Implementing Stable Matching under Private Information through Trust-Minimized Bargaining Protocols

Xiyu Li

Advisor: Axel Anderson

Department of Economics, Georgetown University
Field: Microeconomic Theory — Matching & Dynamic Games

Ethereum Foundation PhD Fellowship Interview

Background: what is a matching market?

A **matching market** pairs agents on two sides who must transact *with each other*, where *who pairs with whom* — and on what terms — is the economic outcome.

Classic examples

- workers & firms (labor)
- borrowers & lenders (credit)
- buyers & sellers (exchange)

On-chain, today

- takers & liquidity providers (DEX)
- borrowers & suppliers (lending pools)
- members & coalitions (DAOs)

Two features make these markets hard — and make them **the natural application layer for decentralized trust**:

- **Private information.** An agent's value to a partner (skill, risk, type) is not fully observed.
- **No auctioneer.** The pairing itself must be negotiated by individuals.

Stability

The central solution concept: a matching is *stable* if no pair would rather abandon their assigned partners and match with each other. Stability is the matching analog of a "no profitable deviation" equilibrium.

Motivation

A strand of recent literature studies **stable and efficient matching under private information**:

- Liu (2020); Liu, Mailath, Postlewaite & Samuelson (2014)
- Chen & Hu (2024); Hu (2025); Kamishiro, Serrano & Vohra (2023)

It tells us **what** outcomes are stable. It is mostly silent on how we reach that outcome:

The open question

Is it possible for privately-informed agents to reach a stable matching?
If so, by what process?

This is a question about the *mechanism of formation*, not about the properties of the endpoint.

One answer: a centralized clearinghouse

The natural first answer is a **central designer** who collects everyone's reports and computes a stable matching. But it has defects. Firstly, it require people to **truthfully** report.

Defect 1: Incentive compatibility

Under incomplete information, requiring the mechanism to be *stable*, *incentive-compatible*, and *efficient* at once is generically impossible as stated in Liu (2020).

- A worker who overstates her type distorts the matching.
- Making truth-telling optimal forces the designer to leave informational rents or to sacrifice stability.

One answer: a centralized clearinghouse

The clearinghouse also assumes a designer who **faithfully runs the rules**. In reality, whoever operates it has **private incentives**.

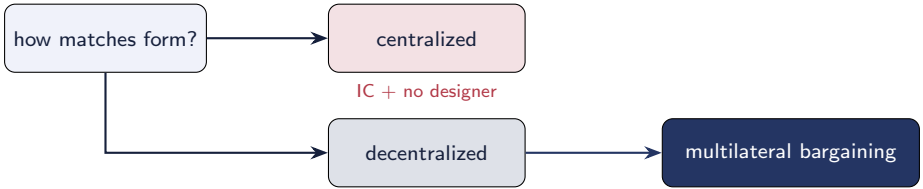
Defect 2: No benevolent designer

- extracts rents from the agents it matches;
- favors fee-paying or preferred participants;
- conceals or reorders information to its own advantage.

Farboodi–Jarosch–Menzio–Wiriadinata (2025): intermediation can be *pure rent extraction* — privately profitable, yet socially wasteful. A truly "benevolent clearinghouse" is a fiction.

Our answer: decentralized implementation

We take the **decentralized** route: no designer, no report. The matching forms through agents' own **bilateral interactions**.



Protocol: Trust-minimized decentralized bargaining (Elliott & Nava, 2019).

- a random proposer makes a directed offer (counterparty + transfer);
- the counterparty accepts or rejects; play continues; decision being exposed to public;
- types are revealed endogenously through *screening*, not reported.

No truthful-reporting requirement: incentives are handled by sequential rationality, not by an IC constraint imposed from outside.

What the bargaining protocol delivers

Target: Bayesian stable matching (Liu, 2020) — no pair (w_i, f_j) can credibly block, where beliefs are Bayes-consistent with the observed matching record.

Equilibrium concept: Markov Perfect Bayesian Equilibrium.

Key Question

As bargaining frictions vanish ($\delta \to 1$), when does the MPBE of the decentralized protocol converge to a Bayesian stable configuration?

My preliminary work suggests *convergence*, under well-behaved surplus functions and other standard assumptions.

Why the Ethereum ecosystem is essential

The decentralized protocol is only *well-defined* on infrastructure that guarantees three things — exactly what a blockchain provides and a centralized platform cannot.

Protocol needs	**Why**	**Ethereum primitive**
Unbiased proposer selection	Payoffs depend on proposer probabilities (p_k)	VRF / RANDAO
Non-discretionary execution	Screening needs offers/transfers to execute as announced	smart contracts
Tamper-proof history	Bayesian beliefs update from public record of offers & rejections	immutable ledger

The constitutive claim

Decentralized trust is not decorative here. These stability outcomes *only exist* on a substrate where the matching history cannot be rewritten and execution cannot be manipulated.

Research plan

1. **Formalize matching on trust-minimized infrastructure.**
2. **Answer the implementation question.**
3. **Quantify the gains from decentralization.**

Contribution

Answer the literature's"how a stable matching could be implemented? "open question with a decentralized protocol, and show that the answer is **native to the Ethereum ecosystem**.

Timeline & closing

Months 1–2	**Model setup.** Formal model with decentralized-trust primitives; literature positioning.
Months 3–5	**Implementation analysis.** Convergence to Bayesian stability; what is clean, what stays open.
Months 6–8	**Centralized counterpart.** Intermediary with private incentives; welfare & rent-extraction wedge.
Months 9–10	**Dissemination.** Present at matching-theory and Ethereum-research venues; explore applications.

Primary deliverable

A working paper answering how stable matches form under private information — through a decentralized bargaining protocol whose guarantees are native to the Ethereum ecosystem.

Thank you. I look forward to your questions.