Misperceptions and Product Choice: Evidence from a Randomized Trial in Zambia*

Jie Bai David Sungho Park Ajay Shenoy†

August 2, 2025

Abstract

Do misperceptions deter small retailers from stocking profitable products? We apply unsupervised machine learning to the inventories of 2000 shops in Lusaka, Zambia, to identify clusters of similar retailers and flag a "target" product—one stocked by some but not all shops in a cluster. Retailers that do not stock the product systematically underestimate its profitability. We test whether misperceptions drive this gap using a randomized trial that temporarily subsidized stocking the product. Treated retailers earn similar profits to peers already selling the target product. They continue stocking it after incentives end, especially when given a longer period to experiment.

JEL Classifications: D22, D83, L26, M21, O12, O33

Keywords: misperception, firm behavior, retail markets, micro and small enterprises

*Bai: Harvard Kennedy School (jie_bai@hks.harvard.edu); Park: KDI School of Public Policy and Management (park@kdis.ac.kr); Shenoy: University of California, Santa Cruz (azshenoy@ucsc.edu). For helpful discussions and comments, we thank seminar participants at the UCSC brown bag workshop. For organizing the data collection, we thank Chisenga Chimenge at BEMAC. We are extremely grateful to all the enumerators who collected this data, though there are too many to list individually. Funding for this project was provided by the Private Enterprise Development in Low-Income Countries (PEDL) and the KDI School of Public Policy and Management. This study has been approved by the UC Santa Cruz Institutional Review Board (Protocol HS-FY2022-265). Trial is registered at AEA RCT Registry under AEARCTR-0011827. Any error is our own.

†Corresponding author

1 Introduction

Firms often face information constraints that lead to suboptimal decisions and hinder growth. Even large, formal firms and exporters can make costly mistakes that result in inefficient outcomes (Dickstein and Morales, 2018; Tanaka et al., 2020). These challenges are likely even more severe for the vast number of small informal firms in developing countries, which account for the bulk of employment (Hsieh and Olken, 2014; Lagakos, 2016). These firms may lack even basic information about local demand, leading to inaccurate beliefs and unprofitable decisions. Providing basic, actionable information, or facilitating learning and information discovery among firms, has the potential to improve decision-making and raise incomes for millions of small entrepreneurs.

This paper examines the role of inaccurate beliefs in shaping a basic but consequential decision faced by nearly all firms: which products to stock. While this choice matters in any context, it is particularly salient for small and micro retailers, who operate under tight constraints and limited information. We combine rich field data with statistical learning tools and a randomized controlled trial to test whether misperceptions about product profitability deter these firms from stocking high-return items.

To systematically identify under-stocked yet potentially profitable products, we collect full-inventory data from nearly 2,000 micro-retailers across 25 retail markets in Lusaka, Zambia. Enumerators photographed each shop's inventory, and these images were hand-coded into machine-readable data. We use a clustering algorithm to group shops with similar products into clusters. Within each cluster we identify a "target" product that was stocked by some but not all shops. A follow-up survey gathered data on whether the shop stocked the target product, the actual retail and wholesale prices (if stocked), and beliefs about those prices (if not). Several key patterns emerge. First, many shops do not stock the target product, even though nearby peers do. Second, among shops that do stock it, the target product tends to yield higher per-unit profits than their main goods (defined as the top three by revenue share). In contrast, shops that do not stock the target product systematically underestimate its profitability. Their beliefs about expected profits are lower than the realized profits reported by those that stock it.

While the descriptive evidence is consistent with our hypothesis that shops misperceive the profitability of the target product, it is only suggestive. An alternative explanation is rational sorting: shops that stock the target product face unobserved demand or supply conditions that make the product more profitable for them, while non-stocking shops face different conditions that make the product less profitable. In this case, beliefs may be accurate, and shops that currently do not stock the product would, if induced to do so, earn lower profits than those who choose to stock it voluntarily.

To distinguish between these explanations, we conduct a randomized controlled trial that exogenously induced shops to stock the target product. The intervention offered two components: a recommendation to stock a cluster-specific product that was commonly sold by peers, and a reimbursement subsidy to lower the upfront cost of trying it. The goal was to create a low-risk opportunity for shops to experiment and learn about the product's profitability. We split a sample of roughly 300 micro-retailers into a control group, which received neither the recommendation nor the subsidy, and two treatment groups. Treated shops received both the recommendation and a subsidy covering the cost of the product for either one or two weeks. The two randomly assigned treatment durations allow us to test whether a longer period of subsidized experimentation is more effective at generating persistent adoption. Critically, treated shops were informed that they could discontinue the product at any time and were encouraged to act in the best interest of their business. This design enables us to observe both initial take-up and medium-run retention in response to the randomized subsidy, and to isolate the causal effect of stocking the target product.

We track stocking behavior during and after the reimbursement period using both follow-up phone surveys and visits by unidentified "mystery shoppers." Initial take-up among treated shops ranged from 10 to 45 percent, depending on treatment arm and measurement method. Importantly, for shops given two weeks of reimbursement, stocking persisted beyond the subsidy period. By the end of the two-month study window, these shops were 37 percent more likely than control shops to be stocking the product. In contrast, shops given only one week of reimbursement stocked at rates statistically indistinguishable from the control group. These results suggest that a longer exposure period allowed shops to experiment, revise beliefs, and discover the product's profitability.

Consistent with misperception rather than rational sorting, we find that shops induced to stock the product through the subsidy earned similar per-unit profits and markups to those in the control group who chose to stock the product on their own. Among treated shops, realized profits were also similar regardless of their initial beliefs about the product's profitability. Furthermore, stocking the target product did not crowd out sales of main goods. Treated shops continued to sell similar volumes of their main products. In fact, based on an independent survey conducted by a separate team, we find that shops in the two-week treatment arm saw more customers than control shops, suggesting that expanding product variety may have attracted additional demand. We also find suggestive evidence of increased overall sales and profits at the store level. Taken together, these findings are difficult to reconcile with rational sorting or accurate beliefs. Instead, they point to a role for misperceptions in driving suboptimal stocking decisions among small retailers, and show that even a brief opportunity to learn through experience can meaningfully shift behavior and improve firm performance.

We further consider two alternative explanations for the persistence in stocking. One is a strategic response to market competition: the subsidy may have induced some shops to stock the product, while prompting nearby control shops to exit or avoid entry in response to increased competition. If so, the profits observed among treated shops during the intervention may not reflect the counterfactual profits they would have earned had they entered without displacing others. In this view, the RCT would then simply reshuffle which shops stock the product without increasing overall adoption. If this mechanism were at play, we would expect control shops to be less likely to stock the product in areas with higher treatment saturation. In practice, we find no such effect: if anything, stocking rates among control shops slightly increase with the share of nearby treated shops. A second alternative explanation is that the subsidy helps shops overcome fixed search costs in sourcing the product. This too is inconsistent with the data: most non-stocking shops report they could source the product from markets they already frequent. These findings support the interpretation that the treatment effect reflects belief updating, rather than changes in market competition or reductions in fixed costs of search and sourcing.

Why, then, do so many retailers hold inaccurate beliefs about these relatively common products, especially when peers nearby are already stocking them? One explanation is self-confirming beliefs: firms that never stock the product have limited opportunities to learn, and their beliefs remain misaligned with actual outcomes. Indeed, we find that among non-stocking shops, expectations about prices and markups are largely uncorrelated with realized outcomes, while stocking shops' predictions show a significantly stronger correlation with actual values. Moreover, shops that say they do not stock the product because it is unprofitable ultimately earn similar profits to those who expressed no such concern, suggesting that beliefs about profitability may be poorly informed.

A second challenge is that learning is inherently noisy in this environment: product-level sales and profitability fluctuate over time, making it difficult to infer profitability from limited experience. Even among shops that always stock the product, we observe meaningful variation in sales from week to week. This noisy environment helps explain why two weeks of experimentation, rather than one, was needed to shift behavior: longer exposure allows shops to accumulate more signals and update their beliefs more confidently. More broadly, these findings suggest that in environments where firms operate with limited information, face uncertain profitability, and where external advisors may not possess better knowledge than the firms themselves, facilitating low-cost experimentation can be an effective policy tool. Subsidizing initial opportunities to try new products—or, more generally, to test new business strategies—may help small firms overcome misperceptions, identify profitable opportunities, and make better-informed decisions.

This paper contributes to a growing literature on information frictions and firm decision-making. Much of this work has focused on how information frictions in search and matching with buyers and suppliers distort firm behavior and market outcomes, particularly in developing country contexts (Jensen, 2010; Aker, 2010; Bergquist et al., 2024; Cai et al., 2024). More recent studies document that even large, formal firms can make costly mistakes due to imperfect information, for example when entering new export markets (Dickstein and Morales, 2018). We focus instead on small, informal firms operating in a much more basic, localized setting: choosing which products to stock. Compared to larger firms, these retailers face tighter constraints and may lack even the most basic information about local demand and profitability. We show that such informational gaps can lead to systematically inaccurate beliefs and, in turn, missed profit opportunities.

Even when information is available, learning can be slow and difficult in noisy environments. The prior literature, set mainly in agriculture, has shown that the adoption of new products and technologies is hampered not only by the lack of exposure but also by the difficulty of inferring outcomes from limited and variable experience (Conley and Udry, 2010; Bold et al., 2017). Firms may also struggle to process the signals they receive or apply them effectively (Hanna et al., 2014). These challenges help explain why interventions that enable direct, low-cost experimentation, rather than abstract third-party advice, often lead to more persistent behavioral change. In our setting, simple stocking suggestions combined with modest subsidies to nudge experimentation were enough to shift beliefs and generate persistent take-up, suggesting that learning through experience can be a powerful tool when information is hard to interpret or act upon.

Our study also contributes to a broader literature on scalable interventions to improve firm performance. While intensive support such as consulting or mentorship can lead to large gains for some businesses (Bloom et al., 2013; Brooks et al., 2018; Bruhn et al., 2018), standard business training programs that focus on general skills tend to have more modest and variable impacts across contexts (McKenzie and Woodruff, 2017). A growing body of evidence suggests that concrete and actionable guidance, whether in the form of simple rules of thumb (Drexler et al., 2014), locally tailored advice (Dalton et al., 2021), or highly specific sourcing recommendations (Brooks et al., 2018), can outperform traditional business training. Our findings align with this pattern, showing that concrete, product-level information can meaningfully shift behavior. At the same time, context and design matter. For example, Banerjee et al. (2022) study a non-randomized, market-level intervention in fruit markets in India and find that shops quickly cease stocking a subsidized product once incentives end, possibly due to social norms around collusion. In contrast, we randomize treatment at the individual firm level and observe persistent take-up after the subsidy is withdrawn, along with no evidence that treatment discourages stocking among nearby shops. These contrasting results highlight the importance of tailoring information interventions to the institutional and market context.

2 Identifying Potential Suboptimality in Stocking

2.1 Digitizing and Clustering Shop Inventories

Our study focuses on small retailers operating in Lusaka, Zambia, originally sampled in a prior study (Samaniego de la Parra and Shenoy, 2024). In September–October 2022, we conducted an inventory survey across 2,000 retailers in 25 markets.[1] As a novel feature of our data collection, enumerators photographed the full inventories of each shop, creating a visual census of firm-level product choices (Figure 1). These images were hand-coded into structured inventory data: each product in each photo was independently classified by two enumerators, with discrepancies resolved by a study coordinator.

Figure 1: Sample Inventory Images

Note: Images taken at the inventory survey in September-October 2022.

[1]A separate research team, as part of (Samaniego de la Parra and Shenoy, 2024), conducted a comprehensive census to map out all retail establishments in 25 major markets in Lusaka, identifying approximately 3,000 shops. For our study, we focused on retailers operating in food or cosmetics/pharmacy sectors, which reduced the sample to around 2,000 shops.

Figure 2: Digitizing Inventories

47131811 - Laundry products
53131602 - Hair care supplies
53131607 - Hand or body lotion or oil
50161509 - Natural sugars or sweetening products
50151513 - Edible vegetable or plant oils
50171551 - Cooking or table salt
53131641 - Petroleum jelly
50221300 - Flour and milled products
51171608 - Glycerine
50192902 - Shelf stable plain pasta or noodles
53102305 - Infant diapers
50193103 - Gravy mix
50181709 - Baking supplies
50201714 - Non dairy creamers
50201709 - Instant coffee
91101601 - Facial or body treatments
26111700 - Batteries and cells and accessories
53131619 - Cosmetics
50201710 - Leaf tea
12131706 - Matches
24111503 - Plastic bags
…

stock1	0
stock2	0
…	
stock12	1
stock13	0
…	
stock15	0
stock30	1
stock31	0
stock32	0
stock33	0
stock34	0
stock35	0
stock36	1

com1	3.426393
com2	-1.07705
com3	-0.66823
com4	-0.09338
com5	-0.90672

Note: The figure illustrates our procedure for digitizing the inventories, based on images collected from roughly 2,000 retailers across 25 markets in Lusaka in September-October 2022. First, from the photographs of each shop's inventory, the enumerators identified individual products appearing in each image using harmonized product codes. To minimize measurement error, two enumerators independently coded each image, with any discrepancies adjudicated by a study coordinator. Each shop's inventory was represented as a binary "product vector," where each entry indicated whether a given product was stocked or not. Given the high dimensionality of the product space relative to the sample size, we applied principal component analysis (PCA) to reduce these vectors into a five-dimensional space. Each shop was then represented as a point in this reduced vector space, where proximity reflected similarity in product mix. Using k-means clustering, we grouped shops into clusters by finding centroids that minimized the average distance between shops and their nearest centroid. This procedure initially identified ten clusters, of which two residual groups were excluded due to lacking a common product stocked by the majority. Ultimately, the analysis focused on eight well-defined clusters.

To identify comparable retailers and systematically uncover products that are stocked by some—but not all—shops, we applied a clustering algorithm. Given the high dimensionality of the product space, we first extracted the first five principal components of the inventory data. Each shop was then represented as a point in this reduced vector space, where proximity reflects similarity in product mix. Figure 2 visualizes how we digitized the shop inventories. We applied k-means clustering to group shops with similar product mix, ultimately identifying eight well-defined clusters (excluding two residual clusters with no shared products).[2]

Within each cluster, we selected one "target" product—defined as a product stocked by approximately 30–40 percent of shops in the group—yielding a natural setting to study variation in adoption and perceived profitability (Table 1). While a small number of shops could not be assigned to any cluster, the final clusters cover approximately 1,000 shops, each with a clearly defined peer group and target product.

This approach combines rich visual inventory data with unsupervised learning to systematically group shops with similar product mix and identify products that are commonly stocked by some but not all peers. By doing so, it allows us to identify potentially under-adopted products in a structured, scalable way. These "target" products—defined relative to a shop's peer cluster—form the basis of our survey and experimental design. To our knowledge, this is one of the first applications of machine learning to analyze microenterprise inventories and to inform the design of a randomized field experiment by leveraging information embedded in firms' actual stocking behavior.

2.2 Stylized Facts

From January to March 2023, we conducted a baseline survey with roughly 1,000 shops across the eight clusters identified through the clustering algorithm discussed above. For each shop, we collected data on whether they stocked the target product, the retail prices charged and order prices paid (if stocked), and beliefs about those prices (if not stocked). From this data, we document two stylized facts:

[2] The algorithm initially identified ten clusters, but two were incoherent "everything else" groups in which no product was stocked by a majority of shops, and were dropped from subsequent analysis.

Fact 1: Many shops do not stock the target product.

As shown in Table 1, the target product was stocked by approximately 30–40 percent of the shops within each cluste, despite otherwise similar product mixes. For example, among shops that mostly sell cosmetics (Cluster 1), only 31% stock deodorants; among shops that mainly sell vegetables (Cluster 2), 40% stock onions.

Table 1: Shop Clusters and Target Products

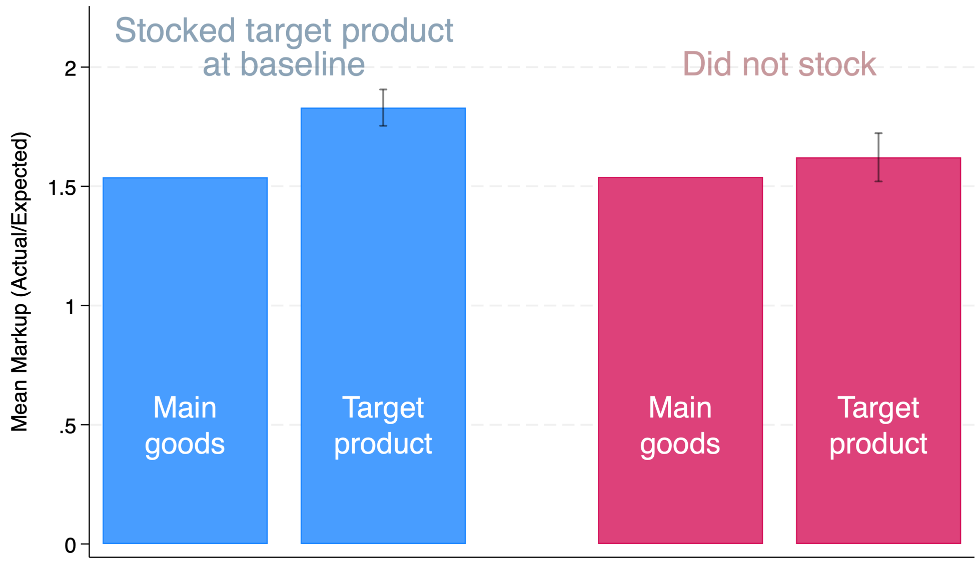
Cluster No.	Common Goods within Cluster	Target Product
1	Nail polish (73%), Hair comb/brush (71%), Makeup kits (71%)	Deodorants (31%)
2	Tomatoes (56%), Onions (40%), Nominant vegetables (38%)	Onions (40%)
3	Dried beans (58%), Shelf stable fish (48%), Mealed vegetable (45%)	Flour and milled products (30%)
4	Laundry products (98%), Shelf stable juice (90%), Sweet biscuits/cookies (85%)	Exercise books (31%)
5	Hair care supplies (55%), Hand or body lotion/oil (46%), Perfumes/colognes/fragrances (37%)	Creams/lotions for common skin ailments (30%)
6	(99%) Laundry products, (99%) Shelf stable juice, (98%) Hand or body lotion/oil	Powdered drink mix (30%)
7	Laundry products (73%), Sugar/sweetening products (60%), Sweet biscuits/-cookies (57%)	Sport or energy drink (30%)
8	Onions (61%), Tomatoes (58%), Peppers (53%)	Ginger root (31%)

Note: The table summarizes the eight identified clusters of shops, derived from the clustering analysis described in Figure 2. For each cluster, Column (2) reports the three most commonly stocked products, along with the percentage of shops in that cluster stocking each product (in parentheses). Column (3) identifies the “target product” for each cluster, defined as a product stocked by approximately 30–40% of shops in the respective cluster. Stocking percentages were based on the inventory survey in September-October 2022.

Fact 2: Shops that stock the target product earn higher per-unit profits on it than on their main goods, while those that do not stock it underestimate its profitability.

As shown in Figure 3, shops that stock the target product report higher per-unit markups on it than on their three main goods (defined as the top three revenue-generating products). In contrast, shops that do not stock the product expect it to yield profits similar to their main goods—substantially below the realized margins reported by the stocking shops. This gap between expected and actual profitability suggests a systematic misalignment in beliefs among non-stocking retailers, potentially deterring them from stocking high-return products.

Figure 3: Markups of Target Product and Main Products



Note: The bars show the average markups for main goods and the target product, separately for two groups: shops that stocked the target product at baseline (left panel), and shops that did not stock the target product at baseline (right panel). Markup is calculated as the selling price per unit divided by the purchase cost per unit. Since selling and purchasing units often differed, we recorded conversion rates (the number of selling units per purchasing unit) to compute per-unit prices and subsequently derive the markup. For shops not stocking the target product at baseline, the figure shows "expected" markups, derived from hypothetical questions regarding expected selling and buying prices if they were to stock the target product. Data from the baseline survey (January-March 2023).

2.3 Misperception or Rational Sorting?

The descriptive patterns above are consistent with firms holding inaccurate beliefs about the profitability of the target product. However, they are also consistent with a full-information, rational sorting explanation, in which potential profits vary across firms, and shops optimally decide whether to stock the product based on their own profit conditions.

Under this alternative explanation, non-stocking shops correctly anticipate lower returns, perhaps due to weaker local demand, different customer preferences, or higher stocking costs. In that case, reported beliefs are accurate, and no firm is making a mistake: those who stock earn high profits, while those who do not would earn less if they did. If this is true, then the gap between beliefs and realized profits simply reflects underlying heterogeneity, not misperception. A further possibility is that stocking the target product could displace inventory or sales of existing products, reducing total profit despite the product being profitable on its own—again, consistent with firms making informed, optimal decisions to avoid stocking.

This distinction has important policy implications. If firms are already optimizing given accurate beliefs, then providing new information or subsidizing product trials should have little effect. But if misperceptions are the barrier, then even brief exposure to a product could shift beliefs and lead to persistent changes in behavior.

To distinguish between these explanations, we design a randomized controlled trial that induces a subset of non-stocking shops to begin stocking their cluster-specific target product. If misperceptions are the binding constraint, we should observe lasting changes in behavior after the subsidy ends, as firms update their beliefs through experience. If instead firms are fully informed, any uptake should be short-lived, and treated shops should earn lower profits than those who stocked voluntarily.

3 Experimental Design and Data

3.1 Experimental Design

We implemented the experiment with 271 small retailers who reported not stocking the target product at baseline. Each shop was assigned a cluster-specific target product using the procedure described in Section 2.1, and was then randomly assigned—with equal probability—to one of three groups: a one-week subsidy treatment, a two-week subsidy treatment, or a control group. Randomization was stratified by product cluster.

In the one-week treatment, shops were offered a fixed reimbursement for purchasing and stocking one or two units of their assigned target product during a specific one-week window.

Shops were instructed to purchase a defined unit (such as a 10 kg bag of onions or a case of 12 energy drinks) and received a fixed payment per unit, up to a maximum amount. The reimbursement was calibrated to cover approximately 100% of the purchase cost. Study staff scheduled a time to verify that the product had been purchased and was in stock.

The two-week treatment followed the same protocol, but extended over two consecutive weeks. To ensure that shops did not simply hold the same stock across both weeks, staff marked the product in the first week and verified that a new unit had been purchased in the second. The comparison of one- and two-week treatments allows us to test whether longer exposure improves learning and leads to more persistent adoption.

All treated shops received the following instruction from enumerators at the outset of the reimbursement period. This framing was designed to minimize experimenter demand and encourage shops to treat the stocking decision as they would for any other business decision.

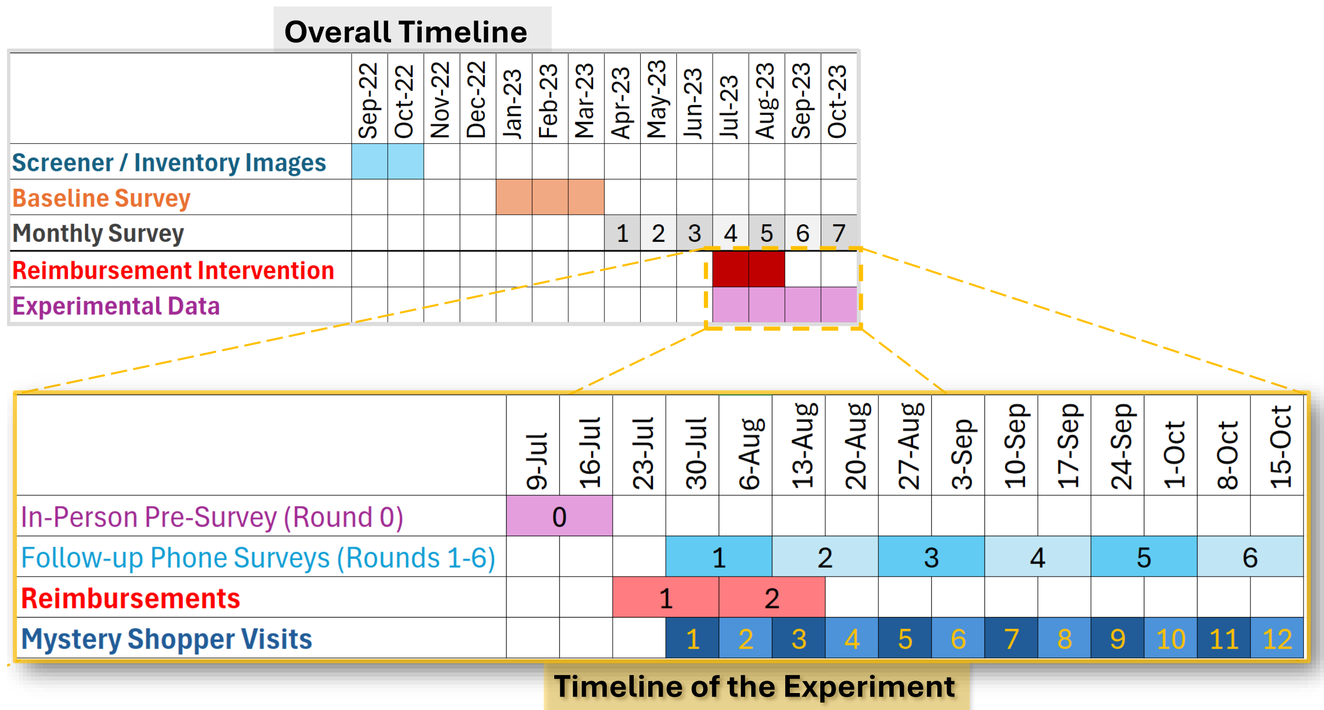
> *. . . you may continue to stock the product at your own expense if you believe it is profitable, or you may stop stocking the product if it is not. We encourage you to do whatever is best for your business. It will not affect your eligibility for any future study.*

3.2 Data

Figure 4 summarizes the data collection timeline and the roll-out of the experimental treatments. The two pre-experiment data sources—the inventory survey and the baseline survey—were completed several months prior to the intervention and are described in detail in Section 2. The experimental sample was drawn from the subset of the baseline survey that (1) was assigned to a product cluster and corresponding target product, and (2) did not stock the target product at baseline. This includes roughly 300 shops; of these, 271 consented to participate in the experiment.

For the 271 shops in the experimental sample, we collected three main types of outcome data before, during and after the intervention: a pre-treatment in-person survey, six rounds of biweekly follow-up phone surveys, and twelve weekly in-person audits by unannounced "mystery shoppers." The timing of each activity is summarized in Figure 4.

Figure 4: Study Timeline



Note: The figure shows the timeline for the intervention and survey data collection. Dates refer to when the vast majority of cases were completed. For all biweekly stages during the experiment (Survey Rounds 0 through 6, and the reimbursements) the sample was divided in half, with one half done in the first week and the second half done in the second week. This odd-even pattern was maintained throughout the study to ensure a consistent gap between survey rounds and between reimbursements.

The pre-treatment survey ("pre-survey" or Round 0) was conducted in person in July 2023, prior to any treatment assignment. At the end of the survey, shops assigned to one of the treatment groups were informed about the reimbursement offer. The survey collected detailed information on stocking behavior, prices, and expectations for the target product, along with two "decoy" products commonly stocked by other shops in the same cluster. The inclusion of decoy products was designed to reduce the chance that shops would start (or stop) stocking the target product simply because we were asking about it. For each product, we asked whether the shop currently stocked it, the supplier and customer prices, and—regardless of current stocking status—expectations about stocking, buying, and selling over the next two weeks. To reduce cognitive burden, non-stocking shops were asked to report expected prices and quantities for a "similar shop" that does stock the product, rather than answer hypothetical questions about their own behavior.

Following the pre-survey and the reimbursement offer, the biweekly phone surveys were administered every two weeks to collect follow-up data on stocking behavior and expectations. These surveys largely mirrored the structure of the pre-survey, with questions on the target and decoy products. To manage field logistics and maintain a consistent gap between rounds, the sample was split in half and surveyed on an alternating two-week cycle.

To verify self-reported stocking behavior in the biweekly phone survey, we conducted weekly in-person audits using "mystery shoppers." These were carried out by field staff with no prior interactions with the shops and no affiliation with the survey team. At no point during their assignment did they reveal their affiliation with the research project. Each shop was visited once per week. Mystery shoppers first attempted to locate the product independently; if it was not visible, they asked whether it was in stock. Once confirmed, they were instructed to remember the product's location and avoid further direct questioning in future visits. These audits provide a more conservative measure of stocking outcomes, mitigating concerns about social desirability or experimenter demand in self-reports (for instance, respondents might overstate stocking due to perceived expectations, despite being explicitly told they could stop stocking after the reimbursement period).

Finally, we supplement our core data with a parallel monthly high-frequency phone survey conducted by a separate enumerator team as part of Samaniego de la Parra and Shenoy

(2024). This survey provides shop-level outcomes, including daily sales, cost of goods sold, and customer counts.[3] The data allows us to compute average daily outcomes for each month, including customer traffic, total sales, and profits.[4] The monthly survey began shortly after the baseline survey and continued through the post-treatment period, spanning eight months in total. Although the monthly cycle does not align exactly with the biweekly treatment follow-up schedule, we map survey months as closely as possible to experimental rounds to construct pre- and post-treatment measures of shop-level outcomes.

3.3 Summary Statistics and Randomization Check

Table 2 reports summary statistics of baseline demographics and shop characteristics, along with balance checks across experimental arms. Panel A shows that the average respondent is 41 years old and 69% are female. Just over half are married, with an average of 2.35 children in the household. In the control group, 29% have completed secondary school and 61% are literate in English. Mobile phone access is nearly universal (94%), and 62% own a smartphone. About 16% operate another business in addition to the shop. Panel B summarizes shop-level characteristics. Shops have operated for an average of 14 years and are open approximately 75 hours per week. The average display area is 9.2 square meters, and 33% have dedicated storage space. About 38% are connected to electricity, and the average value of business assets is $393. Shops employ 0.88 workers on average, serve about 50 customers per day, and report weekly revenue of $189. 20% report a stockout in the past week, and 31% report having regular suppliers. By design, none of these shops stocked the target product at baseline, though 12% report having stocked it in the past.

Overall, treatment and control groups are well balanced: only 2 out of 42 coefficients are statistically significant at the 5% or 10% levels, as shown in Columns 2 and 3 of Table 2. Table A1 shows no differential attribution across treatment groups in the biweekly phone survey and the mystery shopper survey.

[3]The monthly survey includes both the experimental sample and roughly 450 additional shops. Shops were divided into four rotation groups and surveyed one week per month on a fixed schedule. Each day during their assigned week, they answered a brief phone survey about daily operations. If unreachable, enumerators recovered missed data during the next successful call or through an in-person visit if the shop could not be reached for the entire week.

[4]A "customer" is defined as anyone who enters the store, regardless of whether they make a purchase.

Table 2: Baseline Summary Statistics and Randomization Balance Checks

	(1) Control Mean	(2) 1-Week Subsidy - Control	(3) 2-Week Subsidy - Control
Panel A. Demographics of shop owner/manager			
=1 if female	0.69	-0.06	-0.05
		(0.07)	(0.07)
Age	40.93	-3.09*	-1.70
	[12.56]	(1.75)	(1.89)
=1 if married	0.55	-0.00	-0.06
		(0.07)	(0.08)
Number of children living together	2.35	0.28	-0.00
	[1.77]	(0.26)	(0.29)
=1 if completed secondary school	0.29	-0.01	0.04
		(0.07)	(0.07)
=1 if literate in English	0.61	-0.02	0.02
		(0.07)	(0.07)
=1 if has mobile phone	0.94	0.02	-0.00
		(0.03)	(0.04)
=1 if has smartphone	0.62	0.10	0.01
		(0.07)	(0.08)
=1 if operates another business	0.16	0.09	0.06
		(0.06)	(0.06)
Panel B. Shop characteristics			
Years of operation	13.74	-1.84	-2.11
	[10.92]	(1.51)	(1.63)
Number of hours a week the shop is open	74.53	1.94	0.18
	[16.60]	(2.28)	(2.41)
Display area (squared meters)	9.18	0.40	36.49
	[11.02]	(1.81)	(35.64)
=1 if has storage space	0.33	-0.05	-0.07
		(0.07)	(0.07)
=1 if connected to electricity	0.38	-0.02	0.04
		(0.07)	(0.07)
Value of business assets (USD)	393.44	144.20	61.76
	[949.77]	(168.97)	(162.69)
Number of employees	0.88	-0.09	-0.14
	[1.09]	(0.17)	(0.15)
Number of customers a day	50.73	8.26	9.17
	[42.65]	(6.67)	(6.56)
Revenue (past week, USD)	189.40	-7.05	41.00
	[269.09]	(41.43)	(44.85)
=1 if any stockout in past week	0.20	0.05	0.02
		(0.06)	(0.06)
=1 if has regular supplier	0.31	0.01	0.11
		(0.07)	(0.07)
=1 if has stocked the niche product earlier	0.12	0.02	0.12**
		(0.05)	(0.06)
Observations	90	97	84

Note: Column 1 reports means and standard deviations for the control group; Columns 2 and 3 show differences for the treatment groups, with standard errors in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively. Data from the baseline survey (January-March 2023).

3.4 Regression Specifications

We estimate the impact of the intervention on the adoption of the target product in each subsequent round by estimating:

$$Y_{i(c)t} = \sum_{s} \beta_s [Treat\ 1week]_{i(c)} D_{st} + \sum_{s} \gamma_t [Treat\ 2week]_{i(c)} D_{st} + \delta Y_{i(c)0} + \phi_{ct} + \varepsilon_{i(c)t} \quad (1)$$

where $Y_{i(c)t}$ is the outcome for shop i in cluster c at time t. The index t denotes survey rounds for biweekly phone data and weeks for weekly mystery shopper data. $[Treat\ 1week]_{i(c)}$ and $[Treat\ 2week]_{i(c)}$ are indicators for assignment to the 1-week and 2-week subsidy groups, respectively. D_{st} is a binary variable equal to 1 if observation t falls within survey round $s = 1, 2, \ldots, 6$. $Y_{i(c)0}$ denotes the pre-treatment value of the outcome (measured at the pre-survey).[5] Because the mystery shopper data lack pre-treatment observations, we use the pre-treatment in-person survey response as $Y_{i(c)0}$ in Equation 1.[6] ϕ_{ct} is a fixed effect for cluster c interacted with period t, accounting for the stratified randomization by cluster. Standard errors are clustered at store level for all regressions unless otherwise mentioned. In the dynamic specification, we map weekly mystery shopper data to the corresponding two-week survey rounds. Although the pooled regression is estimated at the shop-week level, the dynamic treatment dummies β_s are defined by survey round to ensure comparability with estimates from the phone survey.

We estimate the average post-treatment impact in two panel regressions. The first is comparable to Equation 1 without the dynamic coefficients:

$$Y_{i(c)t} = \beta [Treat\_1week]_{i(c)} + \gamma [Treat\_2week]_{i(c)} + \delta Y_{i(c)0} + \phi_{ct} + \varepsilon_{i(c)t} \quad (2)$$

This specification can be estimated using both the self-reported phone surveys and the mystery shopper reports. As before, we control for the pre-treatment outcome $Y_{i(c)0}$. In some

[5] Observations with missing pre-treatment outcomes are retained by setting $Y_{i(c)0} = 0$ and including a missing indicator in the regression, which effectively excludes them from estimating δ while preserving them for estimating β. Since $Y_{i(c)0}$ is orthogonal to treatment and included only to improve precision, this approach does not introduce bias. Results are nearly identical when excluding these observations entirely, which is unsurprising given that only 2 percent are missing.

[6] Because shops did not yet know the treatment at the pre-survey, they should not have an incentive to over-report, as they might in the phone surveys. Moreover, with the enumerator physically present and able to observe inventory, misreporting was unlikely.

specifications with the mystery shopper data, we omit this control to ensure that differences in measurement do not drive the results. An alternative specification controls for shop fixed effects, λ_i, instead of the pre-treatment outcome:

$$Y_{i(c)t} = \beta[Treat\_1week]_{i(c)} + \gamma[Treat\_2week]_{i(c)} + \lambda_i + \phi_{ct} + \varepsilon_{i(c)t} \tag{3}$$

Finally, we test for heterogeneous treatment effects. Let $Z_{i(c)t}$ be an interaction variable (for example, the size of the shop or the number of nearby competitors). We estimate:

$$\begin{aligned} Y_{i(c)t} = {} & \beta[Treat\_1week]_{i(c)} + \gamma[Treat\_2week]_{i(c)} + \alpha Z_{i(c)t} \\ & + \beta'[Treat\_1week]_{i(c)} Z_{i(c)t} + \gamma'[Treat\_2week]_{i(c)} Z_{i(c)t} + \delta Y_{i(c)0} + \phi_{ct} + \varepsilon_{i(c)t} \end{aligned} \tag{4}$$

where β' and γ' capture any change in treatment effect for every one-unit increase in $Z_{i(c)t}$.

4 Main Results

4.1 Persistent Effects on Stocking After the Subsidy Ends

Our primary outcome, pre-registered before implementation, is whether shops continue to stock the target product two months after the subsidy period ends. Because this two-month point falls around Rounds 5 or 6 for most shops, we focus on outcomes measured in the final rounds of our sample, while also examining how stocking patterns evolve over time.

Figure 5 plots the coefficients and the 95% confidence intervals for $\{\beta_s\}$ and $\{\gamma_s\}$ from Equation 1. Each round covers a two-week period, with Round 1 representing the first two weeks after treatment. Panel (a) shows self-reported stocking from the biweekly phone survey. Both treatment groups report large and statistically significant increases in stocking in Round 1—about 40 percentage points above the control group. While the effect for the 1-week subsidy group declines over time—to around 18 percentage points by Round 5—it remains significant throughout. The 2-week subsidy group shows more persistent effects: by Round 6 (12 weeks after treatment), treated shops are still approximately 30 percentage points more likely to report stocking the product than controls.

Panel (b) shows stocking as measured by mystery shoppers. The 1-week subsidy group shows an initial impact of roughly 20 percentage points in Round 1, but this effect fades quickly and becomes statistically insignificant by Round 2, though point estimates remain positive. In contrast, the 2-week subsidy group shows a more sustained increase: the effect persists through the end of the sample period, declining gradually to 13 percentage points by Round 6, and remains statistically significant throughout.

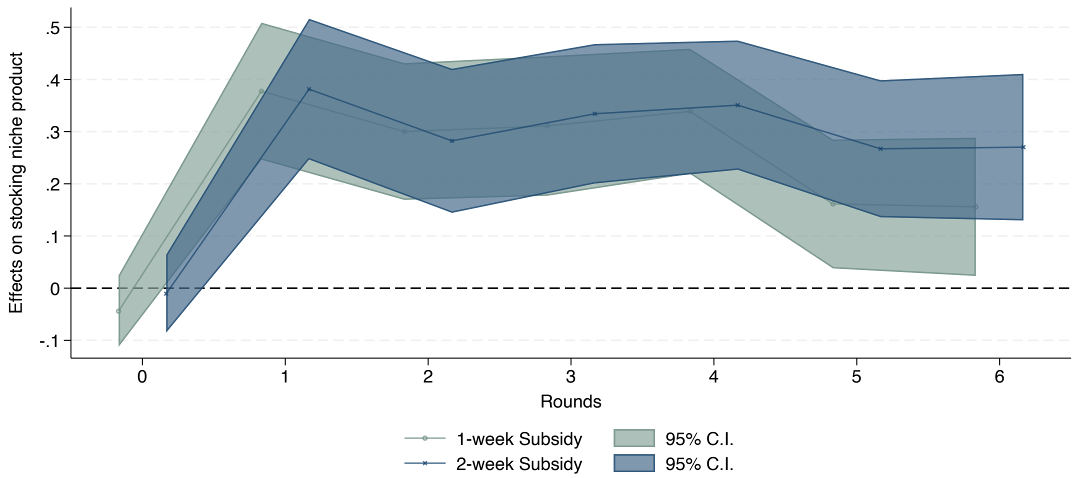
Panel A of Table 3 presents pooled estimates using the full 12-week panel of self-reported stocking outcomes from the biweekly phone surveys. Across all rounds, the 1-week group is 27–29 percentage points more likely to stock the product, and the 2-week group is 30–31 percentage points more likely, with no significant difference between the two groups overall. However, consistent with the dynamics in Figure 5, the 1-week group shows a steeper decline over time, whereas the 2-week group exhibits a lasting and statistically significant increase in stocking that persists through the end of the sample period.

Panel B reports analogous results using the mystery shopper data. Overall effects are smaller but follow a similar pattern. The pooled effect is 8 percentage points for the 1-week group and 16 percentage points for the 2-week group. The 2-week group shows persistent effects that last through the end of the sample period, whereas the impact in the 1-week group appears temporary, becoming statistically insignificant in later rounds.

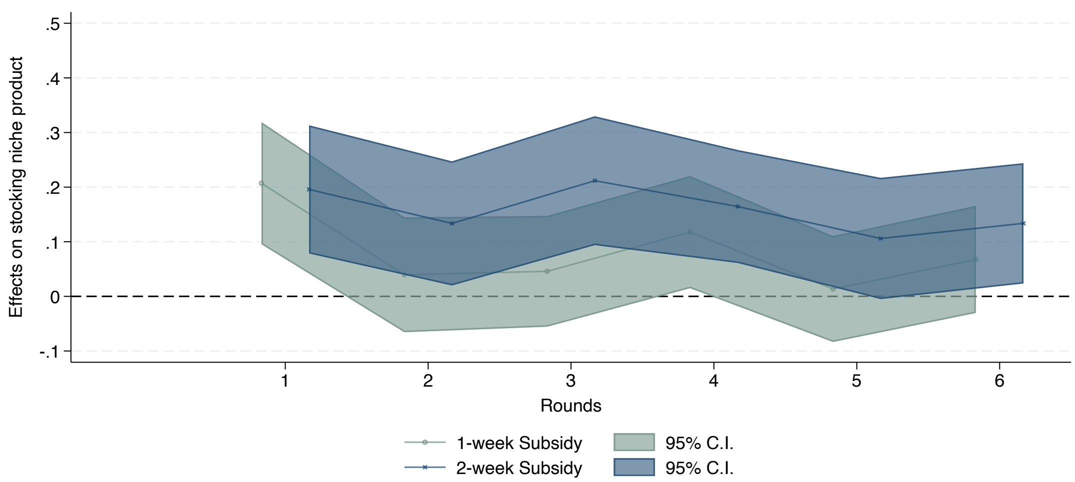
The results show that a temporary subsidy not only shifts stocking decisions in the short run, but can also lead to sustained adoption well beyond the subsidy period, particularly in the two-week treatment where shops had sufficient time to experiment.

Figure 5: Dynamic Treatment Effects on Stocking the Target Product

(a) Data from phone surveys



(b) Data from mystery shoppers



Note: Vertical axes report the coefficients β_s and γ_s and their 95% confidence intervals for each survey round s (two-week intervals) in Equation 1, where regressions include pre-treatment measurement and product×round fixed effects and standard errors are clustered at store level. Panel (a) shows stocking behavior as self-reported by retailers during biweekly phone surveys. Panel (b) displays stocking behavior as verified by mystery shoppers. While mystery shopper visits occurred weekly, data were collapsed into two-week intervals for comparability: a shop is considered to have stocked the product if it was stocked during at least one of the two weeks within each round.

Table 3: Pooled Treatment Effects on Stocking the Target Product

	(1)	(2)	(3)	(4)	(5)	(6)
	Full period		Rounds 1 and 2		Rounds 5 and 6	
Panel A. Phone survey data (bi-weekly)						
1-Week Subsidy (β)	0.27***	0.29***	0.34***	0.36***	0.15***	0.20***
	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)
2-Week Subsidy (γ)	0.31***	0.30***	0.33***	0.35***	0.27***	0.27***
	(0.05)	(0.07)	(0.06)	(0.07)	(0.06)	(0.08)
Stocked at Pre-Survey	0.36***		0.40***		0.33***	
	(0.04)		(0.05)		(0.06)	
Product$\times$Round FEs	X	X	X	X	X	X
Store FEs		X		X		X
p-value ($\beta = \gamma$)	0.408	0.919	0.941	0.869	0.076	0.357
Control mean	0.36	0.35	0.40	0.35	0.35	0.33
Number of unique shops	270	270	259	258	258	256
Observations	1,472	1,736	474	730	496	747
Panel B. Mystery shopper (weekly)						
1-Week Subsidy (β)	0.08**		0.12***		0.04	
	(0.03)		(0.05)		(0.04)	
2-Week Subsidy (γ)	0.16***		0.17***		0.12***	
	(0.04)		(0.05)		(0.05)	
Stocked at Pre-Survey	0.31***		0.31***		0.27***	
	(0.04)		(0.05)		(0.05)	
Product$\times$Round FEs	X		X		X	
p-value ($\beta = \gamma$)	0.063		0.393		0.073	
Control mean	0.31		0.34		0.32	
Number of unique shops	275		268		268	
Observations	2,909		894		1,019	

Note: Odd-numbered columns (1, 3, 5) display coefficients from Equation 2, and even-numbered columns (2, 4, 6) from Equation 3. Panel A uses bi-weekly phone survey data, and Panel B reports results from weekly mystery shopper data. Panel B does not include Equation (3) estimates and uses weekly mystery shopper data without collapsing into two-week intervals. Standard errors clustered at store level in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

4.2 Treated and Voluntary Stocking Shops Earn Similar Profits

A key implication of the rational sorting hypothesis is that shops that chose to stock the target product in the absence of the subsidy should earn higher profits than those induced to stock through the experiment. If profitability varies across shops and firms accurately anticipate their own returns, only high-return shops should self-select into stocking.

To test this, we compare treated shops who began stocking after the subsidy to control group shops who chose to stock voluntarily. Since the treatment increased take-up relative to control, the control group consists entirely of shops that self-selected into stocking, while many treated shops were induced to do so (as shown in Figure 5). Figure 6 shows conditional medians of self-reported markups across groups. We focus on medians due to the presence of outliers in the markup distribution.[7] Median markups are estimated using quantile regressions that control for target product cluster. Across Rounds 1–2 and Rounds 5–6, we find no significant difference in median markups between treated and control shops. Similarly, Table A2 shows no significant differences in average sales, profits, or markups between treatment and control shops, conditional on stocking the target product. These patterns are inconsistent with a rational sorting model, which would predict lower returns among treated shops if they had rationally chosen not to stock due to lower expected profitability.

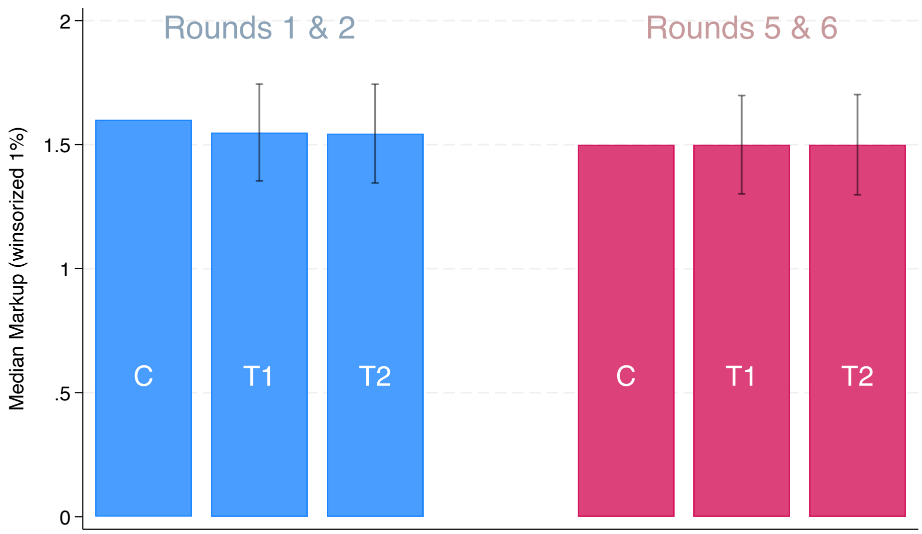
In addition, Table A3 shows that among shops in the treatment groups who are newly induced to stock the target product, profits from the target product are similar regardless of whether they initially reported believing the product to be unprofitable.[8] While this comparison does not exploit experimental variation and should be interpreted as suggestive, the pattern aligns well with the treatment-control comparison discussed above.

Together, these results provide no support for the rational sorting hypothesis. Shops that were experimentally induced to stock the product earned similar margins to those that stocked voluntarily, and prior beliefs about profitability did not predict realized profits.

[7] Outliers can result from misreported conversions between purchase and sale units, such as when shops sell onions in heaps but purchase them in large sacks. Results (not shown) are similar if we we use compare the 25th and 75th percentiles, instead of median.

[8] Shops were classified as viewing the product as unprofitable if they cited reasons such as "too little demand," "net cost exceeds revenue," or "existing competitor makes it unprofitable."

Figure 6: Median Markup for the Target Product by Treatment



Note: This figure plots the conditional median markup for the target product among treatment and control shops that stock it. The control group contains only shops that endogenously chose to stock the target product, while the treated groups contain shops experimentally induced to stock it. We estimate the conditional medians using a quantile regression that controls for product cluster. We plot confidence intervals for the difference between each treated group and the control group. The intervals are calculated as normal approximations based on bootstrapped standard errors (where resampling is done at the level of the shop).

4.3 Stocking the Target Product Increases Customer Traffic and Store Profits Without Displacing Sales of Main Goods

We next examine whether stocking the target product affected overall shop performance using the monthly high-frequency phone survey data. These outcomes, collected independently by a separate team, include daily customer traffic, total sales, and the cost of goods purchased and sold, allowing us to compute profits at both the store and product levels.

Columns 1-4 of Table 4 present the results on store-level customer traffic, sales, and profits measured at the endline monthly survey (corresponding to experimental Rounds 5 and 6). We find a positive, but less precisely estimated, impact on number of customers and store-level sales. Specifically, shops in the two-week subsidy group reported 20 percent more customers than those in the control group, while the one-week group experienced a smaller, statistically insignificant increase of 6 percent. Column 3 shows a positive but statistically insignificant effect on gross profits, defined as sales revenue minus the purchase cost of goods sold during the period. Column 4 reports a sizable (36–39%) increase in net profits, calculated as total sales revenue minus total purchase expenditure over the period, regardless of whether the goods were sold or remain in inventory. The decomposition in Table A4 suggests that this increase in net profits reflects both higher sales and lower overall purchase expenditures (including unsold inventory). While the point estimates are economically meaningful, they should be interpreted with caution given the large standard errors.

We further examine whether the new product displaced sales of existing goods using data collected in Rounds 5 and 6 of the phone survey, which asked shops about both the target product and their top three revenue-generating goods. Column 5 shows a significant increase in profits from the target product, which is expected, since the measure is unconditional—set to zero for shops not stocking the item—and treated shops are more likely to stock it. More importantly, Column 6 shows no significant change in profits from main goods, suggesting that the expansion in product variety did not crowd out existing sales.

Table 4: Treatment Effects on Store-Level and Product-Level Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	Store-level outcomes[a]				Net profits from:[b]	
	log(no. of customers)	log(sales)	log(gross profit)	log(net profit)	Target product	Main good
1-Week Subsidy (β)	0.06	0.13	0.04	0.39***	1.97**	12.79
	(0.07)	(0.11)	(0.10)	(0.13)	(0.79)	(8.26)
2-Week Subsidy (γ)	0.20**	0.17	0.11	0.36***	2.21**	3.64
	(0.09)	(0.10)	(0.10)	(0.12)	(0.94)	(7.64)
p-value ($\beta = \gamma$)	0.107	0.726	0.537	0.836	0.808	0.277
Control mean	3.17	4.49	1.60	4.04	0.16	11.46
Number of unique shops	224	224	224	221	268	268
Observations	224	224	224	221	533	533

Note: Coefficients are estimated from Equation 2, controlling for the pre-treatment value of the outcome and product-by-round fixed effects. Sales and profits are measured in USD. Gross profit is defined as sales revenue minus the purchase cost of goods sold during the period. Net profit is calculated as total sales revenue minus total purchase expenditure over the period, regardless of whether the goods were sold or remain in inventory. Columns (1)–(4) use data from monthly surveys conducted as part of the companion study (Samaniego de la Parra and Shenoy, 2024), matched to the timing of Round 6 of our phone surveys; outcomes are reported as daily averages over the past week. Columns (5) and (6) use data from Rounds 5 and 6 of the bi-weekly phone surveys. Standard errors clustered at the store level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

4.4 Summary

To summarize, the experimental findings provide evidence that misperceptions, rather than rational sorting, deter small retailers from stocking certain products. First, the temporary subsidy caused a substantial and persistent increase in adoption of the target product, particularly in the two-week treatment group. The fact that stocking persisted long after the subsidy ended is inconsistent with full-information sorting and instead points to learning through experience. Second, shops induced to stock the product by the subsidy earned markups comparable to those who had voluntarily chosen to stock it, and these profits did not vary systematically with prior beliefs. Third, we find no evidence that adoption crowded out sales of existing products. Treated shops appear to serve more customers and earn higher profits, suggesting that expanding product variety may have boosted overall demand.

5 Understanding Persistent Misperceptions: The Role of Noisy Learning and Inexperience

Why do retailers hold inaccurate beliefs about relatively common products, especially when those products are profitably stocked by similar shops in the same markets? One reason for the persistence of these misperceptions may be that the profitability of these products varies over time, making it difficult to assess their potential based on casual observation. Among shops that consistently stock the target product (from Round 0 to Round 6), we find a coefficient of variation of 0.3 for markups and 0.7 for sales. The markup variations could shift a product from being marginally unprofitable to profitable. The high volatility in sales makes it particularly challenging for shops that infrequently or newly stock the product to infer true demand from limited experience.

This noisy environment creates scope for self-confirming beliefs. Shops that never stock the product have few opportunities to learn from direct experience. We explore this using questions from the pre-treatment survey (Round 0), where shops were asked to predict the selling price, order price, and conversion rate (between selling and order units) they would expect to face over the next two weeks. If the shop did not stock the product, it was asked to consider a "very similar" shop that did—language designed to help respondents engage with the hypothetical questions.

We then compare predicted and actual prices and markups among shops that ultimately stocked the product in Rounds 1 and 2, just after treatment. Table 5 shows that only shops who were already stocking the product made accurate predictions. Panel A shows that their predicted markups significantly correlate with actual realized markups (Column 1), primarily driven by accurate forecasts of the selling price (Column 2), though their predictions of the order price are also informative (Column 3). By contrast, predictions made by shops that did not stock the product are uncorrelated with actual outcomes (Columns 4–6). Panel B presents similar patterns among treated shops.

Table 5: Correlation Between Pre-treatment Expectations and Realized Markups and Prices

	(1)	(2)	(3)	(4)	(5)	(6)
	Shops that stocked target product at Round 0			Shops that didn't stock target product at Round 0		
	Outcomes for target product at Rounds 1 and 2:					
	Mark-up	Sell price	Order price	Mark-up	Sell price	Order price
Panel A. Entire sample						
Expectations at Round 0	0.73***	0.74***	0.24**	-0.11	0.07	0.03
	(0.26)	(0.13)	(0.11)	(0.18)	(0.13)	(0.13)
Overall R-squared	0.685	0.901	0.246	0.187	0.398	0.454
Within R-squared	0.342	0.688	0.065	0.008	0.004	0.002
Number of unique shops	37	61	61	59	113	113
Observations	48	102	102	73	182	182
Panel B. Treatment group only						
Expectations at Round 0	0.84**	0.70***	0.13	-0.11	-0.01	-0.03
	(0.38)	(0.14)	(0.14)	(0.18)	(0.12)	(0.12)
Overall R-squared	0.561	0.903	0.251	0.212	0.343	0.533
Within R-squared	0.460	0.676	0.022	0.010	0.000	0.004
Number of unique shops	23	38	38	45	94	94
Observations	27	66	66	55	153	153

Note: Columns 1-3 include shops that stocked the target product at Round 0, while columns 4-6 include shops that did not. Regressions include product×round fixed effects. Prices winsorized at the 1st and 99th percentiles; markups winsorized at the 5th and 95th percentiles. Standard errors clustered at shop level in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Taken together, the results suggest that only shops with direct experience can reliably assess profitability. Those without such experience hold beliefs that diverge from reality, likely because they lack the data needed to update their beliefs. This also explains the experimental finding that two weeks of subsidized stocking led to more persistent adoption than one: longer exposure allows retailers to gather more signals in a noisy environment and update their beliefs with greater confidence.

6 Alternative Explanations and Robustness Checks

We now consider two alternative explanations for the persistence in stocking observed in the treatment groups and conduct additional robustness checks. The first is a strategic response to market competition: increased competition from subsidized shops may deter nearby control shops from entering. The second is a fixed cost explanation: the subsidy may have helped shops overcome search or sourcing frictions that would otherwise prevent adoption. While both mechanisms may generate persistence in stocking after subsidies end, additional evidence suggests they are not the primary drivers in our setting.

6.1 Strategic Response to Market Competition

The subsidy may have induced some shops to stock the product, while prompting nearby control shops to exit or avoid entry in response to increased competition. If each additional shop that stocks the product reduces the market power, and hence profitability, of nearby shops, then observed profits among those currently selling the product may not reflect the counterfactual profits of potential entrants. In this view, control shops may rationally anticipate increased competition from subsidized entrants, choose not to stock during the subsidy period, and remain deterred from entry afterward. Under this scenario, the RCT would have merely reshuffled which shops stock the product, without increasing overall adoption.

Under this hypothesis, we would expect the control to be less likely to stock the product in areas with higher treatment saturation. Because treatment was randomized at the individual shop level (within product cluster), the share of nearby shops assigned to treatment—conditional on the total number of nearby shops in the same product cluster—is exogenous. Although the experiment was not explicitly designed to vary treatment saturation, natural variation arises: the median shop has only five other nearby shops within 2km in the same product cluster. We leverage this variation to test whether control shops are less likely to stock the product when a larger share of nearby peers were treated, and whether treated shops are less likely to take up the product under high treatment saturation.

Table 6 estimates Equation 4, interacting treatment assignment with the (log) number of nearby treated shops within 2km in the same product cluster. Table A5 shows similar results using alternative radii (1km or 3km) to define nearby competitors. All specifications control for the level and interaction of the total number of nearby shops, with interaction terms demeaned so that the main treatment effect is interpreted at average saturation. Across specifications, we find no evidence that control shops exited or avoided stocking in response to increased competition from local treatment exposure. The direct effect of nearby treatment saturation is, if anything, significantly positive: control shops are actually more likely to stock the product when surrounded by treated shops. This could reflect information spillovers, consistent with the idea that firms update beliefs by observing others, or strategic responses aimed at retaining customer traffic by stocking the target product. The interaction between own treatment status and treatment saturation is negative but statistically insignificant. Overall, these results indicate that increased local adoption does not crowd out stocking among untreated shops, nor dampen treatment take-up.

Table 6: Effects on Target Product Stocking by Treatment Saturation

	(1)	(2)	(3)	(4)	(5)	(6)
	Full period		Rounds 1 and 2		Rounds 5 and 6	
Number of treatment shops within 2km (log)	0.32**		0.28*		0.26*	
	(0.13)		(0.16)		(0.16)	
1-Week Subsidy	0.22***	0.29***	0.30***	0.35***	0.11*	0.20***
	(0.05)	(0.06)	(0.06)	(0.07)	(0.06)	(0.07)
2-Week Subsidy	0.28***	0.31***	0.32***	0.37***	0.23***	0.28***
	(0.05)	(0.07)	(0.06)	(0.07)	(0.07)	(0.08)
1-Week $\times$ Number of treat shops (log)	-0.24	0.07	-0.21	0.06	-0.20	0.05
	(0.18)	(0.15)	(0.23)	(0.17)	(0.21)	(0.17)
2-Week $\times$ Number of treat shops (log)	-0.19	-0.08	-0.24	-0.13	-0.09	-0.02
	(0.19)	(0.21)	(0.23)	(0.22)	(0.26)	(0.26)
Product$\times$Round FEs	X	X	X	X	X	X
Store FEs		X		X		X
Control mean	0.37	0.35	0.40	0.36	0.35	0.33
Number of unique shops	262	262	256	256	252	252
Observations	1,443	1,705	469	725	486	738

Note: We estimate Equation 4 interacting treatment with the (log) number of nearby shops within the same product cluster that are treated, controlling for the interaction of the (log) number of nearby shops in the same product cluster. We define "nearby" as shops located within 2 kilometers. The outcome is self-reported stocking behavior from the phone surveys (the results are qualitatively similar for the mystery shopper reports of stocking). Standard errors are clustered by shop. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

6.2 Fixed Costs of Stocking New Products

Another possible explanation for the persistent treatment effects is that subsidies helped shops overcome a fixed cost associated with stocking the target product. In this view, although the product may offer attractive per-unit profits, adoption requires an upfront investment—most plausibly a fixed search cost to identify a viable supplier.

This explanation does not align with the data. First, if the relevant cost were a recurring one, such as transport or delivery, it would not explain why stocking persists after subsidies end. Instead, for the fixed-cost explanation to hold, the key friction would have to be informational: that shops initially do not know where to source the product, and the subsidy prompts them to identify a supplier they continue using afterward.

However, from the baseline survey, most shops already know where to source the product. Among those that did not stock at baseline, 90% reported knowing a market where the product could be purchased. Moreover, nearly all shops who adopt the product end up sourcing it from locations they already visit regularly. Among shops that source both the target product and one of their main goods from markets, 84% report using the same market for both. And of the 80 shops that never stocked the target product, 77 already source other goods from a market where another shop buys the target product.

Overall, we find little support for the view that adoption was constrained by fixed sourcing costs. The evidence suggests that the subsidy did not unlock access to new sourcing options, but rather shifted beliefs about whether stocking the product was worthwhile.

7 Conclusion

This paper shows that misperceptions can prevent small firms from adopting profitable products, even when those products are commonly stocked by peers. In a field experiment with small-scale retailers in Zambia, a temporary subsidy led to persistent increases in stocking behavior—especially when firms had more time to experiment—suggesting that firms updated their beliefs through direct experience. Shops induced to stock the product earned comparable profits to those that adopted organically, and we find no evidence of crowd-out from existing goods. These findings point to misperception as a key constraint on firm decision-making.

Misperceptions persist in part because learning is difficult in this environment: sales and profitability fluctuate, and firms that never stock a product lack the feedback necessary to update beliefs. This creates a self-confirming cycle in which firms avoid products they deem unprofitable and never learn otherwise. Temporary interventions that reduce the cost of experimentation can help firms overcome these barriers and make more informed decisions.

This informational constraint may be particularly acute for small firms. Larger firms, by contrast, benefit from scale: they can experiment without taking on excessive risk and observe outcomes across a broader portfolio of products and transactions. Their scale may itself be a source of better decision-making. Small firms lack this informational advantage, and one may speculate that this very constraint contributes to persistent gaps in size and productivity. Policies that help small firms pool information—whether through networks, digital platforms, or peer learning—may help overcome this constraint.

More broadly, our findings suggest that in settings where firms face uncertainty and limited opportunities for feedback, facilitating learning by doing may be more effective than offering abstract advice or external information. Future research can explore scalable ways to reduce misperceptions and test whether collective mechanisms—such as pooling ans sharing experiential data across firms—can achieve the informational benefits of scale.

References

Aker, J. C. (2010): "Information from markets near and far: Mobile phones and agricultural markets in Niger," *American Economic Journal: Applied Economics*, 2, 46–59.

Banerjee, A., G. Fischer, D. Karlan, M. Lowe, and B. Roth (2022): "Do collusive norms maximize profits? Evidence from a vegetable market experiment in India," Tech. Rep. w30360, National Bureau of Economic Research, Cambridge, MA.

Bergquist, L. F., C. McIntosh, and M. Startz (2024): "Search costs, intermediation, and trade: Experimental evidence from Ugandan agricultural markets," Tech. rep., National Bureau of Economic Research.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, J. Roberts, S. Ca, Overland Advisors Llc, S. U. Scid, World Bank, D. U. Bread, and U. Stanford (2013): "Does Management Matter? Evidence From India," *The Quarterly Journal of Economics*, 128, 1–51.

Bold, T., K. C. Kaizzi, J. Svensson, and D. Yanagizawa-Drott (2017): "Lemon technologies and adoption: measurement, theory and evidence from agricultural markets in Uganda," *The Quarterly Journal of Economics*, 132, 1055–1100.

Brooks, W., K. Donovan, and T. R. Johnson (2018): "Mentors or Teachers? Microenterprise Training in Kenya," *American Economic Journal. Applied Economics*, 10, 196–221.

Bruhn, M., D. Karlan, and A. Schoar (2018): "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico," *University of Chicago Press*, accepted: 2019-03-26T13:39:30Z Publisher: University of Chicago Press.

Cai, J., W. Lin, and A. Szeidl (2024): "Firm-to-Firm Referrals," Tech. rep., National Bureau of Economic Research.

Conley, T. G. and C. R. Udry (2010): "Learning about a new technology: Pineapple in Ghana," *American economic review*, 100, 35–69.

Dalton, P. S., J. Rüschenpöhler, B. Uras, and B. Zia (2021): "Curating Local Knowledge: Experimental Evidence from Small Retailers in Indonesia," *Journal of the European Economic Association*.

Dickstein, M. J. and E. Morales (2018): "What do exporters know?" *The Quarterly Journal of Economics*, 133, 1753–1801.

Drexler, A., G. Fischer, and A. Schoar (2014): "Keeping It Simple: Financial Literacy and Rules of Thumb," *American economic journal. Applied economics*, 6, 1–31.

Hanna, R., S. Mullainathan, and J. Schwartzstein (2014): "Learning through noticing: Theory and evidence from a field experiment," *The Quarterly Journal of Economics*, 129, 1311–1353.

Hsieh, C.-T. and B. A. Olken (2014): "The missing "missing middle"," *The journal of economic perspectives: a journal of the American Economic Association*, 28, 89–108.

Jensen, R. T. (2010): "Information, efficiency, and welfare in agricultural markets," *Agricultural Economics*, 41, 203–216.

Lagakos, D. (2016): "Explaining Cross-Country Productivity Differences in Retail Trade," *The journal of political economy*, 124, 579–620.

McKenzie, D. and C. Woodruff (2017): "Business Practices in Small Firms in Developing Countries," *Management Science*, 63, 2967–2981.

Samaniego de la Parra, B. and A. Shenoy (2024): "Measuring and Estimating Retail Productivity," Working paper.

Tanaka, M., N. Bloom, J. M. David, and M. Koga (2020): "Firm performance and macro forecast accuracy," *Journal of monetary economics*, 114, 26–41.

Appendix A

Table A1: Attrition Balance

	(1)	(2)	(3)
	Proportion of observations in:		
	Phone Survey (bi-weekly)	Mystery Shopper (weekly)	
	round level	week level	aggregated round level
1-Week Subsidy Group	-0.01	-0.00	-0.01
	(0.03)	(0.03)	(0.03)
2-Week Subsidy Group	-0.03	-0.00	-0.00
	(0.04)	(0.03)	(0.03)
Control mean	0.90	0.88	0.95
Control SD	0.21	0.18	0.16
Observations	271	271	271

Note: Regressions include target product group fixed effects. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A2: Target Product Profitability by Treatment
(Conditioning on Stocking)

	(1)	(2)	(3)
	Conditional on stocking target product:		
	Sales	Net profit	Mark-up
Panel A. Rounds 1 and 2			
1-Week Subsidy (β)	3.18	0.42	-0.16
	(3.36)	(1.89)	(0.20)
2-Week Subsidy (γ)	-0.87	-1.26	-0.30
	(3.10)	(1.83)	(0.19)
p-value ($\beta = \gamma$)	0.252	0.320	0.376
Control mean	13.90	1.78	2.16
Control SD	19.84	11.64	1.26
Number of unique shops	174	174	164
Observations	284	284	233
Panel B. Rounds 5 and 6			
1-Week Subsidy (β)	9.78**	3.13	-0.21
	(3.98)	(2.09)	(0.21)
2-Week Subsidy (γ)	0.17	2.21	-0.13
	(2.97)	(2.01)	(0.24)
p-value ($\beta = \gamma$)	0.030	0.634	0.642
Control mean	12.21	0.50	2.09
Control SD	15.34	12.02	1.29
Number of unique shops	138	138	138
Observations	231	231	231

Note: This table reports estimated coefficients from Equation 2, comparing profitability outcomes of the target product across treatment and control groups. The sample is restricted to shops that stocked the target product, consistent with the subsample analyzed in Figure 6. All outcomes (sales, net profit, and mark-up) pertain to the target product and monetary values are reported in USD. Standard errors clustered at store level in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A3: Target Product Profitability by Stated Non-stocking Reasons

	(1)	(2)	(3)
	At Round 0, reason for not stocking target product:		
	because unprofitable	other reasons	p-value (mean difference)
Sales (past 2 weeks, USD)	17.87	16.82	0.608
	(31.71)	(25.23)	
	[79]	[75]	
Profits (past 2 weeks, USD)	-2.14	1.52	0.101
	(14.21)	(10.90)	
	[79]	[75]	
Mark-up	1.85	1.55	0.622
	(1.12)	(0.81)	
	[70]	[55]	
Number of unique shops	51	43	

Note: The sample is restricted to treatment shops that did not stock the target product in Round 0 but began doing so in Rounds 1 or 2. Stated reasons for not stocking are measured at Round 0 (pre-survey), prior to the intervention. Sales, profits, and per-unit markups for the target product are measured in Rounds 1 and 2. In columns (1) and (2), means are reported, with standard deviations in parentheses and observations in brackets. Column (3) reports p-value for mean differences, controlling for shop cluster and treatment arm fixed effects.

Table A4: Decomposition of Store-Level Profits

	(1) Sales	(2) Purchase cost of sold goods	(3) Gross profit	(4) Inventory purchases	(5) Net profit
Panel A. Outcomes in levels					
1-Week Subsidy (β)	50.14	3.49	3.24**	-14.09	51.23**
	(31.79)	(3.25)	(1.52)	(14.71)	(25.85)
2-Week Subsidy (γ)	-23.70	-3.60	0.14	-19.01	9.36
	(39.23)	(4.01)	(1.83)	(13.36)	(34.32)
p-value ($\beta = \gamma$)	0.160	0.206	0.178	0.754	0.326
Control mean	197.13	24.62	8.06	70.14	125.12
Number of unique shops	224	224	224	227	224
Observations	224	224	224	227	224
Panel B. Outcomes in logs					
1-Week Subsidy (β)	0.13	-0.01	0.04	-0.02	0.39***
	(0.11)	(0.10)	(0.10)	(0.20)	(0.13)
2-Week Subsidy (γ)	0.17	0.08	0.11	0.08	0.36***
	(0.10)	(0.09)	(0.10)	(0.20)	(0.12)
p-value ($\beta = \gamma$)	0.726	0.376	0.537	0.621	0.836
Control mean	4.49	2.50	1.60	3.35	4.04
Number of unique shops	224	224	224	183	221
Observations	224	224	224	183	221

Note: Outcomes are daily averages for the past week; monthly survey data from companion study (Samaniego de la Parra and Shenoy, 2024) matched to "Round 6" of this study's phone surveys. Gross profit (column 3) is sales (column 1) minus purchase cost of sold goods (column 2); net profit (column 5) is sales minus all inventory purchases in the past week regardless of being sold in the corresponding recall period (column 4). Sales, costs and profits are measured in USD. Standard errors clustered at store level in parentheses. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.

Table A5: Effects on Target Product Stocking by Treatment Saturation: Alternative Radii

	(1)	(2)	(3)	(4)	(5)	(6)
	Full period		Rounds 1 and 2		Rounds 5 and 6	
Panel A. 1km radius						
Number of treatment shops within 1km (log)	0.21*		0.18		0.08	
	(0.13)		(0.18)		(0.18)	
1-Week Subsidy	0.24***	0.28***	0.33***	0.35***	0.14**	0.19***
	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)
2-Week Subsidy	0.30***	0.30***	0.34***	0.36***	0.25***	0.27***
	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)
1-Week × Number of treat shops (log)	-0.08	0.14	-0.20	-0.01	0.08	0.16
	(0.20)	(0.17)	(0.25)	(0.20)	(0.24)	(0.19)
2-Week × Number of treat shops (log)	-0.02	0.01	-0.13	-0.12	0.18	0.07
	(0.20)	(0.21)	(0.25)	(0.22)	(0.28)	(0.27)
Panel B. 3km radius						
Number of treatment shops within 3km (log)	0.17*		0.10		0.12	
	(0.09)		(0.14)		(0.13)	
1-Week Subsidy	0.24***	0.29***	0.31***	0.35***	0.11*	0.20***
	(0.05)	(0.06)	(0.06)	(0.07)	(0.06)	(0.07)
2-Week Subsidy	0.30***	0.31***	0.32***	0.36***	0.24***	0.28***
	(0.05)	(0.07)	(0.06)	(0.07)	(0.07)	(0.08)
1-Week × Number of treat shops (log)	0.02	0.22	0.04	0.16	0.02	0.16
	(0.21)	(0.21)	(0.26)	(0.23)	(0.25)	(0.24)
2-Week × Number of treat shops (log)	-0.06	-0.04	0.06	0.06	-0.02	-0.00
	(0.17)	(0.21)	(0.24)	(0.24)	(0.26)	(0.28)
Product×Round FEs	X	X	X	X	X	X
Store FEs		X		X		X
Control mean	0.37	0.35	0.40	0.36	0.35	0.33
Number of unique shops	262	262	256	256	252	252
Observations	1,443	1,705	469	725	486	738

Note: Similarly to Table 6, we estimate Equation 4 interacting treatment with the (log) number of nearby shops within the same product group that are treated, controlling for the interaction of the (log) number of nearby shops in the same product group. Panel A defines "nearby" as within a 1-kilometer radius, while Panel B uses a 3-kilometer radius. (That is, we include number in the product group and number treated in $Z_{i(c)t}$). The outcome is self-reported stocking behavior from the phone surveys (the results are qualitatively similar for the mystery shopper reports of stocking). Standard errors are clustered by shop. ***, **, and * represent significance at 1%, 5%, and 10%, respectively.