Protocol Risk and Compute Arms Races in AI-Agent OTC Bargaining

Inkee Jang* Hyoung-Goo Kang†

Abstract

This paper studies bilateral OTC bargaining delegated to AI agents. AI delegation creates two technology choices with opposite welfare properties: *protocol intensity*, which reduces costly misunderstanding of negotiated terms, and *outside-option compute*, which improves fallback market execution and strengthens disagreement payoffs. In a CARA-normal Nash bargaining model, protocol investment is privately underprovided because the protocol investor pays the full cost but captures only a bargaining share of the joint surplus. Outside-option compute is privately overprovided because it improves negotiated terms even when fallback execution is rarely realized. The interaction can inefficiently migrate trade from OTC to centralized venues by reducing the market-friction wedge that makes OTC valuable. These wedges survive more general protocol-loss and market-loss schedules under a monotone single-crossing condition. The model yields comparative statics for protocol adoption, compute expenditure, and OTC-to-market migration, and highlights the welfare role of shared validation standards in AI-mediated finance.

Keywords: AI agents; OTC bargaining; protocol risk; outside options; compute arms race; market design.
JEL: C78, D82, G12, G14.

1 Introduction

OTC markets allocate risk and liquidity through bilateral negotiation, dealer intermediation, and relationship-specific execution arrangements (Duffie, 2012; Duffie et al., 2005). As trading workflows become increasingly automated, a growing share of the negotiation stack can be delegated to

*Hanyang University. Email: inkeejang@hanyang.ac.kr.
†Hanyang University. Email: hyoungkang@hanyang.ac.kr.

AI-enabled systems: agents parse requests for quotes, translate contract terms into machine-readable fields, validate tool outputs, recommend execution routes, and fall back to centralized venues when bilateral execution fails. This changes the investment problem around OTC trade. A counterparty can spend resources on protocols that make bilateral agreement safer, or it can spend resources on fallback execution technology that improves its disagreement payoff. The central question is how these two technology margins interact in bargaining.

This setting is not merely a relabeling of standard algorithmic trading. Existing electronic-trading standards, including FIX-style machine-readable trading protocols, show why structured messages and interoperability matter for reliable trade processing (FIX Trading Community, 2026). AI-agent workflows add a further layer: natural-language instructions and model outputs must be converted into schemas, tool calls, confirmations, and external actions. Errors can arise from mis-specified terms, failed validation, tool outages, or autonomous policies optimized for private objectives. Recent policy and research discussions of AI in finance emphasize precisely these issues: AI systems are moving from prediction and workflow support toward tool-using and action-taking architectures, while regulators have raised concerns about model risk, governance, third-party dependencies, market correlations, and conflicts of interest in technology-mediated financial services (Aldasoro et al., 2024; Financial Stability Board, 2024; International Organization of Securities Commissions, 2021, 2025; U.S. Securities and Exchange Commission, 2023).

I develop a tractable model that separates two AI-specific technology choices. The first is *protocol intensity*: structured schemas, verification layers, redundancy, or reconciliation steps that reduce the variance of misinterpreted OTC terms. Protocol intensity is jointly productive because both parties benefit from lower execution risk, but in the baseline the buyer bears the cost as the workflow initiator. The second is *outside-option compute*: routing, impact-control, venue-search, or execution-optimization resources that improve each party's fallback market execution. This compute is privately valuable not only when fallback execution occurs, but also because it strengthens the party's disagreement payoff in OTC bargaining.

The model combines CARA-normal preferences with Nash bargaining over the OTC price. If OTC succeeds, trade occurs at the interpreted bilateral price; if execution fails, trade defaults to the centralized market. The setup yields closed-form expressions for the negotiated price, privately chosen protocol intensity, privately chosen fallback compute, and the threshold for activating OTC trade. The closed form makes transparent a simple wedge:

protocol investment expands the bilateral pie, while outside-option compute changes both the size of that pie and its bargaining division. I also show that the underprovision and overprovision wedges do not depend on the benchmark inverse-protocol-loss and hyperbolic-impact specifications.

Three results follow. First, protocol intensity is underprovided relative to a planner because the protocol investor internalizes only its bargaining share of the joint benefit. Second, outside-option compute is overprovided relative to a planner because it acts as bargaining leverage even when fallback execution is rarely realized. Third, compute can crowd out OTC trade by shrinking the market-friction wedge that OTC would otherwise save. In this sense, AI technology can produce a protocol-compute tension: better fallback execution may make bilateral protocol activation privately unattractive even when coordinated protocol provision would be efficient.

The paper contributes to three literatures. It adds to bargaining models with outside options by endogenizing the technology that determines disagreement payoffs (Binmore et al., 1986; Nash, 1950; Osborne & Rubinstein, 1990; Rubinstein, 1982). It contributes to OTC and market-microstructure research by showing how execution technology can alter the boundary between bilateral and centralized trade (Almgren & Chriss, 2000; Glosten & Milgrom, 1985; Kyle, 1985; Madhavan, 2000). Finally, it links trading-technology arms races to AI-agent governance by separating shared protocol standards from strategic compute investment (Brogaard, 2014; Budish et al., 2015; Hendershott et al., 2011). Methodologically, protocol investment underprovision relates to relationship-specific investment and hold-up (Grossman & Hart, 1986; Grout, 1984; Hart & Moore, 1990), while outside-option compute resembles rent-seeking/contest expenditure (Hirshleifer, 1991; Konrad, 2009; Tullock, 1980).

The remainder of the paper proceeds as follows. The model first derives the certainty equivalents, Nash price, protocol choice, compute choice, and activation threshold. It then characterizes candidate activation regimes, adds governance price bands and functional-form robustness, and closes with comparative statics, observables, institutional implications, and limitations.

2 Model

2.1 Preferences and payoff risk

Two principals $i \in \{B, S\}$ (buyer, seller) trade a fixed quantity $\Delta > 0$ of a risky payoff v at date T. Preferences are CARA:

$$U_i(W) = -\exp(-\gamma_i W), \qquad \gamma_i > 0.$$

Each principal starts with wealth w_i and inventory q_i in the risky payoff:

$$W_i^0 = w_i + q_i v.$$

At the bargaining stage, common information implies

$$v \mid \mathcal{I}^{\text{pub}} \sim \mathcal{N}(\mu, \sigma^2),$$

interpretable as a shared mark (e.g., public mid plus common model inputs).

2.2 Two venues and an endogenous financial outside option

If the OTC negotiation does not execute, the trade executes in a centralized market. Market execution entails (i) a fixed per-unit wedge (fees/half-spread) and (ii) size-dependent impact (Almgren & Chriss, 2000; Glosten & Milgrom, 1985; Kyle, 1985).

The buyer pays per-unit

$$\mu + \tau_B + \alpha_B(e_B)\Delta,$$

and the seller receives per-unit

$$\mu - \tau_S - \alpha_S(e_S)\Delta,$$

where $\tau_i \geq 0$ is an exogenous per-unit cost component and $\alpha_i(e_i)$ is the impact coefficient reduced by *outside-option compute* $e_i \geq 0$:

$$\alpha_i(e_i) = \frac{\bar{\alpha}_i}{1 + e_i}, \qquad \bar{\alpha}_i > 0. \tag{1}$$

Compute costs are linear:

$$C_i(e_i) = \kappa_i e_i, \qquad \kappa_i > 0. \tag{2}$$

Compute choices are observable at the bargaining stage and are sunk before the OTC activation decision. They are interpreted as agent infrastructure, execution analytics, and venue-search capability that improve fallback execution but are not contractible in the bilateral OTC agreement.

2.3 Protocol risk in OTC and protocol intensity

In OTC, the intended offer price p is misinterpreted as

$$\tilde{p} = p + \xi, \qquad \xi \sim \mathcal{N}\left(0, \frac{\sigma_\xi^2}{b}\right),$$

where $\sigma_\xi^2 > 0$ and $b > 0$ is *protocol intensity* (redundancy, structured schemas, verification tokens). Higher b reduces misinterpretation variance. The party implementing the protocol pays

$$K(b) = \lambda \frac{b^{1+\nu}}{1+\nu}, \qquad \lambda > 0,\ \nu > 0. \tag{3}$$

For closed-form exposition, assume the buyer bears $K(b)$ (e.g., initiator/first mover). This is an initiator-pays benchmark: the qualitative underinvestment result extends whenever protocol costs are borne more privately than protocol benefits. Cost sharing through common standards, certification, or venue-provided validation would narrow the wedge between private and planner protocol intensity.

Remark 1 (Protocol transfer risk)**.** *The price error ξ is a transfer between buyer and seller, so aggregate cash flows cancel ex post. It is nevertheless a welfare cost because risk-averse principals bear uninsurable transfer risk before the realization of the interpreted price. Equivalently, the reduced-form noise can be read as mis-specified contract terms, reconciliation costs, or settlement attributes that are locally isomorphic to a risky transfer under CARA-normal certainty equivalents.*

2.4 Execution reliability

Even with agreement, delegated AI execution can fail (timeouts, tool failures). With probability $\rho \in [0, 1)$, OTC execution fails and the trade defaults to market execution. With probability $1 - \rho$, the trade executes OTC at interpreted price $\tilde{p}$.

Assumption 1 (Branch-linear evaluation of execution reliability)**.** *After applying Lemma 1 within each execution branch, agents evaluate the operational failure lottery linearly in branch certainty equivalents. Thus ρ is a reduced-form reliability haircut on the successful-OTC incremental surplus rather than an additional unresolved CARA wealth risk.*

Remark 2. *Without Assumption 1, a literal Bernoulli mixture of market and OTC outcomes would have a log-sum certainty equivalent, for example* $-(1/\gamma_i)\log\{\rho e^{-\gamma_i CE_i^M} + (1-\rho)e^{-\gamma_i CE_i^{OTC}}\}$. *The branch-linear convention keeps the closed-form bargaining and investment expressions transparent.*

2.5 Timing

1. Each side chooses outside-option compute $e_i \geq 0$.
2. The buyer chooses protocol intensity $b \geq 0$; $b = 0$ means no OTC protocol is activated (trade executes in market).
3. If $b > 0$ and OTC is jointly viable, the agents bargain over p via Nash bargaining with buyer bargaining weight $\beta \in (0, 1)$.
4. Execution: OTC with probability $1 - \rho$, otherwise market.

3 Certainty equivalents

Lemma 1 (CARA-normal certainty equivalent)**.** *If* $X \sim \mathcal{N}(m, s^2)$ *and* $U(W) = -e^{-\gamma W}$, *then* $\mathrm{CE}(w + X) = w + m - \frac{\gamma}{2}s^2$.

3.1 Market execution (outside option)

If the trade executes in the market, incremental certainty equivalents relative to no trade are

$$g_B^M(e_B) = -\tau_B\Delta - \alpha_B(e_B)\Delta^2 - \frac{\gamma_B}{2}\Delta(2q_B + \Delta)\sigma^2, \tag{4}$$

$$g_S^M(e_S) = -\tau_S\Delta - \alpha_S(e_S)\Delta^2 - \frac{\gamma_S}{2}\Delta(\Delta - 2q_S)\sigma^2. \tag{5}$$

3.2 Successful OTC execution

If OTC executes at interpreted price $\tilde{p} = p + \xi$, incremental certainty equivalents are

$$g_B^{OTC}(p, b) = \Delta(\mu - p) - \frac{\gamma_B}{2}\Delta(2q_B + \Delta)\sigma^2 - \frac{\gamma_B}{2}\Delta^2\frac{\sigma_\xi^2}{b}, \tag{6}$$

$$g_S^{OTC}(p, b) = \Delta(p - \mu) - \frac{\gamma_S}{2}\Delta(\Delta - 2q_S)\sigma^2 - \frac{\gamma_S}{2}\Delta^2\frac{\sigma_\xi^2}{b}. \tag{7}$$

3.3 The OTC–market wedge (joint gains from OTC)

Relative to market execution, the incremental joint surplus from *successful* OTC is

$$\begin{aligned}\Pi_0(b, e_B, e_S) &:= \big(g_B^{OTC} + g_S^{OTC}\big) - \big(g_B^M + g_S^M\big) \\ &= \Delta(\tau_B + \tau_S) + \Delta^2\big(\alpha_B(e_B) + \alpha_S(e_S)\big) - \frac{\gamma_B + \gamma_S}{2}\Delta^2\frac{\sigma_\xi^2}{b}. \quad (8)\end{aligned}$$

Under Assumption 1, execution reliability scales the successful OTC incremental surplus by $1 - \rho$:

$$\Pi(b, e_B, e_S) = (1 - \rho)\Pi_0(b, e_B, e_S).$$

4 Nash bargaining solution: price and payoffs

Assumption 2 (Nash bargaining over the OTC price)**.** *Conditional on activating OTC ($b > 0$) and joint viability, the agents select p to maximize the Nash product of gains over the market outside option. The protocol cost is sunk at this stage and therefore affects activation but not the price split:*

$$\max_p \quad \big(g_B^{OTC}(p, b) - g_B^M(e_B)\big)^\beta \big(g_S^{OTC}(p, b) - g_S^M(e_S)\big)^{1-\beta}. \quad (9)$$

Proposition 1 (Equilibrium OTC price and surplus split)**.** *Fix (b, e_B, e_S) such that $\Pi_0(b, e_B, e_S) > 0$. Under Assumption 2:*

1. *The negotiated price splits the incremental pie:*

$$\begin{aligned}g_B^{OTC}(p^\star, b) - g_B^M(e_B) &= \beta\,\Pi_0(b, e_B, e_S), \\ g_S^{OTC}(p^\star, b) - g_S^M(e_S) &= (1 - \beta)\,\Pi_0(b, e_B, e_S).\end{aligned}$$

2. *The intended equilibrium price is*

$$\begin{aligned}p^\star = \mu &+ (1 - \beta)\tau_B - \beta\tau_S \\ &+ \Delta\Big((1 - \beta)\alpha_B(e_B) - \beta\alpha_S(e_S)\Big) \\ &+ \frac{\Delta\sigma_\xi^2}{2b}\Big(\beta\gamma_S - (1 - \beta)\gamma_B\Big). \quad (10)\end{aligned}$$

If $\Pi_0(b, e_B, e_S) = 0$, the acceptance interval collapses to this limiting price and both parties are indifferent between OTC and market execution. If $\Pi_0(b, e_B, e_S) < 0$ or $b = 0$, the trade executes in the market.

Remark 3. *Under complete information, Nash allocations can be implemented by stationary alternating-offers equilibria (Binmore et al., 1986; Nash, 1950; Osborne & Rubinstein, 1990; Rubinstein, 1982). The Nash formulation is standard in OTC finance and keeps the letter maximally tractable (Duffie, 2012; Duffie et al., 2005).*

5 Protocol intensity: underprovision and activation

Define

$$A := \frac{\gamma_B + \gamma_S}{2}\Delta^2\sigma_\xi^2, \qquad W(e_B, e_S) := \Delta(\tau_B + \tau_S) + \Delta^2\big(\alpha_B(e_B) + \alpha_S(e_S)\big).$$

Then $\Pi_0(b, e_B, e_S) = W(e_B, e_S) - A/b$.

5.1 Optimal protocol intensity conditional on activation

If the buyer activates OTC, it chooses $b > 0$ to maximize its captured expected incremental surplus net of protocol cost:

$$\begin{aligned} &\max_{b>0}\ \beta(1-\rho)\Pi_0(b, e_B, e_S) - K(b) \\ &= \max_{b>0}\ \beta(1-\rho)\left(W(e_B, e_S) - \frac{A}{b}\right) - \lambda\frac{b^{1+\nu}}{1+\nu}. \end{aligned}$$

Proposition 2 (Closed-form protocol intensity)**.** *Conditional on activation, the buyer's unique interior optimal protocol intensity is*

$$b^\star = \left(\frac{\beta(1-\rho)A}{\lambda}\right)^{\frac{1}{\nu+2}} = \left(\frac{\beta(1-\rho)(\gamma_B+\gamma_S)\Delta^2\sigma_\xi^2}{2\lambda}\right)^{\frac{1}{\nu+2}}. \tag{11}$$

Proposition 3 (Planner protocol and underinvestment)**.** *A planner choosing b to maximize total expected surplus $(1-\rho)\Pi_0(b, e_B, e_S) - K(b)$ sets*

$$b^{SP} = \left(\frac{(1-\rho)A}{\lambda}\right)^{\frac{1}{\nu+2}} = \beta^{-\frac{1}{\nu+2}}\, b^\star,$$

so $b^\star < b^{SP}$ whenever $\beta < 1$.

5.2 Activation decision and a crowd-out threshold

If the buyer does not activate OTC, it sets $b = 0$ and the trade executes in the market. Let

$$\begin{aligned}\mathcal{G}(e_B, e_S) &:= \max_{b>0} \ \beta(1-\rho)\Pi_0(b, e_B, e_S) - K(b) \\ &= \beta(1-\rho)\Pi_0(b^\star, e_B, e_S) - K(b^\star)\end{aligned}$$

be the buyer's net gain from activating OTC (relative to market), evaluated at $b^\star$.

Proposition 4 (OTC activation threshold)**.** *The buyer activates OTC if and only if*

$$W(e_B, e_S) \ \geq \ \underline{W} := \frac{\nu+2}{\nu+1} \cdot \frac{A}{b^\star} = \frac{\nu+2}{\nu+1} A^{\frac{\nu+1}{\nu+2}} \left(\frac{\lambda}{\beta(1-\rho)}\right)^{\frac{1}{\nu+2}}. \tag{12}$$

Moreover, since $W(e_B, e_S)$ decreases in outside-option compute (e_B, e_S), a compute arms race makes OTC activation less likely and can generate endogenous OTC-to-market migration.

6 Outside-option compute: a strategic arms race and overinvestment

6.1 Compute conditional on OTC activation

Fix $b > 0$ and suppose OTC is activated. Using (1)–(2) and Proposition 1, each party chooses e_i anticipating bargaining and execution reliability.

Proposition 5 (Closed-form compute levels (OTC regime))**.** *Fix $b > 0$ and suppose OTC is activated and jointly viable. Then equilibrium outside-option compute satisfies*

$$e_B^{OTC} = \max\left\{0, \ \Delta\sqrt{\frac{\bar{\alpha}_B\big(1-\beta(1-\rho)\big)}{\kappa_B}} - 1\right\}, \tag{13}$$

$$e_S^{OTC} = \max\left\{0, \ \Delta\sqrt{\frac{\bar{\alpha}_S\big(\beta+(1-\beta)\rho\big)}{\kappa_S}} - 1\right\}. \tag{14}$$

Holding impact-reduction productivity and compute costs fixed, a lower buyer bargaining weight increases e_B^{OTC} and decreases e_S^{OTC}. Thus, under similar impact-reduction and cost parameters across parties, the weaker bargaining party invests more in outside-option compute.

6.2 Planner compute and overinvestment

A planner that anticipates OTC execution and treats fallback as an operational failure realized with probability ρ chooses e_i to minimize expected realized impact plus compute cost:

$$\min_{e_i \geq 0} \ \rho\, \alpha_i(e_i)\Delta^2 + \kappa_i e_i.$$

Proposition 6 (Strict overinvestment in outside-option compute)**.** *Under (1)–(2), the planner sets*

$$e_i^{SP} = \max\left\{0,\ \Delta\sqrt{\frac{\rho\,\bar{\alpha}_i}{\kappa_i}} - 1\right\}.$$

Whenever $\rho < 1$ and $\beta \in (0,1)$ and both solutions are interior, equilibrium features strict overinvestment:

$$e_B^{OTC} > e_B^{SP}, \qquad e_S^{OTC} > e_S^{SP}.$$

6.3 Market-regime compute and a two-regime equilibrium characterization

If the buyer does not activate OTC ($b = 0$), the trade executes in the market, so compute only serves to reduce realized impact.

Proposition 7 (Compute in the marketization regime)**.** *If $b = 0$ so the trade executes in the market, each party chooses*

$$e_i^{M} = \max\left\{0,\ \Delta\sqrt{\frac{\bar{\alpha}_i}{\kappa_i}} - 1\right\}.$$

Moreover, $e_i^M \geq e_i^{OTC}$ whenever the latter is interior (since the OTC-regime effective weight on impact is strictly below 1).

Let $s_B := \beta$, $s_S := 1 - \beta$, $\chi_B := 1$, and $\chi_S := 0$. For a fixed counterpart compute choice e_j, define party i's activation-contingent private payoff, net of compute and any protocol cost it bears, as

$$V_i^{OTC}(e_i, e_j) := g_i^M(e_i) + (1-\rho)s_i\Pi_0(b^\star, e_i, e_j) - C_i(e_i) - \chi_i K(b^\star),$$

where g_i^M denotes the relevant market-execution certainty equivalent in (4) or (5). If the trade is routed directly to the market, party i's private payoff is

$$V_i^M(e_i) := g_i^M(e_i) - C_i(e_i).$$

Assumption 3 (Sufficient regime no-deviation conditions)**.** *Let $j \neq i$ and define the deviation sets*

$$\mathcal{D}_i^{OTC\to M}(e_j) := \{e_i \geq 0 : W(e_i, e_j) < \underline{W}\},$$
$$\mathcal{D}_i^{M\to OTC}(e_j) := \{e_i \geq 0 : W(e_i, e_j) \geq \underline{W}\}.$$

At the OTC candidate, no party can profitably deviate to a compute level that would collapse activation:

$$V_i^{OTC}(e_i^{OTC}, e_j^{OTC}) \geq \sup_{e_i \in \mathcal{D}_i^{OTC\to M}(e_j^{OTC})} V_i^M(e_i), \qquad i \in \{B, S\}.$$

At the marketization candidate, no party can profitably deviate to a compute level that would restore activation:

$$V_i^M(e_i^M) \geq \sup_{e_i \in \mathcal{D}_i^{M\to OTC}(e_j^M)} V_i^{OTC}(e_i, e_j^M), \qquad i \in \{B, S\}.$$

The supremum over an empty deviation set is interpreted as $-\infty$.

Corollary 1 (Candidate regimes)**.**
Let $W^{OTC} := W(e_B^{OTC}, e_S^{OTC})$ and $W^M := W(e_B^M, e_S^M)$ with $W^M \leq W^{OTC}$.

1. *If $W^{OTC} < \underline{W}$, then the local OTC-compute candidate cannot support activation.*
2. *If $W^M \geq \underline{W}$, then the marketization candidate is not locally consistent: even market-regime compute leaves sufficient surplus for OTC activation.*
3. *If $W^M < \underline{W} \leq W^{OTC}$, both activation statuses are locally consistent. Under the sufficient no-deviation inequalities in Assumption 3, both pure-strategy candidate regimes are equilibria. This is the protocol-technology trap: high compute lowers market frictions and discourages protocol activation, while low compute supports protocol activation and sustains OTC.*

7 Governance price bands

To capture compliance/fairness constraints in delegated agents, suppose OTC quotes must satisfy

$$|p - \mu| \leq h,$$

where $h \geq 0$ is an exogenous governance tightness parameter. The band is imposed on the intended price p as a pre-trade compliance rule. If instead governance constrained the interpreted price $\tilde{p}$, feasibility would become a chance-constraint problem because protocol noise could move realized terms outside the band.

Proposition 8 (Governance constraints can force market execution)**.** *Define acceptance bounds (conditional on successful OTC) as the set of p such that each party weakly prefers OTC to market:*

$$p_L = \mu - \tau_S - \alpha_S(e_S)\Delta + \frac{\gamma_S}{2}\Delta\frac{\sigma_\xi^2}{b},$$
$$p_U = \mu + \tau_B + \alpha_B(e_B)\Delta - \frac{\gamma_B}{2}\Delta\frac{\sigma_\xi^2}{b}.$$

With the band $[\mu - h, \mu + h]$, OTC execution is feasible if and only if

$$\underbrace{p_L \leq p_U}_{\text{basic OTC viability } (\Pi_0 \geq 0)} \qquad \text{and} \quad \max\{p_L, \mu - h\} \leq \min\{p_U, \mu + h\}.$$

Equivalently, OTC is feasible if and only if $\Pi_0(b, e_B, e_S) \geq 0$ and

$$\begin{aligned} h \;&\geq\; h_{\min}(b, e_B, e_S) \\ &:= \max\left\{0,\; \frac{\gamma_S}{2}\Delta\frac{\sigma_\xi^2}{b} - \tau_S - \alpha_S(e_S)\Delta,\right. \\ &\qquad\qquad \left.\frac{\gamma_B}{2}\Delta\frac{\sigma_\xi^2}{b} - \tau_B - \alpha_B(e_B)\Delta\right\}. \end{aligned}$$

8 Robustness to general loss schedules

The benchmark uses A/b for residual protocol loss and $\bar{\alpha}_i/(1+e_i)$ for market impact to obtain closed forms. The economic wedges do not rely on these particular functions. Let $H(b)$ denote the joint protocol loss under successful OTC, with $H'(b) < 0$, and let $\Psi_i(e_i)$ denote party i's centralized-market execution loss relative to frictionless OTC, with $\Psi_i'(e_i) < 0$. Successful OTC creates joint gains

$$\Pi_0(b, e_B, e_S) = \Psi_B(e_B) + \Psi_S(e_S) - H(b).$$

Assumption 4 (General loss schedules)**.** *H is twice continuously differentiable, strictly decreasing, and weakly convex; K is twice continuously*

differentiable, strictly increasing, and strictly convex; and $K'(b)/[-H'(b)]$ is strictly increasing. For each $i \in \{B, S\}$, Ψ_i is twice continuously differentiable, strictly decreasing, and weakly convex; C_i is twice continuously differentiable, strictly increasing, and strictly convex; and $C_i'(e)/[-\Psi_i'(e)]$ is strictly increasing.

Lemma 2 (Single-crossing investment comparison)**.** *Let $L'(x) < 0$ and $C'(x) > 0$ on an interval, and suppose $R(x) := C'(x)/[-L'(x)]$ is strictly increasing. If two interior choices solve*

$$C'(x(a)) = a[-L'(x(a))], \qquad C'(x(a')) = a'[-L'(x(a'))],$$

with $0 < a < a'$, then $x(a) < x(a')$.

Proposition 9 (Functional-form robustness)**.** *Suppose Assumption 4 holds and compare interior solutions. Conditional on activation and fixed compute, the private and planner protocol choices satisfy*

$$K'(b^\star) = -\beta(1-\rho)H'(b^\star), \qquad K'(b^{SP}) = -(1-\rho)H'(b^{SP}),$$

so $b^\star < b^{SP}$ whenever $\beta < 1$. Conditional on OTC activation and fixed protocol intensity, let $s_B = \beta$ and $s_S = 1 - \beta$. Party i's private compute choice and the planner's compute choice satisfy

$$C_i'(e_i^\star) = \big(1 - s_i(1-\rho)\big)\big[-\Psi_i'(e_i^\star)\big], \qquad C_i'(e_i^{SP}) = \rho\big[-\Psi_i'(e_i^{SP})\big],$$

so $e_i^\star > e_i^{SP}$ for both parties whenever $0 < \rho < 1$ and the compared compute solutions are interior.

Proposition 9 shows that the closed-form model is a tractable special case of a broader delegated-bargaining problem. Productive protocol investment is underprovided when its cost is private and its benefit is shared through bargaining. Outside-option compute is overprovided because the private return includes a bargaining channel in addition to the realized fallback execution channel.

9 Comparative statics and implications

Table 1 summarizes the main comparative statics. The signs should be read as local effects under the maintained regime. Activation effects combine direct threshold changes with induced changes in the market-friction wedge $W(e_B, e_S)$.

Table 1: Comparative statics

Parameter increase	Protocol $b^\star$	Compute	OTC activation
β	Higher	e_B^{OTC} lower; e_S^{OTC} higher	Mixed: protocol incentives rise, but compute incentives shift across parties
ρ	Lower	Both e_B^{OTC} and e_S^{OTC} higher	Harder: reliability falls and compute lowers W
λ	Lower	No direct effect	Harder through a higher activation threshold
σ_ξ^2	Higher conditional on activation	No direct effect	Harder because protocol-risk burden rises
κ_i	No direct effect	Party i computes less	Easier through a larger residual OTC-saving wedge
$\bar{\alpha}_i$	No direct effect	Party i computes more	Usually easier because raw market frictions rise, though compute also responds
$\tau_B + \tau_S$	No direct effect	No direct effect	Easier because OTC saves a larger fixed market wedge

Numerical anchor. Table 2 reports a normalized and illustrative symmetric benchmark with $\Delta = 1$, $\tau_B = \tau_S = 0.30$, $\bar{\alpha}_B = \bar{\alpha}_S = 0.60$, $\kappa_B = \kappa_S = 0.15$, $\beta = 0.5$, $\gamma_B = \gamma_S = 1$, $\sigma_\xi^2 = 1$, $\nu = 1$, and $\lambda = 0.30$. The table varies only the OTC failure probability ρ. The exercise is not a calibration; the parameter values are chosen only to illustrate scale-normalized comparative statics. Higher failure risk lowers protocol intensity, raises fallback compute, lowers the available OTC wedge, and raises the activation threshold.

Observables. The protocol intensity b maps to schema validation, confirmation layers, verification tokens, shared message standards, audit trails, and reconciliation steps. Outside-option compute e_i maps to smart order routing, market-impact models, fallback venue search, broker/algorithm selection, inference spend, simulation budget, and execution analytics. The reliability parameter ρ maps to timeout, tool-failure, rejection, or reconciliation-failure rates. The protocol-noise variance σ_ξ^2 maps to term-misunderstanding, quote-

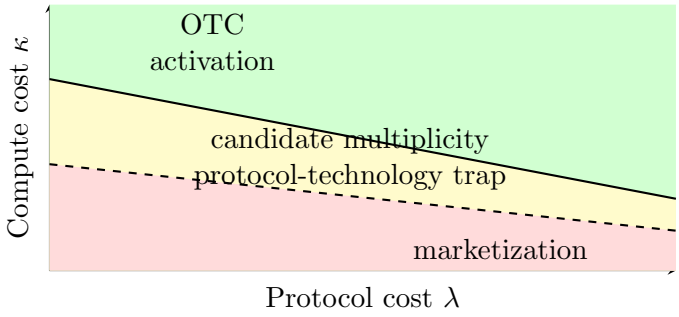


Figure 1: Activation regions under symmetric parameters. Lower protocol costs make OTC activation more attractive, while lower compute costs improve fallback market execution and reduce the OTC-saving wedge. The middle band is a candidate multiplicity region: low compute can sustain OTC activation, whereas high fallback compute can make market execution privately attractive.

Table 2: Illustrative sensitivity to OTC failure risk

ρ	$b^\star$	$e_B^{OTC} = e_S^{OTC}$	W^{OTC}	$\underline{W}$	Candidate regime
0.05	1.166	0.449	1.428	1.287	Multiplicity
0.20	1.101	0.549	1.375	1.363	Multiplicity
0.35	1.027	0.643	1.330	1.461	Market only

discrepancy, message-translation, manual-correction, or contract-field error rates. These observables link the model to work on OTC transaction costs and transparency (Edwards et al., 2007) and algorithmic trading effects on liquidity (Brogaard, 2014; Hendershott et al., 2011). They also suggest two empirical designs for future work: a protocol-adoption test asking whether richer standardized messages and lower quote-revision rates predict a larger OTC share, and a compute-substitution test asking whether improvements in fallback routing or execution analytics reduce OTC use by shrinking the market-friction wedge.

Institutional interpretation. The model distinguishes shared protocol infrastructure from private fallback compute. Standardized schemas, certification, validation layers, auditability, and interoperable rules of engagement lower λ or σ_ξ^2 and raise the value of OTC activation in the model. By contrast, private fallback compute can be useful operationally but excessive strategically because it changes bargaining payoffs even when fallback execution is

off-path.

Limitations. The analysis uses complete information, an exogenous bargaining weight, a reduced-form compute technology, one-sided protocol cost, and the branch-linear CE convention in Assumption 1. It also provides testable comparative statics rather than direct empirical estimates. These assumptions keep the model solvable; future work could endogenize bargaining power, allow two-sided complementary protocol investment, derive exact mixture certainty equivalents, and estimate protocol and compute wedges in transaction data.

10 Conclusion

This paper provides a compact, closed-form theory of AI-agent OTC bargaining that separates a productive protocol margin from a strategic outside-option compute margin. Protocol quality is privately underprovided, fallback compute is privately overprovided, and the interaction can push trade from bilateral OTC negotiation toward centralized market execution. The mechanism is that shared validation standards expand the bilateral surplus, while uncoordinated fallback-compute races can dissipate that surplus through bargaining leverage.

A Proofs

Proof of Lemma 1. For $X \sim \mathcal{N}(m, s^2)$, the moment-generating function gives

$$\mathbb{E}\left[e^{-\gamma X}\right] = \exp\left(-\gamma m + \frac{\gamma^2 s^2}{2}\right).$$

Hence

$$\mathbb{E}\left[-e^{-\gamma(w+X)}\right] = -\exp\left\{-\gamma w - \gamma m + \frac{\gamma^2 s^2}{2}\right\}.$$

The certainty equivalent c solves

$$-e^{-\gamma c} = -\exp\left\{-\gamma w - \gamma m + \frac{\gamma^2 s^2}{2}\right\},$$

so $c = w + m - \gamma s^2/2$. □

Proof of Proposition 1. Let

$$x_B(p) := g_B^{OTC}(p, b) - g_B^M(e_B), \qquad x_S(p) := g_S^{OTC}(p, b) - g_S^M(e_S).$$

For fixed (b, e_B, e_S), $x_B'(p) = -\Delta$, $x_S'(p) = \Delta$, and $x_B(p) + x_S(p) = \Pi_0(b, e_B, e_S)$. The first-order condition for the Nash product is

$$\beta \frac{x_B'(p)}{x_B(p)} + (1-\beta)\frac{x_S'(p)}{x_S(p)} = 0,$$

or equivalently

$$\frac{\beta}{x_B(p)} = \frac{1-\beta}{x_S(p)}.$$

Combining this equation with $x_B + x_S = \Pi_0$ yields

$$x_B(p^\star) = \beta \Pi_0(b, e_B, e_S), \qquad x_S(p^\star) = (1-\beta)\Pi_0(b, e_B, e_S).$$

Using the buyer's split condition,

$$\Delta(\mu - p^\star) + \tau_B \Delta + \alpha_B(e_B)\Delta^2 - \frac{\gamma_B}{2}\Delta^2 \frac{\sigma_\xi^2}{b} = \beta \Pi_0(b, e_B, e_S).$$

Substituting the expression for Π_0 and solving for $p^\star$ gives

$$\begin{aligned} p^\star = \mu &+ (1-\beta)\tau_B - \beta \tau_S \\ &+ \Delta\left((1-\beta)\alpha_B(e_B) - \beta \alpha_S(e_S)\right) \\ &+ \frac{\Delta \sigma_\xi^2}{2b}\left(\beta \gamma_S - (1-\beta)\gamma_B\right). \end{aligned}$$

□

Proof of Proposition 2. Conditional on activation, the buyer solves

$$\max_{b>0} \ \beta(1-\rho)\left(W(e_B, e_S) - \frac{A}{b}\right) - \lambda \frac{b^{1+\nu}}{1+\nu}.$$

The first-order condition is

$$\beta(1-\rho)\frac{A}{b^2} - \lambda b^\nu = 0.$$

Thus

$$b^\star = \left(\frac{\beta(1-\rho)A}{\lambda}\right)^{1/(\nu+2)}.$$

The second derivative is

$$-2\beta(1-\rho)\frac{A}{b^3} - \lambda \nu b^{\nu-1} < 0,$$

and the derivative tends to $+\infty$ as $b \downarrow 0$ and to $-\infty$ as $b \to \infty$, so the stationary point is the unique interior maximizer. □

Proof of Proposition 3. The planner uses full surplus weight $(1-\rho)$ rather than the buyer's captured share:

$$b^{SP} = \left(\frac{(1-\rho)A}{\lambda}\right)^{1/(\nu+2)} = \beta^{-1/(\nu+2)} b^\star.$$

Since $\beta \in (0,1)$, $b^{SP} > b^\star$. □

Proof of Proposition 4. The buyer activates OTC if

$$\mathcal{G}(e_B, e_S) = \beta(1-\rho)\left(W(e_B, e_S) - \frac{A}{b^\star}\right) - \lambda\frac{(b^\star)^{1+\nu}}{1+\nu} \geq 0.$$

From the first-order condition defining $b^\star$,

$$\lambda(b^\star)^{1+\nu} = \beta(1-\rho)\frac{A}{b^\star}.$$

Therefore

$$\mathcal{G}(e_B, e_S) = \beta(1-\rho)\left[W(e_B, e_S) - \frac{A}{b^\star} - \frac{1}{1+\nu}\frac{A}{b^\star}\right].$$

Hence $\mathcal{G} \geq 0$ if and only if

$$W(e_B, e_S) \geq \frac{\nu+2}{\nu+1}\frac{A}{b^\star}.$$

Substituting the expression for $b^\star$ gives the stated threshold. □

Proof of Proposition 5. Conditional on OTC activation, the buyer's payoff terms depending on e_B are

$$g_B^M(e_B) + \beta(1-\rho)\Pi_0(b, e_B, e_S) - \kappa_B e_B.$$

Because

$$\frac{\partial g_B^M}{\partial e_B} = -\alpha_B'(e_B)\Delta^2, \qquad \frac{\partial \Pi_0}{\partial e_B} = \alpha_B'(e_B)\Delta^2,$$

the first-order condition is

$$\left[1-\beta(1-\rho)\right]\left[-\alpha_B'(e_B)\right]\Delta^2 = \kappa_B.$$

Using $\alpha_B(e_B) = \bar{\alpha}_B/(1+e_B)$ gives

$$e_B^{OTC} = \max\left\{0, \Delta\sqrt{\frac{\bar{\alpha}_B(1-\beta(1-\rho))}{\kappa_B}} - 1\right\}.$$

For the seller, the same argument applies with bargaining share $(1-\beta)$:

$$[1-(1-\beta)(1-\rho)]\,\frac{\bar{\alpha}_S\Delta^2}{(1+e_S)^2}=\kappa_S.$$

Since $1-(1-\beta)(1-\rho)=\beta+(1-\beta)\rho$, this gives the stated formula for e_S^{OTC}. □

Proof of Proposition 6. The planner's problem for party i is

$$\min_{e_i\geq 0}\ \rho\frac{\bar{\alpha}_i}{1+e_i}\Delta^2+\kappa_i e_i.$$

The interior first-order condition is

$$\rho\frac{\bar{\alpha}_i\Delta^2}{(1+e_i)^2}=\kappa_i,$$

so

$$e_i^{SP}=\max\left\{0,\Delta\sqrt{\frac{\rho\bar{\alpha}_i}{\kappa_i}}-1\right\}.$$

For the buyer,

$$1-\beta(1-\rho)-\rho=(1-\beta)(1-\rho)>0.$$

For the seller,

$$\beta+(1-\beta)\rho-\rho=\beta(1-\rho)>0.$$

Thus, when both private and planner solutions are interior, private compute strictly exceeds planner compute for both parties. □

Proof of Proposition 7. In the marketization regime, party i solves

$$\min_{e_i\geq 0}\ \frac{\bar{\alpha}_i}{1+e_i}\Delta^2+\kappa_i e_i.$$

The interior first-order condition is

$$\frac{\bar{\alpha}_i\Delta^2}{(1+e_i)^2}=\kappa_i,$$

which gives

$$e_i^{M}=\max\left\{0,\Delta\sqrt{\frac{\bar{\alpha}_i}{\kappa_i}}-1\right\}.$$

The OTC-regime impact weights are $1-\beta(1-\rho)<1$ for the buyer and $\beta+(1-\beta)\rho<1$ for the seller when $\rho<1$ and $\beta\in(0,1)$. Therefore $e_i^M\geq e_i^{OTC}$ when the relevant solutions are interior. □

Proof of Corollary 1. Because $e_i^M \geq e_i^{OTC}$ and $\alpha_i(e_i)$ is decreasing, $W^M \leq W^{OTC}$. If $W^{OTC} < \underline{W}$, even the low-compute OTC candidate does not create enough OTC surplus to meet the activation threshold. If $W^M \geq \underline{W}$, even the high-compute marketization candidate leaves enough OTC surplus for activation, so direct market execution is not locally consistent. If $W^M < \underline{W} \leq W^{OTC}$, the activation status is locally consistent under either candidate compute profile. The first inequality in Assumption 3 rules out profitable unilateral compute deviations from the OTC candidate that would move the trade to market execution. The second inequality rules out profitable unilateral compute deviations from the marketization candidate that would restore OTC activation. Therefore both locally consistent candidate regimes satisfy the relevant compute no-deviation conditions under Assumption 3. □

Proof of Proposition 8. The buyer weakly prefers successful OTC to market when

$$g_B^{OTC}(p, b) - g_B^M(e_B) \geq 0.$$

Solving this inequality for p gives

$$p \leq \mu + \tau_B + \alpha_B(e_B)\Delta - \frac{\gamma_B}{2}\Delta\frac{\sigma_\xi^2}{b} = p_U.$$

The seller weakly prefers successful OTC to market when

$$g_S^{OTC}(p, b) - g_S^M(e_S) \geq 0,$$

which is equivalent to

$$p \geq \mu - \tau_S - \alpha_S(e_S)\Delta + \frac{\gamma_S}{2}\Delta\frac{\sigma_\xi^2}{b} = p_L.$$

Thus OTC is feasible under the governance band $[\mu - h, \mu + h]$ if and only if the acceptance interval $[p_L, p_U]$ intersects the band:

$$p_L \leq p_U, \qquad \max\{p_L, \mu - h\} \leq \min\{p_U, \mu + h\}.$$

Equivalently, $p_L \leq p_U$ and

$$h \geq \max\{0, p_L - \mu, \mu - p_U\}.$$

Substituting p_L and p_U gives the stated expression for $h_{\min}$. □

Proof of Lemma 2. The first-order conditions imply

$$R(x(a)) = a, \qquad R(x(a')) = a'.$$

If $x(a') \leq x(a)$, strict monotonicity gives $R(x(a')) \leq R(x(a))$, contradicting $a' > a$. Therefore $x(a) < x(a')$. □

Proof of Proposition 9. For protocol choice, fixed compute makes $\Psi_B(e_B) + \Psi_S(e_S)$ constant in b. The buyer's first-order condition is

$$-\beta(1-\rho)H'(b^\star) - K'(b^\star) = 0,$$

which gives the private condition $K'(b^\star)/[-H'(b^\star)] = \beta(1-\rho)$. The planner replaces $\beta(1-\rho)$ with $(1-\rho)$. Since $\beta(1-\rho) < 1-\rho$, Lemma 2 applied to $(L, C) = (H, K)$ gives $b^\star < b^{SP}$.

For compute choice, fixed protocol intensity and the other party's compute make party i's relevant private objective

$$-\Psi_i(e_i) + s_i(1-\rho)\Psi_i(e_i) - C_i(e_i),$$

up to constants. The first-order condition is

$$C_i'(e_i^\star) = \big(1 - s_i(1-\rho)\big)\big[-\Psi_i'(e_i^\star)\big].$$

The planner minimizes $\rho\Psi_i(e_i) + C_i(e_i)$, giving $C_i'(e_i^{SP}) = \rho[-\Psi_i'(e_i^{SP})]$. For the buyer, $1 - \beta(1-\rho) - \rho = (1-\beta)(1-\rho) > 0$; for the seller, $1 - (1-\beta)(1-\rho) - \rho = \beta(1-\rho) > 0$. Thus the private coefficient exceeds the planner coefficient for both parties when $0 < \rho < 1$. Applying Lemma 2 to $(L, C) = (\Psi_i, C_i)$ gives $e_i^\star > e_i^{SP}$ for each interior comparison. □

Disclosure of Research Support and Potential Conflicts

The authors report no external research support for this manuscript and no financial relationships or other potential conflicts of interest relevant to this paper.

Disclosure on AI Assistance

During preparation, the authors used AI-assisted tools for grammar, consistency checking, and revision planning. The authors reviewed and edited all AI-assisted output and remain responsible for the manuscript's content, analysis, citations, and conclusions.

References

Aldasoro, I., Gambacorta, L., Korinek, A., Shreeti, V., & Stein, M. (2024). *Intelligent financial system: How ai is transforming finance* (BIS Working Paper No. 1194). Bank for International Settlements.

Almgren, R., & Chriss, N. (2000). Optimal execution of portfolio transactions. *Journal of Risk*, *3*(2), 5–39.

Binmore, K., Rubinstein, A., & Wolinsky, A. (1986). The Nash bargaining solution in economic modelling. *RAND Journal of Economics*, *17*(2), 176–188.

Brogaard, J. (2014). High frequency trading and its impact on market quality. *Review of Financial Studies*, *27*(8), 2267–2306.

Budish, E., Cramton, P., & Shim, J. (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics*, *130*(4), 1547–1621.

Duffie, D. (2012). *Dark markets: Asset pricing and information transmission in over-the-counter markets*. Princeton University Press.

Duffie, D., Gârleanu, N., & Pedersen, L. H. (2005). Over-the-counter markets. *Econometrica*, *73*(6), 1815–1847.

Edwards, A. K., Harris, L. E., & Piwowar, M. S. (2007). Corporate bond market transaction costs and transparency: Evidence from TRACE. *Journal of Finance*, *62*(3), 1421–1451.

Financial Stability Board. (2024, November 14). *The financial stability implications of artificial intelligence*. Financial Stability Board.

FIX Trading Community. (2026). *FIX standards*. https://fixtrading.org/standards/

Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, *14*(1), 71–100.

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

Grout, P. A. (1984). Investment and wages in the absence of binding contracts: A Nash bargaining approach. *Econometrica*, *52*(2), 449–460.

Hart, O., & Moore, J. (1990). Property rights and the nature of the firm. *Journal of Political Economy*, *98*(6), 1119–1158.

Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *Journal of Finance*, *66*(1), 1–33.

Hirshleifer, J. (1991). The technology of conflict as an economic activity. *American Economic Review Papers and Proceedings*, *81*(2), 130–134.

International Organization of Securities Commissions. (2021). *The use of artificial intelligence and machine learning by market intermediaries and asset managers* (Final Report No. FR06/2021). International Organization of Securities Commissions.

International Organization of Securities Commissions. (2025). *Artificial intelligence in capital markets: Use cases, risks, and challenges* (Consultation Report No. CR/01/2025). International Organization of Securities Commissions.

Konrad, K. A. (2009). *Strategy and dynamics in contests*. Oxford University Press.

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, *53*(6), 1315–1335.

Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets*, *3*(3), 205–258.

Nash, J. F. (1950). The bargaining problem. *Econometrica*, *18*(2), 155–162.

Osborne, M. J., & Rubinstein, A. (1990). *Bargaining and markets*. Academic Press.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, *50*(1), 97–109.

Tullock, G. (1980). Efficient rent-seeking. In J. M. Buchanan, R. D. Tollison, & G. Tullock (Eds.), *Toward a theory of the rent-seeking society* (pp. 97–112). Texas A&M University Press.

U.S. Securities and Exchange Commission. (2023, July 26). *SEC proposes new requirements to address risks to investors from conflicts of interest associated with the use of predictive data analytics by broker-dealers and investment advisers* [Press Release 2023-140].