Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Buyer and seller value from optional automated bargaining in electronic commerce

Bankatesh Choudhary

phdmb10007.21@bitmesra.ac.in

Birla Institute of Technology, Mesra

Meenakshi Sharma

Birla Institute of Technology, Mesra

Article

Keywords: Automated bargaining, Electronic commerce, Posted pricing, Buyer and seller value, Interaction costs, Computational modelling

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11271405/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Whether automated bargaining benefits electronic commerce depends on the transaction route it changes, not only on the quality of its negotiated offers. A checkout-relative value model is developed to evaluate an optional seller-side bargaining agent against existing posted-price checkout. The model separates new-sale margin, discounts on otherwise feasible checkout purchases, unrecovered sales and participant interaction costs. It derives conditional cost and fallback limits for seller and buyer outcomes. Public product-price and negotiation inputs support a matched computational comparison of two rules and two frozen fitted Q iteration policies across 193 evaluation products. In the reference configuration, the full-state learner adds 0.0144 seller and 0.0039 buyer listing-price units per original opportunity, but a simpler learner provides higher combined value. Stubborn offers produce seller losses, while low valuations produce buyer losses after interaction costs. These results show how a seller-favoured bargaining policy can differ from a configuration that benefits both participants on average. The contribution is a reproducible model linking executable policy outcomes to the economic consequences of an optional transaction channel. Participation, customer values, fallback and effort costs remain stipulated; the findings are conditional computational evidence, not estimates of actual adoption, customer experience or realised commercial effects. Positive reference averages conceal product heterogeneity: only 82 of 193 product contexts yield positive means for both parties. Structural sensitivities further qualify deployment claims.

Introduction

An electronic commerce platform that introduces automated bargaining changes how customers can transact, not simply how a seller calculates its next offer. A negotiated discount may enable a purchase that would otherwise fail, but it may also reduce the revenue from a customer who would have paid the posted price. An unsuccessful exchange can consume effort and interrupt an otherwise feasible purchase. The relevant economic question is therefore whether bargaining creates additional value for buyers and sellers relative to the checkout route already available.

This question connects pricing institutions with the growing use of artificial agents in commerce. Research on bargaining and posted prices has long examined negotiation costs, competitive selling formats and the coexistence of negotiated and fixed-price channels (Desai and Purohit 2004; Wang 1995; Zeng et al. 2014). Field evidence documents sequential offers and disagreement in online bargaining, while research on negotiation initiation shows why participation cannot be treated as costless (Backus et al. 2020; Jindal and Newberry 2018). These studies establish that bargaining is an economic institution with conditions of use. They do not imply that every capable automated negotiator should be offered to every customer.

Automated-negotiation research supplies concession rules, opponent models and learning methods for choosing offers (Baarslag et al. 2016; Faratin et al. 1998; Jennings et al. 2001). Dialogue systems further distinguish strategic decisions from language generation (He et al. 2018; Xia et al. 2024). Yet a policy

evaluated on the surplus earned within a negotiation has not necessarily been evaluated as an optional transaction channel. Agreement rate, negotiated price and seller reward can omit the value of displaced checkout purchases and the costs borne by unsuccessful participants. They also cannot establish how customers experience an artificial counterpart (Chaudhuri and De Bruyn 2026).

This study develops a checkout-relative value model to connect these levels of analysis. The model evaluates a potential customer opportunity before the customer selects a purchase route. It separates the seller's new-sale margin, discounts on purchases that checkout would have captured, unrecovered checkout margin after failed bargaining, and interaction costs. Buyer outcomes are evaluated against the same original opportunity and their own interaction costs. The model yields conditional cost and fallback limits and an explicit account of why a policy preferred by one participant can differ from the policy preferred on combined value.

The central research question is under what stated conditions optional automated bargaining creates additional seller value without reducing buyers' average net outcomes. Two further questions specify the comparison. Under which route, acceptance and cost assumptions does policy ordering change? Which features of their price, agreement and duration outcomes account for the resulting gains and losses? These are questions about the economic consequences of a digital transaction format, rather than a claim to a new general learning algorithm.

The computational evaluation holds four seller policies fixed: two transparent rules and two previously trained fitted Q iteration policies with different public information states. Public product prices and negotiation records anchor selected inputs; customer values, participation, acceptance, fallback and interaction costs remain stipulated. Policies face matched synthetic opportunities across 193 evaluation products. Retaining fixed policies separates the consequences of channel evaluation from a new round of policy selection after inspecting the results.

The analysis makes three connected contributions. It specifies a common checkout counterfactual and opportunity denominator for comparing participant outcomes. It links economic mechanisms to the outputs of executable policies and derives conditional operating limits with explicit feasibility conditions. Finally, its reproducible comparison shows why a positive bargaining reward is insufficient to establish that an optional service benefits both sides. The reference configuration gives positive average increments, but the seller-favoured learner does not maximise combined value. Other stipulated conditions produce seller or buyer losses. These findings support a conditional account of when bargaining is useful, while leaving actual adoption, perceived fairness and realised commercial effects to subsequent empirical evaluation.

Literature and theoretical positioning

Bargaining as a pricing institution

The choice between bargaining and posted prices has an established economic literature. Wang compares the selling formats dynamically, showing that their attraction depends on costs and other conditions (Wang 1995). Desai and Purohit examine competitive haggling and differences among customer segments (Desai and Purohit 2004). Zeng and colleagues analyse negotiated and no-haggle channels within a dual-pricing strategy, including internal competition and the bargaining price floor (Zeng et al. 2014). These studies preclude treating the coexistence of bargaining and checkout, or the possibility of discount cannibalisation, as a new economic principle.

Experimental and equilibrium research also challenges the assumption that more bargaining necessarily improves market outcomes. In laboratory customer markets with switching costs, haggling generated lower efficiency and higher, stickier seller prices than posted offers under the studied conditions (Cason et al. 2003). Theoretical analyses show that the availability of negotiated discounts can alter initial list-price competition (Raskovich 2007), and that an increased share of bargainers can weaken outside options when firms adjust their posted prices (Gill and Thanassoulis 2009). These mechanisms operate through strategic market responses. They are absent from a single-merchant comparison that deliberately holds the listing price fixed.

That distinction matters for interpretation. In the present model, a common positive uptake probability multiplies the affected outcomes and does not by itself change their signs. This is a conditional implication with fixed prices and potential negotiations, not a general claim that wider bargaining adoption leaves market outcomes unchanged. The study contributes at the configuration level: given a posted price and an executable seller policy, what changes when a customer can choose bargaining before returning to or leaving checkout? It neither solves an equilibrium selling-format problem nor estimates how competitors would respond.

Participation costs and digital market frictions

Lower communication costs do not make electronic transactions frictionless. Electronic-marketplace theory relates search costs to resource allocation and seller power (Bakos 1997). Early online-retail evidence found lower prices alongside substantial dispersion and retailer heterogeneity (Brynjolfsson and Smith 2000), while a subsequent review distinguished dispersion measures and multichannel explanations (Pan et al. 2004). These contributions concern search and price formation across sellers. A historical minimum or maximum in one product record is a different quantity and cannot identify a merchant's acquisition cost or a customer's willingness to pay.

Costly online haggling is especially close to the current question. Hann and Terwiesch estimate frictional costs from repeated bids at a name-your-own-price retailer (Hann and Terwiesch 2003). Spann and colleagues jointly estimate individual willingness to pay and frictional costs from bidding behaviour (Spann et al. 2004). Terwiesch and colleagues connect costly repeated offers with retailer thresholds, interface choices and the welfare consequences of price discrimination (Terwiesch et al. 2005). Thus, adding an effort term to a price model is not itself novel. Those studies identify costs or optimise

mechanisms in their settings; the present comparison applies explicitly assumed costs to the realised trajectories of frozen policies.

Participation must also be separated from performance after entry. Jindal and Newberry examine fixed costs in the decision to initiate a negotiation (Jindal and Newberry 2018). Field bargaining data reveal sequences of offers and disagreement, rather than only the terminal price of successful bargains (Backus et al. 2020). Together, these findings motivate retaining unsuccessful attempts and the original opportunity denominator. They do not supply the uptake, fallback or effort parameters used here. An entry cost and a cost per bargaining round are distinct constructs, and neither should be inferred from an opening bid alone.

The checkout alternative is also a model rather than an observed fact. Research on heterogeneous consumers and platform search demonstrates how search design and evaluation costs affect purchasing processes (Li et al. 2025). Here, checkout eligibility is simplified to simulated value at least listing price, with no search, delivery or payment friction. This makes the counterfactual transparent but limits transportability. A merchant application would need a measured checkout baseline as well as evidence about bargaining participants.

Executable agents and channel evaluation

Automated negotiation provides a mature set of architectures and tactics. Foundational work describes negotiation mechanisms and challenges (Jennings et al. 2001), concession functions (Faratin et al. 1998) and opponent modelling (Baarslag et al. 2016). Goal-deliberated architectures already select among time- and behaviour-dependent strategies in implemented systems (Cao et al. 2015). Tournament analysis further distinguishes winning against a particular opponent set from broader adaptability and robustness (Baarslag et al. 2013). A workflow diagram or a combination of known strategies is therefore insufficient to establish a new bargaining contribution.

Learning and language introduce additional capabilities without resolving the channel question automatically. Tree-based fitted Q iteration learns from batch transitions under a specified reward (Ernst et al. 2005), and ANEGMA uses reinforcement learning with rule-informed initialisation for negotiation in simulated electronic markets (Bagga et al. 2021). Dialogue studies separate strategy from response generation (He et al. 2018), while bargaining benchmarks and recent tool-integrated or feedback-based agents evaluate increasingly capable language-model negotiators (Kwon et al. 2025; Oh et al. 2026; Xia et al. 2024). The present scalar-price policies are not implementations of those conversational systems and are not claimed to outperform them.

The relevant distinction is between a learning objective and the consequences of deploying its policy within an existing transaction route. Algorithmic-pricing research positions automation as a strategic marketing decision rather than a purely technical upgrade (Spann et al. 2026). Research on agentic markets similarly distinguishes technical automation from persistent behavioural and informational

frictions (Bichler 2026). A checkout-relative comparison makes some of these omitted terms explicit for a particular bilateral service. Within model-driven decision support, the model links the decision object, assumptions and computational outputs without establishing that managers or customers have adopted it (Power and Sharda 2007).

Customer experience and heterogeneous responses to artificial agents

Automated bargaining combines a numerical offer with an interaction. General AI-marketing frameworks already distinguish operational, analytical and relational functions (Huang and Rust 2021). An experiential perspective separates consumers' encounters with data capture, classification, delegation and social AI (Puntoni et al. 2021). A price-policy evaluation addresses only part of this domain. The simulated buyer's per-round cost represents an assumed burden; it does not measure perceived control, privacy concerns or interpersonal discomfort.

Evidence also resists a universal assumption about willingness to interact with an algorithm. Forecasting experiments find reduced reliance after people observe algorithmic errors (Dietvorst et al. 2015), whereas other advice-taking experiments find algorithm appreciation under different conditions (Logg et al. 2019). Perceived task subjectivity affects reliance (Castelo et al. 2019), and the relative salience of utilitarian and hedonic attributes affects responses to AI recommendations (Longoni and Cian 2022). These are contextual findings, not interchangeable measures of bargaining adoption. Advice from an agent expected to help a consumer differs from negotiating against a seller's agent.

A recent meta-analysis distinguishes responses to robots, chatbots and algorithms and finds that important contingencies differ across agent types (Gelbrich et al. 2026). Trust research likewise separates cognitive and emotional trust (Glikson and Woolley 2020). The implication for this model is to leave participation and experience unestimated rather than import an average effect from a different task. Variation in assumed uptake can test sensitivity, but it cannot convert the model into a study of trust or establish which interface consumers prefer.

Negotiation-specific evidence strengthens this separation. The Subjective Value Inventory distinguishes feelings about outcomes, the self, the process and relationships (Curhan et al. 2006). Human-bot bargaining research shows why objective performance and subjective evaluation require separate attention (Chaudhuri and De Bruyn 2026). Procurement experiments report different economic and relational consequences of competitively and collaboratively prompted buyer chatbots (Herold et al. 2025). That procurement setting reverses the agent role relative to the seller-side service studied here; its findings motivate an outcome distinction without validating this simulator.

Interface disclosure and pricing rules can matter independently of numerical capability. An outbound-sales field experiment found that chatbot identity disclosure changed purchasing responses (Luo et al. 2019). Experiments with adaptive name-your-own-price thresholds considered transactions, seller

returns and satisfaction together (Hinz et al. 2011). These results support measuring the interface and experience in a prospective test, not concealing agent identity or assuming that any monetary gain implies satisfaction. CaSiNo's human-human resource-allocation dialogues provide supporting experience evidence in another domain, not retail validation of the proposed agent (Chawla et al. 2021).

Simulation evidence and the specific contribution

Simulation credibility requires separating implementation verification from substantive validation. Sargent distinguishes conceptual, data and operational validity from checking whether a model was implemented correctly (Sargent 2013). Research on empirical validation of agent-based models examines the relationship between a model's generating process and observed behaviour (Windrum et al. 2007). A recent methods review similarly treats validation as purpose-specific and supported by complementary approaches (Collins et al. 2024). These distinctions explain why frozen checkpoints, replay agreement and product-bootstrap intervals are useful but insufficient evidence of real customer behaviour.

The present study accordingly treats its public sources as inputs with different roles. Product-price records anchor heterogeneity; Craigslist dialogues inform opening-offer inputs and supporting history analysis; CaSiNo informs the separation of economic and experiential outcomes. They are not three replications of one channel effect. The principal evidence is a matched computational comparison under declared assumptions, including conditions unfavourable to bargaining.

Table 1 summarises the theoretical position. The contribution is the explicit link from executable policy outcomes to the value of an optional checkout alternative, evaluated for both participants over original opportunities. The decomposition and thresholds are conditional accounting implications, not newly discovered laws of bargaining. The computational evidence concerns the behaviour of the evaluated frozen policies within that model and the mechanisms that explain when their conclusions change. This scoped narrative synthesis supports that claim; it is not a systematic review or an assertion that no related evaluation has previously existed.

Table 1
Position of the checkout relative value model within related research

Research stream	Established contribution	Role and boundary of the present model
Pricing formats and outside options (Gill and Thanassoulis 2009; Wang 1995; Zeng et al. 2014)	Bargaining costs, channel interaction and endogenous posted-price responses	Evaluates frozen policies against fixed checkout; does not solve price-setting equilibrium
Costly online offers (Hann and Terwiesch 2003; Terwiesch et al. 2005)	Empirical friction costs and design of repeated-bid mechanisms	Retains failed-attempt costs and original opportunities; does not estimate customer costs or values
Automated negotiation (Baarslag et al. 2013; Bagga et al. 2021; Cao et al. 2015)	Strategy selection, learning and opponent-dependent agent performance	Connects policy outputs to optional-channel value; does not claim a new learning algorithm
Human and artificial-agent interaction (Chaudhuri and De Bruyn 2026; Curhan et al. 2006; Herold et al. 2025)	Distinct economic and subjective outcomes under context-dependent interaction	Reports separate buyer and seller increments; experience, trust and adoption remain unmeasured
Simulation credibility (Collins et al. 2024; Sargent 2013)	Purpose-specific verification and validation of models	Provides reproducible implementation checks and sensitivities, not independent behavioural validation

Note. The table compares selected research streams rather than claiming an exhaustive survey. The present evidence concerns a fixed-price, single-merchant, scalar-price simulation. Literature findings are not treated as calibration data for its participation, fallback or effort assumptions.

Checkout relative value model

Unit of analysis and purchase routes

An opportunity is a potential customer encountering one product before selecting a purchase route. Let the listing price be L, the stipulated merchant floor divided by L be f, and the customer's simulated value divided by L be v. Ordinary checkout produces a sale indicator D_0 equal to one when v is at least one. Its seller and buyer surpluses are S_0 and B_0. For bargaining policy k, the potential negotiation outcome consists of an agreement indicator D_k, a normalised price P_k when agreement occurs, and T_k elapsed rounds, including unsuccessful attempts.

$$D_0 = 1\{v \geq 1\}, \quad S_0 = D_0(1 - f), \quad B_0 = D_0(v - 1).$$

1

For an agreement, S_k equals P_k minus f and B_k equals v minus P_k. All failed-negotiation surpluses are defined as zero without evaluating an undefined transaction price. Seller surplus above f is a modelling outcome, not verified accounting profit. The floor is a stipulated business constraint; historical catalogue prices do not establish merchant acquisition costs or authorize discounts.

Let q_i be the probability of entering bargaining for opportunity i and r the probability of returning to checkout after bargaining fails. A fallback purchase additionally requires checkout eligibility. The possible routes are direct checkout without bargaining, a negotiated sale, checkout following a failed negotiation, and no purchase. The routes are mutually exclusive. Separate costs c_s and c_b accrue per bargaining round for the seller and buyer, including rounds in failed negotiations. The optional-route seller value and its incremental value relative to checkout are

$$V_{sk} = E\left[(1 - q_i)S_0 + q_i S_k + q_i(1 - D_k)rS_0 - q_i c_s T_k\right]$$
$$\Delta V_{sk} = V_{sk} - E\left[S_0\right]$$

2

The buyer counterpart replaces seller surplus by buyer surplus and charges c_b rather than c_s. Expectations average opportunities within each product and then products equally. For a learned policy, outcomes first average its five frozen training seeds within product. Thus changing a route probability does not create additional customers or independent observations.

Figure 1 shows the model's computational structure. It is an evaluation workflow connecting frozen policies to a checkout counterfactual. A real customer-facing interface, inventory transaction and payment system are outside this demonstration.

Mechanisms of incremental value

The incremental seller outcome can be decomposed into new-sale margin, discounts on purchases that checkout would already have captured, unrecovered checkout margin after failed negotiation, and interaction cost:

$$\Delta V_{sk} = M_k - C_k - (1 - r)U_k - c_s H_k.$$

3

Here M_k is the expectation of q_i times (1 minus D_0) times S_k; C_k is the expectation of q_i times D_0 times D_k times (1 minus P_k), with its contribution set to zero for failures. U_k is the expectation of q_i times (1 minus D_k) times S_0, and H_k is the expectation of q_i times T_k. These terms are respectively new-sale margin, discount cannibalization, checkout margin exposed to failed negotiation, and expected rounds per original opportunity. With prices bounded by the listing, C_k is nonnegative. These definitions

give an accounting identity. Their contribution is to make policy-dependent mechanisms auditable, not to introduce an otherwise unknown economic law.

The decomposition makes selection and counterfactual choices explicit. Conditioning on successful bargains omits failed-attempt costs and unrecovered checkout opportunities. Reporting gross negotiated surplus additionally fails to subtract checkout value displaced by successful bargains. Reporting only bargaining participants omits the platform-level question of how many original opportunities are affected. Our primary estimand retains the original opportunity for every policy and route configuration.

Conditional model implications

For compactness define A_{sk} as the expected uptake-weighted difference between negotiation and checkout seller surplus. Define U_{sk} as recoverable checkout seller surplus after negotiation failure and H_k as uptake-weighted duration. The buyer quantities A_{bk} and U_{bk} are defined analogously. The participant index j denotes seller or buyer; the corresponding gross surplus J_{jk} is S_k or B_k, and the checkout quantities are $J_{s0} = S_0$ and $J_{b0} = B_0$. Opportunity indices are suppressed for readability. For fixed potential trajectories and uptake probabilities,

$$A_{jk} = E\left[q_i(J_{jk} - J_{j0})\right], \quad U_{jk} = E\left[q_i(1 - D_k)J_{j0}\right]$$
$$\Delta V_{jk} = A_{jk} + rU_{jk} - c_j H_k$$

4

Proposition 1

concerns variable interaction costs. If H_k is positive, a nonnegative average increment for participant j requires the participant's round cost not to exceed the corresponding threshold in Eq. (5). A negative threshold means that even zero nonnegative variable cost cannot make the average increment nonnegative at that fallback setting. For U_{jk} greater than zero, Eq. (5) also gives the required fallback probability at a specified cost.

$$c_{jk}^{*} = \frac{A_{jk} + rU_{jk}}{H_k}, \quad r_{jk}^{*} = \frac{c_j H_k - A_{jk}}{U_{jk}}.$$

5

A fallback threshold above one is infeasible in this model; a threshold at or below zero means that no positive fallback is required. If U_{jk} is zero, fallback cannot change that participant's increment. For equal but separately charged participant costs, the smaller seller and buyer cost ceiling is the condition for both average increments to be nonnegative. It is not an individual Pareto-improvement criterion: some customers can still lose while the average improves. These statements follow directly by rearranging Eq. (4), with the denominator conditions stated explicitly.

Proposition 2

concerns uptake. If every opportunity enters bargaining with the same positive probability p, all incremental terms are multiplied by p. Changing p alone cannot change the sign of an increment, the variable-cost ceiling or the fallback threshold. A nontrivial uptake threshold requires an additional mechanism, such as changing participant composition or a fixed platform cost. We do not add a fixed platform cost to the evaluated primary model. At p equal to zero, the optional service reduces to checkout and the increment is zero.

Proposition 3

concerns policy rankings. For policies a and b evaluated on the same opportunity population and uptake mechanism, let G_k denote A_{sk} plus r times U_{sk}. Then their seller-value difference and a candidate cost crossing are

$$\Delta V_{sa} - \Delta V_{sb} = (G_a - G_b) - c_s(H_a - H_b)$$

$$c_{ab}{}^{*} = \frac{G_a - G_b}{H_a - H_b}$$

6

If the duration difference is zero, no finite cost crossing is identified from this expression. Otherwise the crossing has substantive meaning only in the stated cost domain, and the direction depends on the sign of the duration difference. At zero cost and zero fallback, with common positive constant uptake, policy ordering is exactly the ordering of unconditional gross negotiation seller surplus. Subtracting the same checkout baseline cannot itself reverse that ordering. This negative control separates a mechanism-based result from a relabelling of the evaluation metric.

These propositions are conditional implications of the specified model. They are not empirical hypotheses about actual adoption or newly estimated consumer behaviour. Computational threshold estimates characterize the evaluated policies; the algebra alone does not establish that any particular crossing occurs within a plausible operating range.

Participation composition as a sensitivity analysis

The primary design uses exogenous uptake. A second design multiplies base participation odds by two below the listing-value boundary and by one half at or above it. This gives different probabilities for customers who could and could not purchase through checkout. It represents a stipulated selection pattern, not a fitted behavioural model or a deployable targeting rule. Private value is available to the simulator but is never passed to the seller policy.

Changing the base propensity in this second design can change the mixture of affected opportunities. Its sign changes therefore need not obey the constant-uptake invariance in Proposition 2. The analysis reports composition coefficients and feasible-domain checks rather than attributing these patterns to

measured preferences. In particular, results cannot justify screening real customers by an unobserved willingness to pay.

Data and computational design

Public data and scope of evidence

Table 2 summarises the role of each public source. No new human participants, observed platform transactions or customer satisfaction observations were collected for this study. The principal evaluation is computational. Public catalogue and dialogue records constrain selected inputs but do not identify all behavioural parameters.

Table 2
Public sources and their distinct evidential roles

Source and processed records	Observed or derived input	Not established by these data
CraigslistBargain (He et al. 2018) 6,449 dialogues; 4,071 scenarios	Recognised price mentions and initial bid ratios from the training split; separate offer-prediction analysis	Private willingness to pay, agent efficacy, platform participation or causal response to a seller quote
AmazonHistoryPrice (Xia et al. 2024) 930 products; 18 categories	Recorded current and historical prices; hashed 555/182/193 product partitions	Actual merchant costs, customer values or live demand. Floor multiplier and value distribution are stipulated
CaSiNo (Chawla et al. 2021) 1,030 dialogues; 2,060 participant-dialogue records	Resource-allocation points and ordered experience responses; supporting associations	Retail customer satisfaction, AI-mediated interaction or independent validation of the channel effect

Note. The sources are not pooled. Counts refer to processed records. The integrated policy-channel experiment uses the 193-product evaluation partition; the other sources do not provide independent replications of that effect.

AmazonHistoryPrice contains 930 product records across 18 categories (Xia et al. 2024). Recorded current price, minimum price and maximum price provide catalogue heterogeneity; current denotes a dataset field rather than a live price. A deterministic identifier hash allocates 555 products to training, 182 to development and 193 to evaluation. Policies were fitted using training products and selected using development outcomes. These partitions are disjoint, but earlier outcomes on the evaluation partition had already been inspected. The integrated experiment is therefore exploratory and does not restore an untouched holdout by drawing additional trajectories.

The reference floor starts from 0.8 times the recorded historical minimum divided by the recorded current price, clipped between 0.2 and 0.95. Multiplication by the listing price gives a currency amount,

which is rounded upward to cents and divided by the listing price again. This is a proxy configuration, not an estimated cost or recommended discount policy. Customer values, deadlines, participation, fallback and per-round burdens remain stated assumptions. Confidence intervals cannot absorb uncertainty about their realism.

The processed CraigslistBargain corpus contains 6,449 dialogues after duplicate and cross-split-scenario exclusions (He et al. 2018). Initial bid ratios are drawn only from its training dialogue records, clipped between 0.30 and 0.85, then capped at 0.9 times simulated value. The previously completed matched-history analysis is retained in the reproducibility archive, rather than used as an independent test of the learned seller. Its target is a future recognized price, not private customer value or the causal effect of a seller quote. CaSiNo's 1,030 dialogues supply only the supporting outcome-measurement analysis (Chawla et al. 2021).

Frozen policy comparison

The primary policy set contains a linear time-concession rule, a buyer-responsive rule, a full-public-state fitted Q iteration policy and a clock-only fitted Q iteration policy. Table 3 records their inputs and selection. All parameters and checkpoints are inherited unchanged from the documented prior development procedure; no policy is selected because it performs best in the channel evaluation.

Table 3
Frozen policy specifications and public inputs

Policy	Inputs	Development setting
Time concession	Round, horizon and configured floor; preceding ask bounds quotation	Linear path to floor; terminal concession 1 and exponent 1.
Buyer responsive	Time-rule inputs and current bid	Response weight 0.25; bid markup 0.20; terminal concession 1 and exponent 1.
Full-state FQI	Nine public state variables: clock, horizon, floor, current/previous bid, previous ask, bid change and running bid extrema	Mixed exploration; 128 episodes per training product; minimum leaf 40; all five training seeds.
Clock-only FQI	Round, horizon, configured floor and preceding ask	Same training budget and selected minimum leaf 40; all five training seeds.

Note. FQI = fitted Q iteration. Each learned checkpoint uses 24 trees, depth 12 and eight fitted iterations. No policy observes private value, deadline or buyer type. Selection preceded the new channel evaluation; prior evaluation outcomes were already inspected.

The learned policies use tree-based fitted Q iteration (Ernst et al. 2005). Each checkpoint uses 24 extremely randomized trees of maximum depth 12, eight fitted iterations and the previously selected minimum-leaf setting of 40. Five training seeds are retained: 271, 811, 1601, 2027 and 4093. Both learned specifications use mixed exploration at 128 episodes per training product. Their state-dependent candidate set contains nine fractions from zero to one in steps of 0.125 of the remaining ask-to-floor gap

and an exact feasible linear-rule quote. This rule-informed candidate and exploration support must be considered when interpreting comparisons with the linear rule; the learners are not independent of that benchmark's structure.

The full public state includes normalised round, platform horizon, floor, current bid, preceding ask, preceding bid, bid change and running bid minimum and maximum. The clock-only policy retains round, horizon, floor and preceding ask. Neither receives private customer value, deadline or behavioural type. The fixed action representation and finite training budget limit conclusions about the value of richer state. The comparison does not establish that a minimal state is generally sufficient or that the full state is universally better.

All asks remain between the merchant floor and preceding ask and are projected upward to the next cent. Customer bids are rounded downward and cannot exceed the listing price or simulated value. A bid meeting the current ask is accepted at the offered bid; otherwise the customer may accept the ask according to the specified mechanism. These details determine realised surplus and are reproduced in the source code.

Matched potential negotiations

The integrated evaluation uses 80 potential customers per evaluation product for each declared environment. Each policy and seed faces the same initial opportunities and abandonment draws. Customers in the reference environment have value-to-listing ratios uniform between 0.60 and 1.15, uniformly sampled deadlines of three, five or eight rounds, and equal representation of linear, early-conceding, late-conceding and reciprocal offer patterns. Reference abandonment probability after an unresolved round is 0.05 and the platform horizon is eight rounds. The stubborn-offer condition holds the opening bid fixed; the low-valuation condition uses values uniform between 0.50 and 0.85. These conditions are retained without stress-specific retraining.

An ask is accepted when it does not exceed the current bid plus an acceptance coefficient times the remaining bid-to-value gap, and does not exceed value. Coefficients zero, 0.25, 0.5 and one separate strict bid acceptance from acceptance anywhere within remaining willingness to pay. The coefficient is a behavioural assumption, not an observed probability. Offer concession and acceptance are deliberately distinct: a customer who does not increase its offer might still accept an ask.

The integration produces 2,223,360 potential negotiation executions across 48 saved policy-condition files. Each condition uses 15,440 original opportunities: 193 products times 80 repetitions. The learned-policy files contain all five checkpoint seeds, whereas each rule has one deterministic policy instance. These repeated executions support matched comparisons; the inferential cluster count remains 193 products. Analytical route accounting evaluates 2,592 policy-channel configurations without creating additional independent observations.

The trained reward remains terminal gross seller surplus in the original reference environment. The integrated evaluation does not optimise a channel-level reward, learn participation or retrain under

alternative acceptance coefficients. It tests the consequences of deploying the frozen policy choice within the specified purchase-route accounting. Earlier evaluation partitions and the reference route stream were already inspected; locally saved protocols document the new analysis choices but are not an external preregistration.

Channel configurations and uncertainty

The primary illustration fixes exogenous opt-in at 0.5, fallback at 0.5, acceptance at 0.25 and a separate cost of 0.005 listing-price units per bargaining round for each participant. A cost of 0.005 means 0.5 percent of the product's listing price per round; it is not a common observed dollar cost across products. Sensitivities vary base opt-in over 0.25, 0.5 and 0.75, fallback over zero, 0.5 and one, and participant round cost over zero, 0.005 and 0.02. Both selection mechanisms and all declared policy and customer conditions are retained.

The analysis integrates route probabilities analytically rather than drawing additional binary participation or fallback outcomes. Results are averages per original opportunity. Within each product, repetitions and the five learned-policy seeds are averaged before comparisons; rule policies have one deterministic instance. Paired differences use the same products and potential customers. The 95 percent percentile intervals use 2,000 product bootstrap resamples. They describe variability over the 193 products conditional on the frozen seeds, customer mechanisms and cost assumptions.

Threshold estimates divide average coefficients, not individual-opportunity ratios. The bootstrap repeats that ratio calculation; undefined denominator cases and infeasible thresholds are retained rather than clipped into a plausible range. Full factorial grids are descriptive, while intervals are reported for the declared primary comparisons and thresholds. No familywise confirmatory interpretation or test-selected best policy is claimed.

Verification and reproducibility

The integration checks frozen checkpoint hashes, product partitions, unique opportunity identifiers, monetary bounds and replay agreement with existing rule and learned-policy implementations. An independently written accounting check enumerates direct, negotiated, fallback and no-purchase routes, then reconstructs participant receipts, proxy costs and benefits. Transfer payments cancel in joint gross surplus; buyer and seller round costs remain separately charged. Tests also check the decomposition, constant-uptake scaling, threshold roots and the zero-cost zero-fallback policy-ranking invariant.

All 24 prior rule-condition replay checks and 30 learned-policy seed-by-regime comparisons passed. Seven integration contract tests, 16 accounting fixtures spanning 90 parameter combinations, and 1,667,520 adjacent-acceptance monotonicity comparisons passed. A complete alternate-directory replay reproduced all 2,223,360 trajectories numerically. The route oracle checks every trajectory under both selection mechanisms. These are separate implementation checks conducted within the same research workflow, not independent external validation.

These tests establish specified code behaviour and arithmetic consistency, not the truth of the customer model (Collins et al. 2024; Sargent 2013; Windrum et al. 2007). The reproducibility archive distinguishes new integrated evaluations from earlier learning, history and experience analyses. Data sources, processing code, selected checkpoints, protocols, numerical results and figure-generation scripts are retained. The paper does not treat millions of simulated executions as millions of independent empirical observations.

Exploratory robustness extension

We examined product-level distributions, all five frozen learner seeds separately and leave-one-category-out means. The last check removes each represented category in turn without retraining; it is an influence diagnostic, not generalisation to an unseen category. Separate seller and buyer cost ceilings were evaluated over fallback probabilities from zero to one. Feasibility regions use point estimates; uncertainty for each marginal boundary is in Supplementary Data 1.

Eight prespecified local sensitivity settings retained the original random stream: the reference replay; floor multipliers 0.65 and 0.95 instead of 0.80; value supports [0.50, 1.05] and [0.70, 1.25] instead of [0.60, 1.15]; beta(2,5) and beta(5,2) shapes scaled to the reference support; and the higher floor combined with the lower support. Quantile ranks, empirical opening draws before the value cap, deadlines and abandonment draws were matched. All negotiations were regenerated, with no policy refitting. These 1,482,240 executions include 185,280 reference replays; they are not additional independent empirical observations. The dated extension protocol preceded these runs but followed inspection of earlier outcomes. Supplementary Tables S1–S5 report settings and checks; all intervals remain exploratory, marginal and conditional on the simulator.

Figures were generated from saved numerical tables using Matplotlib. Vector exports retain text and scientific geometry; no generative imagery or image retouching was used.

Use of generative artificial intelligence

OpenAI Codex assisted with code generation and revision, computational checks, literature organisation, figure-production code, manuscript drafting and interpretation. It was not used as a human participant or an independent annotator. Automated checks do not constitute independent human validation. The authors retain responsibility for verifying sources, data permissions, code, results, interpretations and the submitted text.

Results

Incremental outcomes at the primary configuration

Figure 2 reports seller and buyer increments for the declared primary design.

In the reference environment, full-state fitted Q iteration gives seller increment 0.0144 [0.0102, 0.0189] and buyer increment 0.0039 [0.0021, 0.0057]. These are fractions of the listing price per original opportunity, not percentage increases over baseline profit. Its conversion increment is 0.1560, or 15.60 percentage points under the stipulated checkout rule, not an observed retail conversion uplift.

Table 4 separates policy contrasts from comparisons against checkout. In the reference environment, full-state FQI exceeds the time rule in seller value by 0.0046 [0.0041, 0.0052]. Relative to clock-only FQI, its seller advantage is 0.0030 [0.0026, 0.0033], but its buyer difference is −0.0036 [−0.0044, −0.0029]. The combined difference is −0.0006 [−0.0012, −0.000042]. Thus the seller-favoured learned specification is not the combined-value-favoured specification among the policies evaluated. This result is a stakeholder trade-off, not evidence that richer information is generally harmful.

Table 4
Paired policy differences calculated as full state FQI minus comparator

Environment and comparator	Seller difference [95% interval]	Buyer difference [95% interval]
Reference Full minus Time concession	0.0046 [0.0041, 0.0052]	−0.0010 [−0.0017, −0.0003]
Reference Full minus Clock-only FQI	0.0030 [0.0026, 0.0033]	−0.0036 [−0.0044, −0.0029]
Stubborn offers Full minus Time concession	0.0002 [−0.0007, 0.0011]	−0.0024 [−0.0037, −0.0011]
Stubborn offers Full minus Clock-only FQI	−0.0025 [−0.0031, −0.0019]	−0.0091 [−0.0105, −0.0077]
Low valuations Full minus Time concession	0.0044 [0.0032, 0.0057]	0.0014 [0.0009, 0.0018]
Low valuations Full minus Clock-only FQI	−0.0012 [−0.0018, −0.0007]	−0.0014 [−0.0018, −0.0010]

Note. Differences are full-state FQI minus comparator, per original opportunity in listing-price units. Exogenous uptake 0.5, fallback 0.5, acceptance coefficient 0.25 and a cost of 0.005 per round for each participant. Paired percentile intervals use 2,000 resamples of 193 products, averaging repetitions and frozen seeds within product.

The alternative conditions change the conclusion. Under stubborn offers, every seller increment is negative at the primary configuration. The full-state learner gives −0.0175 [−0.0205, −0.0140], compared with −0.0150 [−0.0183, −0.0112] for the clock-only learner. In the low-valuation condition, full-state FQI produces seller increment 0.0073 [0.0026, 0.0124] but buyer increment −0.0086 [−0.0094, −0.0078]. No customer in that condition is eligible for checkout at the listing price; unsuccessful and prolonged

bargaining still impose buyer costs. Neither higher agreement nor positive seller margin establishes a beneficial service for both sides.

Where the incremental value comes from

Figure 3 decomposes seller increments. For the full-state learner in the reference environment, new-sale margin is 0.03369 per original opportunity. Discount cannibalization subtracts 0.00780, failed-checkout losses subtract 0.00206, and seller interaction costs subtract 0.00942, leaving 0.01441. The gross new-sale term would therefore materially overstate the net channel contribution if presented alone.

For stubborn offers, the same learner generates new-sale margin of 0.01326 but loses 0.00774 through discounts, 0.01233 through unrecovered checkout margin and 0.01065 through interaction costs. Its negative increment is not explained by price concessions alone. In contrast, checkout displacement terms are zero in the low-valuation condition because the simulated values are all below the listing price. There the relevant comparison is newly enabled transactions against failed and successful interaction burdens.

Conditional operating limits

Table 5 reports the full-state learner's cost and fallback limits. In the reference environment the seller can bear up to 0.01265 listing-price units per round on average, but the buyer limit is only 0.00707 [0.00609, 0.00811]. The lower buyer limit determines the common-cost ceiling for simultaneous nonnegative average increments. At the illustrative 0.005 cost, the reference result needs no positive fallback probability under the specified trajectories; this does not establish that fallback is unimportant in other environments.

Table 5
Conditional cost and fallback limits for full-state fitted Q iteration

Environment	Seller cost ceiling	Buyer cost ceiling	Seller fallback floor
Reference	0.0127	0.0071	−3.005
	[0.0103, 0.0153]	[0.0061, 0.0081]	[−4.109, −1.967]
Stubborn offers	−0.0032	0.0056	1.208
	[−0.0044, −0.0017]	[0.0042, 0.0071]	[1.081, 1.317]
Low valuations	0.0083	0.0011	Undefined
	[0.0062, 0.0108]	[0.0008, 0.0014]	Fallback irrelevant

Note. Cost ceilings are listing-price units per bargaining round at fallback 0.5. The seller fallback floor is evaluated at cost 0.005. Uptake is exogenous 0.5 and acceptance is 0.25. Brackets show paired product-bootstrap 95% percentile intervals. Negative cost ceilings indicate infeasibility even at zero cost for that fallback. A negative fallback floor requires no fallback; a floor above 1 is infeasible. Undefined means no eligible failed-checkout margin, not missing data. The smaller participant cost ceiling determines whether both average increments are nonnegative.

Under stubborn offers the seller cost ceiling at fallback 0.5 is negative. At cost 0.005, the required seller fallback probability is 1.208 [1.081, 1.317], beyond the feasible range even at complete fallback. Under low valuations, the buyer cost ceiling falls to 0.00107 [0.00078, 0.00139], well below the primary cost. Fallback cannot repair this result because there are no listing-price-eligible customers. These thresholds describe average outcomes under fixed trajectories; they are not recommended transaction fees or empirically estimated patience limits.

Sensitivity and policy ordering

Figure 4 shows that acceptance behaviour can change both net value and policy ordering.

At acceptance coefficient one under stubborn offers, the time rule yields seller increment 0.02211 and full-state FQI yields 0.01462; their paired full-minus-time difference is −0.00749 [−0.00836, −0.00658]. At the primary coefficient 0.25, however, their difference is only 0.00018 [−0.00073, 0.00115]. We therefore do not describe the latter apparent ordering as a clear learned-policy advantage. A policy that benefits from a particular acceptance mechanism should not be declared robust to acceptance assumptions without this comparison.

Constant exogenous uptake scales increments without changing their signs or thresholds, as required analytically. The stipulated value-dependent selection changes composition instead: in the reference environment the full-state learner's buyer increment becomes 0.00172 [−0.00018, 0.00370], so the interval no longer excludes zero. This does not estimate actual self-selection; it demonstrates dependence on an unmeasured input.

Cost crossings are also conditional. In the reference environment the full-state versus clock-only seller crossing is approximately 0.07498 per round, outside the evaluated cost interval from zero to 0.02. We do not claim that round-cost accounting reverses that particular ranking within the evaluated domain. A stubborn-offer crossing involving the buyer-responsive rule has a highly unstable bootstrap ratio and does not support a precise decision boundary. The negative-control comparison preserves gross-surplus ordering for all 72 policy pairs when cost and fallback are both zero and uptake is common.

Across the twelve regime-by-acceptance conditions at the primary route and cost configuration, each bargaining policy has at least one negative seller increment. Checkout alone has zero increment by definition. Descriptive finite-set losses in the supplement summarise these conditions without assigning probabilities to them. They are not a guarantee against regret, an optimized deployment policy or a reason to choose a policy after inspecting the evaluation outcomes.

Heterogeneity and structural robustness

Figure 5 separates seller and buyer effort costs. Its feasible region requires both average increments to be nonnegative; it does not imply individual benefit. Under stubborn offers, complete fallback creates only a narrow low-seller-cost region. Under low valuations, fallback does not move the boundary because no customer qualifies for full-price checkout.

The reference averages conceal pronounced product heterogeneity (Fig. 6). For full-state FQI, both product mean increments are positive in 82/193 contexts (42.5 percent); seller means are negative in 92 and buyer means in 107. The buyer median is −0.0016, despite its positive overall mean. These are simulated product-context frequencies, not measured customer harm rates or a validated product-selection rule. All products remain in the reported estimates.

Table 6 separates training-seed variability from product composition. All five reference full-state seeds yield positive overall increments for both parties, whereas only three yield positive buyer increments under stubborn offers. Category deletion preserves the reference signs but does not resolve within-category heterogeneity. These checks describe fixed learners and the existing evaluation sample.

Table 6
Full-state policy robustness across seeds and product contexts

Environment	Five-seed range	Products with both positive means	Category-deletion range
Reference	Seller 0.0141 to 0.0148 Buyer 0.0032 to 0.0045	82/193 (42.5%)	Seller 0.0121 to 0.0183 Buyer 0.0028 to 0.0056
Stubborn offers	Seller −0.0185 to −0.0170 Buyer −0.0022 to 0.0040	25/193 (13.0%)	Seller −0.0187 to −0.0145 Buyer −0.0019 to 0.0042
Low valuations	Seller 0.0070 to 0.0076 Buyer −0.0088 to −0.0083	21/193 (10.9%)	Seller 0.0044 to 0.0121 Buyer −0.0092 to −0.0078

Note. Ranges are descriptive minima and maxima, not confidence intervals. Primary accounting and acceptance settings are unchanged. Category deletion removes one represented category at a time and retains equal product weights among remaining products. No retraining or post hoc exclusion is performed. Monetary quantities are listing-price units per original opportunity. Full policy-level results are in Supplementary Tables S3 and S4 and Supplementary Data 1.

Fresh structural simulations further limit generalisation. With the higher floor multiplier, full-state seller and buyer increments are 0.0009 and −0.0040; with lower values they are 0.0204 and −0.0014. Under the lower-skew value distribution they are 0.0205 and −0.0056. Supplementary Table S5 and Supplementary Data 1 retain all eight settings and four policies, including adverse results. The reference replay matches the saved trajectories to an absolute tolerance of 10^{-12}. Changing unobserved valuation assumptions can change the substantive conclusion even when the software and reference averages reproduce exactly.

Discussion

Bargaining capability and the value of a transaction channel

The results distinguish the ability to negotiate from the value of offering negotiation. The full-state learner generates positive average seller and buyer increments in the reference configuration, yet it is not the policy with the highest combined value. Under stubborn offers, all evaluated policies reduce seller value at the primary settings. Under low valuations, enabling some otherwise unavailable transactions does not compensate buyers for the stipulated interaction burden. These are different consequences of the same optional transaction format, not evidence for a universal advantage of automated bargaining.

The decomposition explains these results more clearly than agreement rate alone. A completed transaction with a customer below the listing-value boundary can create new-sale margin. A discount to a customer who could have checked out instead transfers some existing margin to the buyer. An unresolved negotiation can expose a feasible checkout sale to loss, and both successful and unsuccessful exchanges use time. Holding the original customer opportunity fixed makes these mechanisms comparable. A higher number of bargains can therefore coexist with a negative increment for either participant.

This interpretation extends the practical use of pricing-format research without claiming to replace its economic models. Research on haggling costs and dual pricing already explains why negotiated and posted-price channels can interact (Desai and Purohit 2004; Wang 1995; Zeng et al. 2014). The present model supplies a way to evaluate that interaction for particular executable policies whose failures, prices and durations differ. Its contribution depends on those policy-specific outputs and the purchase routes they affect. Subtracting an identical baseline from every policy would not, by itself, create a new ranking.

Buyer and seller gains are distinct outcomes

Separate participant increments reveal tensions that a seller-only reward conceals. In the reference configuration, the full-state learner's seller advantage over the clock-only learner accompanies a lower buyer outcome and a slightly lower combined outcome. This does not show that richer information is intrinsically harmful. It shows that the value of information cannot be assessed independently of the objective used to train and compare policies. A seller-focused learning objective and a channel-level objective for both participants need not select the same specification.

Simultaneous nonnegative averages are a limited decision criterion. They do not guarantee that every individual benefits or that customers regard a discount as fair. Negotiators attach value to process, self-perception and relationships as well as economic outcomes (Curhan et al. 2006). Research on consumer experiences with AI and on human-agent bargaining likewise cautions against equating measured monetary gains with customer experience (Chaudhuri and De Bruyn 2026; Gelbrich et al. 2026; Puntoni et al. 2021). The present buyer cost is an explicit modelling parameter, not a scale measuring frustration, dignity, trust or satisfaction.

Negative buyer increments under assumed uptake are particularly important to interpret correctly. They identify configurations that fail the chosen average-value criterion; they do not establish that informed customers would voluntarily select an unattractive route. Actual participation could depend on expected savings, effort, uncertainty and beliefs about the counterpart. The value-dependent uptake analysis tests one stipulated composition change, but it does not estimate this choice process. It consequently supports caution about service configuration, not a behavioural claim about consumer irrationality.

Implications for an electronic commerce platform

Product-level losses suggest a further deployment requirement: predefine segment-level stop criteria rather than approve the channel from an overall average. The present outcomes cannot be used to select winning products and then serve as an unbiased evaluation of that selection. A prospective sample must test any resulting routing rule.

A platform can use the model in two stages. First it compares candidate bargaining policies under explicit objectives and records their information, training conditions and feasible actions. It then evaluates their outcomes against the existing checkout route on the same opportunity population. The output is a decision record specifying the checkout baseline, policy, affected population, expected value components and conditional cost and fallback limits. This is the operational framework produced by the study.

Before a merchant pilot, the stipulated inputs must be replaced or bounded using evidence about authorised price floors, checkout conversion, negotiation entry, return to checkout and participant effort. The cost ceilings are especially sensitive to how time and failed exchanges are valued. They should not be interpreted as recommended customer fees. The interface should keep direct checkout available, explain whether a displayed quote is an offer or a confirmed order, and allow customers to leave bargaining without ambiguous transaction status.

The existing numerical adapter demonstrates bounded linear-rule quotations, quote version and expiry fields, and confirmation checks. It does not establish a deployed learned-policy interface, inventory concurrency, persistent atomic quote consumption or payment integration. Similarly, the simulator's private customer value must never be interpreted as information available for real customer targeting. The study provides a testable framework for evaluating a proposed service before these engineering and behavioural questions are resolved.

What would establish broader applicability

The next substantive test is whether the specified mechanisms explain outcomes outside the simulator. A prospective comparison with actual checkout would need to measure participation, completion, fallback, effort and reported experience using real incentives and an appropriate ethics determination. Such evidence could change both the numerical thresholds and the assumed relationships between

costs and behaviour. A system in which effort causes earlier abandonment, for example, requires more than applying a larger cost to fixed trajectories.

Reproducibility is useful for that progression but does not substitute for it. Matched opportunities, frozen checkpoints and an independently written accounting check within the research workflow make the present comparison inspectable. They establish what follows from the implementation and its assumptions. External observations are still required to assess whether those assumptions describe a particular merchant's customers. The distinction keeps the model useful as a conditional evaluation framework without presenting a computational result as validated adoption evidence.

Limitations

The main limitation is the synthetic customer environment. The catalogue anchors product heterogeneity, while private values, merchant floors, offer dynamics, acceptance, abandonment, participation and fallback are stipulated. The analytical implications are valid under their definitions, but their numerical thresholds need not transport to another market. Product bootstrap intervals exclude this structural uncertainty.

The checkout benchmark is intentionally simple: value at least listing implies immediate purchase. Actual checkout depends on search, competition, shipping, payment and other frictions. Negotiation can also affect future visits, reference prices and trust; none is modelled. The model excludes multiple merchants, bundles, non-price terms, inventory dynamics, repeated customer interactions and equilibrium price changes. Fixed implementation and maintenance costs would introduce a scale condition absent from our proportional-cost model. Negative buyer increments under assumed uptake are feasibility warnings, not a claim that informed customers would knowingly select a losing route.

The learned policies were trained and evaluated within the same simulator family. New channel accounting and stress conditions do not create an independent external environment. Previously inspected evaluation partitions, rule-informed exploration and a limited set of action and model specifications require exploratory interpretation. The five seeds improve transparency about training variation, but their average is not a deployed ensemble or a learned selector. The study does not show superiority over contemporary conversational bargaining systems.

The sources play different roles. Craigslist price recognition lacks independent semantic validation, and initial-offer distributions can shift across settings. Amazon price extrema do not identify merchant costs. CaSiNo concerns human-human campsite allocation rather than retail use of the proposed agent. The supporting analyses neither pool their units nor establish multi-dataset replication of one effect.

Finally, the contribution is a conditional model and integrated computational comparison. It does not newly establish the economics of haggling costs or dual pricing, both addressed by prior research (Desai and Purohit 2004; Wang 1995; Zeng et al. 2014). Its narrower advance is to make executable policy

outcomes comparable at the channel decision level and to report the boundaries and failures of that comparison in a reproducible form.

Conclusion

Optional automated bargaining changes both the opportunities to transact and the allocation of value between buyers and sellers. This study develops a checkout-relative value model that makes those changes explicit through policy-specific agreement, price and duration outcomes. Its analytical output is a decomposition of incremental value and conditional interaction-cost and fallback limits. Its practical output is a two-stage framework that separates choosing a bargaining policy from deciding whether to offer that policy alongside checkout.

The matched computational comparison supports a conditional conclusion. Both participants gain on average in the reference configuration, but the full-state learner favoured by the seller does not maximise combined value. Stubborn offers and low valuations expose different sources of participant loss. Automated bargaining should therefore be evaluated against the transaction route it changes, with failures, displaced sales and participant burdens retained in the comparison. The present evidence establishes these relationships within a reproducible model; it does not establish actual customer participation or a commercially validated bargaining service. Those empirical tests are the next step towards reproducing useful aspects of shop-floor bargaining in electronic commerce.

Declarations

Competing interests

The authors declare no relevant financial or non-financial competing interests.

Ethical approval

The study uses secondary public research data and synthetic simulations. It includes human dialogue data but recruited no new participants and conducted no intervention. No study-specific institutional approval or formal exemption is claimed for this secondary analysis. The public availability of a dataset is not treated as evidence of an institutional exemption. Demographic and personality fields were not used. No individual participant identities or raw dialogue transcripts are reproduced in the article.

Informed consent

No new participant recruitment or consent collection was undertaken for this secondary analysis. The dialogue corpora were accessed as existing public research resources under their original release arrangements. The CaSiNo source publication reports institutional review and participant consent for the original collection (Chawla et al. 2021). This does not constitute approval of the present secondary

analysis. No new consent procedure or independently verified consent determination for CraigslistBargain is claimed; identifiable participant material is not reproduced.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Author Contribution

B.C contributed conceptualisation, methodology, software, formal analysis, investigation and writing of the original draft. M.S contributed supervision, methodological guidance, interpretation and critical revision. Both authors reviewed and approved the final manuscript.

References

1. Baarslag T, Fujita K, Gerding EH, Hindriks K, Ito T, Jennings NR, Jonker C, Kraus S, Lin R, Robu V, Williams CR (2013) Evaluating practical negotiating agents: Results and analysis of the 2011 international competition. Artif Intell 198:73–103. https://doi.org/10.1016/j.artint.2012.09.004
2. Baarslag T, Hendrikx MJC, Hindriks KV, Jonker CM (2016) Learning about the opponent in automated bilateral negotiation: A comprehensive survey of opponent modeling techniques. Auton Agents Multi-Agent Syst 30(5):849–898. https://doi.org/10.1007/s10458-015-9309-1
3. Backus M, Blake T, Larsen B, Tadelis S (2020) Sequential bargaining in the field: Evidence from millions of online bargaining interactions. Q J Econ 135(3):1319–1361. https://doi.org/10.1093/qje/qjaa003
4. Bagga P, Paoletti N, Alrayes B, Stathis K (2021) ANEGMA: An automated negotiation model for e-markets. Auton Agents Multi-Agent Syst 35(2):27. https://doi.org/10.1007/s10458-021-09513-x
5. Bakos JY (1997) Reducing buyer search costs: Implications for electronic marketplaces. Manag Sci 43(12):1676–1692. https://doi.org/10.1287/mnsc.43.12.1676
6. Bichler M (2026) Agentic markets. Electron Mark 36:55. https://doi.org/10.1007/s12525-026-00906-y
7. Brynjolfsson E, Smith MD (2000) Frictionless commerce? A comparison of Internet and conventional retailers. Manag Sci 46(4):563–585. https://doi.org/10.1287/mnsc.46.4.563.12061
8. Cao M, Luo X, Luo X, Dai X (2015) Automated negotiation for e-commerce decision making: A goal deliberated agent architecture for multi-strategy selection. Decis Support Syst 73:1–14. https://doi.org/10.1016/j.dss.2015.02.012
9. Cason TN, Friedman D, Milam GH (2003) Bargaining versus posted price competition in customer markets. Int J Ind Organ 21(2):223–251. https://doi.org/10.1016/S0167-7187(02)00056-5

10. Castelo N, Bos MW, Lehmann DR (2019) Task-dependent algorithm aversion. J Mark Res 56(5):809–825. https://doi.org/10.1177/0022243719851788
11. Chaudhuri S, De Bruyn A (2026) Bots bargaining with humans: Building AI super-bargainers with algorithmic anthropomorphization. J Mark Res 63(2):234–256. https://doi.org/10.1177/00222437251375323
12. Chawla K, Ramirez J, Clever R, Lucas G, May J, Gratch J (2021) CaSiNo: A corpus of campsite negotiation dialogues for automatic negotiation systems. In: Toutanova K, Rumshisky A, Zettlemoyer L, Hakkani-Tur D, Beltagy I, Bethard S, Cotterell R, Chakraborty T, Zhou Y (eds) Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics, pp 3167–3185. https://doi.org/10.18653/v1/2021.naacl-main.254
13. Collins AJ, Koehler M, Lynch CJ (2024) Methods that support the validation of agent-based models: An overview and discussion. J Artif Soc Soc Simul 27(1):11. https://doi.org/10.18564/jasss.5258
14. Curhan JR, Elfenbein HA, Xu H (2006) What do people value when they negotiate? Mapping the domain of subjective value in negotiation. J Pers Soc Psychol 91(3):493–512. https://doi.org/10.1037/0022-3514.91.3.493
15. Desai PS, Purohit D (2004) Let me talk to my manager: Haggling in a competitive environment. Mark Sci 23(2):219–233. https://doi.org/10.1287/mksc.1040.0045
16. Dietvorst BJ, Simmons JP, Massey C (2015) Algorithm aversion: People erroneously avoid algorithms after seeing them err. J Exp Psychol Gen 144(1):114–126. https://doi.org/10.1037/xge0000033
17. Ernst D, Geurts P, Wehenkel L (2005) Tree-based batch mode reinforcement learning. J Mach Learn Res 6:503–556. https://www.jmlr.org/papers/v6/ernst05a.html
18. Faratin P, Sierra C, Jennings NR (1998) Negotiation decision functions for autonomous agents. Robot Auton Syst 24(3–4):159–182. https://doi.org/10.1016/S0921-8890(98)00029-3
19. Gelbrich K, Roschk H, Miederer S, Kerath A (2026) Automated versus human agents: A meta-analysis of customer responses to robots, chatbots, and algorithms and their contingencies. J Mark 90(2):1–26. https://doi.org/10.1177/00222429251344139
20. Gill D, Thanassoulis J (2009) The impact of bargaining on markets with price takers: Too many bargainers spoil the broth. Eur Econ Rev 53(6):658–674. https://doi.org/10.1016/j.euroecorev.2008.10.004
21. Glikson E, Woolley AW (2020) Human trust in artificial intelligence: Review of empirical research. Acad Manag Ann 14(2):627–660. https://doi.org/10.5465/annals.2018.0057
22. Hann I-H, Terwiesch C (2003) Measuring the frictional costs of online transactions: The case of a name-your-own-price channel. Manag Sci 49(11):1563–1579. https://doi.org/10.1287/mnsc.49.11.1563.20586
23. He H, Chen D, Balakrishnan A, Liang P (2018) Decoupling strategy and generation in negotiation dialogues. In: Riloff E, Chiang D, Hockenmaier J, Tsujii J (eds) Proceedings of the 2018 Conference

on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, pp 2333–2343. https://doi.org/10.18653/v1/D18-1256

24. Herold S, Heller J, Rozemeijer F, Mahr D (2025) Brave new procurement deals: An experimental study of how generative artificial intelligence reshapes buyer–supplier negotiations. J Purch Supply Manag 31(4):101012. https://doi.org/10.1016/j.pursup.2025.101012
25. Hinz O, Hann I-H, Spann M (2011) Price discrimination in e-commerce? An examination of dynamic pricing in name-your-own price markets. MIS Q 35(1):81–98. https://doi.org/10.2307/23043490
26. Huang M-H, Rust RT (2021) A strategic framework for artificial intelligence in marketing. J Acad Mark Sci 49(1):30–50. https://doi.org/10.1007/s11747-020-00749-9
27. Jennings NR, Faratin P, Lomuscio AR, Parsons S, Wooldridge MJ, Sierra C (2001) Automated negotiation: Prospects, methods and challenges. Group Decis Negot 10(2):199–215. https://doi.org/10.1023/A:1008746126376
28. Jindal P, Newberry P (2018) To bargain or not to bargain: The role of fixed costs in price negotiations. J Mark Res 55(6):832–851. https://doi.org/10.1177/0022243718818451
29. Kwon D, Hae J, Clift E, Shamsoddini D, Gratch J, Lucas GM (2025) ASTRA: A negotiation agent with adaptive and strategic reasoning via tool-integrated action for dynamic offer optimization. In: Christodoulopoulos C, Chakraborty T, Rose C, Peng V (eds) Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, pp 16217–16238. https://doi.org/10.18653/v1/2025.emnlp-main.821
30. Li J, He C, Shi Z (2025) Platform search design with heterogeneous consumers. Humanit Soc Sci Commun 12:673. https://doi.org/10.1057/s41599-025-04932-7
31. Logg JM, Minson JA, Moore DA (2019) Algorithm appreciation: People prefer algorithmic to human judgment. Organ Behav Hum Decis Process 151:90–103. https://doi.org/10.1016/j.obhdp.2018.12.005
32. Longoni C, Cian L (2022) Artificial intelligence in utilitarian vs. hedonic contexts: The 'word-of-machine' effect. J Mark 86(1):91–108. https://doi.org/10.1177/0022242920957347
33. Luo X, Tong S, Fang Z, Qu Z (2019) Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. Mark Sci 38(6):937–947. https://doi.org/10.1287/mksc.2019.1192
34. Oh J, Aghazada M, Shin Y, Yun S-Y, Kim T (2026) MERIT feedback elicits better bargaining in LLM negotiators. In: Liakata M, Moreira VP, Zhang J, Jurgens D (eds) Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Association for Computational Linguistics, pp 25202–25223. https://doi.org/10.18653/v1/2026.acl-long.1155
35. Pan X, Ratchford BT, Shankar V (2004) Price dispersion on the internet: A review and directions for future research. J Interact Mark 18(4):116–135. https://doi.org/10.1002/dir.20019
36. Power DJ, Sharda R (2007) Model-driven decision support systems: Concepts and research directions. Decis Support Syst 43(3):1044–1061. https://doi.org/10.1016/j.dss.2005.05.030

37. Puntoni S, Reczek RW, Giesler M, Botti S (2021) Consumers and artificial intelligence: An experiential perspective. J Mark 85(1):131–151. https://doi.org/10.1177/0022242920953847
38. Raskovich A (2007) Competition or collusion? Negotiating discounts off posted prices. Int J Ind Organ 25(2):341–354. https://doi.org/10.1016/j.ijindorg.2006.05.003
39. Sargent RG (2013) Verification and validation of simulation models. J Simul 7(1):12–24. https://doi.org/10.1057/jos.2012.20
40. Spann M, Skiera B, Schäfers B (2004) Measuring individual frictional costs and willingness-to-pay via name-your-own-price mechanisms. J Interact Mark 18(4):22–36. https://doi.org/10.1002/dir.20022
41. Spann M, Bertini M, Koenigsberg O, Zeithammer R, Aparicio D, Chen Y, Fantini F, Jin GZ, Morwitz VG, Popkowski Leszczyc P, Vitorino MA, Williams GY, Yoo H (2026) Algorithmic pricing: Implications for marketing strategy and regulation. Int J Res Mark 43(2):560–579. https://doi.org/10.1016/j.ijresmar.2025.05.001
42. Terwiesch C, Savin S, Hann I-H (2005) Online haggling at a name-your-own-price retailer: Theory and application. Manag Sci 51(3):339–351. https://doi.org/10.1287/mnsc.1040.0337
43. Wang R (1995) Bargaining versus posted-price selling. Eur Econ Rev 39(9):1747–1764. https://doi.org/10.1016/0014-2921(95)90043-8
44. Windrum P, Fagiolo G, Moneta A (2007) Empirical validation of agent-based models: Alternatives and prospects. J Artif Soc Soc Simul 10(2):8. https://www.jasss.org/10/2/8.html
45. Xia T, He Z, Ren T, Miao Y, Zhang Z, Yang Y, Wang R (2024) Measuring bargaining abilities of LLMs: A benchmark and a buyer-enhancement method. In: Ku L-W, Martins A, Srikumar V (eds) Findings of the Association for Computational Linguistics: ACL 2024. Association for Computational Linguistics, pp 3579–3602. https://doi.org/10.18653/v1/2024.findings-acl.213
46. Zeng X, Dasgupta S, Weinberg CB (2014) The effects of a no-haggle channel on marketing strategies. Int J Res Mark 31(4):434–443. https://doi.org/10.1016/j.ijresmar.2014.06.002

Figures

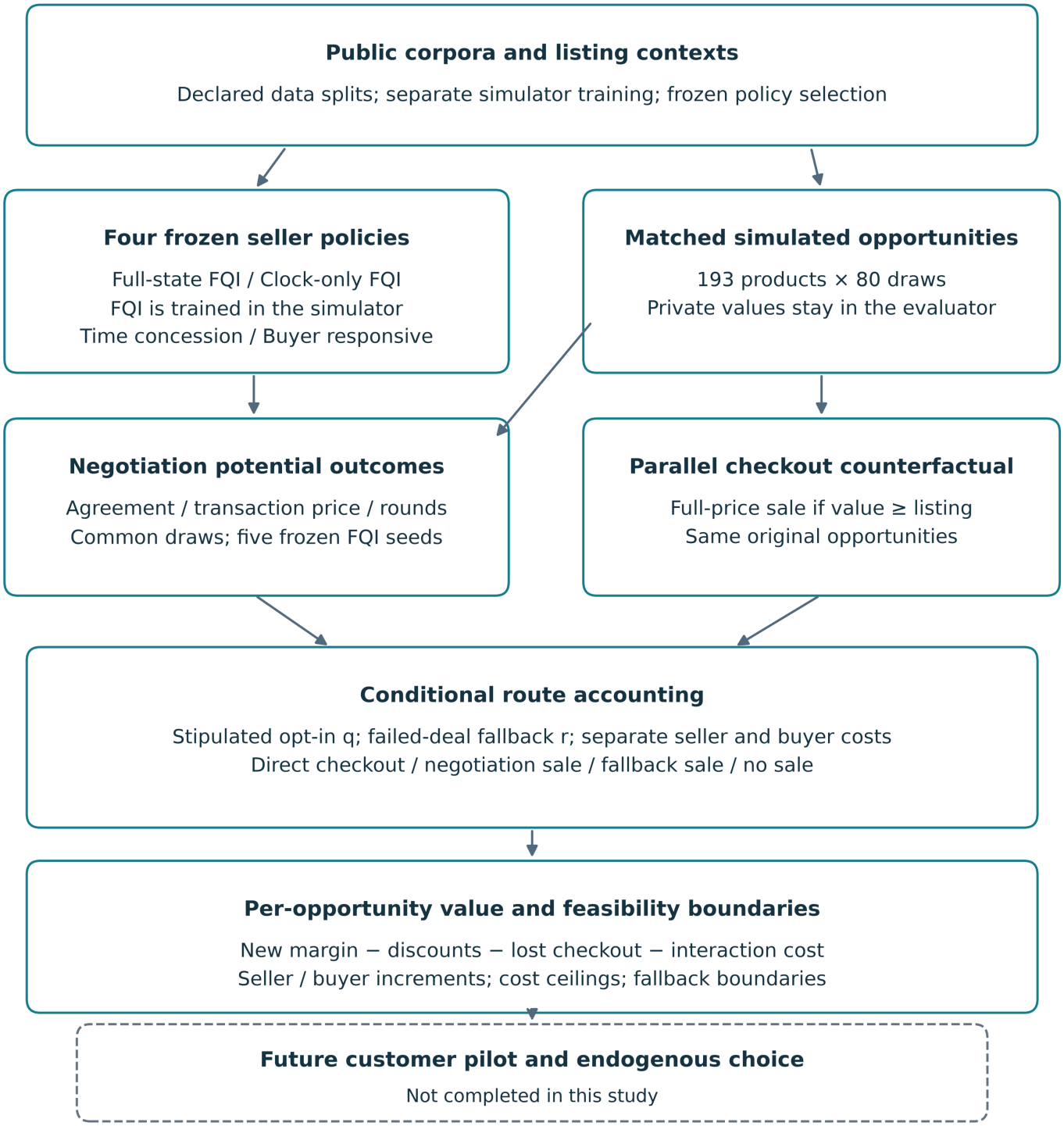


Figure 1

Checkout-relative evaluation workflow. Frozen policy simulations and a checkout counterfactual feed conditional route accounting. The dashed final box identifies future work, not a deployed component. Private synthetic values remain inside the evaluation environment

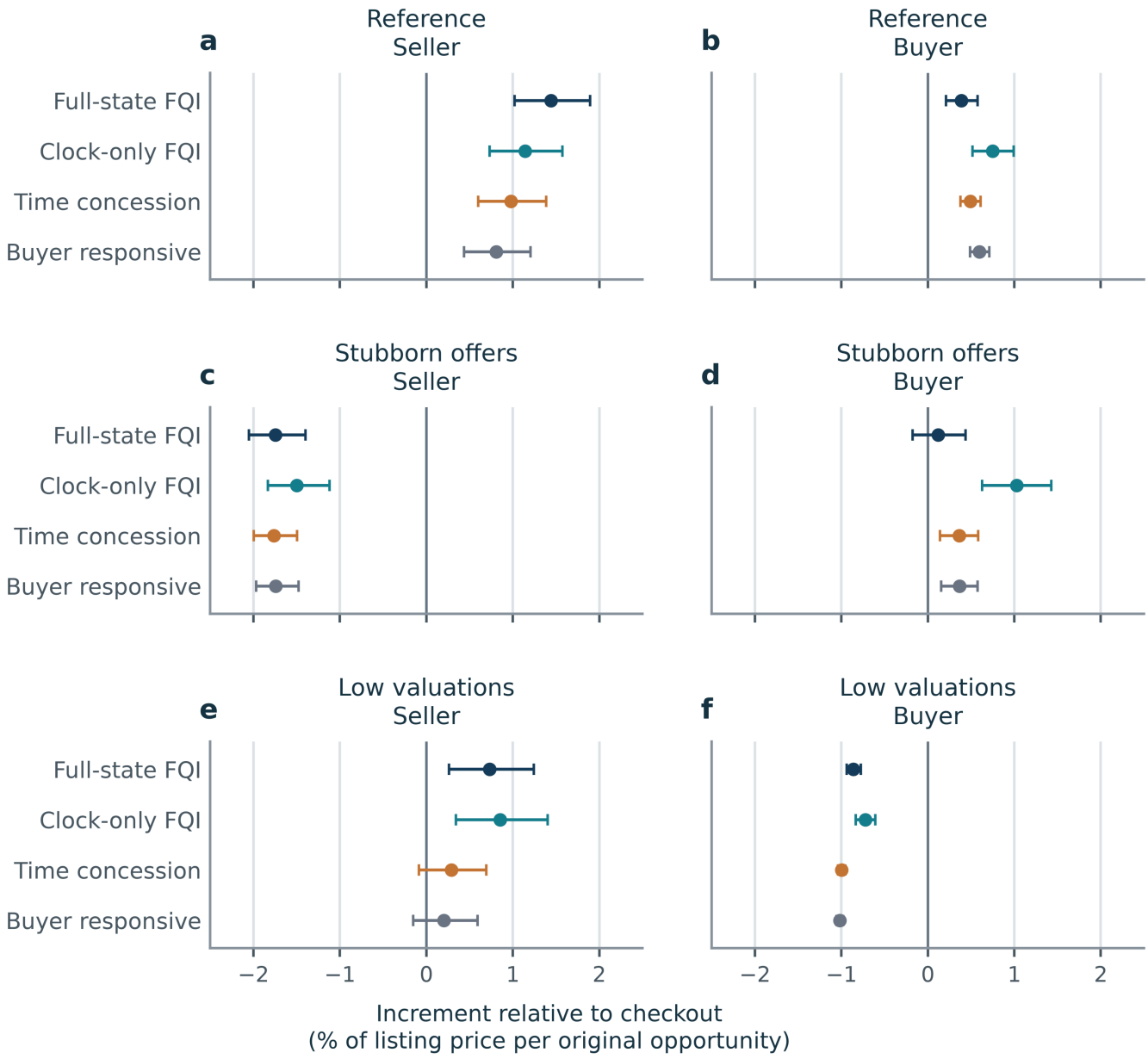


Figure 2

Primary checkout-relative participant increments. Panels a–b show the reference environment, c–d stubborn offers and e–f low valuations; left panels show seller and right panels buyer outcomes. Each symbol is the mean across 193 product means, each averaging 80 opportunities and, for learned policies, five fixed training seeds. Horizontal bars are marginal 95% percentile intervals from 2,000 paired product-bootstrap resamples. The vertical line marks zero. Uptake and fallback are 0.5, acceptance is 0.25, and each participant incurs 0.005 listing-price units per round. Units are percent of listing price per original opportunity



Figure 3

Seller-value decomposition at the primary configuration. Rightward bars are newly enabled margin; leftward bars are discounts, unrecovered checkout margin and interaction costs. Diamonds mark the net increment. All components use percent of listing price per original opportunity; panels share the same scale. Panels a, b and c show reference, stubborn-offer and low-valuation environments. Bars are point estimates over 193 products, with repetitions and learned seeds averaged within product

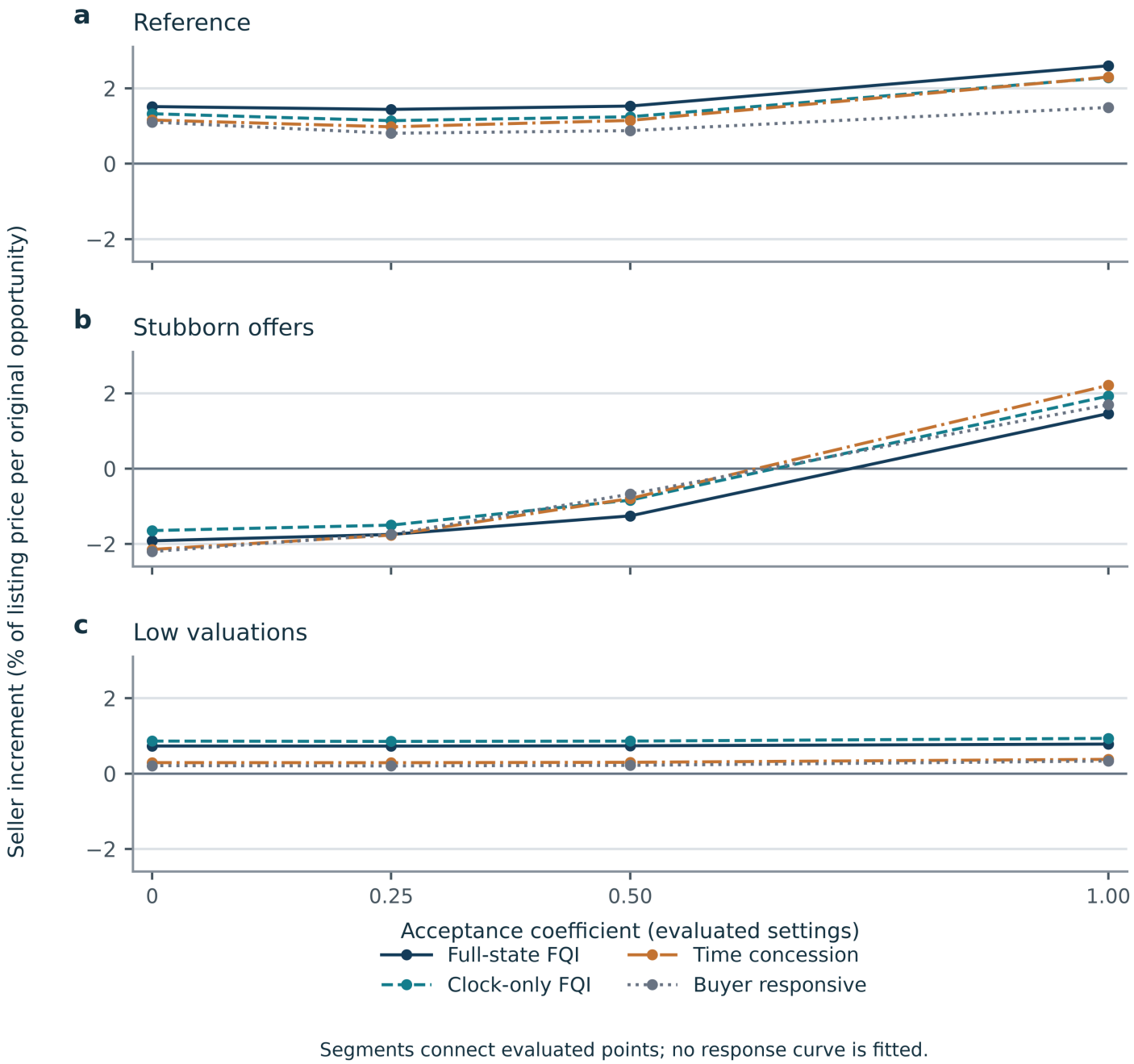


Figure 4

Seller increment across stipulated acceptance coefficients. Lines connect the four evaluated coefficients as a visual guide, not a fitted behavioural function. Uptake and fallback are 0.5 and participant cost is 0.005 per round. These are descriptive point estimates; full product-bootstrap intervals are supplied in Supplementary Table S2 in Supplementary Data 1 (S2_acceptance_intervals.csv). Panels a, b and c show reference, stubborn-offer and low-valuation environments. Means use 193 products and the same within-product averaging as Fig. 2

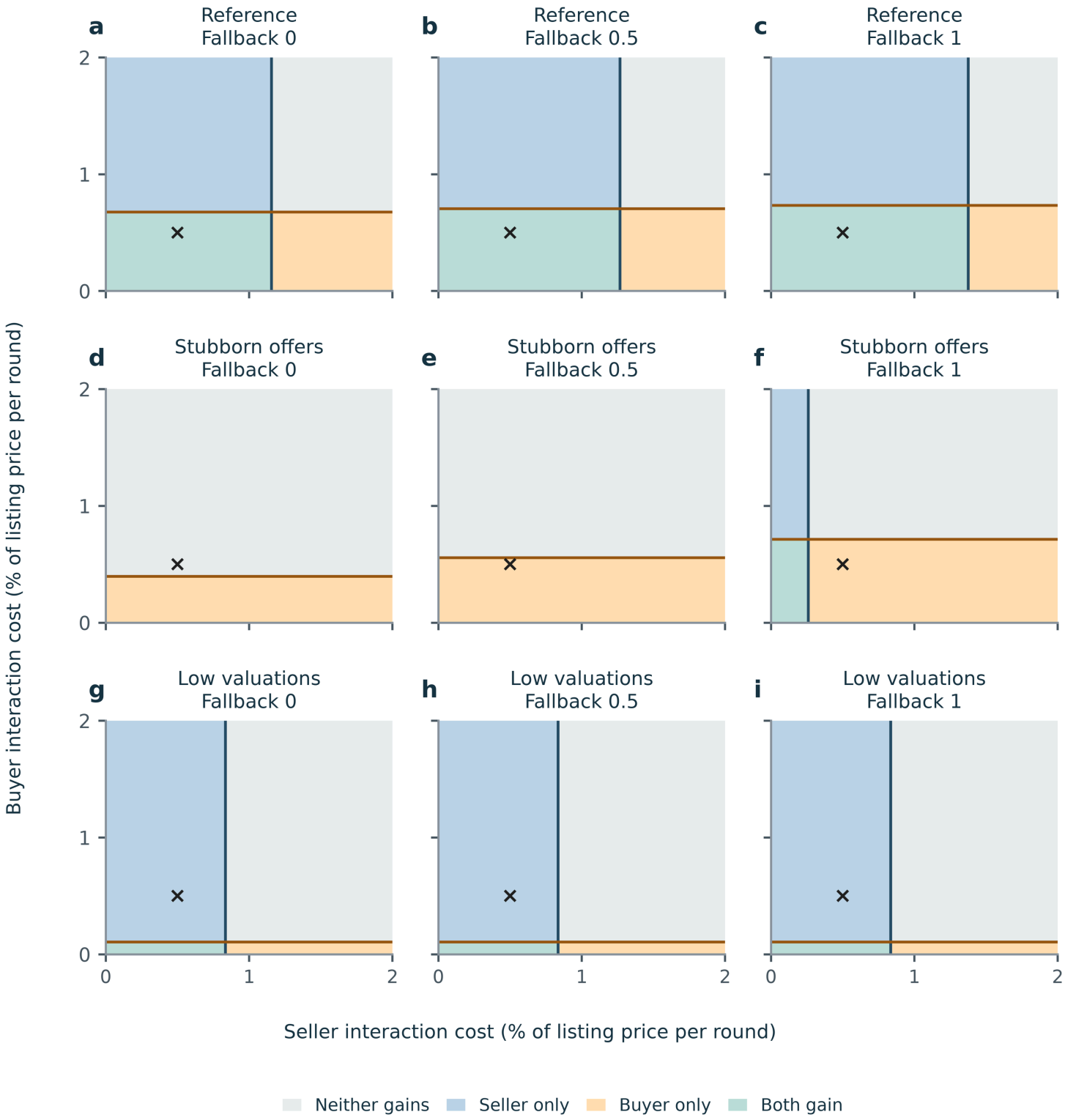


Figure 5

Separate-cost feasibility regions for full-state FQI. Panels a–c, d–f and g–i show reference, stubborn-offer and low-valuation environments; columns vary fallback from 0 to 0.5 to 1. Colours classify the signs of average seller and buyer increments over 193 products, not individual Pareto improvements. Lines mark zero-increment boundaries; crosses mark the primary separate costs of 0.5% of listing price per round. Uptake is exogenous 0.5 and acceptance is 0.25. Regions are point estimates, not confidence

regions; marginal boundary intervals at 101 fallback settings are in Supplementary Data 1 (cost_boundaries.csv). No interpolation of customer behaviour is used

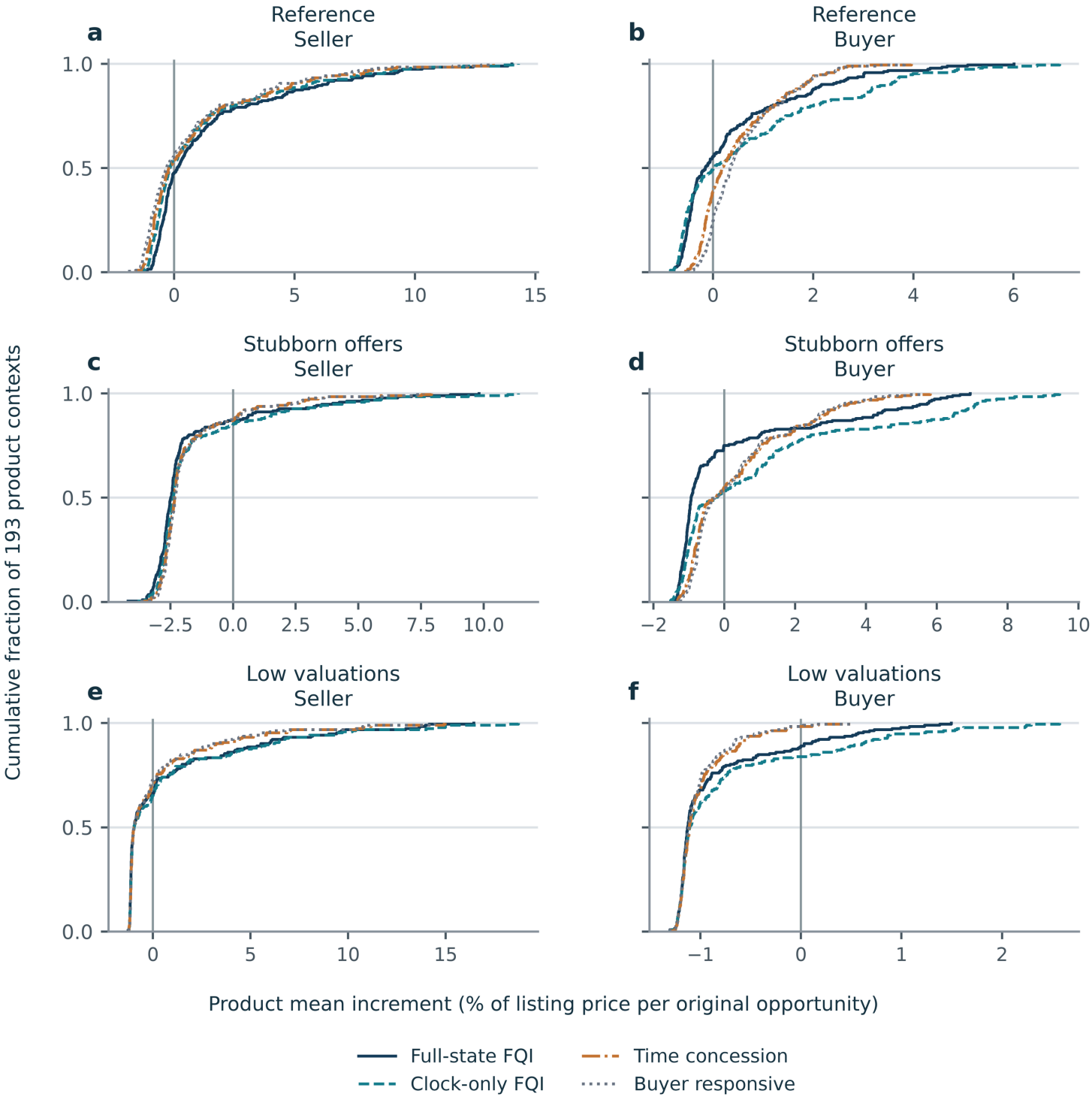


Figure 6

Distributions of product mean increments. Panels a–b show reference, c–d stubborn offers and e–f low valuations; left panels show sellers and right panels buyers. Each curve is the empirical cumulative distribution of 193 product means, averaging 80 opportunities and five frozen seeds for learned policies. Lines distinguish policies; the vertical line marks zero. Axes use panel-specific horizontal ranges.

Primary uptake, fallback, acceptance and separate costs apply. These descriptive product distributions are not confidence intervals or individual customer outcome distributions

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- HSSCSupplementaryInformation.docx
- SupplementaryData1.zip