In-Store Coupons: A Large-Scale Field Experiment

Sebastian Gabel[*], Duncan Simester[†], and Artem Timoshenko[‡]

[*]Rotterdam School of Management, Erasmus University
[†]MIT Sloan School of Management
[‡]Kellogg School of Management, Northwestern University

Abstract

New technologies are transforming grocery coupons by enabling digital delivery via in-store kiosks and mobile apps. We investigate the effectiveness of these "bottom-of-the-funnel" promotions using a large-scale field experiment across 71 product categories. On average, in-store coupons cause a fivefold increase in sales of the promoted brands and drive substantial category expansion. The category expansion is not just temporal substitution. Instead, in-store coupons appear to be effective at inducing customers to switch purchases from competing retailers. This cross-retailer substitution makes in-store coupons less valuable for manufacturers, but especially valuable for retailers.

Keywords: in-store marketing, coupons, field experiments, causal effects, incrementality

1. Introduction

New technologies are transforming the way retailers promote consumer packaged goods (CPG). Retailers such as CVS, Winn-Dixie, Food Lion, and Stop & Shop now distribute coupons via in-store kiosks (Hamstra 2023). Kroger, Target, Publix, and Best Buy have introduced mobile apps that consumers can use to browse coupons in-store (Berthiaume 2022). Other emerging technologies include digital in-store signage and intelligent shopping carts (Grewal et al. 2023). These technologies are disrupting the $430 billion CPG coupon industry by delivering promotions to customers at the very bottom of the purchasing funnel, when they are about to make a purchase.

Whether these "bottom-of-the-funnel" promotions are effective in physical retail remains an open question. On the one hand, distributing coupons in-store reduces the gap between promotional exposure and the purchase decision. Amazon's sponsored search and on-site advertising share this characteristic, and the effectiveness of these lower-funnel interventions has helped Amazon become the world's third-largest advertising company (Statista 2026). On the other hand, because in-store promotions target customers who have already committed to a store visit, they may merely subsidize existing purchase intent. If shoppers simply apply the discounts to planned purchases or use them to pull demand forward from future weeks, the coupons will not generate incremental growth. Determining which of these arguments prevails is an important measurement challenge.

We investigate the effectiveness of in-store coupons using a large-scale field experiment at a major supermarket chain in Germany. The retailer delivers coupons via kiosks located near store entrances, capturing customers at the beginning of their visits. The experiment was conducted in a *business-as-usual* environment by distributing random coupons using the retailer's couponing system. The random coupon assignments exogenously varied the promoted brands and discount depth across different customers. We use the retailer's loyalty program to link coupon exposures to consumers' individual purchases and report the incremental (causal) effects across 101 CPG campaigns spanning 71 categories.

In-store coupons dramatically increase purchases of promoted brands. On average, receiving an in-store coupon causes purchase incidence to increase from 0.33% to 1.61% on the current visit. Even though in-store coupons increase unit sales, we may be concerned that they do not increase revenue because discounts offset the sales lift. Our findings assuage this concern. The sales increases in our study were more than enough to offset the discounts: the average revenue increased from €6.59 to €34.09 per 1,000 coupons (per mille), with no statistically significant revenue losses across any of the 101 campaigns.

In-store coupons also drive category expansion. Across the 101 campaigns, the lift in promoted sales closely aligns with the total category growth, suggesting that the sales growth is largely incremental to the retailer. We investigate whether the category expansion merely reflects temporal substitution, yet find no evidence of a post-promotion dip when averaged across brands. In the subset of categories where a dip does occur, the behavior is consistent with temporal substitution of consumption, meaning the coupons may influence the timing of actual usage rather than induce inventory accumulation. Overall, while temporal substitution occurs in isolated cases, the observed category expansion represents an increase in demand at the focal retailer, rather than a simple acceleration of future sales.

An alternative explanation for category expansion is store switching: customers purchase immediately at the focal store rather than at competing retailers on future visits. We find initial evidence of this in categories where consumption expansion is unlikely, such as toilet paper and detergent. In these categories, we observe no post-promotion dip over the subsequent five weeks, even among customers who had not purchased the category at the focal retailer in the three months before the promotion. To investigate further, we compare stores located near versus far from competitors. While both groups exhibit immediate category expansion, the sales increase persists in stores near competitors, whereas a post-promotion dip occurs in more isolated stores. This pattern is consistent with what we would expect if in-store coupons steal demand from competitors, primarily when those competitors are nearby.

We provide additional evidence of store switching using household panel data. Because the retailer introduced the coupon system in a subset of its stores (in one region), we can use a difference-in-differences analysis to estimate the causal effect of the in-store coupon system on purchases from each retailer. The introduction of the coupon system resulted in customers switching purchases from competing retailers to the focal retailer. Moreover, the categories for which this effect was largest in the panel data are the same categories for which we observe the largest category expansion in our experiments. What makes this pattern especially notable is that it aligns category expansion and store switching evidence obtained from two independent identification approaches: randomized variation within the focal retailer and difference-in-differences analysis of external panel data.

For manufacturers, the evidence of store switching undermines the benefits of the positive brand effects in the focal store. For retailers, the category expansion is unambiguously favorable, particularly given the evidence that temporal substitution plays only a limited role. A marketing promotion that expands category sales by stealing sales from competitors is a very attractive tool for a retailer.

Our study is conducted in a business-as-usual environment, providing causal evidence of what manufacturers and retailers can expect when adopting in-store coupons. However, we caution that these data describe customer behavior at a single supermarket chain and rely on a specific distribution mechanism: coupons delivered at the store entrance, before customers reach the aisle. While this timing allows us to capture "bottom-of-the-funnel" effects, additional research is required to further study generalizability across different retail settings.

Outline of the Paper

The paper proceeds with a review of the related research in Section 2. We describe the data and experimental setting in Section 3 and discuss the identification strategy in Section 4. We document outcomes for the promoted brand in Section 5, and investigate category expansion in Section 6. The paper concludes in Section 7.

2. Related Literature

There is a rich literature studying how consumers respond to coupons and price promotions. We first review the literature on in-store coupons and then discuss research related to category expansion.

2.1 In-Store Coupons

The in-store coupons in this study are distributed when customers start their shopping trips. In contrast, shelf and checkout coupons arrive in the middle and at the end of the trip (respectively). This distinction is important. Coupons distributed at the start of a trip can affect whether a customer visits an aisle, while this is less likely to occur with on-shelf or checkout coupons. We can also contrast these coupons with coupons delivered at home. At-home coupons are received earlier in the purchase funnel and can affect the frequency and choice of which store to visit, which in-store coupons cannot.

Previous investigations of in-store promotions include coupons distributed at the shelf, such as coupons dispensed from J-hooks located directly in front of the promoted product (Dhar and Hoch 1996) or peel-off coupons placed on individual packages (Raju et al. 1994). Dhar and Hoch (1996) conducted two field experiments at 86 Chicago supermarkets. They find that on-shelf coupons yield sales increases similar to price discounts, but because only 55% of purchasers redeemed coupons, the revenue and profit increases were much larger for the coupons. Raju et al. (1994) conducted a series of quasi-experiments at a UCLA student store in which they placed coupons on 14-oz and 20-oz soda-fountain drinks. They conclude that on-pack coupons lead to larger market share and category sales lifts than peel-off and in-pack coupons.

Two studies consider coupons delivered as customers enter a store. Heilman et al. (2002) distributed printed coupons to 105 customers intercepted at the entrance of two St. Louis grocery stores. In a similar study, Hui et al. (2013) distributed printed coupons to 90 customers entering a grocery store in Pittsburgh. Both studies document the impact of the coupons on "unplanned" purchases (in categories not on the shopping list). The interactions between researchers and customers introduce a potential limitation to both studies, by introducing a risk that the interactions themselves could impact customers' behavior in the store (Pollay 1968).[1]

Our paper contributes to this literature by investigating the effectiveness of in-store coupons using a large-scale field experiment. Conducted in a business-as-usual retail environment, the study unobtrusively embeds the intervention into shoppers' purchase journeys without making participants aware of the experimental manipulation, thereby ensuring external validity. The loyalty program data helps us link the coupon distribution to purchase outcomes and evaluate incremental sales by different customer segments. The findings demonstrate that in-store coupons are effective for driving sales of the promoted brands and category expansion at the focal retailer.

[1] In addition, Chiou-Wei and Inman (2008) study a setting in which "consumers browse the Web at home for any suitable coupons available on the store coupon pages and print them out. In the case of forgetting and/or misplacing, they can also get access to the coupon pages by running their shopping cards through terminals at the store and printing out the coupons" (p. 297). The browsing and printing at home is very similar to clip-at-home coupons. An important distinction is that we study the impact of in-store coupons using data from a field experiment.

2.2 Category Expansion

Beyond the effects on the promoted brand, prior research has studied whether coupons can contribute to category expansion by increasing consumption or inducing stockpiling. Chiang (1995) finds that for goods with fixed consumption rates, such as laundry detergent, promotions do not drive category growth. In contrast, Ailawadi and Neslin (1998) demonstrate that for other categories like yogurt, increased household inventory can accelerate consumption rates. Nijs et al. (2001) further observe that price promotions can generate category expansion that persists for up to ten weeks (see also Dekimpe et al. 1998; Mela et al. 1998).

The existing literature on in-store coupons offers limited insight into these category-level dynamics. Heilman et al. (2002) and Hui et al. (2013) both focus on customer basket sizes rather than total category sales. In the Dhar and Hoch (1996) study, the effects on category expansion are confounded by concurrent promotional activity on competing brands during the study period. Raju et al. (1994) also explicitly acknowledges that they do not investigate category expansion.

One explanation for category expansion is that coupons can increase consumption via impulse purchasing. A key motivation for retailers' in-store marketing activities, including in-store coupons, is to increase impulse purchasing (Abratt and Goodey 1990; Zhou and Wong 2004). Impulse purchases are unplanned (see the discussion in Shapiro 2001), and often include a "powerful and persistent urge to buy something immediately" (Rook 1987, p. 191). Bucklin and Lattin (1991) propose a two-state model of purchase incidence and brand choice in which consumers shop in two ways. Shoppers who have planned their shopping trips are unresponsive to in-store promotions. In contrast, shoppers who have not planned their trips make impulsive decisions in the store and respond more strongly to in-store promotions.

Alternatively, retailers are often concerned that in-store coupons merely subsidize purchases that would have occurred anyway, shifting future demand to the present. Literature on price promotions frequently supports this "temporal substitution" hypothesis, documenting a post-promotion dip in sales following a promotion (Gupta 1988; Chiang 1991; Grover and Srinivasan 1992; Bell et al. 1999). Household stockpiling is often cited as the driver of this dip, though empirical evidence linking a product's "suitability for stockpiling" to promotional response remains mixed. Some studies report a positive relationship (Narasimhan et al. 1996; Bell et al. 1999; Ailawadi et al. 2006; Fok et al. 2006), whereas others report a negative or non-significant relationship (Nijs et al. 2001; Lim et al. 2005; Osuna et al. 2016).

Our study systematically investigates how in-store coupons impact sales of both the promoted brand and overall category sales. The scale and scope of the study allow us to document the effects across customers and product categories, providing evidence that the observed category expansion is not merely temporal substitution. Instead, using two different datasets, we provide convergent evidence consistent with store switching. In the next section, we introduce our empirical setting and discuss the design and implementation of the field experiment.

3. Data Overview

The data for our study were provided by a large German grocery retailer. The stores use a traditional supermarket format, and during the experiment, the retailer operated over 500 stores. The product range included over 30,000 SKUs spanning food, beverages, and non-food products, such as laundry detergents and paper towels. The stores use a HiLo price format, and the average price point is similar to its moderately priced competitors. The retailer maintains a loyalty program to collect customer data and distribute in-store coupons. Our dataset contains the purchases of all customers who shopped in the retailer's 147 stores in one of Germany's federal states, including over 250,000 customers in the retailer's loyalty program.[2]

The retailer communicates with loyalty program members through self-service kiosks (see Figure 1). The kiosks allow customers to print in-store coupons and check their loyalty point balances. During the study, the kiosk system was the only touchpoint between the retailer and its loyalty program customers. On average, customers who participate in the loyalty program obtain coupons from an in-store kiosk on 40% of their shopping trips.

Customers typically interact with the kiosks and obtain in-store coupons when entering the store. After customers present their loyalty cards, the kiosk prints seven coupons, each of which offers a discount on a single brand. When customers show their loyalty card at the checkout, they automatically receive these discounts if they purchase any of the brands represented by the coupons.

[2] Loyalty program participants account for 21.1% of the retailer's total revenue. The penetration tends to be larger in suburban residential areas with lower household incomes (up to 43.7%).

Figure 1: In-Store Kiosk and Coupon Printout

Notes: The figure depicts the self-service kiosk (left) and the coupon printout (right). The printout contains seven coupons valid for the same day, regular and discounted prices, and the discount rate in percent.

If customers use the kiosk system multiple times in one day, they receive the same coupons; the coupon assignments are determined the night before and are valid for a day.

The retailer manages coupons in campaigns, where each campaign promotes a specific brand's products to loyalty program members over a set period. In our analysis, we use these campaigns as our primary unit of observation to report how coupons affect sales of the promoted products. We assigned each campaign to a product category using the retailer's internal category map. These category definitions are relatively narrow. Examples include frozen pizza, cat food, deodorant, and juice. On average, a category includes 103 SKUs, and the promoted brands in our experiment represent 13.3% of the revenue in their corresponding categories (see Web Appendix A for a complete list).

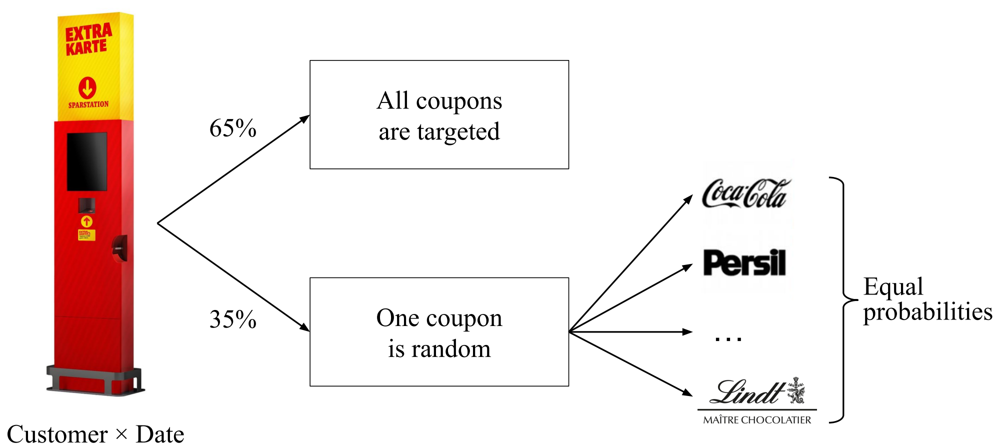
The retailer allocates coupons to customers based on their past spending. For example, the targeting system prioritizes coupons for the brands that a customer has purchased before. We will label coupons that the system chooses to distribute to individual customers as *targeted coupons*. The system ensures customers' daily coupons do not include multiple targeted coupons for the same or competing brands (in the same category).

3.1 Experimental Design

We obtained data from an experiment that we helped design and implement for the retailer, and which became part of the retailer's standard practice. The experiment distributed random coupons to both provide training data and evaluate the performance of the couponing system.

We illustrate the experimental design in Figure 2. By default, all seven coupons are targeted. In the experiment, in 35% of kiosk visits, the couponing system replaces one of the targeted coupons with a random coupon. The random coupons feature a randomly selected brand with a random discount depth ranging from 5% to 50%. The coupon assignments and randomization are implemented daily so that, on a given day, each customer has a 65% chance of receiving only targeted coupons, and a 35% chance of receiving a random coupon.

Figure 2: Experimental Design with Random Coupons



If a customer receives a random coupon, the random coupon replaces one of the seven targeted coupons. The targeted coupon that is replaced is randomly selected. The only exception is if the printout already includes a targeted coupon for the same brand as the random coupon, the random coupon replaces this targeted coupon (ensuring that the printout never includes two coupons for the same brand).

We emphasize that the experiment only includes customers enrolled in this retailer's loyalty program, so our results are limited to this group. The retailer price discriminates by restricting coupon distribution to loyalty program members, deliberately excluding customers who have opted not to join the program from receiving any coupons.

The retailer's couponing system records which coupons are distributed to customers, including the featured brands and discounts. The system also logs which coupons are targeted and which coupons are random (if any). However, the retailer does not record which targeted coupons are replaced by the random coupons, meaning there is no counterfactual policy logging in the system (Johnson et al. 2017).

3.2 Descriptive Statistics

Table 1 reports descriptive statistics. The dataset includes 1,503,000 kiosk visits across 158 different dates (November 2015 to May 2016). Our unit of observation is a customer×date combination with at least one kiosk visit. We focus on 88,840 random coupons distributed over 101 CPG campaigns, which had regular sales prior to the experiment, distributed at least 400 random coupons, and had no eligibility criteria.[3] We provide additional details about sample construction in Appendix A and randomization checks in Web Appendix B. The final sample spans 71 product categories, including food, drinks, and non-food CPG categories. The average discount per random coupon was 23.95%, and the redemption rate for random coupons was 1.61%.

Table 1: Experimental Data Descriptive Statistics

Number of campaigns	101
Number of categories	71
Number of random coupons	88,840
Number of shopping trips with kiosk usage	1,503,000
Average discount per random coupon	23.95%
Redemption rate for random coupons	1.61%
Average face value per redeemed random coupon	€0.84

Our analysis focuses on incremental changes in purchase incidence, quantity, and revenue (see Table 2). Brand managers who execute and evaluate campaigns typically focus on quantity and revenue, and the retailer reports these outcomes in its campaign dashboard. The revenue measure offers the advantage of accounting for price discounts. We consider incidence because we can measure treatment effects for this binary variable with high precision. Comparing multiple

[3]Some brand managers restricted the distribution of coupons to a subset of customers (e.g., customers who had recently purchased in the category). To avoid selection effects and facilitate our comparison of treatment effects, we excluded brands that used such eligibility criteria.

metrics also allows us to evaluate the robustness and consistency of our findings.

Table 2: Performance Metrics

Metrics	**Description**	**Type**
Purchase Incidence	Did a customer purchase or not?	$\{0, 1\}$
Quantity	How many units did a customer purchase?	$\mathbb{N}^0$
Revenue	How much did the customer spend?	$\mathbb{R}^0$

Our primary analysis focuses on the day that the customer used the kiosk to obtain in-store coupons, although in Section 6 we also extend the measurement window to include purchases in subsequent weeks. We next describe an identification approach that estimates the average treatment effect for each coupon campaign.

4. Identification Approach

Our goal is to measure the average treatment effect (ATE) of in-store coupons for loyalty program customers on days that they use the kiosk (customer×date observations). The ATE captures the average impact of providing a coupon for the focal brand, compared to not providing a coupon for the focal brand:

$$\text{ATE} \equiv \mathbb{E}[Y(C=1) - Y(C=0)], \tag{1}$$

where $Y(C=1)$ and $Y(C=0)$ indicate potential outcomes with and without a coupon, and the expectation integrates over customer×date observations. Y refers to the performance metrics described in Table 2. The ATE does not confound in-store coupon effectiveness with the retailer's targeting algorithm.[4]

4.1 Estimator

To estimate Equation (1), we can interpret the random sample of customers who received a random coupon for the focal brand as a "treatment" sample. Outcomes for these customers directly provide an estimate of $\mathbb{E}[Y(C=1)]$. The remaining customers are equivalent to the treatment sample and did not receive a *random* coupon for the focal brand. However, we cannot use their average outcomes

[4] As an alternative, we could study the treatment effects of targeted in-store coupons (ATT), but these findings would be specific to the firm's current targeting algorithm.

to directly estimate $\mathbb{E}[Y(C=0)]$, because some of these customers received a *targeted* coupon for the focal brand.

To account for the focal brand's targeted coupons, our identification approach leverages random coupons for <u>non-focal</u> brands to infer the ATE for the <u>focal</u> brand. We summarize our identification approach here and provide additional details in Appendix B. For each brand, we define three groups of customer×date observations:

1. *Focal*: Observations in which a customer received a random coupon for the focal brand. This is the treatment sample.
2. *NonFocal*: Observations in which a customer received a random coupon for any non-focal brand.
3. *NoRandom*: Observations in which a customer did not receive any random coupons (all coupons are targeted).

The experimental variation ensures that, for each brand, the three groups are equivalent, mutually exclusive, and representative of the overall customer×date population. The *Focal* observations always received a coupon for the focal brand, due to the random assignment. The *NonFocal* and *NoRandom* observations do not receive a random coupon for the focal brand in the experiment, but they could receive a targeted coupon for the focal brand.

We estimate the average treatment effect for the focal brand using the following equation:

$$\widehat{\text{ATE}} = \underbrace{\bar{Y}(\text{Focal})}_{\mathbb{E}[Y(C=1)]} - \underbrace{\Big[\bar{Y}(\text{NoRandom}) - 1/p \cdot \Big(\bar{Y}(\text{NoRandom}) - \bar{Y}(\text{NonFocal})\Big)\Big]}_{\mathbb{E}[Y(C=0)]} \tag{2}$$

where $\bar{Y}(\cdot)$ indicates the average outcome in the corresponding group, and p is the probability that a random *non-focal* coupon replaces a targeted *focal* coupon. This probability is defined over observations in which the customer is assigned to receive both a random non-focal coupon and a targeted focal coupon. In our setting, there are seven targeted coupons that could be replaced by the non-focal coupon, so there is a one in seven chance that the targeted coupon for the focal brand is the one that is replaced ($p = 1/7$).

The treatment effect estimator in Equation (2) contains two components. The first component captures the average outcome in the treatment sample, *Focal*, to estimate $\mathbb{E}[Y(C=1)]$. The second component estimates $\mathbb{E}[Y(C=0)]$ using the outcomes in the *NoRandom* and *NonFocal* conditions. Intuitively, the experimental design ensures the equivalence of observations in these two conditions. However, in *NonFocal*, random non-focal coupons replace p proportion of (any) targeted coupons, so there are fewer focal targeted coupons in *NonFocal* than in *NoRandom*. We know this is the only cause of different outcomes between the two conditions. We can thus recover the effect of the targeted coupons in *NoRandom* by scaling the difference in outcomes between *NoRandom* and *NonFocal* by $1/p$. We obtain $\mathbb{E}[Y(C = 0)]$ by subtracting this effect from $\bar{Y}(\text{NoRandom})$. We illustrate this intuition using an example in Appendix B.

We also note that this adjustment for targeted coupons actually has little impact on the results in our study. If we forgo this adjustment and use the simple difference-in-means estimator, we obtain similar ATE estimates, and our substantive conclusions are unchanged.

4.2 Identification Assumptions

The proposed estimator relies on two assumptions commonly used in the causal inference literature (Angrist et al. 1996; Johnson et al. 2017). First, we make the stable unit treatment value assumption (SUTVA). SUTVA requires that the potential outcomes for any given observation do not vary with the treatment assignment for other observations. In our context, this implies that receiving a coupon does not affect the shopping behavior of other customers in the store at the same time, and receiving a coupon does not change the potential outcomes for that same customer on subsequent dates. Second, we assume that the coupon treatment effects do not extend across brands. This exclusion restriction implies that replacing a focal coupon with a random non-focal coupon shifts a potential outcome from $Y_i(1)$ to $Y_i(0)$. In Appendix B, we show that the proposed estimator is unbiased under these two assumptions and we derive its variance.

In principle, the panel structure of our data and competitive spillovers can introduce threats to these assumptions. Regarding SUTVA, we note that the absence of *cross-customer* spillovers is plausible, but using multiple customer$\times$date observations per customer introduces the possibility of temporal spillovers (such as forward-buying), altering a customer's subsequent potential outcomes. Regarding the exclusion restriction, we recognize that competing brands' coupons could affect sales of the focal brand. Our identification approach naturally partials out these competitive effects in the "control" condition, but targeted competitive coupons can co-occur with random focal coupons, which could introduce bias.

We explicitly address both concerns in Web Appendix C by replicating our analysis with a restricted sample. To rule out temporal spillovers, we isolate only the first customer×date observation for every customer in the experiment, and to rule out competitive spillovers, we restrict the sample to campaigns without competitor coupons. Although this substantially reduces our sample sizes, the key findings all remain robust, and the patterns supporting our proposed mechanisms remain directionally aligned. This empirical stability is consistent with our setting: competitive coupons are rare (co-occurring in $< 1\%$ of "treatment" observations), and their spillover effects are negligible compared to the main focal effects. While temporal spillovers are harder to rule out *a priori*, the evidence from the first-observation sample confirms that potential temporal dependencies do not affect our conclusions.

4.3 OLS Implementation

We implement the proposed identification approach using an OLS regression framework. Specifically, we stack the observations from the *NoRandom*, *NonFocal*, and *Focal* groups, and then estimate the following regression model:

$$Y_i = \alpha + \beta_F \cdot \text{Focal}_i + \beta_{NF} \cdot \text{NonFocal}_i + \boldsymbol{\gamma} \cdot \text{Controls}_i + \varepsilon_i, \tag{3}$$

where Focal_i and NonFocal_i are indicator variables for the respective experimental groups, and Controls represents a vector of covariates. We then calculate the estimated average treatment effect for the focal brand by combining the estimated coefficients:

$$\widehat{\text{ATE}} = \hat{\beta}_F - \frac{1}{p} \cdot \hat{\beta}_{NF}. \tag{4}$$

Without controls, this regression-based approach is mathematically equivalent to the estimator in Equation (2). The baseline group is *NoRandom*, so the intercept α represents the average outcome for customers who received only targeted coupons. Because customers are randomly assigned to groups, $\hat{\beta}_F$ captures the difference $\bar{Y}(\text{Focal}) - \bar{Y}(\text{NoRandom})$, while $\hat{\beta}_{NF}$ captures the difference $\bar{Y}(\text{NonFocal}) - \bar{Y}(\text{NoRandom})$.

To improve estimation precision, we include (customer×date)-specific covariates as controls, which absorb residual variance in the purchase outcomes Y_i. The controls include RFM (Recency, Frequency, Monetary) measures calculated over a 90-day window prior to the coupon receipt. We also cluster standard errors at the customer level. This is necessary because the same customer may appear in the sample on multiple dates, potentially leading to within-customer correlation across observations.

The regression framework supports the three outcome measures introduced in Table 2. In Section 6, we also use this approach to calculate category-level treatment effects. In particular, we estimate separate ATEs with category-level outcome measures for each campaign, and then average these ATEs across campaigns within each category.[5]

5. Brand Treatment Effects

This section evaluates the impact that in-store coupons have on sales of the promoted (focal) brands. We first report the distribution of ATEs across the 101 CPG campaigns. We then investigate how these effects vary depending on customers' previous experience with the brand and category. The section concludes by analyzing how the treatment effects vary with the depth of the discount, offering insight into the primary mechanisms driving in-store coupon effectiveness.

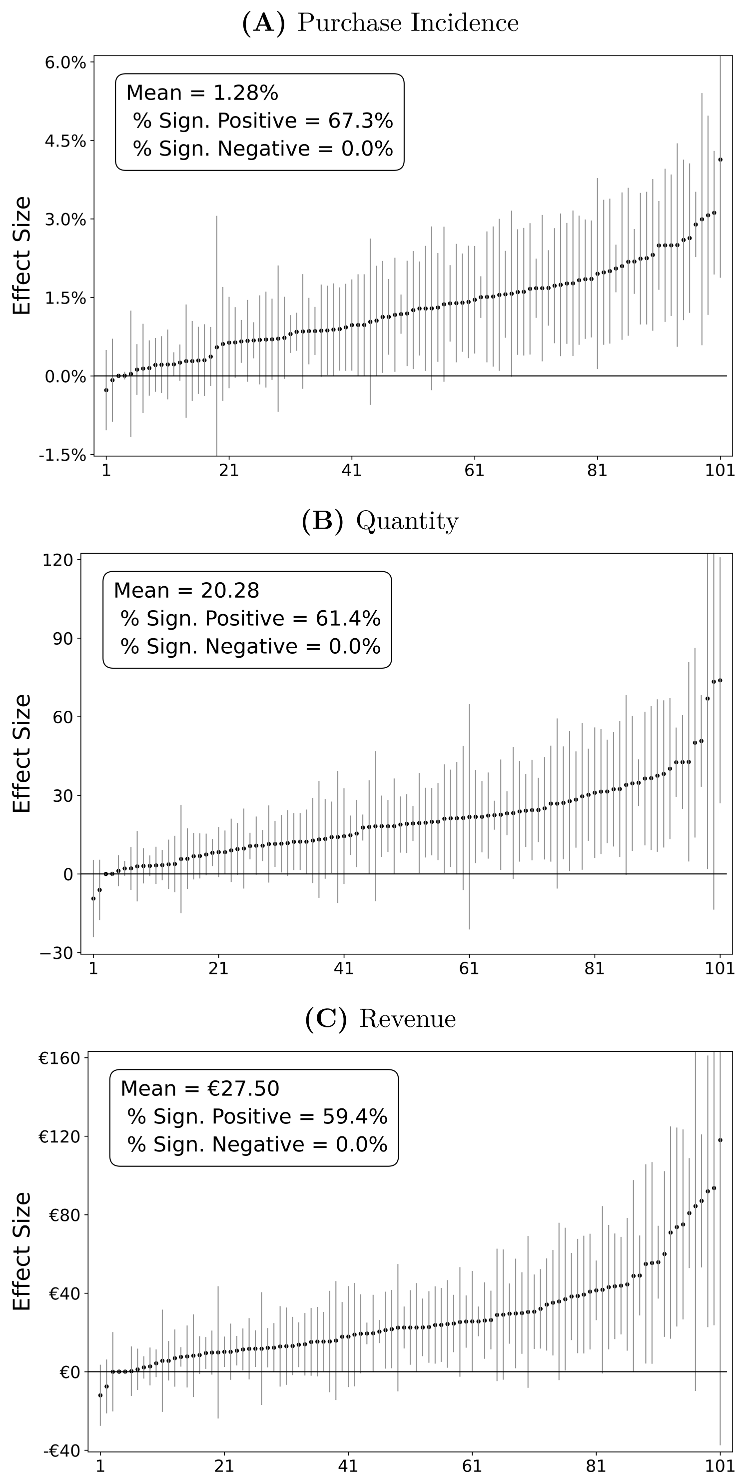
5.1 Distribution of Brand ATEs

In Figure 3, we report the distribution of the ATEs for the promoted brands. Each data point represents a treatment effect estimate for a single campaign on the day that the customer received the coupon, and the error bars indicate 95% confidence intervals. Panel A measures the percentage increase in purchase incidence, while Panels B and C report the incremental quantity (units) and revenue lift (Euros) per 1,000 coupons (per mille). For example, an effect size of €10 indicates that the focal brand's average revenue per 1,000 coupon impressions was €10 higher than the average revenue per 1,000 kiosk visits without a coupon for the focal brand.

In-store coupons drive incremental sales for almost all brands. The purchase incidence treatment effects are positive for 99 of the 101 campaigns, and 67.3% of the effects are statistically significant ($p < 0.05$). We observe a very similar pattern for the quantity and revenue treatment effects. Overall, the average ATE across the 101 brands is 1.28% incremental purchase incidence,

[5] We also considered an alternative approach where we pool observations across campaigns within a category and estimate a single treatment effect. Both approaches yield very similar estimates.

Figure 3: Distribution of Average Treatment Effects (ATEs)



Notes: The figure summarizes the distribution of average treatment effects for the 101 brands. Panel A reports the incremental purchase incidence for each brand (in percentage points). Panel B reports the incremental quantity sold (in units per 1,000 coupons). Panel C reports the incremental revenue (in € per 1,000 coupons). Error bars indicate 95% confidence intervals.

20.28 incremental units, and €27.50 additional revenue per 1,000 coupons. We do not find any negative treatment effects that are statistically significant for any campaign across any measure.

The revenue measure incorporates coupon discounts, which average 23.95% across the sample. To offset this price reduction, coupons need to increase purchase quantities by a factor of 1.31x to ensure that revenue breaks even. This is a substantial threshold, which might make some manufacturers hesitant to allocate budget to this channel. However, despite this breakeven threshold, in-store coupons yield revenue gains for essentially every brand in our sample. In-store coupons drive a 4.9x average increase in purchase incidence, a 6.3x increase in quantity sold, and a 5.2x increase in revenues.

Table 3 illustrates these effects using six globally recognized brands. In-store coupons are effective in driving sales for all six brands. For example, Coke's revenue increases by 86%, and coupons yield even more substantial gains for other brands, with sales lifts for Maggi spice mixes exceeding 16x for quantity and 10x for revenue. As expected, revenue lifts are generally smaller than quantity lifts because the revenue measure accounts for the price reduction from the discount.

The Coke example is particularly interesting. Customers purchase Coke on 1.7% of shopping trips without Coke coupons, which is among the highest purchase rates for brands in our sample. This creates a risk that in-store coupons provide discounts to customers who would purchase anyway. We find that despite the high baseline purchase rate, the incremental demand for Coke is large enough to offset the discounting.

Table 3: Treatment Effects for Six Global Brands

	Incidence		**Quantity**		**Revenue**	
	No Coupon	ATE	No Coupon	ATE	No Coupon	ATE
Ben's Original rice	0.14%	1.38%	1.74	18.25	€3.05	€26.34
Coke soft drinks	1.70%	1.29%	27.09	26.87	€26.00	€22.45
Johnnie Walker whisky	0.13%	0.97%	1.35	9.48	€12.81	€84.36
Lindt chocolates	0.11%	0.86%	1.47	10.83	€4.29	€24.57
Maggi spice mix	0.14%	1.40%	1.90	31.44	€2.19	€23.85
Schauma shampoo	0.08%	0.69%	0.84	9.00	€1.85	€13.03

Notes: The table reports the ATEs and outcomes in the condition without a random coupon for six globally recognized brands.

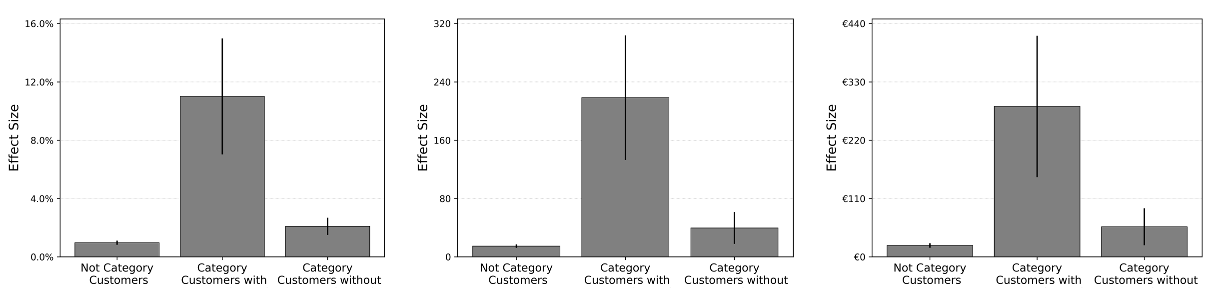
To systematically investigate the variation in ATEs across the 101 campaigns, we analyze how treatment effects correlate with four brand features: penetration, loyalty, price position, and interpurchase time. These features are constructed using the retailer's loyalty card data. The strongest relationship is for brand penetration; brands with higher baseline penetration tend to yield larger ATEs. One explanation is that some customers do not need some categories, and in-store coupons cannot change this. If few customers have pets, then we would expect both a low penetration rate for pet food brands and a relatively small response to pet food coupons. Other brand features are less predictive of the ATEs. We discuss these results in Web Appendix D, where we also provide precise definitions of the four brand features.

5.2 Heterogeneity in Treatment Effects Across Customers

Many retailers allow brands to choose whether in-store coupons are distributed to past brand customers, customers of competing brands, or customers who have not recently purchased in the category. To investigate the variation in treatment effects across these options, we use loyalty data in the three months before the coupon experiment to group customers into three customer segments: *Not Category Customers* (80%), *Category Customers with Promoted Brand* (3%), and *Category Customers without Promoted Brand* (17%). We then calculate conditional average treatment effects (CATEs) separately for each segment, which we then average across the 101 brands. In Figure 4, we report the CATEs for the promoted brand (on the current trip) by each segment.

Figure 4: Heterogeneity in Treatment Effects for the Promoted Brand

(A) Purchase Incidence **(B)** Quantity **(C)** Revenue



Notes: The figures report the average ATEs (on the current trip) averaged across the 101 focal campaigns. Each column represents a different customer segment. Error bars indicate 95% confidence intervals.

Figure 4 reveals an important distinction between segment-level responsiveness and the aggregate composition of the treatment effects. The largest treatment effects occur for the *Category*

Customers with Promoted Brand segment. Among these customers, the promoted brand's revenue increased by over €283 per mille. However, significant lifts also appear among the customers who had only purchased other brands in the category in the previous three months (€57 per mille) or had not purchased in the category in that prior period (€21 per mille). Notably, while existing brand buyers are the most responsive per-coupon, they represent only 3% of the population. Because the *Category Customers with Promoted Brand* and *Not Category Customers* segments are substantially larger, they account for the vast majority of the aggregate revenue lift. These findings suggest that the value of in-store coupons lies in their ability to also drive acquisition and category expansion, rather than just deepening loyalty among existing brand buyers.

5.3 Discount Depth

Recall that our experimental design randomized two coupon features. First, we varied which brand was assigned to individual customers so that one customer received a coupon for Coke, whereas another customer received a coupon for detergent. Second, we also randomized the discount level. For example, one customer received a 15% discount for Coke, whereas another customer received a 30% discount for Coke. We randomly assigned a discount with equal probability from the discount levels that the manufacturer made available for that brand. The exogenous variation in discount levels enables us to measure (causal) treatment effects as a function of discount depth.

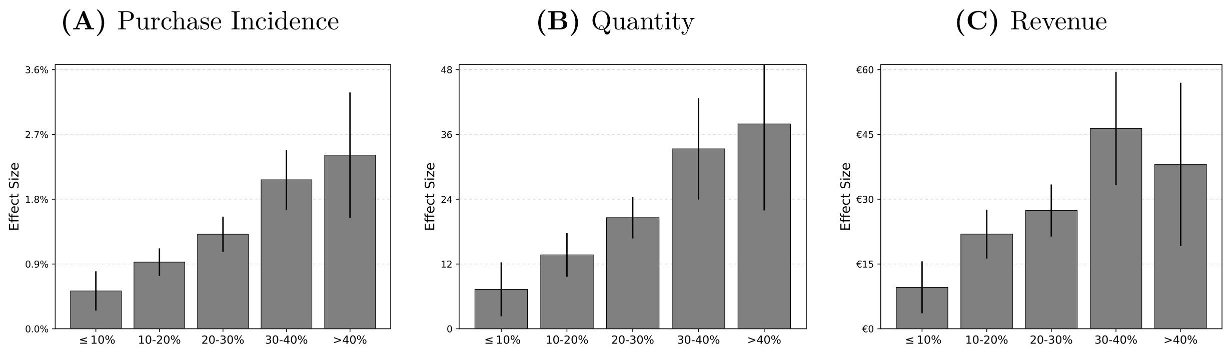
We start by segmenting observations by discount levels. We bin observations into five buckets each spanning 10 percentage points: $\leq$10%, (10%, 20%], (20%, 30%], (30%, 40%], and >40%. Using the identification strategy in Section 4, we calculate campaign-level treatment effects within each bucket and report the average ATEs in Figure 5.[6]

Figure 5 reports positive purchase incidence and quantity treatment effects for all discount levels (Panels A and B). Treatment effects increase monotonically with the discount depth, so that larger discounts lead to larger treatment effects, but this trend flattens at the highest discount levels. The differences in treatment effects are statistically significant between small and large discounts, but not between the two largest discount bins. Previous research has reported similar patterns for coupon redemption rates (Dhar and Hoch 1996; Danaher et al. 2015; Molitor et al.

[6]Some brands do not distribute coupons with the highest and lowest discount values. We present the results for the subset of brands that use all possible discount buckets in Web Appendix E.

2016).

Figure 5: Treatment Effects at Each Discount Level



Notes: The figure depicts treatment effects as a function of the discount depth. Treatment effects within each discount bucket are first calculated at the brand level. The bars represent means calculated by averaging across the brands. Error bars indicate 95% confidence intervals calculated using a non-parametric bootstrap.

The revenue treatment effects are also all positive and statistically significant (Panel C). They increase monotonically for discounts up to 40%. However, for discounts above 40%, revenue treatment effects eventually start to decrease because the gains in purchase rates are offset by customers paying lower prices because of discounts (Neslin 1990).

The evidence that customers are more likely to purchase when they receive larger discounts is consistent with standard arguments that customers are price sensitive. They confirm that the discount itself is an important driver of the increase in sales of the focal brand. However, treatment effects in Figure 5 remain positive even at minimal discount levels (<10%). This suggests that the coupon also serves an "exposure" or "advertising" function that increases brand salience independent of the price incentive. This is consistent with previous research that has argued that price promotions can enhance customer engagement and help to shape consumer behavior through psychological and social mechanisms (Shimp and Kavas 1984; Inman et al. 1990).

Although both mechanisms could contribute to coupon effectiveness, the price effect appears to be the primary driver of the total lift. If we interpret the treatment effect at the lowest discount bin as a proxy for the exposure effect, the substantial and monotonic increase in lift as discounts move toward 40% indicates that the price component accounts for the majority of the incremental sales. We conclude that mere exposure to in-store coupons may serve as a valuable point-of-sale "nudge", but significant shifts in brand performance remain largely dependent on the depth of the price incentive.

5.4 Discussion

Our findings reveal a consistently large increase in purchases of the promoted brand on the day customers received the coupon. The volume increase is sufficiently large to offset the coupon discount, leading to revenue gains for nearly all 101 brands. The consistency of the revenue result is particularly remarkable given that in-store coupons can be exercised immediately, which results in high redemption rates. In our empirical context, coupons are automatically redeemed when customers buy the promoted product. The revenue increase occurs despite this 100% "redemption rate".

Our findings highlight the versatility of in-store coupons. Different brands have different objectives for their promotions. Some are focused solely on expanding top-line sales, while others want to induce brand switching or attract new customers into the category. In-store coupons are an effective tool for achieving these different goals. The sales increases for the promoted brand are not limited to the promoted brand's past customers. Sales of the promoted brand also increase for past customers of other brands, and for customers who had not recently purchased in the category. This suggests that in-store coupons are as much an acquisition and expansion tool as they are a loyalty incentive.

The purchase incidence treatment effects increase monotonically with the depth of the coupon discounts. We interpret this as evidence that customer price sensitivity is a primary driver of these treatment effects. However, positive treatment effects at even nominal discount levels suggest a dual mechanism. While price incentives affect sales most significantly, the coupon's role as a point-of-sale "nudge" likely enhances customer engagement and brand salience, providing a baseline lift independent of the discount magnitude.

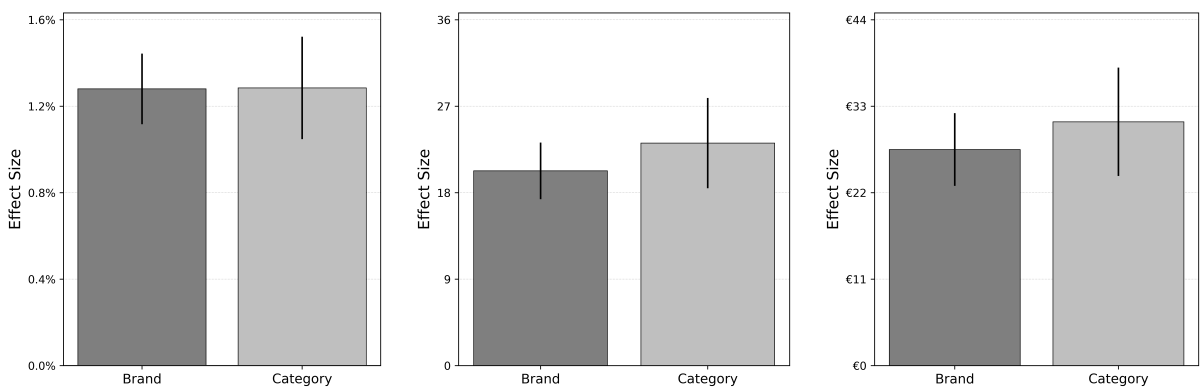
6. Impact of the In-Store Coupons on Category Sales

While increasing a brand's sales is valuable for a manufacturer, retailers care about the performance of the entire category. This section shifts the focus from brand-level outcomes to evaluating whether in-store coupons contribute to category expansion. We use the identification approach for category-level outcomes that we described at the end of Section 4.

Figure 6 provides consistent evidence of category expansion that matches the brand uplifts across all metrics. The increases in the *Quantity* outcome measures are particularly notable. The *Quantity* gains cannot be explained by price differences in the category, either due to the discounts in the treatment condition, or due to switching between higher or lower priced products in the category. The positive Revenue ATEs further confirm that the quantity increase more than compensates for the price reductions offered by the coupons. We provide a complete list of product categories and category-level ATEs in Web Appendix A.

Figure 6: Brand and Category ATEs

(A) Purchase Incidence **(B)** Quantity **(C)** Revenue



Notes: The figures report the average Brand and Category ATEs (on the current trip) averaged across the 71 product categories. Error bars indicate 95% confidence intervals.

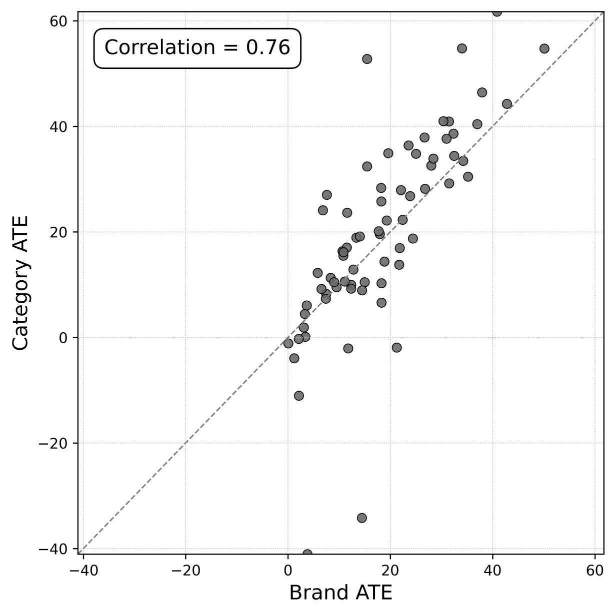
Figure 7 illustrates the relationship between brand-level and category-level quantity ATEs. The strong pairwise correlation ($\rho = 0.76$, $p < 0.01$) suggests that campaigns driving the largest brand gains are also the most effective at expanding the category. The close alignment of the observations around the 45-degree line also highlights that the magnitude of the brand effects is similar to the magnitude of the category effects. This implies that category expansion is a primary source of the incremental brand sales.

In the remainder of this section, we investigate the drivers and boundary conditions of this category expansion. To rule out price effects, we focus our analysis on *Quantity* as the outcome measure.

6.1 Temporal Substitution

A common concern with price promotions is they can result in a "post-promotion dip": immediate sales gains are offset by reduced demand in subsequent weeks (see, for example, Neslin et al.

Figure 7: Scatter Plot of Brand ATEs and Category ATEs



Notes: The figure reports a scatter plot of brand-level ATEs (averaged by category) and category-level ATEs. All ATEs are measured using the *Quantity* outcome measure. The unit of observation is a category.

1985; Neslin and Schneider Stone 1996; Van Heerde et al. 2000; Hendel and Nevo 2003; Macé and Neslin 2004). To investigate whether in-store coupons pulled demand forward resulting in a post-promotion dip, we repeated our category-level analysis using longer measurement windows that combine outcomes on the current visit with outcomes on future visits.[7]

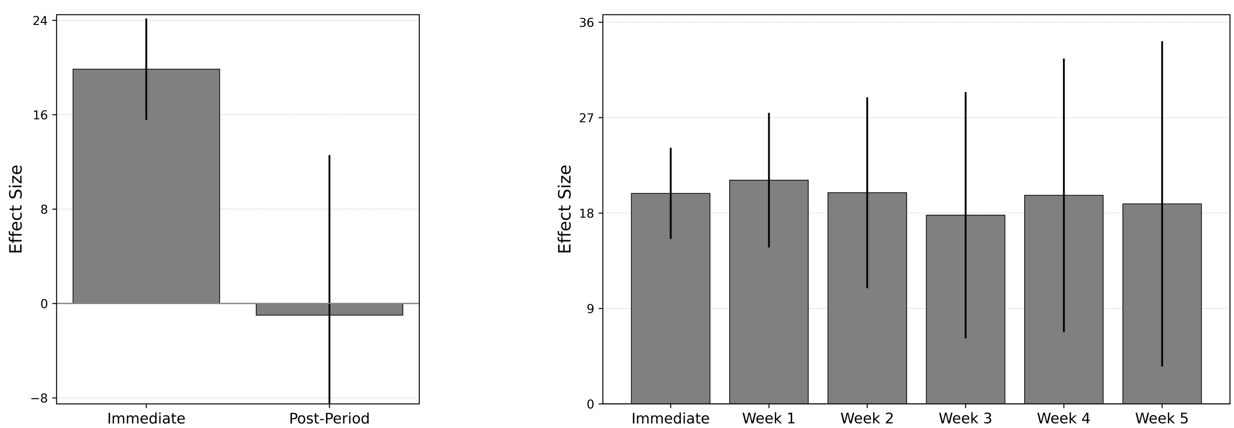
In Figure 8, we report the post-promotion ATEs using two different approaches. Panel A summarizes outcomes for the "Immediate" ATE on the day of the coupon distribution, and the 5-week period starting the day after the distribution date ("Post-Period"). The "Immediate" ATE on the day of coupon distribution is positive and significant, while the "Post-Period" ATE (the following 5 weeks) is essentially zero.

Panel B reports cumulative findings using six measurement windows: the day the coupon is distributed ("Immediate"), the distribution date and the next six days ("One Week"), the distribution date and the next thirteen days ("Two Weeks"), etc. This cumulative analysis shows no

[7]We emphasize that our goal in studying post-treatment effects is to investigate whether the evidence of category expansion in Figure 6 is merely due to temporal substitution (e.g. forward-buying). We recognize that post-treatment effects could also result from changes in firm actions after the treatment, such as more frequent targeted coupons. We are agnostic to the source of these effects, and instead focus on whether they offset the immediate category expansion. We also note that for firms the key question is: if a customer receives a coupon today, what is the total future effect on future purchasing (including changes that result from future firm actions)?

evidence of erosion in the initial sales lift over time, suggesting that the observed category expansion primarily reflects incremental demand rather than simple purchase acceleration.[8]

Figure 8: Temporal Substitution: Immediate and Post-Period Category-Level ATEs



(A) Distinct Periods **(B)** Cumulative Periods

Notes: These figures plot category-level *Quantity* ATEs averaged across the 71 categories. In Panel (A), we report separate findings for the Immediate and Post-Period. In Panel (B), we report cumulative findings. Error bars indicate 95% confidence intervals.

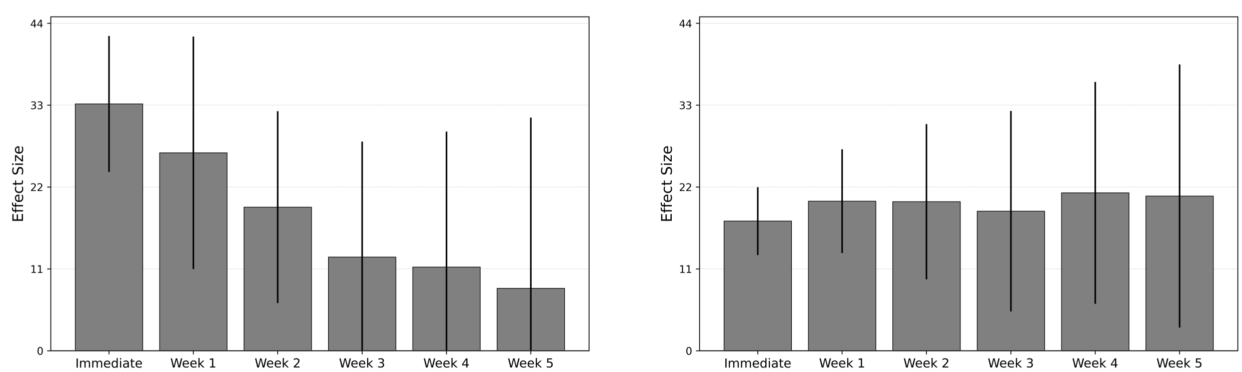
While the aggregate effect shows no dip, we tested whether specific categories prone to stockpiling behaved differently. Using the Narasimhan et al. (1996) two-item *Stockpiling* scale, we found no significant relationship between a category's suitability for storage and its Immediate or Post-Period ATE. We do observe an indication of a post-promotion dip in suburban stores, where stockpiling is presumably more common. In urban stores, it is clear that coupon effectiveness extends beyond any stockpiling effects. These results are reported in Appendix C, and the data collection for the stockpiling variable is described in Web Appendix F.

To explore this further, we compared perishable versus non-perishable categories. The 11 perishable categories include ready-to-eat meals, dairy products, and (cooked) meat sausages (see a complete list in Web Appendix A). Intuitively, shelf-stable goods should be more susceptible to stockpiling and thus more likely to show a post-promotion dip. However, Figure 9 reveals the opposite: in perishable categories (e.g., milk, yogurt, sausages), the immediate gains are almost completely offset in the Post-Period. In contrast, non-perishable categories show no such erosion, with the overall effect remaining stable over time.

[8] At the category level, post-promotion dips remain significant only for two of the 71 categories (using Bonferroni correction): alcohol chocolates and potato pancakes.

This divergence suggests that the post-promotion dip in our setting may be driven by consumption shifts rather than household inventory behavior. An immediate surge in consumption (e.g., eating more chocolate) may lead to subsequent satiation or a desire for variety, reducing demand in the following weeks. For shelf-stable items, customers may maintain large enough inventories to buffer these short-term consumption fluctuations. However, in perishable categories, purchase timing is tightly linked to consumption. While we do not directly observe consumption, products that can be consumed immediately often belong to high-impulse categories (e.g., chocolate). In Web Appendix G, we observe some evidence of a negative correlation between Post-Period ATEs and measures of whether the category is susceptible to impulse purchasing (including the two-item impulse survey scale by Narasimhan et al. 1996).

Figure 9: Category ATEs: Perishable vs. Non-Perishable Categories



(A) Perishable Categories **(B)** Non-Perishable Categories

Notes: The figures report cumulative category-level *Quantity* ATEs averaged across the 11 perishable categories (Panel A) and 60 non-perishable categories (Panel B). Error bars indicate 95% confidence intervals.

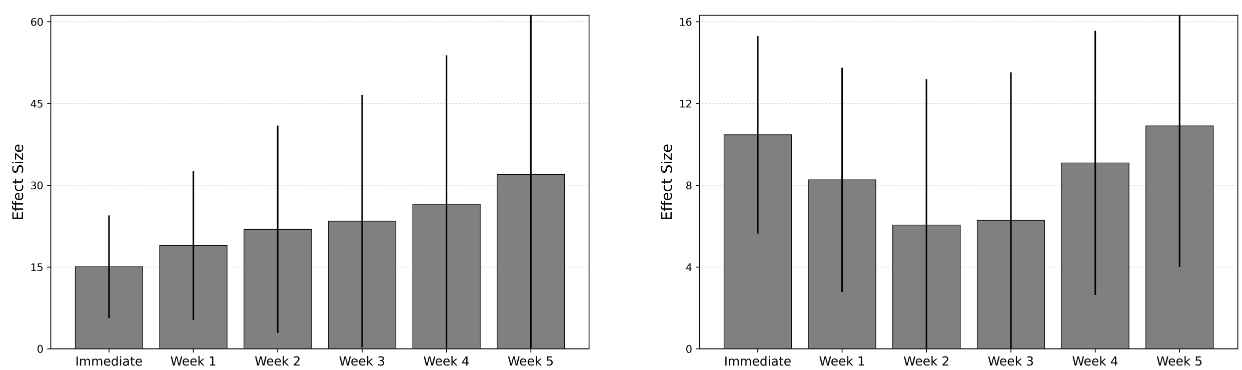
While we do find some examples of temporal substitution, this mechanism cannot fully explain the category expansion we observe on the day the in-store coupons are received. When averaging across all 71 categories, the in-store coupons did not result in a post-promotion dip. We next investigate an alternative explanation for category expansion: customers switching purchases from competing retailers.

6.2 Store Switching

Recall that prior research investigating category expansion distinguishes between products with fixed consumption rates (e.g., laundry detergent) and products for which consumption can expand (e.g. yogurt). To motivate our discussion of store switching, we first examine 13 product categories where consumption appears inelastic: deodorants, facial tissues, hair sprays, paper towels, shampoo, shower gels, toilet paper, toothpaste, dishwasher detergents, fabric softeners, laundry detergents, cat food, and dog food. In all of these categories, increased sales are unlikely to be driven by consumption expansion. For example, a coupon for toothpaste or pet food is unlikely to induce a customer to use more of the product or acquire a new pet.

Figure 10 shows category expansion in these categories. The immediate effects are not followed by a post-promotion dip. This pattern is particularly evident among customers without a category purchase at the focal retailer in the three months before the coupon exposure (Panel B). For this group, the cumulative effect remains positive and significant five weeks post-promotion. This category expansion cannot be attributed to temporal substitution and is unlikely to result from customers increasing their consumption.[9]

Figure 10: Categories in Which We Do Not Anticipate Changes in Consumption



(A) All Customers

(B) Customers Without a Category Purchase

Notes: The figures report cumulative category-level *Quantity* ATEs averaged across nine categories in which we do not expect coupons to cause changes in consumption. Panel (A) includes all customers, Panel (B) includes customers without a prior category purchase at the focal retailer in the 90 days before customers received the in-store coupon. Error bars indicate 95% confidence intervals.

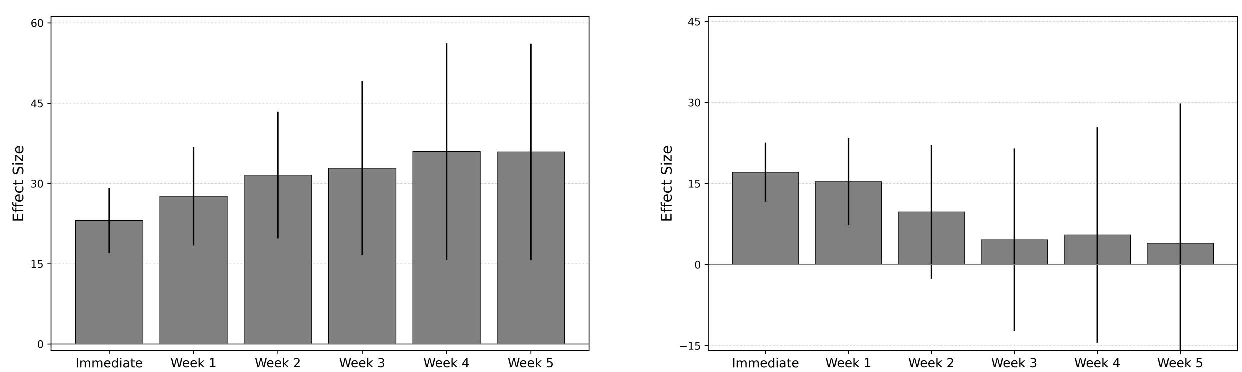
One explanation for these findings is store switching, where in-store coupons enable the focal retailer to capture purchases that would otherwise occur at competing retailers (on future trips). We provide two additional analyses to support this explanation. First, we compare treatment effects calculated using purchases at stores located close to competing stores versus stores not located

[9]Notice that we do not conclude that there is no consumption expansion or temporal substitution, just that these explanations are not sufficient to fully explain the category expansion. Recall also that because we focus on *Quantity* purchased, brand switching is (largely) ruled out as an explanation.

close to competing stores. If store switching plays an important role, its impact should vary with the geographic density of the retail stores. Specifically, in focal stores located near competitors, we expect (a) larger immediate category expansion and (b) a smaller post-promotion dip at the focal retailer, as the "dip" is effectively externalized to the competitor's store. To test this, we divided our stores into two groups based on a median split of the distance to the nearest competing supermarket.[10]

The findings in Figure 11 support these predictions. In stores located near competitors (Panel A), we observe robust category expansion that persists over the extended measurement window. This pattern is consistent with customers shifting purchases from nearby competing stores in response to the in-store coupon. Conversely, in stores farther from competitors (Panel B), the initial category expansion is followed by a notable erosion of the effect. In these more isolated stores, the initial surge in demand is more likely to reflect substitution from the focal retailer's own future sales.

Figure 11: Competing Stores Close or Far From Focal Retailer's Stores



(A) Competing Stores Close

(B) Competing Stores Far

Notes: The figures report category-level *Quantity* ATEs averaged across the 71 categories. In Panel (A), we calculate ATEs using purchases at the focal retailer's stores located close to competing stores. In Panel (B), we calculate ATEs using purchases at stores located far from competing stores. Error bars indicate 95% confidence intervals.

To further investigate store switching, we obtained household panel data covering purchases from grocery retailers for a sample of German households. Unlike our experimental data, this panel

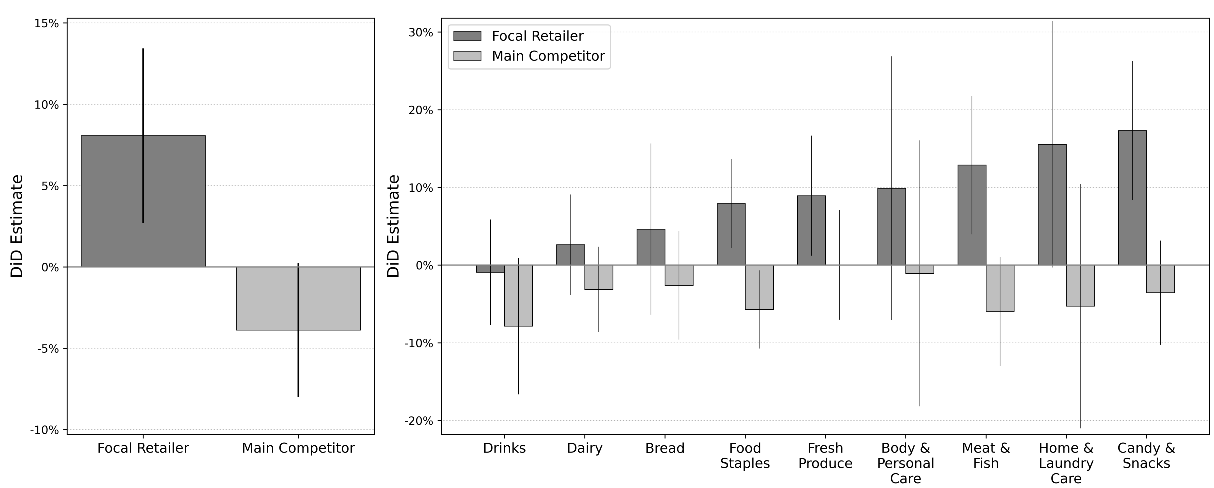
[10] Distance to the nearest competing supermarket is calculated using geographic coordinates (latitude and longitude). We obtain the coordinates for all focal and competing stores, compute the Haversine distance between each focal store and every competing store, and define the distance to the nearest competitor as the minimum of these distances.

captures shopping behavior in stores with and without the coupon system. The time period covered by the panel data spans the introduction of the coupon system at the focal retailer. This provides a quasi-experimental setting to compare quantities sold at the focal retailer before and after the system's introduction, using the stores without the system as a control group.

We conduct a difference-in-differences (DiD) analysis separately for each category to estimate the impact that the coupon system had on sales. We also repeat the analysis for the retailer's main competitor to determine if their sales moved in the opposite direction. Additional details about this DiD analysis are reported in Web Appendix H, where we also include a replication using a placebo date.

Figure 12 reports the impact of the new coupon system on sales. The left panel shows overall sales, while the right panel breaks down the results across the nine product groups used as category definitions (as provided by the panel data supplier). The DiD results reveal consistent evidence that the introduction of the coupon system led to an increase in the focal retailer's quantity sold, while simultaneously decreasing sales at the retailer's main competitor. These findings are consistent with store switching; the coupon system caused customers to switch purchases from the (main) competing retailer to the focal retailer.

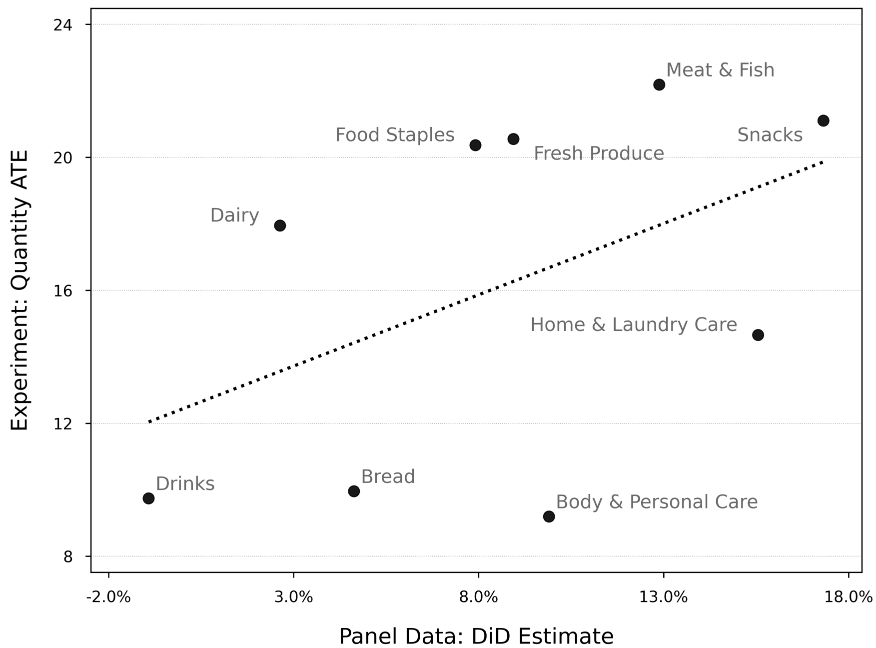
Figure 12: Diff-in-Diff Coefficients From Panel Data



Notes: The figures report diff-in-diff coefficients from the panel data. Error bars indicate 95% confidence intervals. See Web Appendix H for additional details regarding the panel data DiD analysis.

In additional analysis, we investigate whether the categories exhibiting the largest growth in sales (at the focal retailer) in the panel data correspond to those with the largest ATEs in our experiment. In particular, we matched the categories involved in our experiments with the nine product groups in the panel data and averaged the (immediate) category expansion ATEs across the categories within each group. In Figure 13, we compare the panel data DiD coefficients within each product group with the average category expansion ATEs from our experiments.

Figure 13: ATEs from Experiments and Diff-in-Diff Coefficients from the Panel Data



Notes: The figure reports a scatter plot of category-level ATEs from the experiments and diff-in-diff coefficients from the panel data. The nine product groups are defined by the supplier of the panel data.

We observe a strong positive relationship across the product categories. Product groups for which the in-store coupon system caused switching from competing retailers to the focal retailer (e.g., candy and snacks) are the same categories in which we observe the largest category expansion effects in our experiment. Across the nine groups, the (rank-order) correlation between the panel data DiD coefficients and the average experimental ATEs is 0.47.

What makes this relationship especially notable is that the results are obtained from completely separate datasets and using completely different methodologies. Our field experiment captures immediate purchases by customers who are already in the focal store, while the panel data reflects long-term behavior across multiple store visits. In our experiments, we measure the impact of randomly assigned coupons in the focal category, while the panel data measures the impact of (primarily) targeted coupons. Despite these differences, both studies provide convergent evidence of store switching.

6.3 Summary

Our analysis of category-level ATEs provides consistent evidence that in-store coupons increase category sales at the focal retailer on the day that customers receive them. The sales increase is reflected in purchase incidence, quantity sold and revenue. Category expansion cannot occur indefinitely, and so we explored both the mechanisms and boundaries of the effect. Although we do not find evidence of a post-promotion dip when averaging treatment effects across all brands, there is evidence of temporal substitution for perishable products, which may indicate changes in consumption rather than stockpiling. We also find evidence that store switching contributes to category expansion, with customers substituting future purchases at competing retailers for immediate purchases at the focal retailer.

7. Conclusions

A long-standing challenge for retailers is to find an instrument for delivering promotions to customers when they are about to purchase (at the bottom of the purchase funnel). While digital channels have solved this problem through sponsored search and on-site advertising, physical retail has historically resorted to direct mail, newspaper FSIs, and checkout coupons, which are received either before or after a shopping trip. By distributing coupons during shopping trips, in-store kiosks and mobile apps represent a fundamental shift in this landscape.

We investigate the effectiveness of in-store promotions using data from a large-scale field experiment that randomly varied coupon distribution within a business-as-usual environment. Coupons were distributed via kiosks, which loyalty program members typically visited at the start of their trips. The experimental data provide a causal measure of the impact of in-store coupons on customer purchasing. We summarize the key findings in Table 4.

In-store coupons are remarkably effective, causing a nearly fivefold increase in sales for promoted brands on average. The sales lifts are large enough to offset the price discounts, resulting in a net revenue increase for nearly all brands. The promotions successfully reactivated past brand customers, induced trial among competitors' customers, and attracted new customers to the cate-

Table 4: Summary of Key Findings

Promoted Brand Sales: In-store coupons substantially increase sales for promoted brands at the focal retailer. The incremental unit sales offset the price discounts, so that the revenue increases for nearly all campaigns.
Practical Versatility: In-store coupons can be used to increase top-line sales, induce trial among competitors' customers, and attract new customers to the category.
Category Expansion: In-store coupons lead to category expansion, with the magnitude of the category expansion closely matching the increase in sales of the focal brand.
Temporal Substitution: Temporal substitution occurs in some categories but does not appear to stem from stockpiling and is not sufficient to explain the broader category expansion observed across the study.
Store Switching: Store switching appears to contribute to category expansion, with customers purchasing immediate at the focal retailer instead of purchasing at competing retailers (on future visits).

gory. This makes in-store coupons a versatile marketing tool that can be used to both retain past customers and attract new customers.

The impact of in-store coupons extends beyond the promoted brand to drive significant category expansion at the focal retailer. Across the brands, the size of the focal brand effect closely aligns with the magnitude of the category expansion, suggesting a direct relationship between the two. Our analysis investigates two mechanisms that contribute to category expansion. In some categories, we find evidence of temporal substitution; however, this is not driven by traditional stockpiling. Instead, the data suggests a *shift in consumption*, where customers consume more now and less in the future.

The second mechanism contributing to category expansion is *store switching.* Customers purchase immediately from the focal store instead of making purchases at competing retailers on future visits. We provide two sources of evidence for this behavior. First, we show that the coupon effects more pronounced in the focal retailer's stores located close to competing retailers compared to those further away. Second, a difference-in-differences analysis using third-party panel data confirms that the introduction of the couponing system increased sales at the focal retailer at the expense of its main competitor.

These findings have different implications for manufacturers and retailers. For manufacturers, expanding sales of the promoted brand at first appears valuable. However, if the additional sales simply represent a substitution of sales from competing retailers (at possibly higher prices), the overall outcome may yield little improvement in overall unit sales or revenue. In contrast, the results are *positive* for retailers. In-store coupons yield significant category expansion in part by switching purchases from competing stores.

Limitations and Future Research

A limitation in our study is that we do not observe the wholesale prices that the retailer pays to manufacturers for their brands. This means we are unable to report how in-store coupons affect the incremental profits earned by the retailer or the manufacturers. Retailers and manufacturers could combine the treatment effects reported in this paper with their own wholesale price data to estimate changes in profits. A related question is how in-store coupons increase overall channel profits, and how the incremental profits are shared between retailers and manufacturers.

A feature of in-store coupons is that their effectiveness is limited to the customers who actually receive them; in our setting, loyalty program members who visit the kiosk at the start of a shopping trip. We offer two perspectives on this limitation. First, the (very) positive outcomes for the retailer in this study suggest we will continue to see innovations in distribution technologies, so that over time, a larger proportion of customers will receive them. Second, the incomplete adoption of the kiosk system may also be a valuable feature of the system. Like other retailers, this retailer intentionally uses its loyalty program to price discriminate. The coupons allow the retailer to charge lower effective prices to customers who have signed up for the loyalty program. Price discrimination is only effective when some customers do not get the discounts.

Our data describes the response to CPG coupons delivered via in-store kiosks. Many retailers have developed alternative in-store media channels to deliver promotions to customers, including retail apps, in-store video and displays. Examples can be found in consumer electronics (Best Buy), department store (Saks), fast food (Starbucks), home improvement (Home Depot), beauty (Ulta), and other retailers. Future research could study the performance of retail media in these settings.

The panel dataset we use to study store switching is large enough to study changes in aggregate purchases by product group, but it is not large enough to study store switching at the brand level. Data limitations are a common obstacle to measuring store switching. The lack of accurate market-level data means that store switching is often simply ignored in the promotions literature. Additional research investigating how promotions impact customer store visits and their purchases at different retailers would be a valuable contribution to this literature.

References

Abratt R, Goodey SD (1990) Unplanned buying and in-store stimuli in supermarkets. *Managerial and Decision Economics* 11(2):111–121.

Ailawadi KL, Harlam BA, Cesar J, Trounce D (2006) Promotion profitability for a retailer: the role of promotion, brand, category, and store characteristics. *Journal of Marketing Research* 43(4):518–535.

Ailawadi KL, Neslin SA (1998) The effect of promotion on consumption: Buying more and consuming it faster. *Journal of Marketing Research* 35(3):390–398.

Angrist JD, Imbens GW, Rubin DB (1996) Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91(434):444–455.

Bell DR, Chiang J, Padmanabhan V (1999) The decomposition of promotional response: An empirical generalization. *Marketing Science* 18(4):504–526.

Berthiaume D (2022) Survey: Online shoppers search for coupons. https://chainstoreage.com/survey-online-shoppers-search-coupons, accessed: 2025-01-16.

Bucklin RE, Lattin JM (1991) A two-state model of purchase incidence and brand choice. *Marketing Science* 10(1):24–39.

Chiang J (1991) A simultaneous approach to the whether, what and how much to buy questions. *Marketing Science* 10(4):297–315.

Chiang J (1995) Competing coupon promotions and category sales. *Marketing Science* 14(1):105–122.

Chiou-Wei SZ, Inman JJ (2008) Do shoppers like electronic coupons?: A panel data analysis. *Journal of Retailing* 84(3):297–307.

Danaher PJ, Smith MS, Ranasinghe K, Danaher TS (2015) Where, when, and how long: Factors that influence the redemption of mobile phone coupons. *Journal of Marketing Research* 52(5):710–725.

Dekimpe MG, Hanssens DM, Silva-Risso JM (1998) Long-run effects of price promotions in scanner markets. *Journal of econometrics* 89(1-2):269–291.

Dhar SK, Hoch SJ (1996) Price discrimination using in-store merchandising. *Journal of Marketing* 60(1):17–30.

Fok D, Horváth C, Paap R, Franses PH (2006) A hierarchical bayes error correction model to explain dynamic effects of price changes. *Journal of Marketing Research* 43(3):443–461.

Grewal D, Ahlbom CP, Noble SM, Shankar V, Narang U, Nordfält J (2023) The impact of in-store inspirational (vs. deal-oriented) communication on spending: The importance of activating consumption goal completion. *Journal of Marketing Research* 60(6):1071–1094.

Grover R, Srinivasan V (1992) Evaluating the multiple effects of retail promotions on brand loyal and brand switching segments. *Journal of marketing research* 29(1):76–89.

Gupta S (1988) Impact of sales promotions on when, what, and how much to buy. *Journal of Marketing research* 25(4):342–355.

Hamstra M (2023) Stop & shop pilots kiosks for coupons and more. https://www.supermarketnews.com/finance/stop-shop-pilots-kiosks-for-coupons-and-more, accessed: 2025-01-16.

Heilman CM, Nakamoto K, Rao AG (2002) Pleasant surprises: Consumer response to unexpected in-store coupons. *Journal of Marketing Research* 39(2):242–252.

Hendel I, Nevo A (2003) The post-promotion dip puzzle: what do the data have to say? *Quantitative Marketing and Economics* 1:409–424.

Hui SK, Inman JJ, Huang Y, Suher J (2013) The effect of in-store travel distance on unplanned spending: Applications to mobile promotion strategies. *Journal of Marketing* 77:1–16.

Inman JJ, McAlister L, Hoyer WD (1990) Promotion signal: proxy for a price cut? *Journal of consumer research* 17(1):74–81.

Johnson GA, Lewis RA, Nubbemeyer EI (2017) Ghost ads: Improving the economics of measuring online ad effectiveness. *Journal of Marketing Research* 54(6):867–884.

Lim J, Currim IS, Andrews RL (2005) Consumer heterogeneity in the longer-term effects of price promotions. *International Journal of Research in Marketing* 22(4):441–457.

Macé S, Neslin SA (2004) The determinants of pre-and postpromotion dips in sales of frequently purchased goods. *Journal of Marketing Research* 41(3):339–350.

Mela CF, Jedidi K, Bowman D (1998) The long-term impact of promotions on consumer stockpiling behavior. *Journal of Marketing Research* 35(2):250–262.

Molitor D, Reichhart P, Spann M (2016) Location-based advertising: Measuring the impact of context-specific factors on consumers' choice behavior. *In Proceddings of the Thirty Seventh International Conference on Information Systems (ICIS)* .

Narasimhan C, Neslin SA, Sen SK (1996) Promotional elasticities and category characteristics. *Journal of Marketing* 60(2):17–30.

Neslin SA (1990) A market response model for coupon promotions. *Marketing Science* 9(2):125–145.

Neslin SA, Henderson C, Quelch J (1985) Consumer promotions and the acceleration of product purchases. *Marketing Science* 4(2):147–165.

Neslin SA, Schneider Stone LG (1996) Consumer inventory sensitivity and the postpromotion dip. *Marketing Letters* 7:77–94.

Nijs VR, Dekimpe MG, Steenkamp JBE, Hanssens DM (2001) Tracing the impact of price promotions across categories. *Tilburg University the Netherlands Working paper* .

Osuna I, González J, Capizzani M (2016) Which categories and brands to promote with targeted coupons to reward and to develop customers in supermarkets. *Journal of Retailing* 92(2):236–251.

Pollay RW (1968) Customer impulse purchasing behavior: A reexamination. *Journal of Marketing Research* 5(3):323–325.

Raju JS, Dhar SK, Morrison DG (1994) The effect of package coupons on brand choice. *Marketing Science* 13(2):145–164.

Rook DW (1987) The buying impulse. *Journal of Consumer Research* 14(2):189–199.

Rubin DB (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66(5):688.

Shapiro MF (2001) Impulse purchases. *NPN, National Petroleum News* 93(7):40–40.

Shimp TA, Kavas A (1984) The theory of reasoned action applied to coupon usage. *Journal of consumer research* 11(3):795–809.

Statista (2026) Companies with highest digital advertising revenues worldwide in 2025. URL www.statista.com/statistics/1285405/revenues-digital-ad-major-internet-companies, accessed: 2026-02-24.

Van Heerde HJ, Leeflang PS, Wittink DR (2000) The estimation of pre-and postpromotion dips with store-level scanner data. *Journal of Marketing Research* 37(3):383–395.

Zhou L, Wong A (2004) Consumer impulse buying and in-store stimuli in chinese supermarkets. *Journal of International Consumer Marketing* 16(2):37–53.

Appendix A. Sample Selection Criteria

Our analysis uses two sample selection criteria. The first criterion focuses on estimation precision. We restrict our analyses to the brands that had regular sales prior to the experiment and distributed at least 400 random coupons. This criterion eliminates new products and short campaigns (e.g., those lasting only one or two days).

The second sample selection criterion recognizes that brands could define eligibility criteria for their coupons. For example, a pet food brand only distributed coupons to customers who had purchased pet food before. Random coupons were not distributed to non-eligible customers, so we cannot measure *average treatment effects* for these brands. Not all brands used eligibility criteria, and the retailer did not record the eligibility criteria defined by brand managers.

We detect which brands had eligibility criteria by comparing the pretreatment variables across experimental conditions. In particular, we regress coupon assignments on three pretreatment variables, and identify brands for which pretreatment variables are not jointly significant in an F-test with a 95% confidence level. The pretreatment variables include the number of store visits, revenue, and the number of unique products purchased during the 90 days prior to the experiment (from August to October 2015). For each brand, we estimate two regression models that compare assignment into *Focal vs. NonFocal* and *Focal vs. NoRandom* (see Section 4), and we require that neither F-test is statistically significant.

Appendix B. Identification of Treatment Effects

Recall that in the main text, we define an *observation* as a customer×date combination, and we restrict our analysis to customer×date combinations in which the customer used the kiosk system (at least once). We denote a customer×date combination by $i \in 1, .., N$. We also define $Y_i(1)$ and $Y_i(0)$ as potential outcomes for a customer×date observation i, with and without a coupon for the focal brand (Rubin 1974). Our goal is to estimate the average treatment effect (ATE): $\text{ATE} = \mathbb{E}[Y_i(1) - Y_i(0)]$.

The proposed identification approach relies on two assumptions:

Assumption 1 (Stable Unit Treatment Value Assumption; *SUTVA*): The potential outcomes for an observation do not vary with the treatment assignments for other observations.

Assumption 2: Purchases of a focal brand are not affected by coupons from other brands.

Our first assumption recognizes that the same customer could appear in multiple customer×date observations in our data, and we treat these observations as independent. The second assumption implies that only focal coupons are relevant for estimating ATEs (instead of all seven coupons on the printout). Replacing a focal coupon with a random non-focal coupon moves from a potential outcome $Y_i(1)$ to a potential outcome $Y_i(0)$. These assumptions are standard in the causal inference literature.

Proposition: The following estimator provides an unbiased estimate of the average treatment effect for the focal brand:

$$\text{ATE} = \bar{Y}(\text{Focal}) - \left[\bar{Y}(\text{NoRandom}) - 1/p \cdot \left(\bar{Y}(\text{NoRandom}) - \bar{Y}(\text{NonFocal})\right)\right] \tag{B.1}$$

where $\bar{Y}(\cdot)$ indicates the average outcome in the corresponding group, and p captures the probability that a random *non-focal* coupon replaces a targeted *focal* coupon in the NonFocal_b group.

Proof. We can estimate the first term in the ATE as an average monetary outcome in the Focal condition. Focal observations are representative of the customer population, and in our experiment, they always received a coupon for the focal brand:

$$\mathbb{E}[Y(1)] = \mathbb{E}[Y_i \mid \text{Focal}] = \bar{Y}(\text{Focal}) \tag{B.2}$$

To estimate the expected potential outcome without a focal coupon, $\mathbb{E}[Y_i(0)]$, it is helpful to consider a group NoFocal, in which all targeted coupons for a focal brand are replaced by random coupons. Our data does not contain a group NoFocal. However, if we had such data, we could estimate $\mathbb{E}[Y_i(0)]$ as follows:

$$\mathbb{E}[Y_i(0)] = \mathbb{E}[Y_i \mid \text{NoFocal}] = \bar{Y}(\text{NoFocal})$$

The distribution of outcomes in NonFocal is a mixture of outcomes from NoFocal and outcomes from NoRandom, where a share of NoFocal observations is p. Indeed, in the NonFocal group, random coupons replace share p of targeted coupons in the printouts, and the remaining random coupons replace targeted coupons for non-focal brands, which does not change the outcomes for the focal brand. We can thus estimate $\mathbb{E}[Y_i(0)]$ as follows:

$$\mathbb{E}[Y_i \mid \text{NonFocal}] = p \cdot \mathbb{E}[Y_i \mid \text{NoFocal}] + (1-p) \cdot \mathbb{E}[Y_i \mid \text{NoRandom}] \tag{B.3}$$

$$\mathbb{E}[Y_i \mid \text{NoFocal}] = \mathbb{E}[Y_i \mid \text{NoRandom}] - 1/p \cdot (\mathbb{E}[Y_i \mid \text{NoRandom}] - \mathbb{E}[Y_i|\text{NonFocal}])$$

$$\mathbb{E}[Y_i(0)] = \bar{Y}(\text{NoRandom}) - 1/p \cdot (\bar{Y}(\text{NoRandom}]) - \bar{Y}(\text{NonFocal})) \tag{B.4}$$

We obtain Equation (B.1) by combining Equations (B.2) and (B.4). □

Variance

Under the SUTVA assumption, the observations in groups Focal_b, NonFocal_b, and NoRandom_b are independent, so the variance of the proposed estimator is

$$\mathbb{V}(\text{ATE}) = \mathbb{V}\left(\bar{Y}(\text{Focal})\right) + \left(\frac{p-1}{p}\right)^2 \mathbb{V}\left(\bar{Y}(\text{NoRandom})\right) + \left(\frac{1}{p}\right)^2 \mathbb{V}\left(\bar{Y}(\text{NonFocal})\right). \tag{B.5}$$

Illustrative Example

The second term in Equation (B.1) captures the average effect of targeted focal coupons across the entire customer population. It may seem counterintuitive that this term does not depend on the share of targeted focal coupons or the targeting rules themselves. We provide an illustrative example to clarify the underlying intuition for the estimator.

Let λ equal the proportion of targeted focal coupons in the *NoRandom* group, t equal the average treatment effect of a targeted coupon on the targeted population, and p equal a probability that a random coupon replaces a specific targeted coupon on a printout. We notice that t is generally not equivalent to ATE, because targeting rules introduce selection. The average impact of focal targeted coupons across the entire population is λt. We want to demonstrate that we can estimate λt using the difference in (per customer) outcomes between the *NoRandom* and *NonFocal* groups as:

$$\lambda t = \frac{1}{p}\left(\bar{Y}(\text{No Random}) - \bar{Y}(\text{NonFocal})\right) \tag{B.6}$$

Imagine 10,000 customers receive coupons. The platform randomly assigns 500 customers to receive random non-focal coupons, and the targeting algorithm assigns 1,000 customers to receive targeted focal coupons, meaning λ=10%. The remaining coupons received by these customers are targeted non-focal coupons; these coupons have no impact on purchases of the focal brand, and are irrelevant for our example. In our identification approach, these 10,000 customers constitute the *NoRandom* and *NonFocal* groups.

Because of the random assignment, 50 (10%) of the 500 non-focal random coupons are assigned to the same customers as the targeted focal coupons. Among these 50 customers, 1 out of 7 (i.e., 7.1 customers; $p = 1/7$) of the targeted focal customers are replaced by the random non-focal coupon. This results in the following *NoRandom* and *NonFocal* groups:

	NoRandom	**NonFocal**
Total customers	9,500	500
Targeted focal coupon:		
Yes	950 (10%)	42.9 (8.6%)
No	8,550 (90%)	457.1 (91.4%)

Notes: We are rounding $\lambda \cdot (1-p) \cdot 100\% = 0.1 \cdot \frac{6}{7} \cdot 100\% \approx 8.6\%$.

Let b indicate the baseline level of spending per customer without a coupon for the focal brand. The *NoRandom* average per customer outcome is: $b + \lambda t = b + 0.1t$ units, and the *NonFocal* average per customer outcome is: $b + (1-p)\lambda t = b + 0.086t$ units. This yields the following equivalence:

$$\frac{1}{p}\left[\bar{Y}(\text{NoRandom}) - \bar{Y}(\text{NonFocal})\right] = \frac{1}{p}\left[(b + 0.1t) - (b + 0.086t)\right] = \frac{1}{p}\left[\lambda t p\right] = \lambda t \tag{B.7}$$

This example illustrates how our adjustment to the identification in Equation 2 correctly accounts for the treatment effect introduced by targeted coupons. Our identification approach does not require the share of customers targeted by the focal coupons or the average treatment effects on the targeted. It builds on the average outcomes in the *NoRandom* and *NonFocal* conditions, weighted by how frequently targeted coupons are replace by random ones in the *NonFocal* group.

Appendix C. Category Expansion and Stockpiling Measures

In this appendix, we examine how category-level ATEs (Immediate and Post-Period) vary with category-level stockpiling characteristics: *Easy*, *Like*, and the composite *Stockpiling* score (see Web Appendix F). Table C.1 reveals that none of the correlations are statistically significant ($p > 0.10$).

Table C.1: ATE Correlations with Category Stockpiling Scores

(A) Category Purchasers

	Easy	**Like**	**Stockpiling**
Immediate ATE	−0.067 (0.580)	−0.093 (0.444)	−0.088 (0.469)
Post-Period ATE	−0.010 (0.934)	−0.052 (0.668)	−0.038 (0.758)

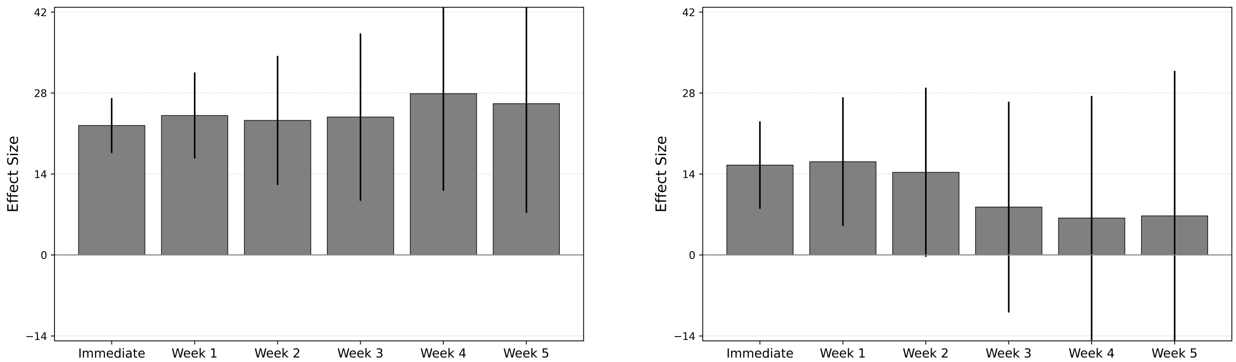
(B) All Respondents

	Easy	**Like**	**Stockpiling**
Immediate ATE	−0.056 (0.644)	−0.093 (0.445)	−0.082 (0.500)
Post-Period ATE	−0.059 (0.625)	−0.080 (0.508)	−0.076 (0.533)

Notes: The table reports the pairwise correlation between each ATE and each category-level characteristic. The sample size for each correlation is 71 (categories), and p-values are reported in parentheses.

In Figure C.1, we compare *Quantity* ATEs between urban and suburban stores. If stockpiling were a major driver of coupon effectiveness, we would expect suburban locations to exhibit larger immediate sales increases and deeper post-promotion dips since these customers have more room to stockpile. Instead, if anything, we observe larger immediate effects in urban locations. There is an indication of a post-promotion dip in suburban stores, but it is not statistically significant, and does not fully offset the initial gains. In urban stores, it is clear that coupon effectiveness extends beyond any stockpiling effects.

Figure C.1: Quantity ATEs in Urban and Suburban Stores



(A) Urban Stores **(B)** Suburban Stores

Notes: The figures report category-level *Quantity* ATEs averaged across the 71 categories. In Panel (A), we calculate ATEs using purchases at the focal retailer's stores located in urban areas. In Panel (B), we calculate ATEs using purchases at stores located in suburban areas. Error bars indicate 95% confidence intervals.

Web Appendix to

In-Store Coupons: A Large-Scale Field Experiment

February 24, 2026

A Category-Level Average Treatment Effects WA-2

B Randomization Checks WA-5

C Robustness to SUTVA and Competitive Coupons WA-6

D Summary Statistics for Brand Features WA-9

E Discount CATEs with Stratified Brand Sample WA-11

F Category-level Survey Scores WA-12

G Correlations: ATEs and Impulse Scores WA-14

H Panel Data Analysis WA-15

Web Appendix A Category-Level Average Treatment Effects

Table WA-A.1: Category-Level Average Treatment Effects

Group	Category	Incidence ATE	Quantity ATE	Revenue ATE
Drinks	**Drinks Alcoholic**			
	Rum blend	0.66%	12.86	€119.16
	Scotch whisky	1.02%	9.52	€82.29
	Spirit-based mixed drinks (premixes)	0.33%	6.06	€14.15
	Drinks Non Alcoholic			
	Carbonated soft drinks (soda)	1.16%	10.45	€11.19
	Energy drinks	–0.07%	8.93	–€2.36
	Fruit juices / sweet musts	1.73%	30.48	€27.17
	Iced tea / iced coffee	–0.02%	9.17	–€6.53
	Sparkling mineral water	–1.21%	–34.19	–€80.98
Food	**Bread Cake**			
	Packaged ready-made cakes	–0.35%	-3.94	–€10.77
	Packaged rusk (zwieback)	0.83%	9.96	€6.52
	Condiments			
	Cooking oil (not refrigerated)	1.50%	24.10	€56.18
	Gourmet sauces	0.68%	0.10	€23.26
	Seasonings (MSG etc.)	0.95%	40.92	€41.98
	Convenience			
	Ready meals [P]	1.80%	54.72	€45.14
	Soups, liquid	1.82%	52.75	€66.74
	Dairy			
	Cheese (self-service) [P]	3.06%	46.42	€50.02
	Dairy products (misc.) [P]	1.02%	33.45	–€0.66
	Fresh milk [P]	1.30%	54.76	€49.68
	Milk-based mixed drinks	2.12%	38.60	€105.61
	Quark / curd cheese [P]	2.01%	40.99	€23.60
	Quark desserts & preparations [P]	3.54%	33.89	€35.77
	UHT milk (long-life milk)	–3.33%	–41.08	–€25.40
	Yogurt (incl. preparations) [P]	0.96%	10.58	€23.44

Continued on next page

Table WA-A.1 – continued from previous page

Group	Category	Incidence ATE	Quantity ATE	Revenue ATE
	Frozen Food			
	Fish fillets (plain)	1.29%	22.16	€93.48
	Leaf & stem vegetables	1.64%	20.11	€29.80
	Pizzas & quiches	2.15%	40.41	€79.01
	Potato pancakes (latkes)	2.02%	22.29	€19.20
	Ready meals & casseroles	2.19%	37.65	€97.06
	Meat Sausages			
	Cooked sausage (sliced etc.) [P]	1.36%	13.78	€9.28
	Meat/ham sausage [P]	1.42%	15.49	€9.29
	Mortadella / bologna sausages [P]	1.41%	28.34	€27.67
	Raw sausage specialties (self-service) [P]	2.21%	32.42	€68.88
	Snacks			
	Chips & snack sticks	1.45%	32.55	€36.46
	Nuts, almonds & seeds/kernels	0.27%	7.34	€14.22
	Roasted peanuts	0.90%	18.73	€12.39
	Snack foods	1.50%	17.03	€19.24
	Staples			
	Baking mixes (household)	0.88%	19.59	€18.08
	Condensed milk	1.42%	34.89	€17.53
	Honey	2.06%	29.18	€62.25
	Jams / preserves	0.69%	–1.91	€5.12
	Long-grain rice	0.85%	10.27	€10.79
	Muesli / granola	0.77%	14.39	€34.08
	Nut & chocolate spreads	2.33%	25.76	€39.99
	Oat flakes / cereal flakes	1.67%	26.78	€24.79
	Pasta (e.g., glass noodles)	1.05%	28.16	€61.37
	Processed rice	1.68%	44.27	€63.62
	Sweets			
	Alcohol chocolates	0.74%	16.33	€32.91
	Butter biscuits	0.43%	6.58	–€3.29
	Chocolate bars (incl. candy bars)	1.48%	16.93	€29.55

Continued on next page

Table WA-A.1 – continued from previous page

Group	**Category**	**Incidence ATE**	**Quantity ATE**	**Revenue ATE**
	Chocolate biscuits/pastries	1.96%	27.03	€30.84
	Chocolate products	2.61%	37.90	€51.97
	Filled chocolate tablets	2.01%	34.39	€44.12
	Sandwich cookies	–0.02%	–1.14	€0.09
	Single chocolates (assorted)	1.04%	16.14	€45.14
	Solid chocolate tablets	1.71%	36.39	€25.37
	Sugar confectionery	1.32%	34.81	€20.40
	Tea Coffee			
	Ground roasted coffee	–0.22%	–2.07	–€33.40
Near Food	**Body Care**			
	Deodorants (aerosol etc.) [F]	1.29%	18.91	€23.69
	Facial tissues [F]	–0.55%	–11.05	–€18.59
	Hair sprays & hair lacquer [F]	0.80%	8.27	€16.31
	Kitchen paper towels [F]	1.42%	11.27	€13.75
	Shampoo / hair wash products [F]	0.89%	10.45	€14.47
	Shower gel / body wash [F]	1.50%	19.15	€25.51
	Toilet paper [F]	1.96%	23.63	€66.03
	Toothpaste [F]	0.67%	9.26	€14.40
	Home Care			
	Dishwasher detergent (machine) [F]	0.48%	4.44	€16.25
	Laundry Care			
	Fabric softener / laundry rinse [F]	2.11%	27.89	€44.30
	Liquid laundry detergent [F]	0.08%	–0.28	€1.49
Pet Care	**Pet Food**			
	Dog snacks & treats	–0.01%	1.88	€1.42
	Wet cat food [F]	1.08%	61.73	€54.15
	Wet dog food [F]	0.29%	12.25	€11.17

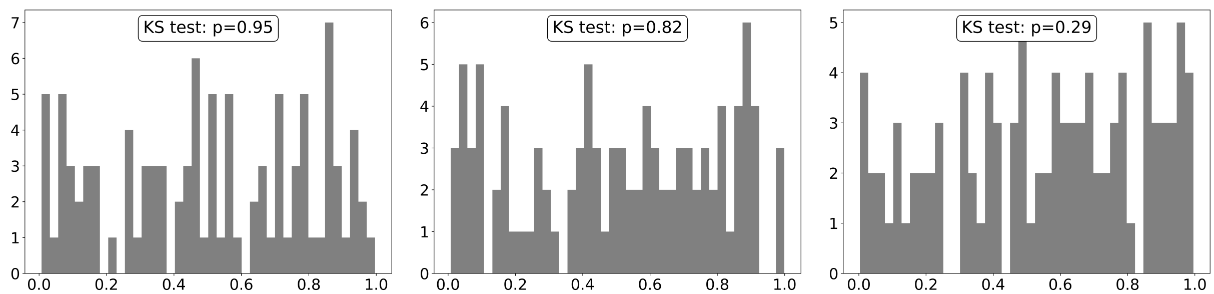
Notes: The table reports the estimated immediate treatment for all product categories in our study. The index [P] denotes perishable product categories, and the index [F] refers to product categories with fixed (inelastic) consumption. The *Quantity* and *Revenue* ATEs are reported per 1,000 coupons (per mille).

Web Appendix B Randomization Checks

The randomization in our experiment is implemented programmatically in the system and was separately verified with the engineering team. Reassuringly, campaigns in our sample pass the randomization tests. We used an F-test with three pretreatment variables to test randomization. The pretreatment variables for include the total spending, recency, and average basket size in 90 days prior to obtaining a coupon (specific to the customer×date observation). These variables are different from the ones used to detect eligibility criteria in Appendix A. We report the distribution of *p*-values across the campaigns for *Focal vs. NonFocal*, *Focal vs. NoRandom*, and *NonFocal vs. NoRandom* in Figure WA-B.1. The distributions are not statistically different from the uniform distribution using the Kolmogorov-Smirnov test. For completeness, we also implemented a randomization check on the discount levels (recall that the experiment randomized both brand assignments and discount levels). The randomization check for the discounts revealed no concerns with the implementation of the randomization.

Figure WA-B.1: Distribution of *p*-values in the Randomization Tests

(A) *Focal vs. NonFocal* **(B)** *Focal vs. NoRandom* **(C)** *NonFocal vs. NoRandom*



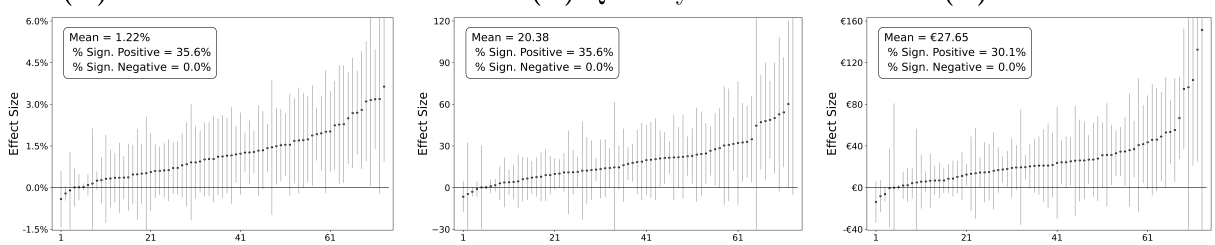
The figure reports histograms of the distribution of *p*-values in randomization tests. The sample size in each histogram is 101 (campaigns). The y-axis is a count of the number of brands, and the x-axis is *p*-values in the F-test with three pretreatment variables.

Web Appendix C Robustness to SUTVA and Competitive Coupons

We replicate our main findings using a reduced data sample that focuses on the first observation per customer (for each campaign), and only includes campaigns which had no competitive coupons during the observation period. The reduced sample rules out potential SUTVA violations and competitive spillover effects. The analysis in this section uses substantially fewer samples: 26,478 random coupons across 73 campaigns, compared to 88,840 random coupons across 101 campaigns in the main text. However, the pattern of results is remarkably consistent with the main text. (The plots in this Appendix follow the key findings summarized in Table 4.)

Figure WA-C.1: Distribution of Average Treatment Effects (ATEs)

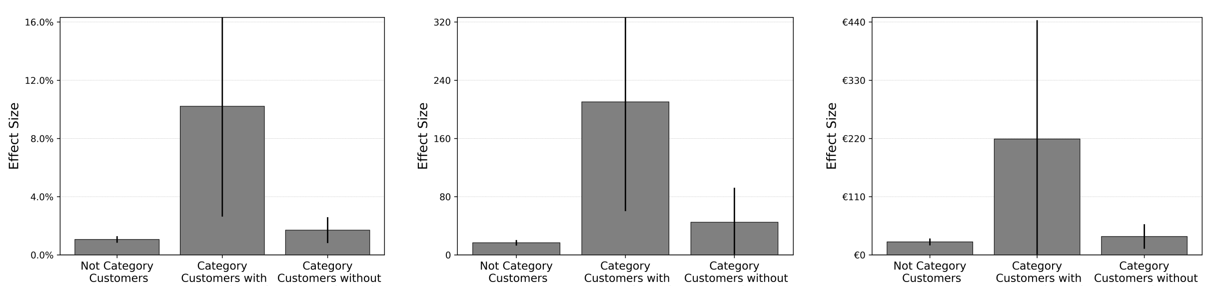
(A) Purchase Incidence **(B)** Quantity **(C)** Revenue



Notes: The figure summarizes the distribution of average treatment effects for the 73 brands that faced no competitive coupons during the experiment. Panel A reports the incremental purchase incidence for each brand (in percentage points). Panel B reports the incremental quantity sold (in units per 1,000 coupons). Panel C reports the incremental revenue (in € per 1,000 coupons). Error bars indicate 95% confidence intervals.

Figure WA-C.2: Heterogeneity in Treatment Effects for the Promoted Brand

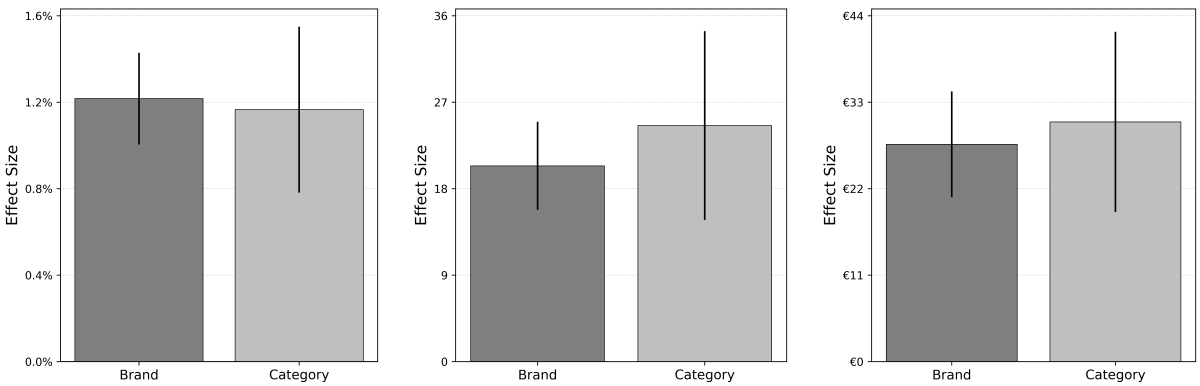
(A) Purchase Incidence **(B)** Quantity **(C)** Revenue



Notes: The figures report the average ATEs (on the current trip) averaged across the 73 focal campaigns that faced no competitive coupons. Each column represents a different customer segment. Error bars indicate 95% confidence intervals.

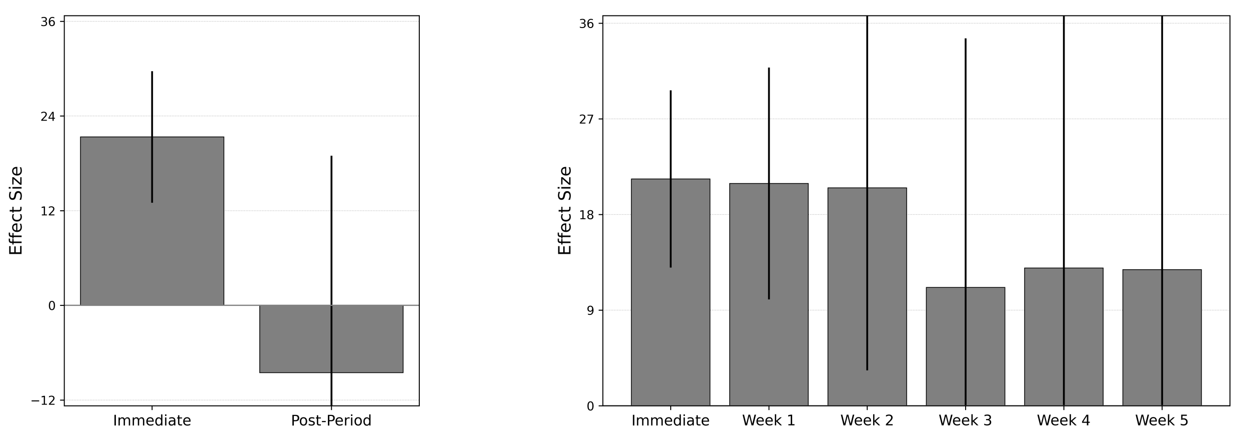
Figure WA-C.3: Compare Brand and Category ATEs

(A) Purchase Incidence **(B)** Quantity **(C)** Revenue



Notes: The figures report the average Brand and Category ATEs (on the current trip) averaged across the 73 focal campaigns that faced no competitive coupons. Error bars indicate 95% confidence intervals.

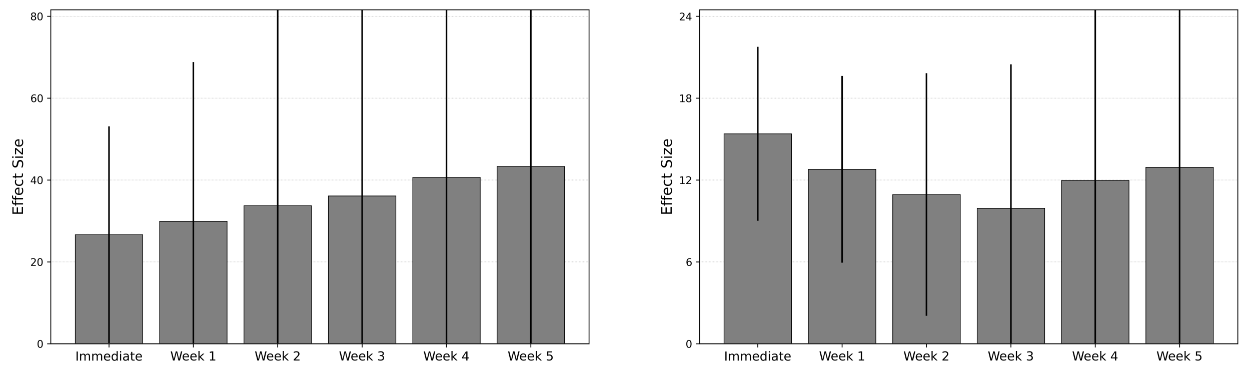
Figure WA-C.4: Temporal Substitution: Immediate and Post-Period Category-Level ATEs



(A) Distinct Periods **(B)** Cumulative Periods

Notes: These figures plot category-level *Quantity* ATEs averaged across the 61 categories where only one brand distributed coupons (no competitive coupons). In Panel (A), we report separate findings for the Immediate and Post-Period. In Panel (B), we report cumulative findings. Error bars indicate 95% confidence intervals.

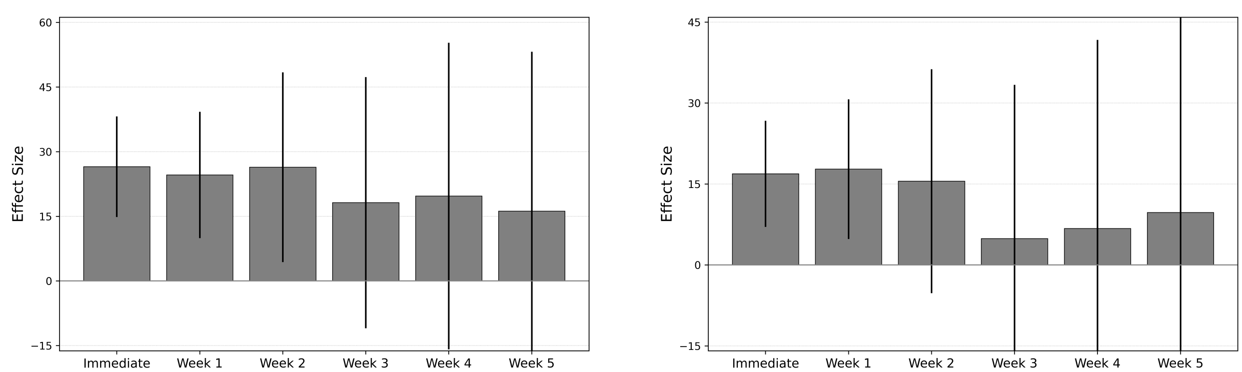
Figure WA-C.5: Categories in Which We Do Not Anticipate Changes in Consumption



(A) All Customers **(B)** Customers Without a Category Purchase

Notes: The figures report cumulative category-level *Quantity* ATEs averaged across categories in which we do not expect coupons to cause changes in consumption. Panel (A) includes all customers, Panel (B) includes customers without a prior category purchase at the focal retailer in the 90 days before customers received the in-store coupon. Error bars indicate 95% confidence intervals.

Figure WA-C.6: Competing Stores Close or Far From Focal Retailer's Stores



(A) Competing Stores Close **(B)** Competing Stores Far

Notes: The figures report category-level *Quantity* ATEs averaged across the 61 categories where only one brand distributed coupons (no competitive coupons). In Panel (A) we calculate ATEs using purchases at the focal retailer's stores located close to competing stores. In Panel (B) we calculate ATEs using purchases at stores located far from competing stores. Error bars indicate 95% confidence intervals.

Web Appendix D Summary Statistics for Brand Features

This appendix provides descriptive statistics and correlation patterns for the brand-level variables used in the main text. Table WA-D.1 reports summary statistics for four brand features that have appeared prominently in the price promotions literature (Narasimhan et al. 1996; Bell et al. 1999; Ailawadi and Neslin 1998):

Brand Penetration: Percentage of customers that purchased at least one item from the brand in the data period.

Brand Loyalty: Spending on the brand as a % of total spending in the category. Calculated first at the customer level (among customers of the focal brand), and then averaged across customers.

Price Position: Average price paid for the brand, divided by average price paid in the category.

Interpurchase Time: Average (across customers) in the number of days between each customer's sequential purchases of the brand.

These features are all constructed using the retailer's transaction data in the three months prior to the in-store coupon experiments.

Table WA-D.1: Brand Characteristics

Variable	**Average**	**Std. Dev**	**$P_{25\%}$**	**$P_{75\%}$**
Brand Penetration	1.27%	1.24%	0.49%	1.47%
Brand Loyalty	0.76	0.14	0.66	0.87
Price Position	1.17	0.64	0.84	1.38
Interpurchase Time	18.53	5.32	15.43	21.13

Notes: *Brand Penetration*, *Brand Loyalty*, *Price Position*, and *Interpurchase Time* are based on the retailer's loyalty card data in the three months prior to the in-store coupon experiments.

In Table WA-D.2, we report the pairwise correlation between the three treatment effects (incidence, quantity, and revenue) and four brand features. *Brand Penetration* explains more variation in in-store coupon treatment effects than brand loyalty, price position, or interpurchase time. Brands that are purchased by more customers tend to have larger ATEs. One explanation for the effect is that some categories are not relevant for certain customers, and coupons do not change this relevance. If few customers have pets, then we would expect both a low penetration rate for pet food brands, and a relatively small response to pet food coupons.

Quantity treatment effects are larger for brands with less loyal customers. A possible interpretation is that when *Brand Loyalty* is low, customers in the focal category are more price-sensitive or have lower switching costs, so they frequently switch brands during a promotion. For these customers, the coupon discount is sufficient to increase preferences for the brand and prompt a purchase.

Table WA-D.2: Pairwise Correlation Between Treatment Effects and Brand Characteristics

	Incidence ATE	**Quantity ATE**	**Revenue ATE**
Brand Penetration	0.41 **	0.46 **	0.20 **
Brand Loyalty	–0.13	–0.28 **	0.12
Price Position	–0.13	–0.18 †	0.04
Interpurchase Time	0.07	–0.09	0.12

The table reports the Spearman pairwise correlations between each brand feature and purchase incidence, purchase quantity, and revenue treatment effects (on the current trip). The unit of analysis is a brand, and the sample size is 101. Standard errors are in parentheses. ** indicates significantly different from zero, $p < 0.01$. * indicates significantly different from zero, $p < 0.05$. † indicates significantly different from zero, $p < 0.10$.

Web Appendix E Discount CATEs with Stratified Brand Sample

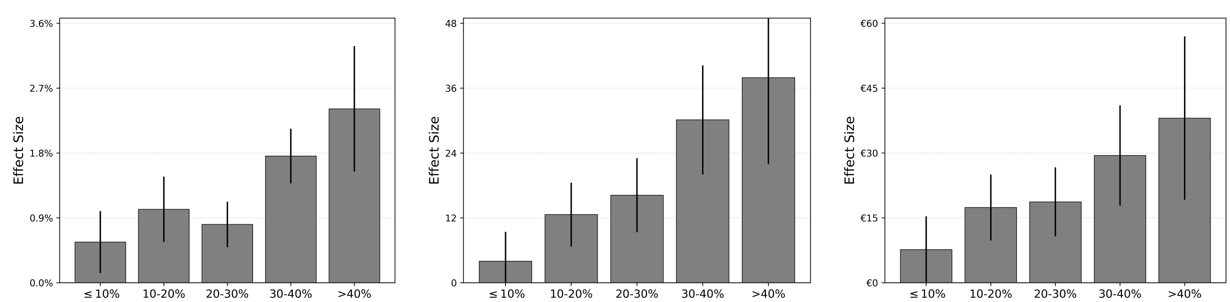
In Section 5.3, we study how coupon treatment effects vary with discount depth by binning coupons into five 10-percentage-point discount buckets and estimating treatment effects by brand and discount bin, then averaging the brand-level treatment effects within each bucket. Because manufacturers do not necessarily offer the full range of discount levels for every brand, some brands have no observations in the lowest and/or highest discount bins. This implies that the set of brands contributing to the averages in Figure 5 can differ across discount bins.

To ensure that comparisons across discount levels are not influenced by changes in the composition of brands across bins, we replicate the discount-depth analysis on a balanced set of 24 brands that have observations in all five discount buckets. We follow the same identification strategy as in the main analysis. Treatment effects are first computed at the brand-by-discount-bin level and then averaged across brands within each discount bin. Figure WA-E.1 reports these averages, with error bars reflecting 95% confidence intervals.

The results in Figure WA-E.1 closely mirror the results in the main text. Purchase incidence and quantity treatment effects are positive across all discount bins and generally increase with discount depth. For revenue, the balanced-sample estimates remain positive and increase with discount depth. In contrast to Figure 5C, where the revenue effect dips in the last discount bin, the results for the stratified brand sample do not show a clear drop for discounts $\geq$40%. Overall, restricting the analysis to brands with support in all discount bins leaves the qualitative conclusion unchanged: deeper discounts generate larger treatment effects, and the main heterogeneity patterns are not driven by shifting brand composition across discount levels.

Figure WA-E.1: Treatment Effects at Each Discount Level

(A) Purchase Incidence **(B)** Quantity **(C)** Revenue



Notes: The figure depicts treatment effects as a function of the discount depth. Treatment effects within each discount bucket are first calculated at the brand level. The black dots represent means calculated by averaging the treatment effects across the brands. Error bars indicate 95% confidence intervals calculated using a non-parametric bootstrap.

Web Appendix F Category-level Survey Scores

To investigate heterogeneity in category expansion, we use category-level measures capturing two behavioral mechanisms emphasized in the past literature: impulse purchasing and stockpiling. These measures are not derived from the retailer's sales data but are instead based on a consumer survey the retailer conducted in July 2017. The survey goal was to elicit category attributes reflecting well-documented differences in how consumers view and approach different product categories.

The stockpiling scores are based on the scale items developed by Narasimhan et al. (1996). Respondents rate each category on a 7-point Likert scale (1 = strongly disagree; 7 = strongly agree). The scale includes two items capturing perceived ease of storing extra quantities at home and the respondent's propensity to stock up when possible:

Easy "It is easy to store extra quantities of this product in my home."

Like "I like to stock up on this product when I can."

We construct a category-level measure of *Stockpilability* by averaging the responses to both of these questions across respondents.

The impulse scores measures are obtained using the two-item Impulse scale from Narasimhan et al. (1996):

Whim "I often buy this product on a whim when I pass by it in the store."

Urge "I typically like to buy this product when the urge strikes me."

Both questions are again measured using a 7-point Likert agreement scale. Similar to the *Stockpilability* measure, we average these two measures across respondents to create a measure of how susceptible a category is to being purchased on impulse.

The retailer recruited a panel of German consumers on Amazon Mechanical Turk (MTurk). A total of 614 respondents participated in the survey, and each respondent evaluated eight product categories that were randomly selected from the full set of categories. Limiting the number of categories per respondent reduced respondent burden and mitigated potential fatigue or wearout effects. In addition to completing the category-rating scales, respondents were asked whether they had purchased from each category within the past six months. Both the order in which categories were presented and the order of the relevant question blocks were randomized to avoid order effects.

From the raw data, we construct category-level scores using (i) all available responses and (ii) the subset of responses from respondents who reported at least one category purchase in the previous six months. Table WA-F.1 reports descriptive statistics for the *Impulse* and *Stockpiling* scales. We note that the retailer conducted the survey using broader category definitions than those analyzed in our main text. Specifically, the retailer collected scores for 43 categories, whereas our study relies on 71. To evaluate the correlation between ATEs and these survey metrics, we map the aggregate scores to our more granular product categories.

Across both samples, the stockpiling items (Easy and Like) exhibit the highest mean levels and relatively low dispersion, indicating broad agreement that many categories are easy to store and suitable for stocking up. Impulse items show more moderate means and greater heterogeneity, with Urge consistently rated higher than Whim. The distributions are broadly similar across the two samples, with slightly lower impulse ratings in Panel (B).

Table WA-F.1: Category-Level Survey Scores

(A) Category Purchasers

Item	**Mean**	**SD**	**Percentiles**				
			10%	25%	50%	75%	90%
Stockpiling							
Easy	5.328	0.567	4.525	4.896	5.245	5.790	6.166
Like	4.562	0.777	3.552	3.797	4.554	5.242	5.489
Impulse							
Urge	4.197	1.053	2.521	3.472	4.542	4.902	5.420
Whim	3.798	0.842	2.513	3.313	3.780	4.480	4.921

(B) All Respondents

Item	**Mean**	**SD**	**Percentiles**				
			10%	25%	50%	75%	90%
Stockpiling							
Easy	5.161	0.620	4.406	4.649	5.141	5.659	5.997
Like	4.287	0.813	3.317	3.625	4.269	4.896	5.393
Impulse							
Urge	3.993	0.992	2.493	3.300	4.085	4.634	5.166
Whim	3.573	0.803	2.454	3.203	3.453	4.191	4.479

Web Appendix G Correlations: ATEs and Impulse Scores

In this appendix, we examine how category-level ATEs (Immediate and Post-Period) vary with category-level characteristics that are related to impulse buying: *Urge*, *Whim*, and *Impulse* (see Web Appendix F). This analysis provides some indication that categories with higher impulse scores exhibit systematically lower Post-Period treatment effects.

Table WA-G.1: ATE Correlations with Category Impulse Scores

(A) Category Purchasers

	Whim	**Urge**	**Impulse**
Immediate ATE	0.187	0.136	0.161
	(0.120)	(0.262)	(0.184)
Post-Period ATE	−0.180	−0.199	−0.192
	(0.137)	(0.099)	(0.112)

(B) All Respondents

	Whim	**Urge**	**Impulse**
Immediate ATE	0.184	0.151	0.167
	(0.127)	(0.213)	(0.167)
Post-Period ATE	−0.193	−0.196	−0.197
	(0.109)	(0.103)	(0.103)

Notes: The table reports the pairwise correlation between each ATE and each category-level characteristic. The sample size for each correlation is 71 (categories), and p-values are reported in parentheses.

Web Appendix H Panel Data Analysis

This appendix provides additional details on the difference-in-differences (DiD) analysis of household panel data that evaluates the impact of the in-store coupon system on retailer sales (see Section 6).

H.1 Approach

We obtained household panel data covering purchases from grocery retailers for a sample of 46,150 German households. Panel participants upload receipts for all their grocery purchases. The data span a period of two years—one year before and one year after the coupon system introduction. For each panel household h, we know in which city the household lives at the time of the coupon system introduction.

Our coupon experiments, as well as the implementation of the coupon system, focus on a single metropolitan area. In contrast, the household panel covers all of Germany. Because the retailer also operates stores elsewhere in Germany without the coupon system, we can use a DiD approach to estimate the causal effect of the coupon system on household spending. We define a treatment indicator $Treated_r$ that equals 1 for households residing in the treated region and 0 for households residing in the untreated region. Furthermore, we define $Post_m$ as an indicator equal to 1 for months after the introduction of the coupon system and 0 for months before. In the DiD analysis, we focus on 3,548 households with at least one purchase at the retailer before the coupon system introduction.

In each of the twelve months before and twelve months after the introduction of the coupon system, we separately aggregate spending at each retailer across households by month m and region r. We denote this aggregate spending as y_{rm}. We then estimate the following DiD specification separately for the focal retailer and for the main competitor:

$$\log y^c_{rm} = \tau^c \cdot (Treated_r \times Post_m) + \sigma_r + \sigma_m + \varepsilon_{rm} \tag{WA-H.1}$$

where c denotes either the focal retail chain or the competing chain, σ_r are region fixed effects, σ_m are month fixed effects, and ε_{rm} is a normally distributed error term.

The coefficient of interest is τ^c, which captures the change in spending for treated households at chain c after the introduction of the in-store coupon system, relative to the corresponding change for control households. We report DiD estimates as percent changes and construct 95% confidence intervals using the regression standard errors. In addition to overall spending, we repeat the analysis for the nine product groups shown in the main text. Table WA-H.1 summarizes the results.

H.2 Balance Check

A potential concern in the panel-data DiD design is that we did not randomly assign households to treatment and control. Consequently, households in the treated region could differ systematically from households in the untreated region. We implement a balance check by estimating a logit model

Table WA-H.1: Difference-in-Differences Estimates

	Focal Retailer			Main Competitor		
	$\hat{\tau}$	SE	p-value	$\hat{\tau}$	SE	p-value
All	8.07%	2.74%	0.007	-3.88%	2.10%	0.077
Body & personal care	9.90%	8.66%	0.265	−1.06%	8.73%	0.905
Bread	4.63%	5.62%	0.419	−2.61%	3.56%	0.471
Candy & snacks	17.32%	4.55%	0.001	−3.55%	3.42%	0.309
Dairy	2.63%	3.30%	0.433	−3.14%	2.81%	0.276
Drinks	−0.92%	3.46%	0.793	−7.85%	4.48%	0.093
Food staples	7.91%	2.92%	0.013	−5.71%	2.57%	0.036
Fresh produce	8.94%	3.94%	0.033	0.04%	3.61%	0.992
Home & laundry care	15.55%	8.09%	0.067	−5.28%	8.03%	0.517
Meat & fish	12.88%	4.55%	0.009	−5.94%	3.58%	0.110

Notes: The table reports coefficients and standard errors from estimating Equation WA-H.1. Each treatment effect $\hat{\tau}$ is estimated in a separate regression. The unit of observation is region×month, and the sample size in each model is 48.

where the dependent variable equals one if the household resides in the treated region (and zero otherwise). The explanatory variables are measured before the introduction of the coupon system and include both pre-period purchasing behavior and household demographics: the household's mean and standard deviation of spending (in euros), the Herfindahl index of households' spending concentration across retailers, and demographic indicators (age groups, household size groups, and social classes).

Table WA-H.2 reports the estimated coefficients. Consistent with random differences in observables, none of the pre-treatment covariates is statistically significant at conventional levels. Overall, this balance check suggests that treated vs. control status is not predicted by observable pre-period purchasing patterns or demographic characteristics. This supports the comparability of the treated and control samples.

H.3 Placebo Test Results

A central identifying assumption in our panel-data DiD design is that, absent the introduction of the in-store coupon system, outcomes for households in the treated region would have evolved in parallel to outcomes for households in the control region. A standard diagnostic is a *placebo timing test*: we assign a *fake* introduction date in a period when the coupon system was not yet operating and re-run the DiD. In this test, we should not find systematic treatment effects.

We implement the placebo by replicating the main DiD specification but shifting the post indicator to a pre-treatment month. Specifically, we keep the treatment definition unchanged, focus on purchases at the focal retailer, and aggregate purchases to the month×treatment-status level (overall and separately for each of the nine product groups used in the panel data). We then define a placebo post indicator that is 0 in the first six months and 1 in the last six months of the original

Table WA-H.2: Balance Check Results

Variable	Coefficient	Std. Err.	p-value	95% CI
Intercept	0.3612	0.222	0.104	[−0.074, 0.797]
Spending average	−0.0315	0.153	0.837	[−0.331, 0.268]
Spending variability	−0.1326	0.099	0.179	[−0.326, 0.061]
Spending concentration	0.0815	0.234	0.727	[−0.376, 0.539]
Age low	0.2414	0.263	0.360	[−0.275, 0.758]
Age high	0.0829	0.114	0.468	[−0.141, 0.307]
Large households	0.1164	0.320	0.716	[−0.510, 0.743]
Low social class	−0.0943	0.134	0.481	[−0.356, 0.168]
High social class	−0.0333	0.129	0.796	[−0.286, 0.220]

Notes: The table reports coefficients and standard errors from estimating a logit model where the dependent variable equals one if the household resides in the treated region (and zero otherwise). The unit of analysis is a household, and the sample size is 1,862.

time window (i.e., before the actual introduction of the couponing system). We then estimate the same DiD regression as in the main analysis.

The results in Table WA-H.3 show that the placebo estimates provide no evidence of an effect around the fake introduction date. The overall placebo DiD coefficient is small (0.67%) and statistically insignificant ($p = 0.871$). At the product-group level, placebo coefficients fluctuate in sign and magnitude but are uniformly imprecise and not statistically significant at conventional levels (all p-values exceed 0.10). The absence of systematic or statistically significant placebo effects supports the interpretation that the main DiD results are driven by the actual introduction of the coupon system rather than differential pre-trends or unrelated shocks in the treated region.

Table WA-H.3: Placebo Difference-in-Differences Estimates

	$\hat{\tau}$	**Std. Err.**	p**-value**
All	0.67%	3.99%	0.871
Body & personal care	−15.07%	8.72%	0.118
Bread	12.95%	8.32%	0.154
Candy & snacks	5.05%	6.56%	0.461
Dairy	−4.62%	4.64%	0.345
Drinks	3.34%	4.34%	0.461
Food staples	1.96%	4.84%	0.695
Fresh produce	3.06%	6.56%	0.653
Home & laundry care	0.14%	12.87%	0.991
Meat & fish	−3.26%	4.62%	0.497

Notes: The table reports coefficients and standard errors from estimating Equation WA-H.1 using a placebo treatment date. Each treatment effect $\hat{\tau}$ is estimated in a separate regression. The unit of observation is region×month, and the sample size is 24.