DIM-WM: FORECAST-INFORMED SERVICES IN AI-AGENT DATA MARKETS

Anonymous authors
Paper under double-blind review

ABSTRACT

Can platform forecasts improve data-market decisions when participants retain final control? We study this question in an auditable simulated data market with LLM buyers and sellers used without further training, whose decisions can respond to observations and accumulated interaction history. Task utility is measured on real classification data. DIM-WM organizes five conditional prediction interfaces: buyer choice (Natural), legal-action outcomes (Controlled), task gain (Technical), 14-day seller outcomes (Core14), and 28-/42-day seller outcomes (Dynamic42). Buyer net technical surplus is realized task gain minus payment on an experimental accounting scale, with nontransactions counted as zero. Predictive performance is target-dependent, and stronger public-input baselines outperform some evaluated components. Fixed-support development selects a quantile LightGBM gain predictor and finds a positive value-information signal using archived action outcomes. In an 80-block online format-control study, individualized descriptive forecasts do not establish a buyer-surplus benefit over same-format training-reference statistics or no advice; both point estimates are negative. A separate 72-blueprint online study gives individualized descriptive forecasts to both advised arms and varies the buyer-value package. Individualizing that package increases mean 42-day buyer surplus by 41.16 accounting units per market relative to reference value information (95% CI [27.68, 55.08]). The full service increases surplus by 46.44 units relative to no advice (95% CI [29.35, 64.68]), with fewer transactions and lower seller gross revenue on average. The results support a buyer-surplus benefit from the evaluated value-information service while distinguishing predictive quality, fixed-state decision value, and online market effects.

1 INTRODUCTION

Data buyers often decide whether to purchase a product before knowing whether it will improve their own task. Because gains depend on the buyer's existing data, predicting purchase behavior is not the same as estimating the value of a purchase. Prior work studies data purchasing, utility appraisal, value signaling, and marketplace interaction (Azcoitia & Laoutaris, 2022; Xu et al., 2022; Peng et al., 2025; Wu et al., 2026). We focus on the platform as an information provider: can its forecasts improve decisions when buyers and sellers retain final control?

This question requires separating prediction targets from information use. *Descriptive forecasts* concern choices, transaction outcomes, and seller trajectories, whereas *explicit value information* relates predicted task gain to a visible offer. The same prediction can support a fixed scoring rule or a nonbinding advisor, and in continuing markets advice changes later public histories. We therefore evaluate predictive quality, fixed-state decision value, and online service effects separately.

DIM-WM studies these distinctions in a simulated market with LLM buyers and sellers used without further training, whose decisions can respond to observations and accumulated interaction history. An independent evaluator measures task gain on four real classification data lineages; buyer net technical surplus is realized task gain minus payment on an experimental accounting scale, with nontransactions contributing zero. The five conditional interfaces predict buyer choice (Natural), legal-action outcomes (Controlled), task gain (Technical), 14-day seller outcomes (Core14), and 28-/42-day seller outcomes (Dynamic42). The service is designed around platform-visible information; private buyer data are used to construct utility labels but are not supplied prospectively to predic-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

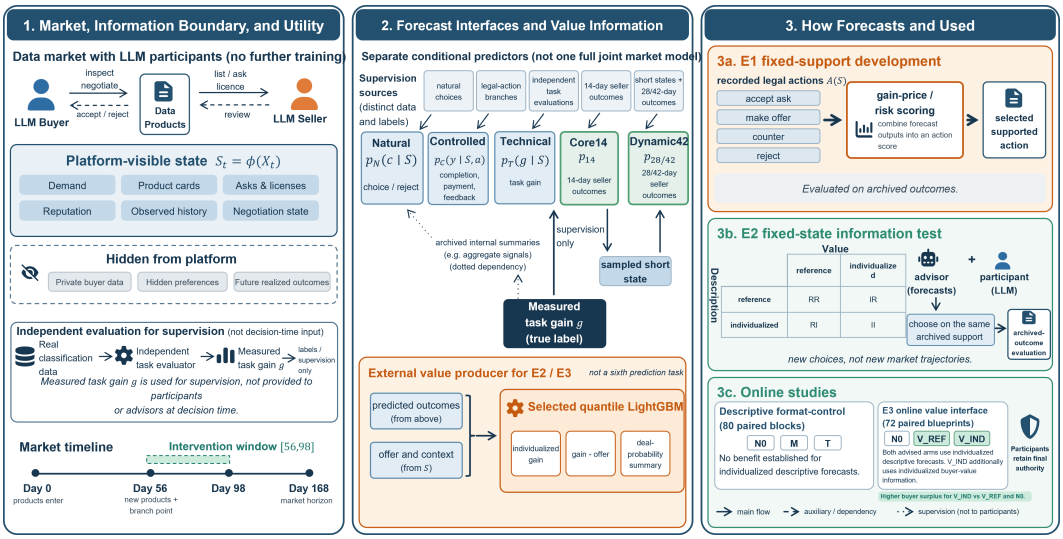

Figure 1: DIM-WM evaluation structure. An independent evaluator measures task gains in a market with LLM participants used without further training. Five conditional interfaces organize the prediction targets, with the selected quantile LightGBM supplying task-gain summaries for explicit value information. E1/E2 use archived action outcomes, whereas the online studies evaluate continuing market branches. In advice-based evaluations, participants retain final decision authority.

tors. Historical temporal evaluations additionally condition on trajectory-arrival capacities, and the available records do not fully establish the mapping of those conditions to all online temporal inputs (Appendix B). Figure 1 summarizes the market, information interfaces, and evaluation settings.

We ask whether participants respond to market conditions (RQ1), which prediction targets contain usable information (RQ2), and whether forecast-informed scoring or advice improves buyer surplus (RQ3). We distinguish fixed-support choices evaluated on archived action outcomes from online advice evaluated in paired continuing markets.

Predictive performance is target-dependent, with stronger baselines outperforming some original components. On shared 60-blueprint support, E1 selects a quantile LightGBM gain predictor under a fixed gain–price rule, and E2 finds a positive effect of individualized value information on new participant choices evaluated against the same archived outcomes. Both are development analyses, not independent confirmation.

An 80-block online format-control study does not establish a buyer-surplus benefit from individualized descriptive forecasts over same-format training-reference statistics or no advice; both primary point estimates are negative. A separate 72-blueprint study (E3) keeps individualized descriptive forecasts in both advised arms while changing the source of buyer-value information. Individualizing this package increases mean 42-day buyer surplus by 41.16 units per market relative to the reference-value arm; the full service increases surplus by 46.44 units relative to no advice, with fewer transactions and lower seller gross revenue on average. The two online studies are not a matched test of descriptive versus value-interface superiority.

Our empirical contribution is a controlled evaluation of forecast services for autonomous LLM market participants, separating estimator substitution under a fixed rule, information-source substitution at fixed states, and paired interventions in continuing markets. The principal online finding concerns the incremental effect of a buyer-value package over an already individualized descriptive service, together with its transaction and revenue tradeoffs.

2 Related Work

Data markets and value signals. Data-market research studies pricing, allocation, and pre-purchase information under heterogeneous buyer tasks. Try-Before-You-Buy estimates task performance before purchase, while data appraisal and utility-signaling methods study value evidence

under sharing or privacy constraints (Azcoitia & Laoutaris, 2022; Xu et al., 2022; Peng et al., 2025). Other work develops algorithmic data marketplaces and model-based pricing (Agarwal et al., 2019; Chen et al., 2019), as well as contribution-based valuation using Shapley-style or learned estimators (Ghorbani & Zou, 2019; Jia et al., 2019; Kwon & Zou, 2022; Yoon et al., 2020). DUPRE predicts utility to reduce expensive valuation calls (Pham et al., 2025), and recent marketplace designs learn from observed bids or simulate strategic data-market agents (Gao et al., 2026; Wu et al., 2026). Our study contributes an evaluation design for forecast and value information as market services: estimator replacement under a fixed action rule, information-source replacement at fixed states, and paired service interventions in continuing markets.

Prediction, optimization, and deployment feedback. Predict-then-optimize and decision-focused learning examine the mismatch between predictive losses and downstream decision loss (Elmachtoub & Grigas, 2022; Donti et al., 2017; Wilder et al., 2019; Mandi et al., 2024). Prescriptive analytics studies the use of predictions in decision making, while performative prediction considers feedback from deployed decisions to the environment being predicted (Bertsimas & Kallus, 2020; Perdomo et al., 2020). DIM-WM holds predictors fixed within each evaluation and separately examines action scoring, nonbinding advice, and continuing-market outcomes. The shared-support E1/E2 analyses are treated as development evidence, consistent with concerns about post-selection evaluation (Cawley & Talbot, 2010).

LLM agents, advice, and reliance. Generative-agent and LLM-simulation studies motivate using language models as interactive participants, while requiring behavioral validation rather than assuming human equivalence (Park et al., 2023; Aher et al., 2023; Argyle et al., 2023; Binz & Schulz, 2023). Human–AI decision-support research shows that explanations and advice format can affect reliance and team performance (Bansal et al., 2021; Buçinca et al., 2021; Schemmer et al., 2023). We therefore examine participants' responses to controlled conditions and evaluate advice through realized market outcomes. Advice remains nonbinding, and aggregate service effects do not identify recommendation adoption or its behavioral mechanism.

Modular prediction and world models. Related world-model research studies compositional or causal dynamics in other domains (Zhang et al., 2025; Lei et al., 2022; Zhang et al., 2026). DIM-WM organizes heterogeneous market targets into five conditional prediction interfaces, each with its own inputs and evaluation conditions. The empirical question is how these interfaces behave when used as decision information: which predictive targets are informative, which fixed-state substitutions change selected actions, and which online services improve or redistribute market outcomes.

3 Problem Setup and Research Design

3.1 Market Dynamics and Participants

The continuing market contains 16 buyers and eight sellers across four classification lineages: electricity, MiniBooNE, pol, and nomao. Each seller introduces one product on day 0 and another on day 56, adding products rather than participants. Buyers may inspect products, negotiate, accept, reject, and later review completed transactions; sellers set and revise quotes under legal action rules. Buyers and sellers are LLM participants used without further training, whose decisions can respond to observations and accumulated interaction history. Exogenous arrivals create decision opportunities, not guaranteed purchases. The market runs for 168 days, the online interventions cover days $[56, 98)$, and prediction horizons are measured relative to their checkpoints. Appendix A specifies the arrival process, product construction, and event dynamics.

3.2 Task Utility and Market Accounting

Task utility is measured independently of the dialogue. For a recorded buyer dataset and candidate product, the evaluator fits the same logistic-regression procedure before and after concatenating the product, then evaluates both models on the same utility pool. With baseline and augmented log losses L_{base} and L_{aug}, define

$$g = 100 \left(1 - \frac{L_{\text{aug}}}{L_{\text{base}}}\right), \qquad U = D(g - p_{\text{paid}}), \tag{1}$$

Table 1: Evaluation settings and analysis units. Repeated decisions, rollouts, and continuing branches are not interchangeable independent-market observations.

Study	Scope	Setting
RQ1 behavior	17,856 LLM decisions	Matched worlds and controlled behavioral observations
RQ2 prediction	Five conditional prediction tasks	Task-specific records and splits
E1	60 blueprints; 239 parents; 3,830 actions	Fixed-rule replacement on archived outcomes
E2	Same 60 blueprints; 956 parent–arm cells	New choices scored on the same archived outcomes
Historical online	206 paired blueprints; seven arms	Continuing markets; historical advice package
Descriptive online	80 paired blocks; 320 branches	Continuing markets; same-format reference control
E3 value online	72 paired blueprints; 216 branches	Continuing markets; reference or individualized value

where $D \in \{0, 1\}$ indicates transaction completion and p_{paid} is the realized payment. Task gain g can be negative. Nontransactions contribute zero realized gain, payment, and surplus, without assigning zero potential gain to an unpurchased product.

Payments and gains share an experimental accounting scale. Day-0 launch prices are scaled using positive task gains from training-only buyer–product pairs. Subsequent quotes and negotiations remain within lineage bounds, so price–value comparisons are conditional on an informative initialization rather than a proof of price discovery without a prior utility scale. Seller gross revenue is not profit, because comparable seller costs are not modeled.

Online outcomes accumulate over the 42-day intervention window before averaging across market blueprints. For blueprint b and arm h, let $G_b(h)$ and $R_b(h)$ denote total realized task gain and seller gross revenue. The total buyer net technical surplus is

$$Y_b(h) = G_b(h) - R_b(h). \tag{2}$$

These are market-level accounting totals, not per-transaction averages or final-model performance after pooling all purchases. Appendix A.2 documents the recorded buyer-data snapshots and the remaining uncertainty about post-purchase state updates and utility caching.

3.3 Information Boundary

Let X_t denote the simulator state and $S_t = \phi(X_t)$ its platform-visible projection, including demand regime, product cards, asks, licenses, reputation, and visible transaction and negotiation history. Predictors use specified subsets and transformations of these fields; S_t is not assumed to be a sufficient Markov state. The public-input contract excludes private buyer data, hidden preferences, future realized outcomes, test aggregates, and empirically best actions, although the independent evaluator uses buyer data to construct labels. The service is designed around platform-visible state. Historical temporal evaluations additionally condition on actual trajectory-arrival capacities; the available records do not fully establish the mapping of those conditions to all online temporal inputs. Appendix B records this qualification.

3.4 Evaluation Settings and Statistical Units

RQ1 evaluates behavior in matched worlds and controlled persona, economic, and history conditions. RQ2 evaluates task-specific prediction records. RQ3 distinguishes fixed-support legal actions scored against recorded outcomes from online advice that generates continuing market trajectories.

A market blueprint identifies the instance used to group recorded states or matched continuing branches. E1 and E2 share 60 blueprints, 239 parent states, and 3,830 legal parent–action pairs with 30 recorded rollouts per action; neither creates new market trajectories. The three online studies retain separate input pools, treatments, and comparison families (Table 1).

Predictors and their designated calibration maps remain fixed within each evaluation. Matched predictive comparisons use the same target and evaluation pool; historical Test-access qualifications and exploratory temporal analyses remain separate from later Validation comparisons. Online primary

effects are paired differences in 42-day market totals. Reported intervals use paired bootstrap replicates, and online p values use paired sign-flip calculations with study-specific Holm families (Efron, 1979; Holm, 1979). Intervals are not simultaneous family-wise intervals, and sign-flip inference retains its exchangeability or sign-invariance assumptions. Appendix E specifies the procedures and comparison families.

4 DIM-WM AND INFORMATION INTERFACES

4.1 PREDICTION OBJECTS

DIM-WM organizes five conditional prediction interfaces:

$$\mathcal{M} = \left\{ p_{\mathrm{N}}(c \mid S), p_{\mathrm{C}}(y \mid S, a), p_{\mathrm{T}}(g \mid S, j), p_{14}, p_{28/42} \right\}, \tag{3}$$

where c is a buyer choice, a is a legal action, y contains its applicable outcomes, and j identifies the focal product.

Natural predicts visible buyer choice or rejection; Controlled predicts completion, conditional payment, and applicable feedback for a legal action; Technical predicts independently measured task gain with price and action excluded. Core14 predicts 14-day seller outcomes, and Dynamic42 propagates sampled short states to 28- and 42-day seller counts and revenue. The interfaces have separate inputs, targets, splits, and calibration procedures rather than a single joint full-market model. Natural's internal Technical features and the explicit value summaries below are distinct uses of task-gain information. Appendix B gives the model definitions and version-specific limitations.

4.2 EXPLICIT BUYER-VALUE SUMMARIES

The explicit value path uses the quantile LightGBM predictor selected in E1 (Ke et al., 2017). It predicts task gain from the 64-dimensional structured public buyer–product representation, without text features, price, or action. Five quantile regressors are fitted. For each query, their predicted quantiles are sorted, linearly interpolated with constant endpoint extension, and averaged on a fixed 192-point grid. No additional post-hoc calibration is applied. Appendix B gives the target, quantile levels, hyperparameters, and exact summary equation.

This predictor is an alternative estimator for the Technical task, not a sixth interface. It supplies explicit value information in E2 and E3, but does not replace Natural's historical Technical features or regenerate the descriptive-advice studies. Explicit value information relates a gain summary to a visible supported offer, distinct from a choice forecast or realized-payment estimate. E3's reference and individualized fields are specified in Table 43.

4.3 SUPPORTED-ACTION SCORING AND NONBINDING ADVICE

Supported-action scoring. At a recorded parent state S, a fixed-support policy selects from the legal action set $\mathcal{A}(S)$. For product action a, let $j(a)$ identify its product and $\mathrm{offer}(a)$ its visible supported offer. The gain–price rule is

$$s_{\mathrm{GP}}(S, a) = \widehat{g}_\theta(S, j(a)) - \mathrm{offer}(a). \tag{4}$$

Rejection has score zero, and the rule selects a non-rejection action only when its score is strictly positive. Gain–Price is a fixed supported-action heuristic rather than an optimizer of negotiation success: for a given product, it ranks eligible offers solely by their displayed prices because the gain term is unchanged across those offers. E1 varies the gain predictor while holding the action set, offers, scoring rule, and evaluation outcomes fixed.

For comparison, historical Expected Net Benefit and risk policies use a finite operational surplus distribution $\widehat{F}$ and score

$$s_\lambda(S, a) = \mathbb{E}_{\widehat{F}}[U] - \lambda \max\{0, -\mathrm{CVaR}^{-}_{0.10}(\widehat{F})\}, \tag{5}$$

with default $\lambda = 0.50$; $\lambda = 0$ gives Expected Net Benefit. The finite-distribution construction, lower-tail convention, and incomplete later Controlled correction are detailed in Appendix B; gain–price scoring does not require that payment mixture.

Nonbinding advice. Advice uses a different decision path: a local orchestrator supplies role-specific public context and the designated forecast or reference panel, the advisor returns a recommendation, and the participant submits its own legal action. Buyer descriptive panels use Natural and Controlled summaries, while seller panels use the designated temporal path; E3 additionally gives buyer advisors explicit value information. E2 restricts selected actions to recorded support and archived outcomes, whereas online studies execute actions and let market states evolve. Reference and individualized conditions align field roles and structure, but not every numerical value, source annotation, or rendered token.

5 EXPERIMENTS

5.1 RQ1: BEHAVIORAL FOUNDATIONS

Behavioral evaluations ask whether participants respond to conditions rather than acting as constant policies. On 12 common worlds, DeepSeek spreads buyer actions across Buy A, Buy B, and Wait, while Qwen mostly buys and rarely chooses No purchase. Matched economic-update responses exceed replay and surface-change controls, although some isolated license-response intervals include zero. Persona-label permutation analyses also show observed persona statistics above the corresponding null means for buyers and sellers, under the stated within-block exchangeability assumption. Recorded DeepSeek sequences show history-dependent actions; Qwen has no sequential evaluation. Together, these results support studying conditional behavior, not human equivalence or optimality. Appendix A reports the behavioral definitions and complete controls.

5.2 RQ2: PREDICTIVE QUALITY

Predictive performance is target-dependent (Figure 2). On matched Validation, the original Natural predictor has NLL 0.5264 versus 0.4922 for the five-member DeepSets ensemble. Controlled has completion Brier 0.0611 versus 0.0649 for the full Platt-calibrated LightGBM baseline, while Technical has gain CRPS 5.1562 versus 3.2067 for quantile LightGBM and 3.5187 for linear Student-t.

The Technical ranking changes across criteria. Quantile LightGBM has sign log loss 1.8817, compared with 0.5091 for the linear model, and its nominal 90% coverage is 0.8005. Lower gain CRPS therefore does not establish uniformly better sign prediction or interval calibration.

Temporal diagnostics retain their additional input conditions: on 20 Validation blueprints, shuffled and mean-only short-state substitutions worsen the raw joint score by 0.0812 and 3.3903, whereas observed-state replacement improves it by 0.1075. These fixed-predictor substitutions condition on recorded trajectory-arrival support, and observed-state access is diagnostic rather than prospectively available. Appendix C reports the remaining baselines, reputation contrasts, temporal diagnostics, and parameterization studies.

5.3 RQ3-A: FIXED-SUPPORT DEVELOPMENT—E1 AND E2

E1 and E2 use the same 60-blueprint action-outcome library. Selected-action outcomes are summarized within blueprint and equally averaged; they are fixed-state choices, not 42-day market totals.

Historical scoring rules. The historical policy comparison illustrates the role of the action objective. Natural imitation has mean buyer surplus 0.5412, compared with 2.0550 for Expected Net Benefit and 2.3687 for Technical Gain–Price. The default risk planner reduces negative-net frequency to 0.0042 but trades in only 0.1436 of rollout rows, and its surplus contrast against imitation does not meet the historical Holm criterion ($p_{\mathrm{Holm}} = 0.0624$). Thus its low loss frequency accompanies substantially reduced trading, rather than a conditional completed-trade loss rate.

E1: replacing the gain predictor. E1 holds the gain–price rule fixed and replaces its gain summary (Figure 3). Quantile LightGBM raises mean surplus from 2.3687 to 2.6734 and deal rate from 0.5431 to 0.6483 relative to the historical Technical predictor. Its grouped conditional negative rate is 0.1611, compared with 0.1668 for the original predictor. Linear Student-t trades more often, with deal rate 0.8264, but lowers surplus to 2.2189 and raises the conditional negative rate to 0.3122.

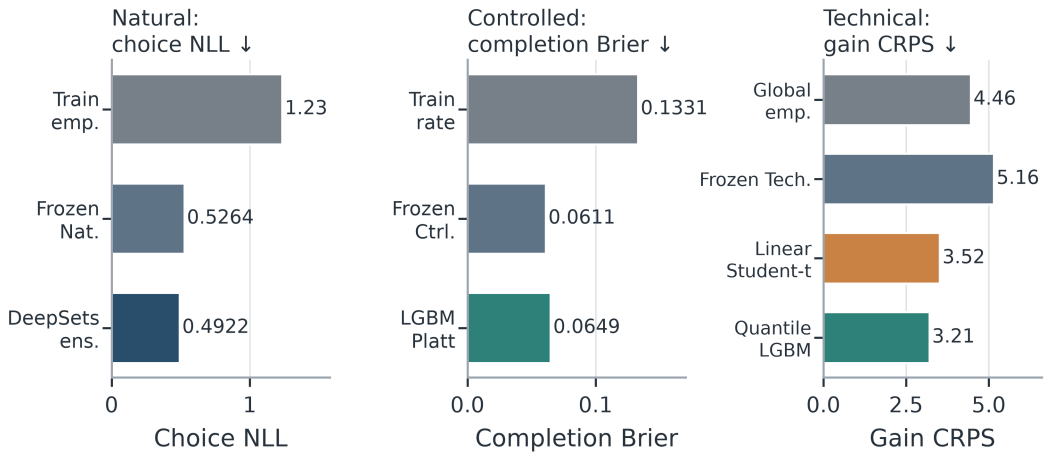

Figure 2: Buyer-side predictive comparisons on Validation. Panels report choice NLL, completion Brier score, and task-gain CRPS on separate scales; lower is better. These scores do not define an aggregate ranking of all five interfaces.

Thus lower Validation CRPS does not necessarily translate into a better outcome under the same supported-action rule.

The development rule fixed before E2 selects the highest-surplus available gain–price candidate whose grouped conditional negative rate is no more than 0.02 above the original Gain–Price value, yielding quantile LightGBM. These are development point estimates; no new paired uncertainty intervals are reported for the predictor-replacement contrasts. Appendix D gives the recorded selection rule, grouped conditional-rate definition, and unavailable comparator definitions.

E2: supplying value information to participants. E2 tests whether explicit value information helps when an advisor and participant, rather than a numerical rule, produce the final choice. It crosses reference versus individualized descriptive information with reference versus individualized value information, producing arms RR, IR, RI, and II; reference information comes from training statistics, and individualized value uses the E1-selected quantile LightGBM predictor. New choices are validated against the legal action set and joined to archived outcomes. The primary effect averages RI–RR and II–IR over equal-weight 60-blueprint surplus aggregates.

The design has 956 parent–arm cells and 2,867/2,868 complete request units. Available repetitions are averaged within parent–arm cell, then within blueprint, and finally across blueprints; the failed repetition is not imputed.

Individualized value information increases the fixed-state surplus aggregate by 1.95 accounting units (95% CI [1.53,2.39]), using a 10,000-replicate blueprint bootstrap without Holm adjustment. The estimate is not a percentage, per-call outcome, or 42-day market effect; E1 and E2 share support and are not independent post-selection confirmation.

5.4 RQ3-B: Online Descriptive Advice

The official-API format-control study evaluates 80 paired blocks and 320 continuing branches across no advisor (N_0), language-only bilateral advice (L), same-format training-reference descriptive statistics (M), and individualized descriptive forecasts (T). Neither M nor T exposes the explicit buyer-value interface, and M is informative rather than an empty panel. The primary endpoint is 42-day total buyer net technical surplus. The paired effects are $T-M = -4.60$ (95% CI $[-14.06, 4.68]$, $p_{\text{Holm}} = 0.3374$) and $T - N_0 = -12.56$ (95% CI $[-23.44, -1.68]$, $p_{\text{Holm}} = 0.0512$) within this study's two-test primary family. Individualized descriptive forecasts do not establish a buyer-surplus benefit over either reference. The $T - N_0$ point estimate is unfavorable and its unadjusted interval lies below zero, although the Holm criterion is not met at 0.05. Non-rejection is not evidence of equivalence or absence of harm.

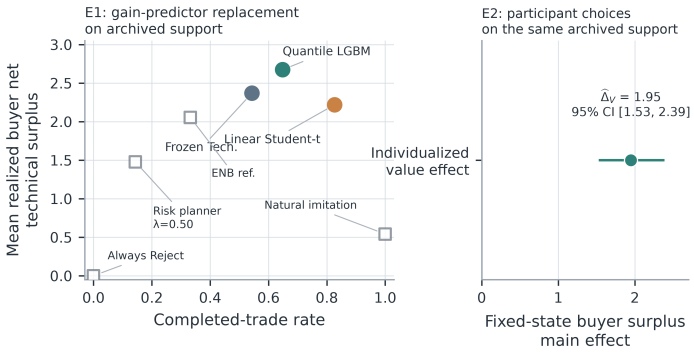

Figure 3: E1/E2 fixed-support development on the same 60-blueprint archive. Left: E1 uses the same Gain–Price rule; gray squares are historical references. Right: E2 value main effect. No E1 replacement-contrast error bars are reported.

For historical context, a separate 206-blueprint JDCloud seven-arm study improves task gain and buyer surplus relative to language-only bilateral advice, but does not establish improvement relative to no advice; because hosted-model conditions, treatments, and input pools differ, the studies are not pooled. Appendix E reports all means and comparison families.

5.5 RQ3-C: Online Value-Interface Evaluation—E3

E3 evaluates the value interface in a separate official-API continuing-market study with 72 paired blueprints, balanced between 36 LOW- and 36 HIGH-demand blueprints, and 216 complete branches. The arms are no advisor (N_0), individualized descriptive forecasts with training-reference value information (V_{REF}), and the same descriptive source with individualized value information (V_{IND}).

Information supplied to each arm. Both advised arms activate buyer and seller advisors. Buyer advisors receive descriptive forecasts and a value table; seller value rows are role-not-applicable. The advised arms share the descriptive source and field structure, although later branch states can differ. V_{REF} uses training-reference gains and deal probabilities, whereas V_{IND} uses the LightGBM gain summary and current descriptive deal probabilities. Appendix E gives the full field contract. Thus $V_{\mathrm{IND}} - V_{\mathrm{REF}}$ estimates the effect of individualizing the buyer-value package alongside the designated descriptive service, not gain prediction alone; V_{REF} is not format-control M, and V_{IND} is not descriptive T.

Market outcomes and primary contrasts. For reference $r \in \{V_{\mathrm{REF}}, N_0\}$, the primary paired estimate averages the 72 blueprint-level differences in total buyer net technical surplus over days $[56, 98)$. Table 2 reports the corresponding arm means.

Table 2: E3 online arm means. Entries are 42-day market totals averaged with equal weight over 72 paired blueprints. Buyer surplus denotes buyer net technical surplus; seller revenue is gross revenue.

Arm	Transactions	Task gain	Buyer surplus	Seller revenue
N_0	39.58	98.30	26.76	71.54
V_{REF}	24.60	76.55	32.05	44.51
V_{IND}	28.19	126.35	73.20	53.14

The paired buyer-surplus effects are $V_{\mathrm{IND}} - V_{\mathrm{REF}} = 41.16$ (95% CI $[27.68, 55.08]$) and $V_{\mathrm{IND}} - N_0 = 46.44$ (95% CI $[29.35, 64.68]$), both meeting E3's two-test Holm criterion with $p_{\mathrm{Holm}} = 2.0 \times 10^{-5}$. Figure 4 shows these effects separately from the descriptive study. Relative to V_{REF}, V_{IND} is higher on all four displayed means; relative to N_0, it has higher task gain and buyer surplus but 11.39 fewer transactions and 18.39 less seller gross revenue. These secondary means are not additional primary-family tests, and $V_{\mathrm{IND}} - N_0$ is the total bilateral-service effect, not an isolated numerical-field effect.

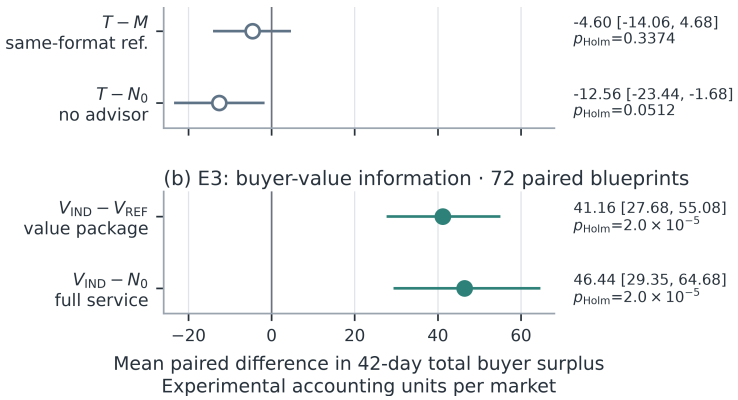

Figure 4: Primary effects on 42-day total buyer net technical surplus (experimental accounting units per market). In E3, both advised arms use individualized descriptive forecasts. Error bars are unadjusted 95% paired-bootstrap intervals; filled markers meet the study-specific Holm criterion.

6 Discussion and Limitations

Predictive rankings vary across targets and metrics, and E1, E2, and E3 use different outcome units and feedback settings. Reference arms are informative rather than empty: format-control M supplies training statistics, and E3's V_{REF} already contains individualized descriptive forecasts. The E3 contrasts therefore concern either a multi-field value package or the complete service, not a unique mechanism of advice adoption, bargaining, or low-value rejection.

The participants are LLM agents, not human traders, and Qwen lacks sequential behavioral evaluation. The task evaluator, training-derived price scale, and market rules do not establish human willingness to pay, commercial profitability, or real-market welfare. Post-purchase buyer-data updates and utility-cache keys are not fully recovered; cumulative task gain is transaction-level accounting over recorded buyer-data snapshots, not a verified sequence of marginal gains from continuously accumulated purchased data. E3 improves buyer surplus while reducing transactions and seller gross revenue relative to no advice; directions reverse relative to V_{REF}, and the accounting quantities are not additive welfare components. Entity-disjoint splits, exact-hash checks, the external LightGBM value summary, the Controlled correction, and temporal input provenance retain the qualifications detailed in the appendices; E3 is not validation of every component.

7 Conclusion

DIM-WM evaluates forecast services for LLM-agent data markets by separating three uses of prediction: estimator substitution under a fixed rule, information-source substitution at fixed states, and paired service interventions in continuing markets. This design makes prediction quality, participant-controlled information use, and online market effects auditable without treating them as the same estimand.

The empirical results are target- and condition-specific. The descriptive format-control study does not establish a buyer-surplus benefit from individualized descriptive forecasts. The strongest online finding is that, over an already individualized descriptive service, an individualized buyer-value information package increases buyer net technical surplus in E3; relative to no advice, the complete service has higher mean buyer surplus and task gain, with lower mean transactions and seller gross revenue.

AI Use Statement

LLMs are components of the experimental system and generate participant decisions and platform advice without further training. The online studies generate market continuations, while E2 generates new choices evaluated against archived action outcomes. Separately, generative AI tools were used as drafting and implementation aids, including support for code and test scaffolding, source search, translation, English editing, LaTeX and plotting-code assistance, and revisions to the con-

ceptual overview figure. Numerical findings derive from experimental records and independent task evaluations; generated illustrations are not statistical evidence. The authors remain responsible for the interpretation, references, artifacts, and final scientific review.

Ethics Statement

The computational participants do not provide evidence about the preferences or behavior of human traders. Task gains and payments use an experimental accounting scale, and the findings do not establish commercial profitability or real-market welfare. Advice can affect participation and the distribution of gains. Seller gross revenue is not profit, and task gain, buyer surplus, and seller revenue are not three additive welfare components. The information restrictions are experimental boundaries rather than formal privacy or security guarantees. Source-dataset attribution, exact local versions, and applicable redistribution permissions require version-specific verification before derived assets are released. No institutional ethics approval is claimed.

Reproducibility Statement

The appendices describe market construction, prediction inputs, model and calibration versions, supported-action evaluation, and the separate online comparison families. Recorded artifacts include manifests, configurations, prompt and protocol files, selected actions, outcomes, analysis outputs, and API usage and recovery logs. They distinguish archived-outcome reuse from new decisions and market continuations.

Appendix F specifies the different requirements for figure reconstruction, numerical reanalysis, fixed-predictor inference, retraining, and online reruns. Incomplete historical data and model mappings, unavailable per-query outputs for the Controlled correction, and restrictions on artifact release limit complete end-to-end reproduction. Credentials and personal paths are excluded from publication materials. Hosted-model aliases do not identify immutable weight snapshots, and server-side sampling is not fully controlled.

REFERENCES

Anish Agarwal, Munther Dahleh, and Tuhin Sarkar. A marketplace for data: An algorithmic solution. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pp. 701–726. Association for Computing Machinery, 2019. doi: 10.1145/3328526.3329589. URL https://doi.org/10.1145/3328526.3329589.

Gati V. Aher, Rosa I. Arriaga, and Adam Tauman Kalai. Using large language models to simulate multiple humans and replicate human subject studies. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 337–371, 2023. URL https://proceedings.mlr.press/v202/aher23a.html.

Lisa P. Argyle, Ethan C. Busby, Nancy Fulda, Joshua R. Gubler, Christopher Rytting, and David Wingate. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3):337–351, 2023. doi: 10.1017/pan.2023.2. URL https://www.cambridge.org/core/journals/political-analysis/article/out-of-one-many-using-language-models-to-simulate-human-samples/035D7C8A55B237942FB6DBAD7CAA4E49.

Santiago Andrés Azcoitia and Nikolaos Laoutaris. Try Before You Buy: A practical data purchasing algorithm for real-world data marketplaces. In *Proceedings of the 1st International Workshop on Data Economy*, pp. 11–18. Association for Computing Machinery, 2022. doi: 10.1145/3565011.3569054. URL https://doi.org/10.1145/3565011.3569054.

Gagan Bansal, Tongshuang Wu, Joyce Zhou, Raymond Fok, Besmira Nushi, Ece Kamar, Marco Tulio Ribeiro, and Daniel S. Weld. Does the whole exceed its parts? the effect of AI explanations on complementary team performance. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 2021. doi: 10.1145/3411764.3445717. URL https://dl.acm.org/doi/10.1145/3411764.3445717.

Dimitris Bertsimas and Nathan Kallus. From predictive to prescriptive analytics. *Management Science*, 66(3):1025–1044, 2020. doi: 10.1287/mnsc.2018.3253. URL https://pubsonline.informs.org/doi/10.1287/mnsc.2018.3253.

Marcel Binz and Eric Schulz. Using cognitive psychology to understand GPT-3. *Proceedings of the National Academy of Sciences*, 120(6):e2218523120, 2023. doi: 10.1073/pnas.2218523120. URL https://www.pnas.org/doi/10.1073/pnas.2218523120.

Zana Buçinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 2021. doi: 10.1145/3449287. URL https://dl.acm.org/doi/10.1145/3449287.

Laurent Candillier and Vincent Lemaire. Nomao. UCI Machine Learning Repository, 2012. URL https://archive.ics.uci.edu/dataset/227/nomao. Dataset.

Gavin C. Cawley and Nicola L. C. Talbot. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11:2079–2107, 2010. URL https://jmlr.org/papers/v11/cawley10a.html.

Lingjiao Chen, Paraschos Koutris, and Arun Kumar. Towards model-based pricing for machine learning in a data marketplace. In *Proceedings of the 2019 International Conference on Management of Data*, pp. 1535–1552. Association for Computing Machinery, 2019. doi: 10.1145/3299869.3300078. URL https://dl.acm.org/doi/10.1145/3299869.3300078.

Priya L. Donti, Brandon Amos, and J. Zico Kolter. Task-based end-to-end model learning in stochastic optimization. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/3fc2c60b5782f641f76bcefc39fb2392-Abstract.html.

Bradley Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979. doi: 10.1214/aos/1176344552. URL https://projecteuclid.org/journals/annals-of-statistics/volume-7/

issue-1/Bootstrap-Methods-Another-Look-at-the-Jackknife/10.1214/aos/1176344552.full.

Adam N. Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68(1): 9–26, 2022. doi: 10.1287/mnsc.2020.3922. URL https://pubsonline.informs.org/doi/10.1287/mnsc.2020.3922.

Yingqi Gao, Wenlu Xu, Jin J. Zhou, Hua Zhou, Yong Chen, and Xiaowu Dai. Learn then Decide: A learning approach for designing data marketplaces. *arXiv preprint arXiv:2503.10773*, 2026. doi: 10.48550/arXiv.2503.10773. URL https://arxiv.org/abs/2503.10773v2. Version 2, revised April 3, 2026; preprint.

Pierre Geurts, Damien Ernst, and Louis Wehenkel. Extremely randomized trees. *Machine Learning*, 63(1):3–42, 2006. doi: 10.1007/s10994-006-6226-1. URL https://link.springer.com/article/10.1007/s10994-006-6226-1.

Amirata Ghorbani and James Zou. Data shapley: Equitable valuation of data for machine learning. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 2242–2251, 2019. URL https://proceedings.mlr.press/v97/ghorbani19c.html.

Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007. doi: 10.1198/016214506000001437. URL https://sites.stat.washington.edu/raftery/Research/PDF/Gneiting2007jasa.pdf.

Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979. URL https://www.jstor.org/stable/4615733.

Ruoxi Jia, David Dao, Boxin Wang, Frances Ann Hubis, Nick Hynes, Nezihe Merve Gürel, Bo Li, Ce Zhang, Dawn Song, and Costas J. Spanos. Towards efficient data valuation based on the shapley value. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pp. 1167–1176, 2019. URL https://proceedings.mlr.press/v89/jia19a.html.

Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html.

Yongchan Kwon and James Zou. Beta shapley: A unified and noise-reduced data valuation framework for machine learning. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pp. 8780–8802, 2022. URL https://proceedings.mlr.press/v151/kwon22a.html.

Anson Lei, Bernhard Schölkopf, and Ingmar Posner. Variational Causal Dynamics: Discovering modular world models from interventions. *arXiv preprint arXiv:2206.11131*, 2022. doi: 10.48550/arXiv.2206.11131. URL https://arxiv.org/abs/2206.11131v1. Version 1; preprint.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=Bkg6RiCqY7.

Jayanta Mandi, James Kotary, Senne Berden, Maxime Mulamba, Victor Bucarey, Tias Guns, and Ferdinando Fioretto. Decision-focused learning: Foundations, state of the art, benchmark and future opportunities. *Journal of Artificial Intelligence Research*, 80:1623–1701, 2024. doi: 10.1613/jair.1.15320. URL https://www.jair.org/index.php/jair/article/view/15320.

OpenML. electricity: Dataset 151, version 1. OpenML dataset catalog, 2014a. URL https://www.openml.org/d/151. Versioned catalog record; year denotes catalog upload, not original data collection.

OpenML. pol: Dataset 722, version 2. OpenML dataset catalog, 2014b. URL https://www.openml.org/d/722. Versioned catalog record; binarized regression dataset; year denotes catalog upload.

Joon Sung Park, Joseph C. O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. Association for Computing Machinery, 2023. doi: 10.1145/3586183.3606763. URL https://dl.acm.org/doi/10.1145/3586183.3606763.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(85): 2825–2830, 2011. URL https://jmlr.org/papers/v12/pedregosa11a.html.

Li Peng, Jiayao Zhang, Yihang Wu, Weiran Liu, Jinfei Liu, Zheng Yan, Kui Ren, Lei Zhang, and Lin Qu. Reliable and private utility signaling for data markets. *Proceedings of the ACM on Management of Data*, 3(6), 2025. doi: 10.1145/3769821. URL https://dl.acm.org/doi/10.1145/3769821.

Juan C. Perdomo, Tijana Zrnic, Celestine Mendler-Dünner, and Moritz Hardt. Performative prediction. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 7599–7609, 2020. URL https://proceedings.mlr.press/v119/perdomo20a.html.

Kieu Thao Nguyen Pham, Rachael Hwee Ling Sim, Quoc Phong Nguyen, See Kiong Ng, and Bryan Kian Hsiang Low. DUPRE: Data utility prediction for efficient data valuation. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1557–1565, 2025. URL https://www.ifaamas.org/Proceedings/aamas2025/pdfs/p1557.pdf.

Byron Roe. MiniBooNE particle identification. UCI Machine Learning Repository, 2005. URL https://archive.ics.uci.edu/dataset/199/miniboone+particle+identification. Dataset; year follows the repository's recommended citation.

Max Schemmer, Niklas Kühl, Carina Benz, Andrea Bartos, and Gerhard Satzger. Appropriate reliance on AI advice: Conceptualization and the effect of explanations. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*, pp. 410–422. Association for Computing Machinery, 2023. doi: 10.1145/3581641.3584066. URL https://dl.acm.org/doi/10.1145/3581641.3584066.

Bryan Wilder, Bistra Dilkina, and Milind Tambe. Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 1658–1665, 2019. doi: 10.1609/aaai.v33i01.33011658. URL https://ojs.aaai.org/index.php/AAAI/article/view/3982.

Pangjing Wu, Peter Q. Chen, Xiaodong Li, Wenqi Fan, and Li Qing. Datamart-Agent: LLM-driven game-theoretic agent for data marketplace modeling. In *Findings of the Association for Computational Linguistics: ACL 2026*, pp. 32509–32531, San Diego, California, United States, July 2026. Association for Computational Linguistics. doi: 10.18653/v1/2026.findings-acl.1628. URL https://aclanthology.org/2026.findings-acl.1628/.

Xinlei Xu, Awni Hannun, and Laurens Van Der Maaten. Data appraisal without data sharing. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pp. 11422–11437, 2022. URL https://proceedings.mlr.press/v151/xu22e.html.

Jinsung Yoon, Sercan O. Arik, and Tomas Pfister. Data valuation using reinforcement learning. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 10842–10851, 2020. URL https://proceedings.mlr.press/v119/yoon20a.html.

Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Ruslan Salakhutdinov, and Alexander J. Smola. Deep sets. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper/2017/hash/f22e4747da1aa27e363d86d40ff442fe-Abstract.html.

Boxuan Zhang, Weipu Zhang, Zhaohan Feng, Wei Xiao, Jian Sun, Jie Chen, and Gang Wang. Mixture-of-World Models: Scaling multi-task reinforcement learning with modular latent dynamics. *arXiv preprint arXiv:2602.01270*, 2026. doi: 10.48550/arXiv.2602.01270. URL https://arxiv.org/abs/2602.01270v1. Version 1; preprint.

Hongxin Zhang, Zeyuan Wang, Qiushi Lyu, Zheyuan Zhang, Sunli Chen, Tianmin Shu, Behzad Dariush, Kwonjoon Lee, Yilun Du, and Chuang Gan. COMBO: Compositional world models for embodied multi-agent cooperation. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=YXRyYkb1im.

A Market Construction and Behavioral Validation

This appendix specifies the continuing market, the independently measured task-utility labels, and the controlled behavioral evaluations. The continuing market supplies trajectories for prediction and service evaluation; the behavioral scenarios characterize responses under specified persona, economic, and history conditions. Their participant counts, legal actions, clocks, and replication units are not interchangeable. The later persona-null analysis reuses recorded behavioral actions rather than adding new LLM decisions.

A.1 Participants, Products, and Event Dynamics

Population and product entry. The continuing market runs for 168 simulated days across four classification lineages: electricity, MiniBooNE, pol, and nomao. Each lineage contains four buyers and two sellers, giving 16 buyers and eight sellers overall. Buyers begin with 250, 500, 750, or 1,000 existing observations and differ in bargaining style and review propensity. Sellers differ in pricing style. Each seller introduces two products, one on day 0 and one on day 56. The second entry therefore adds products, not new seller identities. The LLMs are used without further training, while their decisions can respond to observations and accumulated interaction history.

Demand and legal decisions. Under low demand, each buyer receives exogenous opportunities at a rate of 0.75 per 14 days. High demand adds an independent process at the same rate, giving a total rate of 1.50 opportunities per 14 days. These rates describe decision opportunities, not completed purchases. Product entry changes the available choice set, while an arrival creates an opportunity to act; neither event determines the participant's choice.

Participants select products, quote, make offers or counteroffers, accept, and reject when those actions are legal. Quotes are reconsidered every 14 days. Following completion, the simulator evaluates the task effect, and a review scheduled 14 days later can update public reputation. Task evaluation and reputation updating are simulator events rather than freely selectable participant actions. Rejecting an opportunity in the continuing market preserves the possibility of acting at later arrivals. It is not identified with the explicit `WAIT` action used in the controlled behavioral experiment.

Prediction and service clocks. The 14-, 28-, and 42-day prediction horizons are measured relative to a checkpoint. The online advice studies evaluate continuations over days $[56, 98)$ within the 168-day market. These horizons are distinct from the round indices of the sequential behavioral tests and from predictor-training iterations. Study-specific branch construction and online comparison arms are described in Appendix E.

Platform-visible information. The public projection $S_t = \phi(X_t)$ includes demand information, product declarations and verified cards, asks, licenses, reputation, and visible transactions and statistics available by time t. Each predictor uses its specified subset and transformations of these fields; the projection is not a claim that every model receives every public field.

Full private buyer data, hidden participant types or undisclosed preferences, realized future events, test aggregates, and empirically best actions are excluded from predictor features. Persistent identifiers support record linkage, splitting, and auditing, rather than learned entity features. Persona conditions controlled by the behavioral researcher are likewise not automatically public conditioning variables for DIM-WM. An evaluator can access a private buyer dataset to produce a task label without exposing that dataset or its future measured outcome to a predictor.

No Markov sufficiency is assumed for S_t. Sampling from a known exogenous arrival mechanism is distinguished from reading the actual future realization. The archived temporal evaluations additionally condition on trajectory arrival capacities; those conditions are kept separate from the contemporaneous public state in Appendix B.

Collection semantics and grouping. Natural-choice records retain the visible observation, complete candidate set, and chosen option, including rejection. Controlled records restore a pre-action parent snapshot, execute a specified legal action, and record the resulting completion, payment, applicable feedback, and branch metadata. The specified action is not thereby converted into a

natural-preference label. Independent task evaluations associate a buyer–product snapshot with its measured task effect, while seller windows associate checkpoints with time-indexed outcomes.

Candidate-set, parent, buyer–product, and market identities are retained when records are joined. These routes separate supervision meanings, not statistical dependence: they may share products, states, and trajectories. Repeated snapshots of a buyer–product pair do not create independent utility labels. Behavioral evaluation records and fixed-support decision outcomes are not automatically additional predictor-training examples.

A.2 Data Pools, Product Variation, and Task Utility

Distinct data roles. For each lineage, the construction specifies a 3,000-row independent utility-evaluation pool, a 400-row buyer-visible validation pool, and separate 1,400-row buyer/seller pools for Train, Calibration, Validation, and Test. The buyer-visible validation summary may enter a public buyer description. It is distinct from the predictor-development split named Validation and from the independent pool used to measure task utility. Dataset provenance and remaining local-version bindings are documented in Appendix F.

Product recipes and public descriptions. Products contain 250, 500, or 1,000 sampled observations. For each size, the construction includes a clean recipe and four quality families—label noise, feature noise, random missingness, and duplication—at severities of 5%, 10%, 15%, and 20%. This gives

$$3\,(1 + 4 \times 4) = 51 \tag{6}$$

recipes per lineage. These are product-construction templates, not 51 independent datasets or markets. The 750-row buyer tier does not introduce an additional product-size tier.

Seller declarations and verified cards are separate public objects. Declarations may differ from delivered quality, whereas cards expose protocol-established fields. Later consistency feedback may concern the agreement between declaration and delivery. Declaration language neither certifies product quality nor assigns task utility.

Independent task evaluator. For a recorded buyer dataset and its concatenation with a candidate product, the evaluator fits the same logistic-regression pipeline twice. Numeric features use median imputation and standardization. Categorical features use most-frequent imputation and one-hot encoding with unknown-category handling. The classifier uses L2 regularization, $C = 1$, and the `lbfgs` solver. The procedure and hyperparameters are fixed, but the fitted coefficients and preprocessing statistics can differ between the baseline and augmented fits.

Both fitted pipelines predict the same independent utility pool. For binary evaluation labels y_i and fitted probabilities $\pi_i^{(v)}$, where $v \in \{\mathrm{base}, \mathrm{aug}\}$, define

$$L_v = -\frac{1}{n_{\mathrm{eval}}} \sum_{i=1}^{n_{\mathrm{eval}}} \left[y_i \log \pi_i^{(v)} + (1 - y_i) \log \left(1 - \pi_i^{(v)} \right) \right]. \tag{7}$$

The implementation clips probabilities to $[10^{-7}, 1 - 10^{-7}]$ for this loss calculation and checks that the losses are finite and the baseline loss is positive. A convergence warning triggers a deterministic retry with the iteration limit increased from 2,000 to 10,000. Persistent failure is recorded as an evaluation failure, not a task gain of zero. These numerical conventions belong to the evaluator; they do not truncate the Technical predictor's Student-t law.

Task gain and realized buyer surplus follow Eq. (1). The task gain can be negative and depends on the buyer's existing data. Independent pair evaluations may supply labels even when the product was not naturally selected, but those labels are not available as realized future gains for prospective decisions. Non-transactions contribute zero realized gain, payment, and surplus, without assigning zero potential gain to the unpurchased product.

Post-purchase data state and utility caching. The available market specification establishes the buyer-data snapshot used for each recorded buyer–product evaluation. It does not establish whether a completed purchase mutates the buyer's private training data for later opportunities, whether repeated purchases of the same buyer–product pair are excluded by state, or whether utility caches are

keyed by buyer identity and product or by buyer-state snapshot and product. Reported cumulative task gains are therefore transaction-level accounting totals over recorded evaluation snapshots, not verified marginal gains along a fully specified evolving private dataset.

The accounting scale is neither currency nor LLM billing tokens. Seller gross revenue is not profit because comparable seller costs are not modeled. Task gain, buyer surplus, and seller revenue are distinct reported outcomes, not three additive components of social welfare. Measured classification effects also do not validate real demand curves, human willingness to pay, or commercial profitability.

Training-derived launch prices. A lineage-specific launch bank establishes the initial price scale. The bank first selects 1,000 unique training buyer–product pairs per lineage and then retains their strictly positive measured gains. It is therefore not a sample of 1,000 positive pairs. The construction requires at least 200 positive pairs and a viable upper quantile.

For lineage ℓ, let $q_{\ell,.05}$, $q_{\ell,.50}$, $q_{\ell,.95}$, and $q_{\ell,.99}$ be linearly interpolated empirical quantiles of the retained positive gains. In experimental token units, define

$$\epsilon_\ell = \max\{0.01,\, 0.01 q_{\ell,.50}\}, \qquad p_{\max,\ell} = \max\{1,\, 10 q_{\ell,.99}\}. \tag{8}$$

The legal cent-grid maximum is the rounded representation of $p_{\max,\ell}$.

Day-0 products receive approximately 90% core, 5% low-tail, and 5% high-tail assignments. The implementation allocates integer quotas within split, lineage, and seller slot, with semantically keyed randomization assigning products to those quotas. It does not independently sample a three-category assignment for each listing.

Core prices are sampled from the empirical positive-gain bank restricted to $[q_{\ell,.05}, q_{\ell,.95}]$. This is an empirical draw, not a uniform draw over the numerical interval. The guarded tail intervals are

$$\begin{aligned} I_{\text{low},\ell} &= [\epsilon_\ell,\, \max\{\epsilon_\ell, 0.75 q_{\ell,.05}\}]\,, \\ b_\ell &= \min\{p_{\max,\ell}, 1.25 q_{\ell,.95}\}, \\ I_{\text{high},\ell} &= [b_\ell,\, \max\{b_\ell,\, \min\{p_{\max,\ell}, 4 q_{\ell,.95}\}\}]\,. \end{aligned} \tag{9}$$

Tail prices are log-uniform on these intervals and are then converted to the legal cent grid within the inherited price bound. A degenerate interval contributes its single endpoint.

Only day-0 products receive prices from this routine. A seller provides the initial quote for its day-56 product within the same lineage bound. Later quote revisions, negotiation, and acceptance determine payments. The launch bank neither uses Validation/Test utility outcomes nor assigns a product the gain of a particular future transaction. Nevertheless, price–value comparisons are conditional on this informative initialization, rather than demonstrating price discovery without a prior utility scale.

A.3 Behavioral Protocols and Controls

Evaluation scope. The behavioral experiments examine action non-degeneracy, natural-language persona conditioning, responses to economic changes, and adaptation to realized history. They do not optimize or rank participants by return. The static protocol uses eight economic persona conditions per role. These are controlled prompt conditions, not the 16 buyer identities in the continuing market.

The archived primary configuration is DeepSeek-V4-Flash, with 24 static worlds per role and three sampling repeats identified by seeds 11, 29, and 47. The secondary configuration is Qwen-3.6-23B, with 12 static worlds per role. Cross-model common-support summaries use the 12 shared worlds; the additional DeepSeek worlds remain a separate full-support evaluation. Static experiments use temperature 0.1. The archived continuing-market configuration records temperature 0.70 and top-p 0.95, so behavioral and continuing-market evaluations are not collapsed into one unspecified sampling configuration. These historical settings are not substituted for the separately recorded official-API supplement settings.

For DeepSeek, matched-change support contains 24 update, 24 license, and 16 combined-change scenarios per role. Qwen covers 12, 12, and eight, respectively. The sequential design selects four matched sets per change type and role, uses four balanced persona conditions, and records four

Table 3: Recorded behavioral decisions by design. Counts include repeated observations and both evaluated roles; they are not independent-world counts. Qwen has no sequential evaluation. The design-specific persona and scenario descriptions are retained separately rather than inferred from these totals.

Design	DeepSeek	Qwen
Static heterogeneity (`HET`)	1,152	576
Matched economic changes (`ADAPT`)	6,144	3,072
Sequential behavior (`DYNAMIC`)	6,912	Not evaluated
Total completed LLM decisions	14,208	3,648

buyer rounds or eight seller rounds. A separate seller-market control has a 40-round setting. Qwen was not evaluated in the sequential design.

The matched-change documentation has an unresolved persona-count discrepancy: its narrative specifies eight conditions, while the `ADAPT` inventory records four profiles. Scenario counts and completed-decision totals are therefore reported directly rather than reconstructed from an assumed common profile count.

Execution and controls. The behavioral execution report records 17,856 completed LLM decisions: 14,208 from DeepSeek and 3,648 from Qwen. It reports no fallback decisions or format-contract failures. Four failed API attempts were retried within the same frozen calling plan; they are not additional behavioral replicates. Table 3 preserves the design-level decision counts.

Generic and type-specific rule and reinforcement-learning (RL) controls use the same frozen evaluation inputs. The archived RL configuration uses a separate 48-world development distribution per role and five seeds: 11, 29, 47, 71, and 89. It records 100,000 buyer static steps, 100,000 buyer dynamic steps, and 150,000 seller sequential steps; the buyer network has two hidden layers of width 128. These describe the evaluated implementations, not all possible rule-based or RL participants. Their results are descriptive controls, not pass/fail thresholds for LLM eligibility.

Action distributions and the archived persona statistic. For one fixed role and agent, let a_{wpr} denote the action in world w, persona condition p, and repeat r. Using the observed repeat count R_{wp}, define

$$\widehat{\pi}_{wp}(a) = \frac{1}{R_{wp}} \sum_{r=1}^{R_{wp}} \mathbf{1}\{a_{wpr} = a\}, \qquad d_{\mathrm{TV}}(\pi, \pi') = \frac{1}{2} \sum_{a \in \mathcal{A}} |\pi(a) - \pi'(a)|. \tag{10}$$

The categorical support is the union of the observed legal action labels for the comparison.

With P_w persona conditions in world w, the archived statistic is

$$\begin{aligned} B_w &= \binom{P_w}{2}^{-1} \sum_{p<p'} d_{\mathrm{TV}}\left(\widehat{\pi}_{wp}, \widehat{\pi}_{wp'}\right), \\ N_w &= \frac{1}{P_w} \sum_{p=1}^{P_w} \left[1 - \max_a \widehat{\pi}_{wp}(a)\right], \\ H_w &= B_w - N_w, \qquad \overline{H} = \frac{1}{W} \sum_{w=1}^{W} H_w. \end{aligned} \tag{11}$$

Thus N_w is within-persona modal disagreement, not mean pairwise TVD between repeated actions. World-level intervals describe the sampling variation of the reported statistic. Because H_w is not generally zero-centered under identical persona action distributions, an interval excluding zero is not by itself a valid no-persona-effect test. Section A.4 reports the additional null analysis without changing the archived definition.

Matched economic changes. Update and license experiments retain different reference cells. For update, B denotes baseline, U the substantive update, and S, D the two replay or surface-change

Table 4: Static behavioral summaries on the 12 common worlds. Panel (a) reports buyer action percentages without newly added confidence intervals. Panel (b) retains the archived seller *price fraction* and *update adoption* summaries and their 95% intervals. These measures have different meanings and are not combined into a behavioral quality score.

(a) Buyer action composition (%)

Agent	Buy A	Buy B	Wait	No purchase
DeepSeek	43.75	26.74	29.51	0.00
Qwen	63.19	34.38	2.08	0.35
RL generic	0.00	0.00	58.33	41.67
RL typed	0.00	0.00	58.33	41.67
Rule generic	33.33	66.67	0.00	0.00
Rule typed	2.08	0.00	1.04	96.88

(b) Seller summaries

Agent	Price fraction	95% CI	Update adoption	95% CI
DeepSeek	0.3307	$[0.2725, 0.3958]$	0.0903	$[0.0000, 0.2118]$
Qwen	0.3550	$[0.2925, 0.4184]$	0.4514	$[0.2604, 0.6597]$
RL generic	0.2708	$[0.2500, 0.3042]$	0.0000	$[0.0000, 0.0000]$
RL typed	0.2698	$[0.2500, 0.3094]$	0.0000	$[0.0000, 0.0000]$
Rule generic	0.3750	$[0.3125, 0.4375]$	0.0000	$[0.0000, 0.0000]$
Rule typed	0.4375	$[0.3333, 0.5521]$	0.8750	$[0.7812, 0.9583]$

references. For license, L denotes the substantive license change and S, L_0 its reference cells. Within a matched scenario, the reported movements are

$$\begin{aligned} H_{\text{update}} &= d_{\text{TV}}(\widehat{\pi}_B, \widehat{\pi}_U) - \frac{1}{2}\left[d_{\text{TV}}(\widehat{\pi}_B, \widehat{\pi}_S) + d_{\text{TV}}(\widehat{\pi}_B, \widehat{\pi}_D)\right], \\ H_{\text{license}} &= d_{\text{TV}}(\widehat{\pi}_B, \widehat{\pi}_L) - \frac{1}{2}\left[d_{\text{TV}}(\widehat{\pi}_B, \widehat{\pi}_S) + d_{\text{TV}}(\widehat{\pi}_B, \widehat{\pi}_{L_0})\right]. \end{aligned} \tag{12}$$

These cell labels are local to the behavioral design; D here does not denote the transaction-completion variable.

The combined-change design contains B, U, L, UL and reports

$$H_{\text{combined}} = d_{\text{TV}}(\widehat{\pi}_B, \widehat{\pi}_{UL}). \tag{13}$$

This is an unsubtracted distance, not a sham-adjusted response or a factorial interaction. Recorded action repeats are pooled within the corresponding cells and then summarized over matched scenarios. Persona-specific summaries remain distinct from those pooled summaries.

Sequential feedback. Sequential statistics describe actions conditional on realized history, including purchasing after waiting and next-quote changes following zero or positive demand. Feedback is not independently randomized while the entire preceding trajectory is held fixed. The results therefore describe history-dependent behavior under frozen parameters, not a causal effect of feedback, optimal adaptation, or parameter learning. Repeated rounds and eligible transitions remain observations within their matched trajectories.

A.4 Behavioral Results and Persona Null Sensitivity

Action composition and seller summaries. Table 4 reports buyer action composition and the archived seller summaries on the 12 common worlds. Both LLM buyers use more than one action. Several controls have concentrated action distributions, but this concerns the specified implementations and inputs rather than an impossibility result for their algorithmic families. Buyer `WAIT` and `no purchase` are retained as separate labels.

Seller price fraction and update adoption also retain their original measurement identities. They are not transaction surplus or evidence that a particular price is optimal. Figure A1 summarizes common-support action composition and matched economic responses; all displayed values come from the corresponding recorded summaries.

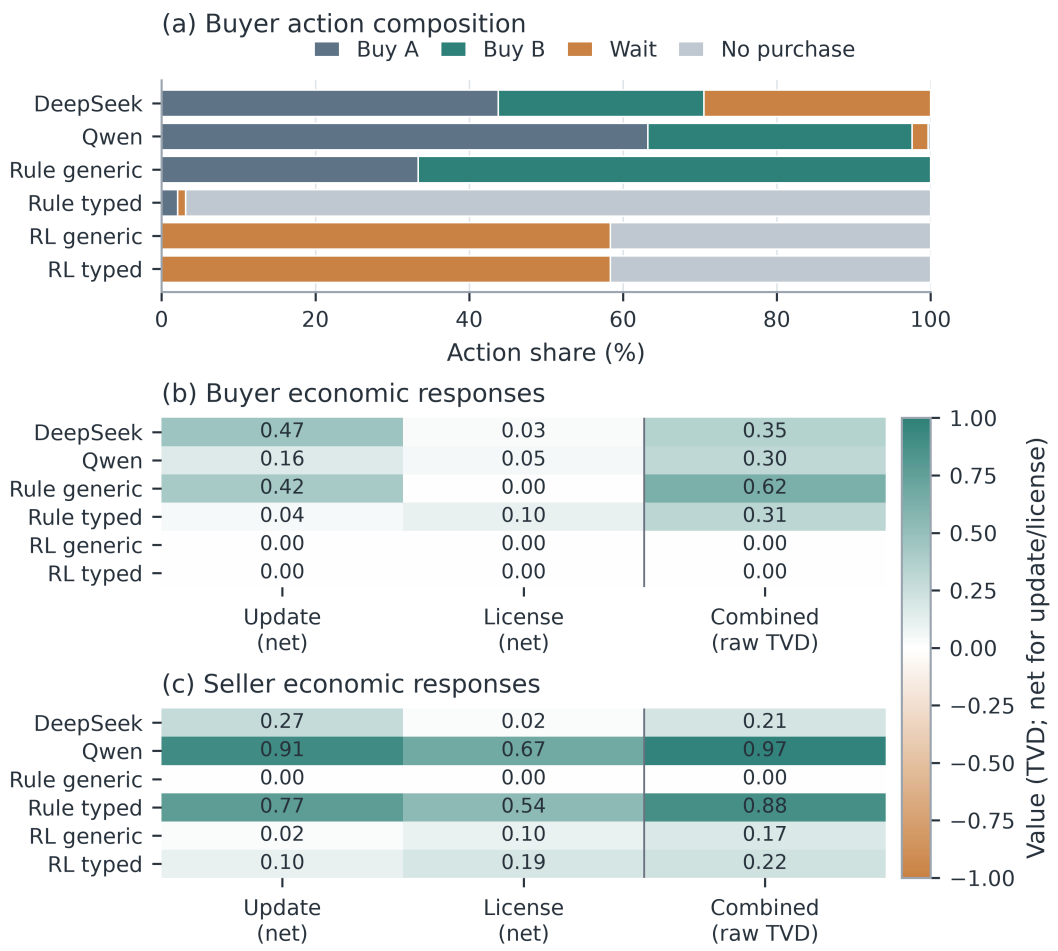

Figure A1: Behavioral evidence on common support. Buyer action composition retains all evaluated LLM, rule, and RL controls. Model rows are ordered consistently across panels. Matched economic-response panels separately display update and license responses after subtraction of their specified references, and raw baseline-to-combined-change TVD. The latter is not a sham-adjusted response or an interaction effect. Heatmap color encodes the reported response statistic only, not significance or agent quality. No new uncertainty bands are inferred for action proportions.

The seller measure name *price fraction* is retained from the archived output. Its normalization is not rebound here to the original implementation, so it is not reinterpreted as payment-to-ask, seller profit, or a new economic ratio.

Archived persona summaries. Table 5 preserves the original $\overline{H}$ estimates and world-level intervals on common and full support. Common support contains 12 worlds for each listed agent. Full support contains 24 worlds except for Qwen, whose evaluation remains at 12. The two archived summaries remain separate even where point estimates coincide. Their interval endpoints are not replaced by a pooled estimate.

The seller Rule-typed control, for example, has a substantial positive archived persona statistic. This prevents interpreting a positive H as a property unique to LLMs. Conversely, negative or zero control values describe the recorded statistic on the specified conditions; they do not establish a universal failure of rules or RL.

Why zero is not the no-effect reference. Consider a toy setting in which every persona independently produces one of two actions with probability $1/2$, with three repeats per persona. For two persona conditions, let $K, L \overset{\text{iid}}{\sim} \text{Binomial}(3, 1/2)$ count the first action. Their empirical TVD is

Table 5: Archived persona statistic $\overline{H}$ on common and full support. Common support uses 12 worlds throughout; full support uses 24 worlds except for Qwen, which uses 12. Intervals are the archived 95% world-level intervals for $\overline{H}$, not null-distribution ranges or tests against a correctly centered no-persona-effect reference. Small differences between archived Qwen interval endpoints are retained rather than reconciled by a new calculation.

Role	Agent	Common H	95% CI	Full H	95% CI
Buyer	DeepSeek	0.3919	$[0.3001, 0.4692]$	0.3413	$[0.2599, 0.4177]$
Buyer	Qwen	0.3105	$[0.1746, 0.4435]$	0.3105	$[0.1736, 0.4479]$
Buyer	RL generic	0.0000	$[0.0000, 0.0000]$	0.0000	$[0.0000, 0.0000]$
Buyer	RL typed	0.0000	$[0.0000, 0.0000]$	0.0000	$[0.0000, 0.0000]$
Buyer	Rule generic	0.0000	$[0.0000, 0.0000]$	0.0000	$[0.0000, 0.0000]$
Buyer	Rule typed	0.0625	$[0.0000, 0.1250]$	0.0952	$[0.0357, 0.1682]$
Seller	DeepSeek	0.0635	$[0.0064, 0.1528]$	0.1012	$[0.0432, 0.1684]$
Seller	Qwen	0.2297	$[0.1032, 0.3740]$	0.2297	$[0.1031, 0.3760]$
Seller	RL generic	-0.0333	$[-0.0833, 0.0000]$	-0.0750	$[-0.1333, -0.0250]$
Seller	RL typed	0.0190	$[0.0000, 0.0571]$	0.0176	$[0.0021, 0.0396]$
Seller	Rule generic	0.0000	$[0.0000, 0.0000]$	0.0000	$[0.0000, 0.0000]$
Seller	Rule typed	0.4405	$[0.2798, 0.5833]$	0.3869	$[0.2812, 0.4896]$

$|K - L|/3$. The within-persona modal disagreement is $\min(K, 3 - K)/3$. Consequently,

$$\begin{aligned} \mathbb{E}[B_w] &= \mathbb{E}\left[\frac{|K - L|}{3}\right] = \frac{5}{16}, \\ \mathbb{E}[N_w] &= \mathbb{E}\left[\frac{\min(K, 3 - K)}{3}\right] = \frac{1}{4}, \\ \mathbb{E}[H_w] &= \frac{1}{16} > 0. \end{aligned} \tag{14}$$

The same expectations apply to equal averages over persona pairs and persona conditions in this toy design. Thus a positive expected H_w is possible even when the underlying action distribution does not depend on persona. The value $1/16$ is a counterexample to zero-centering, not a universal correction to subtract from the empirical studies.

Blockwise permutation sensitivity. The later analysis reuses the archived static actions and performs 10,000 profile-label permutations. Labels are permuted within role, model, world, and repetition blocks, after which persona-level empirical distributions and the statistic in Eq. (11) are recomputed. This changes the grouping of repeated actions into persona conditions; merely renaming complete persona groups would leave the pairwise statistic unchanged.

Let $\overline{H}^{(m)}$ denote the world-averaged statistic for permutation m, with $M = 10{,}000$. The reported null mean and excess are

$$\widehat{\mu}_0 = \frac{1}{M} \sum_{m=1}^{M} \overline{H}^{(m)}, \qquad \widehat{H}_{\text{excess}} = \overline{H}_{\text{obs}} - \widehat{\mu}_0. \tag{15}$$

The displayed null range consists of the 2.5th and 97.5th percentiles of the permuted statistics. It is not a confidence interval for the true persona effect.

Table 6 and Figure A2 report the available LLM results. Observed statistics exceed their permutation references for both roles and both models. The reported right-tail probabilities are approximately 10^{-4}; their equal displayed values do not imply equal effect sizes or a mathematically zero probability. The primary-model analysis uses 24 worlds per role and the secondary-model analysis uses 12, rather than a newly constructed common-support comparison.

Interpretation depends on exchangeability of profile labels within the specified blocks. We therefore present this result as a sensitivity analysis, not a confirmatory test that is valid without assumptions about the design. It does not measure persona realism or identify latent psychological mechanisms. No additional permutation results are inferred for the rule or RL controls.

Table 6: Later blockwise persona-label permutation sensitivity analysis. Primary and secondary records correspond to the archived DeepSeek and Qwen configurations, respectively. The null range is the central 95% range of permuted world-averaged statistics, not an effect confidence interval. Probabilities are the reported right-tail values.

Role	Model	Worlds	Observed H	Null mean	Excess	Null 95% range	p
Buyer	DeepSeek	24	0.341270	0.074954	0.266316	$[0.030258, 0.119792]$	0.000100
Buyer	Qwen	12	0.310516	0.062232	0.248283	$[0.002976, 0.119048]$	0.000100
Seller	DeepSeek	24	0.101190	0.033800	0.067390	$[0.005456, 0.061756]$	0.000100
Seller	Qwen	12	0.229663	0.041645	0.188018	$[-0.009921, 0.093254]$	0.000100

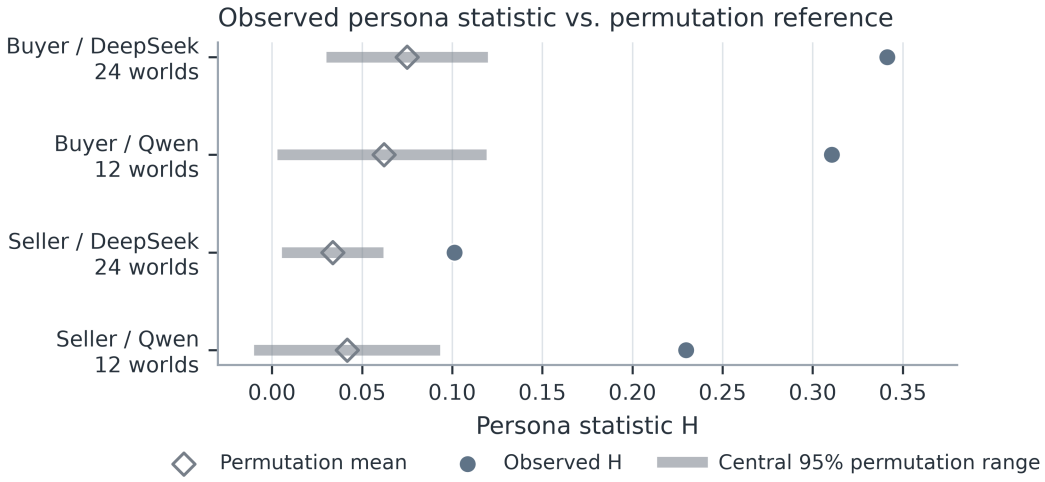

Figure A2: Persona statistics relative to their blockwise permutation references. For each role and model, show the observed world-averaged H, the permutation-null mean, and the central 95% null range. DeepSeek uses 24 worlds per role and Qwen uses 12. The null ranges are not confidence intervals for an effect. Positive excess is interpreted under within-block profile-label exchangeability and does not establish human-like personas.

Matched economic responses. Tables 7 and 8 retain common- and full-support results, including the specific rule and RL controls. On common support, DeepSeek's update-response intervals are positive for both roles, whereas its isolated license-response intervals include zero. Qwen sellers show large changes under update, license, and combined conditions. The larger DeepSeek full-support results differ from the common-support results: in particular, its full-support license intervals exclude zero. The common-support statement is therefore not generalized to all archived scenarios.

The update and license columns use the subtracted quantities in Eq. (12); the combined column uses the raw distance in Eq. (13). Their adjacent presentation does not make them the same estimand or turn the combined response into evidence of a synergistic interaction. All intervals below retain their archived source identities.

Sequential behavior under frozen parameters. Table 9 retains the archived transition and quote summaries. DeepSeek buyers purchase after waiting with estimated probability 0.2478, while seller next-quote changes average -2.2708 after zero demand and 0.3163 after positive demand. These observations show history-dependent actions in the recorded scenarios, without identifying a causal feedback effect or optimal adaptation.

Eligible matched-set counts differ across conditional outcomes and controls. The price-change rows are in experimental price units, whereas switch and wait-to-buy rows are proportions. Qwen has no sequential evaluation, and absent rows are not interpreted as zero responses.

Taken together, the action, economic-response, and sequential summaries support studying conditional behavior in these simulated participants. The persona-null sensitivity strengthens the interpretation of the recorded persona contrasts under its stated exchangeability assumption, without

Table 7: Matched economic responses on common support. Each entry is an estimate followed by its archived 95% interval. Update and license each use 12 matched scenarios per role and agent; combined change uses eight. Update/license are reference-subtracted responses, while combined change is raw baseline-to-combination TVD.

Role	Agent	Update	License	Combined
Buyer	DeepSeek	0.4653 [0.3681, 0.5590]	0.0278 [−0.0104, 0.0694]	0.3542 [0.1771, 0.5312]
Buyer	Qwen	0.1597 [0.0625, 0.2674]	0.0521 [0.0139, 0.0972]	0.3021 [0.0938, 0.5521]
Buyer	RL generic	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]
Buyer	RL typed	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]
Buyer	Rule generic	0.4167 [0.1667, 0.6667]	0.0000 [0.0000, 0.0000]	0.6250 [0.2500, 0.8750]
Buyer	Rule typed	0.0417 [0.0000, 0.1250]	0.1042 [0.0000, 0.2292]	0.3125 [0.1250, 0.5000]
Seller	DeepSeek	0.2743 [0.0729, 0.4965]	0.0208 [−0.0174, 0.0764]	0.2083 [0.0521, 0.3750]
Seller	Qwen	0.9097 [0.8194, 0.9861]	0.6736 [0.4583, 0.8611]	0.9688 [0.9375, 0.9896]
Seller	RL generic	0.0167 [0.0000, 0.0500]	0.1000 [0.0500, 0.1500]	0.1750 [0.1000, 0.2500]
Seller	RL typed	0.1042 [0.0749, 0.1375]	0.1875 [0.1083, 0.2708]	0.2188 [0.0875, 0.3563]
Seller	Rule generic	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]
Seller	Rule typed	0.7708 [0.5833, 0.9375]	0.5417 [0.3333, 0.7500]	0.8750 [0.7812, 0.9688]

Table 8: Matched economic responses on full support, with archived 95% intervals. DeepSeek and the controls use 24 update scenarios, 24 license scenarios, and 16 combined-change scenarios per role. Qwen retains 12, 12, and eight. This table is not pooled with the common-support table; its source-specific interval endpoints are preserved.

Role	Agent	Update	License	Combined
Buyer	DeepSeek	0.4219 [0.3194, 0.5226]	0.0469 [0.0069, 0.0920]	0.3854 [0.2708, 0.4948]
Buyer	Qwen	0.1597 [0.0590, 0.2674]	0.0521 [0.0139, 0.0972]	0.3021 [0.0938, 0.5521]
Buyer	RL generic	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]
Buyer	RL typed	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]
Buyer	Rule generic	0.2917 [0.1250, 0.5000]	0.0000 [0.0000, 0.0000]	0.4375 [0.1875, 0.6875]
Buyer	Rule typed	0.0729 [0.0208, 0.1354]	0.1042 [0.0417, 0.1667]	0.2812 [0.1562, 0.4062]
Seller	DeepSeek	0.1771 [0.0503, 0.3142]	0.0885 [0.0278, 0.1649]	0.1875 [0.0990, 0.2865]
Seller	Qwen	0.9097 [0.8194, 0.9861]	0.6736 [0.4653, 0.8611]	0.9688 [0.9375, 0.9896]
Seller	RL generic	0.0500 [0.0167, 0.0917]	0.0667 [0.0333, 0.1083]	0.1000 [0.0375, 0.1625]
Seller	RL typed	0.1187 [0.0917, 0.1458]	0.1729 [0.1250, 0.2229]	0.2375 [0.1562, 0.3187]
Seller	Rule generic	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]	0.0000 [0.0000, 0.0000]
Seller	Rule typed	0.7812 [0.6458, 0.8958]	0.5625 [0.4167, 0.7083]	0.8750 [0.8125, 0.9375]

replacing the original statistic or its archived intervals. None of these behavioral results establishes predictive accuracy, beneficial advice, or a welfare optimum. Those questions are evaluated separately in Appendices C, D, and E.

B Predictor Specifications and Mathematical Conventions

This appendix specifies the conditioning variables, predictive supports, training objectives, sampling procedures, and calibration conventions of the evaluated models. Natural uses Technical-derived features, and Dynamic42 conditions on sampled short-period outcomes. These dependencies do not define a single end-to-end likelihood or an identified factorization of the full market distribution. The historical probabilistic models, the later Technical baselines, and their uses in information interfaces retain separate identities.

B.1 Inputs, Supervision, and Data Partitions

Module-specific conditioning. Natural conditions on public buyer data-size and validation summaries, demand regime, observed market history, and a complete candidate set. Each candidate contains verified structural attributes, declaration features, ask, reputation, and a visibility mask. Controlled additionally receives negotiation stage, legal action, offer, public message, and current negotiation history. Technical receives buyer and product summaries, reputation, and declarations, with price and action structurally excluded.

Table 9: Sequential behavioral summaries with archived 95% intervals. "Sets" counts the eligible matched scenarios for each statistic, not individual action transitions. Buyer/seller switch and wait-to-buy outcomes are proportions; quote changes use experimental price units. Qwen was not evaluated. Rows not reported in the archive are not imputed.

Agent	Sets	Estimate	95% CI
Buyer: action switch			
DeepSeek	12	0.6447	$[0.5642, 0.7344]$
RL typed	12	0.2000	$[0.2000, 0.2000]$
Rule generic	12	0.6944	$[0.6667, 0.7361]$
Rule typed	12	0.0833	$[0.0312, 0.1424]$
Buyer: wait to buy			
DeepSeek	12	0.2478	$[0.1001, 0.4264]$
RL typed	12	0.0000	$[0.0000, 0.0000]$
Rule generic	12	0.0556	$[0.0000, 0.1389]$
Seller: action switch			
DeepSeek	12	0.0288	$[0.0099, 0.0516]$
RL typed	12	0.0192	$[0.0123, 0.0262]$
Rule generic	12	0.0000	$[0.0000, 0.0000]$
Rule typed	12	0.0253	$[0.0119, 0.0387]$
Seller: next-quote change after zero demand			
DeepSeek	9	-2.2708	$[-4.0685, -0.5331]$
RL typed	10	-10.8230	$[-23.9309, -2.6722]$
Rule generic	9	0.0000	$[0.0000, 0.0000]$
Rule typed	9	0.0000	$[0.0000, 0.0000]$
Seller: next-quote change after positive demand			
DeepSeek	12	0.3163	$[0.0296, 0.7246]$
RL typed	11	-0.0521	$[-0.4467, 0.3391]$
Rule generic	11	0.0000	$[0.0000, 0.0000]$
Rule typed	11	0.0000	$[0.0000, 0.0000]$

Seller predictors receive focal-product, public-market, quote, and candidate-message conditions. Core14 Simulator V3 removes stream-origin information from the shared actor-query representation, represents candidate quotes relative to the checkpoint quote, and receives arrival capacities as support conditions. Dynamic42 additionally uses period indicators and previous count, revenue, and quote summaries.

Persistent entity identifiers are used for record linkage, grouping, and split assignment rather than as predictive features. Private buyer data may be accessed by the independent evaluator to construct a task-gain label without being supplied to the predictor. The distinction between a generated exogenous query and an actual future arrival realization is maintained throughout this appendix.

Supervision and effective weights. Natural supervision concerns one complete candidate-set decision, including rejection. Controlled supervision concerns an executed legal action from a shared pre-action parent. Technical supervision concerns an independently evaluated buyer–product pair, possibly attached to several recorded snapshots. Temporal supervision concerns outcomes within specified periods and their associated market conditions. These records may share underlying products or trajectories; separating their target meanings does not make all compiled rows independent.

Let w_n denote the effective weight attached to a compiled example. Each objective uses the weights associated with its own supervision object, rather than treating every joined row or candidate slot as an equally weighted independent observation. For Technical, repeated snapshots of one buyer–product pair receive inverse-multiplicity weights, giving that pair total weight one within the compiled selection.

Partitions and feature provenance. Train, Calibration, Validation, and historical Test records preserve their recorded roles. Calibration maps are fitted on the designated Calibration records and then frozen. An entity-split audit of 200 blueprint manifests reports no overlap of literal blueprint, buyer, seller, or product identities across these partitions, and no cross-split exact observation-hash or normalized blueprint-vector duplicates. Types, personas, product categories, and templates can

nevertheless overlap. The split is therefore entity-disjoint rather than type-disjoint or mechanism-OOD.

Natural's repair and selection path generates held-out market-group features with Technical producers trained without those groups. However, the final all-Train feature file is not fully linked row by row to its originating producers. Consequently, the verified out-of-fold selection procedure does not establish out-of-fold provenance for every formal training row. Frozen inference uses the recorded producers without reading target gains. Remaining fit-count and artifact-identity qualifications are collected in Section B.6.

B.2 Natural and Controlled

Natural: masked candidate-set prediction. Natural accepts four candidate slots, at least one visible candidate, and an explicit reject-all output. The public context has 26 dimensions. Each candidate has 59 base dimensions and four raw Technical summaries, giving 63 dimensions after augmentation. A shared candidate encoder and a separate context encoder each produce 48-dimensional representations.

For candidate mask $m_j \in \{0,1\}$, context x_{ctx}, and candidate features x_j, the computational form is

$$\begin{aligned} h_j &= f_{\mathrm{cand}}(x_j), \qquad & h_{\mathrm{ctx}} &= f_{\mathrm{ctx}}(x_{\mathrm{ctx}}), \\ \bar{h} &= \frac{\sum_{j=1}^{4} m_j h_j}{\sum_{j=1}^{4} m_j}, \qquad & \ell_j &= f_{\mathrm{score}}(h_j, h_{\mathrm{ctx}}, \bar{h}), \\ \ell_{\mathrm{reject}} &= f_{\mathrm{reject}}(h_{\mathrm{ctx}}, \bar{h}), \qquad & p_{\mathrm{N}} &= \mathrm{softmax}(\ell). \end{aligned} \tag{16}$$

Invisible candidate logits are set to $-\infty$, while rejection remains legal. Jointly permuting candidate features and masks permutes candidate probabilities and preserves the rejection probability through shared encoding and pooling. Adding a candidate can still change the probabilities assigned to existing candidates.

For example n, let $\mathcal{L}_n$ contain its visible candidates and rejection, and let c_n be the observed choice. Label smoothing is applied only to this legal support, with $\epsilon = 0.10$:

$$\mathcal{L}_{\mathrm{N}} = -\frac{1}{\sum_n w_n} \sum_n w_n \sum_{j \in \mathcal{L}_n} \left[(1-\epsilon)\mathbf{1}\{j = c_n\} + \frac{\epsilon}{|\mathcal{L}_n|} \right] \log p_{\mathrm{N},n}(j). \tag{17}$$

Invalid slots receive neither target mass nor loss. Effective weights apply to candidate-set decisions, not individual slot rows. The final ensemble equally mixes three categorical predictive distributions; it does not average parameters or select the best seed.

The four internal Technical features. For the three frozen Technical members, let (μ_i, s_i, ν_i) denote their log-loss-ratio parameters. Define

$$\begin{aligned} \bar{\mu} &= \frac{1}{3} \sum_{i=1}^{3} \mu_i, \\ v_R &= \frac{1}{3} \sum_{i=1}^{3} \left(s_i^2 \frac{\nu_i}{\nu_i - 2} + \mu_i^2 \right) - \bar{\mu}^{\,2}. \end{aligned} \tag{18}$$

The stored feature vector is

$$z_{\mathrm{T}} = \left(\bar{\mu}, \log \sqrt{v_R}, \frac{1}{3} \sum_{i=1}^{3} F_{t,\nu_i}\left(-\frac{\mu_i}{s_i} \right), \sqrt{\frac{1}{3} \sum_{i=1}^{3} (\mu_i - \bar{\mu})^2} \right). \tag{19}$$

Here $F_{t,\nu}$ is the standard Student-t cumulative distribution function with ν degrees of freedom. The implementation protects the variance calculation using $\max\{\nu_i - 2, 10^{-6}\}$ in the denominator and floors the resulting variance at 10^{-8} before taking its square root. These protections do not truncate the predictive support.

The first coordinate is a log-loss-ratio location summary, not expected transformed task gain. The third uses the identity $G > 0 \iff R < 0$ for the continuous Technical law. All four coordinates remain raw under the frozen calibration contract: calibrated gain samples are not fed back to Natural.

Controlled: completion and conditional payment. Controlled maps its public/action representation through layers of width 256 and 128 to a 64-dimensional hidden vector. Its completion head is Bernoulli.

Conditional payment is represented on the legal cent grid. Let A_c be the current ask in integer cents and let $c \in \{1, \dots, A_c\}$ be the payment in cents. For $A_c > 1$, define

$$\begin{aligned} l_c &= \max\left\{0,\ \frac{c - \frac{1}{2} - 1}{A_c - 1}\right\}, \\ u_c &= \min\left\{1,\ \frac{c + \frac{1}{2} - 1}{A_c - 1}\right\}, \\ b_{A_c}(c; \alpha, \beta) &= F_{\mathrm{Beta}(\alpha,\beta)}(u_c) - F_{\mathrm{Beta}(\alpha,\beta)}(l_c). \end{aligned} \tag{20}$$

The payment mass function is

$$p_{\mathrm{pay}}(c \mid D = 1, x) = \rho\, \mathbf{1}\{c = A_c\} + (1 - \rho) b_{A_c}(c; \alpha, \beta). \tag{21}$$

The Beta shape parameters use $\mathrm{softplus}(\cdot) + 0.20$, and ρ is a sigmoid output. At $A_c = 1$, payment is deterministic. At the upper endpoint, both the ordinary Beta cell and the explicit point mass contribute. Likelihood evaluation retains the implementation's numerical probability floor.

This is a payment distribution conditional on completion. It is not an unconstrained continuous amount and is not the buyer's initial offer.

Applicable feedback. Four conditional Bernoulli heads predict whether a review is made, whether an existing review is positive, and whether its reason is `HONESTY` rather than `EFFICACY` within each polarity. Let their probabilities be r, v, h_+, and h_-. The five mutually exclusive categories, conditional on completion, have probabilities

$$\pi_{\mathrm{review}} = \left(1 - r,\ rvh_+,\ rv(1 - h_+),\ r(1 - v)h_-,\ r(1 - v)(1 - h_-)\right). \tag{22}$$

The category order is no review, positive–honesty, positive–efficacy, negative–honesty, and negative–efficacy. `HONESTY` concerns agreement between declaration and delivery, not whether the reviewer is truthful. An absent review has no polarity or reason. Termination is retained as branch metadata but has no separate learned head.

With observed completion d, review indicator b, polarity v^{obs}, and reason h^{obs}, the row loss follows only applicable branches:

$$\begin{aligned} \ell_C = & -\log p_D(d \mid x) \\ & + d\Big[-\log p_{\mathrm{pay}}(c \mid D = 1, x) - \log p_B(b \mid D = 1, x) \\ & \quad + b\big\{ -\log p_V(v^{\mathrm{obs}} \mid B = 1, D = 1, x) \\ & \qquad - \log p_H(h^{\mathrm{obs}} \mid V = v^{\mathrm{obs}}, B = 1, D = 1, x)\big\}\Big]. \end{aligned} \tag{23}$$

The training objective is $\sum_n w_n \ell_{C,n} / \sum_n w_n$. The indicators denote applicability masks: unavailable targets are not evaluated and then multiplied by zero. Payment and feedback terms are absent on non-deals; polarity and reason terms are absent on non-reviewed deals. Independent Technical labels are not an additional Controlled loss term.

Generated arrival queries in the historical implementation. For each row and Controlled member, the collector constructs 256 queries using next-14-day arrival summaries generated under the recorded demand regime and checkpoint time. They integrate a known exogenous mechanism rather than representing 256 observed future trajectories. If q_b and π_b are the completion probability and the vector in Eq. (22) for query b, the historical collector uses

$$\bar{q} = \frac{1}{256} \sum_{b=1}^{256} q_b, \qquad \bar{\pi} = \frac{1}{256} \sum_{b=1}^{256} \pi_b. \tag{24}$$

The averaged five-category vector is converted back into the stored conditional feedback heads. Neither feedback aggregation nor payment-query selection reweights queries by their completion probabilities.

For payment, a deterministic permutation selects 64 of the 256 query indices. A separately permuted grid of 64 midpoint probabilities selects values from the endpoint-inflated Beta predictions. Values are rounded onto the legal cent support and converted to the experiment's token scale. Completion probabilities are averaged over queries before member mixing and frozen binary calibration. Thus 256 is a mechanism-query count, while 64 is a retained payment-sample count per member.

Conditional-target correction and its implementation status. The later mathematical check distinguishes the historical average from a mixture conditional on completion. For a fixed state and action, let π_z be the probability of arrival condition z, let $q_z = P(D = 1 \mid S, a, z)$, and let F_z be the payment distribution conditional on completion and z. For $\bar{q} = \sum_z \pi_z q_z > 0$, the corresponding conditional mixture is

$$F_{\text{pay}|D=1,S,a} = \sum_z \frac{\pi_z q_z}{\sum_{z'} \pi_{z'} q_{z'}} F_z. \tag{25}$$

The same conditioning weights apply when mixing completed-transaction feedback categories. When $\bar{q} = 0$, these conditional distributions are undefined and are masked rather than assigned a fabricated payment or review distribution.

A two-condition unit example uses $\pi = (0.5, 0.5)$, $q = (0.9, 0.1)$, and deterministic conditional payments $(10, 100)$ in a common illustrative unit. The completion probability is 0.5, conditional weights are $(0.9, 0.1)$, the conditional payment mean is 19, and expected realized payment is 9.5. The unweighted conditional average would instead be 55.

This correction has been formulated, implemented as a numerical utility, and tested, but the per-member, per-arrival-query outputs required for full prediction regeneration and recalibration were not recovered. A globally calibrated completion probability does not recover those query-level quantities. Consequently, the reported historical payment, feedback, and Expected Net results retain Eq. (24); they are not results from a fully corrected model. Gain–price scoring does not require this conditional-payment mixture.

B.3 Technical Predictors and Operational Value Summaries

Historical log-loss-ratio model. Technical receives 16 buyer-structure dimensions, 48 product-structure dimensions, and a 384-dimensional declaration embedding. Three trainable branches project these inputs to 32, 64, and 32 dimensions, respectively. Their concatenation is a 128-dimensional representation. Two further layers produce the location, scale, and degrees of freedom of the log-loss ratio:

$$\begin{aligned} R &= \log \frac{L_{\text{aug}}}{L_{\text{base}}}, \\ R \mid x &\sim t_\nu(\mu, s), \\ s &= \text{softplus}(s_0) + 0.05, \qquad \nu = 2 + \text{softplus}(\nu_0). \end{aligned} \tag{26}$$

Its objective is the effective-weighted negative Student-t log density of the measured R. Inverse-multiplicity weighting gives each compiled buyer–product group total weight one; it is not class balancing or an inferred importance-sampling correction. Compilation requires a finite positive augmented-to-baseline loss ratio. That requirement constrains valid labels, not predictive support.

Finite retained predictions. For each member and prediction row, the historical collector retains $K = 64$ midpoint quantiles in a deterministic row/member-specific permutation σ_i:

$$\begin{aligned} u_k &= \frac{k - \frac{1}{2}}{K}, \\ R_{ik} &= \mu_i + s_i F^{-1}_{t,\nu_i}\left(u_{\sigma_i(k)}\right), \\ G_{ik} &= 100\left(1 - \exp(R_{ik})\right), \qquad k = 1, \ldots, K. \end{aligned} \tag{27}$$

These are finite quantile evaluations, not independent realized trades or identically distributed Monte Carlo draws. Split-specific collectors retain their corresponding semantic randomization identities. Nonfinite transformed values cause a failed numerical check rather than silent clipping into a newly truncated model. Subsequent marginal calibration acts on the finite sample representation.

The untruncated transformed mean. For positive scale and finite degrees of freedom, a Student-t density has a polynomial positive tail. There exist constants $c > 0$ and r_0 such that $f_R(r) \geq cr^{-(\nu+1)}$ for $r \geq r_0$. Therefore,

$$\mathbb{E}[\exp(R)] \geq c \int_{r_0}^{\infty} \exp(r) r^{-(\nu+1)}\, dr = \infty, \qquad \mathbb{E}[100(1 - \exp(R))] = -\infty. \tag{28}$$

The second expectation is extended-real: G is bounded above by 100, so its positive part is integrable, while its negative part has infinite expectation. The restriction $\nu > 2$ provides finite variance of R, not a finite exponential moment.

When numerical checks pass, the retained values in Eq. (27) have finite means and tail summaries. Those functionals describe the retained empirical prediction. They do not estimate a finite mean of the untruncated transformed law, because that mean does not exist. Finite-sample calibration does not by itself establish truncation of the original law or a general tail guarantee.

Later Technical baselines and the value interface. The baseline archive separately identifies `lightgbm_quantile` and `linear_student_t` under the `structured_no_text` input variant. The listed LightGBM Validation record uses five quantile regressors with seed 11; the linear baseline is recorded as deterministic. These entries are not members of the historical three-member neural Technical mixture.

E1 compares their recorded gain summaries with the frozen Technical summary under the same gain–price action rule. The selected quantile LightGBM predictor supplies individualized value information in E2 and E3. The historical gain–price rule uses the raw finite Technical gain summary, rather than the calibrated 576-element net-surplus construction below. Replacing the external gain summary in E1–E3 does not replace the four internal features in Eq. (19), retrain Natural, or regenerate historical advice.

The executed LightGBM value producer trains directly on the transformed task-gain target

$$g = 100\{1 - \exp(\rho)\}, \tag{29}$$

where ρ is the recorded log-loss ratio used in the Technical structured data. The feature vector is the 64-dimensional structured Technical feature representation produced from the buyer public profile and public product fields, using a zero text encoder. Those public product fields include the verified-card flag, listing claim, completed and pending sales, review counts, product and seller reputations, and review-coverage summaries. The fitted models are LightGBM quantile regressors at

$$\tau \in \{0.05, 0.25, 0.50, 0.75, 0.95\},$$

with objective `quantile`, 240 estimators, learning rate 0.04, 31 leaves, maximum depth 7, minimum child samples 40, and random seed 11 (Ke et al., 2017). For one buyer–product pair, the five predicted quantiles are sorted to remove crossing for the displayed summary, linearly interpolated on the fixed grid $t_m = (m - 1/2)/192$ with constant endpoint extension, and averaged:

$$\widehat{g}_{\mathrm{LGBM}}(S, j) = \frac{1}{192} \sum_{m=1}^{192} \widetilde{q}(t_m). \tag{30}$$

No additional post-hoc calibration is applied to this value producer before the gain–price subtraction in E1 or before E3 value-table construction.

In E1 the score for a product action is $\widehat{g}_{\mathrm{LGBM}}(S, j(a)) - \mathrm{offer}(a)$, with reject scored as zero and a positive score required to trade. For a fixed product, this rule mechanically prefers the lower supported offer among legal offers because the gain term is unchanged; it does not model the completion-probability effect of offering less. In E3 the buyer-facing value table uses the same gain summary for V_{IND}. Rows contain the legal-action label, product reference, supported offer, predicted technical gain, gain minus offer, an estimated deal probability, the value source, and a short note. Accept rows use the current public ask as the supported offer; counter rows, when legal, display a 65% ask anchor clamped to the legal counter interval. These rows inform the advisor prompt; they do not force the participant to choose the displayed action or price.

For V_{REF}, the value table keeps the same schema but uses training-reference values. The reference technical gain is the mean training-split technical gain within product lineage, falling back to the

global training mean when no lineage value is available. Reference deal probabilities come from the training-rollout buyer scenario summaries for accepting the current ask or countering the observed train price. For V_{IND}, the displayed deal probability is the mean transaction probability from the current individualized descriptive table. Thus the E3 treatment changes the source of the buyer-value information package, including gain and deal-probability sources, while preserving the value-field contract and the individualized descriptive prediction source. It does not isolate the gain field alone. Different later decisions can produce different market states and descriptive values despite a common prediction source. The remaining provenance items are the artifact hashes, feature-builder version, and train/reference files listed in Appendix F; they are not inferred from aggregate CRPS values.

B.4 CORE14 AND DYNAMIC42

Core14 count, timing, and revenue. Core14 Simulator V3 conditions on public context, a candidate message and quote, and supplied arrival support. For J opportunities in the 14-day window, the count distribution is beta-binomial:

$$p(N = n \mid J, x) = \binom{J}{n} \frac{\mathrm{B}(n + \alpha, J - n + \beta)}{\mathrm{B}(\alpha, \beta)}, \qquad n = 0, \ldots, J, \tag{31}$$

where $\mathrm{B}(\cdot, \cdot)$ denotes the beta function. First-sale hazards use seven bins with masks determined by the supplied arrival summary. Revenue is conditioned on count and first-sale timing using a legal cent-grid law with endpoint mass. No sale implies zero revenue. The joint sampling correspondence is retained when these outputs become short states for Dynamic42; independently resampling the marginals would define a different state input.

Anchored quote representation. A candidate quote p is represented relative to the checkpoint quote p_0. For $\xi = \log(p/p_0)$, the basis is

$$\mathcal{B}(\xi) = \left(\xi,\, [\xi - \kappa_1]_+ - [-\kappa_1]_+,\, \ldots,\, [\xi - \kappa_4]_+ - [-\kappa_4]_+\right), \qquad \mathcal{B}(0) = 0, \tag{32}$$

with knots $(-0.40, -0.20, 0.20, 0.40)$ and $[z]_+ = \max\{z, 0\}$. A rank-eight interaction multiplies projected context and basis coordinates; the basis and interaction feed the count, timing, and revenue components. Candidate-message effects remain separate, so a natural action and a neutral-message intervention at the same quote need not have identical predictions. The anchor is a representation constraint, not a guarantee of monotone or causally identified quote effects.

Temporal conditions and factorization. Let Z collect current public focal-product, market, and action information. Let $A_{1:3}$ denote the supplied arrival-support conditions for three successive 14-day periods. A period outcome V_r contains its count increment, revenue increment, and quote. Within a specified Simulator/S1 member pair, the computational chain is

$$\begin{aligned} p(V_1, V_2, V_3 \mid Z, A_{1:3}) = {} & p_{14}(V_1 \mid Z, A_1) \\ & \times p_{\mathrm{S1}}(V_2 \mid V_1, Z, A_{1:2}) \\ & \times p_{\mathrm{S1}}(V_3 \mid V_1, V_2, Z, A_{1:3}). \end{aligned} \tag{33}$$

S1 shares parameters across periods two and three, distinguished by period indicators. The archived conditional evaluation supplies actual trajectory arrival capacities, so this factorization does not establish prediction from contemporaneous Z alone. The full mapping from those conditions to generated online inputs remains incompletely bound in the historical record.

Two history slots. Each slot encodes period count n, capacity c, ask a in cents, revenue r in cents, and validity flag m:

$$\begin{aligned} H(n, c, a, r, m) = m\Bigg(& \frac{n}{\max(c, 1)},\, \frac{\log(1 + n)}{4},\, \frac{r}{\max(na, 1)},\, \frac{\log(1 + r)}{12}, \\ & \frac{\log(1 + a)}{12},\, \frac{a}{a + 1000},\, \mathbf{1}\{n = 0\},\, m \Bigg). \end{aligned} \tag{34}$$

Here a is the slot's ask, not the action variable elsewhere. The two slots supply 16 features to a bias-free $16 \to 64$ projection that is initialized to zero and subsequently learned. Its output is added to the base state before GELU. Period two uses the first-period slot and an inactive second slot; period three uses both preceding periods. Cached first-sale timing is not read by S1.

Teacher forcing and generated paths. Training constructs history slots from observed adjacent-period outcomes. Price, count, and conditional revenue log probabilities are computed under those histories with effective weights. The implementation returns separate loss components, but not all auxiliary coefficients and preprocessing versions are fully bound in the available record. No universal joint loss with guessed coefficients, scheduled-sampling procedure, or joint update of Core14 is implied.

Within a generated later period, the quote is sampled first, then count, then conditional revenue. The quote law contains atoms at the initial quote, one cent, and the price ceiling, together with two discretized Beta components on the legal quote grid. Count has beta-binomial support $0, \dots, J_r$. For positive count N_r and generated ask Q_r^{c} in cents, revenue lies on the cent grid between N_r and $N_r Q_r^{\mathrm{c}}$, using an endpoint-inflated Beta law. Zero count implies zero revenue.

Each raw Core14 joint sample supplies first-period count and revenue, the initial quote, and capacity. S1 generates period two from that sample, then constructs period-three history using the same path's first two periods. The same S1 member advances that path through both later periods. For count increments (N_1, N_2, N_3) and token-revenue increments (R_1, R_2, R_3), the reported cumulative vector is

$$Y = (N_1 + N_2,\ N_1 + N_2 + N_3,\ R_1 + R_2,\ R_1 + R_2 + R_3)\,. \tag{35}$$

The order is count28, count42, revenue28, and revenue42. Quotes are not added to revenue. Observed histories in training and generated histories in inference are different conditions; prediction error can therefore propagate through the chain.

Mixture members and loaded producers. Three Simulator V3 members produce 128 raw joint paths each, giving 384 short paths. Each of three S1 members advances all 384 paths. The nine Simulator/S1 combinations contribute equally, producing 1,152 equally weighted paths per row on the bound final Calibration/Validation route and the archived 57-market evaluation. Five additional Core14 summary producers are feature dependencies, not extra Cartesian mixture axes.

The separate 60-market post-hoc sensitivity descriptor reports 1,200 draws. Its underlying sample-generation mapping was not independently recovered; that record retains its separate 1,200-draw identity.

Core14 Simulator V3 and the separate Core14 summary-producing model also have different functions. The seller-scoring and temporal short-state roles share the recorded C6 checkpoint identities, but the short-state cache retains raw joint draws even when seller-score reporting uses seller-specific calibration. A shared family name does not make checkpoints, calibrators, or sampling roles interchangeable.

B.5 Calibration, Finite Composition, and Risk

Calibration scope. Historical calibration maps are fitted on Calibration and then frozen. Natural applies a positive global categorical temperature after mixing member probabilities. Binary heads use positive-slope Platt maps on mixed log odds:

$$C_{\mathrm{bin}}(p) = \mathrm{sigmoid}\left(a_{\mathrm{cal}}\,\mathrm{logit}(p) + b_{\mathrm{cal}}\right), \qquad a_{\mathrm{cal}} > 0, \tag{36}$$

with the recorded endpoint protections.

Gain and conditional payment use 20-knot weighted empirical probability-integral-transform (PIT) maps. Feedback is mixed as a five-category distribution, converted back into conditional heads, and then processed by the applicable binary maps. These are descriptions of the historical calibration routes, not assumed calibration settings for every later baseline.

Marginal transformation and restored indices. For a finite vector $x_1, \dots, x_M$, let π be its stable sorting permutation. Let Q_x interpolate the sorted values linearly at midpoint positions $v_j = (j - \frac{1}{2})/M$. For a frozen monotone PIT map C, the transformed values satisfy

$$\widetilde{x}_{\pi(j)} = Q_x\big(C^{-1}(v_j)\big)\,, \qquad j = 1, \dots, M. \tag{37}$$

The inverse-axis interpolation deduplicates repeated values by retaining the first knot position. Outside the empirical midpoint range, interpolation uses endpoint values.

Crucially, transformed sorted values are scattered back to their original sample columns. Sorting is a marginal calibration operation, not a final instruction to pair gain and payment by common quantile rank. Repeated values retain their multiplicities.

The 576-element net-surplus construction. For a fixed product–action prediction row, the raw Technical and Controlled arrays each have three members and $K = 64$ retained samples. Write their values as G_{ik} and P_{jk}. The construction crosses gain and payment members but retains only the matched sample index within each member pair:

$$\begin{aligned} m &= K\,[3(i-1)+(j-1)]+k, \\ g_m &= G_{ik}, \qquad p_m = P_{jk}, \\ u_m &= \widetilde{g}_m - \widetilde{p}_m, \qquad i,j \in \{1,2,3\}, \quad k \in \{1,\dots,K\}. \end{aligned} \tag{38}$$

Each expanded margin has $9K = 576$ entries. The gain and payment margins undergo Eq. (37) separately before subtraction. The construction is not a 64×64 product within each member pair.

The direct component-scoring route instead mixes $3 \times 64 = 192$ samples. Already calibrated component-scoring arrays are not inputs to the 576-column net routine; substituting them would change the calibration order and could apply calibration twice. The 256 arrival-query count is not another multiplier in this construction.

Let q be the completion probability after query averaging, member mixing, and frozen binary calibration. The finite operational law is

$$\widehat{F}_U = (1-q)\delta_0 + \frac{q}{576}\sum_{m=1}^{576}\delta_{u_m}. \tag{39}$$

Its atoms represent a computational composition, not shared realized future trajectories. In particular, the construction does not identify the true dependence between gain, payment, and completion. Technical is not trained only on completed trades, so its pair-level gain law is not established as the gain law after selection into completion. For non-product actions, the compiler retains its separate zero realized-value handling without assigning a buyer–product gain prediction.

Mean and event-probability outputs. Under Eq. (39),

$$\begin{aligned} \widehat{\mu}_U &= \frac{q}{576}\sum_{m=1}^{576} u_m, \\ p_+^{\mathrm{emp}} &= \frac{q}{576}\sum_{m=1}^{576} \mathbf{1}\{u_m > 0\}, \\ p_-^{\mathrm{emp}} &= \frac{q}{576}\sum_{m=1}^{576} \mathbf{1}\{u_m < 0\}, \\ p_+^{\mathrm{out}} &= c_+\big(p_+^{\mathrm{emp}}\big)\,. \end{aligned} \tag{40}$$

The separately fitted positive-event map c_+ changes that scalar output, not the support or weights of $\widehat{F}_U$. Policy preparation reads the net samples and completion probability; it does not replace every risk functional with p_+^{out}. The negative-event probability is not $1 - p_+^{\mathrm{out}}$: non-completion and completed zero-surplus outcomes contribute zero mass, and the positive-event calibration is separate.

The unweighted mean of the 576 differences is conditional only within this operational completed-branch construction. Multiplying it by q gives the finite expected realized surplus. Neither quantity resolves the conditional-gain or dependence limitations of the underlying model.

Weighted lower-tail risk. For sorted support atoms $(u_{(j)}, w_{(j)})$, including the non-completion zero atom, define

$$W_j = \sum_{\ell=1}^{j} w_{(\ell)}, \qquad W_0 = 0, \qquad j^\star = \min\{j : W_j \geq \alpha\}. \tag{41}$$

The lower-tail mean is

$$\mathrm{CVaR}_{\alpha}^{-}(\widehat{F}_U) = \frac{\sum_{j<j^\star} w_{(j)} u_{(j)} + (\alpha - W_{j^\star-1})\, u_{(j^\star)}}{\alpha}. \tag{42}$$

Only the required part of the boundary atom is included. If negative mass is less than α, part of the lower tail can be zero rather than exclusively a completed-trade loss. Ties and repeated samples retain their recorded weights.

The historical supported objective uses $\alpha = 0.10$ and $\lambda = 0.50$ in Eq. (5). At $\lambda = 0$, it becomes Expected Net Benefit. The legal support and deterministic tie rule are unchanged, and the utility-scored rule rejects when no strictly positive eligible score remains.

Realized-rollout tail summaries use a different convention. Within a selected parent/action and evaluation repeat, they average the lowest $\lceil \alpha n \rceil$ observed outcomes, with $n = 30$ for ordinary policies and $n = 15$ for an Oracle evaluation half. Those summaries are averaged within blueprint and then across blueprints. They are not Eq. (42) applied to predictive atoms, nor the CVaR of one pooled cross-market distribution. The archived negative-net frequency also includes all selected rollouts, not only completed trades. Conditional-loss summaries retain their own grouping and eligible denominators in Appendix D.

Temporal calibration and score normalization. The frozen v2 temporal calibrator uses shrinkage monotone piecewise-linear PIT maps fitted on Calibration. It transforms mixed cumulative predictions while preserving $\mathrm{count42} \geq \mathrm{count28}$ and $\mathrm{revenue42} \geq \mathrm{revenue28}$. No-sale probability has a separate map. Raw short-state generation, calibrated cumulative prediction, and training climatology remain different output identities.

For a scalar empirical distribution with equally weighted samples $x_1, \ldots, x_M$, the scored CRPS is

$$\mathrm{CRPS}(\widehat{F}, y) = \frac{1}{M} \sum_{m=1}^{M} |x_m - y| - \frac{1}{2M^2} \sum_{m=1}^{M} \sum_{n=1}^{M} |x_m - x_n|. \tag{43}$$

Efficient implementations need not construct the full pairwise matrix. The score does not permit recovery of a predictive distribution from its reported mean value.

For the four temporal targets in Eq. (35), the training-only weighted empirical scale is $s_h = \max\{1, \mathrm{IQR}_{\mathrm{Train},h}\}$. The recorded joint score is

$$\begin{aligned} \mathrm{Joint} = \frac{1}{3}\Bigg[& \frac{1}{4} \sum_{h=1}^{4} \frac{\mathrm{CRPS}_h}{s_h} + \mathrm{Brier}_{\text{no sale42}} \\ & + \mathrm{Energy}\left(\widehat{F}_{Y/s}, Y^{\mathrm{obs}}/s\right) \Bigg]. \end{aligned} \tag{44}$$

Energy uses the recorded average of 16 fixed sample-pairing estimates. Rows are aggregated with their effective weights. The four marginal CRPS terms share one score family; they do not each receive one third of the total weight. No-sale probability is the raw zero-count fraction or its frozen mapping, according to the evaluated row.

Core14's baseline-relative skills have different units and conventions from these losses. All task-specific metrics and comparison pools are preserved in Appendix C, rather than collapsed into a universal model ranking.

B.6 Training Settings and Interface Versions

Recorded formal training settings. Table 10 lists the bound settings of the five historical main model families. Each uses three members with seeds 20260723, 20260724, and 20260725. The later baseline models do not inherit this three-member configuration.

Natural, Technical, final S1, and Controlled use AdamW (Loshchilov & Hutter, 2019); the Controlled run configuration does not explicitly bind its weight decay. Core14 C6 FULL CF separately records 37 batches per epoch, 100 context-only steps, rank eight, and bf16 precision. Its two held-out price actions are 0.67 and 1.25 times the checkpoint ask.

Table 10: Recorded formal training settings. NR denotes a value not explicitly bound in the inspected run configuration, not zero or an assumed library default. "Fixed all-Train" retains the archived fit designation. Core14's batch is a panel batch, not a count of independent markets.

Model	Steps	Batch	Learning rate	Weight decay	Selection
Natural	300	256	3×10^{-4}	0.003	Fixed all-Train
Controlled	400	128	3×10^{-4}	NR	Fixed all-Train
Technical	400	128	5×10^{-4}	10^{-4}	Fixed all-Train
Core14 C6	592	128 panels	3×10^{-4}	10^{-4}	Frozen $t^\star = 16$
Dynamic42 S1	400	256	3×10^{-4}	10^{-4}	Fixed all-Train

Table 11: Model roles and their sampling or calibration identities. Loaded roles are not independent experiments. The external E1/E2/E3 value predictor is not a replacement for Natural's internal Technical feature producers.

Role	Model identity	Output and calibration boundary
Natural output	Formal set-choice model	Three members; mixed categorical output and temperature
Natural value features	Raw historical Technical producers	Three producers; four raw features; no calibrated-gain feedback
Controlled output	Formal buyer-outcome model	Three members; 256 queries and 64 retained payment samples per member
Technical output	Formal log-loss-ratio model	Three members; 64 gain samples per member; finite marginal calibration
Seller short score	C6 FULL CF Simulator V3	Three members; 384 paths; seller-specific calibration
Temporal short states	Same C6 checkpoint identities	Three members; raw joint cache of 384 paths
Later periods	V12 final S1	Three members; nine combinations; 1,152 paths on the bound route
Temporal summary features	Separate Core14 summary model	Five producers; distinct from Simulator V3
E1/E2/E3 explicit value	Selected quantile LightGBM baseline	External gain summary for action scoring and value information; separate from the historical internal feature path

The common-feature structure comparison is a different experiment: it uses 21-dimensional inputs, 260 AdamW steps, learning rate 10^{-3}, and weight decay 10^{-4}. It is not a parameter-matched training comparison of the complete probabilistic models. Its architectures and scores belong to Appendix C.

Fit counts and unresolved mappings. Natural's formal fit contains 38,751 candidate sets and 180 effective market-group weight units. Controlled contains 60,834 compiled fit rows with a recorded input-array hash; the separately reported count of 1,440 parents is not independently bound to that formal file. Technical contains 14,293 compiled rows and 5,760 pair groups, whereas the historical 19,070-snapshot description has an unresolved different counting scale. Final Dynamic42 fitting uses 5,290 rows, with reported Calibration and Validation scales of 3,560 and 720. These are record or effective-weight counts, not interchangeable independent-market counts.

The final Natural feature-producer mapping, some temporal auxiliary coefficients and preprocessing versions, and parts of the online producer-to-checkpoint correspondence remain incomplete. Technical's Calibration and Test descriptors identify the same member checkpoints and logical mixture but have different descriptor-file hashes whose byte-level difference remains unexplained. The frozen calibration-package identity matches the Test raw-prediction identity. These are documented provenance limits; they are not resolved by altering parameters or relabeling old predictions.

Loaded roles and prediction outputs. Table 11 distinguishes trained members, feature producers, raw caches, and calibrated outputs. In the historical seller path, three Simulator members, three S1 members, and five summary producers constitute eleven loaded roles. The nine predictive combinations are the 3×3 Simulator/S1 pairs, not nine extra models in addition to those roles.

Information exposure by study. For the historical descriptive-advice study, a local orchestrator constructs the prescribed public context and obtains forecast summaries before one advisor completion. The buyer advisor receives Natural and Controlled context, while the seller advisor receives the designated Dynamic42 V12 route. The three Technical producers remain internal to Natural. Explicit Technical outputs and automatic net-surplus optimization are not exposed through this path. The orchestrator is not an autonomous multi-round agent deciding which tools to invoke, and advice does not replace participants' final actions.

The later official-API format-control study retains the descriptive versus reference-information distinction, while E2 and E3 separately expose explicit value information using the selected external predictor. In E3, V_{REF} and V_{IND} use the same individualized descriptive prediction source and field contract, while differing in the source of buyer-facing value information. These treatments do not retroactively modify the historical forecasts, calibrators, or prompts. The historical JDCloud runtime and the official-API experiments preserve their separate hosted-model and execution identities.

Historical treatment notes also belong to their original contexts. In particular, the approximately 44% seller-score reduction quoted in a historical tool note comes from the 60-market post-hoc pool, not the qualified 57-market result. It is not relabeled as a different accuracy result in retrospect. The designed no-OOD scope is not an online OOD detector or a certificate of generalization.

Detailed arm definitions, prompt identities, failures and retries, and online analyses appear in Appendix E. Fixed-support predictor selection and the E2 interface evaluation appear in Appendix D. Appendix F records the remaining asset and release requirements.

C Complete Predictive Results and Diagnostics

This appendix reports the predictive comparisons underlying RQ2. It distinguishes whether the available inputs contain predictive information from how well a particular estimator extracts that information. The consequences of using predictions for action selection and advice are evaluated separately in Appendices D and E. Model definitions, calibration procedures, and implementation limitations appear in Appendix B. All results below retain their original evaluation pools and prediction versions.

C.1 Evaluation Pools and Metrics

Evaluation pools. Historical buyer and Core14 scores retain their documented Test-access qualifications. The later Natural, Controlled, and Technical baseline comparisons use Validation. Core14 baseline results are divided into Validation and exploratory Test records. Comparisons are made within these pools, rather than treating scores from different splits as matched observations.

The temporal analyses have four distinct settings. The 57-market evaluation follows integrity isolation and supports an exploratory held-out prediction claim. The 60-market complete-pool analysis is post-hoc sensitivity, not an independent confirmatory evaluation. Validation short-state substitutions and grouped Train-inner comparisons have separate purposes and sample definitions. These temporal evaluations condition on trajectory-arrival capacities in addition to current public information; they do not establish unrestricted forecasting from the public checkpoint state alone.

Targets and metric scales. Natural is evaluated using listwise negative log-likelihood (NLL), Top-1 and Top-2 accuracy, expected calibration error (ECE), multiclass Brier score, and rejection-event Brier score where available. Controlled separates completion probability, payment conditional on completion, and applicable feedback. Technical gain, the sign of potential gain, and positive realized net surplus are different prediction targets.

CRPS and the normalized Dynamic42 Joint score follow Eqs. (43) and (44). CRPS is evaluated on the corresponding predictive representation using the probabilistic scoring convention of Gneiting & Raftery (2007). Lower losses and higher accuracy or baseline-relative skill are preferable. Coverage is interpreted relative to its nominal level, rather than maximized independently of interval width or prediction quality. Historical Core14 skill, the Core14 baseline-comparison Joint column, and Dynamic42 Joint have different definitions. Payment and revenue scores retain their original units and normalizations.

Weights, seeds, and unavailable metrics. Each target uses its designated eligible records and effective weights. Candidate slots, action branches, repeated buyer–product snapshots, and sampled prediction paths are not independent market replications. Uncertainty intervals, where reported, retain their analysis-specific grouping and resampling procedures.

NR denotes an unreported metric, except in the structure tables where the captions identify targets absent from a training regime. It never denotes zero. Individual seed results are reported separately. An ensemble score evaluates the ensemble prediction, not the arithmetic mean of member scores. "Frozen" identifies the original prediction outputs held fixed for the corresponding evaluation; later baselines are not substituted into historical services.

The N, C, T, K, S, R, and D identifiers below retain the archived row identities for Natural, Controlled, Technical, Core14, structure, reputation, and temporal diagnostics, respectively. Repeated or coincident records are retained for traceability, but are not counted as independent replications.

C.2 Natural and Controlled Comparisons

Natural choice. The historical Natural evaluation reports listwise NLL 0.469832, Top-1 accuracy 0.831728, Top-2 accuracy 0.960367, and ECE 0.013221. These historical scores are not the baseline against which the later Validation rows are interpreted.

On matched Validation, the frozen Natural model has NLL 0.526403. The stored five-seed DeepSets ensemble (Zaheer et al., 2017) has NLL 0.492231, while the full no-ID LightGBM candidate ranker (Ke et al., 2017) has NLL 0.513164 (Table 12). The seed-specific DeepSets rows also remain visible. Thus the frozen model contains useful choice information without being the strongest estimator on every listed metric.

Empirical, grouped-label-shuffle, and ID-only controls have their own recorded meanings. The ID-only diagnostic does not make persistent entity identifiers admissible inputs to DIM-WM. Its result also does not replace the broader split and feature-provenance checks in Appendix F. None of these later Natural baselines is retroactively substituted into the frozen historical service.

Table 12: Natural-choice Validation results, overall slice. N03 is the frozen three-member v2-calibrated prediction. N04 is the stored five-seed ensemble; N05–N09 retain its individual seeds. All LightGBM rows use seed 11. R1–R7 denote the recorded feature variants defined in Table 24. Remaining empirical/logit controls are recorded as deterministic. MC Brier is multiclass Brier; Reject Brier concerns rejection. No new intervals or across-seed averages are added.

ID and method	NLL	Top-1	Top-2	ECE	MC Brier	Reject Brier
N01 Candidate-count empirical	1.190213	0.348830	0.677149	0.004776	0.671382	0.007706
N02 Conditional logit, no IDs	0.524546	0.809614	0.954886	0.011463	0.279300	0.008388
N03 Frozen Natural	0.526403	0.805956	NR	NR	0.280570	NR
N04 DeepSets ensemble	0.492231	0.817589	0.959711	0.012688	0.266121	0.007319
N05 DeepSets, seed 11	0.492598	0.819752	0.960701	0.015066	0.266073	0.007352
N06 DeepSets, seed 29	0.497143	0.815753	0.959703	0.013641	0.267373	0.007407
N07 DeepSets, seed 47	0.494869	0.816412	0.960270	0.008787	0.266562	0.007586
N08 DeepSets, seed 71	0.507174	0.812273	0.956987	0.011382	0.273463	0.007522
N09 DeepSets, seed 89	0.499482	0.816974	0.958036	0.014483	0.267729	0.007467
N10 ID-only frequency	1.191768	0.333259	0.649218	0.006144	0.671996	0.007715
N11 LightGBM, grouped label shuffle	1.200497	0.355431	0.674528	0.027851	0.671577	0.008192
N12 LightGBM, R1	0.513164	0.816900	0.957921	0.022314	0.274399	0.007460
N13 LightGBM, R2	0.514216	0.812967	0.958750	0.014545	0.275719	0.007462
N14 LightGBM, R3	0.513320	0.814966	0.958759	0.020040	0.274586	0.007476
N15 LightGBM, R4	0.519163	0.812809	0.959463	0.018523	0.277446	0.007476
N16 LightGBM, R5	0.512964	0.815106	0.958262	0.020213	0.274679	0.007430
N17 LightGBM, R6	0.538273	0.806308	0.957516	0.012307	0.286924	0.007615
N18 LightGBM, R7	0.537098	0.806954	0.956875	0.010027	0.286660	0.007603
N19 Train empirical	1.234618	0.333259	0.646074	0.057315	0.692107	0.007714

Completion probability. The historical completion Brier score is 0.068991 and log loss is 0.239807. On the later Validation comparison, frozen Controlled has Brier 0.061082 and log loss 0.217972. The full Platt-calibrated LightGBM row has corresponding scores 0.064854 and 0.247093. Table 13 retains raw and calibrated alternatives, feature removals, and the parent-grouped label-shuffle control. The values are numerical comparisons on the specified completion target, not a test of every action consequence.

Table 13: Controlled completion: complete overall Validation metrics. C01 is the frozen three-member v2-calibrated model. LightGBM rows use seed 11. C02 shuffles labels by parent groups. C03 and C04 preserve calibrated and raw full-model identities. R2–R7 refer to Table 24. Logistic and empirical controls are recorded as deterministic. Predicted and observed rates are retained only where reported.

ID and method	Brier	Log loss	ECE	Predicted rate	Observed rate
C01 Frozen Controlled	0.061082	0.217972	NR	NR	NR
C02 LightGBM, parent-label shuffle	0.130295	0.425767	0.011966	0.852735	0.841958
C03 LightGBM, full Platt	0.064854	0.247093	0.012327	0.851864	0.841958
C04 LightGBM, full raw	0.125325	0.407084	0.025157	0.867115	0.841958
C05 LightGBM, R2	0.064829	0.247151	0.012095	0.851784	0.841958
C06 LightGBM, R3	0.064914	0.247832	0.012139	0.851700	0.841958
C07 LightGBM, R4	0.065002	0.248881	0.011986	0.851935	0.841958
C08 LightGBM, R5	0.065002	0.248881	0.011986	0.851935	0.841958
C09 LightGBM, R6	0.064888	0.247767	0.011973	0.851547	0.841958
C10 LightGBM, R7	0.064854	0.247093	0.012327	0.851864	0.841958
C11 Logistic, Platt	0.068582	0.270958	0.038762	0.851449	0.841958
C12 Logistic, raw	0.124551	0.406840	0.119647	0.849948	0.841958
C13 Train empirical rate	0.133130	0.436660	0.008060	0.850018	0.841958

Feedback and payment. Historical feedback NLL, Top-1, Top-2, and ECE are 1.283096, 0.500870, 0.732398, and 0.046896, respectively. In the later Validation records, frozen feedback NLL is 1.227418, compared with 1.280353 for the LightGBM multiclass baseline. The scored categories and applicable records follow the review cascade in Appendix B; they are not unconditional predictions of all future reviews.

The conditional-payment Validation CRPS is 20.458364 for frozen Controlled and 21.270163 for the five-quantile LightGBM baseline (Table 14). The separate historical payment CRPS is 0.140725 on its recorded evaluation scale. No numerical before/after comparison is made between these scales.

The completion-weighted mixture in Eq. (25) was not fully regenerated and recalibrated. The frozen rows here therefore retain the historical query aggregation. The mathematical correction does not authorize relabeling these scores as corrected-model results.

C.3 Technical Distributional Trade-offs

The historical Technical evaluation reports gain CRPS 5.961552 and sign Brier score 0.330425. Table 15 reports the distinct later Validation comparison. Frozen Technical has CRPS 5.156179, compared with 3.206748 for quantile LightGBM and 3.518700 for the linear Student-t baseline. The global empirical distribution also has lower CRPS than frozen Technical on this comparison. Because the alternatives use public inputs, the frozen model's error cannot be attributed entirely to hidden buyer information.

The metric ordering is not uniform. Quantile LightGBM has the lowest listed CRPS and MAE, but its sign log loss is 1.881663, compared with 0.509097 for the linear model. Its nominal 90% coverage is 0.800456, whereas the linear model's is 0.965430. Higher coverage alone is not necessarily better: both undercoverage and overcoverage must be interpreted relative to the nominal level and the predictive distribution. These summaries do not identify which individual predictions drive the differences.

Task-gain sign versus positive realized surplus. The remaining Technical-dictionary row, T03, is the frozen Technical-plus-Controlled positive-net interface. It reports positive-net Brier 0.007322

Table 14: Controlled feedback and conditional payment on Validation. Panels retain different targets and metric scales. Frozen rows use the recorded three-member v2-calibrated outputs. LightGBM rows use seed 11; other listed baselines are deterministic. Payment width and coverage retain their recorded definitions. NR denotes an unreported metric, not zero or failure.

(a) Applicable feedback categories

ID and method	NLL	Top-1	Top-2	ECE	MC Brier
C14 Frozen Controlled	1.227418	0.530966	NR	NR	0.630155
C15 LightGBM multiclass	1.280353	0.500889	0.764202	0.076002	0.655827
C16 Multinomial logistic	1.298148	0.494075	0.755051	0.064892	0.666650
C17 Train empirical distribution	1.554385	0.465142	0.755051	0.242578	0.777655

(b) Payment conditional on completion

ID and method	CRPS	MAE	Cov.90	Width90
C18 Ask-stratified empirical	31.407399	47.104630	1.000000	185.607476
C19 Frozen Controlled	20.458364	NR	NR	NR
C20 Global empirical distribution	31.598034	47.483212	1.000000	185.607476
C21 LightGBM, blueprint/price-stratified label shuffle	31.628598	47.058099	0.920123	184.313012
C22 LightGBM, five quantiles	21.270163	27.823097	0.930757	130.370302
C23 Linear residual distribution, ridge	30.220835	45.334414	0.893976	178.657256

Table 15: Technical gain: complete overall Validation metrics. T05 uses structured inputs without text and seed 11; T06 uses the corresponding recorded structured-no-text variant. T01, T02, T04, and T06 are marked deterministic in the comparison dictionary. This seed-field designation does not change the underlying historical three-member Technical definition. Missing fields remain NR.

ID and method	CRPS	MAE	Sign Brier	Sign log loss	Cov.90
T01 Buyer/product-type empirical	5.303423	8.760241	0.294231	0.792395	0.878906
T02 Frozen Technical, calibrated v2	5.156179	NR	0.298272	NR	NR
T04 Global empirical distribution	4.455399	4.941280	0.249116	0.691378	0.961719
T05 Quantile LightGBM	3.206748	4.491038	0.218842	1.881663	0.800456
T06 Linear Student-t	3.518700	4.565686	0.169004	0.509097	0.965430

and log loss 0.026259 on its Validation record. It does not report a gain CRPS, gain MAE, or gain-sign score. The target is positive realized surplus, not positive potential task gain, and its output includes the separately calibrated event interface described in Appendix B.

The historical buyer record separately reports positive-net Brier 0.212765, positive-net log loss 0.722534, and net CRPS 6.302743. These records are not merged with T03 or interpreted as an improvement from retraining. A favorable binary event score does not establish a calibrated gain distribution or an identified gain–completion–payment dependence.

Relationship to downstream evaluation. Figure C1 displays the differing prediction criteria without constructing an aggregate ranking. E1 subsequently compares the recorded gain summaries under the same supported-action rule, and E2 uses the selected external value predictor. Those results belong to Appendix D. The predictive comparison here neither substitutes new Technical features into historical Natural nor demonstrates a dynamic-market benefit.

C.4 Temporal Prediction and Short-State Diagnostics

Core14 historical skill and later baseline comparisons. The historical Core14 seller evaluation reports composite skill 0.674647, count skill 0.682368, revenue skill 0.666926, and price-response composite skill 0.407654. Its reported one-sided composite lower bound is 0.621207. These are the designated historical baseline-relative quantities, not the raw losses of the later baseline tables.

Tables 16 and 17 separate the later Core14 Validation and exploratory Test records. The full LightGBM Validation row has Energy 0.495361, compared with 0.613362 for training climatology and 0.619769 for the blueprint-safe label-shuffle row. On the exploratory Test records, frozen FULL-CF

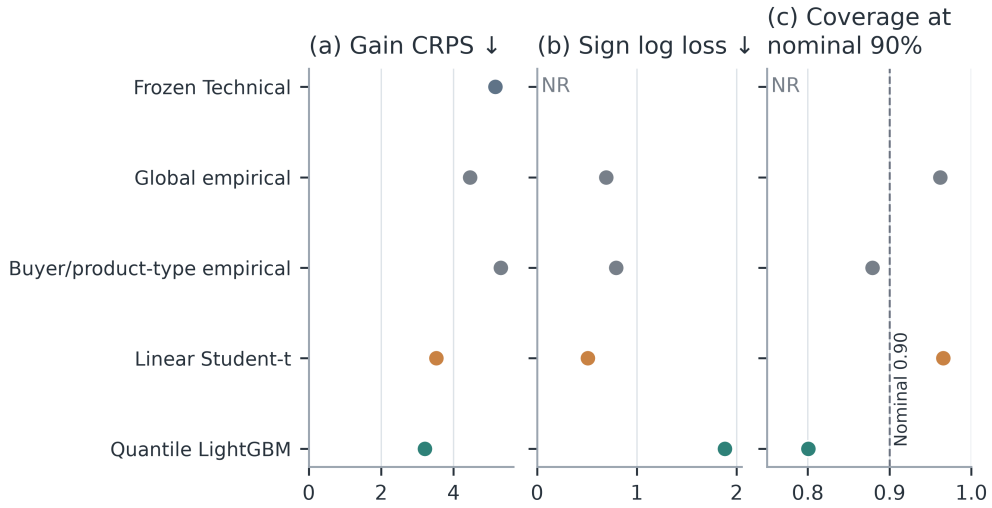

Figure C1: Technical predictive trade-offs on Validation. Separate panels show recorded gain CRPS, sign log loss, and coverage at nominal 90%. The coverage panel includes the 0.90 nominal reference. Missing metrics are omitted rather than plotted at zero. Each displayed point corresponds to its recorded predictor, variant, and seed identity; no new intervals, seed averages, or composite scores are inferred.

has count and revenue CRPS 0.244731 and 73.912024, whereas the matching full LightGBM row has 0.802910 and 226.959401.

These statements compare each model only within its recorded pool. The collection ablations and frozen Core14 rows in the exploratory Test table are not relabeled as Validation results. Their missing joint or coverage fields do not imply zero or failure.

Table 16: Core14 Validation: complete recorded overall metrics. K07 is a hurdle logistic/negative-binomial model with a Tweedie revenue GLM. K08 and K10–K16 are LightGBM hurdle-distribution variants. All rows retain seed field 11 from the comparison dictionary. R1–R7 are defined in Table 24. Joint is the recorded Core14 baseline-comparison column, not the historical composite skill or the Dynamic42 score.

ID and method	Count CRPS	Revenue CRPS	No-sale Brier	Energy	Joint	Count Cov.90	Revenue Cov.90
K07 Hurdle GLM	1.066887	387.854940	0.244145	0.630338	0.383504	0.939555	0.894727
K08 Label shuffle	1.087718	376.981663	0.250222	0.619769	0.383735	0.936217	0.901172
K10 LightGBM R1	0.894193	283.971625	0.215857	0.495361	0.305258	0.969122	0.832971
K11 LightGBM R2	0.892752	286.791264	0.214590	0.496082	0.306055	0.968006	0.831994
K12 LightGBM R3	0.884465	287.866751	0.213082	0.493109	0.304685	0.971549	0.826665
K13 LightGBM R4	0.882485	283.571292	0.212265	0.489801	0.302566	0.971159	0.833668
K14 LightGBM R5	0.889169	282.295351	0.214591	0.492374	0.303511	0.969959	0.827753
K15 LightGBM R6	0.885108	288.003510	0.214031	0.493802	0.304878	0.970852	0.832580
K16 LightGBM R7	0.880160	285.829318	0.212801	0.490786	0.302951	0.971075	0.835538
K18 Train climatology	1.104035	367.548609	0.250923	0.613362	0.383556	0.922080	0.908398

The two long-horizon pools. The 57-market evaluation contains 3,150 rows and follows isolation of three markets from the original collection. The recovered records identify a pre-execution Test-access/search incident before final long-horizon reporting, but the paper assets do not reconstruct the exact query contents or every downstream choice made after that access. The remedy recorded in the analysis separates three affected markets from the 60-market complete pool and reports the remaining 57-market route as exploratory held-out evidence rather than as a confirmatory Test claim. Available summaries do not indicate that the incident retrained Dynamic42, refit calibration maps, or changed the E1/E3 value-interface selection rule; they also do not justify treating the 57-market temporal result as a clean confirmatory Test result. Raw, frozen-v2 calibrated, and training-climatology predictions have Joint scores 0.463465, 0.430062, and 0.715459, respec-

Table 17: Core14 exploratory Test: complete recorded overall metrics. K01–K05 use their frozen-calibration configuration identities. K06 is the hurdle GLM and K09 the full no-ID LightGBM hurdle-distribution baseline. All rows retain comparison seed field 11. These are not merged with the Validation table.

ID and method	Count CRPS	Revenue CRPS	No-sale Brier	Energy	Joint	Count Cov.90	Revenue Cov.90
K01 Core14 B1	0.770487	221.908734	0.211798	NR	NR	NR	NR
K02 CF size-matched	0.289599	88.863280	0.086180	NR	NR	NR	NR
K03 FULL-CF	0.244731	73.912024	0.079375	NR	NR	NR	NR
K04 Natural bootstrap	0.279158	94.124310	0.084066	NR	NR	NR	NR
K05 Natural only	0.283927	93.155750	0.084570	NR	NR	NR	NR
K06 Hurdle GLM	0.862814	278.391428	0.215984	0.487003	0.296271	0.947083	0.935556
K09 LightGBM full	0.802910	226.959401	0.215279	0.429890	0.263072	0.949491	0.828380
K17 Train climatology	0.815628	269.452893	0.212751	0.448217	0.282541	0.945463	0.932037

tively. The recorded calibrated-minus-climatology difference is -0.285397, with bootstrap interval $[-0.431972, -0.166003]$.

The complete 60-market sensitivity analysis contains 3,240 rows. Its corresponding Joint values are 0.549901, 0.526021, and 0.945608. The direction of the within-pool comparison remains favorable to the model, but revenue errors increase. The pools overlap and are not independent confirmations.

Table 18 preserves scaled and unscaled marginal CRPS. The 57-market sampling identity is 1,152 paths per row. The separate 60-market descriptor reports 1,200 draws; its underlying generation mapping was not independently recovered. The draw count and pool identities are retained separately rather than harmonized across tables. E3 uses new-seed continuing-market blueprints and official-API execution; it is not pooled with these temporal Test records, and the temporal incident is not treated as evidence for or against the E3 online effect.

Table 18: Dynamic42 long-horizon results in two distinct pools. The 57-market evaluation is exploratory after integrity isolation; the 60-market complete-pool analysis is post-hoc sensitivity. Energy uses the archived scaled target vector. Marginal raw-unit CRPS and scaled CRPS remain separate. All columns are losses.

	57 markets; 1,152 paths			60 markets; 1,200 declared draws		
Metric	**Raw**	**Cal. v2**	**Climate**	**Raw**	**Cal. v2**	**Climate**
Joint	0.463465	0.430062	0.715459	0.549901	0.526021	0.945608
Scaled Energy	0.826813	0.784574	1.336880	1.021413	1.000068	1.842942
No-sale42 Brier	0.235181	0.191476	0.259619	0.231438	0.188325	0.258040
Count28 CRPS	0.865764	0.865249	1.709812	0.876612	0.876216	1.786800
Count42 CRPS	1.550277	1.542548	2.611366	1.561242	1.554313	2.719321
Revenue28 CRPS	228.998026	224.497499	406.815023	300.743973	300.041564	602.350659
Revenue42 CRPS	427.980433	387.417012	630.771771	539.417309	516.714555	915.912156
Count28 scaled CRPS	0.173153	0.173050	0.341962	0.175322	0.175243	0.357360
Count42 scaled CRPS	0.193785	0.192818	0.326421	0.195155	0.194289	0.339915
Revenue28 scaled CRPS	0.449016	0.440191	0.797677	0.589694	0.588317	1.181080
Revenue42 scaled CRPS	0.497652	0.450485	0.733456	0.627229	0.600831	1.065014

Short-state substitutions on Validation. The frozen S1 diagnostic uses 20 Validation blueprints and 1,152 paths per row. Its full raw Joint score is 0.624773. The separately archived S1 Validation record has raw Joint 0.625508 and calibrated Joint 0.597486; the diagnostic full route instead has calibrated Joint 0.603121. These are distinct stored producer/output identities, not rounding variants of one score.

Table 19 retains the no-state, mean-only, shuffled, and observed-state substitutions. The changes concern the frozen predictor's history input rather than separately retrained alternatives. Relative to

the full raw prediction, the paired Joint changes are

$$\begin{aligned} \Delta_{\text{no}} &= 0.111117 \quad [-0.002337,\ 0.266248], \\ \Delta_{\text{mean}} &= 3.390289 \quad [2.736387,\ 4.059041], \\ \Delta_{\text{shuffle}} &= 0.081211 \quad [0.029379,\ 0.139461], \\ \Delta_{\text{oracle}} &= -0.107498 \quad [-0.187033,\ -0.016026]. \end{aligned} \tag{45}$$

The brackets are the recorded 95% intervals from 10,000 paired bootstrap replicates.

Shuffling correspondence and replacing predictions with observed states support, respectively, information use and error propagation in this interface. The severe mean-only substitution result does not establish that every predictor trained to use a mean summary must perform poorly. Observed-state replacement is diagnostic access to an unavailable future state, not a deployable forecast.

The improvement is also metric dependent. For example, the raw observed-state route improves Joint but has revenue42 CRPS 639.559125, compared with 621.680285 for the raw full route. It is therefore not described as improving every marginal outcome.

Table 19: Validation temporal predictions and frozen short-state substitutions. D01–D03 preserve the separate Archive-S1 identity. D16–D17 and D30–D37 are the Frozen-S1 diagnostic routes. All use 1,152 recorded paths; raw and v2-calibrated predictions remain separate. Energy uses the archived scaled vector; marginal CRPS columns retain their raw units.

ID and condition	Joint	Energy	No-sale Brier	Count28 CRPS	Count42 CRPS	Rev.28 CRPS	Rev.42 CRPS
D01 Archive raw	0.625508	1.203221	0.192194	1.151876	1.899980	373.509226	622.810631
D02 Archive cal.	0.597486	1.155190	0.172283	1.151367	1.894119	367.904041	577.508797
D03 Archive climate	1.042477	2.049249	0.244578	1.896997	2.863521	703.118542	1047.833956
D16 Full raw	0.624773	1.201731	0.192087	1.151962	1.898841	372.997353	621.680285
D17 Full cal.	0.603121	1.154353	0.171736	1.151427	1.892939	367.453616	576.553287
D30 No state raw	0.735890	1.464619	0.170322	1.235468	2.027448	448.550477	783.352514
D31 No state cal.	0.707647	1.401592	0.152949	1.235197	2.024234	440.847588	726.136473
D32 Mean only raw	4.015062	8.650776	0.327451	2.055883	3.856422	2278.774613	5939.520623
D33 Mean only cal.	3.313440	7.035944	0.311236	2.053057	3.832348	2083.234418	4589.931291
D34 Shuffled raw	0.705984	1.354150	0.221292	1.321627	2.229104	413.476677	702.049362
D35 Shuffled cal.	0.681316	1.300083	0.196544	1.320740	2.220738	406.988650	651.593284
D36 Oracle raw	0.517275	1.033013	0.128012	0.604250	1.364668	269.324602	639.559125
D37 Oracle cal.	0.478205	0.944402	0.120104	0.603894	1.359673	259.422271	560.256253

Coverage. Table 20 reports the available 80% and 90% coverage points for these Validation state conditions. They are not enough to reconstruct a complete calibration or PIT curve. The 57-market historical record separately reports count42 coverage near 0.728 and 0.864 at those nominal levels. Those rounded historical values are not replaced by the different Validation rows.

The mean-only route has particularly low coverage in the recorded substitution experiment. Conversely, some full-state Validation coverages exceed their nominal levels. Average loss, marginal coverage, and joint dependence therefore remain distinct aspects of forecast quality.

Direct42 comparisons. The direct baseline is an ExtraTrees-based prediction route (Geurts et al., 2006; Pedregosa et al., 2011). Its full-summary input includes short-state summaries, so it is not a baseline with no short-term information. The stored three-seed ensemble has raw Joint 0.872422, compared with 0.624773 for the full Frozen-S1 route. Individual direct seeds retain 128 draws and the ensemble 384, whereas Frozen-S1 uses 1,152 on the comparison route. The model and sampling budgets are not matched.

Table 21 retains every direct seed and condition. The full, no-state, mean-only, shuffled, and Oracle rows are not collapsed to the best condition. In particular, the direct baseline's Oracle rows use unavailable observed-state information and remain diagnostics. These results support a qualified comparison of the recorded routes, not universal superiority of temporal composition.

Table 20: Recorded Validation interval coverage for full and substituted short-state inputs. Each outcome has only the two shown nominal levels. These are coverage estimates, not confidence limits. Values are not pooled with the 57-market historical coverage summary.

	Count28		Count42		Revenue28		Revenue42	
Condition	**80%**	**90%**	**80%**	**90%**	**80%**	**90%**	**80%**	**90%**
Full raw	0.908333	0.951389	0.877778	0.931944	0.909722	0.961111	0.920833	0.959722
Full cal.	0.908333	0.951389	0.880556	0.931944	0.902778	0.962500	0.891667	0.963889
No state raw	0.829167	0.873611	0.779167	0.831944	0.813889	0.898611	0.805556	0.888889
No state cal.	0.829167	0.873611	0.781944	0.833333	0.809722	0.898611	0.768056	0.868056
Mean raw	0.586111	0.740278	0.473611	0.531944	0.541667	0.700000	0.401389	0.483333
Mean cal.	0.587500	0.740278	0.479167	0.533333	0.551389	0.713889	0.425000	0.518056
Shuffled raw	0.855556	0.905556	0.783333	0.863889	0.862500	0.933333	0.837500	0.915278
Shuffled cal.	0.855556	0.905556	0.787500	0.863889	0.858333	0.929167	0.820833	0.919444
Oracle raw	0.925000	0.973611	0.841667	0.888889	0.900000	0.965278	0.837500	0.894444
Oracle cal.	0.926389	0.973611	0.841667	0.888889	0.900000	0.963889	0.805556	0.898611

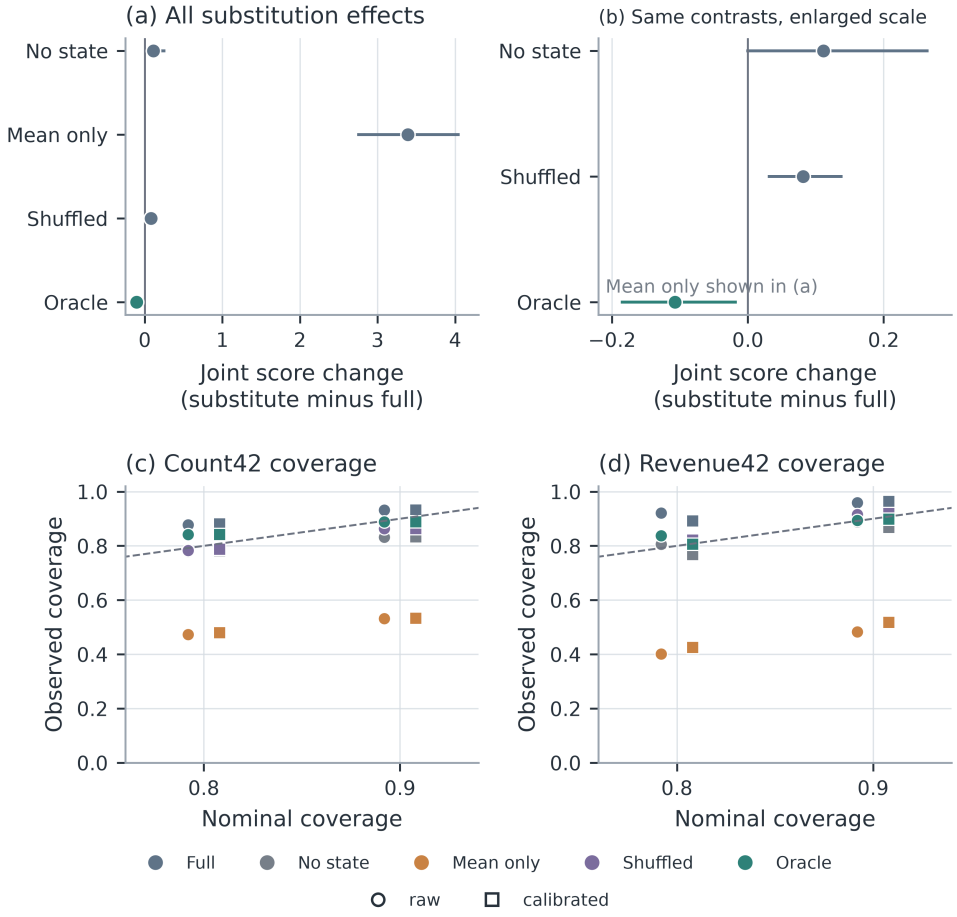

Figure C2: Short-state information and temporal coverage. Panel (a) shows all four raw Joint differences from the full Frozen-S1 route, with the archived paired 95% intervals in Eq. (45); panel (b) repeats the same contrasts on an enlarged scale and does not replace the full range. Coverage panels show the recorded 80% and 90% count42 and revenue42 points for raw and calibrated routes. The horizontal offsets in those panels are only for visibility: the true nominal levels are 0.80 and 0.90, and no curve is inferred between them. All panels use the designated Validation diagnostic rather than pooling the 57- and 60-market evaluations. Oracle states are diagnostic, not deployment inputs.

Table 21: Direct42 Validation comparisons. Each condition has seeds 11, 29, and 47, followed by the stored ensemble (Ens.). Individual seeds use 128 draws; ensembles use 384. All rows are raw predictions. The full condition uses short-state summaries and is not a no-short-state baseline. Additional recorded marginal fields retain their identities in the machine-readable temporal records.

Condition	Seed	Joint	Energy	No-sale Brier	Count42 CRPS	Revenue42 CRPS
Full summary	11	0.866958	1.724876	0.187035	2.304549	891.612678
Full summary	29	0.883041	1.757547	0.189546	2.322767	906.541023
Full summary	47	0.872236	1.734292	0.186847	2.332067	891.067260
Full summary	Ens.	0.872422	1.737762	0.186866	2.309728	892.639064
No short state	11	0.857444	1.701950	0.188715	2.329039	866.692517
No short state	29	0.866656	1.719157	0.190903	2.339699	878.250135
No short state	47	0.866850	1.723133	0.187613	2.312618	882.548990
No short state	Ens.	0.862087	1.713980	0.188349	2.316693	871.799409
Mean only	11	0.852556	1.695564	0.186764	2.289695	861.355159
Mean only	29	0.858162	1.704976	0.186468	2.317468	872.967170
Mean only	47	0.850746	1.686385	0.187804	2.331010	859.759774
Mean only	Ens.	0.853006	1.696819	0.186065	2.302484	861.304470
Shuffled	11	0.848435	1.685204	0.180748	2.322066	860.224089
Shuffled	29	0.866873	1.723344	0.186878	2.334646	877.808761
Shuffled	47	0.858506	1.704768	0.186851	2.325529	870.927846
Shuffled	Ens.	0.857030	1.705290	0.183979	2.317048	866.182818
Oracle	11	0.526427	1.023355	0.147314	1.669158	557.116572
Oracle	29	0.532866	1.042223	0.140262	1.682253	568.286124
Oracle	47	0.535532	1.044310	0.145916	1.685293	569.460359
Oracle	Ens.	0.530818	1.036786	0.143715	1.671812	562.516425

Member and fold heterogeneity. The complete predictive mixture does not choose a best member after observing Validation scores. Table 22 retains the per-S1 and individual Simulator/S1 combinations. Several constituent routes have lower Joint than the full mixture, while others have higher Joint. This prevents interpreting the mixture as uniformly superior to every component.

The grouped Train-inner S0/S1 comparison also retains all five folds (Table 23). Its recorded average Joint changes from 0.743748 to 0.698660. Fold 2 is unfavorable: Energy increases from 0.980002 to 1.007273, Joint increases by 0.002360, and revenue42 CRPS increases from 540.846576 to 583.387858. Fold identities are not successive training epochs. These rows use 768 draws and seed 20260723, not the Validation ensemble's 1,152-path identity.

Earlier development failures remain distinct from final-model results: Natural M3 V1 passed two of five folds, and an earlier PIT version improved its CvM criterion while failing the NLL guard. They are not deleted, relabeled as final-model failures, or counted as independent positive replications.

C.5 Reputation and Parameterization Ablations

Reputation variants. Table 24 defines the recorded reputation variants used by the LightGBM comparison families. Their question is whether the available representation supplies predictive information on the evaluated records. They do not intervene on a participant's reputation while holding the market's entire causal history fixed.

The reputation analysis uses 20 Validation blueprints. For losses, the reported difference is full minus variant, so a negative value favors the full representation. The R8 record is the separately archived incremental R1–R6 contrast, not an eighth trained representation.

Reputation results and source-specific intervals. For Natural, the R6 contrast is -0.025109 with archived interval $[-0.030577, -0.019762]$. The separate R8 record has the same point estimate but interval $[-0.030463, -0.019834]$. Both records are preserved in Table 25; their intervals are not reconciled by assumption.

Table 22: Frozen-S1 member diagnostics on Validation. Seed suffixes 23–25 denote 20260723–20260725. "All C" advances all three Core14 members with the indicated S1 member. Each single pair has 128 paths; each All-C route has 384. The separately reported full mixture uses all nine pairs. The table reports selected recorded diagnostic metrics without selecting a deployment member.

ID	S1 / Core14	Paths	Joint	Energy	Revenue42 CRPS
D18	23 / All C	384	0.619187	1.185182	605.042398
D19	23 / 23	128	0.624884	1.194049	615.879518
D20	23 / 24	128	0.619668	1.182667	606.569166
D21	23 / 25	128	0.625963	1.196512	616.871744
D22	24 / All C	384	0.671138	1.302753	703.871690
D23	24 / 23	128	0.675337	1.309751	712.588996
D24	24 / 24	128	0.673355	1.303434	709.660644
D25	24 / 25	128	0.680835	1.318151	719.582571
D26	25 / All C	384	0.604811	1.158804	590.534663
D27	25 / 23	128	0.610049	1.166242	599.922352
D28	25 / 24	128	0.605521	1.156549	593.991947
D29	25 / 25	128	0.613108	1.173353	603.408923

Table 23: Grouped Train-inner S0/S1 comparisons. Individual fold rows use seed 20260723 and 768 draws. Mean rows are the stored fold summaries, not predictions from an additional fitted ensemble. Joint, Energy, and the two displayed 42-day marginals retain separate scales. Fold 2 is retained despite being unfavorable to S1 on several criteria.

Fold	Model	Joint	Energy	No-sale Brier	Count42 CRPS	Revenue42 CRPS
0	S0	0.907717	1.799225	0.228442	2.158009	931.712995
0	S1	0.845759	1.689938	0.192865	2.041450	855.259650
1	S0	0.724992	1.403374	0.217600	2.072463	721.547252
1	S1	0.677328	1.322729	0.184870	1.955495	678.232568
2	S0	0.528440	0.980002	0.220968	1.828630	540.846576
2	S1	0.530800	1.007273	0.192358	1.675696	583.387858
3	S0	0.974346	1.966785	0.198344	2.168113	957.663334
3	S1	0.885181	1.786345	0.175383	2.012495	843.758072
4	S0	0.583248	1.072174	0.249770	2.048371	582.769047
4	S1	0.554234	1.030790	0.219977	1.968528	554.773018
Mean	S0	0.743748	1.444312	0.223025	2.055117	746.907841
Mean	S1	0.698660	1.367415	0.193090	1.930733	703.082233

Table 24: Reputation-variant dictionary. R1–R7 identify recorded feature configurations. R8 identifies a separate incremental contrast. The composite representation and the different removals are not combined into a single generic "no reputation" condition.

Code	Recorded meaning
R1	Full representation; no persistent entity IDs
R2	Composite reputation representation
R3	No coverage features
R4	No product-reputation features
R5	No seller-reputation features
R6	No reputation features
R7	Stratified reputation shuffle
R8	Separately archived incremental R1–R6 contrast; not an additional model

The completion reputation table has full-model score 0.064811, whereas the task-baseline table reports 0.064854 for its full Platt-calibrated LightGBM row. These retain separate source-table identities. No unverified rounding or preprocessing explanation is supplied.

The independent reputation contribution is not uniformly established across tasks. Completion and Core14 full-minus-no-reputation intervals include zero, and their reported Holm-adjusted comparisons do not reject. Some other unadjusted intervals exclude zero without passing their stored Holm criterion. The Natural result therefore does not justify a general statement that every market predictor benefits from every reputation component.

Table 25: Recorded reputation diagnostics on Validation. Differences are full minus variant. Intervals are unadjusted archived intervals; Holm probabilities retain the corresponding source family. For R8, the "Score" field is itself the recorded incremental contrast, not an absolute model loss. $0^{\dagger}$ preserves a source display value of zero; it is not a claim that the mathematical probability is exactly zero. R1 self-comparisons and their degenerate intervals are reference rows, not evidence about statistical uncertainty.

Task	ID	Score	Difference	CI low	CI high	Holm p
Natural	R1	0.513164	0.000000	0.000000	0.000000	1.000000
Natural	R2	0.514216	-0.001052	-0.004318	0.002111	1.000000
Natural	R3	0.513320	-0.000157	-0.001613	0.001256	1.000000
Natural	R4	0.519163	-0.005999	-0.009430	-0.002759	0.002400
Natural	R5	0.512964	0.000200	-0.001851	0.002126	1.000000
Natural	R6	0.538273	-0.025109	-0.030577	-0.019762	$0^{\dagger}$
Natural	R7	0.537098	-0.023934	-0.029153	-0.018778	$0^{\dagger}$
Natural	R8	-0.025109	-0.025109	-0.030463	-0.019834	$0^{\dagger}$
Deal	R1	0.064811	0.000000	0.000000	0.000000	1.000000
Deal	R2	0.064787	2.44×10^{-5}	-3.73×10^{-5}	0.000102	1.000000
Deal	R3	0.064871	-6.00×10^{-5}	-0.000116	-4.45×10^{-6}	0.170000
Deal	R4	0.064959	-0.000148	-0.000242	-3.82×10^{-5}	0.063600
Deal	R5	0.064959	-0.000148	-0.000242	-4.10×10^{-5}	0.057400
Deal	R6	0.064847	-3.53×10^{-5}	-0.000110	4.70×10^{-5}	1.000000
Deal	R7	0.064811	0.000000	0.000000	0.000000	1.000000
Deal	R8	-3.53×10^{-5}	-3.53×10^{-5}	-0.000109	4.93×10^{-5}	1.000000
Core14	R1	0.495361	0.000000	0.000000	0.000000	1.000000
Core14	R2	0.496082	-0.000722	-0.010210	0.008112	1.000000
Core14	R3	0.493109	0.002252	-0.004148	0.007932	1.000000
Core14	R4	0.489801	0.005560	0.000982	0.009837	0.148400
Core14	R5	0.492374	0.002986	-0.000322	0.006375	0.464400
Core14	R6	0.493802	0.001559	-0.009596	0.012000	1.000000
Core14	R7	0.490786	0.004575	-0.006024	0.014250	1.000000
Core14	R8	0.001559	0.001559	-0.009336	0.012068	1.000000

Common-feature parameterization study. The structure experiment compares independent experts, a shared encoder with task heads, and a single unified model under a common 21-dimensional input. Its recorded training budget is 260 AdamW steps, learning rate 10^{-3}, and weight decay 10^{-4}. These are not the formal probabilistic-model budgets in Appendix B.

Natural uses candidate binary training with listwise evaluation. Other targets use their recorded standardized point-loss construction. The all-task parameter counts are 8,965 for independent experts, 5,893 for the shared encoder, and 5,953 for the unified model. Thus the experiment is not parameter matched, and its point-prediction scores are not a replacement for the distributional comparisons above. The "size-matched" regime refers to the recorded data/training condition, not equal parameter counts.

Table 26 specifies all 24 identities. Tables 27 and 28 preserve the two score blocks. The full-CF rows S13–S15 coincide with S04–S06 in the stored metric blocks. They retain their recorded regime names but are not counted as new independent evidence for the same score pattern.

Table 26: Dictionary of all recorded 21-dimensional structure runs. Within each listed three-ID range, IDs correspond in order to seeds 11, 29, and 47. Parameter counts are per recorded configuration, not a matched budget across architectures.

IDs	Recorded regime	Architecture	Parameters
S01–S03	All tasks, common public subset	Independent experts	8,965
S04–S06	All tasks, common public subset	Shared encoder + task heads	5,893
S07–S09	All tasks, common public subset	Single unified model	5,953
S10–S12	Controlled only	Shared encoder + task heads	5,633
S13–S15	Natural–Controlled full CF	Shared encoder + task heads	5,893
S16–S18	Natural–Controlled size-matched	Shared encoder + task heads	5,698
S19–S21	Natural only	Shared encoder + task heads	5,633
S22–S24	Natural plus Controlled	Shared encoder + task heads	5,698

Table 27: Complete archived structure-score block for choice, completion, and Technical MAE. IDs resolve to Table 26. NR in this table denotes a target absent from the recorded regime. These are task-specific point estimates; no cross-task total or new significance test is constructed.

ID	Natural NLL	Natural Top-1	Deal Brier	Deal log loss	Technical MAE
S01	0.849015	0.640120	0.134092	0.440114	7.579183
S02	0.833261	0.660183	0.134123	0.440362	7.571718
S03	0.821012	0.658208	0.133561	0.438359	7.498254
S04	0.783727	0.679592	0.133908	0.439623	7.848859
S05	0.785049	0.678182	0.133633	0.438174	7.485657
S06	0.795335	0.680535	0.134421	0.441039	7.630033
S07	0.816021	0.677946	0.136608	0.448760	7.645599
S08	0.834187	0.669670	0.135561	0.445711	7.493796
S09	0.820003	0.658128	0.136461	0.448123	7.456047
S10	NR	NR	0.133985	0.440703	NR
S11	NR	NR	0.134182	0.441147	NR
S12	NR	NR	0.134007	0.440150	NR
S13	0.783727	0.679592	0.133908	0.439623	7.848859
S14	0.785049	0.678182	0.133633	0.438174	7.485657
S15	0.795335	0.680535	0.134421	0.441039	7.630033
S16	0.825283	0.663782	0.133711	0.439295	NR
S17	0.823134	0.660721	0.133890	0.439150	NR
S18	0.807480	0.674882	0.133712	0.438737	NR
S19	0.839679	0.645929	NR	NR	NR
S20	0.843356	0.640783	NR	NR	NR
S21	0.829104	0.661436	NR	NR	NR
S22	0.820417	0.672241	0.133508	0.438291	NR
S23	0.827418	0.659401	0.133691	0.438491	NR
S24	0.798324	0.684952	0.133858	0.439597	NR

The structure results describe task-dependent sharing trade-offs. For the recorded all-task seeds, the shared encoder has lower Natural NLL than the independent experts, but completion and Technical errors do not improve uniformly. The unified model is likewise not uniformly better. Unequal parameter counts, distinct training regimes, and point-target losses prevent promoting this study into a general architectural superiority claim for the complete DIM-WM probability interfaces.

Across the appendix, the strongest supported interpretation is therefore target- and condition-specific predictive information. The availability of a favorable loss, calibration summary, or feature contrast is not substituted for the supported-decision and online service evaluations that follow.

D Fixed-Support Decisions and Value-Interface Evaluation

This appendix evaluates three uses of the same recorded action-outcome library. The historical analysis compares numerical objectives and heuristics on fixed legal support. E1 replaces the gain predictor within a fixed gain–price rule and selects an external value producer for development. E2

Table 28: Recorded scalar-error block for the structure experiment. S10–S12 and S16–S24 have no applicable reported targets in this block and are omitted here rather than displayed as zero. Their identities and available choice/completion results remain in Tables 26 and 27. Units and normalization follow the recorded common-feature study, not the main probabilistic-score tables.

ID	Technical RMSE	Count MAE	Count RMSE	Revenue MAE	Revenue RMSE
S01	9.676520	0.564082	0.699585	2.356510	2.778749
S02	9.785856	0.562519	0.697560	2.342823	2.795650
S03	9.620128	0.572358	0.698510	2.314229	2.773947
S04	10.025536	0.566202	0.699663	2.323747	2.786165
S05	9.525542	0.558506	0.691626	2.316228	2.765308
S06	9.698530	0.565974	0.695775	2.347034	2.786232
S07	9.823232	0.579730	0.699510	2.373346	2.799828
S08	9.714054	0.585747	0.696567	2.401307	2.805818
S09	9.663116	0.577457	0.699865	2.361755	2.814660
S13	10.025536	0.566202	0.699663	2.323747	2.786165
S14	9.525542	0.558506	0.691626	2.316228	2.765308
S15	9.698530	0.565974	0.695775	2.347034	2.786232

supplies different information to an advisor and participant, generates new choices, and evaluates the selected actions against the same archived outcomes. These analyses distinguish action scoring from participant-controlled information use, but do not constitute independent dynamic-market evaluations.

D.1 Archived Action Support and Outcome Evaluation

Recorded states and legal actions. The library contains 239 pre-action parent states belonging to 60 blueprints. Across these states, 3,830 legal parent–action pairs each have 30 recorded rollouts, giving 114,900 outcome rows. A state determines its own legal support $\mathcal{A}(S)$. An action outside this support has no measured value in this evaluation; it is not assigned an interpolated, extrapolated, or zero-valued outcome.

Ordinary policies select actions using their permitted public information and frozen predictions. The selected action is then joined to the corresponding outcome records for evaluation. The outcome lookup does not supply the policy's ordinary action score. The Oracle diagnostics in Section D.2 are explicitly different because they use a designated selection subset of outcomes.

Reanalysis integrity. The later fixed-support audit reconnects the archived policy choices and rollout records. It reports 10,516 policy-action records, no duplicate rollout row identifiers, and no selected actions missing from the outcome library. The policy-action count includes the recorded policy and evaluation identities; it is not the number of independent parent states.

The audit also reports a maximum absolute discrepancy of zero for the stored accounting identity

$$\texttt{net_benefit} = \texttt{technical_gain} - \texttt{paid_price_tokens}.$$

These stored quantities are realized outcomes, with non-transaction gain and payment recorded as zero. They are consistent with the distinction in Eq. (1) between a potential buyer–product gain and a realized transaction surplus.

This audit is a supplementary reanalysis of archived fixed support. It does not generate new market trajectories, repair all historical prediction mappings, or independently validate the simulated participants.

Blueprint aggregation. Let $Y_b(\pi)$ denote the outcome of policy π after the recorded within-blueprint aggregation of selected parent/action rollouts and, where applicable, evaluation repeats. The reported mean and a paired policy contrast have the form

$$\widehat{Y}(\pi) = \frac{1}{60}\sum_{b=1}^{60} Y_b(\pi), \qquad \widehat{\Delta}(\pi, \pi_0) = \frac{1}{60}\sum_{b=1}^{60} \left[Y_b(\pi) - Y_b(\pi_0)\right]. \tag{46}$$

The notation preserves the existing within-blueprint aggregation rather than assigning equal weight to every raw row. The 30 rollouts of one action are not 30 independent markets. Nor are repeated policy evaluations or Oracle splits independent blueprints.

Coverage and negative outcomes. For the selected realized rollouts, let v_r denote the normalized weights induced by the recorded aggregation, with $\sum_r v_r = 1$. Transaction coverage and negative-net frequency are

$$\widehat{C}(\pi) = \sum_r v_r D_r, \qquad \widehat{N}(\pi) = \sum_r v_r \mathbf{1}\{U_r < 0\}. \tag{47}$$

Non-transactions contribute $D_r = 0$ and $U_r = 0$. Consequently, a lower negative-net frequency can result from reduced trading, not only from better selection among completed trades.

The later reanalysis and E1 additionally report a grouped `conditional_negative_rate`. For the selected parent–action cell q, let

$$e_q = \sum_{r \in q} D_r, \qquad c_q = \frac{\sum_{r \in q} \mathbf{1}\{D_r = 1,\ U_r < 0\}}{e_q}$$

when $e_q > 0$. Cells with no completed transactions have undefined c_q and are omitted from this conditional statistic rather than assigned zero. The reported E1 value first averages c_q over eligible selected parent–action cells within a blueprint and then averages the non-missing blueprint values. It is therefore not reconstructed by dividing the published average negative frequency by the published average deal rate. The same grouped statistic is used by the E1 development selection rule in Section D.3; historical rows without this reconstructed field retain their recorded summaries.

Realized lower-tail summaries. The historical outcome field `cvar10` uses a finite realized sample convention. Within a policy, evaluation repeat, parent, and selected action, order the n observed outcomes as $U_{(1)} \le \cdots \le U_{(n)}$ and compute

$$\widehat{T}_{0.10} = \frac{1}{\lceil 0.10n \rceil} \sum_{j=1}^{\lceil 0.10n \rceil} U_{(j)}. \tag{48}$$

Here $n = 30$ for ordinary policy evaluation and $n = 15$ for an Oracle evaluation half. The resulting parent/repeat summaries are aggregated within blueprint and then across the 60 blueprints.

This statistic includes non-transaction zeros. It is not the predictive weighted-atom calculation in Eq. (42), which takes exactly the specified probability mass and permits a fractional boundary atom. It is also not the lower-tail mean of one pooled distribution across all states and markets.

Action regret and false rejection retain the archived evaluation reference and grouping. They are not recomputed from the displayed policy means. In particular, the sample-split Oracle is not substituted as a common scalar benchmark from which every regret value is obtained.

D.2 Historical Policies, Risk, and Oracle Diagnostics

Policy identities and action mappings. Table 29 retains all 28 historical policy names. Natural-choice imitation ranks products by predicted choice probability, compares rejection separately, and maps a chosen product to `ACCEPT_ASK` when available or otherwise to its lowest supported offer. Gain-only and reputation-based product rankings use the same product-to-action mapping. This is a defined product-ranking comparator, not a complete replay of the original participant's negotiation policy. Its difference from Gain–Price therefore includes action mapping and rejection logic, not only predictive information quality.

Technical Gain–Price instead ranks actions using the raw finite Technical gain summary minus the supported offer, as in Eq. (4). It does not subtract an observed future payment. Because the gain term is product-level, two legal actions for the same product are ranked only by their supported offers; the rule does not trade off a lower offer against a lower predicted completion probability. Expected Net Benefit uses the calibrated empirical gain–payment construction in Eq. (39). The default DIM-WM planner applies the lower-tail penalty in Eq. (5).

The positive-score rejection gate applies to the utility-scored rules. It is not imposed retrospectively on every heuristic: Natural compares rejection probability, while the product-ranking and Cheapest policies retain their recorded selection rules. Action identifiers resolve ties.

Table 29: Historical policy dictionary. Names identify the archived rows rather than 28 independent experiments. P05 is the default risk planner with $\lambda = 0.50$. P06 and P13 name the same expected-net objective. P12 and P14 are explicit Oracle aliases. Other coincident outcome rows remain distinct named settings without being counted as independent replications.

ID	Archived policy	ID	Archived policy
P01	Always Reject / Conservative	P15	Highest Composite Reputation
P02	CVaR05 maximization	P16	Highest Predicted Technical Gain
P03	CVaR10 maximization	P17	Highest Product Reputation
P04	Cheapest	P18	Highest Seller Reputation
P05	DIM-WM Deterministic Planner, $\lambda = 0.50$	P19	Natural Choice imitation
P06	DIM-WM Planner, $\lambda = 0.00$	P20	Negative Net Probability Constraint 0.10
P07	DIM-WM Planner, $\lambda = 0.25$	P21	Negative Net Probability Constraint 0.25
P08	DIM-WM Planner, $\lambda = 1.00$	P22	Negative Net Probability Constraint 0.40
P09	Deal Probability Constraint 0.25	P23	Oracle Technical + Predicted Controlled
P10	Deal Probability Constraint 0.50	P24	Predicted Technical + Oracle Controlled
P11	Deal Probability Constraint 0.75	P25	Reputation Constraint 0.40
P12	Empirical Oracle	P26	Reputation Constraint 0.50
P13	Expected Net Benefit	P27	Reputation Constraint 0.60
P14	Fully Oracle	P28	Technical Gain–Price

Discrete risk and constraint settings. The evaluated risk weights are $\lambda \in \{0, 0.25, 0.50, 1.00\}$. At $\lambda = 0$, the risk objective is Expected Net Benefit. Connecting these settings in a figure indicates their parameter order, not a continuous efficient frontier, a monotonicity guarantee, or a newly optimized hyperparameter.

Constraint variants maximize expected net benefit subject to their recorded completion-probability, negative-net, or reputation tests. For a saved prediction row, the negative-event field is

$$\widehat{p}_{-} = \frac{q}{576} \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{64} \mathbf{1}\left\{ \widetilde{G}_{ijk} - \widetilde{P}_{ijk} < 0 \right\}. \tag{49}$$

The corresponding prediction fields are subsequently averaged into the action table across its recorded prediction rows. This is not a product of independently averaged components, nor one minus the separately calibrated positive-event output. Non-completion and completed zero-surplus outcomes contribute zero mass.

The three completion thresholds give identical displayed outcomes on this support. This does not establish that the constraint is inactive in every possible market. Similarly, matching displayed numbers for other settings do not by themselves establish identical actions or independent replication.

Complete outcome summaries. Table 30 preserves all historical outcome rows, including unfavorable heuristics. Natural imitation has surplus 0.541153, Expected Net Benefit 2.055032, and Technical Gain–Price 2.368721. Cheapest and Highest Seller Reputation have negative mean surplus, -1.148779 and -0.968400, respectively. Ranking a product by one public characteristic therefore does not uniformly produce high realized buyer surplus.

The default risk planner yields surplus 1.476307, deal rate 0.143611, and negative-net frequency 0.004167. Its small negative frequency accompanies substantially lower coverage than Gain–Price or Natural imitation. The later grouped conditional negative-rate summaries are reported separately rather than added retrospectively to every historical row.

The historical P01 row reports regret 3.570162. The later compact reanalysis summary leaves the Always Reject regret field unreported. The historical value is retained here under its original identity; the later summary's missing field remains missing under its own record identity.

Absolute means and paired differences. Table 31 distinguishes an interval for a policy's absolute mean from an interval for its paired difference from Natural imitation. Expected Net Benefit minus

Table 30: All archived fixed-support policy outcomes. Each outcome is aggregated under its recorded protocol and equally across 60 blueprints. Negative rate includes all selected rollouts. Revenue is seller gross revenue, not profit. The `cvar10` column uses the realized convention in Eq. (48). Oracle rows retain their sample-split evaluation identity.

ID	Net surplus	Regret	False reject	Negative rate	Deal rate	Revenue	cvar10
P01	0.000000	3.570162	0.756944	0.000000	0.000000	0.000000	0.000000
P02	0.566741	3.003421	0.727778	0.000000	0.029167	0.064625	0.566741
P03	0.902554	2.667608	0.706944	0.000000	0.050000	0.116833	0.902554
P04	-1.148779	4.718941	0.000000	0.508565	0.857130	1.472535	-2.930129
P05	1.476307	2.093855	0.611111	0.004167	0.143611	0.253763	1.381480
P06	2.055032	1.515130	0.419444	0.024167	0.331806	0.478281	1.728369
P07	1.789418	1.780744	0.531944	0.011250	0.218333	0.368344	1.571972
P08	1.222965	2.347197	0.652778	0.004167	0.102361	0.196689	1.169316
P09	2.055032	1.515130	0.419444	0.024167	0.331806	0.478281	1.728369
P10	2.055032	1.515130	0.419444	0.024167	0.331806	0.478281	1.728369
P11	2.055032	1.515130	0.419444	0.024167	0.331806	0.478281	1.728369
P12	3.511807	0.072461	0.005000	0.013944	0.733389	1.126699	2.973877
P13	2.055032	1.515130	0.419444	0.024167	0.331806	0.478281	1.728369
P14	3.511807	0.072461	0.005000	0.013944	0.733389	1.126699	2.973877
P15	0.812306	2.757856	0.000000	0.427778	1.000000	2.803542	0.812306
P16	1.147343	2.422819	0.000000	0.402778	1.000000	2.796667	1.147343
P17	0.573094	2.997068	0.000000	0.427778	1.000000	2.800042	0.573094
P18	-0.968400	4.538562	0.000000	0.552778	1.000000	2.783972	-0.968400
P19	0.541153	3.029008	0.000000	0.461111	1.000000	2.792528	0.541153
P20	1.151464	2.418698	0.669444	0.004167	0.087500	0.192247	1.148919
P21	1.658048	1.912114	0.590278	0.008333	0.163194	0.301097	1.508146
P22	2.041369	1.528793	0.452778	0.020278	0.294583	0.471018	1.715674
P23	3.481227	0.103041	0.019167	0.009167	0.710222	1.151457	3.236247
P24	2.355798	1.228471	0.243056	0.084833	0.551278	0.697076	2.029583
P25	2.055032	1.515130	0.419444	0.024167	0.331806	0.478281	1.728369
P26	2.061189	1.508973	0.423611	0.020000	0.323611	0.464332	1.741393
P27	1.746989	1.823173	0.481944	0.011667	0.273056	0.332971	1.622681
P28	2.368721	1.201441	0.230556	0.087778	0.543056	0.685175	1.560148

Natural is 1.513879, with interval $[0.851714, 2.251478]$. Gain–Price minus Natural is 1.827567, with interval $[1.187292, 2.524859]$.

The default risk planner has a positive absolute mean and a positive unadjusted difference interval, but its net-surplus contrast does not pass the archived Holm criterion ($p_{\mathrm{Holm}} = 0.0624$). Overlap of separate absolute intervals does not replace the paired comparison. Likewise, visualization does not create a new significance test.

The original interval archive contains 220 absolute or paired records across net surplus, action regret, negative-net frequency, and realized `cvar10`. The net-surplus records are reproduced below; selected secondary contrasts are shown in Table 32. The remaining archived interval records retain their source definitions and are not recomputed from these tables. Each metric retains its own recorded correction family. None borrows the historical online eight-test family or the later format-control family.

The source interval archive contains the complete machine-readable records, including raw probabilities and metric-specific families. The table preserves source-displayed zeros rather than constructing unreported positive probabilities from them.

Component-replacement Oracle diagnostics. Five fixed splits separate each action's 30 outcomes into 15 selection outcomes and 15 evaluation outcomes. A partial Oracle substitutes the selection half's realized component summary while retaining the other component's frozen prediction. The chosen action is evaluated using the other half. No predictive component is retrained.

With Oracle Technical and predicted Controlled, mean surplus is 3.481227. With predicted Technical and Oracle Controlled, it is 2.355798. The full empirical Oracle yields 3.511807. This asym-

Table 31: All archived net-surplus intervals. Policy means are listed in Table 30. Δ is policy minus Natural imitation. Intervals and Holm probabilities retain their historical analysis identities. $0^{\dagger}$ denotes a source-displayed zero probability, not an assertion of a mathematically zero probability. A dash marks the inapplicable reference-to-itself contrast.

ID	Absolute 95% CI	Δ	Paired 95% CI	Holm p
P01	$[0.000000, 0.000000]$	-0.541153	$[-1.452277, 0.407347]$	1.000000
P02	$[0.128650, 1.114315]$	0.025588	$[-0.764283, 0.848516]$	1.000000
P03	$[0.322378, 1.593718]$	0.361400	$[-0.417133, 1.200895]$	1.000000
P04	$[-2.082808, -0.250795]$	-1.689932	$[-2.389013, -0.959071]$	$0^{\dagger}$
P05	$[0.855349, 2.178379]$	0.935153	$[0.237820, 1.690443]$	0.062400
P06	$[1.406577, 2.775250]$	1.513879	$[0.851714, 2.251478]$	$0^{\dagger}$
P07	$[1.153126, 2.498794]$	1.248265	$[0.559929, 2.007047]$	0.004800
P08	$[0.613277, 1.919278]$	0.681812	$[-0.069727, 1.477044]$	0.491400
P09	$[1.406577, 2.775250]$	1.513879	$[0.851714, 2.251478]$	$0^{\dagger}$
P10	$[1.406577, 2.775250]$	1.513879	$[0.851714, 2.251478]$	$0^{\dagger}$
P11	$[1.406577, 2.775250]$	1.513879	$[0.851714, 2.251478]$	$0^{\dagger}$
P12	$[2.831247, 4.257991]$	2.970654	$[2.388181, 3.622778]$	$0^{\dagger}$
P13	$[1.406577, 2.775250]$	1.513879	$[0.851714, 2.251478]$	$0^{\dagger}$
P14	$[2.831247, 4.257991]$	2.970654	$[2.388181, 3.622778]$	$0^{\dagger}$
P15	$[-0.325566, 1.876003]$	0.271152	$[-0.299999, 0.881809]$	1.000000
P16	$[0.156576, 2.104049]$	0.606189	$[0.161316, 1.117570]$	0.037800
P17	$[-0.634287, 1.706215]$	0.031940	$[-0.541184, 0.601850]$	1.000000
P18	$[-2.180560, 0.146563]$	-1.509553	$[-2.261381, -0.793091]$	$0^{\dagger}$
P19	$[-0.407347, 1.452277]$	—	—	—
P20	$[0.560371, 1.838675]$	0.610310	$[-0.147955, 1.423714]$	0.654000
P21	$[1.008979, 2.366583]$	1.116894	$[0.436726, 1.872025]$	0.010000
P22	$[1.393436, 2.758847]$	1.500215	$[0.836858, 2.238630]$	$0^{\dagger}$
P23	$[2.801908, 4.227982]$	2.940074	$[2.350403, 3.602390]$	$0^{\dagger}$
P24	$[1.697307, 3.059449]$	1.814644	$[1.135367, 2.549539]$	$0^{\dagger}$
P25	$[1.406577, 2.775250]$	1.513879	$[0.851714, 2.251478]$	$0^{\dagger}$
P26	$[1.404871, 2.778109]$	1.520036	$[0.851416, 2.257900]$	$0^{\dagger}$
P27	$[1.133964, 2.448365]$	1.205835	$[0.509834, 1.970634]$	0.004800
P28	$[1.725777, 3.075017]$	1.827567	$[1.187292, 2.524859]$	$0^{\dagger}$

metry indicates that Technical estimation is an important constraint for the recorded states, legal actions, and component substitution procedure.

The separation of selection and evaluation avoids reporting the same finite-sample outcome used to select an action. It does not make the substituted information prospectively available, guarantee a pointwise upper bound on every evaluation sample, or identify the best action outside the library. The five splits reuse the same recorded outcomes and are not five independent market samples.

Figure D1 combines the discrete risk settings and the component-replacement summaries. Ordinary policies use 30 outcomes per selected action, whereas Oracle values use their held-out 15-outcome evaluation halves; the figure retains this distinction.

D.3 E1: Predictor Replacement and Development Selection

Question and unchanged support. E1 asks whether an alternative Technical predictor changes decision quality when the action-selection rule is held fixed. It reuses the 239 parent states, 3,830 legal actions, and 114,900 archived outcomes. No new LLM interaction or market rollout is generated for E1.

The principal comparison changes the gain summary in Eq. (4) among the frozen Technical predictor, quantile LightGBM, and linear Student-t. Legal actions, visible offers, the scoring rule, and the evaluation outcome library are unchanged. The historical Expected Net Benefit route is retained as a separate decision-rule reference, not treated as another gain–price predictor replacement.

The corresponding predictive-quality comparisons are in Appendix C. A better CRPS is not used as a substitute for evaluating the selected actions.

Table 32: Selected secondary paired contrasts against Natural imitation. The selection displays the main risk, expected-value, gain–price, and component-Oracle comparisons discussed in the text. It does not redefine the original correction families. Negative differences favor a policy for regret and negative-net frequency; positive differences favor it for the realized lower tail. $0^{\dagger}$ retains a source-displayed zero.

Metric	Policy	Difference	Paired 95% CI	Holm p
Regret	P05	-0.935153	$[-1.691150, -0.225831]$	0.070400
Regret	P13	-1.513879	$[-2.240281, -0.837186]$	$0^{\dagger}$
Regret	P28	-1.827567	$[-2.522233, -1.184495]$	$0^{\dagger}$
Regret	P23	-2.925967	$[-3.587673, -2.324507]$	$0^{\dagger}$
Regret	P24	-1.800538	$[-2.524492, -1.117783]$	$0^{\dagger}$
Regret	P12	-2.956548	$[-3.616625, -2.359479]$	$0^{\dagger}$
Negative rate	P05	-0.456944	$[-0.513889, -0.400000]$	$0^{\dagger}$
Negative rate	P13	-0.436944	$[-0.494306, -0.381667]$	$0^{\dagger}$
Negative rate	P28	-0.373333	$[-0.431253, -0.316806]$	$0^{\dagger}$
Negative rate	P23	-0.451944	$[-0.510000, -0.395000]$	$0^{\dagger}$
Negative rate	P24	-0.376278	$[-0.437335, -0.316832]$	$0^{\dagger}$
Negative rate	P12	-0.447167	$[-0.503779, -0.392164]$	$0^{\dagger}$
cvar10	P05	0.840327	$[0.129415, 1.609288]$	0.145600
cvar10	P13	1.187215	$[0.489563, 1.936405]$	0.012600
cvar10	P28	1.018994	$[0.301110, 1.782418]$	0.048400
cvar10	P23	2.695093	$[2.088419, 3.358754]$	$0^{\dagger}$
cvar10	P24	1.488429	$[0.784622, 2.241001]$	$0^{\dagger}$
cvar10	P12	2.432723	$[1.828901, 3.107870]$	$0^{\dagger}$

Recorded decision outcomes. Table 33 reports every completed E1 row. Quantile LightGBM increases the numerical mean surplus from 2.368721 to 2.673447 relative to frozen Technical under Gain–Price. Its deal rate increases from 0.543056 to 0.648287, false rejection decreases from 0.230556 to 0.125000, and the reported grouped conditional negative rate is 0.161077 rather than 0.166761.

The linear Student-t alternative has a higher deal rate, 0.826435, but lower surplus, 2.218891, and higher conditional negative rate, 0.312232. Consequently, more frequent trading does not imply greater net value in this comparison. Although this linear predictor has lower Validation CRPS than frozen Technical, its mean Gain–Price decision outcome is lower. The predictive and decision rankings are therefore not identical.

These are development point estimates. The E1 analysis does not provide new paired intervals for the predictor-replacement contrasts, and no such intervals or significance claims are inferred from the reported means.

Table 33: Completed E1 comparisons on archived fixed support. GP denotes Gain–Price; ENB denotes Expected Net Benefit. The conditional-negative column is the separately reported grouped conditional statistic, not negative frequency divided by average deal rate. Values are development point estimates without reported uncertainty intervals.

Technical / prediction source	Rule	Net surplus	Deal rate	Conditional negative rate	False rejection	Regret
Current frozen Technical	GP	2.368721	0.543056	0.166761	0.230556	1.201441
Quantile LightGBM	GP	2.673447	0.648287	0.161077	0.125000	0.896714
Linear Student-t	GP	2.218891	0.826435	0.312232	0.025000	1.351271
Frozen Technical + historical Controlled	ENB	2.055032	0.331806	0.100680	0.419444	1.515130

Unavailable comparisons. Two planned rows were not completed. The completion-weighted Controlled Expected Net route is marked `BLOCKED_RAW_QUERY_OUTPUTS`, because the per-member, per-arrival completion and conditional-payment predictions were not recovered. The direct-net out-of-fold comparator is marked `BLOCKED_ASSET_NOT_FOUND`. Their scores remain unavailable, rather than zero or an inferred failure to outperform a baseline.

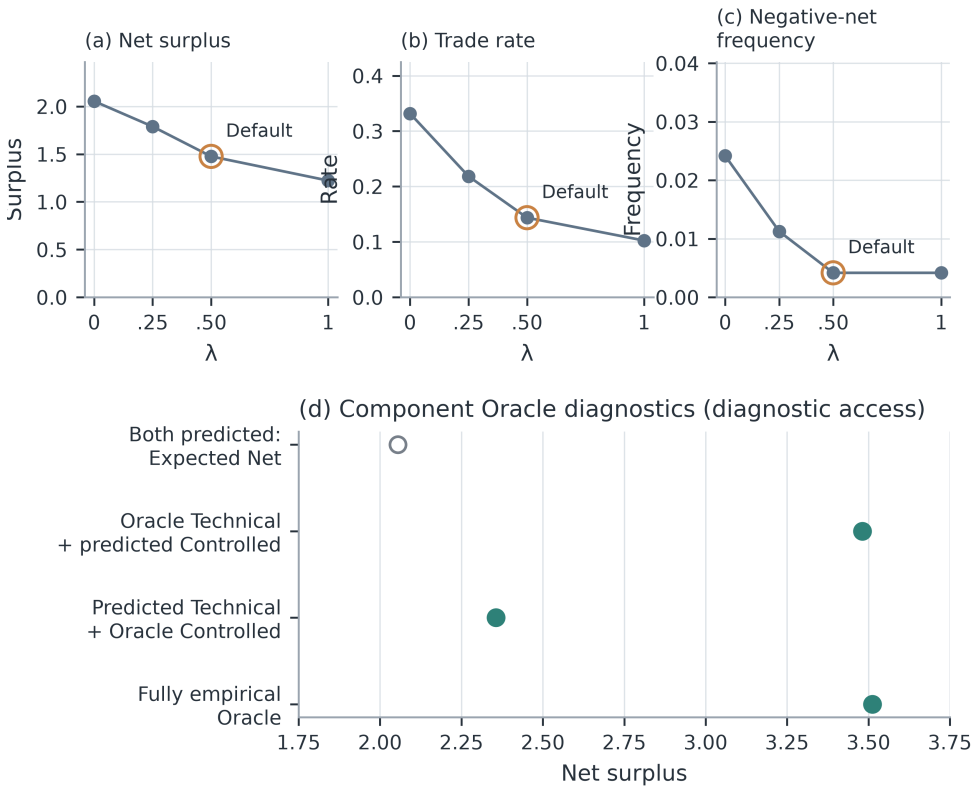

Figure D1: Historical fixed-support risk and Oracle diagnostics. Risk panels show realized net surplus, deal rate, and negative-net frequency at the four evaluated weights $\lambda \in \{0, 0.25, 0.50, 1.00\}$. Negative-net frequency is computed over selected realized rollouts; nontransactions contribute zero net surplus and therefore are included as zero, not as negative outcomes. Connecting these points indicates discrete parameter order, not a continuous optimal frontier. The component panel shows Oracle-Technical, Oracle-Controlled, and full empirical-Oracle outcomes under five fixed 15/15 selection/evaluation splits. All results retain blueprint-level aggregation. Oracle access is diagnostic; marker styling in that panel does not denote a significance test, and ordinary-policy and Oracle evaluation sample conventions are not conflated.

The conditional-mixture correction in Eq. (25) was formulated and tested, but not followed by complete regeneration and calibration of the affected Controlled outputs. E1 therefore does not establish the result of a fully corrected Expected Net model. Using Gain–Price avoids dependence on that conditional-payment mixture; it does not repair the historical Controlled implementation.

Development selection. The completed E1 report records selection of quantile LightGBM with Gain–Price under the frozen development rule. This identifies the external value producer used by E2 and E3. It does not replace the frozen Technical summaries already embedded in Natural, or regenerate the historical online advice treatment.

Selection uses the same development support on which the E1 outcomes are reported. The selected predictor's performance is therefore not an independent post-selection confirmation. Neither the selection nor the reported point estimates establish universal superiority across risk preferences or market mechanisms. The frozen protocol specifies the selection predicate. Eligible candidates are the completed, unblocked Gain–Price rows: the original frozen Technical predictor, quantile LightGBM, and linear Student-t. Expected Net rows are references, and the blocked corrected-Controlled and direct-net rows are not eligible. Let $c_0 = 0.1667608286$ denote the original Gain–Price grouped conditional negative rate defined in Section D.1. The rule sets the admissible threshold to $c_0 + 0.02 = 0.1867608286$, keeps candidates with conditional negative rate no larger than that threshold, and selects the remaining candidate with the largest mean net surplus. If no candidate

passes the threshold, the fallback is to retain the original Gain–Price row. In the completed E1 table, quantile LightGBM passes with conditional negative rate 0.161077 and mean surplus 2.673447; the original Technical row also passes but has lower surplus, and linear Student-t fails the threshold with conditional negative rate 0.312232. No tie affected the recorded selection. The full model, feature, preprocessing, and summary identities required to regenerate the candidate scores remain the artifacts specified in Appendix B; aggregate E1 scores do not reconstruct them.

D.4 E2: Factorial Design, Prompts, and Aggregation

Two independently varied information sources. E2 distinguishes descriptive information from explicit value information at fixed recorded states. Each source is either a training reference, denoted `R`, or individualized, denoted `I`. In the two-letter arm name, the first letter denotes descriptive information and the second denotes value information:

Arm	Descriptive source	Value source
RR	Reference	Reference
IR	Individualized	Reference
RI	Reference	Individualized
II	Individualized	Individualized

(50)

This is the fixed-state design in Eq. (50), not four additional arms pooled with the online studies.

Descriptive information concerns the designated market predictions. Value information relates predicted task gain to visible supported action prices. The individualized value source is the E1-selected quantile LightGBM predictor. Reference information comes from training statistics, not from outcomes of the actions being evaluated.

Executed prompt and action interface. The interface design aligns field roles and presentation across the reference and individualized conditions. Public state and supported action prices remain available under the corresponding frozen interface. An advisor produces a recommendation, and a participant retains the final action choice. An automatic gain–price maximizer does not replace that participant decision in E2.

The executed advisor system instruction begins: "You are a non-binding market advisor," and requires compact JSON with `recommended_action_id`, `confidence`, and `rationale`, recommending one exact legal `action_id`. The participant system instruction begins: "You are the buyer participant in a fixed-state archived market decision," and requires compact JSON with `action_id` and `rationale`. The advisor user payload contains `task`, `state`, and `interface_panels`; the participant payload adds `nonbinding_advisor`. The panels contain legal-action rows with action identifiers, product references, supported offers, descriptive predictions, conditional payment summaries, value sources, predicted task gains, gain minus offer, and estimated deal probabilities. Requests use JSON-object responses, a 120-second timeout, and up to three accepted-response attempts; the advisor and participant calls use 256 and 192 maximum output tokens, respectively. The first accepted valid response for a logical task is persisted.

These excerpts are taken from the executed interface code. They do not by themselves reproduce every rendered token. Exact field order, numeric formatting, reference-group construction, sampling settings inherited by the transport, and configuration hashes remain dependencies of the executed artifacts, not properties recoverable from the outcome table alone. Likewise, aligned information fields do not prove equality of every rendered token, numerical value, or semantic cue.

The selected legal action is linked to its archived outcomes only for evaluation. Oracle action values, realized future payments, and the empirical best action are not supplied as prospective value-panel inputs. Choices outside the recorded legal support cannot be assigned a measured outcome by this design.

The executed E2 interface is implemented in `src/iclr_budget1000/value_interface.py`. It serializes compact JSON messages for a nonbinding advisor and then for the buyer participant, requests JSON-object responses, validates the selected action identifier against the legal support, and persists the first accepted valid response by logical task identifier. The panel fields report the source of descriptive and value quantities, supported offers, action identifiers, choice and completion summaries, conditional payment, predicted technical gain, gain minus offer, and an ex-

plicit common-NA seller temporal field for this fixed-state support. These implementation artifacts, rather than a reconstructed prose template, define the executed prompt and validation contract.

Completed units and the missing repetition. The intended matrix contains $239 \times 4 = 956$ parent–arm cells, with three repetitions per cell, or 2,868 planned request units. The completed report contains 2,867 complete request units. All 956 parent–arm aggregate cells and all four arms within each of the 60 blueprints are represented.

Completeness at the parent–arm or blueprint–arm level therefore does not mean that every planned repetition completed. The missing request unit is parent `full:id_test:bp-0032:...:buyer_initial`, arm RR, repeat 2. It terminates as `FAILED` with a JSON decoding error after the recorded attempts. The analysis uses the 2,867 rows with status `COMPLETE`; the failed row has no selected action and is not imputed. A request unit and its associated API attempts also do not represent an independent market.

Grouping and estimands. New advisor and participant choices are evaluated against the old action-outcome library. The design aggregates repeated choices within parent–arm cells and then combines the corresponding parent summaries within each blueprint. Let $Z_b(h)$ denote the recorded blueprint-level surplus aggregate for arm h, and define

$$\overline{Z}_h = \frac{1}{60} \sum_{b=1}^{60} Z_b(h). \tag{51}$$

The implemented aggregation first averages available complete repetitions within each parent–arm cell, records the number of completed repeats, and then averages parent–arm rows within each blueprint and arm. The one failed repeat therefore changes only the available-repeat average for its affected parent–arm cell; all 60 blueprint–arm combinations remain present.

The primary individualized-value effect averages over the descriptive-information factor:

$$\widehat{\Delta}_V = \frac{1}{2} \left[\left(\overline{Z}_{\mathrm{RI}} - \overline{Z}_{\mathrm{RR}} \right) + \left(\overline{Z}_{\mathrm{II}} - \overline{Z}_{\mathrm{IR}} \right) \right]. \tag{52}$$

The exploratory descriptive main effect and interaction are

$$\begin{aligned} \widehat{\Delta}_D &= \frac{1}{2} \left[\left(\overline{Z}_{\mathrm{IR}} - \overline{Z}_{\mathrm{RR}} \right) + \left(\overline{Z}_{\mathrm{II}} - \overline{Z}_{\mathrm{RI}} \right) \right], \\ \widehat{\Delta}_{DV} &= \left(\overline{Z}_{\mathrm{II}} - \overline{Z}_{\mathrm{IR}} \right) - \left(\overline{Z}_{\mathrm{RI}} - \overline{Z}_{\mathrm{RR}} \right). \end{aligned} \tag{53}$$

The interaction is a difference between the two value-information contrasts. It is not the combined main effect and is not divided by two.

The relevant replication unit for uncertainty is the blueprint, not each model call, selected action, or reused rollout. For the reported interval, the implemented analysis forms the 60 blueprint-level $\widehat{\Delta}_V$ values and draws 10,000 bootstrap resamples of those 60 values with replacement, using random seed 20260925. The endpoints are the 0.025 and 0.975 sample quantiles computed by `numpy.quantile`. No sign-flip probability or Holm adjustment is attached to this E2 development interval.

D.5 E2 Results, Secondary Effects, and Development Gate

Reported factorial effects. The individualized-value main effect is 1.948255, with reported 95% interval $[1.527040, 2.385438]$. This supports a positive fixed-state development signal for the individualized value information supplied through the evaluated advisor/participant interface.

The exploratory descriptive-information effect is -0.058160, and the interaction is -0.484524. No additional confidence intervals or hypothesis-test probabilities are reported for these two quantities. Their signs are therefore not promoted into separate significance claims.

Table 34 preserves the available estimates. The four absolute arm means are not reconstructed from the three reported effects: the effects do not identify the common outcome level. Only the available factorial contrasts are reported.

Table 34: Completed E2 factorial-effect summary. Values are rounded from the reported analysis output. Only the value main effect has a reported 95% interval. NR means not reported here; it does not indicate a zero interval or a nonsignificant test. Effects concern fixed-state choices evaluated against archived action outcomes.

Estimand	Estimate	Reported 95% interval	Role
Individualized value, $\widehat{\Delta}_V$	1.948255	$[1.527040, 2.385438]$	Primary
Individualized description, $\widehat{\Delta}_D$	-0.058160	NR	Exploratory
Interaction, $\widehat{\Delta}_{DV}$	-0.484524	NR	Exploratory

Table 35: E2 execution and development diagnostics as reported. Rate and contribution fields retain their originating analysis definitions. The gate result is a development decision, not a new significance statement.

Recorded field	Value
Parent–arm aggregate cells	956
Complete request units	2,867
Blueprints with all four arms	60
Individualized-value action-change rate	0.797071
Conditional negative-rate worsening	-0.164365
Top-five absolute-contribution fraction	0.219322
Development GO gate passed	True

Development diagnostics. The reported individualized-value action-change statistic is 0.797071. The conditional negative-rate worsening statistic is -0.164365, and the top-five-blueprint absolute-contribution fraction is 0.219322 (Table 35).

These fields retain their analysis-specific grouping. The action-change statistic is not an advice-adoption probability and is not converted into a count of distinct parents without its actual denominator. The negative worsening value is a difference in rates, not a negative probability. The contribution fraction is not converted into a new concentration definition or an influence-adjusted estimate.

`go_gate_passed=True` under the prespecified development rule for the recorded E2 workflow. The frozen predicate requires $\widehat{\Delta}_V > 0.5$, the bootstrap lower endpoint at least -0.25, individualized-value action change at least 0.10, conditional negative-rate worsening at most 0.03, and top-five absolute blueprint contribution fraction at most 0.50. This is a decision about the development workflow, not an additional statistical test, an independent replication, or proof of a dynamic-market benefit.

The action-change statistic is computed from complete parent-state–repeat rows by comparing RI with RR and II with IR where the paired arm choices are available. The conditional negative-rate worsening is computed after blueprint–arm aggregation as $0.5(\mathrm{RI}+\mathrm{II})-0.5(\mathrm{RR}+\mathrm{IR})$, with missing values skipped by the stored mean operation. The contribution fraction is the sum of the five largest absolute blueprint-level value effects divided by the sum of all absolute blueprint-level value effects.

Interpretation and relationship to other evidence. E2 changes the information delivered to a choice-producing interface while evaluating the resulting legal actions on recorded consequences. It therefore differs from E1, which directly scores actions with a specified numerical rule. Agreement in the direction of their development results does not make them independent tests: predictor selection and choice evaluation reuse the same action-outcome library.

The positive value main effect also does not by itself establish a dynamic ranking of value advice over descriptive advice. The completed online studies in Appendix E allow market states and subsequent decisions to evolve, whereas E2 holds the evaluated parent states and action consequences fixed. Its outcome unit and feedback structure are different. Effect magnitudes or significance decisions across these studies are not a matched test of interface superiority.

The combined fixed-support evidence identifies a useful development direction: alternative gain estimates and explicit value information can change supported choices and their recorded returns.

It does not establish generalization to new market mechanisms, optimal purchasing outside the legal support, or beneficial feedback through a continuing market. Those claims require their own completed evaluations and cannot be supplied by the development gate.

E Online Advice Studies and Statistical Analyses

This appendix reports three completed continuing-market studies. The historical Advisor42 study compares seven advice conditions on 206 paired blueprints using a JDCloud-hosted model. The official-API format-control study compares four conditions on 80 paired blocks, including a same-format training-reference control. The online value-interface study, E3, compares three conditions on 72 paired blueprints and introduces explicit buyer-value information.

Each study retains its own hosted-model configuration, information package, input pool, and primary comparison family. The fixed-state experiment E2 in Appendix D is different: it generates new choices but evaluates them against archived action outcomes. The online studies instead execute the selected actions and allow market states and subsequent decisions to evolve.

E.1 Shared Market Outcomes and Intervention Mechanics

Continuing-market outcomes. All three online studies use the market rules and task-utility definition in Appendix A. The intervention window is $[56, 98)$, a 42-day continuation within the 168-day market. Outcomes accumulate over the evaluation window before averaging across blueprints. They are not per-transaction averages, fixed-parent action values, or 14-day prediction scores.

The four recorded market outcomes are completed transactions, realized technical gain, buyer net technical surplus, and seller gross revenue. Let $C_b(h)$, $G_b(h)$, $U_b(h)$, and $R_b(h)$ denote these totals for blueprint b under arm h. Under the stated transaction accounting,

$$U_b(h) = G_b(h) - R_b(h). \tag{54}$$

Non-transactions contribute zero realized gain, payment, and surplus, without assigning zero potential value to an unpurchased product. The stored window and event-inclusion conventions determine which transactions contribute to each total.

Experimental gain and payment share the specified accounting scale, not a monetary interpretation. Seller gross revenue is not profit because comparable seller costs are not modeled. Gain, buyer surplus, and seller revenue are also not three additive components of a welfare score. Displayed means retain their source precision; small discrepancies caused by rounding are not repaired by changing a reported component.

Matched continuations and participant authority. Pairing is defined within each study. In the historical study, a common day-56 checkpoint is copied into seven branches, which continue under matched exogenous semantic streams. Matching exogenous conditions does not force identical participant decisions, identical numbers of subsequent prompts, or identical LLM outputs. Once actions differ, the market states encountered later can differ.

The advisor receives its permitted role-specific public context and, when applicable, the designated prediction summaries before one advisor completion. The local orchestrator obtains those summaries; the advisor does not autonomously choose tools through an open-ended interaction loop. Advice remains optional. Participants retain their legal action interface and final decision authority, rather than being replaced by the offline optimizer.

The online market is not restricted to evaluating actions by lookup in the archived fixed-support outcome library. Its consequences arise from the continuing simulation. This distinguishes its market-level outcomes from the fixed-state E1 and E2 results.

Forecast exposure. In the descriptive-advice paths, buyer advisors receive Natural and Controlled summaries, while seller advisors receive the designated Dynamic42 forecasts. Technical-derived summaries remain internal features of Natural. Explicit Technical-value outputs and automatic net-surplus optimization are not exposed to these advisors. In E3, the advised value arms instead expose

buyer-facing value information from either training-reference values or the selected external individualized value predictor, alongside individualized descriptive forecasts.

Predictors and their designated calibration versions remain frozen within an evaluation. The external quantile LightGBM predictor selected for the value interfaces is not retrospectively inserted into the descriptive online studies. The model, calibration, and temporal-input qualifications in Appendix B consequently remain applicable.

Scope of the no-OOD designation. The historical no-OOD protocol specifies a designed market regime, rather than operating an out-of-distribution detector. Development and main markets use the designated MarketEngine, LOW/HIGH regimes, and generation rules. The reported protocol does not execute the earlier OOD, nearest-neighbor, robust-IQR, or trajectory-assessment machinery. An in-scope compatibility flag therefore does not establish generalization to new market mechanisms.

The format-control study likewise retains the identity of its archived sealed no-OOD input pool. Generating new API continuations does not turn that pool into a new mechanism-OOD population.

E.2 Executed Interaction Protocol

Advisor objective and participant-visible context. In advised online branches, the local orchestrator builds the public market page and the study-specific information tables before one advisor completion. The participant-visible advisor panel is then inserted into the participant's ordinary decision context; it does not replace the participant's legal action interface. For E3, the executed advisor system instruction begins: "You are a nonbinding platform advisor in a simulated data market," requires compact JSON with exactly six keys, and instructs the advisor not to claim the model is proven to improve profit. The required keys are `headline`, `public_evidence`, `nonbinding_guidance`, `guidance_code`, `guidance_subject_ref`, and `warnings`. Buyer guidance codes are `BUYER_CONSIDER_ACCEPT`, `BUYER_CONSIDER_COUNTER`, `BUYER_CONSIDER_REJECT`, and `BUYER_NO_DIRECTION`; seller guidance codes analogously cover lowering, holding, raising, or giving no direction.

The E3 user payload contains `advisor_role`, `arm`, `public_market_page`, `descriptive_table`, and `value_table`. The E3 advisor panel inserted into the participant request contains the descriptive and value tables, the nonbinding narrative, the guidance code, role, arm, decision stage, and source labels. Thus the participant can directly receive the numeric descriptive and value tables as well as the advisor's text; the effect is a service and information-package effect, not necessarily one mediated only through the textual recommendation. Audit fields such as prompt hashes, public-input hashes, tool-input hashes, panel hashes, latency, usage, replay status, and fallback flags are returned in a separate audit object and recorded in result metadata; they are not fields of the participant-visible panel. For seller decisions, the value table is a role-not-applicable empty table. For buyer decisions, the value table columns are legal action, product reference, supported offer, predicted technical gain, gain minus offer, estimated deal probability, value source, and note.

The snippets above identify the executed interaction contract more directly than the arm labels alone. They still do not replace the full prompt artifacts: complete rendered prompts, field ordering, reference-statistic construction, repair and retry records, and prompt hashes remain dependencies of the recorded execution.

E.3 Historical Seven-Arm Study

Study and model identity. The historical fresh-main record contains 206 blueprints, with 103 in each demand regime and one complete run of each arm per blueprint. Both participants and advisors use the recorded JDCloud-hosted DeepSeek V4 Flash runtime, whose request-model identifier is `/mnt/model-1`. This identifies the historical execution; it is not a claim that the hosted weights are identical to those behind a later official API alias.

The seven arms are

$$\mathcal{H}_{\text{hist}} = \{N_0,\, L_S, L_B, L_{SB},\, T_S, T_B, T_{SB}\}\,. \tag{55}$$

Table 36: Executed role, model, and output contracts for the later value interface and online API studies. The table gives recorded settings, not inferred provider defaults. Historical JDCloud arms retain the separate execution identity in Section E.3.

Study	Generated object	Recorded contract
E2 fixed-state	Advisor and buyer participant	Official-API transport with JSON-object responses and 120-second timeout; advisor output validates `recommended_action_id`; participant output validates `action_id`; maximum output tokens are 256 and 192.
Format-control online	Bilateral advisor service	Official DeepSeek API `deepseek-flash`; thinking disabled; descriptive fields aligned between reference and individualized arms; literal prompt and reference-construction details remain stage artifacts.
E3 online value interface	Buyer and seller advisors	Official DeepSeek API `deepseek-flash`; thinking disabled in the study configuration; advisor calls use JSON-object responses, maximum 384 output tokens, temperature 0.70, top-p 0.95, and 120-second timeout.
E3 online market actions	Buyer and seller participants	Official DeepSeek API `deepseek-flash`; participant calls use JSON-object responses, maximum 64 output tokens, temperature 0.70, top-p 0.95, thinking disabled, and the recorded 120-second timeout; advised branches receive the participant-visible advisor panel described above.

N_0 supplies no advisor. L supplies language-only advice, and T adds the designated frozen forecast context. Subscripts S, B, and SB indicate advice to sellers, buyers, or both. Serving one side means serving the designated participants on that side, not randomly selecting a fraction of them.

The buyer context contains the qualified descriptive choice, completion, conditional-payment, and public-reputation information. The seller context uses the designated Dynamic42 V12 route. The treatment changes the complete forecast-context package, including its presentation. It does not independently randomize numerical forecast content while holding all context length and wording fixed.

Historical tool notes retain their actual treatment identity. In particular, the approximately 44% seller-score reduction quoted in the historical context comes from the 60-market post-hoc predictive pool, not the qualified 57-market evaluation. It is not relabeled in retrospect. The online study itself does not re-estimate all Core14 or Dynamic42 prediction accuracies.

All arm means. Table 37 retains the four outcomes for all seven arms. One-sided service arms provide context but do not introduce additional primary tests or establish which recipient arrangement is optimal.

Table 37: Historical online arm means. Each entry averages 42-day blueprint totals over the same 206 paired blueprints. L and T denote language-only and tool-augmented advice; subscripts identify the served participants. These point estimates do not add new pairwise significance tests.

Arm	Transactions	Task gain	Buyer surplus	Seller revenue
N_0	52.820388	176.629832	59.151482	117.478350
L_S	53.033981	175.000341	58.016069	116.984272
L_B	52.830097	162.093100	43.176304	118.916796
L_{SB}	52.990291	155.832129	37.166886	118.665243
T_S	53.067961	175.300117	66.002253	109.297864
T_B	52.567961	173.842856	53.521302	120.321553
T_{SB}	52.451456	172.519936	56.760518	115.759417

The two primary references. The historical contrast $T_{SB} - L_{SB}$ measures the incremental effect of the complete forecast-context package within bilateral advice. The contrast $T_{SB} - N_0$ measures the complete bilateral tool-advice service relative to no advice. These are different questions, and neither reference is discarded.

Table 38: Complete historical eight-test primary family. Every effect is T_{SB} minus the indicated reference. Intervals are unadjusted 95% paired percentile bootstraps. Holm probabilities retain the original eight-comparison family and source precision. Adjustment is not recomputed from the rounded probabilities displayed here.

Reference	Outcome	Effect	95% interval	Raw p	Holm p
N_0	Transactions	-0.368932	$[-0.611650, -0.116505]$	0.004050	0.020250
N_0	Task gain	-4.109896	$[-14.041962, 6.059433]$	0.435016	0.870031
N_0	Buyer surplus	-2.390964	$[-12.146192, 7.389413]$	0.637534	0.870031
N_0	Seller revenue	-1.718932	$[-4.179829, 0.564909]$	0.162728	0.488185
L_{SB}	Transactions	-0.538835	$[-0.742718, -0.330097]$	0.000010	0.000080
L_{SB}	Task gain	16.687807	$[6.635589, 27.715253]$	0.001540	0.009240
L_{SB}	Buyer surplus	19.593632	$[9.582992, 30.332259]$	0.000120	0.000840
L_{SB}	Seller revenue	-2.905825	$[-5.382189, -0.646436]$	0.012840	0.051359

The two references and four outcomes form one eight-test primary family. Table 38 gives every recorded effect, unadjusted bootstrap interval, raw probability, and Holm-adjusted probability. The corresponding procedures are specified in Section E.6.

Relative to language-only bilateral advice, the tool-context arm increases task gain and buyer surplus and reduces transactions under the historical Holm criterion. The seller-revenue estimate is negative and its unadjusted interval excludes zero, but its Holm probability is 0.051359 and does not meet the family criterion.

Relative to no advice, the transaction decrease passes adjustment, whereas the task-gain, buyer-surplus, and seller-revenue comparisons do not. Thus the historical evidence establishes some incremental value over the language-only advisor, not value improvement over the unadvised market. Non-rejection does not establish equivalence or absence of harm.

E.4 Four-Arm Format-Control Study

Frozen design and input-pool identity. The completed study is identified as `ADVISOR42_FORMAT_CONTROL_BUDGET1000_V1`. It generates new official-API continuations on an archived sealed no-OOD main input pool. Its identity is therefore new API branches on recorded inputs, not a newly generated mechanism-OOD blueprint population. No cross-study independence claim is inferred merely from a new execution identifier.

A four-block development pilot, balanced with two LOW and two HIGH blocks, completes 16 branches and checks execution, information interfaces, and budget feasibility. Pilot blocks are excluded from main inference. The main protocol is frozen before its formal execution.

The final main contains 80 blocks, with 40 per demand regime and four arms per block. All 320 branches complete. The outcome table contains 320 complete block–arm rows: 80 per arm and 40 per regime–arm cell. Sample-size selection follows the pilot-based budget rule, rather than a guarantee of statistical power or a rule to continue until an effect becomes significant.

Official API configuration. The recorded request and response model identifier is `deepseek-flash` on the official DeepSeek API. Requests explicitly disable thinking through `thinking={"type":"disabled"}`. The model identifier is an execution record, not a guarantee that the provider's underlying weights are permanently immutable. The experiment does not have a verifiably fixed server-side sampling seed.

The historical JDCloud request identifier and this official identifier are not treated as equivalent weight snapshots. Other exact request settings, prompt versions, and serialization rules remain attached to the frozen execution configuration rather than being inferred from provider defaults.

Reference statistics and individualized descriptions. Table 39 maps the manuscript's short arm names to their recorded identifiers. N_0 has no advisor generation. L supplies bilateral language-only advice from the permitted public context, without an additional prediction table. M and T supply aligned descriptive fields but use different sources for their numerical content.

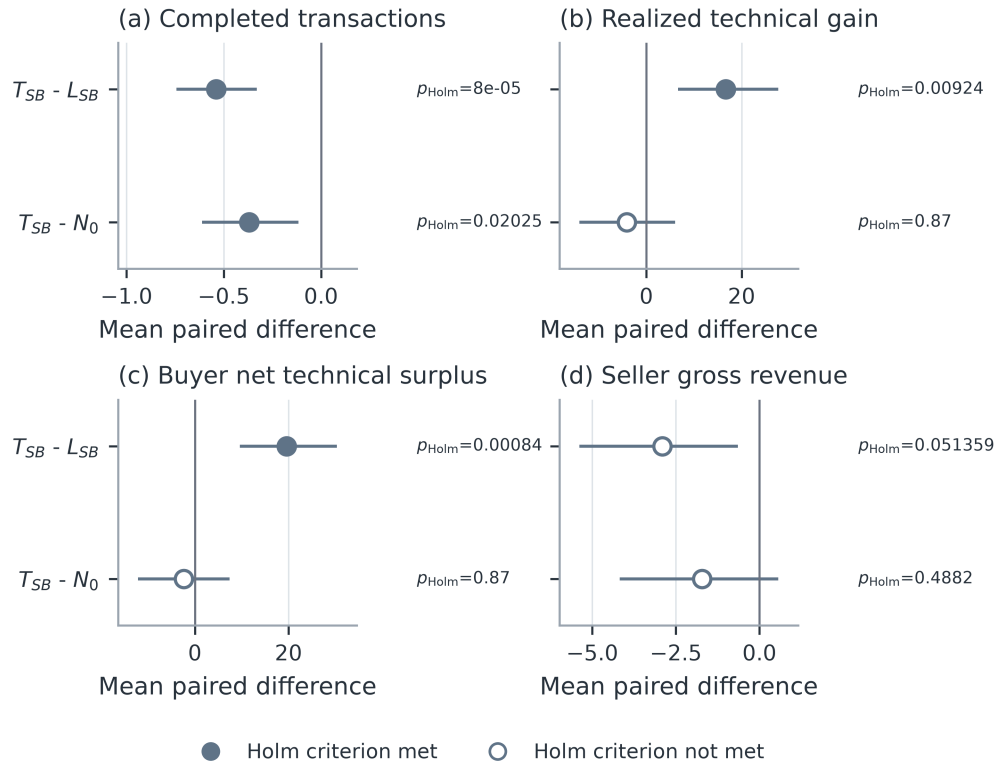

Figure E1: Historical online effects across 206 paired blueprints. Separate panels retain the four 42-day outcomes and their units, with both $T_{SB} - L_{SB}$ and $T_{SB} - N_0$ displayed. The horizontal axis is the mean paired difference in the outcome's own units. Intervals are unadjusted 95% paired bootstraps. Filled markers indicate comparisons passing the original eight-test Holm criterion, and displayed probabilities are Holm-adjusted. Seller revenue relative to L_{SB} does not pass adjustment despite its negative unadjusted interval. These historical effects are not pooled with the official-API format-control study.

Table 39: Official-API format-control arms. All advisor conditions preserve their permitted public context and participant decision authority. M and T align the descriptive field structure; neither exposes the explicit value interface evaluated in E2 and E3.

Arm	Recorded identifier	Advice condition
N_0	`N0_NO_ADVISOR`	No advisor generation
L	`L_SB`	Language-only bilateral advice
M	`M_SB_REFERENCE`	Same-format training-reference descriptive statistics
T	`T_SB_DESCRIPTIVE`	Individualized descriptive predictions

The M arm uses training-only and Train-rollout reference statistics mapped to the corresponding fields. It does not read main outcomes or provide a focal-object individualized prediction. The T arm uses frozen descriptive predictions for the current admissible conditions. It does not expose explicit Technical values or invoke the automatic net-surplus optimizer.

The resulting $T - M$ contrast tests individualized descriptive predictions relative to a same-format reference-information package. It does not isolate numerical values from every source label, semantic difference, or presentation cue. Nor is M an information-free control: its reference statistics can themselves affect advice.

The completion report establishes the arm definitions and aligned-field contract, but does not reproduce the full literal prompts and all reference-construction details. Exact reference grouping and fallback rules, units, missing-value handling, numerical formatting, system instructions, and config-

Table 40: Official-API format-control means over 80 paired blocks. Each entry is a mean of 42-day block totals. Values retain the four-decimal precision of the completion report. The full-precision paired analysis, rather than subtraction of rounded means, supplies the primary contrast table.

Arm	Transactions	Task gain	Buyer surplus	Seller revenue
N_0	40.2500	155.3991	74.7895	80.6096
L	41.0125	157.4102	73.7205	83.6898
M	37.7625	140.0531	66.8221	73.2310
T	39.6000	141.9174	62.2259	79.6915

Table 41: Complete primary family of the format-control study. Both comparisons use buyer net technical surplus and 80 paired blocks. Intervals are unadjusted 95% stratified paired-bootstrap intervals. The two Holm probabilities belong to this study's primary family, not to the historical eight-test family.

Comparison	Effect	95% interval	Raw p	Holm p
$T - M$	-4.5962	$[-14.0565, 4.6803]$	0.3374	0.3374
$T - N_0$	-12.5635	$[-23.4380, -1.6811]$	0.0256	0.0512

uration hashes remain dependencies of the executed interface artifacts. A reconstructed illustrative prompt is not treated as the prompt used in the experiment.

Four-arm outcomes. Table 40 reports all four recorded outcome means. The primary endpoint is buyer net technical surplus over the 42-day continuation. The other three outcomes remain visible as secondary descriptions; their point ordering does not add primary hypothesis tests.

The two primary comparisons are $T - M$ and $T - N_0$ on buyer surplus. Table 41 retains their reported paired effects and the common two-test Holm family. Individualized descriptive forecasts do not establish an improvement over reference statistics: the mean effect is -4.5962, with interval $[-14.0565, 4.6803]$ and Holm probability 0.3374.

The no-advice comparison has mean effect -12.5635 and an unadjusted interval entirely below zero. Its raw probability is 0.0256, but its Holm probability is 0.0512. The point estimate is unfavorable, while the prespecified family-wise rejection criterion is not met. Neither the negative direction nor the failure to meet that criterion is described as evidence of a benefit.

Descriptively, T has slightly greater task gain than M but also greater payment, leaving lower buyer surplus. Relative to N_0, it has lower task gain and buyer surplus. Transaction counts in T are below those of L and N_0 but above those of M. Thus the historical statement that transactions decline against both historical references is not transferred to the new $T - M$ comparison.

These findings concern the evaluated descriptive interface and comparison packages. They do not establish that forecasts have no possible value, that the arms are equivalent, or that the intervention cannot cause harm. They also do not identify the outcomes of the separate explicit value-interface design.

Relationship to the historical study. The historical study contrasts a complete tool-context treatment with language-only advice and no advice. The format-control study adds a same-format training-reference condition, uses the official API, and has its own sealed input-pool and execution identities.

The difference between their findings cannot be attributed solely to format matching. Hosted-model conditions and treatment packages also differ. Historical and new means, confidence intervals, and significance decisions are therefore reported within their respective studies rather than combined into a pooled effect or a larger nominal sample size.

E.5 Three-Arm Online Value-Interface Study

Design and arms. The completed online value-interface study has evidence identity `E3_VALUE_INTERFACE_NEW_OFFICIAL_DEEPSEEK_API_ON_NEW_SEED_BLUEPRINTS`. It contains 72 formal paired blueprints, with 36 LOW and 36 HIGH demand-regime blueprints,

and 216 complete blueprint–arm rows. The study uses new official-API continuations on a new-seed blueprint pool; it is not a continuation of the 80-block format-control sample and is not a fixed-support replay of E2.

Table 42 records the arm definitions. Both advised value arms use the same individualized descriptive prediction source and field contract. Because the market branches can diverge after earlier decisions, this does not mean that later prompts in the two arms contain identical state rows or identical numerical forecasts. They differ in the source of the buyer-facing value information: V_{REF} supplies training-reference value fields, whereas V_{IND} supplies individualized value fields from the selected external value producer. The no-advisor arm preserves the participants' ordinary public observations and final decision authority; it does not remove the market itself.

Table 42: Online value-interface arms. V_{REF} is not the format-control M arm because it already includes individualized descriptive information. V_{IND} is not the descriptive T arm because it additionally exposes individualized value information.

Arm	Descriptive information	Explicit value information
N_0	No advisor panel	No advisor panel
V_{REF}	Individualized descriptive forecasts	Training-reference value fields
V_{IND}	Individualized descriptive forecasts	Individualized value fields

In the advised E3 arms, both buyer and seller advisor providers are active before their corresponding participant decision. The seller advisor receives individualized descriptive forecasts and a role-not-applicable empty value table. The buyer advisor receives individualized descriptive forecasts plus the buyer-facing value table described in Table 43. The advisor prompt, not an automatic optimizer, receives those tables and returns a nonbinding panel with evidence and guidance. Participants retain the legal action interface and final decision authority.

Table 43: E3 value-interface field contract. The table records what differs by arm in the executed value panel; Appendix B defines the LightGBM gain summary and reference construction.

Field group	V_{REF}	V_{IND}
Descriptive table	Individualized current-state descriptive forecasts	Same prediction source and field contract
Technical gain	Training-split mean by product lineage, with global fallback	Individualized quantile LightGBM gain summary
Supported offer	Reject row at zero; current ask for accept; legal 65% ask anchor for counter rows	Same displayed-offer rule
Gain minus offer	Reference gain minus displayed supported offer	Individualized gain minus displayed supported offer
Deal probability	Training-rollout buyer scenario reference probability	Mean transaction probability from the current descriptive table
Seller value table	Role-not-applicable empty table	Role-not-applicable empty table

Arm means and primary contrasts. Table 44 reports all four 42-day outcome means. The primary endpoint is buyer net technical surplus. The other outcomes are descriptive and retain their accounting identities; seller gross revenue is not profit.

Table 44: Online value-interface means over 72 paired blueprints. Values are 42-day blueprint totals averaged with equal blueprint weight.

Arm	Transactions	Task gain	Buyer surplus	Seller revenue
N_0	39.583333	98.295901	26.759096	71.536806
V_{REF}	24.597222	76.554604	32.045160	44.509444
V_{IND}	28.194444	126.346254	73.203338	53.142917

The primary family contains two buyer-surplus comparisons, shown in Table 45. Relative to V_{REF}, individualized value information raises buyer surplus by 41.158178. Relative to N_0, the complete value-interface service raises buyer surplus by 46.444242. Both comparisons pass the stored two-test Holm family criterion.

Table 45: Primary family of the online value-interface study. Both comparisons use buyer net technical surplus and 72 paired blueprints. The family is separate from the historical and format-control online families.

Comparison	Effect	95% interval	Raw p	Holm p	n
$V_{\mathrm{IND}} - V_{\mathrm{REF}}$	41.158178	$[27.682780, 55.076469]$	0.000010	0.000020	72
$V_{\mathrm{IND}} - N_0$	46.444242	$[29.345599, 64.680318]$	0.000010	0.000020	72

Secondary means show tradeoffs rather than a uniform market improvement. Relative to V_{REF}, V_{IND} has 3.597222 more transactions and 8.633473 more seller gross revenue. Relative to N_0, it has 11.388889 fewer transactions and 18.393889 lower seller gross revenue. No additional primary-family test is inferred from those secondary mean differences, and the accounting identity $U = G - R$ is not a behavioral mechanism decomposition.

Completion and recovery boundary. The final E3 analysis is defined by the completed outcome and contrast files, together with the frozen protocol and rescue protocol. The formal target remains 72 paired blueprints. The rescue continuation filled planned formal slots that were stopped by authentication failures; it is not counted as extra independent blueprints and failed attempts are not converted to market rejections or zero-surplus observations. The request-identity audit reports no mismatches between persisted request identities and their expected logical tasks. These execution facts support the final branch-status table; they do not make the hosted model weights or server-side sampling exactly replayable.

E.6 Statistical Procedures and Comparison Families

Paired estimands. Let $s \in \{\mathrm{hist}, \mathrm{fmt}, \mathrm{val}\}$ denote the historical study, format-control study, or online value-interface study, with $B_{\mathrm{hist}} = 206$, $B_{\mathrm{fmt}} = 80$, and $B_{\mathrm{val}} = 72$. For outcome k, arm h, and blueprint/block b, write $Y_{b,k}^{(s)}(h)$ for the recorded 42-day total. The active arm is $h_{\mathrm{hist}} = T_{SB}$, $h_{\mathrm{fmt}} = T$, or $h_{\mathrm{val}} = V_{\mathrm{IND}}$. For a specified reference r,

$$\begin{aligned} d_{b,r,k}^{(s)} &= Y_{b,k}^{(s)}(h_s) - Y_{b,k}^{(s)}(r), \\ \widehat{\Delta}_{r,k}^{(s)} &= \frac{1}{B_s} \sum_{b=1}^{B_s} d_{b,r,k}^{(s)}. \end{aligned} \tag{56}$$

The historical contract combines replications within a blueprint when applicable, but the fresh-main record has one run per arm and blueprint. Transactions, advisor messages, and API attempts are not additional independent units in Eq. (56).

Historical paired percentile bootstrap. For one historical outcome/reference comparison, abbreviate $d_{b,r,k}^{(\mathrm{hist})}$ as d_b. Each of 10,000 bootstrap replicates resamples 206 blueprint effects with replacement:

$$\widehat{\Delta}_{\mathrm{hist}}^{*(u)} = \frac{1}{206} \sum_{j=1}^{206} d_{I_j^{(u)}}, \qquad I_j^{(u)} \sim \mathrm{Unif}\{1, \ldots, 206\}. \tag{57}$$

The historical implementation sorts the replicate estimates and linearly interpolates at zero-based position $(R_{\mathrm{boot}} - 1)p$ for quantile p. Its displayed interval is

$$\mathrm{CI}_{0.95} = \left[Q_{0.025}\left(\widehat{\Delta}^*\right), Q_{0.975}\left(\widehat{\Delta}^*\right) \right], \qquad R_{\mathrm{boot}} = 10{,}000. \tag{58}$$

This is the recorded paired bootstrap over the full blueprint set, not a newly imposed demand-stratified variant.

Demand-stratified bootstraps for the official-API studies. The format-control protocol retains 40 LOW and 40 HIGH blocks. Bootstrap resampling occurs within demand regime and preserves that allocation and the within-block arm pairing. For primary paired differences d_b, its resampled

mean has the form

$$\widehat{\Delta}^{*(u)}_{\mathrm{fmt}} = \frac{1}{2} \sum_{\ell \in \{\mathrm{LOW},\mathrm{HIGH}\}} \frac{1}{40} \sum_{j=1}^{40} d_{I^{(u)}_{\ell,j}}, \qquad I^{(u)}_{\ell,j} \sim \mathrm{Unif}(\mathcal{B}_\ell), \tag{59}$$

where $\mathcal{B}_\ell$ contains the 40 blocks in regime ℓ. The reported intervals use 10,000 such paired, stratified replicates. The equal regime weights here agree with the balanced equal-block target; they are not chosen after observing outcomes.

The format-control analysis configuration remains the source for its numerical quantile conventions and seeds. The historical base seed and interpolation details are not automatically assigned to this later analysis.

The online value-interface study uses the same stratified paired blueprint logic with its own 36 LOW and 36 HIGH blueprints:

$$\widehat{\Delta}^{*(u)}_{\mathrm{val}} = \frac{1}{2} \sum_{\ell \in \{\mathrm{LOW},\mathrm{HIGH}\}} \frac{1}{36} \sum_{j=1}^{36} d_{I^{(u)}_{\ell,j}}, \qquad I^{(u)}_{\ell,j} \sim \mathrm{Unif}(\mathcal{B}^{\mathrm{val}}_\ell). \tag{60}$$

Its analysis file records 10,000 bootstrap replications, sorts the replicate means, and linearly interpolates endpoints at zero-based position $(R_{\mathrm{boot}}-1)p$ for $p = 0.025$ and 0.975. The value-interface randomization calculations use 100,000 sign-flip replications and statistical seed 2026092503. Those records, rather than the historical or format-control defaults, define the E3 analysis identity.

Two-sided sign-flip calculations. The completed online contrast tables report 100,000 paired randomization or sign-flip replicates and an add-one probability. For a particular study and comparison, let

$$T^{(u)} = \frac{1}{B_s} \sum_{b=1}^{B_s} \epsilon^{(u)}_b d_b, \qquad P(\epsilon^{(u)}_b = 1) = P(\epsilon^{(u)}_b = -1) = \frac{1}{2}. \tag{61}$$

If K_{extreme} is the count of replicate statistics whose absolute value is at least as extreme as the observed mean under the recorded comparison rule, then

$$\widehat{p}_{\mathrm{raw}} = \frac{1 + K_{\mathrm{extreme}}}{R_{\mathrm{flip}} + 1}, \qquad R_{\mathrm{flip}} = 100{,}000. \tag{62}$$

The historical implementation uses $|T^{(u)}| + 10^{-15} \geq |\widehat{\Delta}|$ for the extreme-value comparison. Its contract records base seed 20260812 and stable per-comparison seeds. These are historical settings, not automatically assigned to the new official-API studies. With 100,000 Monte Carlo replicates and the add-one rule, the smallest displayable raw probability is $1/100001 \approx 0.000010$. The E3 raw probabilities at that value therefore indicate the recorded Monte Carlo resolution, not an exact tail probability of zero.

Sign flipping requires the relevant within-pair exchangeability or sign-invariance conditions under the tested null. Pairing alone does not establish those conditions, and an arbitrary mean-zero null need not imply sign invariance. The calculations are not presented as a randomized experiment on human traders or as proof of a causal mechanism of advice adoption. Matched local exogenous streams also do not certify identical server-side LLM randomness.

The format-control completion report specifies pairing, demand stratification, resampling counts, and add-one treatment, but does not fully reproduce its numerical quantile conventions, seed mapping, or sign-flip tolerance and tie handling. Those details remain dependencies of the executed analysis configuration. Neither the historical settings nor the E3 seed is assigned to it by assumption.

Separate Holm families. The historical primary family contains two references and four outcomes:

$$\mathcal{F}_{\mathrm{hist}} = \{N_0, L_{SB}\} \times \{\mathrm{transactions}, \mathrm{gain}, \mathrm{buyer\ surplus}, \mathrm{seller\ revenue}\}. \tag{63}$$

It has $m = 8$ hypotheses. The format-control study instead has one primary endpoint and two references:

$$\mathcal{F}_{\mathrm{fmt}} = \{(M, \mathrm{buyer\ surplus}), (N_0, \mathrm{buyer\ surplus})\}, \tag{64}$$

with $m = 2$. The online value-interface study likewise has one primary endpoint and two references:

$$\mathcal{F}_{\text{val}} = \{(V_{\text{REF}}, \text{buyer surplus}), (N_0, \text{buyer surplus})\}, \tag{65}$$

also with $m = 2$. One-sided historical service arms and secondary outcomes do not enlarge or replace these frozen primary families.

For ordered raw probabilities $p_{(1)} \leq \cdots \leq p_{(m)}$, Holm-adjusted probabilities are

$$\widetilde{p}_{(i)} = \max_{j \leq i} \min \left\{1, (m - j + 1)p_{(j)}\right\}. \tag{66}$$

The corresponding step-down procedure uses family level $\alpha = 0.05$, comparing successive ordered probabilities with $\alpha/(m-i+1)$ and stopping at the first failure (Holm, 1979). Adjusted probabilities are calculated from the stored unrounded probabilities, not the rounded table displays.

Intervals and testing answer different reporting questions. The displayed 95% intervals are not simultaneous family-wise intervals. Holm adjustment changes the testing decision, not the endpoints of those separately reported intervals. Consequently, a negative interval can coexist with an adjusted probability above 0.05.

This occurs for historical seller revenue relative to L_{SB} and for the format-control buyer surplus relative to N_0. Their reported adjusted probabilities are 0.051359 and 0.0512, respectively. The raw interval is not used to override the prespecified family decision. Conversely, non-rejection is not interpreted as equivalence, no effect, or no harm.

All formulas in this section document the associated analyses. They do not imply that manuscript preparation independently repeated the original bootstrap or sign-flip computations.

E.7 Execution Integrity and Descriptive Analyses

Historical completion and recovery. The historical execution report records all 206 main blueprints with seven complete branches. No-advisor parity checks pass, and N_0 has zero advisor calls. All 41,954 advisor calls in the tool arms have recorded delivery of the designated frozen prediction context.

The cumulative historical resource record contains 441,091 remote generations, 1,798,484,276 input tokens, and 41,870,199 output tokens. It preserves 190 engineering-failure attempts and reports no missing final tasks, selective deletion, or outcome-driven rerunning. Recovery retains the original task, prompt, and seed identities. These cumulative counts are not counts of independent main blueprints or unique decisions.

The earlier withdrawn Advisor42 V1 treatment is not pooled with the fresh no-OOD main study. Likewise, successful forecast-context delivery does not independently validate every forecast's accuracy or show that participants followed its advice.

Official-API completion. The format-control study's terminal record contains 80 completed four-arm blocks and 320 complete branches. The formal branch matrix and finalized outcome files, rather than intermediate runner output, define the reported result.

Live requests use a common transport and budget ledger. The first accepted valid response is persisted so that resumption does not create a new independent repetition of an already accepted decision. Infrastructure failures such as SSL or connection errors are retried under the same frozen protocol. They are not converted into market rejection actions, zero task gains, or zero-surplus observations.

Uncertain request costs remain uncertain in the ledger. A request with unknown billing status is not automatically a missing market branch, just as a successful request is not an independent blueprint. The finalized branch-completion status and the accounting ledger answer different integrity questions.

The authoritative main artifacts include the block manifest, frozen protocol, the completed block–arm outcome table, the contrast table, final block records, and the runner ledger. Their final source mapping is retained in Appendix F. The completion records and local code tests support execution auditing, not independent replication of every scientific result.

The final artifact map binds this study to finalized outcome, contrast, protocol, block-record, and ledger files rather than partial terminal output. The similarly named zero-byte runner-state path reported under the stage outputs is not the authoritative runner database. Exact hashes and permitted release paths belong to the release manifest and reproducibility records, without credentials or personal absolute paths.

Historical concentration and coverage. The exploratory historical analysis describes participation and distribution in addition to the four primary totals. Table 46 preserves the recorded concentration and coverage contrasts. These intervals are the archived descriptive normal intervals, not paired-bootstrap intervals or additional members of the eight-test primary family.

For example, the seller-sales Gini difference is 0.039812 relative to no advice and 0.034436 relative to language-only bilateral advice. These positive differences describe greater concentration under the recorded measure. They do not by themselves establish improved or reduced social welfare.

Conversion, review coverage, terminal product-review coverage, Gini, and HHI retain their different definitions and eligible denominators. They are not summed into a composite score, and the intervals are not promoted into unreported Holm-adjusted decisions.

Table 46: Complete recorded historical exploratory concentration and coverage contrasts. Each difference is T_{SB} minus the reference. Intervals are archived descriptive normal intervals, not primary-family bootstrap intervals. Metric definitions and eligibility remain specific to the corresponding analysis.

Reference	Measure	Difference	Descriptive 95% interval
N_0	Conversion rate	-0.005999	$[-0.010590, -0.001408]$
N_0	Review coverage	-0.005780	$[-0.011132, -0.000428]$
N_0	Seller sales Gini	0.039812	$[0.030663, 0.048962]$
N_0	Seller sales HHI	0.018459	$[0.014524, 0.022394]$
N_0	Seller revenue Gini	0.026692	$[0.018662, 0.034721]$
N_0	Seller revenue HHI	0.030827	$[0.018421, 0.043233]$
N_0	Terminal product review coverage mean	-0.066125	$[-0.074444, -0.057805]$
L_{SB}	Conversion rate	-0.009293	$[-0.013143, -0.005444]$
L_{SB}	Review coverage	-0.003799	$[-0.008648, 0.001050]$
L_{SB}	Seller sales Gini	0.034436	$[0.025089, 0.043782]$
L_{SB}	Seller sales HHI	0.016395	$[0.012396, 0.020394]$
L_{SB}	Seller revenue Gini	0.026750	$[0.018332, 0.035169]$
L_{SB}	Seller revenue HHI	0.033187	$[0.019401, 0.046973]$
L_{SB}	Terminal product review coverage mean	-0.046325	$[-0.054436, -0.038214]$

Later advice-chain audit. A later audit of the historical advice records supplies the linkage counts in Table 47. This adds traceability information beyond the aggregate market outcomes, but does not identify a causal advice- adoption rate.

An advice record linked to a transaction does not establish that the participant would otherwise have acted differently, that every recommendation was followed, or that the transaction was improved by that advice. An unlinked record is not automatically an ignored recommendation. The reported fields are also not assumed to form an exhaustive, disjoint partition without their underlying join definitions.

Table 47: Later linkage audit of historical advice records. Counts retain the reported field meanings and audit scope. They are not independent-market counts, adoption probabilities, or a certified partition of all events. The tool-context delivery count reported separately concerns tool-arm advisor calls.

Recorded field	Count
Advisor events	83,908
Advisor events with a project identifier	44,356
Advisor events linked to a transaction	43,433
Advisor events without a transaction link	40,301
Advisor events with a recorded reject action	174

The cumulative generation totals, role-specific deduplicated records, tool-delivery counts, and linkage-event counts have different scopes. They are not added together or divided by a common denominator to manufacture a utilization or adoption statistic. Exact join keys and coverage limitations belong to the audit artifacts.

Exact join keys, deduplication rules, scope filters, and the treatment of rejection or no-transaction records remain dependencies of the audit artifacts. The aggregate counts do not justify a reconstructed dialogue, timestamped case study, recommendation-adoption rate, or mediation effect.

Explanatory limits. Reduced transactions with greater aggregate value relative to a particular advisor is compatible with several mechanisms. The summaries do not identify which opportunities were rejected, which low-value trades were removed, or whether any particular recommendation caused an action change. The later linkage audit does not remove this identification limit.

The completed format-control study more directly evaluates individualized descriptions against a reference-information package, but its aggregate outcomes likewise do not identify a unique behavioral mechanism. Its unfavorable point estimates are reported alongside the full comparison design, rather than used to conclude that every predictive information service is ineffective.

The online evidence therefore supports study-specific statements about the evaluated descriptive and value-interface services and their references. E3 supplies online buyer-surplus evidence for the evaluated value-interface service, but it does not validate an optimal policy, production-market welfare, or a general claim that all forecast advice improves all participants' outcomes.

F Reproducibility, Versioning, and Evidence Boundaries

This appendix maps the reported evaluations to their recorded artifacts, distinguishes reproduction tasks, and consolidates remaining data, model, prompt, and execution limitations. Historical predictions and market outcomes, later reanalyses, and newly generated API decisions retain separate identities. Their numerical results appear with the corresponding methods in Appendices A–E.

An artifact location identifies a recorded source. It does not by itself establish public availability or successful reproduction on another system. The distinction between presentation, numerical reanalysis, prediction regeneration, and new experimental execution is therefore explicit throughout this appendix.

F.1 Artifact Map and Reproduction Levels

Archive organization and entry points. The project handoff separates readable implementation copies under `code/`, data and model assets under `data/`, reports and integrity records under `logs/`, and original archives under `raw/`. Inventory and digest records are organized under `00_raw_audit/`. The later manuscript organization uses `paper/`; the older `ICLR/` directory is a document project rather than a complete training package.

The recorded code entry points are `code/ex1_agent_validation/` for behavioral evaluation, `code/core14_ex2/` for the short-horizon and buyer predictors, `code/dynamic42/` for temporal modeling, `code/dimwm_nonapi/` for non-API analyses, and `code/advisor42_no_ood/` for the historical online study. The later studies are organized under `experiments/iclr_budget1000_v1/`.

The organization report describes `code/` as independent copies, while parts of `data/` and `raw/` share hard-linked files. Directory separation therefore does not imply independent file contents. Any modifying reproduction should use independent working copies rather than overwrite those linked archival assets. Some entry points also retain assumptions about the original archive layout; relocating readable code alone does not establish that every pipeline runs unchanged.

Different reproduction tasks. Table 48 distinguishes the artifacts needed for different operations. Reading stored metrics, recomputing an estimator, regenerating a forecast, and obtaining a new hosted-model response are different procedures.

For example, rescoring E2's saved choices against the archived outcome library does not require generating new advice. Repeating E2's choice-generation stage does require the appropriate hosted-model interface. Neither procedure is a new dynamic-market continuation. Likewise, the non-API pipeline includes baseline and analysis stages as well as presentation stages; invoking an entire pipeline is not synonymous with read-only figure reconstruction.

Table 48: Reproduction levels and their distinct requirements. The table specifies what each operation depends on; it does not certify that every dependency is completely released or independently verified. Later levels are not established merely by completing an earlier one.

Operation	Required artifacts	What it establishes
Document and figure reconstruction	LaTeX sources, bibliography, specified numerical inputs, plotting code, and final figure exports	Presentation of the supplied records, not new statistical or experimental evidence
Numerical reanalysis	Saved predictions or outcomes, record identities, weights, selection rules, and analysis configuration	Recomputed summaries or contrasts on the same recorded evidence
Frozen inference	Checkpoints, fitted preprocessing, feature producers, calibration maps, and sampling identities	Predictions from the designated frozen implementation, subject to complete dependency recovery
Predictor retraining	Versioned data, partitions, supervision compilation, training settings, seeds, and software environment	A new fit under a specified protocol, not necessarily byte-identical parameters
New decisions or market continuations	Market states, participant and advisor prompts, forecast interfaces, hosted-model access, and execution/recovery configuration	New generated behavior under the specified protocol, not guaranteed replay of archived API responses

Sources for the completed additions. Within the later-study workspace, the 80-block online study is linked to `outputs/online_by_blueprint.csv` and `outputs/online_contrasts.csv`. Its frozen protocol, finalized pilot record, block outputs, and runner state belong to the stage `stages/live_integration_20260924_165618/`. The authoritative runner database is the stage's `ledger/main_runner_state.sqlite`, not the similarly named zero-byte file under its `outputs/` directory. The finalized pilot report is stage-local; the earlier top-level pilot-status document is not its replacement.

The completed E1 and E2 summaries are linked to `stages/value_interface_e1e2_20260925_v1/outputs/e1_policy_results.csv` and `stages/value_interface_e1e2_20260925_v1/outputs/e2_effects.csv`. These summaries identify reported results but do not alone reconstruct fitted predictors, literal prompts, or the complete statistical procedure.

The completed E3 online value-interface study is linked to `stages/value_interface_e3_20260925_v2/outputs/e3_online_by_blueprint.csv`, `stages/value_interface_e3_20260925_v2/outputs/e3_online_contrasts.csv`, `stages/value_interface_e3_20260925_v2/outputs/e3_analysis_summary.json`, and `stages/value_interface_e3_20260925_v2/outputs/e3_main_run_result.json`. Its frozen and rescue protocols remain distinct execution records; the rescue fills planned formal slots rather than adding a new statistical sample.

The later fixed-support and persona checks have separate outputs, including `outputs/offline_policy_recompute_integrity.json`, `outputs/offline_policy_recomputed.csv`, `outputs/persona_toy_null.json`, and `outputs/persona_permutation_summary.csv`. The read-asset manifest and claim-update ledger preserve their own source scope. Their existence does not convert reanalysis of archived outcomes into new market data.

Presentation provenance. Each numerical figure and table is associated with an experiment, evaluation pool, model or arm identity, outcome definition, aggregation unit, and uncertainty source. Historical intervals, newly reported intervals, missing metrics, and source-specific precision remain distinguishable. Aggregate scores are not expanded into unavailable calibration curves, individual examples, or confidence intervals.

Figure 1 is a conceptual, image-assisted overview rather than statistical evidence. Its labels and dependencies must agree with the methods, but its appearance does not validate an implementation. The remaining figures use numerical records. A plotting placeholder in the LaTeX source is not a completed data-figure export, and an archived PDF digest does not certify the later rewritten manuscript.

Figure and table reconstruction requires the corresponding source files, filters, metric columns, grouping rules, and uncertainty fields. A filename or matching figure number is insufficient to establish that an export uses the intended result. The digest of an earlier manuscript build identifies that earlier build, not a subsequently revised PDF.

F.2 Data, Model, and Prompt Provenance

Dataset attribution and local versions. The source bibliography identifies four data lineages. The archived electricity reference is OpenML dataset 151, version 1, describing the ELEC2 binary price-direction task. The pol reference is OpenML dataset 722, version 2, whose classification target is derived by binarizing an earlier telecommunications regression target at its mean (OpenML, 2014a;b). The original-source chain for pol remains incompletely documented.

The MiniBooNE reference is UCI dataset 199, attributed in the recorded citation to Byron Roe, 2005; the archive distinguishes that citation year from the later donation date. The Nomao reference is UCI dataset 227, attributed to Candillier and Lemaire, 2012 (Roe, 2005; Candillier & Lemaire, 2012). These references describe the task lineages, not necessarily the exact bytes used by every local preprocessing or product construction step.

The archived attribution notes record the string `Public` in the two OpenML license fields and CC BY 4.0 on the two UCI pages. These are recorded metadata observations, not a new legal clearance of the complete derived package. Exact local source files, transformations, attribution notices, and applicable redistribution permissions still require a version-specific mapping. This appendix does not replace any experimental dataset or recompute task gains using a newly downloaded version.

Local predictor identities. A frozen prediction depends on more than a checkpoint name. Its identity also includes input features and fitted preprocessing, upstream feature producers, calibration maps, retained-sample construction, and row alignment. The same model family can serve different roles with different raw or calibrated outputs. The raw Technical summaries consumed by Natural, calibrated Technical gain samples, and the external LightGBM value summaries used by E2 are therefore not interchangeable artifacts.

Similarly, the Core14 seller-scoring model, raw short-state producer, and additional temporal summary producers retain their specified roles. Matching checkpoint identities does not imply that a calibrated seller score and a raw joint-state cache are the same prediction. The explicit differences and incomplete mappings are documented in Appendix B.

The non-API archive identifies environment records under `data/dimwm_nonapi/environment/`. The documentation does not establish a single fully verified software and hardware environment for all historical and later stages. Exact package versions, numerical settings, and hardware requirements must remain attached to the corresponding run. They are not supplied by assuming contemporary library defaults or by copying one experiment's environment to another.

Hosted models and prompts. The historical Advisor42 run records a JDCloud-hosted DeepSeek V4 Flash runtime with request identifier `/mnt/model-1`. The completed official-API format-control study records `deepseek-flash` as both request and response identifier and explicitly disables thinking. These are execution identities, not evidence that the two providers served identical immutable weight snapshots.

The behavioral experiments also retain their separately recorded model and sampling configurations. A shared model-family name is insufficient to merge static behavior, continuing-market, and later API conditions. Local semantic seeds and paired exogenous streams do not establish control over all server-side generation randomness.

Prompt provenance includes the system and role instructions, public-state serialization, forecast or reference fields, numerical formatting, sampling settings, action validation, and retry behavior.

Table 49: Current evidence-availability status for the manuscript snapshot. "Recorded path" means the manuscript identifies an internal archive location; it is not itself a public-access guarantee.

Component	Current status	Scope
Manuscript text, bibliography, rendered figures, and reported tables	In manuscript materials	Provide presentation of the recorded analyses and results.
Aggregated E1/E2/E3 and format-control result files	Recorded paths listed in this appendix	Support numerical reanalysis when the corresponding archive is available; the PDF alone does not contain row-level records.
Fixed-support action tables, selected actions, rollout outcomes, and saved prediction arrays	Internal archive dependencies	Required to rerun the fixed-support and prediction analyses; not reconstructed from aggregate tables.
Model checkpoints, fitted preprocessing, compiled supervision, and local dataset derivatives	Restricted or incomplete mappings	Needed for fixed-predictor inference or retraining; exact local hashes and attribution mappings remain dependencies.
Hosted-model prompt/response logs and recovery ledgers	Internal records requiring review	Needed for complete execution auditing; not distributed here without credential and identifying-content checks.
Credentials, personal absolute paths, administrative records, and unapproved repository addresses	Excluded	Not part of the experimental evidence release.

Appendix D.4 and Appendix E.2 provide executed prompt excerpts and field schemas for E2 and E3. For the same-format study, the source and construction of training-reference statistics are part of the treatment. For E2, the source of explicit value information is additionally part of the treatment. These excerpts do not recover the exact rendered prompts in every branch, and complete literal templates remain dependencies of the execution artifacts that contain them.

Release scope and confidentiality. Internal archive availability is distinct from public release. Redistribution is limited by the applicable data, model, software, and hosted-service terms, as well as by the assets actually included in the release. No public repository address or unconditional release of every raw record is asserted here.

Third-party reference papers and layout-study screenshots are not part of the experimental source release. Undocumented software settings and hardware records remain run-specific dependencies rather than being replaced by inferred defaults.

F.3 Execution Costs and Integrity

Historical inventory and numerical audits. The September 24 organization report records 25,680 files in the raw inventory and 21,131 files in the organized main view. It identifies `00_raw_audit/raw_sha256sums.txt` and `00_raw_audit/organized_file_inventory.tsv` as the associated digest and inventory records. These counts describe file views, including linked archival content, not independent experimental observations.

A separate preceding document audit reports matching hashes for 1,533 original/snapshot files and 14 generated tables, agreement of 90 score pairs, and verification of the 239-parent, 3,830-action, 114,900-rollout support. It also reports reconstruction of 28 historical online arm–outcome means, eight paired effects, and eight Holm adjustments. That audit did not reproduce the historical bootstrap intervals, raw sign-flip probabilities, model training, or all market trajectories.

The non-API audit contains 45 completed tasks and seven skipped tasks because API execution was outside its scope. Its 90 frozen-metric comparisons match within 10^{-12}, with maximum absolute discrepancy 1.137×10^{-13}. It did not rerun official Test inference. These audit summaries can concern overlapping assets or calculations; their counts are not added as independent replications.

Later execution checks. The four-arm format-control study reports 89 passing local tests at its completion snapshot, while the final supplement report records 95 passing tests for the later combined project snapshot. The tested functionality includes budget reservations and recovery, response persistence, information-interface constraints, finite-probability calculations, grouped statistics, and completion-matrix checks. Those results do not certify every later code revision or every implementation detail.

The later fixed-support reanalysis reports no duplicate rollout identifiers, no selected actions missing from the outcome library, and zero discrepancy in the realized net-surplus accounting identity. The format-control study reports all 80 four-arm blocks complete. E2 separately reports all 956 parent–arm aggregate cells but only 2,867 of 2,868 planned request units complete. E3 reports all 72 formal value-interface blueprints complete, with 216 block–arm rows and 36/36 LOW/HIGH balance. Thus branch completeness, cell completeness, and repetition completeness remain different integrity conditions.

Reading these reports during manuscript preparation is not an independent rerun of their checks. Hash equality establishes byte correspondence within the checked scope; it does not by itself establish correct labels, proper statistical assumptions, or predictive generalization.

Accounting snapshots. Table 50 separates reported snapshots of the later project ledger. At completion of the four-arm format-control study, the ledger contains 123,762 request records, including connectivity, pilot, and main execution. It reports 538,130,056 prompt tokens and 14,156,342 completion tokens, with 87,874,446 cache-hit prompt tokens and 450,255,610 cache-miss prompt tokens.

Charges are computed from recorded usage and the applicable pricing snapshot, not from an independently reconciled provider invoice. The pricing configuration identified for this execution is `configs/deepseek_pricing_20260924.json`. It is an archived configuration, not a statement of current API prices.

The final supplement snapshot includes the completed fixed-state work and E3 online value-interface execution now reported in Appendix E. It is cumulative for the later supplement workspace, not a final cost for the whole historical paper project, and its values are not added to the earlier cumulative snapshot.

Table 50: Reported API-accounting snapshots in RMB. "Usage-based cost" is the ledger's `actual_yuan` field, not an independently reconciled invoice. Unknown charges and active reservations are separate accounting fields. The final snapshot is cumulative for the later supplement workspace and includes the completed E3 value-interface study; it is neither additive to the earlier snapshot nor a provider invoice.

Recorded snapshot and scope	Usage-based cost	Unknown charges	Active reservations
Four-arm completion snapshot, September 25, 2026	508.685187	0.624836	0.000000
Final supplement report, September 26, 2026, 14:12 Asia/Shanghai	893.648061	2.616829	0.000000

The original project budget is 1,000 RMB. An uncertain request charge is not released as zero. An active reservation concerns an in-flight request; completed and unknown records may still retain historical reservation fields. Summing all historical reservation fields would therefore misstate the amount currently reserved. Only records still in the active reservation state contribute to that snapshot field.

Historical JDCloud resource counts belong to a different execution scope. They are not billed retrospectively using the later official API's price configuration. Likewise, the available execution records do not establish complete CPU/GPU costs, wall-clock costs, or a single monetary total for all earlier model training and market generation. Those quantities are not inferred from API token counts.

Failure and recovery semantics. Infrastructure failures are not market actions. The later transport and budget ledger preserve uncertain charges and accepted responses across recovery. A failed API attempt is not converted into participant rejection, zero gain, or a new independent replication. The

Table 51: Remaining reproduction and interpretation boundaries. Each entry points to the corresponding detailed appendix. Reported completion of a later experiment does not automatically resolve an earlier model or data-provenance limitation.

Item	Current boundary	Details
Data provenance	Public dataset references do not completely establish local-byte identity, all transformations, or redistribution clearance.	A, F.2
Behavioral records	Matched-change profile-count descriptions and the archived inventory need a complete mapping; persona permutation inference retains its exchangeability assumption.	A
Natural and formal fits	The final feature-to-producer manifest and some historical row/group count mappings remain incomplete; verified selection-stage cross-fitting is not universal provenance for every final fit row.	B
Controlled correction	Completion-weighted conditional mixing was analyzed and tested, but per-query assets were unavailable for full prediction regeneration and recalibration.	B, D
Technical and temporal models	Finite Technical summaries do not repair the untruncated transformed mean; temporal arrival conditions, producer mappings, and distinct draw/pool identities retain their qualifications.	B, C
E1 value producer	Quantile files, action-level summaries, and the frozen selection predicate are linked to the executed value-interface stage; full public release still requires feature/preprocessing and checkpoint mapping.	B, D
E2 analysis	One RR repeat for `bp-0032` failed and is excluded from complete-row aggregation; the bootstrap seed, repeat averaging, and gate thresholds are now linked to execution records, while literal prompt and release metadata remain artifact-level dependencies.	D
Hosted online execution	Model aliases are not immutable weight snapshots; exact prompt and recovery identities are study-specific, rescue execution fills planned slots rather than adding samples, and server-side randomness is not fully controlled.	E

first accepted valid response is persisted so that resuming an already completed decision retains the accepted output.

The historical and later executions retain their own recovery records and protocol identities. Intermediate runner output, an unfinished block, or a similarly named empty database does not replace the finalized branch-completion record. Recovery logs and successful completion establish execution facts, not a mechanism explaining why advice changes behavior.

The accounting snapshot used for a manuscript claim must preserve its timestamp, project or stage scope, usage fields, pricing configuration, active reservations, and unknown charges. Stage-specific costs are not inferred from a wider cumulative snapshot without stage-level allocation. The 89-test result belongs to its stated format-control snapshot; later revisions are covered only by their own recorded checks.

F.4 Unresolved Issues and Claim Boundaries

Remaining implementation and analysis links. Table 51 consolidates the principal open items already identified in the preceding appendices. It distinguishes a missing artifact or execution mapping from a mathematical limitation and from an external-validity restriction. None is closed merely by changing terminology in the manuscript.

The unavailable DirectNet out-of-fold comparator remains `BLOCKED_ASSET_NOT_FOUND`. The corrected Controlled route remains `BLOCKED_RAW_QUERY_OUTPUTS` for its full evaluated prediction product. Neither status is a numerical score or a failed superiority test.

Similarly, documentation gaps are not resolved by assigning plausible hyperparameters, borrowing another study's seed, reconstructing individual records from rounded means, or treating a proposed protocol as the executed one. The observed E1 selection and E2 gate result remain reported outcomes of their development workflow, tied here to their frozen artifacts rather than promoted into independent confirmation.

Claim-to-evidence map. Behavioral statements are supported by the controlled scenarios, recorded actions, and qualified null analyses in Appendix A. They establish neither human equivalence nor optimal market behavior.

Predictive statements use the conditional interfaces in Appendix B and task-matched comparisons in Appendix C. Historical Test-access qualifications, Validation comparisons, the 57-market exploratory pool, and the 60-market post-hoc sensitivity retain their separate identities. Entity-disjoint splits and matching exact-hash checks do not establish mechanism-OOD generalization or exclude every possible leakage route.

Fixed-support claims use the action/outcome library and selection procedures in Appendix D. E1 and E2 reuse that library, so their agreement is not independent post-selection confirmation. The new E2 choices are also not new market trajectories.

Online service claims use the three completed studies in Appendix E. Historical tool-context effects, official-API format-control effects, online value-interface effects, and fixed-state value effects are not pooled. Their separate endpoints, comparators, confidence intervals, and correction families determine the scope of each reported conclusion.

Completed value-online continuation. The value-online continuation is now included only through its finalized E3 result files and protocol records. Earlier interim or partial status records are not used as scientific results. The E3 inclusion does not retroactively pool its 72 blueprints with the 80-block format-control study, the 206-blueprint historical study, or the fixed-support E2 development results.

Final reproduction boundary. The manuscript distinguishes historical experimental records, completed later reanalyses and API executions, and editorial reorganization. Preparing the text, reading completion reports, and reconstructing figures do not constitute independent experimental replication.

Source digests, software tests, execution manifests, and grouped numerical checks provide complementary evidence within their reported scopes. Hash equality establishes byte correspondence, not correct labels or valid statistical assumptions. Tests apply to their recorded code snapshots, and successful compilation establishes a document build.

Complete end-to-end reproduction would require the remaining data, model, prompt, environment, and analysis dependencies to be available and successfully exercised. That broader claim is not made here.