Consumer Responses to 'Favorite' Product Removal: Evidence from Beverage Vending Machines

March 27, 2026

Abstract

How do consumers react when their favorite product is removed from their regular shopping outlet? While conventional wisdom suggests a loss in sales or substitution towards similar/familiar products, by analyzing highly granular individual purchase data from Tokyo train stations vending machines, we uncover a more complex and surprising response: this removal intervention drives consumers to purchase categories that are new to them. Remarkably, this surge in exploration persists even when consumers switch to a different machine where their favorite product is still available. We argue that the removal may act as a disruption to low-involvement habits, forcing consumers out of "autopilot" and into high-involvement decision process. Once the habit is broken, consumers pay more attention to the full range of available options, leading them to explore and learn about products they previously ignored. The impact, however, is not uniform. We find a systematic divide along consumer variety-seeking tendency: high variety-seekers are more likely to choose a substitute at their favorite vending machine, whereas low variety-seekers tend to search for their favorite product at other locations.

Keywords: favorite product removal, purchase substitution, habit, autopilot behavior, experimentation, exploration, learning, variety-seeking, synthetic difference-in-differences

1 Introduction

The consumer packaged goods market is characterized by frequent introductions of new products (MacGregor, 2024). To make room for these new products, managers often need to decide which old products to discontinue. Although under-performed products are natural candidates for removal, discontinuation decisions are complicated by heterogeneous consumer preferences – even low-sales products may be the most frequently chosen product for certain consumers. When a favorite product is removed, rational consumers may simply switch to a close substitute. But highly loyal customers, who have emotional attachment to their favorite product, may react strongly by abandoning the brand or retailer altogether.[1] In contrast, the most frequently chosen product could also be an outcome of low-involvement habitual behavior (Celsi and Olson, 1988) – in this case we have little guidance from theories about how such consumers would react.

Understanding how consumers react to this situation will not only shed light on the mechanisms behind why consumers keep choosing the same product, but also help managers to make product removal decisions. However, empirical evidence remains scant regarding consumer responses to the removal of their "favorite" products in retail environments. Within the literature on product assortment changes, the studies most closely related to ours are Borle et al. (2005), Sloot et al. (2006), and Boatwright and Nunes (2001), which used field experiments to provide evidence that a large-scale assortment reduction can lead to short-term sales declines. Our paper is also related to Anupindi et al. (1998) and Conlon and Mortimer (2013), which studied how demand changes in response to temporary product stock-outs in vending machines. However, none of these papers investigate how individual consumers react when their favorite product is discontinued.

This research gap persists, at least in part, due to limited availability of individual-level

[1] One recent example is TaB, which had suffered low sales performance for many years before it was discontinued in 2020. Coca-Cola did not terminate it earlier because TaB drinkers were very loyal to this drink, and the company was concerned about their reactions. `https://www.nbcnews.com/think/opinion/coca-cola-s-tab-soda-has-been-discontinued-its-retirement-ncna1243950`, accessed on Oct 24, 2024.

consumer data required to analyze such discontinuation effects. For instance, Anupindi et al. (1998) and Conlon and Mortimer (2013) cannot study how individual consumers respond to the removal of their favorite product because they only observed product-level sales. Moreover, the commonly used A.C. Nielsen scanner panel data only provide information on purchase decisions at the household level, instead of the individual level.[2] Furthermore, consumer purchase decisions at grocery stores can be subject to stockpiling behavior, especially during sales (Erdem et al., 2003; Hendel and Nevo, 2006; Ching and Osborne, 2020). Such stockpiling behavior can delay consumer response to product removal because they can keep consuming the removed product from their own inventory for a while. In addition, other confounding factors, such as price fluctuations, advertising, and other promotions, can further complicate the identification problem.

In this research, we contribute to the literature by providing the first causal evidence examining how consumers respond to the removal of their most frequently purchased product in a retail environment. We overcome the aforementioned challenges by analyzing a unique dataset from Japanese Railway beverage vending machines. This setting offers several distinct advantages for investigating consumer responses to the discontinuation of their favorite product. First, the immediate consumption of purchased items and the flat pricing policy allow us to track consumption response more accurately and bypass the complication due to consumer stockpiling at grocery stores. Second, our detailed individual-level purchase history data allow us to identify one's favorite product. Third, this market has minimal advertising, and this avoids the usual confounding factors present at grocery stores. Fourth, each vending machine's product assortment is determined by overall demand and broader market factors, making product assortment changes exogenous to individual consumers, a key feature for establishing causality between product removal and consumer response.

Another challenge of establishing causal inference is finding a suitable control unit that satisfies the parallel trends assumption. To address this, we apply the synthetic difference-

[2]For households with only one member, the purchase decisions recorded in Nielsen represent individual-level data. However, such individuals may not be representative of the general population.

in-differences (synthetic DID) method (Arkhangelsky et al., 2021). For each treated unit, this approach constructs a synthetic control unit by taking a weighted average of potential control units, ensuring that the synthetic control follows the same pre-treatment trend as the treated unit. Our analysis focuses on a 14-day time window before and after a favorite product is removed. Although this time frame limits our inference to short-term consumer reaction, it helps minimize the influence of other events or policy changes that could occur over a longer period.

Our results show that while consumers reduce their purchases at their favorite vending machines by 16% (-0.052 units) after their favorite product has been removed, such reduction is offset by the increase in purchases at their non-favorite machines, resulting in their overall purchases at favorite stations remain unchanged. Interestingly, we also observe increased purchases in the same category as the favorite product's and other categories by 99% (+0.070 units) and 78% (+0.052 units), respectively, at their favorite machines. At non-favorite machines, we observe a 71% rise (+0.037 units) in demand for favorite products, a 74% increase (+0.034 units) for other products belonging to the same category of the favorite product, and a 44% gain (+0.028 units) for products in other categories.

It is conceivable that some consumers decide to search for their favorite product at their non-favorite machines. However, it is surprising that even when the favorite product is present at a non-favorite machine at which consumers ended up buying, we still observe a large increase in purchases of non-favorite products. This led us to investigate and discover that this intervention also results in increased consumer exploration and learning behavior: daily probability of purchasing of new products surged by 73% (+0.027 units), while the weekly product switching rate rose by 23% (+0.046 units). We posit that prior to the removal, consumers purchased their favorite product, at least to some extent, out of low involvement habit buying (i.e., an autopilot mode) (Celsi and Olson, 1988; Webb et al., 2024); the intervention disrupted this routine, forcing them to pay attention to other products and prompting the exploration of new options. This is further supported by the evidence that

consumers increase their purchase for new products at non-favorite machines where their favorite product is available.

The increase in consumer exploration/experimentation might have additional long-term benefits. By triggering consumers to try new products, the intervention could help them discover better matches, potentially enhancing overall satisfaction in the long run. To our knowledge, this is the first paper that points out this potential benefit of discontinuing some existing products.

Lastly, to gain additional insights, we examine how the treatment effect varies with consumers' variety-seeking tendencies. We find that high variety-seeking consumers are more likely to choose a substitute at their favorite vending machine, whereas low variety-seeking consumers tend to seek their favorite product at their non-favorite vending machines. Although exploration behavior increased uniformly for both groups, product switching rose significantly only for low variety-seekers.

The remainder of the paper is structured as follows. Section 2 describes the institutional background and our datasets. Section 3 lays out the empirical strategy and discusses the identification assumptions. We report the estimated average treatment effects in Section 4, followed by the heterogeneous treatment effects by variety-seeking tendency in Section 5. Section 7 concludes by discussing the practical implications of our findings.

2 Institutional Background and Data

2.1 Beverage Vending Machine Business

Our study is based on point-of-sale data in beverage vending machines located at railway stations in the eastern part of the main island of the Japanese archipelago, covering both urban and intercity transportation. These vending machines are managed by East Japan Railway, the largest railway company in Japan. The company owns all vending machines in its stations, and no other company is allowed to operate the beverage vending machine

business in these stations. In the Tokyo metropolitan area alone, there are approximately 17 million passengers per day, and 41% of them use the service of East Japan Railway.[3]

Three parties are involved in the operation of this vending machine business: a central company, third-party operators, and local workers. The central company, which is a subsidiary of East Japan Railway, sets prices and general product assortments for vending machines. The price of each product is common across vending machines and fixed over time. Thus, there is no temporary price promotion. Moreover, there is minimal advertising in this market.[4] The on-site management of vending machines, such as maintenance, restocking, and product assortment, is outsourced to third-party operators. These operators hire and assign local workers to the vending machines for on-site operation. Local workers' compensation partly depends on the sales of their assigned vending machines, and hence they have an incentive to raise their overall sales.

The central company divides a year into the Spring/Summer and Fall/Winter seasons. The Spring/Summer season is from April 1st to August 31st (because most Japanese schools and companies start the fiscal year in April), and the Fall/Winter season is from September 1st to March 31st. The product assortment policy is planned for each season. The product assortment for the Spring/Summer (Fall/Winter) season is determined by around the end of February and communicated to the third-party operators by March (August). Consequently, most product changes occur around April-June and September-December for the Fall/Winter and Spring/Summer seasons, respectively.

The central company's policy can be summarized as follows. They categorize the vending machines into several groups based on the traffic volume and location (e.g., a high weekday traffic machine near the office district). They then assign a list of products that must be placed in each category of vending machines, which usually takes up approximately 70% of

[3]JEKI Media data 2022 https://www.jeki.co.jp/transit/mediaguide/data/, accessed on April 22, 2022.

[4]The product recommendations examined in Kawaguchi et al. (2019, 2021) are specific to a distinct subset of vending machines, representing only 3% of the total machines. Consequently, their influence on the current analysis is minimal. Furthermore, these recommendations follow a consistent pattern across machines and remain unchanged throughout the sample period. As a result, any potential effect is controlled by synthetic control introduced in the subsequent section of this study.

the machine's slots. The local workers who are responsible for the vending machines are expected to comply with the policy and replace the products by the deadline, either at the end of June or at the end of November. However, it is worth noting that the local workers have some discretion about the exact timing of implementing the change, and which products to remove.

2.2 Data Description

We have detailed vending machine characteristics, including their location (which station and whereabouts at the station) and the number of slots. We also have the point-of-sale data at the consumer level. For each transaction, we observe the time of purchase, the product purchased, the vending machine, and the corresponding slot at which the purchase happened. If the payment is made by the electric commuter card of East Japan Railway, SUICA, the card ID is recorded. Typically, each commuter has their own SUICA card. The data is from January 1st, 2018 to August 31st, 2018, involving over 10 million unique consumer IDs, 6,434 vending machines, 896 stations, and 480 products. The average product price is 143.7 JPY, and it ranges from 100 to 210 JPY.

Table 1: Station and Vending Machine Characteristics

	Min	Median	Mean	Max	Sd
# machines in a station	1.0	2.0	5.2	173	9.1
# slots in a machine	12.0	36.0	35.5	42	4.3
# distinct products in a machine	10.5	34.2	33.8	42	4.1

Note: We summarize the number of distinct products per machine at the machine–day level to capture time variation.

Table 1 shows the summary statistics across dates at the station and vending machine levels. The average number of vending machines in a station is 5.2, with a minimum of 1 and a maximum of 173. The average number of slots in a vending machine is 35.5, and the average number of distinct products in a vending machine is 33.8, suggesting that a product

typically occupies one slot per machine. The large number of products in a machine also suggests that consumers might not pay attention to all of them as Tokyo train stations are very busy and commuters are usually in rush.

Table 2: Average Daily Sales Decomposition by Product Categories

	Sales (unit)	Share (%)
Vending machine level		
All products	53.73	100.00
Other tea	10.59	19.71
Coffee	9.44	17.58
Mineral water	6.53	12.14
Carbonated drink	6.06	11.28
Others	5.94	11.05
Energy drink	5.87	10.93
Green tea	4.26	7.93
Juice	3.94	7.33
Unknown	0.59	1.10
Station level		
All products	384.49	100.00
Other tea	76.07	19.78
Coffee	67.82	17.64
Mineral water	46.86	12.19
Carbonated drink	43.53	11.32
Others	42.63	11.09
Energy drink	42.17	10.97
Green tea	30.61	7.96
Juice	28.27	7.35
Unknown	4.24	1.10

Note: The average daily sales is first aggregated over each vending machine or station, then average across all vending machines or stations, and day from our entire sample period. "Unknown" category refers to products with missing product descriptions.

Among the 480 unique products, we have a detailed description of 283 products for the 2017 Fall/Winter and 2018 Spring/Summer seasons, including the price, category, date of release, and packaging information. The remaining 197 products have missing product descriptions, and we classify them as the "unknown" category. Note that the "unknown"

category market share is very small (1.1%), and none of the consumer's favorite products in our final sample belongs to it.[5] There are eight known product categories: carbonated drinks, coffee, energy drinks, green teas, other teas, juices, mineral waters, and others. Table 2 shows the average daily sales volume across different product categories and their sales at the level of vending machines and stations based on the eight-month data sample. The top four product categories at both levels are other teas, coffee, mineral water, and carbonated drinks.

2.3 Product Assortment Changes

We construct the product assortment of each vending machine from the point-of-sale data. A product is considered available from a slot of the vending machine on a given day if a purchase is made from that slot on the same day. However, about 31% of slot-machine-date combinations have no purchase. If there is no purchase observed on a given day, we use a forward-filling method to infer its availability, i.e., we assume the slot carries the most recently observed product before that day. Based on the constructed daily product assortment, we identify which product is removed and which is added as a replacement on each date for each vending machine.[6]

We find that during our data period, product assortments changed frequently, and involved a large number of regular and seasonal products. About 63% of unique products were replaced during our study period, affecting a substantial number of consumers with varying preferences. This high rate of product turnover provides a robust dataset to analyze the effects of product changes on consumer behavior. Figure 1 displays the daily proportion of vending machines that experienced changes in product assortment. The first peak of product

[5]Because the market share of the unknown category is very small, we will not consider it when we study cross-category substitution analysis. On average, the probability of choosing the unknown category products increases from 0.002 to 0.003 after the treatment.

[6]To check the robustness, we also construct the vending machine product assortment data by backward filling, with which we replicate our analyses. The results are very similar for these two methods. Hence, without loss of generality, we only report the results based on the forward-filling method. The construction of other variables specific to each analysis is discussed in the corresponding sections later.

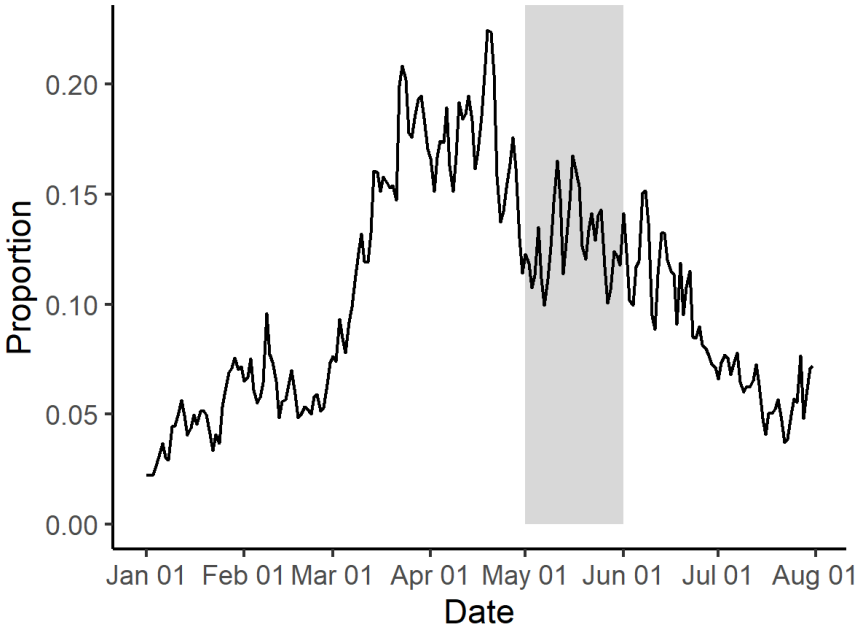


Figure 1: Daily Proportion of Machines with Assortment Changes

Note: The figure shows the daily proportion of vending machines experiencing product assortment changes during the entire sample period. The shaded area highlights the period that we focus on.

changes occurs from late March to late April, reaching a maximum of 25%. It is followed by a second peak (around 15%) in May, as indicated by a grey shaded area. From March to April, Tokyo experiences a notable shift in average temperature, significantly influencing beverage demand. Additionally, demand increases during this period due to the start of the fiscal and school years in April, introducing an influx of new consumers. Consequently, product changes in this period, including the removal of favorite products, are primarily seasonal transitions from hot to cold beverages, which customers generally anticipate; this may violate the "no anticipation" assumption required for difference-in-differences analysis. This issue is mitigated in May when most products are cold beverages for the Spring and Summer seasons. Therefore, we decide to use the product removal events happened in May to study how this intervention affects consumer choice.

To shed light on the sales difference between removed and retained products, we compare their average weekly sales per vending machine in April, which is prior to the treatments

Table 3: Sales Comparison between Removed and Retained Products Prior to Treatment

Products	Mean	SD	Median	Min	Max
Removed	12.02	12.94	8	0	226
Retained	16.00	19.39	10	0	867

Note: The table reports average weekly unit sales per machine in April (pre-treatment) for products that were removed in May and for products that were retained in May.

that happened in May. Table 3 displays the results. The average weekly sales of the removed and retained products are 12 and 16 units, respectively. The median sales of the removed and retained products are 8 and 10 units, respectively. Although the removed products are less popular in general, their demand are still comparable to the retained products. Hence, it is not surprising that they could be someone's favorite product.

2.4 Regular Consumer Sample

We define a consumer's favorite product as the product purchased most frequently from January to March, before the seasonal product assortment changes happened in April. We define a consumer's favorite vending machine and favorite station in the same manner. It follows that a consumer's favorite vending machine is always located in their favorite station. Henceforth, we will omit mentioning "the favorite station" when discussing purchases at one's favorite vending machine. It is important to note that we use the term "favorite product" to capture one's most frequently chosen product. The mechanism behind this "favorite" choice could reflect either preferences, habit, or a mixture of these two. We will explore their empirical implications in the econometric exercises.

To properly identify a consumer's favorite choice, we focus on a set of consumers who purchased regularly. We define *regular consumers* as consumers whose average inter-purchase time interval is less than 2 days and the average monthly number of purchases is larger than 15. This selection rule gives us a sample of 8,654 regular consumers. Table 4 reports their

Table 4: Summary Statistics of Regular Consumers' Daily Purchases at Vending Machines

	Mean	Sd	Median	Min	Max
Buying probability (morning)	0.54	0.08	0.53	0.2	0.98
Buying probability (other times)	0.11	0.12	0.06	0.0	0.79
Demand at favorite machine per consumer-day	0.47	0.54	0.00	0.0	13.00
Demand at favorite station per consumer-day	0.59	0.58	1.00	0.0	13.00
# products per purchase	1.03	0.18	1.00	1.0	8.00

Note: Based on January to March data. Demand is measured in units.

summary statistics of regular consumer daily purchases at the vending machine level from January to March. Most of the purchases happened in the morning, with an average daily buying probability of 0.54 (versus 0.11 at other times and 0.35 not buying at all). In fact, these consumers usually commuted to the center of Tokyo by train, and they purchased beverages on their way during the morning rush hour. The average individual daily demand at the favorite vending machine and favorite station are 0.47 and 0.59 units, respectively. Conditioning on buying, the average number of units per purchase is 1.03, and the 97th percentile is close to 1. This strongly supports our claim that there is almost no stockpiling behavior for regular consumers in our data.

2.5 Treatments Under Study

To achieve causal inference, we focus on *single treatment* events for each regular consumer. We have three selection rules. First, we only consider the *first* removal of the favorite product from each consumer's favorite vending machine. When there are multiple interventions in a unit, the exogeneity of the subsequent interventions is harder to justify than that of the first intervention because they can be an endogenous response to the outcome of the first intervention. By defining the first intervention as the treatment, we can avoid this complication.[7] Second, we only include events where exactly one product is removed – that

[7]In this case, the estimand is interpreted as the effect of intervening in the unit at least once, including the effects of potentially multiple interventions. Nevertheless, we expect that the difference between this

is, we exclude observations where multiple products are replaced on the same day at the same machine. Third, we only study consumers who were first treated in May to avoid confounding seasonal effects explained earlier. These selection rules lead to a total of 157 treated consumers. We find that 15 of the 157 treated consumers made no purchases in the 14 days prior to treatment. We exclude these consumers because, without recent purchase activity, it is difficult to argue that they would notice the removal of their favorite product. This leaves 142 treated consumers in the final sample. Our control consumers are those who never had their favorite products removed from their favorite vending machines from January to August (i.e., the never-treated consumers). We end up with 4,218 control consumers.

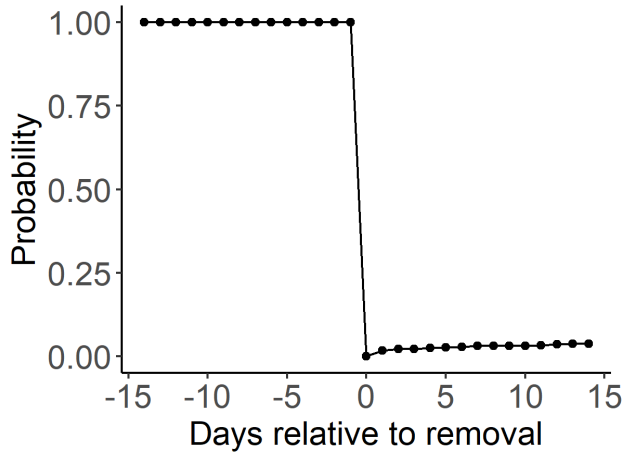


Figure 2: Non-reversibility of the Treatment

Note: Y-axis is the probability of finding the favorite product in the favorite vending machine for treated consumers.

In our estimation, we focus on a short window of 14 days before and after each treatment event. The high-frequency nature of our purchase data allows us to use this locally randomized variation of treatment timing. Moreover, as shown in Figure 2, treated consumers' favorite product is hardly placed back in our 14-day window after it is removed from the favorite vending machine ($< 4\%$). We include this small group of consumers who experience the reintroduction of their favorite product because it is a post-treatment endogenous

estimand and the effect of the first intervention, excluding the effects of the subsequent interventions, is negligible because there are very few consumers who are intervened more than once.

variable. In a robustness check, we exclude these consumers and our main findings remain unchanged.

When a product is removed from a machine, its slot will be filled by another product (around 92% of the replacement products are different from the existing products in the machine prior to the treatment). A potential concern is that our estimates may capture the combined effect of both the removal of a consumer's favorite product and the introduction of a replacement product. To investigate this concern, we examine to what extent treated consumers purchase the replacement product following the treatment. We find that 147 out of 157 treated consumers (93.6%) did not purchase the replacement during the 14-day window.[8] Given that the vast majority of treated consumers did not adopt the replacement, our results should be primarily driven by the removal of their favorite products.[9]

We also compare the characteristics of the treated and control consumers from January to March, prior to our synthetic DID analysis window. Table 5 shows that their purchase share of favorite products and favorite machines, as well as the number of purchases at their favorite machine, are all very similar. Note that the mean purchase share of favorite product is less than 0.80 for both the treated and control consumers. This indicates that consumers occasionally choose products other than their favorite product. This is an aspect that we will explore further when we consider consumer variety-seeking tendency.

[8]This fact is very important when we discuss the increase in exploration behavior. One might conjecture that low-involvement decision-making consumers would simply take the replacement product placed in the same slot of the favorite product's and that can explain the increase in "exploration." But this fact rule this out. It's likely that the consumers pay attention to slots other than what the favorite product used to occupy.

[9]Among the ten treated consumers who adopted the replacement product, purchase shares vary significantly: seven consumers allocated 3–40% of their purchases to the replacement, one consumer allocated 90%, and two consumers fully substituted (100%).

Table 5: Summary Statistics of Treated and Control Consumers

	Treated		Control		Treated - Control	
	Mean	Sd	Mean	Sd	Mean	t-stats.
Purchase share of favorite product	0.780	0.207	0.750	0.238	0.030	1.801
Purchase share at favorite machine	0.781	0.188	0.765	0.208	0.016	1.069
# total purchases at favorite machine	42.184	11.829	42.178	14.734	0.006	0.005

Note: Based on January to March purchases. Control consumers refers to consumers who have never been treated in our observation periods. Total purchases are measured in units.

3 Empirical Strategies

3.1 Synthetic DID Model

Let $Y_{i,t}$ be consumer i's purchases on date t. Each consumer only experiences at most one treatment in our sample, but possibly at different dates. We let g_i denote the date that treatment happens for consumer i. Then, consumers are characterized by the date when they receive the treatment. We refer to a group of consumers with $g_i = g$ as treatment cohorts g ($g = 1$ corresponds to consumers treated on May 1, and $g = 31$ corresponds to consumers treated on May 31 in our application). If a consumer is never treated in our sample period, we set $g_i = \infty$. In other words, all control consumers belong to cohort ∞.

Treatment effects may vary across treatment cohorts. If one ignores this issue and estimates a model assuming the treatment effect is homogeneous, it could lead to a biased estimate in the average effect.[10] Therefore, we allow the treatment effect to be heterogeneous across treatment cohorts. That is, instead of modeling a uniform treatment effect across consumers treated at different dates, we model the heterogeneous treatment effect, τ^g, for each treatment cohort g, and then define the average treatment effect τ^{ATE} by taking

[10]Goodman-Bacon (2021) suggests that one should estimate a model with heterogeneous treatment effects and aggregate the estimates to obtain the average treatment effect, rather than estimating the model with a homogeneous treatment effect, even if the average treatment effect is the primary interest.

the weighted average as follows.[11]

$$\tau^{ATE} = \sum_{g=1}^{31} \tau^g \cdot Prob(\text{a consumer is treated in date } g). \tag{1}$$

The purchase of consumer i on date t is modeled as,

$$Y_{i,t} = \sum_{g=1}^{31} \tau^g \cdot \mathbf{1}\{g_i = g\} \cdot \mathbf{1}\{t \geq g_i\} + \xi_{i,t} + \epsilon_{i,t}, \tag{2}$$

where τ^g is the treatment effect for cohort g, $\xi_{i,t}$ is an unobserved fixed effect, and $\epsilon_{i,t}$ is an idiosyncratic shock. The model includes the common trend assumption as a special case when $\xi_{i,t} = \mu_i + \lambda_t$. We focus on the window $t = \{g - T_0, g - T_0 + 1, \ldots, g, \ldots, g + T_0 - 1\}$ with $T_0 = 14$. In our analysis, our first treatment cohort is May 1, and the last treatment cohort is May 31, so our calendar date spans from April 17 to June 13.

3.2 Estimation

We apply the synthetic difference-in-differences (DID) method (Arkhangelsky et al., 2021), which is robust to certain types of violations of the common trend assumption. The main idea of the synthetic DID is to construct a synthetic control that achieves the common pre-trend for the treated unit and use it to construct the counterfactual outcome of the treated unit in the post-treatment period. Intuitively speaking, compared to the DID estimator, the synthetic DID requires more time-series observations to construct the synthetic control but fewer cross-sectional observations, in order to achieve a consistent estimate. Because we have both a large number of cross-sectional and time-series observations, this requirement

[11] A full heterogeneous model would assume that the treatment effects are heterogeneous over both the treatment cohort and the elapsed time from treatment. However, we find that estimating the specification with full heterogeneity using our proposed estimation method (synthetic DID, described in the later section) is prohibitively time-consuming. Therefore, we assume that the treatment effects are homogeneous over post-treatment days. As a robustness check, we estimated the fully heterogeneous model with a computationally efficient multi-period DID method (Callaway and Sant'Anna, 2021), but the results are similar because we focus on a narrow 28-day window around the treatment. The multi-period DID results are provided on request.

should be satisfied here.

We now describe the synthetic DID method in more detail. We first construct the estimator of the average treatment effect on consumers of cohort g (i.e., τ^g) and then aggregate them to construct the average treatment effect τ^{ATE}. Let $\mathcal{S}^g$ and $\mathcal{S}^\infty$ denote the set of consumers belonging to the treatment cohort g and the control group, respectively. We assign weights $\hat{\omega}_i^g$ for consumer units $i \in \mathcal{S}^\infty$, and $\hat{\lambda}_t^g$ to pre-treatment dates $t = g-14, \cdots, g-1$ to construct the weighted within-difference in means before and after date g for $i \in \mathcal{S}^g$ by

$$\hat{\tau}^g = \frac{1}{N_1^g} \sum_{i \in \mathcal{S}^g} \hat{\delta}_i^g - \sum_{i \in \mathcal{S}^\infty} \hat{\omega}_i^g \hat{\delta}_i^g, \tag{3}$$

$$\hat{\delta}_i^g = \frac{1}{14} \sum_{t=g}^{g+13} Y_{i,t} - \sum_{t=g-14}^{g-1} \hat{\lambda}_t^g Y_{i,t}, \tag{4}$$

where N_1^g is the number of consumers in treatment cohort $\mathcal{S}^g$.

We obtain the weights by solving the constrained minimization problem as follows,

$$\begin{aligned} \min_{\omega_0^g, \omega_i^g} \quad & \sum_{t=g-14}^{g-1} \left(\omega_0^g + \sum_{i \in \mathcal{S}^\infty} \omega_i^g Y_{i,t} - \frac{1}{N_1^g} \sum_{i \in \mathcal{S}^g} Y_{i,t} \right)^2 + \zeta^2 \cdot 14 \sum_{i \in \mathcal{S}^\infty} (\omega_i^g)^2 \\ s.t. \quad & \omega_i^g \geq 0, \ \sum_{i \in \mathcal{S}^\infty} \omega_i^g = 1, \end{aligned} \tag{5}$$

and

$$\begin{aligned} \min_{\lambda_0^g, \lambda_t^g} \quad & \sum_{i \in \mathcal{S}^\infty} \left(\lambda_0^g + \sum_{t=g-14}^{g-1} \lambda_t^g Y_{i,t} - \frac{1}{14} \sum_{t=g}^{g+13} Y_{i,t} \right)^2 \\ s.t. \quad & \lambda_t^g \geq 0, \ \sum_{t=g-14}^{g-1} \lambda_t^g = 1, \end{aligned} \tag{6}$$

where ζ is a regulation parameter to prevent the norm of the weight ω from being too large. The solution to this problem gives the weights for the control unit and the weights for pre-treatment periods by which the resulting synthetic control has an approximately common pre-trend with the treatment units.

We estimate the average treatment effect by

$$\hat{\tau}^{ATE} = \sum_{g=1}^{31} \hat{\tau}^g \frac{N_1^g}{\sum_{g=1}^{31} N_1^g}, \tag{7}$$

where $\hat{\tau}^g$ is the estimated treatment effect for treatment cohort g, and N_1^g is the number of consumers in treatment cohort g. We vary the time window and show the results in the robustness check section. We use bootstrapping to obtain the standard error of $\hat{\tau}^{ATE}$, relying on random resampling of the original data set with replacement 100 times.

3.3 Identification Assumptions

Synthetic DID assumes that the combination of double differencing, unit weights, and time weights can remove the bias due to the unobserved fixed effect $\xi_{i,t}$. If the common trend assumption ($\xi_{i,t} = \mu_i + \lambda_t$) holds, then the double differencing with equal unit and time weights is sufficient for the bias removal. Even if the common trend assumption is violated, as long as the weights of the units and times create a synthetic control that satisfies the common trend assumption, the synthetic DID estimator is still consistent. In our study, we indeed find such weights that lead to the common trend, as our control consumers share many similarities with the treated consumers, and we only require the common trend assumption to hold for a short window around each treatment date.

One key identification assumption for both DID and synthetic DID is no anticipation of the treatment by consumers. In our context, although local workers may remove certain products from vending machines based on factors like past aggregate sales, the decision is never based on individual consumers' preferences. Therefore, we believe it is unlikely for consumers to anticipate that their favorite product will be removed from their favorite vending machine on a particular date.

We should also comment on the possible spillover effect across vending machines. That is, the treatment of a vending machine could affect the consumer demand at other vending ma-

chines in the same station. More specifically, when a consumer's favorite product is removed from a vending machine, they may walk to another vending machine at the same station to make a purchase, and this might affect other consumers' buying behavior if it increases wait time or temporary stock-out. We believe that such problems are likely very minor, because the percentage of consumers who are treated is very small (only 2%). Nevertheless, we are being conservative, and our potential control units exclude all never-treated consumers whose favorite stations are the same as any treated consumer's favorite station. In a robustness check, we use all never-treated consumers (including those never-treated consumers from the same station) as the control and find the results remain largely unchanged.

4 Average Treatment Effects

4.1 Main Results

Table 6 presents the average treatment effects of removing a consumer's favorite product (Row 1), alongside the pre-treatment means of the dependent variables (Row 2) as baselines for comparison. We first discuss how the treatment affects one consumer's daily demand at various locations (favorite station, favorite machine, and non-favorite machines). For each location, daily demand is defined as the number of purchases made by the consumer on that day, with non-purchase days coded as zero. The discussion of this subsection will focus on this table unless stated otherwise.

4.1.1 Within Favorite Vending Machine Substitution

The treatment leads to a significant reduction in individual demand at their favorite vending machine, decreasing by 0.052 units ($p < 0.01$, Column 2) per day. This decline represents 16% of the pre-treatment mean (0.331 units), and is clearly driven by the demand for the

favorite product dropping from 0.20 units [12] However, this drop is accompanied by notable substitution effects within the same machine. Specifically, demand for products of the same category as the removed favorite product increases by 0.070 units ($p < 0.001$, Column 3), a 99% rise relative to the pre-treatment mean (0.030 units). Moreover, demand for products in other categories rises by 0.052 units ($p < 0.001$, Column 4), equivalent to an 78% increase over the baseline. However, even though the demand for same and other categories products increases, it cannot compensate for the loss from the favorite product, leading to the overall demand decrease at the favorite machine level.

Interestingly, while individual demand at their favorite machine decreases, it remains unchanged at the station-level (Column 1); this suggests that consumers offset their reduced purchases at the favorite machine by increasing demand at non-favorite machines within the same station, rather than reducing overall consumption. This leads us to examine individual between-machine substitution patterns next.

Table 6: Average Treatment Effects of Favorite Product Removals on Individual Demand

	Station	Favorite machine			Non-favorite machines		
	All products	All products	Same category products	Other categories products	Favorite products	Same category products	Other categories products
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ATE	0.010	-0.052**	0.070***	0.052***	0.037**	0.034***	0.028*
	(0.016)	(0.017)	(0.010)	(0.010)	(0.011)	(0.006)	(0.011)
Mean before treatment	0.493	0.331	0.071	0.067	0.052	0.046	0.063

Note: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. "Same category products" refers to non-favorite products in the same category as the consumer's favorite product, and "Other categories products" refers to products in different categories from the consumer's favorite product. Here, we ignore the "unknown" category because its sales are so small that it does not have any material impact. Standard errors are obtained by bootstrapping and reported in the parentheses. The mean is calculated as the average of the dependent variables across 14 days prior to the treatment. Demand is measured in units.

[12] The pre-treatment mean of the demand for favorite product at the favorite machine is 0.20 = 0.331 (col 2) - 0.071 (col 3) - 0.067 (col 4).

4.1.2 Between Vending Machine Substitution

When one's favorite product is not available at their favorite machine, consumers may look for it at their non-favorite machines. But when they are at a non-favorite machine, which they may be unfamiliar with its product assortment and placement, what would consumers actually choose? To shed light on this question, we estimate the average treatment effects on the demand at non-favorite machines, distinguishing between: (i) favorite products, (ii) products of the same category as the favorite product (excluding the favorite product), and (iii) products from other categories. The results are reported in Columns 5-7.[13]

Our analysis reveals several patterns. First, there is a significant increase in demand for favorite products at non-favorite machines (0.037 units, $p < 0.01$), representing 71% of the pre-treatment level. To understand the significance of this finding, we should highlight that the favorite product also becomes more scarce among non-favorite machines after the treatment date – the percentage of non-favorite machines that carry a treated unit favorite product drops from 58% to 40%. Although it is less likely for a consumer to find their favorite product in a non-favorite machine, we still observe that more purchases of favorite products happened at non-favorite machines, consistent with the hypothesis that consumers make an effort to find their favorite product after the intervention.

The demand for same category products (excluding favorites) also rises significantly by 0.034 units ($p < 0.001$), or 74% of the pre-treatment mean. These results indicate that while treated consumers may initially seek their favorite products at other machines, some eventually purchase substitutes within the same product category.

Surprisingly, we also find a significant increase in demand for other categories products (0.028 units, $p < 0.05$), corresponding to 44% of the pre-treatment level.[14]

[13]For the demand of the favorite products at non-favorite vending machines, we do not conditional on the purchases on a non-favorite vending machine where the favorite product is available. By doing this, we are estimating the joint effect of the treatment by affecting availability of the favorite product at non-favorite vending machines and consumer demand of their favorite products at non-favorite vending machines conditional on favorite product is available.

[14]The combined ATE for same and other categories is 0.034+0.028 which exceeds the increase in demand at the favorite machine (-0.052) by roughly 0.010, consistent with the ATE on the demand at the favorite

The cross-category spillover effect at both favorite machine and non-favorite machines is surprising: Why would removal of one's favorite product lead to more purchases in other categories products? Is it possible that it triggers consumers to pay more attention to products that they typically ignore? We investigate this further in the next subsection.

4.1.3 Increased in Exploration

Due to limited mental resources, consumers often make low-involvement decisions (Celsi and Olson, 1988; Webb et al., 2024). Such decision mode is common in familiar environments like a frequented store or vending machine here . This is also a type of exploitation behavior: they go directly to their known, familiar option to minimize effort on decision-making. The removal of that favorite product may act as a "jolt" that breaks this routine. It forces the consumer out of a passive low-involvement exploitation mode into a more active high-involvement exploration mode. Suddenly, they must actively scan the options they typically ignore, making them more likely to notice and consider products in categories they would have otherwise skipped. This transition from habit driven to deliberate conscious decision-making can explain the observed demand spillovers across product categories following the treatment.

The above discussion suggests that if the low-involvement habitual buying mechanism is partly responsible for frequent purchases of the same item, removing a consumer's "favorite" product should lead to increased exploratory and learning behavior. To test this hypothesis, we make use of a synthetic DID design and a linear probability model to examine whether treated consumers are more likely to buy products which they have not tried before. We use a dummy variable to capture exploration behavior at the daily level, which equals 1 if the consumer buy a new product at date t, and 0 otherwise.[15] A *new product* to a consumer at date t is defined as a product that the consumer has not purchased up to date $t-1$, since

station.

[15]Note that this definition will code "no purchase" to be zero as well. Hence, one can interpret our results as a lower bound and a conservative estimate for the exploration behavior.

Janurary 1st, 2018. Using this measure as the dependent variable, we find that the average treatment effect at the station level is significantly positive with a point estimate at 0.027 units, around 73% of its before-treatment mean value (column 1 of Table 7). This provides support for the presence of low-involvement habitual buying behavior.

Table 7: Average Treatment Effects of Favorite Product Removals on Product Experimentation

	New product purchase probability				
	Station	Favorite machines	Non-favorite machines	Non-favorite machines (favorite product available)	Product switching rate
	(1)	(2)	(3)	(4)	(5)
ATE	0.027***	0.015***	0.016***	0.003**	0.046+
	(0.005)	(0.003)	(0.004)	(0.001)	(0.026)
Mean before treatment	0.037	0.018	0.020	0.006	0.203

Note: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. "New product" at date t is defined as the product a consumer never purchased before date t since January 1st, 2018. "Product switching rate" is defined as the weekly proportion of purchases that involve a switch of products. Standard errors are obtained by bootstrapping and reported in parentheses. The mean is calculated as the average of the dependent variables across 14 days prior to the treatment.

To further investigate the learning hypothesis, we examine consumer product switching behavior. A consumer may or may not like a new drink that they explore. Hence, if their choice reflects learning from their consumption experiences of new products, we expect to see more product switching patterns (e.g., Crawford and Shum, 2005; Tehrani and Ching, 2024). We define *product switching rate* as the weekly proportion of purchases that involve a product switch (i.e., the product purchased this time is different from the last purchase). Because our research design considers 14 days before and after the treatment and the product switching rate is a weekly observation, the DID regression for product switching rate is identified by two observations before and after the treatment. Column 5 of Table 7 presents the synthetic DID estimate using the switching rate as the dependent variable at the station level.[16] We find that the average treatment effect to be marginally significant: it increases the product

[16]If a consumer did not purchase for a week, we cannot construct their weekly switching rate and must drop that treated consumer in synthetic DID. Consequently, the number of treated consumers drops from 157 to 107, and the smaller sample size may lead to noisier estimates, and it is indeed what we find.

switching rate by 0.046, around 23% of its pre-treatment mean value at 0.203. This result provides further support that consumers become more active in learning after their favorite product has been removed.[17]

The results of both daily demand for new products and product switching rate consistently indicate that consumers become more inclined to learn and explore/experiment with new products after their favorite products are removed from their favorite machine. We interpret the results as evidence supporting that the habitual behavior is one of the mechanisms in consumers' persistent selection of the same item here.

We also examine the daily probability of purchasing a new product at the favorite machines and non-favorite machines separately (Column 2 and 3 of Table 7). We find that the new product demand increases at not only the favorite machine (point estimate at 0.015 units, 80% of its before-treatment mean value), but also the non-favorite machines (point estimate at 0.016 units, 80% of the before-treatment mean value). One possible explanation for the latter is that when deprived of their favorite product, consumers first engage in search behavior across machines to locate it. However, at a non-favorite machine, they encounter unfamiliar assortments that require them scan through the products and pay more attention to the full product array. As a result, this may trigger them to explore new products.

This mechanism suggests another testable hypothesis: even when a treated consumer lands at a non-favorite machine which has their favorite product, they may still explore more. To test this hypothesis, we examine purchases at non-favorite machines where the consumer's favorite product is available (Column 4 of Table 7). Here, the dependent variable is a dummy variable which equals 1 if a consumer purchases new products at a non-favorite vending machine where her favorite product is available, and 0 otherwise.[18] The average treatment

[17] We have also estimated the average treatment effect on switching rate at the favorite machine and non-favorite machines. The results are presented in Appendix C. Because we observe even fewer purchases in such breakdown, there are more missing switching rate observations, leading to even fewer treated units. We still find positive and marginally significant result at the favorite machine, but the results at non-favorite machines are not significant.

[18] A value of zero for the dummy includes three cases: (i) the consumer purchases at her favorite vending machine, (ii) she purchases at a non-favorite machine where her favorite product is not available, and (iii) she does not purchase. With this definition, we compare pre- and post-treatment unconditional probabilities

effect is 0.003 and significant, representing 50% of its pre-treatment mean. This finding provides support for this hypothesis and additional evidence for our proposed mechanism.

4.1.4 Remarks

Our analysis reveals a noteworthy finding: consumers become more likely to experiment with previously untried products when their "favorite" option is removed. This result not only helps explain the increase in cross-category demand, but also points to a potential long-term gain for consumers. Prior to the treatment, some consumers may be locked into habitual autopilot buying mode. Yet when this routine is disrupted, they do not merely switch to other products that they have tried before. Instead, we observe a marked increase in trials of new products. This behavioral shift suggests that the increase in exploration induced by the removal of favorite products may help consumers discover alternatives that are better aligned with their preferences. We will provide additional descriptive evidence on this in section 6.

4.2 Parallel Trend and Robustness Checks

In this section, we show the parallel trend achieved by the synthetic control units and conduct several robustness checks for the main findings.

Parallel Trend In Appendix A, we show the mean difference of each dependent variable used in the main results between the treated units and the synthetic control units we constructed, for each day before and after the date of the treatment within the 28-day window. The figures show that our constructed control units sucessfully match the pre-treatmenet trend of the treatment units.

of purchasing a new product at a non-favorite machine where the favorite product is available, rather than probabilities conditional on purchasing at such a machine. We adopt this approach for practical reasons. Conditioning on purchases at the relevant non-favorite machine would generate many missing daily observations for each consumer (on days when she does not purchase there). This would yield an unbalanced panel, making it difficult to implement the synthetic difference-in-differences analysis.

Longer Time Window We extend the pre- and post-treatment window of our analysis from 14 to 28 days. Table B1 in Appendix B shows the results remain robust.

Excluding Cases with Favorite Products Returned Approximately 4% of treated consumers in our sample experience the return of their favorite product during the treatment window. While these consumers are included in our baseline analysis, we conduct a robustness check excluding them in the sample. Table B2 in Appendix B shows that our main results remain robust.

Including All Never-Treated Consumers in Control Group As noted above, our baseline specification excludes never-treated consumers who share the same favorite station as treated consumers, in order to mitigate potential spillover effects. However, these individuals may be more appropriate candidates to construct the synthetic control, as they operate in the same environment. As a robustness check, we rerun the synthetic DID analysis by including them to form the synthetic controls. Table B3 in Appendix B presents the results, which remain largely consistent with our main findings.

5 The Role of Variety-seeking Tendency

Building on our preceding analysis, we hypothesize that variety-seeking tendency may play a moderating role in the treatment effect.[19] We propose two behavioral implications.

- High variety-seekers: These consumers maintain a broader portfolio of preferred products. When their favorite product is unavailable, they can more readily substitute it with another product from their portfolio at the same vending machine.

- Low v ariety-seekers: Consumers with limited variety-seeking tendencies rely on a

[19]Note that the concept of a consumer's variety-seeking tendency is closely related to a consumer's loyalty to a subset of products. For instance, variety-seeking behavior could be an outcome of a consumer having similar brand loyalty or strong preferences for several products. For a review of the variety-seeking concept, see McAlister and Pessemier (1982).

narrower set of preferred products. When faced with favorite product removal, they should be more willing to search across locations (i.e., across vending machines) rather than substitute within their favorite machine, reflecting stronger attachment to specific products.

5.1 Measures of Consumer Variety-Seeking Tendency

To test these behavioral implications, we consider four ways to measure variety-seeking tendency: (i) portfolio size, (ii) portfolio instability, (iii) switching tendency, and (iv) non-favorite product share. These measures are closely related, but they capture different dimensions of variety-seeking tendencies.

Table 8: Summary Statistics of Variety-Seeking Measures from January to March

	Mean	Sd	10th Pctl	Median	90th Pctl	Min	Max
Portfolio size	1.576	0.613	1	1.375	2.375	1	4.714
Portfolio instability	1.082	1.065	0	0.812	2.625	0	5.250
Switching tendency	0.248	0.239	0	0.188	0.596	0	1.000
Non-favorite product share	0.206	0.206	0	0.123	0.508	0	0.781

Note: "10th Pctl" and "90th Pctl" report the 10th percentile and 90th percentile of the variables, respectively.

Portfolio size is defined as the average weekly number of distinct products purchased by a consumer. *Portfolio Instability* is the average difference between one's weekly portfolio size and monthly portfolio size.[20] *Switching tendency* is defined as the average weekly proportion of purchases involved switching, where a switch happens when the current purchase is different from the last purchase.[21] Finally, *non-favorite product share* measures how often a

[20]Let N_m be the number of distinct products that a consumer purchases in month m, N_{wm} be the number of distinct products the consumer purchases in week w within month m, and $\bar{N}_{wm}$ be the average of N_{wm} in month m. We use $\Delta N_m := N_m - \bar{N}_{wm}$ as the measure of portfolio instability for month m. Our measure of portfolio instability is the average ΔN_m across m, $\overline{\Delta N_m}$. Intuitively, this captures the idea that variety-seeking consumers may want to try a new set of products from time to time.

[21]The switching tendency is closely related to the product switching rate used in our product experimentation analysis, but they differ in the following aspects: For switching tendency, we use consumer purchase data from January to March and take the average across weeks to measure the tendency at the consumer

consumer purchases their non-favorite products. It is measured as the average of weekly purchase share of the non-favorite products. All four variety-seeking measures are constructed based on consumer purchases from January to March. For all four metrics, a higher value implies a stronger tendency to seek variety. Table 8 reports their summary statistics. By examining their means and standard deviations, we can see that there is decent consumer heterogeneity in variety-seeking tendency (e.g., the average switching tendency is 0.248 with a SD of 0.239).

5.2 Substitution Patterns by Variety-seeking

Our *first hypothesis* is that high variety-seeking consumers will exhibit a larger increase in demand for non-favorite products at their favorite machine following the treatment. The rationale is that these consumers have a broader set of acceptable options and are therefore more likely to find a substitute without changing vending machines. Extending this reasoning, our *second hypothesis* is that low variety-seeking consumers, who have fewer acceptable options, are more likely to switch from their favorite machine to a non-favorite machine to seek their favorite product.

To test these two hypotheses, we median-split our sample into high and low variety-seeking groups and apply synthetic DID to estimate their average treatment effects separately. The results are reported in Table 9. The findings are consistent across all measures of variety-seeking tendencies.

Panel 3 reveals that although consumer demand for non-favorite products at favorite machines increases for both types of consumers, the increase is significantly larger for the high variety-seeking consumers. This provides support for our first hypothesis that high variety-seeking consumers are more likely to find a substitute at their favorite machine.

Turning to panels 2 and 4, it shows that low variety-seeking consumers have a significantly

level. For the product switching rate, we use consumer purchase data within the 14-day time window from the treatment, and it is measured weekly, and hence it varies by week. This measure has been widely used in the variety-seeking literature (e.g., Givon, 1984; Bawa, 1990).

Table 9: Treatment Effects on Individual Demand by Variety-Seeking Tendencies

	High (1)	Low (2)	High - Low (3)
Demand in the favorite station			
Portfolio size	0.013	0.000	0.012
	(0.009)	(0.012)	(0.014)
Portfolio instability	0.021+	-0.006	0.027+
	(0.011)	(0.010)	(0.016)
Switching tendency	0.013	-0.002	0.014
	(0.008)	(0.011)	(0.013)
Non-favorite product share	0.031	-0.005	0.036
	(0.023)	(0.012)	(0.025)
Demand in the favorite machine			
Portfolio size	0.008	-0.059***	0.067***
	(0.007)	(0.009)	(0.011)
Portfolio instability	0.010	-0.059***	0.069***
	(0.010)	(0.009)	(0.014)
Switching tendency	0.008	-0.059***	0.067***
	(0.006)	(0.008)	(0.011)
Non-favorite product share	0.016	-0.060***	0.076***
	(0.019)	(0.009)	(0.020)
Demand in the favorite machine for non-favorite products			
Portfolio size	0.075***	0.042***	0.034*
	(0.008)	(0.010)	(0.013)
Portfolio instability	0.073***	0.043***	0.030*
	(0.009)	(0.008)	(0.014)
Switching tendency	0.071***	0.043***	0.028**
	(0.008)	(0.006)	(0.010)
Non-favorite product share	0.145***	0.042***	0.104***
	(0.020)	(0.009)	(0.022)
Demand in the non-favorite machines			
Portfolio size	0.013+	0.071***	-0.058***
	(0.007)	(0.010)	(0.011)
Portfolio instability	0.017**	0.065***	-0.048***
	(0.006)	(0.010)	(0.012)
Switching tendency	0.014*	0.068***	-0.055***
	(0.006)	(0.010)	(0.012)
Non-favorite product share	0.031*	0.066***	-0.035*
	(0.013)	(0.012)	(0.016)

Note: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. "High" and "Low" refer to consumers with the above-median and below-median variety-seeking measures. Demand is measured in units.

larger decrease (increase) in their demand at favorite (non-favorite) machine. This provides support for our second hypothesis that low variety-seeking consumers are more likely to buy at non-favorite machines after treatment.

For completeness, Panel 1 presents the results for the demand at the favorite station. Consistent with our main findings, both high and low variety-seeking consumers did not change their overall demand at the station level after the removal of their favorite product.

This empirical finding points to a potential link between consumers' variety-seeking tendency and their within- and across-outlet search behavior. While prior research has largely examined these two forms of search separately, our results suggest they may be fundamentally connected.

Table 10: Treatment Effects on Experimentation/Learning by Variety-Seeking Tendencies

	High (1)	Low (2)	High - Low (3)
New product purchase probability			
Portfolio size	0.020***	0.016***	0.004
	(0.004)	(0.004)	(0.006)
Portfolio instability	0.018***	0.017***	0.001
	(0.005)	(0.004)	(0.006)
Switching tendency	0.019***	0.016***	0.003
	(0.004)	(0.004)	(0.005)
Non-favorite product share	0.047***	0.012**	0.035**
	(0.010)	(0.004)	(0.011)
Product switching rate			
Portfolio size	0.025	0.020	0.005
	(0.018)	(0.013)	(0.021)
Portfolio instability	0.036+	0.007	0.029
	(0.020)	(0.014)	(0.023)
Switching tendency	0.026	0.017	0.008
	(0.018)	(0.014)	(0.023)
Non-favorite product share	0.055	0.011	0.044
	(0.041)	(0.012)	(0.044)

Note: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. "High" and "Low" refer to consumers with the above-median and below-median variety-seeking measures. The total number of treated consumers is 142 for the upper panel, while the total number of treated consumers is 106 due to the removal of missing observations for the lower panel.

5.3 Exploration and Learning by Variety-seeking

We also examine whether consumer experimentation and learning behaviors vary by variety-seeking tendency. Results are reported in Table 10. Panel 1 uses the daily new product purchase probability as the dependent variable, while panel 2 focuses on the product switching rate.

Focusing on panel 1, we find that the treatment effect on daily new product purchase probability is positive and statistically significant for both consumer types, ranging from 0.018 to 0.047 for high variety-seeking consumers and from 0.012 to 0.017 for low variety-seeking consumers. The difference between the two groups is generally not statistically significant, except for using non-favorite product share as the measure. These results suggest that the habit disruption mechanism discussed earlier applies broadly: both high and low variety-seeking consumers become more open to trying something they have not purchased before when their favorite product is removed.

When we split the sample by variety-seeking tendencies, the estimated treatment effects on product switching rates remain positive in both the high and low groups, but they are not statistically significant. This can happen because splitting the sample reduces the effective sample size, which leads to a loss in the statistical power. Therefore, the bootstrap standard errors become larger even if the underlying effect is similar. Regardless of significancy, we find no statistically significant difference between the treatment effects for the high and low variety-seeking groups.

6 Longer Term Descriptive Evidence

We have provided evidence that consumers become more willing to explore during the 14 days after their favorite product has been removed. Because our DID analysis focuses on the short-term outcome, a natural questions are: Did consumers find a better match from this increase in exploration behavior? Are there any structural changes in their purchase

behavior?

To shed light on the first question, we look into the extent to which the new products that they explored during the 14-day window become part of their regular product portfolio. More specifically, we examine whether the new products tried within 14-day window in May after treatment account for a meaningful share of consumers' purchases over a longer horizon from June through the end of August (which is still within our observation window). For each consumer, we compute the purchase share of each new product tried within 14 days after treatment from June to August, and then take the maximum of these shares across the consumer's new products. The first row of Table D1 reports the results. Across consumers, this maximum share has a mean of 0.26 and a standard deviation of 0.32. Moreover, the 90th percentile is 0.82 and the maximum is 1, indicating that at least 10% of consumers like a new product discovered right after the intervention so much that it becomes their favorite product. This pattern is more consistent with habitual behavior before treatment than with pure substitution. Under proportional substitution, the removal of one product reallocates demand toward existing alternatives in roughly fixed proportions, implying that products with negligible pre-treatment choice probabilities should remain niche even post-treatment. In contrast, we find that many consumers not only sample a new product shortly after treatment, but subsequently allocate a substantial fraction of their purchases to that product over the remaining sample period. This persistence is consistent with habit purchases before the treatment.

We also examine treated consumers' variety-seeking measures from June through August after the treatment, as a comparison to Table 8. Rows 2 to 5 of Table D1 reports the results. The patterns are consistent with a sustained increase in variety seeking tendencies: mean portfolio size rises from 1.568 to 1.604; mean portfolio instability rises from 1.059 to 1.296; mean switching tendency rises from 0.250 to 0.333; and mean non-favorite product share rises from 0.206 to 0.355. While these comparisons are descriptive rather than causal, they are suggestive evidence that removing a favorite product may induce consumers to explore

more

7 Conclusion

This study examines consumer responses to the removal of favorite products at their favorite vending machines. Leveraging an unusually detailed micro level data set of beverage vending machines at Japanese railway stations and applying synthetic DID, we document four key findings: Following the treatment, (i) consumers decrease purchases at their favorite machines, (ii) the purchase reduction at their favorite machines is largely offset by increased purchases at alternative machines, (iii) consumers become more likely to explore previously untried products and (iv) consumer's product switching also increases. In particular, we interpret the last two findings through the level of involvement in attention theory (Celsi and Olson, 1988; Webb et al., 2024): the product removal disrupts habitual choice, shifting consumers from an automated, low-involvement state to an active, high-involvement evaluation state. This suggests that apparent "favorites" may partly reflect low-involvement habitual choice rather than true preferences.

These insights are potentially important for modeling consumer learning behavior. This literature assumes that consumers are always in high-involvement decision mode and actively balancing exploration and exploitation incentives in their choices (Erdem and Keane, Tehrani and Ching, etc.). Hence, in these models, the reason why a consumer keeps choosing the same product must be because its exploitation value is higher than the exploration value of other unfamiliar choices. The insights we learn from this study challenges this view – some consumers may just use low-involvement decisions by default, and only start considering exploring unfamiliar options when something triggers it.

These insights are also useful for managers grappling with the difficult decision of whether to remove an under-performing product from a portfolio. Traditionally, this decision is weighed through a narrow cost-benefit analysis that considers only the immediate loss of

sales from that product and the potential for customer attrition. Our findings, however, highlight a significant and previously overlooked potential benefit: the strategic elimination of a consumer's "favorite" product can act as a powerful catalyst for experimentation. By disrupting the consumer's low involvement habitual choice pattern, the removal forces a cognitive re-engagement with the category. This can lead customers to discover new, better-matched products that they otherwise would not have tried due to habitual inertia, thereby potentially increasing their long-term satisfaction and their lifetime value. Furthermore, this forced experimentation can serve as an accelerator for the diffusion of new products, as a segment of now-active explorers provides early adoption and word-of-mouth marketing. Therefore, managers should re-conceptualize product removal not merely as a cost-saving measure but as a potential growth strategy and a targeted intervention to revitalize consumer engagement within the brand's ecosystem.

Our analysis further reveals an interesting heterogeneous consumer response based on their variety-seeking tendencies. We find that high variety-seeking consumers, when confronted with the removal of their favorite product, are much more likely to substitute within the same favorite machine. In contrast, low variety-seeking consumers are significantly more likely to respond by switching to a different machine. We also see evidence that consumers become more likely to explore new products regardless of their variety-seeking tendency. But low variety-seeking consumers switch products more often, suggesting that they are more selective and it is harder for them to find a right match. The change in switching behavior is largely insignificant for high variety-seeking consumers, indicating that they are more easily satisfied with new products. This suggests that the high variety-seekers may benefit more from the boosted exploration, as it might help them to find a wider set of products that match their preferences.

If we consider vending machines as counterpart for retail outlets such as grocery stores or convenience stores, consumer substitution pattern across stores might pose a strategic risk. When consumers visit another store to look for their favorite product, it could lead

to defection to a competitor. This implies that the decision to remove a product is not merely a category management choice, but also a strategic calculus that must account for heterogeneous consumer reactions and the competitive structure of the market.

Funding and Competing Interests

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no funding to report.

References

Anupindi, Ravi, Maqbool Dada, and Sachin Gupta, "Estimation of Consumer Demand with Stock-Out Based Substitution: An Application to Vending Machine Products," *Marketing Science*, 1998, *17* (4), 406–423.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager, "Synthetic Difference-in-Differences," *American Economic Review*, 2021, *111* (12), 4088–4118.

Bawa, Kapil, "Modeling Inertia and Variety Seeking Tendencies in Brand Choice Behavior," *Marketing Science*, 1990, *9* (3), 263–278.

Boatwright, Peter and Joseph C. Nunes, "Reducing Assortment: An Attribute-Based Approach," *Journal of Marketing*, 2001, *65* (3), 50–63.

Borle, Sharad, Peter Boatwright, Joseph B. Kadane, Joseph C. Nunes, and Shmueli Galit, "The Effect of Product Assortment Changes on Customer Retention," *Marketing Science*, 2005, *24* (4), 616–622.

Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Celsi, Richard L. and Jerry C. Olson, "The Role of Involvement in Attention and Comprehension Processes," *Journal of Consumer Research*, 1988, *15* (2), 210–224.

Ching, Andrew T. and Matthew Osborne, "Identification and Estimation of Forward-Looking Behavior: The Case of Consumer Stockpiling," *Marketing Science*, 2020, *39* (4), 707–726.

Conlon, Christopher T. and Julie Holland Mortimer, "Demand Estimation under Incomplete Product Availability," *American Economic Journal: Microeconomics*, 2013, *5* (4), 1–30.

Crawford, Gregory S. and Matthew Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, *73* (4), 1137–1173.

Erdem, Tülin, Susumu Imai, and Michael P. Keane, "Brand and Quantity Choice

Dynamics Under Price Uncertainty," *Quantitative Marketing and Economics*, 2003, *1* (1), 5–64.

Givon, Moshe, "Variety Seeking through Brand Switching," *Marketing Science*, 1984, *3* (1), 1–22.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Hendel, Igal and Aviv Nevo, "Measuring the Implications of Sales and Consumer Inventory Behavior," *Econometrica*, 2006, *74* (6), 1637–1673.

Kawaguchi, Kohei, Kosuke Uetake, and Yasutora Watanabe, "Effectiveness of Product Recommendations Under Time and Crowd Pressures," *Marketing Science*, 2019, *38* (2), 253–273.

_ , _ , and _ , "Designing Context-Based Marketing: Product Recommendations Under Time Pressure," *Management Science*, 2021, *67* (9), 5642–5659.

MacGregor, Jim, "The Beverage Industry in Mid 2024: Trends, Challenges, and Future Outlook for 2025," *Whale Pod*, 2024. Accessed May 12, 2025, `https://whalepodshipper.com/blogs/whale-pod/the-beverage-industry-in-mid-2024-trends-challenges-and-future-outlook-for-2025`.

McAlister, Leigh and Edgar Pessemier, "Variety Seeking Behavior: An Interdisciplinary Review," *Journal of Consumer Research*, 1982, *9* (3), 311–322.

Sloot, Laurens M., Dennis Fok, and Peter C. Verhoef, "The Short- and Long-Term Impact of an Assortment Reduction on Category Sales," *Journal of Marketing Research*, 2006, *43* (4), 536–548.

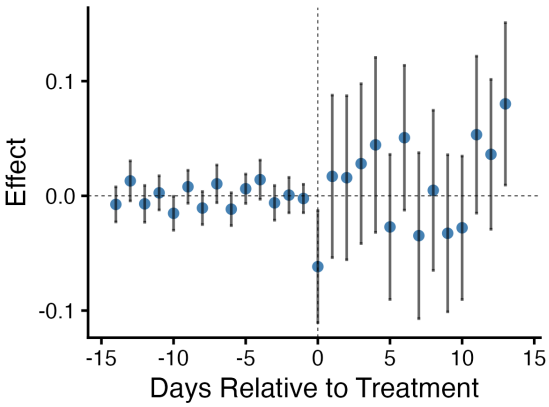
Tehrani, Shervin Shahrokhi and Andrew T. Ching, "A Heuristic Approach to Explore: The Value of Perfect Information," *Management Science*, 2024, *70* (5), 3200–3224.

Webb, Ryan, Jessica Fong, Asaf Mazar, Julia Levine, Alexandra Steiny Wellsjo, Olivia Natan, Clarice Zhao, Phillippa Lally, Sanne de Wit, John Odoherty, Andrew T. Ching, Raphael Thomadsen, Matthew Osborne, Peter Landry,

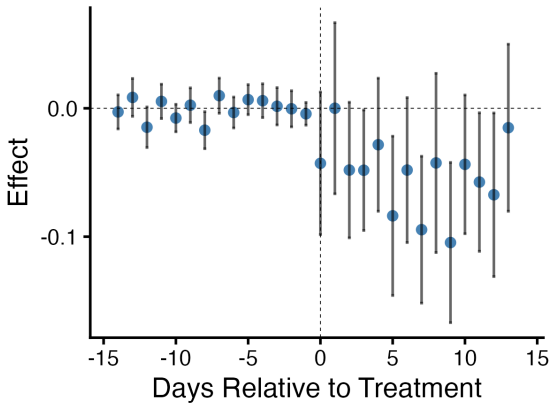
Mark Bouton, Wendy Wood, and Colin Camerer, "Integrating Neuro-Psychological Habit Research into Consumer Choice Models," *SSRN Working Paper*, 2024. `https://papers.ssrn.com/abstract=4853969`.

Appendix

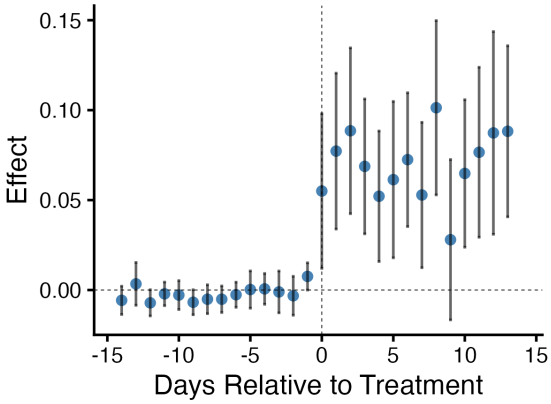
A Parallel Pretend Checks for Synthetic Control Units



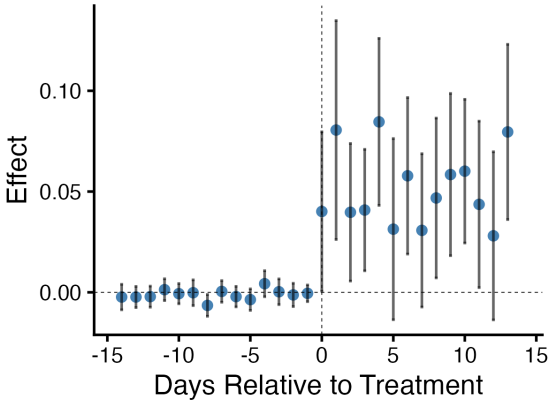
(a) Favorite Station



(b) Favorite Machine



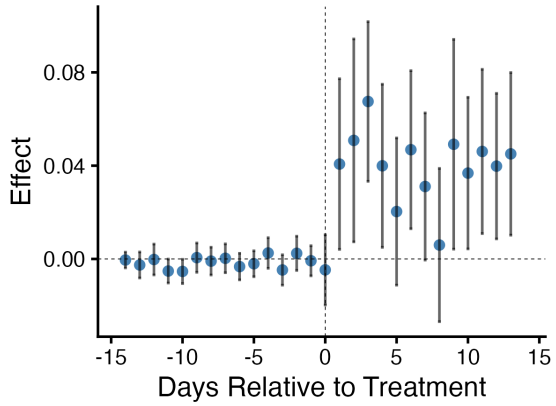
(c) Same Category Products at Favorite Machines



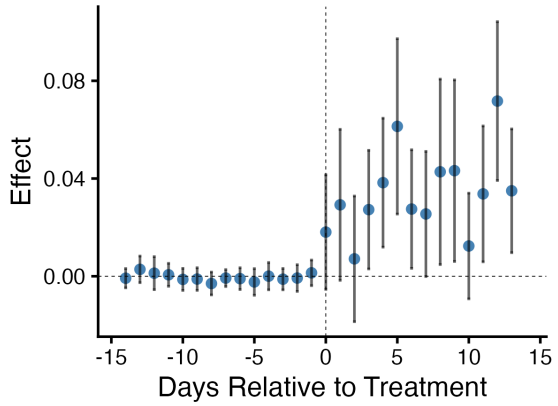
(d) Other Categories Products at Favorite Machines

Figure A1: Pre-Trend Checks for Individual Demand at Station and Favorite Machine

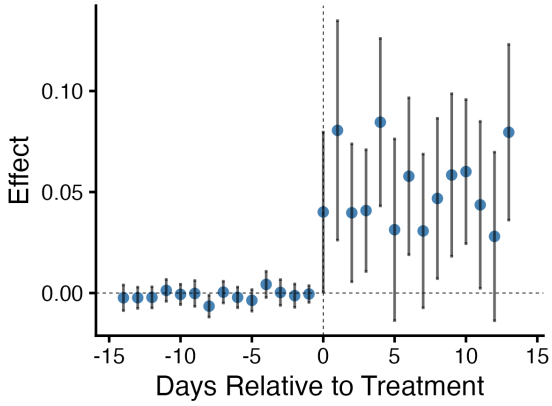
Note: The figures show the mean difference of the dependent variable between the treated and synthetic control units with 95% confidence interval at each period.We first computed the mean difference for a given treatment cohort (e.g. consumers treated on May 1st), and then take the weighted average across the treatment cohort as shown in Equation (7). The standard error is calculated through bootstrapping.



(a) Favorite Products at Non-favorite Machines



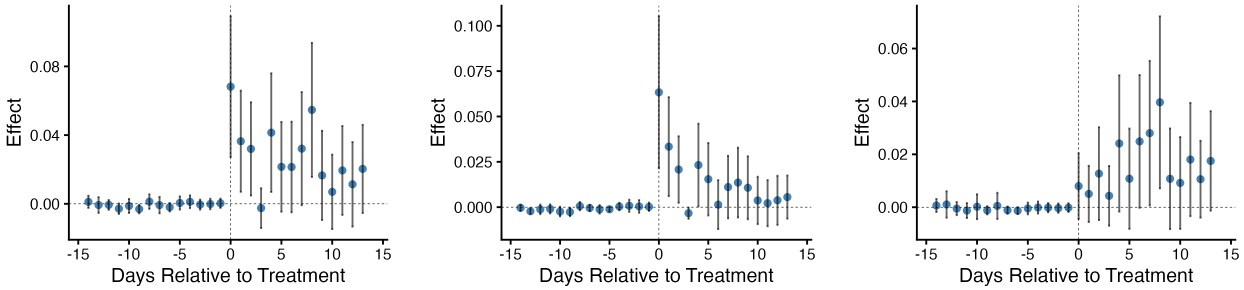
(b) Same Category Products at Non-favorite Machines



(c) Other Categories Products at Non-favorite Machines

Figure A2: Pre-Trend Checks for Individual Demand at Non-favorite Machines

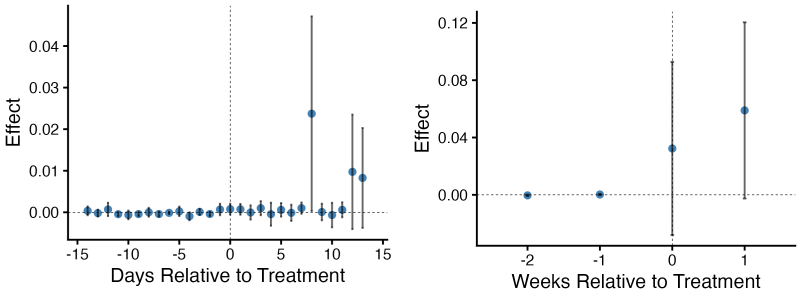
Note: The figures show the mean difference of the dependent variable between the treated and synthetic control units with 95% confidence interval at each period. We first computed the mean difference for a given treatment cohort (e.g. consumers treated on May 1st), and then take the weighted average across the treatment cohort as shown in Equation (7). The standard error is calculated through bootstrapping.



(a) New Products at Favorite Station

(b) New Products at Favorite Machine

(c) New Products at Non-favorite Machines



(d) New Products at Non-favorite Machines with Favorite Product Available

(e) Product Switching Rate at Favorite Station

Figure A3: Pre-Trend Checks for Product Experimentation Studies

Note: The figures show the mean difference of the dependent variable between the treated and synthetic control units with 95% confidence interval at each period. The first four panels present the mean difference in new product purchasing probability for each location. We first computed the mean difference for a given treatment cohort (e.g. consumers treated on May 1st), and then take the weighted average across the treatment cohort as shown in Equation (7). The standard error is calculated through bootstrapping.

B Robustness Checks for Average Treatmemt Effects

Table B1: Longer Treatment Period

	Individual demand							Product experimentation	
	Station	Favorite machine			Non-favorite machines				
	All products	All products	Same category products	Other categories proodducts	Favorite products	Same category products	Other categories products	New product purchasing probability	Product switching rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ATE	0.004	-0.074***	0.065***	0.034***	0.035**	0.039***	0.031**	0.018***	0.050+
	(0.016)	(0.018)	(0.011)	(0.009)	(0.011)	(0.007)	(0.010)	(0.004)	(0.028)

Note: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. The analysis extends the post-treatment (and pre-treatment) window from 14 to 28 days. Standard errors are obtained by bootstrapping and reported in the parentheses. Demand is measured in units.

Table B2: Average Treatment Effects Without Return of Favorite Product

	Individual demand							Product experimentation	
	Station	Favorite machine			Non-favorite machines				
	All products	All products	Same category products	Other categories products	Favorite products	Same category products	Other categories products	New product purchasing probability	Product switching rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ATE	0.004	-0.059***	0.066***	0.053***	0.033**	0.036***	0.026*	0.029***	0.045+
	(0.015)	(0.014)	(0.010)	(0.011)	(0.011)	(0.008)	(0.011)	(0.005)	(0.025)

Note: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. Standard errors are obtained by bootstrapping and reported in the parentheses. Demand is measured in units.

Table B3: Average Treatment Effects With All Never-Treated Consumers

	Individual demand							Product experimentation	
	Station	Favorite machine			Non-favorite machines				
	All products	All products	Same category products	Other categories products	Favorite products	Same category products	Other categories products	New product purchasing probability	Product switching rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ATE	0.011	-0.049***	0.070***	0.051***	0.036**	0.033***	0.026*	0.027***	0.046+
	(0.017)	(0.013)	(0.011)	(0.010)	(0.011)	(0.007)	(0.011)	(0.004)	(0.026)

Note: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$. Standard errors are obtained by bootstrapping and reported in the parentheses. Demand is measured in units.

C Switching Rate

This section reports treatment effects on product switching rates, constructed using alternative definitions. We define the weekly product switching rate as the number of switches divided by the total number of purchases in that week. When a week has zero purchases, the switching rate is undefined and recorded as missing. Table C1 excludes these missing observations, whereas Table C2 imputes zero switching rates for weeks with no purchases.

Table C1: Average Treatment Effects on Product Switching Rate I

	Product switching rate	Product switching rate favorite machine	Product switching rate other machines	Product switching rate other machines (favorite product available)
	(1)	(2)	(3)	(4)
ATE	0.046+	0.035	0.033	-0.033
	(0.026)	(0.024)	(0.043)	(0.027)
Number of treated consumers	106	63	42	15

Note: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$.

Table C2: Average Treatment Effects on Product Switching Rate II [to be removed]

	Station	Favorite machine	Non-favorite machines	Non-favorite machines (favorite product available)
	(1)	(2)	(3)	(4)
ATE	0.032+	-0.001	0.026+	0.008
	(0.018)	(0.010)	(0.013)	(0.005)
Number of treated consumers	142	142	142	142

Note: $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$.

D After Treatment Variety-seeking Tendencies

Table D1: Summary Statistics of Variety-Seeking Measures From June to August

	Mean	Sd	10th Pctl	25th Pctl	Median	75th Pctl	90th Pctl	Min	Max
Maximum new product purchase share	0.264	0.320	0.016	0.024	0.125	0.425	0.820	0.013	1.000
Summed new product purchase share	0.277	0.318	0.016	0.024	0.135	0.425	0.820	0.013	1.000
# New product tried (if any)	1.519	1.038	1.000	1.000	1.000	2.000	3.000	1.000	6.000
Portfolio size	1.604	0.648	1.000	1.083	1.455	1.885	2.361	1.000	4.333
Portfolio instability	1.296	1.101	0.250	0.562	1.076	1.579	2.530	0.000	6.031
Switching tendency	0.333	0.290	0.000	0.042	0.313	0.532	0.746	0.000	1.000
Non-favorite product share	0.355	0.245	0.000	0.108	0.397	0.556	0.642	0.000	0.838

Note: "10th Pctl" and "90th Pctl" report the 10th percentile and 90th percentile of the variables, respectively. "New products" are those that treated consumers tried for the first time within 14 days after treatment. The first row reports summary statistics for each consumer's maximum purchase share among these new products. The second row reports summary statistics for each consumer's total purchase share across all new products.

Table D2: Summary Statistics of Variety-Seeking Measures from January to March

	Mean	Sd	10th Pctl	25th Pctl	Median	75th Pctl	90th Pctl	Min	Max
Portfolio size	1.568	0.575	1	1.125	1.402	1.875	2.375	1	3.625
Portfolio instability	1.059	1.041	0	0.375	0.750	1.469	2.613	0	5.250
Switching tendency	0.250	0.240	0	0.043	0.188	0.382	0.621	0	1.000
Non-favorite product share	0.206	0.207	0	0.031	0.135	0.355	0.523	0	0.781

Note: "10th Pctl" and "90th Pctl" report the 10th percentile and 90th percentile of the variables, respectively.