Online Appendix for
"Returnless Partial Refunds:
A Large-Scale Field Experiment in E-Commerce"

Lu Fang   Chunhua Wu   Zhe Yuan   Kaifu Zhang

A Additional Context Details

Traditional Return vs. Returnless Partial Refund Processes. Figure A1 contrasts the two workflows. Under the traditional return process, customers obtain a shipping label, send the item back, and the platform adjudicates eligibility before issuing a refund. Under the returnless partial refund process, the multi-agent AI system evaluates claims in real time using multimodal evidence. The first decision node is whether a case is eligible for a partial refund; eligible cases receive an immediate offer, which customers may accept (keeping the item with a partial refund) or decline. Both non-acceptance branches—ineligible cases and declined offers—route back to the traditional return process. The returnless partial refund path therefore augments rather than replaces standard returns. This streamlined routing reduces hassle and improves consistency while preserving the traditional pathway as a fallback. In rare cases (less than 1%), a shipped-back item may receive no refund (e.g., when the return is not received or there is a severe policy violation); this reflects exceptional enforcement rather than the standard adjudication outcome.

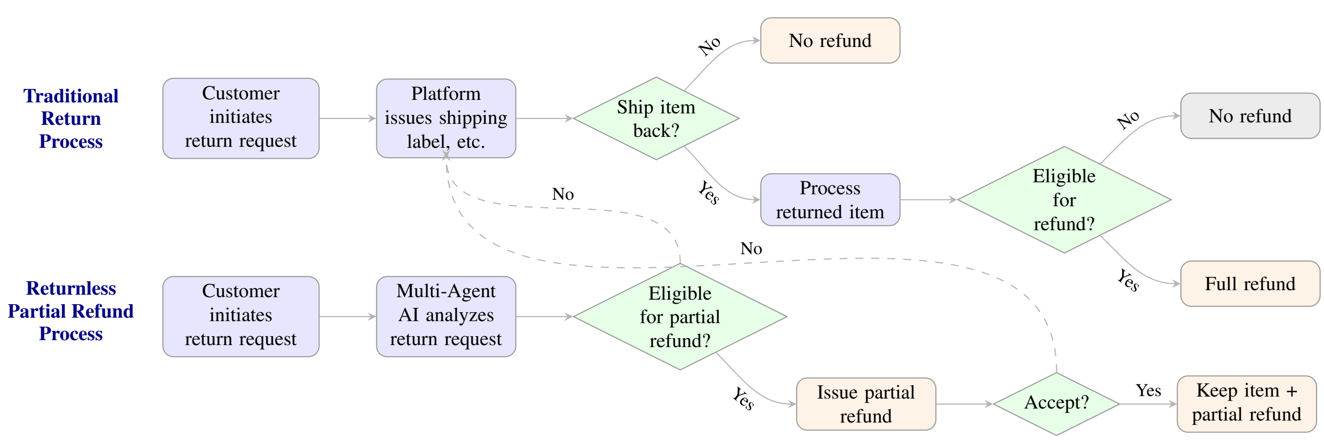


Figure A1: Comparison of Traditional Return and Returnless Partial Refund Processes

Multi-Agent Policy Implementation. The multi-agent AI system in our study represents a real-world field case of agentic AI deployment at scale. The system comprises three specialized AI agents: the *Customer Intent Agent*, the *Evidence Analysis Agent*, and the *Resolution Decision Agent*. The Customer Intent Agent specializes in natural language processing (NLP) to translate and interpret customer statements and classify intent. The Evidence Analysis Agent performs multimodal analysis, comparing customer-submitted visual evidence and textual claims against product information stored on the platform. Finally, the Resolution Decision Agent integrates these outputs to reach a final judgment, covering reason validity, issue severity, and product category.

The full pipeline can be illustrated with a representative case (Figure A2). A customer reported that a purchased white T-shirt was "totally transparent." The Intent Agent comprehended and classified the issue as "color not as described." The Evidence Agent then verified the claim by comparing the customer's image, which confirmed the sheerness, with the opaque appearance in the product photo. Finally, the Resolution Agent synthesized the findings, confirmed the issue, rated it as "medium severity," and issued a 40% partial refund by referencing the predefined mapping table. The resulting offer is presented to the customer through a return interface (Figure 1 in the main paper), allowing them to either accept the offer and keep the item or decline it and proceed with a standard physical return.

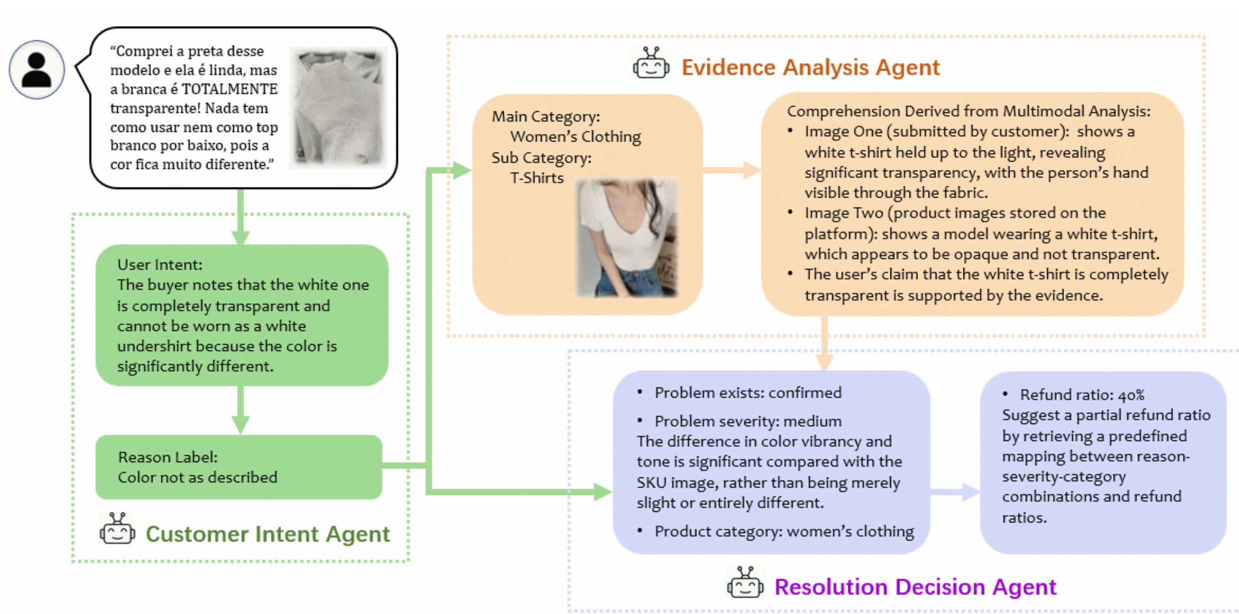


Figure A2: Multi-Agent Implementation of the Returnless Partial Refund Policy

Technical Description of the Multi-Agent AI System. From a technical standpoint, this agentic architecture operates as a sequential pipeline, where each agent functions as a distinct module specialized in a particular task. The output of one agent serves as the input for the next, allowing the system to avoid the inefficiencies of feeding all inputs into a single, general-purpose model, and ensuring modular coordination across the process. Before the experiment began, the agents were guided by carefully engineered prompts, iteratively refined to ensure accurate intent recognition, robust evidence analysis, and consistent decision judgments. The system is built upon and fine-tuned based on the open-source multimodal foundation model Ovis[1]. During the experiment, the system's prompts, parameters and evaluation logic remained fixed, without feedback adaptation from observed customer behavior. This design feature ensures that any systematic behavioral shifts reflect customers adapting to a static policy rather than the AI adapting to them, allowing for a consistent benchmark for return legitimacy and claim quality. Hence, systematic decline in the approval rate tends to stem from deteriorating claim quality due to genuine customer adaptation rather than changes in AI's adjudication standards.

B Parametric Specifications and Empirical Implementation

We estimate linear response functions for the behavioral parameters:

$$r_a^P(\alpha) = \kappa_0^a + \kappa_1^a \alpha, \tag{B.1}$$

$$r_s^P(\alpha) = \kappa_0^s + \kappa_1^s \alpha, \tag{B.2}$$

$$\delta(\alpha) = \kappa_0^\delta + \kappa_1^\delta \alpha, \tag{B.3}$$

where $\kappa_1^a > 0$ (higher refund increases acceptance) and $\kappa_1^s < 0$ (higher refund decreases ship-backs).

Based on 2SLS estimation of Equation (7) in the main paper, Table B1 reports the effect of the initial offer refund level α on the acceptance rate (r_a^P), ship-back rate (r_s^P), and long-term net customer value impact (δ), corresponding to the slope parameters κ_1^a, κ_1^s, and κ_1^δ in Equations (B.1)–(B.3). The estimates confirm the expected signs: a 10-percentage-point increase in the offered refund level raises the acceptance rate by 5.0 percentage points ($\hat{\kappa}_1^a = 0.502$), lowers the ship-back rate by 1.7 percentage points ($\hat{\kappa}_1^s = -0.172$), and increases subsequent 30-day net transaction value by \$0.94 ($\hat{\kappa}_1^\delta = 9.425/LCV$). Refund generosity is thus a strong driver of immediate customer response and has a positive, though noisier, effect on long-term value. The intercepts in Equations (B.1)–(B.3) are recovered from the sample means of the dependent and independent variables:

$$\hat{\kappa}_0^a = \bar{r}_a^P - \hat{\kappa}_1^a \bar{\alpha}, \quad \hat{\kappa}_0^s = \bar{r}_s^P - \hat{\kappa}_1^s \bar{\alpha}, \quad \hat{\kappa}_0^\delta = \bar{\delta} - \hat{\kappa}_1^\delta \bar{\alpha}. \tag{B.4}$$

Substituting the estimated response functions into the platform payoff in Equations (1)–(3) in the main

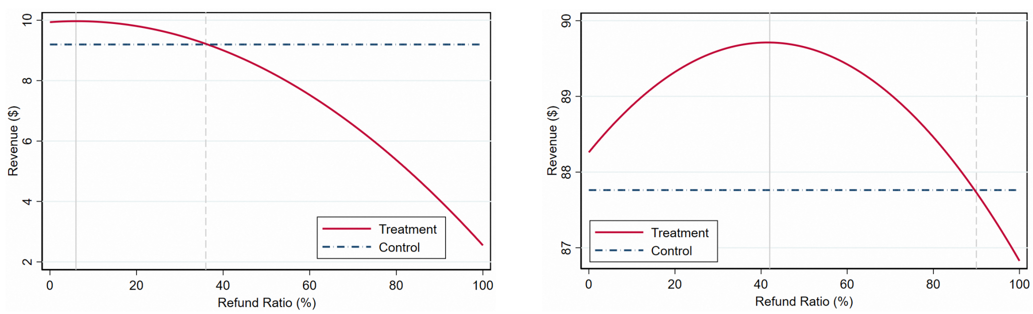
[1]Further details on Ovis are available at: https://github.com/AIDC-AI/Ovis.

Table B1: Immediate and Long-run Effects of Offer Refund Ratio using IV Estimation

	(1) **Acceptance Rate of Initial Offer** r_a^P	(2) **Ship Rate of Initial Return** r_s^P	(3) **Long-term Net Transaction Value** $LCV \cdot \delta$
Initial Offer Refund Ratio (α)	0.502***	-0.172***	9.425*
	(0.009)	(0.010)	(4.868)
Constant	-0.032	0.119***	-140.0***
	(0.033)	(0.037)	(18.91)
Observations	256,382	256,382	256,382
R-squared	0.076	0.075	0.626
Kleibergen-Paap Wald rk F statistic	119,830	119,830	119,830

Note: Standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1. Column 1-3: treatment group who received a returnless partial refund offer in their initial returns, and IV estimation is conducted using high/low period as an instrument. Since the estimation is restricted to the treatment group, each customer's ship rate and net transaction value are demeaned by the corresponding control-group mean within the same return date, return reason, and product category cell to control for seasonality. Logistic regression is also employed to estimate the acceptance rate, producing consistent results. Controls include customer demographics, previous 30-day purchase and return history, and initial return characteristics (reason, product category).

paper yields the platform revenue curves in Figure B1.



(a) Immediate Impact of the Initial Return (b) Incorporating Long-term Customer Value

Figure B1: Revenue Comparisons between Returnless Partial Refund and Traditional Return Processes

Note: Solid and dashed vertical lines mark the revenue-maximizing and break-even refund ratios, respectively.

C Extended 90-Day Analysis

We use the 30-day window as the primary analysis horizon due to its superior experimental design properties: minimal cross-phase contamination, clean causal attribution, and conservative estimation. The extended 90-day analysis serves as supplementary evidence that the policy's positive effects persist beyond the short term. Tables C1, C2, C3, and C4 report the 90-day results for customer spending, key metrics, strategic behavior and the cost-benefit analysis, respectively. Results are qualitatively consistent with the 30-day baseline: the policy continues to enhance customer engagement and generate positive subsequent gains while inducing strategic customer behavior. Aggregating all components, the net transaction value is \$4.41 and the net profit is \$1.23 per customer. Assuming a residual product value of 50% of the transaction value, retained customer value is approximately \$1.89, and sample-wide reduction in physical returns yields carbon emission savings equivalent to the sequestration of roughly 3,000 mature trees.

Table C1: Subsequent Impact of Returnless Partial Refund Policy on Customer Spending (90 Days)

Variables	(1) **Transaction Value**	(2) **Purchase Probability**	(3) **Purchase Frequency**	(4) **Category Expansion**	(5) **Price**
Treatment	4.885***	0.005***	0.332**	0.140***	-0.033
	(1.717)	(0.001)	(0.130)	(0.032)	(0.062)
Constant	-52.73***	0.617***	2.413***	5.770***	14.560***
	(7.129)	(0.001)	(0.866)	(0.068)	(0.084)
% of change	2.1%	0.7%	1.7%	1.3%	-0.2%
Observations	853,521	853,521	853,521	853,521	651,948
R-squared	0.542	0.113	0.363	0.175	0.035

Note: Same specification and definitions as Table 3 in the main paper. 90-day post-return window.

Table C2: Subsequent Impact of Returnless Partial Refund Policy on Key Metrics (90 Days)

Variables	(1) **Transaction Value**	(2) **Refund Amount**	(3) **Processing Cost**	(4) **Net Transaction Value (LCV)**	(5) **Order Return Rate**
Treatment	4.885***	0.269	-0.074**	4.690***	0.002***
	(1.717)	(0.179)	(0.032)	(1.668)	(0.0001)
Constant	-52.73***	6.331***	1.080***	-60.14***	0.099***
	(7.129)	(0.433)	(0.097)	(7.014)	(0.0001)
% of change	2.1%	1.7%	-2.8%	2.2%	3.1%
Observations	853,521	853,521	853,521	853,521	17,078,759
R-squared	0.542	0.108	0.101	0.542	0.091

Note: Same specification and definitions as Table 5 in the main paper. 90-day post-return window.

D Stability of Experimental Entry by Assignment

A potential concern in our experimental design is that information spillovers may contaminate the treatment-control comparison. For example, control customers may learn about the returnless partial refund policy through word-of-mouth or other channels and strategically increase their return activity, resulting in differential entry into the experiment relative to the treatment. Given that customers were unaware of the new policy before initiating their first return during the experimental period, their entry into the experiment, triggered by this return event, should be random with respect to treatment assignment. Accordingly, the share of control group customers among all experimental entrants should remain stable over time. Figure D1 plots this share by entry date (i.e., initial return date) and shows it remains stable at approximately 50% throughout the study. Consistent with this, we detect no significant differences in pre-experiment demographics and behaviors between the groups (Table D1), nor in the transaction value of the initial return (Column 1 of Table 2 in the main paper). Together, these findings support the validity of the randomization in experimental

Table C3: Offer Refund Ratio and Strategic Behavior using IV Estimation (90 Days)

Variables	(1) **Number of Subsequent Returns**	(2) **Order Return Rate**	(3) **Approval Rate**	(4) **Ship Rate Conditional On No Offer**
Initial Offer Refund Ratio (α)	0.152***	0.009***	-0.055***	-0.189***
	(0.017)	(0.001)	(0.015)	(0.048)
Constant	-0.112**	0.078***	0.738***	0.833***
	(0.045)	(0.003)	(0.060)	(0.187)
Observations	256,382	5,369,776	49,624	8,692
R-squared	0.030	0.057	0.023	0.124
Kleibergen-Paap Wald rk F statistic	119,830	1,662,054	23,198	4,481

Note: Same specification and definitions as Table 7 in the main paper. 90-day post-return window.

Table C4: Cost-Benefit Analysis (Per Customer Within 90 Days of Initial Return Request)

Metric ($)	**Difference**	**% Change**	**p-value**
(1) Costs:			
Increased Refund Amount:			
– Initial Return Request	(0.63)	(4.2%)	<0.01
– Subsequent Return Requests	(0.27)	(1.7%)	0.132
Reduced Processing Cost:			
– Initial Return Request	0.35	11.8%	<0.01
– Subsequent Return Requests	0.07	2.8%	0.02
(2) Benefits:			
Increased Revenue:			
– Transaction Value of Initial Return Request	0.00	0.0%	0.95
– Transaction Value of Subsequent Purchases	4.89	2.1%	<0.01
(3) Platform Net Transaction Value Impact	4.41	2.0%	<0.01
Back-Of-The-Envelope Calculation			
(4) Platform Net Profit Impact (35% Margin)	1.23	2.2%	
(5) Customer Retained Value	1.89	17.2%	
(6) Total Carbon Footprint Saving	3000 trees		

Note: Same specification and definitions as Table 8 in the main paper.

entry.

Table D1: Covariate Balance and Randomization Check

Variable	**Control Group** (N = 427,943)	**Treatment Group** (N = 425,578)	**P-value of Difference**
Customer Characteristics			
Gender	1.000 (0.001)	1.001 (0.001)	0.622
Age Tier	1.000 (0.0004)	0.999 (0.0004)	0.223
Registered Years	1.000 (0.001)	1.001 (0.001)	0.563
Country Index	1.000 (0.001)	1.001 (0.001)	0.261
Language Index	1.000 (0.001)	0.998 (0.001)	0.163
Behavior in Past 30 Days			
Transaction Value ($)	1.000 (0.005)	1.004 (0.005)	0.564
Refund Amount ($)	1.000 (0.009)	0.999 (0.010)	0.912
Processing Cost ($)	1.000 (0.011)	1.003 (0.012)	0.835
Net Transaction Value ($)	1.000 (0.005)	1.004 (0.005)	0.556

Note: For each variable, the mean and standard error in parentheses are shown for both the control and treatment groups, with all values normalized by dividing by the variable's control group mean. P-values from t-tests confirm successful randomization. Gender: platform-predicted (1=male). Age Tier: 7-point scale from 1 (youngest) to 7 (oldest). All financial metrics from 30 days pre-experiment. Net Transaction Value = Transaction Value - Refund Amount - Processing Cost.

Figure D1: Share of Control Group Customers among All Customers Over Time (%)