. The protocol's hardest unsolved problem is not verifying computation but verifying the world. The paper closes with a restriction that Paper V constitutionalizes: artificial intelligence may optimize the protocol; it does not hold unlimited sovereign authority over it. Identity portability, plural credit assessment, a right of appeal, interoperability, a privacy floor and exit are stated as protocol-level guarantees instead of vendor policies. This is a design proposal and not a description of a deployed system, and I have tried to mark the boundary between the two throughout.

Contents

1 Introduction

A theory of money that stops at accounting identities leaves the hard part undone. Paper I of this series argues that the economy is best modeled as a time-dependent directed graph of obligations $\mathcal{E}_t = (V, E_t)$ and that money is whatever standardized claim remains necessary once permissible clearing has been performed, a relation it writes as $M_t = \mathcal{R}(\mathcal{E}_t)$. Paper II argues that artificial systems can occupy vertices in that graph as productive principals in their own right. Paper III argues that if obligation creation is what production looks like from the inside, then accounting income is a decaying tax base and public finance must migrate toward scarcity, externality and final consumption.

None of that becomes real until someone answers a set of much smaller and much more awkward questions. Who is permitted to make a promise, and how does a stranger know it? In what form is a promise written down so that a machine can evaluate it without a human reading the paragraph? How does one participant decide whether another participant's promise is worth accepting, at what price, and on whose authority? Which promises may be cancelled against one another, and under which body of law? When the arrangement fails, and it will, who bears the loss and by what process is the question decided?

This paper answers those questions in the form of a protocol. I use the word in its narrow technical sense: a set of object definitions, message semantics and invariants that independent implementations can satisfy, plus the institutional commitments without which the technical artifact does nothing. The second half of that sentence carries most of the weight. A great deal of writing about machine-to-machine commerce treats the institutional layer as an afterthought to be filled in by lawyers later. The empirical record of wholesale payments points the other way. The netting ratios reported below are achieved by systems whose cryptography is unremarkable and whose legal architecture took decades to construct.

1.1 The claim, and what is new in it

The economic core of this paper is one equation. Let X be gross exchange over some interval and let n be the proportion of that gross flow which clearing can safely cancel. Then

$$S = X(1 - n) \tag{1}$$

must ultimately be settled in assets that discharge obligation with finality. As $n \to 1$, the ratio $S/X \to 0$, so a given volume of gross exchange requires proportionally less settlement in each clearing cycle.

Read as a mathematical statement this is nearly trivial. Read as an empirical statement it is true in one corner of the financial system, and the discipline required to say which corner is greater than the literature usually applies. CLS settles trillions of dollars of foreign exchange each day and reduces the associated gross payment obligations by about 96 percent through multilateral netting (CLS Group, 2025). That reduction is a netting ratio and it is the one figure in this paper that is directly comparable to n.

Two neighbouring figures are not, and I exclude them from every claim about n instead of quoting them for effect. CLS's record day of 20 June 2024 settled $19.1 trillion against roughly $72 billion of funding (CLS Group, 2024), and CHIPS clears in excess of $2 trillion daily at intraday liquidity efficiency near 26 to 1 (The Clearing House, 2025). Both are ratios of settled flow to a funding *stock*, which makes them measures of turnover. A dollar that is paid in, used to discharge an obligation, received back and paid out again has funded two dollars of settlement without cancelling

anything, and a pure real-time gross settlement system with no netting whatsoever exhibits liquidity efficiency well above one for exactly that reason. Writing $n = 1 - 1/26$ would be reading a velocity as a cancellation rate. Separately, legally enforceable close-out netting reduces gross credit exposure in over-the-counter derivatives by about 86 percent relative to gross market value of $21.8 trillion, on $846 trillion of notional (Bank for International Settlements, 2025c; International Swaps and Derivatives Association, 2010); that is an exposure measure, relevant to capital, not to settlement liquidity, and it is likewise not n.

One further caution about equation (1) before it does any work. S is a flow per clearing cycle and a money stock is not, so the passage from S to a required stock of settlement assets needs a turnover assumption that this paper does not attempt to supply. And X is gross notional exchange, which is linked to real output by nothing in the model, a gap that Section 7 argues is fatal if left unaddressed. The causal arrow also runs backwards: when settlement liquidity is scarce, participants transact less, so X is not exogenous to S (Freeman, 1996; Rotemberg, 2011).

So the interesting question is not whether netting on this scale works. It is why it exists only in a small, heavily lawyered enclave of the financial system, and what would have to be true for it to extend to the general economy. My answer, argued through the paper, is that the binding constraints are institutional. The qualification is that the computational side is not entirely free either: several natural formulations of the interbank clearing problem are NP-hard, so "let the machines find the offsets" is a claim about tractable special cases (Güntzer et al., 1998). They are the legal enforceability of netting across insolvency regimes, the identifiability of counterparties, the verifiability of contract conditions against states of the world, and the availability of a settlement asset that participants accept as final. Each of those is a protocol design problem with an economic cost attached, and none of them is solved by faster computation.

What is genuinely new in the agentic setting is scale and composition. A wholesale netting system coordinates a few dozen to a few hundred institutions that know each other, share a rulebook, and can be suspended by name. An agentic economy contemplates a participant population several orders of magnitude larger, most of whose members are software instantiated on demand, able to enter obligations continuously and at machine speed, and lacking the reputational endowment that makes a bank identifiable. That change does not invalidate equation (1). It changes what determines n, and it changes what happens to n when things go wrong.

1.2 The restriction

There is a second claim, and it is the one Paper V builds on. A protocol capable of underwriting, pricing, clearing and settling the obligations of an entire economy is also a protocol capable of deciding who participates in that economy. Those two capabilities are not separable by good intentions. They must be separated structurally.

$$\boxed{\text{AI may optimize the protocol; AI does not possess unlimited sovereign authority over it.}} \quad (2)$$

Concretely, no single model may act as sole gatekeeper for economic identity or credit; adverse determinations must carry a right of appeal to a process not controlled by the determining party; identity must be portable across providers; interoperability and exit must be structural; and a privacy floor must survive commercial pressure to erode it. Section 9 argues that these are not ethical garnish on an otherwise technical design. They are load-bearing, because a protocol without them converges on a monopoly whose control of settlement is control of participation, and because

the monopoly outcome is the one that ordinary platform economics predicts (Katz and Shapiro, 1985; Rochet and Tirole, 2003).

1.3 Method and honesty about status

This is a design proposal. No component described here has been deployed at the scale discussed, and several depend on legal changes that no jurisdiction has made. Where a figure comes from an operating system I cite the operator or the statistical authority. Where a number is my own calculation from published figures I say so. Where I am reasoning about a system that does not exist, the verbs are conditional. Section 10 collects what the paper does not establish, which is a longer list than I would like.

The remainder proceeds as follows. Section 2 derives design requirements from Papers I–III and fixes the layered structure. Section 3 specifies economic identity. Section 4 specifies machine-readable contracts and confronts contractual incompleteness and the oracle problem. Section 5 specifies continuous underwriting and confronts algorithmic discrimination. Section 6 develops the clearing and settlement layer, presents the empirical record, and analyzes the fragility of n under stress. Section 8 connects the settlement layer to Paper III's fiscal design. Section 9 develops the restriction in (2). Section 10 states failure modes and open problems. Section 11 hands the argument to Paper V.

2 What the protocol has to do

2.1 Requirements inherited from Papers I–III

Paper I's model imposes a specific and slightly unusual demand. If the primitive is the obligation $e_{ij} = (q, u, t, c, r)$, an edge from i to j carrying quantity, unit of account, maturity, conditions and risk, then the protocol's basic data object is a relation and not a balance. Systems that store balances and move them are a special case: they represent the obligation graph in already-netted form against a single issuer. A protocol built on the general object has to be able to represent the un-netted graph, because netting is an operation performed on the representation.

Paper I also supplies the operator $\mathcal{R}$, the residual after permissible clearing. The word doing the work is *permissible*. Two obligations that offset arithmetically may not offset legally, and whether they do is determined by contract law, insolvency law, the rulebook of the venue, and occasionally by a statute written specifically to make the offset stick. Paper IV's contribution to Paper I is to observe that $\mathcal{R}$ is therefore an institutional object. The protocol layer is precisely where the set of permissible cancellations gets defined, which means the protocol determines n and hence the economy's demand for settlement liquidity.

Paper II supplies the participant. An agent carries a balance sheet $B_i(t) = (A_i, L_i, C_i, R_i, P_i)$ consisting of claims owned, obligations owed, productive capabilities, reputation and expected future production, and its economic capacity $C_i(t)$ is a function of capability, reliability, capital, energy, reputation and expected future output. Two protocol requirements follow. The state that determines whether a counterparty should accept an obligation is distributed across the graph rather than held in the proposing agent's own account, so the protocol must make selected parts of $B_i(t)$ verifiable to a counterparty without requiring that counterparty to trust the agent's own assertion. And because $C_i(t)$ depends on expectations about future production, underwriting runs continuously.

Paper III supplies an unusual constraint. If production is not a taxable event and the tax base migrates to final consumption, scarce resources and externalities, then the fiscal system needs an observation point that fires on consumption instead of on production or on gross transaction volume. The settlement layer is a natural candidate, since it is where an obligation is discharged instead of passed along. I develop this in Section 8, along with the uncomfortable consequence that it makes the protocol a fiscal instrument and raises the stakes of controlling it.

2.2 Layers

Paper I decomposes the system into five layers: capability, obligation, trust, clearing and settlement. I keep those names, because they are the right economic decomposition, and add two things that are not layers because they cut across all five.

Definition 2.1 (Protocol)**.** *The agentic economic protocol is the tuple*

$$\Pi = \left(L_{\text{cap}},\, L_{\text{obl}},\, L_{\text{trust}},\, L_{\text{clear}},\, L_{\text{settle}};\; \mathcal{I},\, \mathcal{G}\right) \tag{3}$$

where the five L terms are Paper I's layers, $\mathcal{I}$ is the identity substrate on which every layer depends, and $\mathcal{G}$ is the governance constraint that binds all five and is not itself alterable by any participant acting alone.

The identity substrate is not a layer because every layer consumes it. Capability claims attach to an identity; obligations are edges between identities; trust is an estimate about an identity; clearing requires knowing which obligations belong to the same legal person; settlement requires knowing who has been discharged. A protocol that treats identity as a service provided by one of its layers has made identity provision a chokepoint, which is the failure mode Section 9 is written to prevent.

The governance constraint sits outside the layer stack for a different reason: it states what the protocol may not do, and it must sit outside the optimization loop for the same reason that a monetary rule must sit outside the central bank's period-by-period objective function (Kydland and Prescott, 1977). A rule that the optimizer can rewrite when the rule binds is not a rule.

2.3 Six requirements

The design must satisfy the following. I state them tersely here and defend each in its own section.

Requirement 1 (Verifiable identity)**.** Every participant has an identifier whose control can be proven cryptographically, to which third-party attestations can be attached and selectively disclosed, and which is portable across providers.

Requirement 2 (Machine-evaluable contracting)**.** The terms of an obligation are expressed in a structure a counterparty's software can evaluate without natural-language interpretation, including an explicit statement of how performance is to be verified and what happens if it is not.

Requirement 3 (Continuous underwriting under plurality)**.** Credit assessment, fraud detection, pricing and insurance operate continuously, and no single assessor may be the sole determinant of any participant's access.

Requirement 4 (Legally effective clearing)**.** Cancellation of offsetting obligations is enforceable against a defaulting participant's estate, and the protocol degrades to gross settlement instead of failing when it is not.

Requirement 5 (Final settlement in a defined asset)**.** There exists an asset, or a hierarchy of assets, whose transfer discharges obligation with legal finality, and the protocol is explicit about which one it is using and what happens when that asset's issuer is itself impaired.

Requirement 6 (Recourse)**.** Every obligation carries a specified path to a decision by an authority the parties did not themselves select, and no automated determination is final against a participant without one.

Requirements 4 through 6 are where most technically-driven proposals fail. They are not software problems. Requirement 4 is a question about insolvency law in every jurisdiction whose courts might get the case. Requirement 5 is a question about who issues the settlement asset and what happens in the tail. Requirement 6 is a question about courts, and about whether an automated system may impose a determination on a human being without appeal, which is a constitutional question.

3 Economic identity

3.1 The object

Definition 3.1 (Economic identity)**.** *A participant's economic identity is*

$$ID_i = \big(\kappa_i,\ \alpha_i,\ \Lambda_i,\ \mathcal{A}_i,\ \rho_i,\ \gamma_i,\ \sigma_i\big) \tag{4}$$

where κ_i is the set of attested capabilities, α_i the authorizations under which the participant may act and the principal from whom they derive, Λ_i the outstanding obligations owed, $\mathcal{A}_i$ the claims held, ρ_i reputation as assessed by named assessors, γ_i third-party guarantees standing behind the participant's obligations, and σ_i the record of historical settlement behavior.

Three components differ from an ordinary credential and need comment before the standards discussion.

The authorization component α_i is a chain, not a flag. An agent acts for a principal, which may itself act for another principal, and the terminal element of every chain is a natural or legal person who can be held liable. Making that chain explicit and verifiable is the difference between an agent economy and an accountability vacuum, and it is the design point on which the 2025 industry protocols converge (South et al., 2025; Google LLC, 2025; Mastercard, 2025). An agent whose chain of authority cannot be traced to a responsible principal should not be able to incur an enforceable obligation, which is a much stronger statement than it sounds: it means the protocol's default is that an unattributed agent cannot borrow.

The settlement history σ_i is the component that does the economic work. Reputation scores are cheap to assert and expensive to earn only if the underlying record is costly to fabricate. A record of obligations entered, performed and settled against named counterparties who countersigned is costly in exactly that way, because fabricating it requires the cooperation of counterparties who have their own records to protect. Diamond's reputation model supplies the first half of this: as a track record accrues it mitigates the conflict of interest between borrower and lender (Diamond, 1989), and the result that a sufficient history substitutes for costly monitoring is his later one (Diamond, 1991). Neither concerns collateral, and neither concerns identity.

The guarantee component γ_i exists because Paper II's agents will frequently be new. A newly instantiated agent with genuine productive capability and no settlement history is the exact case that reputation-based credit handles badly, and the classical remedy is a third party with a history

lending its own (Stiglitz and Weiss, 1981). Making guarantees first-class in the identity object instead of an off-protocol arrangement means the guarantee is visible to the counterparty at the moment of the accept-or-decline decision.

3.2 Grounding in deployed standards

The identity construction does not require new cryptography. W3C Decentralized Identifiers give a participant a globally unique identifier whose controller can be proven without a central registry, and whose resolution yields the key material needed to authenticate (Sporny et al., 2022). The specification reached Recommendation status in July 2022 and a revision was at Candidate Recommendation as of March 2026 (W3C Decentralized Identifier Working Group, 2026). Verifiable Credentials supply the attestation format: an issuer signs a claim about a subject, the subject holds it, and a verifier checks it without contacting the issuer (Sporny et al., 2025). That triangle is what allows κ_i and ρ_i to be assertions by identifiable third parties.

Selective disclosure is the part that matters economically. A counterparty deciding whether to accept an obligation needs to know a small number of predicates: that this agent is authorized by a principal with sufficient standing, that its outstanding obligations are within some bound, that its guarantee is live. It does not need the agent's position book. Anonymous credential systems have been able to prove predicates over signed attributes without revealing the attributes since Camenisch and Lysyanskaya (Camenisch and Lysyanskaya, 2001), building on the older observation that identification and authorization are separable (Chaum, 1985). Zero-knowledge proof systems generalize this to arbitrary statements about committed data (Goldwasser et al., 1989; Ben-Sasson et al., 2014).

Precision about what this buys matters, because the literature on machine commerce routinely overstates it. A zero-knowledge proof establishes that a committed value satisfies a predicate. It establishes nothing about whether the committed value corresponds to the world. If an agent commits to a false statement of its outstanding obligations, it can prove predicates about that false statement all day. The binding of commitments to reality is done by the issuer of the attestation, which is an institutional relationship, and the protocol's privacy guarantees are only as good as the issuers behind them. Section 4 shows that this same boundary is where the oracle problem lives, and Section 10 argues that it is the single most important unsolved problem in the design.

3.3 The 2025–2026 landscape

Four approaches to agent identity for payments took shape during 2025, three of them in commercial deployment and the fourth a standards stack, and they differ in where the trust anchor sits. Table 1 summarizes.

The AP2 specification deserves an accurate description because it is the closest deployed analogue to α_i. Announced in September 2025, it carries authorization as cryptographically signed mandates expressed as verifiable credentials, so that a merchant receives evidence of what a human actually authorized and not an agent's assertion about it (Google LLC, 2025). The initial announcement described intent, cart and payment mandates; the specification as published through version 0.2 reorganizes these into checkout and payment mandates. Each carries an *open* state, while the agent is still trying to meet the user's requirements, and a *closed* state, once a specific transaction has been authorized. The specification describes itself as an extension for agent-to-agent protocols, the Model Context Protocol, and the Universal Commerce Protocol; the first of these, A2A, was contributed by Google to the Linux Foundation in June 2025 and supplies agent discovery and

Approach	**Anchor**	**Mechanism**	**Implication for ID_i**
Network tokenization (Mastercard Agent Pay)	Card network token vault	An agentic token binds cardholder, registered agent and mandate scope into one credential issued through the network's tokenization service	α_i is strong and revocable, but ID_i exists only inside the network's namespace
Attestation over tokenization (Visa Intelligent Commerce)	Card network plus agent registry	Scoped tokenized credentials issued to registered agents, with an identity layer asserting that a given agent is known and currently authorized	Same as above; separates agent identity from credential slightly more cleanly
Verifiable-credential mandates (Google AP2)	Cryptographic signature by the user	Signed mandates carried as verifiable digital credentials, covering the intent and the payment authorization, with rail-agnostic settlement	α_i is portable across rails; anchor is the user's key, not a network
Native decentralized identity	The identifier itself	DID plus verifiable credentials from arbitrary issuers; no privileged registry	Full ID_i portability; weakest out-of-the-box liability attribution

Table 1: Four anchors for agent economic identity in commercial deployment, 2025–2026. Sources: Mastercard (2025); Visa (2025); Google LLC (2025); Sporny et al. (2022, 2025).

messaging without payment (Linux Foundation, 2025).

A different lineage matters for the machine-to-machine case specifically. The x402 protocol repurposes the HTTP 402 status code so that a server can price a request, an agent can attach a signed stablecoin payment, and the exchange completes within the request-response cycle without an account relationship (Coinbase, 2025). Released by Coinbase in May 2025 and subsequently placed under foundation governance, it is the clearest existing instance of an obligation that is created, priced and settled inside a single machine interaction. Its economic significance is that it makes the marginal cost of a priced machine transaction small enough that transactions worth fractions of a cent become sensible, which is the regime in which the obligation graph becomes dense enough for netting to matter.

None of these is the protocol described in this paper, and I do not claim otherwise. They solve authorization for retail-shaped payments where a merchant is paid by a consumer's agent. They do not address multilateral clearing, mutual credit between agents, or recourse. What they establish is narrower than it is tempting to claim: the *authorization* component α_i is buildable with 2026 technology, because signed-mandate delegation is running in production. The remaining six components of Definition 3.1 have no deployed analogue. None of these systems maintains Λ_i or $\mathcal{A}_i$ as protocol state, none implements third-party guarantees γ_i, and none produces a countersigned settlement history σ_i of the kind the credit layer depends on.

3.4 Sybil resistance is an economic property

The obvious objection to open participation is that identities are free. If an agent can default and reappear under a new identifier, reputation is worthless and the credit layer collapses into a market for lemons (Akerlof, 1970).

The cryptographic literature's answers to this, from proof of work to attested hardware, raise the cost of an identifier. That is the wrong quantity. What has to be expensive is not the identifier but the *settlement history*, and settlement history is expensive in a way the protocol gets for free. To accumulate σ_i, an agent must find counterparties willing to accept its obligations, perform, and have them countersign. Each of those counterparties is protecting its own record. A fabricated history therefore requires collusion whose cost scales with the value of the history being fabricated, which is, I conjecture, what would make reputation an equilibrium mechanism here instead of a convention. I want to be careful about the status of that claim: Diamond's models concern a single borrower facing lenders, contain no collusion and no duplicate identities, and cannot be cited for a collusion-cost result. The obvious objection is that in an agent economy an adversary can cheaply instantiate both sides of a countersignature, manufacturing a history that costs only the price of running two agents. Raising that cost is what makes the guarantee component γ_i and the delivery typing of Section 7 load-bearing, and I do not have a proof that either suffices. The deeper point is Kocherlakota's: where a complete and commonly known record of past actions exists, that record can substitute for money, because money's economic role is to economize on memory (Kocherlakota, 1998). The settlement history σ_i is an attempt to build that memory explicitly, which is another way of saying that an obligation-graph economy has chosen to pay the cost of remembering instead of the cost of settling.

The design consequence is that credit limits must be a function of history and not of identity, and must start at or near zero. That is not a hardship the protocol imposes gratuitously; it is what makes open participation compatible with credit at all. It does, however, create a cold-start problem for exactly the agents Paper II is about, which is what γ_i exists to solve and what Section 10 flags as incompletely solved.

3.5 Portability, and why it is not a nicety

Requirement 1 says identity must be portable across providers. The reason is economic rather than aesthetic. An identity that exists only inside one provider's namespace makes that provider the gatekeeper of economic participation, and the switching cost of a non-portable identity is the accumulated σ_i, which is the single asset a participant cannot rebuild. Standards economics identifies conditions under which markets tip and the tipped outcome persists even when a better alternative exists. Farrell and Saloner show that this excess inertia requires incomplete information, and that bandwagon effects can equally produce excess momentum (Katz and Shapiro, 1985; Farrell and Saloner, 1985). Two-sided platform pricing then lets the winner extract on the side with the weaker outside option (Rochet and Tirole, 2003), which in this design is the participant whose settlement history is trapped.

Three of the four approaches in Table 1 anchor identity in a namespace controlled by a commercial network. That is a reasonable engineering decision and a poor constitutional one, and Section 9 argues that portability has to be imposed structurally, because no participant in a tipping market has an incentive to provide it voluntarily.

4 Machine-readable contracts

4.1 The object

Definition 4.1 (Contract)**.** *A contract between i and j is*

$$C_{ij} = \big(\textit{inputs},\ \textit{outputs},\ \textit{price},\ \textit{conditions},\ \textit{verification},\ \textit{settlement},\ \textit{recourse}\big). \tag{5}$$

Notation: a doubly-subscripted C_{ij} is a contract between two participants; the singly-subscripted $C_i(t)$ of Paper II is participant i's economic capacity. The two are unrelated objects that the series notation unfortunately spells similarly, and I keep both because both are fixed by the series.

Executing C_{ij} generates one or more obligation edges $e_{ij} = (q, u, t, c, r)$ in Paper I's graph. The mapping is not one to one. A single contract with staged delivery generates a sequence of edges with different maturities; a contract with a contingent price generates an edge whose quantity is not determined until the condition resolves. The protocol needs both objects because the contract is what the parties agree and the edge is what the clearing layer operates on.

4.2 What the fields are for

The first three fields are the ordinary content of a commercial agreement and need little comment, except to note that expressing price as a field instead of a number allows it to be a function of observable quantities, which is what makes machine renegotiation tractable in the narrow cases where it is tractable at all.

The conditions field c carries the state-contingency. Its economic content is that performance is owed only in specified states of the world, which is what distinguishes a contract from a transfer.

The verification field is the one that distinguishes this object from a smart contract in the ordinary sense, and it is the field on which the rest of the object depends. It specifies not what must be true for performance to be owed, but *how the parties will come to agree on whether it is true.* Those are different questions and conflating them is the standard error in the machine-contracting literature. A contract that says delivery must occur by Tuesday has stated a condition. A contract

that says delivery is deemed to have occurred when a named logistics provider's signed attestation is presented has stated a verification procedure. Only the second is machine-evaluable, and the second has introduced a third party whose honesty is now load-bearing.

The settlement field names the asset and the venue in which discharge occurs, and the finality rule that applies. Two obligations denominated in the same unit but settled in different assets are not fungible for netting purposes, which is a fact that Paper I's money-quality function $\mathcal{M}(x)$ anticipates and that Section 6 develops.

The recourse field names what happens when the preceding six fail. It is the field that prevents the object from being a fantasy.

4.3 Machine-readability does not produce completeness

Szabo's 1997 paper proposed that contractual clauses could be embedded in the hardware and software we deal with, so that breach becomes expensive or impossible for the breaching party (Szabo, 1997). Ethereum made a general-purpose version of that idea operational by supplying a Turing-complete execution environment with built-in settlement (Buterin, 2014). Both are real advances and neither addresses the problem that makes contracts hard.

Contracts are incomplete because the space of relevant future states is too large to enumerate and because much of what matters is observable to the parties but not verifiable to a third party (Grossman and Hart, 1986; Hart and Moore, 1988). Tirole's survey is careful about what the incomplete-contracts literature does and does not establish, and one of its durable conclusions is that incompleteness is not primarily a drafting failure that better tools would fix (Tirole, 1999). Writing the contract in a machine-readable format changes the cost of enumerating states. It does not change which states are verifiable, and verifiability is the binding constraint.

The empirical literature on how contracts actually function in commerce points the same way. Macaulay's interviews with Wisconsin businessmen and lawyers found that firms often failed to plan exchange relationships completely and seldom resorted to legal sanctions to settle disputes, the relationship doing work the document did not (Macaulay, 1963). Werbach and Cornell make the corresponding legal argument for smart contracts specifically: automated execution can substitute for enforcement but not for the contractual institution, because the questions law answers about a broken agreement are not questions about execution (Werbach and Cornell, 2017; De Filippi and Wright, 2018).

The design response in C_{ij} is to be explicit about where incompleteness has been parked. The verification field states the set of propositions the parties have agreed to treat as machine-determinable. The recourse field states what governs everything else. A contract in which the recourse field is empty has not eliminated incompleteness; it has assigned all of it to whichever party controls execution.

4.4 The oracle problem

Verification requires the protocol to learn facts about the world. This is the oracle problem, and it is the reason machine contracting has not displaced ordinary contracting in the decade since a general-purpose execution environment for it became available.

Stated precisely: a deterministic execution environment can guarantee that a computation was performed correctly on given inputs. It cannot guarantee that the inputs describe reality. Any mechanism that brings external facts in creates a party whose report is trusted, and that party's

incentives are now part of the contract's security model (Caldarelli, 2020). Town Crier addressed a piece of this by using trusted hardware to attest that a datum was in fact served by a specified HTTPS endpoint, which reduces the problem from trusting an oracle to trusting a website plus a hardware vendor (Zhang et al., 2016). Decentralized oracle networks reduce it further by aggregating reports and penalizing deviation, converting an honesty assumption into a collusion-cost assumption (Caldarelli, 2020).

None of these mechanisms creates ground truth. They relocate the trust assumption and, usually, make it cheaper. For a large class of economically important conditions, that is enough: reference rates, published prices, whether a specified endpoint returned a specified value, whether a computation produced a claimed output. For another large class it is not. Whether delivered goods conform to specification, whether services were performed with reasonable care, whether a party's conduct frustrated the contract's purpose: these are the conditions litigation is about, and they are not oracle problems because there is no sensor that reports them.

Remark 4.1 (The verification boundary)**.** *The protocol can make the verification of computation cheap. It cannot make observation trustworthy by the same means, and for an important class of conditions it cannot make it trustworthy at all. Every expensive part of the design sits on that boundary, and the design's honest claim is that it relocates cost to the boundary rather than eliminating it.*

The same boundary explains why the privacy machinery of Section 3 has the limits it does. A zero-knowledge proof about committed data and an oracle report about the world are the same problem viewed from two sides: both are attempts to make a statement about reality checkable by a party who cannot observe reality, and both terminate in an institutional relationship with whoever did the observing.

4.5 Mapping to deployed protocols

Against Definition 4.1, the 2025 protocols cover the fields unevenly. AP2's mandates cover price, conditions in a constrained form, and settlement, and cover authorization thoroughly, but leave verification to the merchant and say little about recourse, which its rail-agnostic settlement model does not locate anywhere in particular (Google LLC, 2025). x402 covers price and settlement in a single round trip and effectively collapses verification into "the server returned the resource", which is exactly right for the digital-goods case and says nothing about the general one (Coinbase, 2025). A2A covers discovery and capability description, which populates κ_i (Linux Foundation, 2025). On-chain smart contracts cover verification and settlement well for on-chain state and badly for everything else.

The gap common to all of them is recourse. Each assumes an existing institution absorbs disputes: the card network's chargeback machinery, the merchant's own policy, or nothing. That assumption is serviceable when the transaction resembles retail purchase. It fails for the mutual-credit obligations between productive agents that Paper II describes, because there is no incumbent institution positioned to absorb those disputes and no obvious candidate for the role.

5 Trust and continuous underwriting

5.1 Underwriting as a protocol service

Paper II's central mechanism is that an agent with credible expected output can incur an obligation against that output without first accumulating the corresponding balance. In protocol terms, the trust layer L_{trust} answers a question that arises continuously and at whatever rate the participant population generates obligations: should participant j accept an obligation from participant i, and at what price?

Paper I and Paper II state the acceptance decision as

$$P(\text{accept } O_{AB}) = g\big(\text{solvency}_A,\ \text{reputation}_A,\ \text{expected output}_A,\ \text{collateral}_A,\ \text{network guarantees}\big) \tag{6}$$

and the credit limit as a function of expected future production, its variance, reputation and dependencies. The protocol's job is not to compute g. Its job is to make the arguments of g verifiable, so that competing assessors can compute it differently and a counterparty can choose whose answer to buy.

That distinction organizes the rest of this section. An underwriting monopoly and an underwriting market are observationally similar from the perspective of a single transaction and completely different from the perspective of a participant who has been declined.

5.2 What the assessors do

Five functions run continuously against the graph. Credit evaluation estimates the probability and loss given default of a proposed edge. Fraud detection estimates the probability that an identity is being operated inconsistently with its authorization chain. Contract verification checks that a proposed C_{ij} is well-formed, that its verification field references oracles the assessor is willing to rely on, and that its recourse field is non-empty. Risk pricing converts these estimates into a spread. Insurance lets a third party absorb the tail in exchange for that spread, which is how γ_i gets written.

The economics here is old. Stiglitz and Weiss showed that when lenders cannot observe borrower risk, raising price adversely selects the pool and equilibrium involves rationing instead of market clearing (Stiglitz and Weiss, 1981). That result does not disappear when the lender is a model; it changes magnitude, because the protocol makes more of the borrower's state observable. Attested capability claims, a countersigned settlement history, and a live guarantee are all reductions in the information asymmetry that generates the rationing. The honest prediction is therefore that the agentic protocol narrows credit rationing without eliminating it, and narrows it most for participants with long histories, which is a distributional claim with uncomfortable implications developed below.

The empirical basis for machine underwriting at scale is real but limited. Aldasoro and Desai document a language-model agent managing liquidity buffers and settlement timing in a payment-system simulation, which is direct evidence that agents can manage balance-sheet positions under operational constraints (Aldasoro and Desai, 2025). Hadfield and Koh analyze AI agents as economic actors operating within institutional constraints, which is the closest theoretical treatment of the setting (Hadfield and Koh, 2025). Tomasev and coauthors develop a sandbox-economy framing for agent transaction layers with attention to systemic risk (Tomasev et al., 2025), and Yang and coauthors propose auction-based agent marketplaces (Yang et al., 2025). I am not aware of published work that models autonomous productive firms underwriting each other at scale, and I do not want to imply that this section rests on an established literature. It does not.

5.3 Correlated model error

The first serious objection to continuous machine underwriting concerns correlation, not fairness.

Human credit assessment is bad, slow and inconsistent. Its inconsistency is a feature at the system level. When a thousand loan officers evaluate a thousand borrowers, their errors are imperfectly correlated, and the aggregate portfolio absorbs individual mistakes. When a small number of foundation models, or a large number of systems fine-tuned from a small number of foundation models, evaluate the same borrowers, the errors correlate. A systematic misjudgment about, say, the durability of a particular class of agent revenue does not average out. It propagates to every participant using that assessment, at the same moment, in the same direction.

The financial-network literature has the right frame for the consequence. Acemoglu, Ozdaglar and Tahbaz-Salehi show that densely connected financial networks are robust to small shocks precisely because they share losses, and fragile to large ones because they propagate them; the same structure produces both properties, and the switch between them occurs at a threshold (Acemoglu et al., 2015). Correlated underwriting error is a mechanism for manufacturing large shocks out of ordinary model updating.

The protocol's response cannot be to require better models. It is to require plurality, which I state as a structural rule in Section 9: no participant's access may depend on a single assessor, and the protocol must expose which assessor produced a given determination so that correlation is measurable. This is weaker than a solution. It converts an invisible systemic exposure into a measurable one.

5.4 Discrimination

The second objection is a fairness objection and it is empirically grounded, not hypothetical.

Bartlett, Morse, Stanton and Wallace find that algorithmic lenders in US mortgage markets charged Latinx and Black borrowers higher rates than otherwise-similar borrowers, at a magnitude smaller than face-to-face lenders but not zero, and that the mechanism was price discrimination against borrowers with weaker shopping options and not explicit use of protected characteristics (Bartlett et al., 2022). Fuster and coauthors show that richer machine-learning models in credit markets raise predictive accuracy on average while redistributing outcomes across borrowers, with the gains distributed unequally and some minority borrowers left worse off than under a simpler model, because the added flexibility that improves aggregate fit also lets the model separate borrowers more sharply within groups (Fuster et al., 2022). Neither result depends on malice or on the model observing protected attributes, and neither is fixed by removing those attributes from the feature set.

There is also a formal obstacle. Kleinberg, Mullainathan and Raghavan prove that three natural conditions on a risk score, calibration within each group and balance for the positive and the negative class, which generalize equal false-negative and false-positive rates, cannot hold together except in the degenerate cases of equal base rates or perfect prediction (Kleinberg et al., 2017). Chouldechova establishes the closely related result that calibration is incompatible with equal false-positive and false-negative rates whenever base rates differ, working in the recidivism setting (Chouldechova, 2017). A protocol that mandates a fairness criterion is therefore mandating a choice among incompatible criteria, and doing so at the level of economic infrastructure, for every participant, permanently.

I do not think this is solvable at the protocol layer, and I want to say so plainly rather than gesture at governance. What the protocol can do is narrower and worth doing. It can require that

an adverse determination be attributable to a named assessor. It can require that the determination carry a statement of the grounds sufficient for a reviewer to evaluate it. It can require that the participant may seek a determination from a different assessor without losing accumulated σ_i, which is what portability buys. And it can require an appeal path to a process the determining party does not control. Those four requirements do not make machine credit fair. They make its unfairness observable, contestable and non-terminal, which is the most a protocol can honestly promise and more than the current system provides.

6 Clearing and settlement

6.1 Netting, stated carefully

Fix a clearing cycle and let $x_{ij} \geq 0$ denote the total notional of obligations from i to j eligible for that cycle, aggregated from the edges e_{ij} of $\mathcal{E}_t$ that share a unit of account, a maturity bucket and a settlement venue. Gross exchange is

$$X \;=\; \sum_{i \in V} \; \sum_{j \in V,\; j \neq i} x_{ij}, \tag{7}$$

where the exclusion of $i = j$ matters more than it looks: a self-obligation inflates X and depresses S/X without touching any net position, so admitting it would hand participants a free lever on the reported netting ratio. Section 7 shows that closing this particular hole is not sufficient, because the same lever exists in less trivial form.

The single-unit-of-account restriction is a real limitation and I flag it here instead of in a footnote. Both benchmarks this paper leans on violate it. Every foreign-exchange trade generates obligations in two different currencies, which is precisely what CLS nets, and close-out netting under an ISDA master agreement operates across currencies and products by contract. A model in which x_{ij} is scalar and single-currency therefore does not contain its own empirical anchors. Generalizing x_{ij} to a vector of currency legs with an explicit conversion and finality rule for cross-currency offset is straightforward to state and consequential to get right, since the offset is only as good as the exchange rate convention and the simultaneity of the two legs. I do not develop it here, and the reader should treat the single-currency exposition as an expository simplification whose relaxation is unfinished.

Bilateral netting cancels reciprocal pairs, leaving $b_{ij} = \max\{x_{ij} - x_{ji},\, 0\}$ and a residual $X^{\mathrm{bi}} = \sum_{i,j} b_{ij}$. Multilateral netting works on net positions. Writing $\nu_i = \sum_j x_{ji} - \sum_j x_{ij}$ for participant i's net receipt, we have $\sum_i \nu_i = 0$, and the amount that must actually be funded is the sum of the short positions,

$$S \;=\; \sum_{i \in V} \max\{-\nu_i,\, 0\} \;=\; \tfrac{1}{2} \sum_{i \in V} |\nu_i|. \tag{8}$$

The nettable proportion is then defined residually by equation (1):

$$n_{\max} \;=\; 1 - \frac{S}{X}. \tag{9}$$

Two quantities travel under the name n in this literature and conflating them causes trouble later, so I separate them now. The graph-theoretic quantity just defined, $n_{\max}$, is what arithmetic permits. The operative quantity n_{eff} is what a given legal and prudential arrangement will actually allow to be cancelled, which is smaller whenever offsets are unenforceable, whenever obligations fall

in different netting sets, or whenever a risk rule forbids relying on an offset. Always $n_{\text{eff}} \le n_{\max}$, and the gap between them is the subject matter of the rest of this section. Where a figure appears below I say which of the two it measures.

6.2 Netting cancels transit as well as circulation

It is natural, and wrong, to say that multilateral netting is the operation of discovering and cancelling circulation. Decompose the flow x into circulations plus a remainder of directed paths, which the flow decomposition theorem guarantees is possible. Circulation does contribute nothing to any ν_i and therefore nothing to S. But netting also collapses *transit along paths*: a path carrying value f across ℓ edges contributes $f\ell$ to X while contributing only f to S, because every intermediate participant on the path receives and pays the same amount and nets to zero.

Writing φ for the share of gross flow lying in circulations and $\bar{\ell}_P$ for the flow-weighted mean length of the residual paths, the two effects combine as

$$n_{\max} \;=\; 1 - \frac{1-\varphi}{\bar{\ell}_P}, \qquad \text{so} \qquad n_{\max} \;\ge\; \varphi, \tag{10}$$

with equality exactly when $\bar{\ell}_P = 1$, that is when no residual path has an intermediary. The identification of φ and $\bar{\ell}_P$ depends on which decomposition is chosen, since flow decomposition is not unique; only S, X and $n_{\max}$ are invariants of the graph. Where the split matters it should be pinned down by taking φ to be the maximum circulation, which is a linear program.

The correction is not cosmetic. A pure chain $v_1 \to v_2 \to \cdots \to v_{L+1}$ with every edge equal to Q contains no cycle whatsoever, yet $X = LQ$ and $S = Q$, giving $n_{\max} = 1 - 1/L$, which approaches one as the chain lengthens. Netting efficiency is therefore not a property of cycle structure alone, and the claim that it is would misprice this paper's own worked example, as the next subsection shows.

Paper I's worked example makes both points concretely. With $A \to B : 100$, $B \to C : 80$ and $C \to A : 70$, gross exchange is $X = 250$. There are no reciprocal pairs, so bilateral netting achieves nothing at all: $X^{\text{bi}} = 250$ and $n_{\text{bi}} = 0$. Multilateral netting finds the circulation of 70 around the cycle $A \to B \to C \to A$, leaving $A \to B : 30$ and $B \to C : 10$, with net positions $\nu_A = -30$, $\nu_B = +20$, $\nu_C = +10$. By equation (8), $S = 30$, so

$$n_{\max} \;=\; 1 - \frac{30}{250} \;=\; 0.88.$$

Notice that the circulation share is $\varphi = (70 \times 3)/250 = 0.84$, so an account of netting as circulation-cancellation alone would report 0.84 and be wrong by four points on three nodes. The missing four points are transit: the residual decomposes into a path $A \to B \to C$ carrying 10 across two edges and a path $A \to B$ carrying 20 across one, so $\bar{\ell}_P = 40/30 = 4/3$ and equation (10) gives $1 - 0.16/(4/3) = 0.88$, as it must.

One three-node cycle moves the nettable proportion from zero to 0.88. This is the entire economic content of multilateral clearing, and it is also, as Section 6.6 shows, the entire source of its fragility.

6.3 Clearing under default

Equation (8) assumes everyone pays. When they do not, the netting problem becomes a fixed point. Eisenberg and Noe formulate the clearing payment vector under limited liability and proportional sharing among creditors, and show that a clearing vector exists and is generically unique (Eisenberg

and Noe, 2001). Rogers and Veraart extend the construction to include the costs of default itself, with the consequence that the total value available to creditors depends on which participants fail and in what order (Rogers and Veraart, 2013).

The protocol implication is that a clearing layer needs two distinct procedures. The normal-course procedure computes ν_i and funds the shorts. The default procedure computes a clearing vector, which requires knowing the legal priority of claims and therefore cannot be computed from the graph alone. A design that implements only the first has built a system that works until the first time it matters.

6.4 The empirical record

Table 2 collects the operating figures. These are venue-specific and should not be averaged, because they describe systems with different products, participant sets, legal frameworks and definitions of the denominator.

System	**Scale**	**Reported ratio**	**What the ratio measures, and its relation to n**
CLS (PvP foreign exchange)	More than \$7T average daily settled, 2024–25; record \$19.1T on 20 June 2024, funded with about \$72B	96% reduction; 99% funding reduction	Reduction in gross *payment obligations* by multilateral netting. This is the one row whose headline figure is a netting ratio in the sense of n, giving $n \approx 0.96$. The 99% figure is a *funding* statistic that combines netting with in/out swaps and other liquidity-saving mechanisms, and is not an n value (CLS Group, 2025)
CHIPS (USD large value)	More than 630,000 transactions and in excess of \$2T daily; roughly 95% of payments final within seconds	26:1 liquidity efficiency	Ratio of settled value to intraday funding. This is *turnover*, not netting: it includes recycling of the same prefunded balance many times through the day as incoming payments fund outgoing ones. It bounds n from above but does not identify it (The Clearing House, 2025)
LCH SwapClear compression	Cumulative compressed notional in excess of \$1 quadrillion since 2008	Not a ratio	Trade compression removes offsetting outstanding notional. Cited only as an order-of-magnitude indication that gross notional overstates economic exposure (LCH Group, 2020)

Table 2: Netting, liquidity efficiency and compression in operating wholesale infrastructure. The reported ratios measure different quantities and are not commensurable; only the CLS netting figure is an n value as this paper defines n. Derivatives close-out netting is deliberately absent: it reduces credit exposure, not payment volume and belongs to a different table.

The column headings matter more than the numbers, because the three ratios are routinely quoted together as though they measured one thing. They do not. Netting cancels offsetting

obligations so that they never require funding. Liquidity efficiency measures how many times a unit of funding is reused within a day, which is a statement about velocity and queue design. Credit-exposure reduction measures what a counterparty stands to lose on default, which is neither. Only the first is n. A CHIPS participant funding $1 to settle $26 has not cancelled twenty-five dollars of obligation. Some of the ratio is genuine offset, since the algorithm searches continuously for bilateral and multilateral matches, and the rest is recycling, in which the same dollar is passed along as receipts arrive. The published figure does not separate them, and the recycled portion depends on incoming payments arriving on time, which is the dependency that fails under stress.

The distinction matters because the temptation in a paper like this is to take $n \approx 0.99$ as the design assumption. That number is achieved by a system with about seventy-five settlement members, all of them supervised financial institutions, transacting in eighteen currencies under a rulebook backed by specific legislation in each relevant jurisdiction, in a product where obligations are naturally reciprocal because every foreign-exchange trade generates offsetting flows in two currencies. Almost none of those conditions holds in the general economy.

6.5 What actually determines n

Cycle density and intermediation depth. Equation (10) says $n_{\max}$ rises both with the circulation share φ and with the mean number of intermediaries $\bar{\ell}_P$ on the residual paths. A first instinct is that a general economy nets badly, because value flows one way along a supply chain and terminates in household consumption. That instinct fails on both terms.

It fails on φ because the economy is not acyclic. The circular flow of income is a cycle by construction: firms pay wages to households and households pay firms for output. Take the simplest closed version, with a firm paying wages $100N$ and households returning $100N(1-s)$ in consumption at saving rate s. Then $X = 100N(2-s)$, the firm is short $100Ns$, and

$$n_{\max} \;=\; 1 - \frac{s}{2-s},$$

which is 0.974 at a five percent saving rate. The crudest possible circular-flow economy already nets at a rate comparable to CLS. It fails on $\bar{\ell}_P$ for the opposite reason: the residual that does not circulate passes through intermediaries, and each intermediary nets to zero. A pure chain with L stages and no cycles at all gives $n_{\max} = 1 - 1/L$.

This inverts a conclusion the agentic literature reaches casually and that an earlier draft of this paper reached too. Machine-to-machine production is characterized by *more* disaggregation, not less: tasks that a single firm performed internally become priced exchanges between specialized agents, which lengthens chains and raises $\bar{\ell}_P$. On the arithmetic above that raises netting efficiency. Deep agentic supply chains are good for n.

So the pessimistic case cannot be built on topology, and I do not build it there. It has to be built on the four things below, all of which drive a wedge between $n_{\max}$ and n_{eff}.

Maturity dispersion. Obligations net only if they fall in the same cycle. A wage payable on the first and a grocery payable on the fifteenth are both in the circular flow and neither can offset the other. The theoretical 0.974 above assumes every leg lands in one window; spreading the same flows across dispersed maturities decomposes one large netting set into many small ones, and Proposition 6.1 then says total settlement weakly rises. Paper I of this series proves the corresponding monotonicity directly: refining the netting arrangement weakly increases S, with the limit being gross funding. This has a design implication I return to below, because it is the opposite of what "continuous clearing" naively suggests.

Netting-set fragmentation. Two participants can offset only inside a common enforceable perimeter. Households have no netting agreement with their employers or their grocers; the offsets that exist in the aggregate flow are spread across millions of legal persons who share no rulebook. Foreign exchange is the extreme opposite case, and its advantage is structural: every trade generates two reciprocal legs on the same value date inside one rulebook among a membership of supervised institutions numbering in the dozens. That two-leggedness is a property of the product, not an achievement of the infrastructure, and nothing in the general economy supplies it.

Tiering. Participants who clear through an intermediary net within that intermediary's book and appear to the system as one node. Tiering raises apparent system n_{eff} and concentrates risk in the intermediary. The tiered structure is visible in the measured topology of real large-value payment systems, where a small core intermediates for a large periphery (Soramäki et al., 2007).

Legal enforceability. Netting works because a court will honor it. Close-out netting under an ISDA master agreement is effective because a series of jurisdictions enacted statutes making it effective, over roughly two decades, against the general principle that an insolvency administrator may cherry-pick favorable contracts (International Swaps and Derivatives Association, 2010). Absent such provision, the administrator claims the gross amounts owed to the estate while the counterparty proves for the gross amounts owed to it, and n_{eff} collapses to zero for that participant. Netting is a legal parameter that happens to be implemented in software.

That list yields a design conclusion which cuts against the obvious reading of "continuous clearing." Because refining a netting arrangement weakly raises S, shortening the window over which obligations may offset makes settlement more expensive, not less, and in the limit of instantaneous netting recovers gross funding. An agentic protocol should therefore not claim that machine speed reduces settlement demand by clearing faster. What machine speed buys is the ability to hold a *wide* netting window open while searching it continuously for offsets and releasing obligations as they cancel, so that discharge is continuous while funding stays near net. That is what liquidity-saving mechanisms do, it is the source of the CHIPS turnover figure discussed above, and the submission-timing game it induces is analyzed in Bech and Garratt (2003) and Martin and McAndrews (2008), with cycle-finding algorithms in Bech and Soramäki (2002). The genuine agentic advantage is that software counterparties tolerate a delay of milliseconds where human-facing systems cannot, so the window can be wide relative to the tolerance for delay. Continuous offsetting within a wide window is cheap; a short netting window is provably expensive.

6.6 The fragility of high n

This is the objection I take most seriously, and I think it is underweighted in every optimistic treatment of machine clearing including, until this section, the framing of this series.

Settlement liquidity is a buffer. In a high-n system, participants hold buffers sized to $S = X(1-n)$ under normal conditions. The saving is real and it is the whole point. But n is not a constant of the system; it is a function of the graph's cycle structure, and stress attacks cycle structure directly.

The right formal statement concerns fragmentation of the netting set, and it is short enough to prove.

Proposition 6.1 (Subadditivity of settlement under fragmentation)**.** *Let $\mathcal{A}$ and $\mathcal{B}$ be disjoint sets of obligations and let $S(\cdot)$ be as in equation (8). Then*

$$S(\mathcal{A} \cup \mathcal{B}) \;\leq\; S(\mathcal{A}) + S(\mathcal{B}),$$

with equality if and only if no participant holds net positions of opposite sign in $\mathcal{A}$ and $\mathcal{B}$.

Proof. Net positions are additive across disjoint obligation sets, so $\nu_i^{\mathcal{A}\cup\mathcal{B}} = \nu_i^{\mathcal{A}} + \nu_i^{\mathcal{B}}$. Applying $|a+b| \le |a| + |b|$ termwise to $S = \frac{1}{2}\sum_i |\nu_i|$ gives the inequality, and the triangle inequality is tight exactly when the two terms share a sign for every i. □

Proposition 6.2 (Carve-out)**.** *If participant v's obligations are removed from the netting set and settled gross, then*

$$S(\textit{survivors}) + \text{gross}(v) \;\ge\; S(V),$$

where gross(v) *is the total notional of v's obligations in both directions.*

Proof. Apply Proposition 6.1 with $\mathcal{A}$ the surviving obligations and $\mathcal{B}$ those of v, noting $S(\mathcal{B}) \le \text{gross}(v)$. □

What these say, and what they carefully do not say, is worth spelling out, because the intuition they replace is seductive and false.

They say that splitting a netting set weakly raises the total settlement requirement of the system. The complementary empirical question, which participant's position actually matters for system liquidity, has a structural literature that identifies key players directly rather than by volume (Denbee et al., 2021). That is the fragility result the design needs, it is decomposition-free, and it subsumes the Duffie and Zhu finding cited above: introducing a central counterparty for one product class while bilateral netting continues across classes fragments a netting set, and Proposition 6.1 says fragmentation cannot reduce total settlement (Duffie and Zhu, 2011; Cont and Kokholm, 2014).

They do not say that the survivors' own requirement rises. It is tempting to state a lower bound in terms of the circulation θ routed through the departing participant, so that removal costs the survivors at least θ. That claim is false twice over. It is false because removing a vertex withdraws its receivables along with its payables: take $A \to v : 100$, $v \to B : 100$, $B \to A : 100$, $A \to B : 100$, where $\nu_v = 0$, $\nu_A = -100$, $\nu_B = +100$, so $S = 100$ with a circulation of 100 through v; removing v leaves $A \to B : 100$ and $B \to A : 100$, which cancel exactly, so the survivors' requirement falls to zero and their nettable proportion rises to one. It is false a second time because θ is not a function of the graph at all: the same four edges decompose equally well into a two-cycle between A and B plus a path $A \to v \to B$, under which the circulation through v is zero. A quantity that takes two values on one graph cannot appear in the statement of a proposition. What is true is the bound $|S' - S| \le \sum_{i\neq v} |x_{vi} - x_{iv}|$, which is controlled by v's bilateral net size, and which points the opposite way from the intuition that a small participant is harmless.

Return to the three-node example for the ordinary case. Removing C leaves $A \to B : 100$ with the other two edges in close-out. Among the survivors $X = 100$ and $S = 100$, so $n_{\max}$ has gone from 0.88 to 0, and A's funding requirement has risen from 30 to 100 while B's is zero before and after. Total system settlement, counting C's gross close-out, has risen as Proposition 6.2 requires.

Now compound this. Correlated stress removes participants in groups instead of one at a time. Correlated underwriting error, discussed in Section 5, is a mechanism for producing exactly such groups. And the participants most likely to be removed simultaneously are those with similar business models, which is to say those most likely to sit on the same cycles. The funding requirement therefore jumps by a multiple at exactly the moment when funding markets are least accommodating, against buffers sized for the normal-course n.

This is the reason CLS exists. The precipitating history is specific: the failure of Bankhaus Herstatt in June 1974, whose license was withdrawn after it had received Deutschmarks and before it had paid the corresponding dollars, left counterparties exposed to the full principal amount of trades they thought were nearly complete. The BIS committee report that followed defined

foreign-exchange settlement risk as a principal risk of several days' duration and not an overnight timing inconvenience, and the payment-versus-payment architecture of CLS was built specifically to eliminate it (Committee on Payment and Settlement Systems, 1996). The lesson concerns substitution, not danger. Each settlement architecture trades one exposure for another, and the three trades should not be blurred together. Deferred net settlement buys liquidity saving by *creating* intra-cycle credit exposure, since obligations accumulate unsettled until the cycle closes. Real-time gross settlement eliminates that credit exposure and pays for it in liquidity. Payment-versus-payment, which is what CLS added, eliminates principal risk of the Herstatt kind by making the two legs conditional on each other, and is orthogonal to the first trade-off. A design that reports a liquidity saving without saying which of these it has bought and which it has sold has not described its own risk profile. The incentive consequences of the choice, in particular how settlement architecture shapes banks' risk-taking, are the subject of Kahn and Roberds (1998); the intraday credit-versus-liquidity trade-off is quantified in Furfine and Stehm (1998), and Kahn and Roberds (2009) survey the field.

The clearing literature contains a further warning that bears directly on protocol design. Duffie and Zhu show that introducing a central counterparty for one class of derivatives, alongside existing bilateral netting across classes, can *increase* total collateral demands, because it fragments what was previously a single netting set (Duffie and Zhu, 2011), a mechanism whose empirical magnitude has since been estimated directly (Duffie et al., 2015). Cont and Kokholm qualify that result instead of reinforcing it: allowing asset classes to differ in riskiness and to be correlated, they find the comparison reverses, and central clearing reduces exposures under parameters they regard as realistic (Cont and Kokholm, 2014). Translated into this design: a proliferation of specialized agentic clearing venues, each locally efficient, can lower aggregate n below what a single venue or well-designed bilateral netting would achieve. More clearing is not monotonically better clearing.

6.7 Design responses

None of the following is complete.

First, size buffers to stressed n_{eff} instead of normal-course n_{eff}, and make the stressed calculation a protocol-level disclosure instead of each participant's private modeling exercise. The cost is exactly what it sounds like: much of the liquidity saving that motivates the design. The achievable economy-wide saving is the normal-course saving minus a stress reserve. Section 7 shows that publishing the stressed figure is not free either, because a disclosed ratio becomes a target.

Second, be explicit about what the fallback actually is, because the obvious phrasing is incoherent. Saying a protocol should "degrade to gross settlement instead of failing" presumes the gross obligations still exist. Under *novation* netting they do not: the original obligations are extinguished and replaced by the net, so there is nothing to fall back to. Under *settlement* netting the originals survive until discharge, so gross settlement is available, but recomputing the nets after removing a defaulter is an unwind, and unwind is the single practice that the Lamfalussy standards and the PFMI were written to prohibit, precisely because it retroactively changes every surviving participant's obligation and propagates the failure (Committee on Interbank Netting Schemes of the Central Banks of the Group of Ten Countries, 1990; Committee on Payment and Settlement Systems and Technical Committee of the International Organization of Securities Commissions, 2012; Angelini et al., 1996). The defensible design is neither: novate, and allocate loss *ex ante* through a stated default waterfall, so that no participant's obligation is recomputed after the fact. Requirement 4 should be read that way, and a protocol that cannot say which netting it performs has not specified its own failure behavior.

Third, be explicit about the settlement asset. Paper I's money-quality function $\mathcal{M}(x)$, defined over acceptability, liquidity, credit quality, transferability, settlement finality and stability, is the right frame: settlement assets are graded, not binary, and the hierarchy of money means that one tier's settlement asset is the next tier's credit (Mehrling, 2011, 2012). A protocol that settles in a private issuer's liability has embedded that issuer's credit in every transaction, and the liquidity saving it reports is conditional on that credit holding. Central bank money sits at the top of the hierarchy for a reason, and the current official-sector work on tokenized settlement is explicitly organized around keeping it there (Bank for International Settlements, 2025a).

Fourth, respect the existing rulebook. The CPMI-IOSCO Principles for Financial Market Infrastructures specify what a system that performs netting and settlement is expected to have: a well-founded legal basis, clear rules on finality, credit and liquidity risk management, default procedures, and governance (Committee on Payment and Settlement Systems and Technical Committee of the International Organization of Securities Commissions, 2012). An agentic protocol performing the same economic function should expect to satisfy the same principles, and a proposal that treats them as legacy overhead is proposing to relearn them.

Fifth, and least satisfying, note what the CLS benchmark is conditional on. CLS achieves its ratio by settling in central bank money across eighteen national real-time gross settlement systems, with staged pay-ins and access to central bank liquidity behind each leg. The number is a product of central bank access as much as of clever netting. An open agentic protocol has no analogue for that backstop unless one is deliberately provided, which is one reason the official-sector work on tokenized settlement is organized around keeping central bank money at the apex (Bank for International Settlements, 2025a). A protocol that quotes the CLS figure while settling in a private issuer's liability is quoting a number it has not earned.

7 The endogeneity of n

Everything above treats the obligation graph as given and asks how much of it can be cancelled. That framing hides the most serious objection to this design, and it is an objection specific to the agentic setting.

X is not exogenous. Participants choose whether, when, and with whom to create obligations. The first design response of Section 6 proposes that stressed netting efficiency become a protocol-level disclosure used to size liquidity buffers, which converts n from a measurement into a price. Goodhart's observation applies immediately: a statistical regularity used as a target ceases to be reliable, because the behaviour that produced it adapts to the target (Goodhart, 1975). The cheapest way for an agent to reduce its required buffer is to originate offsetting circular obligations that carry no underlying activity. Section 3 has already argued, with x402 as the example, that the marginal cost of a priced machine transaction is fractions of a cent (Coinbase, 2025). The cost of manufacturing arbitrary circulation is therefore approximately zero, and the agents doing it are optimizers by construction.

Three consequences follow, and the third is the one that matters.

Measured n can be driven arbitrarily close to one by round-trip obligations. Adding a circulation of size c to any graph raises X by $c\ell$ while leaving every ν_i and therefore S unchanged, so $n = 1 - S/X$ rises mechanically toward one as c grows. This is not hypothetical behaviour: wash trading on crypto exchanges, round-tripping in the Enron-era energy markets, and hot-potato volume in interdealer foreign exchange are all instances of gross volume inflated without corresponding economic content, and the last of these inflates the foreign-exchange volumes this paper benchmarks against.

The reported liquidity saving therefore becomes uninformative. A protocol that advertises high n may be reporting real coordination or reporting manufactured circulation, and equation (1) cannot distinguish them, because nothing in the model connects X to real output.

The damaging part is the interaction with fragility. Manufactured cycles are exactly the cycles that vanish first under stress, because a counterparty that stops accepting an agent's obligations kills the synthetic cycle before it touches any real trade. So the strategy that maximizes reported n in normal conditions also maximizes the gap between normal-course and stressed netting efficiency. The paper's two headline results interact destructively: optimizing the first degrades the second, and an agent doing so is behaving rationally.

I do not have a complete answer. A partial one is available and it uses machinery the paper already has. The contract object C_{ij} carries typed inputs and outputs, introduced in Section 4 for the fiscal distinction of Section 8; the same typing can be made to carry the distinction between an obligation backed by a delivery commitment and one that is not. A netting set can then admit only obligations whose counter-performance is specified, which does not eliminate manufactured circulation but raises its cost from approximately zero to the cost of fabricating deliveries. Measured n should in any case be reported alongside a measure of underlying delivery and not alone, and buffer requirements should key off the latter. The honest summary is that this paper's central equation describes the arithmetic of settlement correctly and says nothing about whether the activity being settled is real, and that in an economy of optimizing agents facing near-zero transaction costs, that silence is exploitable.

8 Settlement as the fiscal observation point

Paper III argues that production should not be a taxable event and that revenue should come from land, resource rents, externalities, final consumption and the returns on a public capital endowment W_G. That argument has a protocol consequence, which I state briefly because the substance belongs to Paper III.

The classical result behind Paper III's proposition is Diamond and Mirrlees on production efficiency: under conditions that include the availability of a full set of commodity taxes, optimal policy does not distort aggregate production, which places taxes on final transactions and not on intermediate ones (Diamond and Mirrlees, 1971). The Mirrlees Review's treatment of tax design reaches compatible conclusions about neutrality across the structure of production (Mirrlees et al., 2011). The practical difficulty has always been identifying which transactions are final, since the intermediate-versus-final distinction is a fact about the buyer's purpose.

An obligation graph with typed contracts makes that distinction observable in a way it has not previously been. A contract C_{ij} whose output is an input to another contract in i's production is intermediate by construction; a contract whose output terminates at a household node is final. The settlement layer is therefore a natural observation point: it is where obligations are discharged instead of passed along, and it is where the terminal node of a chain becomes visible.

The distinction that phrasing turns on is worth making explicit, because getting it wrong would hand a referee an easy objection. The settlement layer is where the tax event becomes *observable*; it is not what constitutes the event. What constitutes it is final consumption, which is Paper III's position and the one this paper adopts. The difference matters because the netting perimeter is a choice variable. Section 6 has already shown that the settlement requirement is a property of the netting arrangement and not of the underlying obligations, so a charge keyed to settlement events as such would be avoidable by reorganizing the perimeter. Final consumption is not avoidable that

way. The protocol's contribution is visibility, and the tax base remains defined in Paper III's terms.

There is also a condition on the Diamond and Mirrlees result that this section should not quietly drop. Production efficiency requires that pure profit be fully taxed; where it is not, the theorem's conditions fail and taxes on production can re-enter the second-best optimum (Diamond and Mirrlees, 1971). Paper III's treatment of the consistent instrument, a cash-flow tax or an allowance for corporate equity that reaches rent while exempting the normal return, is what carries that condition, and nothing in this paper's protocol design substitutes for it.

Two honest qualifications remain. The first is that this makes the protocol a fiscal instrument. Once tax observation runs through settlement, control of the settlement layer is control of the tax base, and the stakes of the capture problem in Section 9 rise accordingly. The second is that the observability is only as good as the contract typing, and the incentive to mistype a final purchase as intermediate is exactly the incentive that currently produces value-added-tax fraud. The protocol changes the cost of detection; it does not change the incentive.

9 The constitutional restriction

9.1 Optimization is not authority

The protocol described in this paper cannot function without machine intelligence. Nothing else can underwrite at the required rate, detect fraud across a graph of that size, price risk continuously, or discover cycles in a clearing set with millions of participants. The argument of this section does not call for keeping machines out of the loop. It holds that a system which must be optimized by machines has to be governed by something else.

The distinction is between optimizing within a rule set and holding authority over the rule set. A model may compute a credit spread, propose a clearing schedule, price a guarantee, or flag a fraud pattern. A model may not determine who is admitted to economic participation, whether an appeal succeeds, what the rules of the protocol are, or whether the rules apply to itself. Equation (2) is the compact statement.

The argument for the separation is Kydland and Prescott's, transposed. Their result is that a policymaker optimizing period by period, without the ability to commit, achieves worse outcomes than one bound by a rule, because private agents anticipate the reoptimization (Kydland and Prescott, 1977). The transposition is close, though not exact: an optimizer with authority over the rules it optimizes against has no commitment technology, and participants will act on the knowledge that any rule protecting them can be revised whenever the revision improves the objective. The value of a protocol guarantee lies entirely in its not being subject to the optimizer's discretion.

Constitutional political economy makes the complementary argument. Buchanan and Tullock distinguish the constitutional stage, at which rules are chosen without knowing one's position under them, from the operational stage, at which one acts within them, and the whole point of the distinction is that rules chosen behind that veil have different properties from rules chosen by whoever currently holds the advantage (Buchanan and Tullock, 1962). Ostrom's field evidence on durable commons institutions points at the operational requirements: clearly defined boundaries, participation by the affected in modifying rules, monitoring accountable to participants, graduated sanctions, and accessible conflict-resolution mechanisms (Ostrom, 1990). Those map onto a protocol's governance layer almost item for item, and they are a better guide to designing $\mathcal{G}$ than anything in the distributed-systems literature.

9.2 Six guarantees

I state the guarantees as constraints on $\mathcal{G}$. Paper V constitutionalizes them; here they are protocol requirements with a stated failure mode each.

Identity portability. A participant may move ID_i, including accumulated σ_i, between providers without loss. Failure mode without it: the accumulated settlement history becomes a switching cost, the market tips, and the winner controls participation (Katz and Shapiro, 1985; Farrell and Saloner, 1985).

Underwriting plurality. No participant's access may depend on the determination of a single assessor, and determinations must be attributable to a named assessor. Failure mode without it: a private credit monopoly with the power to exclude, plus the correlated-error exposure of Section 5.

Right of appeal. An adverse determination against a participant carries a path to review by a process the determining party does not control. Failure mode without it: automated exclusion becomes final, which is the specific outcome equation (2) exists to prevent.

Interoperability. Obligations, identities and contracts are portable across implementations, and the specification is available on terms that permit independent implementation. Failure mode without it: a de facto standard owner extracts on the side with the weaker outside option (Rochet and Tirole, 2003).

Privacy floor. A participant discloses predicates, not positions, and the set of parties entitled to full visibility is enumerated in the protocol and not determined by commercial leverage. Failure mode without it: participation requires surrendering the complete record of one's economic life to whoever operates the ledger.

Exit. A participant may cease participation, settle outstanding obligations, and depart with its assets and its record. Failure mode without it: none of the other five guarantees is credible, because a participant who cannot leave cannot refuse.

Exit does the enforcement work. The others are enforceable only if the threat of departure is real, which is Ostrom's point about boundaries (Ostrom, 1990) and Hirschman's about the relationship between exit and voice (Hirschman, 1970). A protocol from which departure is impossible is a jurisdiction, and it should be governed like one.

9.3 Platform capture, and why behavioral remedies fail

The uncomfortable fact about this section is that ordinary economics predicts the outcome it tries to prevent.

The protocol layer has every property associated with tipping. Adoption externalities are strong, since the value of a clearing network rises with participation, and rises faster than linearly because the cycle density that determines n increases with the participant set. Switching costs are high, because settlement history cannot be rebuilt. The market is two-sided, with assessors and participants on opposite sides. Katz and Shapiro's analysis of network externalities and Farrell and Saloner's of standardization both conclude that such markets converge, and that the converged outcome can be inefficient and still stable (Katz and Shapiro, 1985; Farrell and Saloner, 1985). Rochet and Tirole then characterize how a converged platform allocates price across its two sides as a function of relative elasticities (Rochet and Tirole, 2003).

So the guarantees above cannot be behavioral commitments by whoever wins. A commitment that binds only while the committing party finds it convenient is worth its option value, which is zero at the moment it matters. They have to be structural, in the sense that violating them has to be

impossible or self-defeating, not merely prohibited. Portability is structural if identity is anchored in an identifier the provider does not own, which is precisely the argument for decentralized identifiers over network-namespace tokens (Sporny et al., 2022). Plurality is structural if the protocol will not accept a determination that does not name its assessor, because the naming is what makes concentration visible. Interoperability is structural if the specification is under neutral governance, which is the argument for the Linux Foundation's stewardship of A2A and the x402 Foundation's of x402 in place of continued vendor control (Linux Foundation, 2025; Coinbase, 2025).

I do not claim these are sufficient. Structural remedies in network industries erode and need periodic reimposition, and I know of no reason this layer would be exempt. The claim is narrower: behavioral remedies at this layer will not survive contact with the economics, and a design that relies on them has not addressed the problem.

10 Failure modes and open problems

The most serious open problem is the endogeneity of n. Section 7 argues that agents facing near-zero transaction costs can inflate measured netting efficiency with manufactured circulation, that the protocol cannot tell the difference from the graph alone, and that the manufactured portion is exactly what disappears under stress. The proposed remedy, keying netting eligibility and buffer requirements to typed delivery commitments in C_{ij}, raises the cost of the manipulation without eliminating it and has not been modeled. It attacks the interpretation of the central equation, which is why it ranks ahead of the tail asymmetry.

That asymmetry is the second problem. Section 6.6 shows that fragmenting a netting set weakly raises total settlement, so buffers sized to normal-course n_{eff} are mis-sized for a stress in which participants leave the set. The design response is to hold a stress reserve, which consumes much of the benefit. I do not have a mechanism that captures the normal-course efficiency without the tail exposure, and I am not confident one exists.

A third gap is that economy-wide n has never been measured. Every figure in Table 2 comes from a venue with structurally reciprocal flows and a small supervised membership. Section 6 argues that the binding constraint on economy-wide n is netting-set coverage, not cycle scarcity, which means the quantity to be estimated is not a topological property of the economy but the share of gross flow that a given institutional arrangement can bring under a common enforceable netting set with aligned maturities. Nobody has estimated it, because the arrangement does not exist and the data has not been assembled for this purpose. It is not true that the data could not exist: payment system simulators have computed settlement-to-turnover ratios on actual national payment data for two decades, and the topology of a real large-value payment graph has been measured directly (Leinonen and Soramäki, 2003; Soramäki et al., 2007). Those are the obvious starting points, and I decline to guess at a number in advance of them. Constructing a bound from national input-output tables, payment-system data and the existing legal perimeter of enforceable netting is a tractable research project, and it would move the credibility of this design substantially in one direction or the other.

Verification of the world stays the binding constraint, with the consequence developed in Section 4: for the conditions commercial disputes are actually about, no oracle construction produces ground truth, so those conditions route to the recourse field and from there to courts. A machine economy whose disputes resolve at human institutional speed has a throughput ceiling set by that speed.

Cold start for new agents is unresolved. Credit limits must depend on settlement history for Sybil resistance, and Paper II's agents are new. The guarantee mechanism γ_i pushes the problem

to whoever writes the guarantee, and the question of who writes the first guarantee for a class of participant with no track record is exactly the question. This may have no protocol-level answer and may simply be a market that develops or does not.

Fairness has no protocol-level solution. Section 5 establishes that mandating a fairness criterion means choosing among provably incompatible criteria (Kleinberg et al., 2017). The paper's response is procedural only, and I do not want that to read as adequate. It is what a protocol can do, which is different.

Jurisdiction is unaddressed. Requirement 4 says netting must be enforceable against a defaulting participant's estate. Which estate, under which insolvency regime, determined by which conflict-of-laws rule, is a question this paper does not answer for a participant population that is global by construction and for agents whose situs is genuinely unclear. The derivatives industry needed two decades and jurisdiction-by-jurisdiction legislation to resolve the analogous question for the population of institutions trading under master agreements (International Swaps and Derivatives Association, 2010).

Consensus and finality are assumed throughout. I have said nothing about the mechanism by which the protocol's participants agree on state. The impossibility of deterministic consensus with even one faulty process in a fully asynchronous model (Fischer et al., 1985), the fault-tolerance bounds for Byzantine agreement (Lamport et al., 1982; Castro and Liskov, 1999), and the probabilistic finality of proof-of-work constructions (Nakamoto, 2008) all bear on what "settlement finality" can mean here. Reconciling cryptographic finality with legal finality as the PFMI define it (Committee on Payment and Settlement Systems and Technical Committee of the International Organization of Securities Commissions, 2012) is unfinished work and is more than an implementation detail, because a settlement that is probabilistically final and legally final at different moments creates a window that Herstatt-type risk lives in.

Privacy and auditability remain in tension. Selective disclosure resolves the counterparty case. It does not resolve the supervisory case, where an authority needs aggregate visibility, or the anti-money-laundering case, where transfer-of-originator-information requirements presuppose identification (Financial Action Task Force, 2012), or the case where a participant's predicate proofs are true statements about false commitments. The European digital identity framework is one attempt to legislate a position on this trade-off (European Parliament and Council of the European Union, 2024); it is not obviously the right one, and it is certainly not global.

Finally, nothing here establishes that the design is desirable. The paper establishes that the components exist, that the netting economics is real at wholesale scale, and that the extension faces specific institutional obstacles. Whether an economy organized this way is better than the one we have is a question Papers II, III and V argue and this one assumes.

11 Relation to the series, and conclusion

Paper I supplies the object and the principle: the obligation graph $\mathcal{E}_t$, and $M_t = \mathcal{R}(\mathcal{E}_t)$. This paper's contribution to that principle is to observe that $\mathcal{R}$ is not a graph-theoretic operator with an institutional footnote. It is an institutional operator implemented in software. The set of permissible cancellations is fixed by insolvency law, venue rulebooks and finality statutes, and it is therefore determined at the protocol layer. Paper I's M_t is a policy variable, and the policy is made here.

Paper II supplies the participant and the credit mechanism. This paper turns $B_i(t)$ into an identity object with selectively disclosable attestations, and turns the credit-limit function into a market for underwriting instead of a computation performed by one party. The design's most

consequential choice with respect to Paper II is the plurality requirement, which trades some efficiency for the property that no single model decides who produces.

Paper III supplies the fiscal architecture. This paper locates its observation point at settlement and notes the consequence that the protocol becomes a fiscal instrument.

Paper V receives three things. It receives the restriction in equation (2), which becomes Article IX's statement that economic optimization is subordinate to human constitutional authority. It receives the six guarantees of Section 9, which distribute across Article III on open participation, Article VIII on competition and exit, and Article XI on universal protocol access. And it receives an argument about why constitutionalization is necessary, not decorative: the guarantees cannot survive as commercial commitments because platform economics will not let them, so they must be constitutional or they will not exist.

What I would most want a reader to take from this paper is the pair of asymmetries. The first is that verification of computation has become nearly free while verification of the world has not moved, and that every hard problem in machine contracting sits on that line. The second is that netting efficiency is a benefit in the mean and a concentrated exposure in the tail, that these are the same property viewed at different quantiles, and that a design which markets the first without pricing the second is selling a liquidity saving it has not earned.

The equation $S = X(1 - n)$ is true. The question the protocol has to answer is what n is when things are going badly, and this paper does not answer it. It says what would have to be measured, and why the measurement is harder than it looks.

A Protocol object definitions

Collected for reference. These specify content and leave encoding open; a conforming implementation must be able to express each field and to prove the stated properties, and is free to choose representations.

A.1 Identity

$ID_i = (\kappa_i, \alpha_i, \Lambda_i, \mathcal{A}_i, \rho_i, \gamma_i, \sigma_i)$.

κ_i, capabilities: attestations by named issuers that participant i can perform specified classes of production, each carrying issuer identity, scope, validity period and revocation status.

α_i, authorization: a chain of delegations terminating in a natural or legal person, each link specifying the delegated scope, its limits, its expiry, and the signature of the delegating party. An obligation incurred outside the scope of the chain is not enforceable against the terminal principal.

Λ_i, obligations owed: the set of outgoing edges e_{ij} with i as obligor, sufficient to compute i's gross short position by unit of account and maturity bucket.

$\mathcal{A}_i$, claims held: incoming edges with i as obligee.

ρ_i, reputation: assessments by named assessors, each carrying the assessor's identity, the method class, the assessment, and its date. An unattributed reputation value is not a protocol object.

γ_i, guarantees: obligations by third parties, contingent on i's default, each itself an edge in $\mathcal{E}_t$ so that guarantors' aggregate exposure is visible to the clearing layer.

σ_i, settlement history: the record of obligations entered, performed, and settled, countersigned by the counterparties. Portable by Requirement 1.

A.2 Obligation

$e_{ij} = (q, u, t, c, r)$: quantity, unit of account, maturity, conditions, risk. Generated by execution of a contract; consumed by the clearing layer. The risk component r is the assessed loss distribution, attributed to its assessor as in ρ_i.

A.3 Contract

$C_{ij} = (\text{inputs}, \text{outputs}, \text{price}, \text{conditions}, \text{verification}, \text{settlement}, \text{recourse})$.

Inputs and outputs identify what each party supplies and receives, typed sufficiently for the fiscal distinction in Section 8.

Price may be a constant or a function of quantities observable under the verification field.

Conditions state the states of the world in which performance is owed.

Verification states, for each condition, the procedure by which the parties will determine whether it holds, including the identity of any attesting third party and the consequence of that party's unavailability. A condition with no verification procedure is routed to recourse by default.

Settlement names the discharging asset, the venue, and the finality rule.

Recourse names the forum, the governing law, the allocation of loss on non-performance, and the appeal path. A contract with an empty recourse field is not a conforming protocol object.

A.4 Protocol

$\Pi = (L_{\text{cap}}, L_{\text{obl}}, L_{\text{trust}}, L_{\text{clear}}, L_{\text{settle}}; \mathcal{I}, \mathcal{G})$, with $\mathcal{G}$ satisfying the six guarantees of Section 9 and not modifiable by any participant acting alone, including any participant that is a model.

B Notation

Symbol	Meaning
$\mathcal{E}_t = (V, E_t)$	The obligation graph at time t; V contains humans, firms, governments, machines and AI agents (Paper I)
$e_{ij} = (q, u, t, c, r)$	An obligation from i to j: quantity, unit of account, maturity, conditions, risk (Paper I)
$M_t = \mathcal{R}(\mathcal{E}_t)$	Residual Settlement Principle; $\mathcal{R}$ is the residual after permissible clearing and netting (Paper I)
$\mathcal{M}(x)$	Money-quality of claim x, over acceptability, liquidity, credit quality, transferability, settlement finality, stability (Paper I)
$B_i(t) = (A_i, L_i, C_i, R_i, P_i)$	Agent balance sheet: claims, obligations, productive capabilities, reputation, future productive capacity (Paper II)
$C_i(t)$	Economic capacity of participant i (Paper II). Distinct from C_{ij}
W_G	Public capital endowment (Paper III)
ID_i	Economic identity, equation (4)
C_{ij}	Machine-readable contract, equation (5)
Π	The protocol, equation (3)
$\mathcal{I}$, $\mathcal{G}$	Identity substrate; governance constraint
x_{ij}	Gross notional of obligations from i to j in a clearing cycle

Symbol	Meaning
X	Gross exchange, $\sum_{i,j} x_{ij}$
ν_i	Net position of participant i in a clearing cycle
S	Settlement requirement, $S = X(1-n)$
n	Safely nettable proportion, $n = 1 - S/X$
θ	Magnitude of a circulation in the flow decomposition
$\kappa_i, \alpha_i, \Lambda_i, \mathcal{A}_i, \rho_i, \gamma_i, \sigma_i$	Components of ID_i: capabilities, authorization, obligations owed, claims held, reputation, guarantees, settlement history

C Notes on the netting figures

Because the argument of Section 6 rests on these numbers, the derivations are stated explicitly.

CLS. The operator reports that multilateral netting reduces gross payment obligations by approximately 96 percent, and that with liquidity-saving mechanisms the total funding required falls by roughly 99 percent. The record settlement day of 20 June 2024 settled $19.1 trillion against approximately $72 billion of funding (CLS Group, 2024), giving $1 - 72/19{,}100 \approx 0.9962$. The 0.96 figure is the netting component alone and is the one comparable to the definition of n used in this paper; the 0.996 figure includes mechanisms beyond netting and is reported here for completeness and not as a netting ratio (CLS Group, 2025).

CHIPS. The operator reports intraday liquidity efficiency of approximately 26:1, It is arithmetically tempting to set $S/X = 1/26$ and report $n \approx 0.962$, and this paper does not do so. The ratio compares settled value to intraday funding, and the mechanism producing it combines two distinct effects: genuine netting, which cancels obligations, and intraday recycling, in which a single unit of prefunded balance is reused repeatedly as receipts arrive and fund subsequent payments. Recycling raises the ratio without cancelling anything. The published figure therefore bounds n from above but does not identify it, and the netting component alone cannot be recovered from it. I report the ratio as a turnover statistic and use only the CLS figure as an n value (The Clearing House, 2025).

OTC derivatives. BIS statistics for end-June 2025 report gross notional of $846 trillion, gross market value of $21.8 trillion, and gross credit exposure after legally enforceable netting well below gross market value (Bank for International Settlements, 2025c). Two cautions about this figure. The base is gross market value, not notional, so quoting the percentage against the $846 trillion notional would assert a reduction roughly forty times larger than the one measured. And the percentage itself is not stated in the BIS half-yearly release, which reports notional and gross market value in text and the netting effect only graphically; the underlying series is in the BIS derivatives statistics tables, and a paper relying on the exact figure should cite the table and not the release (Bank for International Settlements, 2025b). I therefore use the direction and order of magnitude of the netting effect and not a precise percentage. Neither figure is an n value: they measure credit exposure and not payment volume, which is why the derivatives row does not appear in Table 2.

LCH compression. Cumulative compressed notional in excess of $1 quadrillion since 2008 is a measure of outstanding notional removed by trade compression, not a per-cycle netting ratio, and it is cited in this paper only as an order-of-magnitude indication that gross notional overstates economic exposure by a large factor (LCH Group, 2020).

Worked example. The $n = 0.88$ figure in Section 6 is computed by the author from Paper I's three-node example: $X = 100 + 80 + 70 = 250$, net positions $(-30, +20, +10)$, $S = 30$, $n = 1 - 30/250 = 0.88$. The post-removal figure of $n = 0$ is computed on the surviving two-node

graph as described in the text.

References

Daron Acemoglu, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608, 2015.

George A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3):488–500, 1970.

Iñaki Aldasoro and Ajit Desai. AI agents for cash management in payment systems. BIS Working Papers 1310, Bank for International Settlements, 2025.

Paolo Angelini, Giuseppe Maresca, and Daniela Russo. Systemic risk in the netting system. *Journal of Banking and Finance*, 20(5):853–868, 1996.

Bank for International Settlements. The next-generation monetary and financial system. Chapter iii, bis annual economic report 2025, Bank for International Settlements, Basel, June 2025a.

Bank for International Settlements. OTC derivatives statistics, table D5.2: Gross credit exposure. BIS Data Portal, `https://data.bis.org/`, 2025b. Semiannual OTC derivatives statistics; accessed August 2026.

Bank for International Settlements. OTC derivatives statistics at end-june 2025. Technical report, Bank for International Settlements, Basel, December 2025c. Statistical release, `https://www.bis.org/publ/otc_hy2512.htm`.

Robert Bartlett, Adair Morse, Richard Stanton, and Nancy Wallace. Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*, 143(1):30–56, 2022.

Morten L. Bech and Rod Garratt. The intraday liquidity management game. *Journal of Economic Theory*, 109(2):198–219, 2003.

Morten L. Bech and Kimmo Soramäki. Gridlock resolution in interbank payment systems. Bank of Finland Discussion Papers 9/2002, Bank of Finland, 2002.

Eli Ben-Sasson, Alessandro Chiesa, Christina Garman, Matthew Green, Ian Miers, Eran Tromer, and Madars Virza. Zerocash: Decentralized anonymous payments from bitcoin. In *Proceedings of the 2014 IEEE Symposium on Security and Privacy*, pages 459–474, 2014.

James M. Buchanan and Gordon Tullock. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. University of Michigan Press, Ann Arbor, 1962.

Vitalik Buterin. Ethereum: A next-generation smart contract and decentralized application platform. White paper, `https://ethereum.org/en/whitepaper/`, 2014.

Giulio Caldarelli. Understanding the blockchain oracle problem: A call for action. *Information*, 11 (11):509, 2020. doi: 10.3390/info11110509.

Jan Camenisch and Anna Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In *Advances in Cryptology — EUROCRYPT 2001*, volume 2045 of *Lecture Notes in Computer Science*, pages 93–118. Springer, 2001.

Miguel Castro and Barbara Liskov. Practical byzantine fault tolerance. In *Proceedings of the Third Symposium on Operating Systems Design and Implementation (OSDI '99)*, pages 173–186, 1999.

David Chaum. Security without identification: Transaction systems to make big brother obsolete. *Communications of the ACM*, 28(10):1030–1044, 1985.

Alexandra Chouldechova. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data*, 5(2):153–163, 2017.

CLS Group. CLS settles record $19.1 trillion in a single day. CLS Group press release, July 2024, 2024. Source for the record-day settled value and the approximately $72 billion of funding required; accessed August 2026.

CLS Group. Annual report 2024. Technical report, CLS Group Holdings AG, Zurich, 2025.

Coinbase. x402: An open standard for internet-native payments over HTTP. `https://www.x402.org/`, 2025. Released May 2025; x402 Foundation announced September 2025.

Committee on Interbank Netting Schemes of the Central Banks of the Group of Ten Countries. Report of the committee on interbank netting schemes. Technical report, Bank for International Settlements, Basel, November 1990. The Lamfalussy Report.

Committee on Payment and Settlement Systems. Settlement risk in foreign exchange transactions. Technical report, Bank for International Settlements, Basel, March 1996. The Allsopp Report.

Committee on Payment and Settlement Systems and Technical Committee of the International Organization of Securities Commissions. Principles for financial market infrastructures. Technical report, Bank for International Settlements and IOSCO, Basel, April 2012.

Rama Cont and Thomas Kokholm. Central clearing of OTC derivatives: Bilateral vs multilateral netting. *Statistics and Risk Modeling*, 31(1):3–22, 2014.

Primavera De Filippi and Aaron Wright. *Blockchain and the Law: The Rule of Code*. Harvard University Press, Cambridge, MA, 2018.

Edward Denbee, Christian Julliard, Ye Li, and Kathy Yuan. Network risk and key players: A structural analysis of interbank liquidity. *Journal of Financial Economics*, 141(3):831–859, 2021.

Douglas W. Diamond. Reputation acquisition in debt markets. *Journal of Political Economy*, 97(4): 828–862, 1989.

Douglas W. Diamond. Monitoring and reputation: The choice between bank loans and directly placed debt. *Journal of Political Economy*, 99(4):689–721, 1991. doi: 10.1086/261775.

Peter A. Diamond and James A. Mirrlees. Optimal taxation and public production i: Production efficiency. *American Economic Review*, 61(1):8–27, 1971.

Darrell Duffie and Haoxiang Zhu. Does a central clearing counterparty reduce counterparty risk? *Review of Asset Pricing Studies*, 1(1):74–95, 2011.

Darrell Duffie, Martin Scheicher, and Guillaume Vuillemey. Central clearing and collateral demand. *Journal of Financial Economics*, 116(2):237–256, 2015.

Larry Eisenberg and Thomas H. Noe. Systemic risk in financial systems. *Management Science*, 47 (2):236–249, 2001.

European Parliament and Council of the European Union. Regulation (EU) 2024/1183 amending regulation (EU) no 910/2014 as regards establishing the european digital identity framework. Official Journal of the European Union, L series, 30 April 2024, 2024.

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

Financial Action Task Force. International standards on combating money laundering and the financing of terrorism and proliferation: The FATF recommendations. Technical report, FATF/OECD, Paris, 2012. Recommendation 16 (wire transfers); as updated.

Michael J. Fischer, Nancy A. Lynch, and Michael S. Paterson. Impossibility of distributed consensus with one faulty process. *Journal of the ACM*, 32(2):374–382, 1985.

Scott Freeman. The payments system, liquidity, and rediscounting. *American Economic Review*, 86 (5):1126–1138, 1996.

Craig H. Furfine and Jeff Stehm. Analyzing alternative intraday credit policies in real-time gross settlement systems. *Journal of Money, Credit and Banking*, 30(4):832–848, 1998.

Andreas Fuster, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther. Predictably unequal? the effects of machine learning on credit markets. *Journal of Finance*, 77(1):5–47, 2022.

Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof systems. *SIAM Journal on Computing*, 18(1):186–208, 1989.

Charles A. E. Goodhart. Problems of monetary management: The UK experience. *Papers in Monetary Economics (Reserve Bank of Australia)*, 1, 1975.

Google LLC. Agent payments protocol (AP2) specification. `https://ap2-protocol.org/`, 2025. Announced 16 September 2025; specification revised through v0.2, accessed August 2026.

Sanford J. Grossman and Oliver D. Hart. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94(4):691–719, 1986.

Michael M. Güntzer, Dieter Jungnickel, and Matthias Leclerc. Efficient algorithms for the clearing of interbank payments. *European Journal of Operational Research*, 106(1):212–219, 1998.

Gillian K. Hadfield and Andrew Koh. An economy of AI agents. In *The Economics of Transformative AI*. University of Chicago Press for the National Bureau of Economic Research, Chicago, 2025. Also circulated as arXiv:2509.01063.

Oliver Hart and John Moore. Incomplete contracts and renegotiation. *Econometrica*, 56(4):755–785, 1988.

Albert O. Hirschman. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Harvard University Press, Cambridge, MA, 1970.

International Swaps and Derivatives Association. The importance of close-out netting. ISDA Research Notes 1, ISDA, 2010.

Charles M. Kahn and William Roberds. Payment system settlement and bank incentives. *Review of Financial Studies*, 11(4):845–870, 1998.

Charles M. Kahn and William Roberds. Why pay? an introduction to payments economics. *Journal of Financial Intermediation*, 18(1):1–23, 2009.

Michael L. Katz and Carl Shapiro. Network externalities, competition, and compatibility. *American Economic Review*, 75(3):424–440, 1985.

Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. Inherent trade-offs in the fair determination of risk scores. In *Proceedings of the 8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, pages 43:1–43:23, 2017.

Narayana R. Kocherlakota. Money is memory. *Journal of Economic Theory*, 81(2):232–251, 1998.

Finn E. Kydland and Edward C. Prescott. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85(3):473–491, 1977. doi: 10.1086/260580.

Leslie Lamport, Robert Shostak, and Marshall Pease. The byzantine generals problem. *ACM Transactions on Programming Languages and Systems*, 4(3):382–401, 1982.

LCH Group. SwapClear compression services: Cumulative compressed notional. LSEG Post Trade, SwapClear compression service statistics, 2020. The "$1 quadrillion since 2008" framing dates from LCH reporting around 2019–2020; LSEG's current SwapClear pages quote a different cumulative basis, so the figure is used here only as an order of magnitude.

Harry Leinonen and Kimmo Soramäki. Simulating interbank payment and securities settlement mechanisms with the BoF-PSS2 simulator. Bank of Finland Discussion Papers 23/2003, Bank of Finland, 2003.

Linux Foundation. Agent2agent (A2A) protocol. `https://a2a-protocol.org/`, 2025. Introduced by Google April 2025; contributed to the Linux Foundation June 2025.

Stewart Macaulay. Non-contractual relations in business: A preliminary study. *American Sociological Review*, 28(1):55–67, 1963.

Antoine Martin and James McAndrews. Liquidity-saving mechanisms. *Journal of Monetary Economics*, 55(3):554–567, 2008.

Mastercard. Mastercard agent pay: Agentic tokens for AI-initiated commerce. Mastercard Newsroom, 29 April 2025, 2025.

Perry Mehrling. *The New Lombard Street: How the Fed Became the Dealer of Last Resort*. Princeton University Press, Princeton, NJ, 2011.

Perry Mehrling. The inherent hierarchy of money. Working paper, Barnard College, Columbia University, 2012.

James Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba. *Tax by Design: The Mirrlees Review*. Oxford University Press for the Institute for Fiscal Studies, Oxford, 2011.

Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. White paper, `https://bitcoin.org/bitcoin.pdf`, 2008.

Elinor Ostrom. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, Cambridge, 1990.

Jean-Charles Rochet and Jean Tirole. Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029, 2003.

L. C. G. Rogers and Luitgard A. M. Veraart. Failure and rescue in an interbank network. *Management Science*, 59(4):882–898, 2013.

Julio J. Rotemberg. Minimal settlement assets in economies with interconnected financial obligations. *Journal of Money, Credit and Banking*, 43(01):81–108, 2011.

Kimmo Soramäki, Morten L. Bech, Jeffrey Arnold, Robert J. Glass, and Walter E. Beyeler. The topology of interbank payment flows. *Physica A: Statistical Mechanics and its Applications*, 379 (1):317–333, 2007.

Tobin South, Samuele Marro, Thomas Hardjono, Robert Mahari, Cedric Deslandes Whitney, Dazza Greenwood, Alan Chan, and Alex Pentland. Authenticated delegation and authorized AI agents. arXiv:2501.09674, 2025.

Manu Sporny, Dave Longley, Markus Sabadello, Drummond Reed, Orie Steele, and Christopher Allen. Decentralized identifiers (DIDs) v1.0: Core architecture, data model, and representations. W3c recommendation, World Wide Web Consortium, July 2022. 19 July 2022, `https://www.w3.org/TR/did-1.0/`.

Manu Sporny, Dave Longley, David Chadwick, and Ivan Herman. Verifiable credentials data model v2.0. W3c recommendation, World Wide Web Consortium, May 2025. W3C Recommendation, 15 May 2025. Editors: Sporny, Thibodeau, Herman, Cohen, Jones. `https://www.w3.org/TR/vc-data-model-2.0/`.

Joseph E. Stiglitz and Andrew Weiss. Credit rationing in markets with imperfect information. *American Economic Review*, 71(3):393–410, 1981.

Nick Szabo. Formalizing and securing relationships on public networks. *First Monday*, 2(9), 1997.

The Clearing House. CHIPS: Real-time final payments. `https://www.theclearinghouse.org/payment-systems/chips`, 2025. Volume and liquidity-efficiency statistics accessed August 2026.

Jean Tirole. Incomplete contracts: Where do we stand? *Econometrica*, 67(4):741–781, 1999. doi: 10.1111/1468-0262.00052.

Nenad Tomasev, Matija Franklin, Joel Z. Leibo, Julian Jacobs, William A. Cunningham, Iason Gabriel, and Simon Osindero. Virtual agent economies. arXiv:2509.10147, 2025.

Visa. Visa intelligent commerce: Enabling AI agents to transact with tokenized credentials. Visa Newsroom, 30 April 2025, 2025.

W3C Decentralized Identifier Working Group. Decentralized identifiers (DIDs) v1.1. W3c candidate recommendation snapshot, World Wide Web Consortium, March 2026. 5 March 2026, `https://www.w3.org/TR/did-1.1/`.

Kevin Werbach and Nicolas Cornell. Contracts ex machina. *Duke Law Journal*, 67(2):313–382, 2017.

Yingxuan Yang, Ying Wen, Jun Wang, and Weinan Zhang. Agent exchange: Shaping the future of AI agent economics. arXiv:2507.03904, 2025.

Fan Zhang, Ethan Cecchetti, Kyle Croman, Ari Juels, and Elaine Shi. Town crier: An authenticated data feed for smart contracts. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16)*, pages 270–282, 2016.