Broad Promotions, Narrow Responses: Decomposing Causal Effects in Large Product Assortments

Khaled Boughanmi*, Raghuram Iyengar, and Young-Hoon Park

Abstract

Broad-scope promotions give consumers flexibility over where to redeem a discount. For managers, however, evaluating these promotions requires answering two questions: How much incremental spending do they generate, and where within the assortment does that spending originate? Randomized field experiments answer the first question but leave the second unresolved because explaining the incremental lift depends on how products are grouped for analysis. We study these questions using a randomized field experiment involving a broad-scope coupon at a retailer with a large product line. The experiment establishes that the promotion increases spending, while a retailer-category analysis suggests that these gains are broadly distributed across the assortment. This explanation relies on the retailer's product taxonomy, which may not reflect how consumers organize their purchases. Instead, we infer behavioral product groupings from customers' pre-treatment purchase histories and decompose the experimentally identified treatment effect across them. Our analysis reveals a markedly different picture of where promotional gains originate. The inferred groupings frequently cut across retailer categories and, viewed through these groupings, incremental spending is concentrated within a small number of them rather than broadly distributed across the assortment. These findings suggest that broad-scope promotions create broad choice but may generate surprisingly narrow behavioral responses.

Keywords: Broad-scope promotions; causal decomposition; Bayesian nonparametrics; latent demand components; treatment-effect heterogeneity; large product assortments.

*Khaled Boughanmi is an Assistant Professor of Marketing, Samuel Curtis Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853. Email: kb746@cornell.edu. He is the corresponding author. Raghuram Iyengar is the Miers-Busch W'1885 Professor and Professor of Marketing, Department of Marketing, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104. Email: riyengar@wharton.upenn.edu. Young-Hoon Park is the Sung-Whan Suh Professor of Management and Professor of Marketing, Samuel Curtis Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853. Email: yp34@cornell.edu.

1 Introduction

A growing share of promotional activity in retail takes the form of broad-scope offers that can be redeemed across large product assortments. Examples include Amazon's sitewide credits, Target's basket-level discounts (e.g., "$X off $Y"), and platform-wide incentives offered by services such as Uber and DoorDash. Unlike traditional product- or category-specific promotions, these offers do not prescribe where consumers can redeem the incentive. Instead, consumers decide which products, categories, or combinations of products receive the discount. Firms use broad-scope promotions because they are simple to administer while preserving consumer flexibility. At the same time, they raise an important managerial question: where within the assortment do broad-scope promotions actually generate incremental demand?

Answering this question is more difficult than determining whether the promotion works. A randomized field experiment can establish whether a broad-scope promotion increases overall spending, providing a clean estimate of its average causal effect. Explaining where that incremental spending originates, however, requires moving beyond the aggregate analysis. It requires organizing products into meaningful groups and then decomposing the treatment effect across these groups. This distinction becomes consequential in modern retail settings, where assortments routinely contain tens of thousands of products.

For large assortments, existing approaches offer only partial answers on how to group products. Product-level analyses avoid the grouping problem altogether, but are difficult to execute because purchases are sparse across thousands of products, and even where estimates are available, thousands of separate product-level effects are difficult to make sense of (Kalyanam, Borle, and Boatwright, 2007). Retailer-defined categories provide a practical alternative and therefore dominate both managerial practice and academic research. These categories, however, organize products for merchandising, inventory management, and reporting rather than for explaining consumer purchasing behavior. As a result, they may obscure (and sometimes even mislead) where promotional effects actually arise. The promotions literature does not resolve this issue either: studies of product-specific discounts and targeted promotions have substantially advanced our understanding of promotional effectiveness (e.g., Ailawadi et al., 2007; Anderson and Simester, 2004; Blattberg, Briesch, and Fox, 1995; Leeflang and Parreño-Selva, 2012; Nijs et al., 2001; Van Heerde, Gupta, and Wittink, 2003; Van Heerde, Leeflang, and Wittink, 2004), but typically evaluate effects at predefined levels of aggregation, such as product-level sales, or retailer-defined categories. These studies establish how to identify promotional effects, but not how to organize the products across which those effects unfold.

Guidance on this question instead comes from two streams of literature outside of promotions. One concerns the resolution at which demand is analyzed: inference changes with analytical resolution (Hui and Bradlow, 2012), and managerial conclusions depend on the granularity at which marketing data and model parameters are defined (Kim, Bradlow, and Iyengar, 2026). A second concerns the structure of consumer purchasing behavior, which frequently transcends retailer-defined product categories: consumer purchases cut across administrative taxonomies (Broniarczyk, Hoyer, and McAlister, 1998; Jacobs, Fok, and Donkers, 2021; Jacobs, Donkers, and Fok, 2016; Rooderkerk and Lehmann, 2021), and can be predictively grouped using scalable methods that recover product relationships from purchasing data (Jiang, Li, and Zhang, 2025; Ruiz, Athey, and Blei, 2020). Other research has examined consumer choice (Gabel and Timoshenko, 2022) and competition (Ringel and Skiera, 2016) in large product assortments, but has not studied the impact of broad promotions in these settings. Collectively, this literature suggests that retailer-defined categories may not provide the most behaviorally meaningful way to organize demand. What remains unaddressed is how such groupings can be integrated with randomized experiments to explain, rather than predict, where promotional treatment effects originate.

We address this challenge by developing a Bayesian causal decomposition framework that constructs meaningful product groupings directly from customers' pre-treatment purchase histories. We refer to these groupings as latent demand components. Rather than relying on retailer-defined categories, the framework learns a small number of components that summarize recurring patterns of purchasing behavior across the assortment. We then decompose the experimentally identified treatment effect across the inferred components. As the components are learned exclusively from pre-treatment purchasing behavior, they remain independent of treatment assignment and post-treatment outcomes, preserving the causal interpretation afforded by the randomized experiment.

We apply the framework to a randomized field experiment conducted by a major Asian personal-care retailer with a large product line. Customers were randomly assigned to receive a broad-scope coupon or to a control condition with no coupon. This setting is particularly well suited to our research question because the size of the assortment, typical for large retailers, makes product-level analyses impractical. At the same time, since the randomized design identifies the causal effect of the promotion, our analysis can focus on how that effect is distributed across alternative representations of demand.

The experiment establishes that the promotion increases spending. A category-level analysis suggests these gains are broadly distributed across major merchandise categories. The inferred components from our analysis, however, frequently cut across retailer categories and represent purchasing behavior substantially better than the retailer's taxonomy does. Decomposing the same treatment effect across these components reveals a markedly different picture of where promotional gains originate: incremental spending is concentrated within a small number of them, and consumers increase purchases primarily within components in which they were already active before treatment. Broad-scope promotions thus create broad choice, but that flexibility does not translate into broad behavioral change.

The remainder of the paper is organized as follows. We begin with the empirical setting and show the aggregate- and firm imposed category-level treatment effects. We discuss why product grouping can change our understanding about the origin of promotional effects. We then introduce the notion of latent demand components and our proposed modeling framework. Next, we show the treatment effects across the inferred demand components and how they present a different view of where the promotional gains are originating from. Finally, we discuss the managerial implications of the findings and conclude.

2 Empirical Setting

We use data from a randomized field experiment involving coupon promotions, originally described by Gopalakrishnan and Park (2021). The experiment was conducted by a leading Asian personal care retailer to evaluate the impact of mobile coupons on customer behavior. Customers were randomly assigned to receive a coupon or to a control condition that received no communication.

This setting is particularly well suited to our research question because it combines randomized treatment assignment with a large retail assortment. The retailer carries a large product line spanning a broad range of personal care products and organizes them into a conventional merchandise hierarchy in which each product belongs to a single retailer-defined category (e.g., skincare or makeup). The randomized design identifies the causal effect of the promotion, while the size and complexity of the assortment make it challenging to understand where within the assortment those effects originate. Our analysis focuses on comparing customers assigned to receive a coupon ($10 off a $20 purchase) with those assigned to the control condition.[1] The resulting dataset contains 4,247 customers, of whom 3,446 received the coupon and 801 were assigned to the control condition. Randomization checks reported by Gopalakrishnan and Park (2021) (Table 1, Panel B) support the validity of the experimental design.

The data consist of detailed transaction records at both the customer and product levels, including purchase timestamps, spending, product descriptions, brands, and retailer-defined categories, together with customer demographics such as age and gender. We observe purchasing behavior during a 12-month pre-treatment period and throughout the one-week campaign period. Our primary outcome is each customer's total spending during the campaign.[2]

2.1 Treatment Effects

Table 1: Aggregate- and Retailer-Category-Level Treatment Effects

Outcome	Density	Control Mean	Treated Mean	ATE	SE	p
Aggregate outcome						
Total Spending	1.000	1.199	3.142	1.943	0.349	<0.001
Retailer-category outcomes						
Skin Care	0.388	0.848	1.972	1.124	0.283	<0.001
Makeup	0.186	0.166	0.611	0.444	0.109	<0.001
Other	0.162	0.054	0.237	0.183	0.079	0.021
Bath & Body Care	0.140	0.055	0.193	0.138	0.053	0.009
Hair Care	0.124	0.076	0.129	0.053	0.046	0.250

[1] We analyze the treatment group in Gopalakrishnan and Park (2021) consisting of high-value customers who received the premium coupon, as this group is of the greatest managerial importance; their study includes additional treatment groups that we do not analyze here.

[2] All transactions were recorded in the local currency. Purchase amounts are converted to U.S. dollars using the average exchange rate over the sample period.

Table 1 reports treatment effects at both the aggregate and retailer-category levels. At the aggregate level, replicating the results of Gopalakrishnan and Park (2021), the coupon increases total post-treatment spending by $\$1.943$ ($p < 0.001$), confirming a substantial overall promotional effect.

Decomposing the response across the retailer's administrative categories reveals that the effect is broadly distributed throughout the assortment. Skin Care accounts for the largest increase in spending, with an average treatment effect of $\$1.124$ ($p < 0.001$), followed by Makeup ($0.444, $p < 0.001$). Smaller but statistically significant increases are observed for Other ($0.183, $p = 0.021$) and Bath & Body Care ($0.138, $p = 0.009$), whereas the effect for Hair Care is positive but not statistically distinguishable from zero ($0.053, $p = 0.250$).

Comparing the treatment effects with the category densities suggests that the promotion generates broad gains across the retailer's assortment. A manager relying on this analysis would likely conclude that the coupon stimulated demand across most major merchandise categories, implying that its effects were broadly distributed throughout the assortment.

3 Product Grouping and the Origin of Promotional Effects

The category-level analysis suggests that the promotion generates broad gains across the retailer's assortment, at least when products are grouped by the retailer's product taxonomy. As we show next, this conclusion depends critically on how products are grouped in the first place.

Retailer-defined categories are the conventional way of organizing products within an assortment and therefore form the basis for most managerial analyses. Alternative groupings may instead organize products according to shared functional or behavioral relationships. As a result, the same experimentally identified treatment effect may be distributed differently depending on the grouping that is used. To build intuition for this idea, suppose an analyst knows exactly why each customer purchases each product. Consider the following stylized example involving a small assortment. The limited number of products allows the underlying behavioral pattern to be identified through direct inspection before turning to the additional challenges posed by large assortments.

Stylized Example: Broad Effects Across Retail Categories and Concentrated Effects Within a Usage Routine Consider the example in Table 2, with $I = 2$ customers and $J = 15$ beauty products. Grouping 1 follows the retailer's five merchandise categories, whereas Grouping 2 organizes products according to three consumer usage routines. Let τ_{ij} denote the treatment effect for customer i on product j.

The product-level average effects are

$$\bar{\tau}_1 = \bar{\tau}_4 = \bar{\tau}_7 = \bar{\tau}_{10} = \bar{\tau}_{13} = 2, \qquad \bar{\tau}_j = 0 \quad \text{otherwise},$$

where $\bar{\tau}_j = I^{-1} \sum_{i=1}^{I} \tau_{ij}$. For grouping $h \in \{1, 2\}$, define the effect attributed to group g as

$$\tau_{1,g}^{(h)} = \frac{1}{I} \sum_{i=1}^{I} \sum_{j=1}^{J} \mathbb{I}_g^{(h)}(j) \tau_{ij},$$

Table 2: Retail Categories, Usage Routines, and Customer-Level Treatment Effects

j	Product	Retail Category	Usage Routine	τ_{1j}	τ_{2j}
1	Cucumber face mask	Skin Care	Hydration	3	1
2	Brightening serum	Skin Care	Appearance	0	0
3	Foaming cleanser	Skin Care	Cleansing	0	0
4	Hydrating primer	Makeup	Hydration	3	1
5	Lip color	Makeup	Appearance	0	0
6	Makeup remover	Makeup	Cleansing	0	0
7	Honey lip balm	Other	Hydration	3	1
8	False eyelashes	Other	Appearance	0	0
9	Cotton cleansing pads	Other	Cleansing	0	0
10	Yogurt body cream	Bath & Body Care	Hydration	3	1
11	Body shimmer	Bath & Body Care	Appearance	0	0
12	Citrus body wash	Bath & Body Care	Cleansing	0	0
13	Honey hair mask	Hair Care	Hydration	3	1
14	Hair color cream	Hair Care	Appearance	0	0
15	Clarifying shampoo	Hair Care	Cleansing	0	0

where $\mathbb{I}_g^{(h)}(j)$ indicates that product j belongs to group g under grouping h.

Under the retailer grouping, each of the five categories contains one responsive product, yielding

$$\tau_{1,g}^{(1)} = 2, \qquad g = 1, \ldots, 5.$$

The response therefore appears broadly distributed across Skin Care, Makeup, Other, Bath & Body Care, and Hair Care. Under the usage-routine grouping, however, all responsive products belong to the Hydration routine, yielding

$$\tau_{1,1}^{(2)} = 10, \qquad \tau_{1,2}^{(2)} = \tau_{1,3}^{(2)} = 0.$$

Thus, the same response appears broad across the retailer's categories but fully concentrated within a single consumer routine. Both groupings nevertheless recover the same aggregate treatment effect:

$$\tau_1^a = \sum_{g=1}^{5} \tau_{1,g}^{(1)} = \sum_{g=1}^{3} \tau_{1,g}^{(2)} = 10.$$

The stylized example establishes a key point: the concentration of an experimentally identified treatment effect depends on how products are grouped. Even in this simplified setting, the choice of grouping can generate materially different inventory decisions. The retailer-category decomposition would encourage managers to distribute incremental inventory and replenishment capacity across all five categories, since each appears to respond to the promotion, yet this would allocate inventory to many products with no incremental demand. The behavioral grouping instead identifies Hydration as the sole source of the response, directing inventory toward responsive hydration products regardless of retailer category. Although both groupings recover the same aggregate treatment effect, they lead to fundamentally different conclusions about how resources should be allocated.

The example, however, intentionally sets aside two complications that arise in real-world empirical settings. First, it assumes purchasing routines are both known and common across customers: an analyst is assumed to know exactly why each customer purchases each product, and that reason is assumed to be shared across customers. In practice, neither assumption may hold: these relationships are latent and must be inferred from purchasing data, and they can vary across customers. Second, with only fifteen products, an analyst can still inspect product-level treatment effects directly and recognize that the response concentrates within a hydration routine. Large assortments remove this option altogether: as assortments expand to thousands of products, examining product-level responses one by one is no longer possible, and even perfect knowledge of every product-level treatment effect would not, by itself, reveal the behavioral patterns underlying the aggregate response. Products are the objects we observe while consumer demand is the behavior we seek to understand. Connecting products to demand therefore requires organizing them into behaviorally meaningful groupings, inferred without assuming common routines across customers, and decomposing the experimentally identified treatment effect across those groupings.

The discussion thus far establishes why product grouping matters and why large assortments make it difficult to identify appropriate groupings through direct inspection. The next section develops a data-driven approach for learning behaviorally meaningful product groupings directly from customers' pre-treatment purchase histories. These groupings provide the foundation for decomposing the experimentally identified treatment effect across the underlying patterns of consumer demand.

4 Latent Demand Components

Motivated by the previous section, we infer behaviorally meaningful product groupings directly from customers' purchase histories. We refer to these groupings as latent demand components. A demand component represents a pattern of demand shared across products in the assortment (e.g., Jacobs, Donkers, and Fok, 2016; Urban and Hauser, 2004). Each customer is characterized by a different combination of these components, which together summarize their purchasing preferences.

Recovering these components from data, however, imposes several requirements on the model used for the subsequent causal decomposition. First, it should produce components that are interpretable from a managerial perspective. Second, it should let the data determine the number of components rather than fixing it in advance, allowing the approach to adapt to different assortments and demand structures. Third, it should allow both customers and products to be probabilistically associated with multiple components, reflecting the overlapping nature of purchasing routines in large assortments. Fourth, it should support an additive decomposition of the aggregate treatment effect into component-level effects. Finally, it should preserve the causal interpretation of the subsequent treatment-effect decomposition.

The hierarchical Dirichlet process (HDP; Teh et al., 2006) provides a framework that aligns well with these requirements. Unlike finite mixture models (Blei, Ng, and Jordan, 2003), the HDP allows the effective number of components to be inferred from the data while retaining a mixed-membership representation. This differs from structured choice models such as SHOPPER (Ruiz, Athey, and Blei, 2020), which model item-level interactions directly rather than producing a small number of interpretable, reusable components. The HDP's component-specific product distributions facilitate substantive interpretation by identifying the products that most strongly characterize each latent demand component. In sum, the HDP therefore provides a flexible and interpretable intermediate level of analysis between coarse retailer-defined categories and sparse product-level outcomes. Realizing this benefit for causal analysis, however, requires estimating the HDP exclusively on pre-treatment data, with its posterior predictive distributions subsequently used to construct component-level outcomes and estimate their treatment effects.

4.1 Data Generating Process of Pre-treatment Purchases

We represent a demand component k by a probability distribution $\boldsymbol{\phi}_k = (\phi_{k1}, \ldots, \phi_{kJ})$ over the J products. The element ϕ_{kj} measures the strength of the association between component k and product j. Products that characterize a component receive relatively large values of ϕ_{kj}, whereas products with weaker associations receive smaller values.

Assuming a countably infinite number of demand components a priori, let G_0 denote the global distribution over demand components. We set a Dirichlet Process prior (DP) over it such that

$$G_0 \sim \mathrm{DP}(\gamma, H). \tag{1}$$

The parameter γ governs how population-level probability mass is distributed across demand components. Smaller values of γ favor representations in which the mass is concentrated on relatively few components, whereas larger values allow it to be distributed more broadly. The base distribution H specifies the prior distribution over the component-specific product probabilities, $\boldsymbol{\phi}_k$. We set H to a symmetric Dirichlet distribution, $H = \mathrm{Dirichlet}(\eta, \ldots, \eta)$, where $\eta > 0$ controls the dispersion of product probabilities within each demand component.

Using the Sethuraman stick-breaking construction (Sethuraman, 1994), the population-level distribution over demand components is written as

$$G_0 = \sum_{k=1}^{\infty} \beta_k \delta_{\boldsymbol{\phi}_k}, \tag{2}$$

with

$$\beta_k = \beta'_k \prod_{\ell=1}^{k-1} (1 - \beta'_\ell), \qquad \beta'_k \sim \mathrm{Beta}(1, \gamma). \tag{3}$$

Here, $\delta_{\boldsymbol{\phi}_k}$ denotes a point mass at component $\boldsymbol{\phi}_k$, and the weights satisfy $\sum_{k=1}^{\infty} \beta_k = 1$.

Each customer i has a distinct distribution over the shared demand components. To ensure customers share the same demand components as the G_0, we set

$$G_i \sim \mathrm{DP}(\alpha, G_0) \tag{4}$$

which in turn can also be written as $G_i = \sum_{k=1}^{\infty} \theta_{ik}\delta_{\phi_k}$, where the weights satisfy $\sum_{k=1}^{\infty} \theta_{ik} = 1$. The k-th element of $\boldsymbol{\theta}_i = \{\theta_{ik}\}_{k=1}^{\infty}$ denotes the importance of component k in customer i's purchasing behavior. Customers may place substantial probability on multiple components and are therefore represented through mixed membership rather than assigned to a single segment. The parameter α governs how customer-level probability mass is distributed across demand components.

For each pre-treatment purchase instance ℓ made by customer i, a latent component assignment $\zeta_{i\ell}$ is first drawn according to the customer-specific weights:

$$P(\zeta_{i\ell} = k \mid \boldsymbol{\theta}_i) = \theta_{ik}. \tag{5}$$

Conditional on the selected component, the purchased product $v_{i\ell}$ is drawn from the corresponding component-specific distribution:

$$P(v_{i\ell} = j \mid \zeta_{i\ell} = k, \boldsymbol{\phi}_k) = \phi_{kj}. \tag{6}$$

This construction allows the same product to contribute to multiple demand components and permits customers to combine the shared components in different proportions.

Figure 1 provides a schematic representation of the HDP. The figure illustrates both the hierarchical distribution over latent demand components and the generative process through which consumers' observed purchases arise.

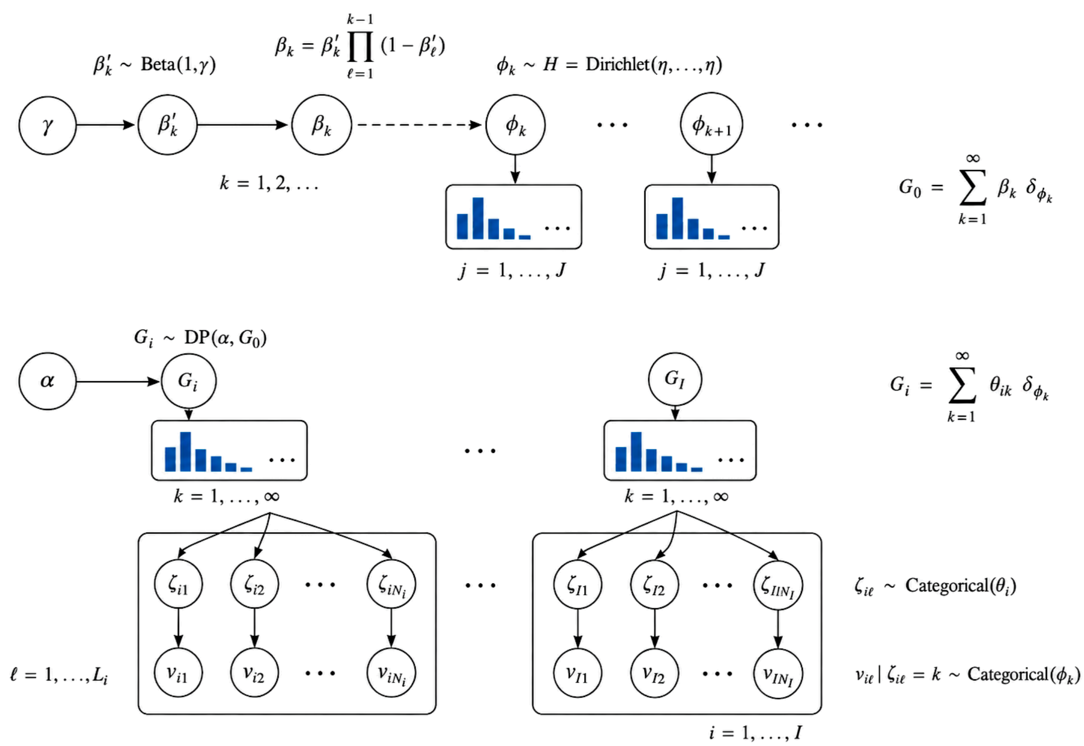

Figure 1: Illustrative representation of the HDP model for demand components

4.2 Product-Component Association

Based on the conditional independence implied by the HDP model, the probability that product j is associated with component k for customer i does not depend on the specific purchase instance ℓ and is given by

$$P(z_{ij} = k \mid \boldsymbol{\theta}_i, \boldsymbol{\phi}) = P(\zeta_{i\ell} = k \mid v_{i\ell} = j, \boldsymbol{\theta}_i, \boldsymbol{\phi}) = \frac{\theta_{ik} \cdot \phi_{kj}}{\sum_{k'} \theta_{ik'} \cdot \phi_{k'j}}, \tag{7}$$

where z_{ij} denotes the latent demand component associated with product j for customer i. Unlike $\zeta_{i\ell}$, which indexes the component assignment of a specific purchase instance ℓ, z_{ij} summarizes the customer-specific posterior association between product j and component k.

Having introduced the framework for modeling latent demand components and characterized the resulting product–component association probabilities, we next turn to causal inference for component-level treatment effects.

5 Causal Analysis Using Demand Components

Our goal is to jointly identify demand components and determine the causal impact of the broad-scope coupon at the component level. Let $y_{i,j}$ denote the outcome (i.e., spending) for product j purchased by customer i after treatment. If no purchases of product j are made, then $y_{i,j} = 0$. Let y_i^a denote the spending for customer i which can be expressed as:

$$y_i^a = \sum_{j=1}^{J} y_{i,j}.$$

Additionally, let $y_{i,k}^c$ denote the spend associated with the k-th component. Since the demand component allocations z_{ij} are probabilistic, we use the probabilities given by Equation 7 to construct the component outcomes $y_{i,k}^c$. Consequently, the component outcomes are obtained by integrating over z_{ij} using:

$$y_{i,k}^c = \sum_{j=1}^{J} \left[\int_{z_{ij}} \mathbb{I}_k(z_{ij}) dF_{\boldsymbol{\theta}_i, \boldsymbol{\phi}}(z_{ij}) \right] y_{i,j} = \sum_{j=1}^{J} \Pr(z_{ij} = k \mid \boldsymbol{\theta}_i, \boldsymbol{\phi})\, y_{i,j}, \tag{8}$$

where $\mathbb{I}_k$ is an indicator function that returns one if $z_{ij} = k$, zero otherwise.

Under this formulation, y_i^a the total spending of customer i across all products, can be expressed as

$$y_i^a = \sum_{k=1}^{\infty} y_{i,k}^c, \tag{9}$$

where the relationship between component-level spending and total spending can be confirmed by expanding $y_{i,k}^c$ such that

$$y_i^a = \sum_{k=1}^{\infty} y_{i,k}^c = \sum_{k=1}^{\infty} \sum_{j=1}^{J} \Pr(z_{ij} = k \mid \boldsymbol{\theta}_i, \boldsymbol{\phi})\, y_{i,j} = \sum_{j=1}^{J} y_{i,j}. \tag{10}$$

5.1 Component-level Treatment Effects

Recall that the allocation probabilities $\Pr(z_{ij} = k \mid \boldsymbol{\theta}_i, \boldsymbol{\phi})$ are inferred exclusively from pre-treatment purchases and are therefore predetermined with respect to treatment assignment. Therefore, component-level outcomes are constructed by applying these pre-treatment probabilities to post-treatment product outcomes. Thus, randomization ensures that treated and control customers are comparable in expectation, allowing both aggregate and component-level treatment effects to be identified from observed outcomes.

The aggregate treatment effect, τ_1^a, and the treatment effect for demand component k, $\tau_{1,k}^c$, are identified as the differences in mean observed outcomes between the treatment and control groups:

$$\begin{cases} \tau_1^a = \mathbb{E}[y_i^a \mid w_i = 1] - \mathbb{E}[y_i^a \mid w_i = 0]\,, \\ \tau_{1,k}^c = \mathbb{E}\big[y_{i,k}^c \mid w_i = 1\big] - \mathbb{E}\big[y_{i,k}^c \mid w_i = 0\big]\,. \end{cases}$$

Here, w_i is the treatment indicator, equal to one if customer i is treated and zero otherwise. An important property of the proposed framework is that the decomposition preserves the experimentally identified aggregate treatment effect. Rather than changing the causal effect identified by the experiment, it simply distributes that effect across the underlying demand components. Consequently, the aggregate treatment effect is exactly recovered by summing the component-level treatment effects, as by construction,

$$\begin{aligned} \tau_1^a &= \mathbb{E}[y_i^a \mid w_i = 1] - \mathbb{E}[y_i^a \mid w_i = 0] \\ &= \sum_{k=1}^{\infty} \Big[\mathbb{E}\big[y_{i,k}^c \mid w_i = 1\big] - \mathbb{E}\big[y_{i,k}^c \mid w_i = 0\big]\Big] \\ &= \sum_{k=1}^{\infty} \tau_{1,k}^c. \end{aligned}$$

6 Model Estimation

Estimating the model raises several considerations. First, inferring an unbounded number of demand components may seem conceptually difficult, but this is not a practical concern for estimation: only a finite number K emerges from the HDP for the causal analysis (Boughanmi and Ansari, 2021). Two genuine challenges remain, however: propagating uncertainty in the latent components into the causal estimates, and estimating heterogeneity in component-level treatment effects. We address both using a two-stage Bayesian procedure with an intentionally unidirectional flow of information, from the component model to the causal model (Plummer, 2015). The first stage estimates the latent demand representation exclusively from pre-treatment purchases, preventing treatment assignment and post-treatment outcomes from influencing component formation (Angrist and Pischke, 2009; Imbens and Rubin, 2015; Rosenbaum, 1984). The second stage uses each retained posterior draw of the latent representation to reconstruct component-level outcomes and pre-treatment component spending, after which a collection of component-specific Bayesian causal regressions is estimated conditional on that draw. Repeating this procedure across posterior draws propagates uncertainty from the demand model into the causal estimates while preserving this unidirectional flow (Murphy and Topel, 2002).

Treatment Effects Estimation The treatment effect for component k is estimated using the following:

$$y^c_{i,k} = \tau^c_{0,k} + \tau^c_{1,k} w_i + \varepsilon^c_{i,k}, \tag{11}$$

where $\tau^c_{0,k}$ is the expected component-level outcome for the control group and $\tau^c_{1,k}$ is the component-level average treatment effect. The error term is assumed to follow $\varepsilon^c_{i,k} \sim \mathcal{N}\big(0, (\sigma^c_k)^2\big)$.

Equation 11 can be extended to allow treatment effects to vary across customers. Let $\boldsymbol{x}_{1,i}$ denote observed pre-treatment customer characteristics (e.g., frequency of purchases) and let $\boldsymbol{x}_{2,i}$ denote customer i's pre-treatment spending across the K demand components. Note that the latter is inferred based on model estimates. Specifically, the k-th element of $\boldsymbol{x}_{2,i}$ is given by

$$x_{2,i,k} = \sum_{j=1}^{J} \Pr(z_{ij} = k \mid \boldsymbol{\theta}_i, \boldsymbol{\phi})\, s^{\text{pre}}_{ij}, \tag{12}$$

where s^{pre}_{ij} is customer i's pre-treatment spending on product j. We combine these two sources of heterogeneity in a single vector $\boldsymbol{x}_i = \big(\boldsymbol{x}^\top_{1,i}, \boldsymbol{x}^\top_{2,i}\big)^\top$.

For each component k, we estimate the following specification:

$$y^c_{i,k} = \tau^c_{0,k} + \boldsymbol{\tau}^{c,\text{het}\top}_{0,k} \boldsymbol{x}_i + \left(\tau^c_{1,k} + \boldsymbol{\tau}^{c,\text{het}\top}_{1,k} \boldsymbol{x}_i\right) w_i + \varepsilon^{c,\text{het}}_{i,k}, \tag{13}$$

with $\varepsilon^{c,\text{het}}_{i,k} \sim \mathcal{N}(0, (\sigma^{c,\text{het}}_k)^2)$. The implied customer-specific component-level treatment effect is given by

$$\tau^{c,\text{het}}_{1,i,k} = \tau^c_{1,k} + \boldsymbol{\tau}^{c,\text{het}\top}_{1,k} \boldsymbol{x}_i. \tag{14}$$

This specification allows component-level treatment responses to vary with observed customer characteristics and pre-promotion demand spending on components.

Uncertainty Propagation Using the direct-assignment Gibbs sampler (Heinrich, 2011; Teh et al., 2004), let n^θ_{ik} denote the number of customer i's purchases assigned to component k, and let n^ϕ_{kj} denote the number of purchases of product j assigned to component k. Conditional on these assignments, the customer mixture weights and component-specific product distributions follow the conjugate Dirichlet full conditionals

$$\boldsymbol{\theta}_i \mid \boldsymbol{\zeta}, \alpha, \boldsymbol{\beta} \sim \text{Dirichlet}\left(n^\theta_{i1} + \alpha\beta_1, \ldots, n^\theta_{iK} + \alpha\beta_K\right), \tag{15}$$

where $\boldsymbol{\beta}$ denotes the global demand-component weights. Similarly,

$$\boldsymbol{\phi}_k \mid \boldsymbol{\zeta}, \eta \sim \text{Dirichlet}\left(n^\phi_{k1} + \eta, \ldots, n^\phi_{kJ} + \eta\right). \tag{16}$$

These posterior distributions quantify uncertainty in customers' demand-component mixtures and in the product composition of each component. For posterior draw d, we calculate the customer-specific product–component association probability

$$\Pr\left(z_{ij} = k \mid \boldsymbol{\theta}_i^{(d)}, \boldsymbol{\phi}^{(d)}\right) = \frac{\theta_{ik}^{(d)} \cdot \phi_{kj}^{(d)}}{\sum_{k'=1}^{K} \theta_{ik'}^{(d)} \cdot \phi_{k'j}^{(d)}}. \tag{17}$$

Consequently, these probabilities are used to construct the post-treatment component outcomes in Equation 8 and the pre-treatment spending allocated to component k in Equation 12. Thus, component-level outcomes and pre-treatment spending are recomputed for every retained HDP draw, and leveraged as entries for the MCMC estimation of Equations 11 and 13. Averaging the resulting estimates across draws numerically integrates over posterior uncertainty in $\boldsymbol{\theta}_i$ and $\boldsymbol{\phi}$.

In sum, for each retained HDP draw, we reconstruct the component-level post-treatment outcomes and pre-treatment component-spending covariates using the corresponding draws of $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$. Conditional on these draws, we estimate a separate Bayesian regression for each demand component using Gibbs sampling (Gelman et al., 1995). Each component has its own coefficient vector and residual variance with posterior summaries over the resulting draws from both stages.

Identification. Identification relies on both the latent demand model and the causal design. The HDP recovers demand components from systematic variation in customers' product co-purchases: products that repeatedly appear together across customers are assigned to common latent components, whereas distinct co-purchase patterns generate separate components. As with finite-mixture models, the components are identified only up to label permutations (Stephens, 2000). To maintain consistent component identities across posterior draws, we align components based on the similarity of their specific product distributions while estimating the treatment effects. Simulation evidence in Boughanmi and Ansari (2021) further shows that the HDP can recover the number and composition of the underlying components under conditions similar to those considered here.

Treatment effects are identified by random assignment, which makes treatment independent of both observed customer characteristics and pre-treatment latent demand. Heterogeneous effects are identified from treatment interactions with pre-treatment characteristics, provided that these characteristics exhibit sufficient variation and are not perfectly collinear. Because the demand components are estimated exclusively from pre-treatment purchases, their construction cannot be influenced by treatment assignment or post-treatment outcomes.

We fit the HDP model on the entire pre-treatment data. The sampler is run for $10{,}000$ MCMC iterations, with the first $8{,}000$ iterations discarded as burn-in. At the end of burn-in and following Boughanmi and Ansari (2021), the active set of demand components is fixed for downstream causal estimation. We then use the $2{,}000$ retained posterior draws to propagate uncertainty from the component model into the component-level treatment-effect estimates. The final $1{,}000$ posterior draws are then used for treatment-effect reporting. Details on the prior and hyperparameter values as well as convergence checks are provided in Web Appendix A.

Next, we describe the demand components identified by the HDP model and then provide a more detailed analysis of the causal effects of the broad coupon at the component level.

7 Demand-Component Results

Table 3 summarizes the inferred components using anonymized generic product descriptions and average price.[3] The components capture both narrowly defined beauty-care routines, such as Intensive Hydration (C1) and Routine Mask Care (C2), and broader regions of the assortment spanning hair, body, oral, nail, and men's care. They therefore organize demand not only according to conventional product categories, but also according to consumption routines, complementary use, specialization, and price tier.

Table 3: Latent Demand Components and Representative Products

Nb.	Name	Representative Generic Product Types	Avg. Price
C1	Intensive Hydration	Energy masks; hydrogel masks; watery moisturizing creams; facial peeling products; hydrating skin water; ...	\$5.87
C2	Routine Mask Care	Essence sheet masks; flower essence masks; lotus moisture masks; skin-fit masks; camellia-oil nourishing masks; ...	\$6.14
C3	Daily Beauty Routine	Cushion puffs; specialized sun-cushion puffs; rose-water toner; air-cushion compacts; creamy tinted lip color; ...	\$16.85
C4	Household Personal Care	Hair-loss prevention shampoo; tartar-control toothpaste; daily-care toothpaste; fresh toothpaste; lavender bar soap; ...	\$12.55
C5	Corrective Treatment	Bio-essence facial masks; blackhead nose patches; men's bio-essence; men's anti-aging emulsion; oil-blotting paper; ...	\$19.48
C6	Beauty & Wellness	UV mist cushions; diet-support supplements; BB cushions; calming botanical skin water; collagen cushion compacts; ...	\$31.99
C7	Skincare Maintenance	Botanical skin water; watery moisturizing creams; hydrating emulsions; facial peeling products; air-cushion compacts; ...	\$31.05
C8	Hair & Body Repair	Hair-repair masks; essence pouch masks; oil-control films; repair shampoos; brightening facial cleansers; ...	\$16.55
C9	Appearance Renewal	Nail color; gray-coverage hair dye; shine-enhancing hair essence; herbal gray-coverage hair dye; permanent hair color; ...	\$23.27
C10	Everyday Care	Green-tea toothpaste; rich hand cream; lip gloss; moisturizing face cream; false eyelashes; ...	\$24.36
C11	Complete Moisture Care	Multi-cushion compacts; morning cleansers; watery moisturizing creams; intensive moisturizing creams; anti-aging cleansing oils; ...	\$24.54
C12	Men's Grooming	Men's skin toner; men's emulsion; men's skincare sets; floral hair spray; men's black skin toner; ...	\$26.51

Average prices also vary considerably across components.[4] The two mask-oriented components, Intensive Hydration (C1) and Routine Mask Care (C2), are the least expensive, with average prices

[3] Component names are assigned by inspecting the highest-weight products in each component-specific distribution, ϕ_k. To comply with the non-disclosure agreement, product names are anonymized and reported as generic product descriptions.

[4] Average price is the component-probability-weighted mean price across products.

ranging from approximately $\$5.87$ to $\$6.14$. Household Personal Care (C4), Hair & Body Repair (C8), and Daily Beauty Routine (C3) occupy intermediate price tiers of approximately $\$12.55$ to $\$16.85$. The highest-priced components, Beauty & Wellness (C6) and Skincare Maintenance (C7), average approximately $\$31$ to $\$32$. Overall, the latent demand components summarize the assortment along managerially meaningful dimensions of function, consumption routine, breadth, specialization, and price.

7.1 Demand-Component Weights

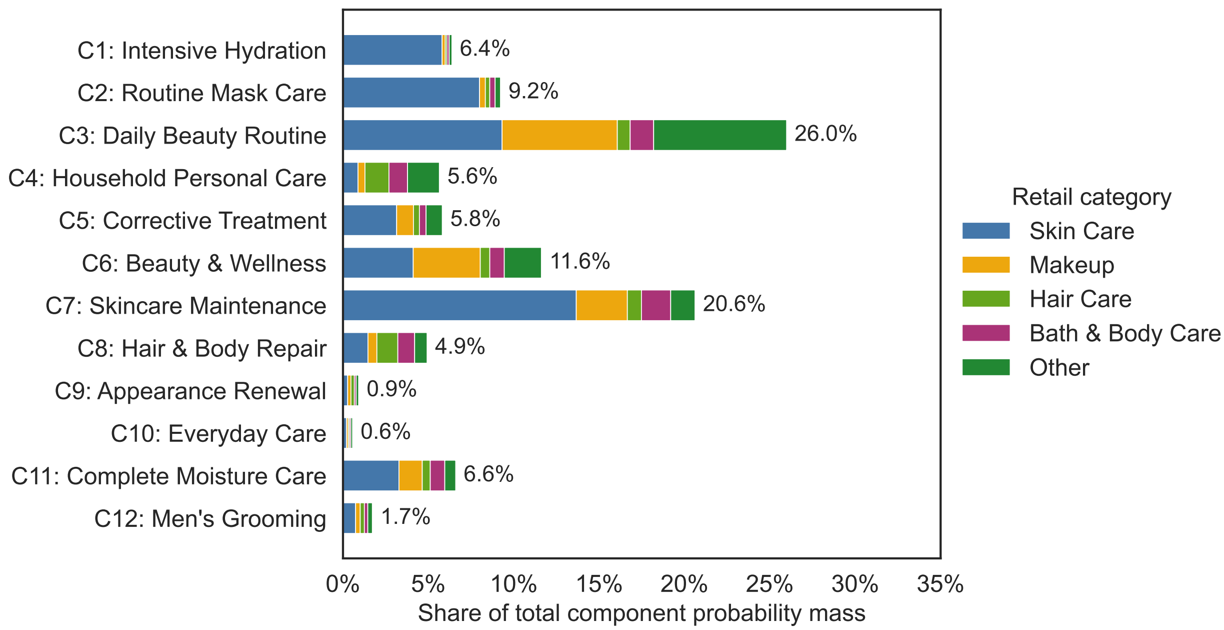

Figure 2: Distribution and Retail-Category Composition of the HDP Demand Components

Figure 2 displays the posterior mean global weights of the twelve latent demand components ($\boldsymbol{\beta}$). Pre-treatment purchasing is concentrated in a small number of recurring routines. Daily Beauty Routine (C3) receives the largest global weight, accounting for approximately 26.0% of the population-level component mass, followed by Skincare Maintenance (C7), which accounts for approximately 20.6%. Together, these two components represent approximately 46.6% of the global component mass. Beauty & Wellness (C6) and Routine Mask Care (C2) form a second tier, with weights of approximately 11.6% and 9.2%, respectively. In contrast, Appearance Renewal (C9), Everyday Care (C10), and Men's Grooming (C12) receive relatively small global weights.

Figure 2 also shows the mapping between categories and components.[5] It is clear that the latent demand components differ from retailer-defined categories as the recovered purchasing routines only partially overlap with the retailer's administrative taxonomy. Intensive Hydration (C1) and Routine Mask Care (C2) are both dominated by Skin Care products, with shares of 91.0% and 86.7%, respectively, yet they capture distinct demand structures. Intensive Hydration (C1) is anchored

[5] For each component, we aggregate posterior mean product probabilities by retailer category, normalize within component, and scale the resulting shares by the component's posterior mean global weight, β_k.

by a narrow set of concentrated hydration products, whereas Routine Mask Care (C2) represents a broader recurring mask-based routine. Skincare Maintenance (C7) is also primarily composed of Skin Care products, with a share of 66.2%, but includes meaningful shares of Makeup and Bath & Body Care products, at 14.5% and 8.3%, respectively. These patterns show that multiple behaviorally distinct demand components can emerge within the same retailer-defined category, while individual components may also combine products from several administrative categories.

Our framework allows customers to exhibit heterogeneous weights across demand components. Web Appendix B provides further details on this heterogeneity by examining the distribution of customers' component weights ($\boldsymbol{\theta}_i$).

7.2 Overlapping Product Roles Across Demand Components

Figure 3 shows that our inferred representation does not impose a hard partition of products into mutually exclusive groups. Instead, products can be associated with multiple components, with the allocation depending on both the product and the customer's mixture. Using a threshold of average $\Pr(z_{ij} = k \mid \hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\phi}}) > 0.05$, the median product is associated with 3 components, with some products associated with as many as 8 components. This overlap is substantively important

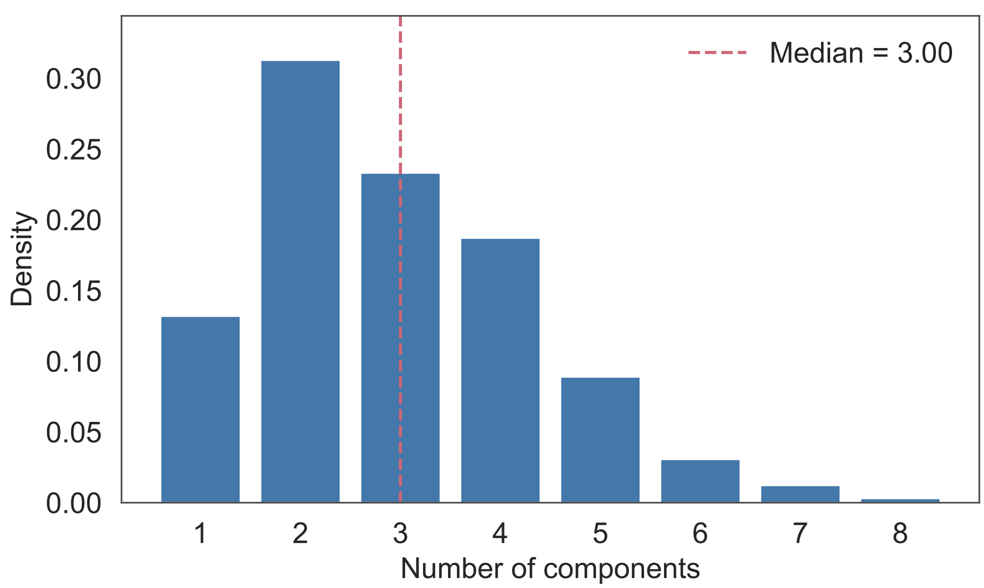

Figure 3: Distribution of the Number of Components Associated with Each Product

in the retail setting because many personal-care products can serve multiple demand routines. For example, cushion products contribute to both Daily Beauty Routine (C3) and Complete Moisture Care (C11), while cleansing and treatment products create overlap among Complete Moisture Care (C11), Hair & Body Repair (C8), and Corrective Treatment (C5). The latent representation captures these overlapping roles directly, allowing the causal decomposition to measure how promotions shift interconnected purchasing routines rather than forcing each product into a single mutually exclusive category.

7.3 Model Validation

The proposed causal decomposition can, in principle, be applied to any partition of product demand, including an arbitrary partition. Thus, its usefulness depends on whether the retained structure provides a valid and interpretable representation of customers' purchasing behavior. Although the primary purpose of the HDP is to reveal the latent demand structure underlying the effects of broad-scope promotions, the resulting managerial insights are credible only if the model adequately captures observed purchasing patterns. Predictive validity thus complements interpretability as an important criterion for evaluating the decomposition.

We validate the HDP model by comparing its out-of-sample predictive performance with that of three informative benchmarks: a naive uniform model that assigns equal probability to every product, a global product-popularity model that predicts purchases using aggregate product frequencies, and an observed-category model that combines customers' category preferences with product popularity within the retailer's predefined categories.

As noted earlier, to preserve the validity of the causal analysis, demand components must be estimated without using any post-treatment outcomes. Thus, we perform model estimation and validation using only pre-treatment purchase data. We split each customer's pre-treatment purchases into 80% training and 20% test samples, retaining at least one purchase in the training set for every customer. All models are estimated on the same training data and evaluated on the same held-out purchases.

Table 4: Out-of-Sample Predictive Performance

Model	log-PD	Recall@20	NDCG@20
Naive Uniform	-7.356	0.013	0.006
Global Product Popularity	-5.455	0.223	0.128
Observed-Category Popularity	-5.138	0.265	0.163
HDP	-4.648	0.344	0.219

The predictive-validation results support the empirical relevance of the latent demand structure recovered by the HDP. As shown in Table 4, on the held-out pre-treatment purchases, the HDP achieves the highest average log predictive density (log-PD) among the models considered. It also provides the strongest ranking performance, with Recall@20 of 0.344 and NDCG@20 of 0.219.[6] The improvement over the observed-category model is notable as it indicates that the demand components capture systematic purchasing relationships that are not adequately represented by the retailer's predefined product categories.

[6]Recall@20 measures the proportion of products a customer actually purchases that appear among the model's top 20 recommendations. NDCG@20, or Normalized Discounted Cumulative Gain at 20, additionally accounts for ranking by assigning greater weight to purchased products that appear closer to the top of the recommendation list.

8 Demand-Component Treatment Effects

We now turn to the causal results at the demand-component level. We first assess the causal validity of the component-level decomposition and then examine the estimated treatment effects.

8.1 Pre-Treatment Balance and Causal Validity

Figure 4 compares the control and treatment groups for observed customer characteristics and the twelve inferred demand-component memberships ($\boldsymbol{\theta}_i$).[7]

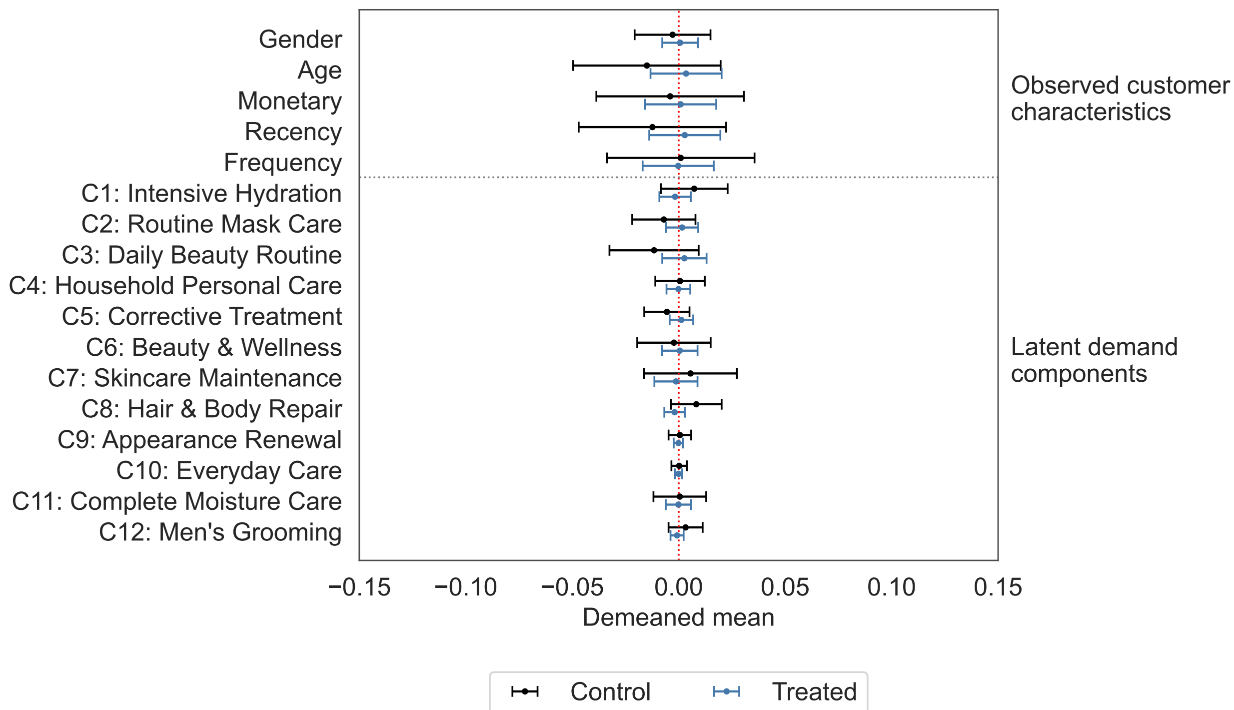

Figure 4: Pre-Treatment Balance Across Experimental Groups

The groups are well balanced: none of the observed characteristics differs significantly across conditions, and component-membership differences range from -0.010 to 0.014, with all p-values above 0.12. Because the components are learned exclusively from pre-treatment behavior and are balanced across randomized treatment, the subsequent decomposition preserves the causal interpretation of the original experiment. Estimating the HDP using post-treatment purchases introduces post-treatment bias and invalidates the causal interpretation of the latent demand components. Evidence is provided in Web Appendix D.

[7] We consider five observed customer characteristics: (1) age, (2) gender, (3) recency, (4) frequency, and (5) monetary. Gender is coded as a binary indicator (1 if female), while the remaining variables are dichotomized based on median splits to enable consistent interpretation. For recency, customers with more recent interactions are positively coded.

8.2 Component-level Treatment Effects

Table 5 reports the posterior treatment effect associated with each demand component.[8] Because the decomposition is additive, the component-level treatment effects sum to the aggregate treatment effect. Accordingly, the sum of the posterior means reported in the table closely matches the aggregate estimate, with a small relative discrepancy of approximately 1.5% attributable to Monte Carlo error.

Table 5: Bayesian Posterior Treatment-Effect Summaries

Nb.	Name	Post. Mean	Post. SD	p_{Bayes}
C1	Intensive Hydration	0.127	0.101	0.198
C2	Routine Mask Care	0.140	0.089	0.122
C3	Daily Beauty Routine	0.256	0.178	0.152
C4	Household Personal Care	0.071	0.057	0.202
C5	Corrective Treatment	0.060	0.090	0.504
C6	Beauty & Wellness	0.421	0.179	0.020
C7	Skincare Maintenance	0.670	0.257	0.006
C8	Hair & Body Repair	0.027	0.086	0.728
C9	Appearance Renewal	0.027	0.048	0.458
C10	Everyday Care	0.004	0.043	0.964
C11	Complete Moisture Care	0.161	0.091	0.070
C12	Men's Grooming	−0.051	0.079	0.466

The largest treatment effects arise in three demand components centered on skincare maintenance and premium beauty routines. Skincare Maintenance (C7) exhibits the largest posterior mean effect (0.670, $p_{\text{Bayes}} = 0.006$), followed by Beauty & Wellness (C6; 0.421, $p_{\text{Bayes}} = 0.020$). Complete Moisture Care (C11) also displays a positive effect (0.161, $p_{\text{Bayes}} = 0.070$), although the posterior evidence is weaker.

The treatment effect is substantially concentrated across demand components. The two components with the strongest posterior evidence, Beauty & Wellness (C6) and Skincare Maintenance (C7), account for 56.2% of the aggregate treatment effect while receiving only 32.2% of the posterior mean global component mass, as measured by β_k. Including Complete Moisture Care (C11) raises these shares to 64.4% of the aggregate treatment effect and 38.8% of the global component mass. Thus, although the coupon applies broadly across the assortment, most of its incremental impact is transmitted through a relatively small set of latent purchasing routines.

To assess whether the modeling framework mechanically generates nonzero treatment effects even in the absence of genuine treatment variation, we randomly reassign treatment status and re-estimate the aggregate and component-level effects. Web Appendix E shows that none of the resulting

[8] We report the two-sided posterior probability $p_{\text{Bayes}} = 2\min\{\Pr(\tau > 0 \mid \mathcal{D}), \Pr(\tau < 0 \mid \mathcal{D})\}$, which equals twice the posterior mass in the direction opposite to the predominant sign of the effect. Smaller values indicate stronger posterior evidence that the effect is either positive or negative (Shi and Yin, 2021).

placebo estimates is statistically distinguishable from zero.

8.3 Heterogeneity in Demand-Component Treatment Effects

Figure 5 examines how the promotion's effect within each demand component varies with customers' observed characteristics and pre-treatment purchasing patterns. Each row corresponds to a post-treatment demand component, while the columns report heterogeneity coefficients for gender, age, monetary value, recency, and frequency, followed by pre-treatment spending in each of the twelve demand components. The values in parentheses are the corresponding p_{Bayes} values.

The observed customer characteristics provide little evidence of systematic treatment-effect heterogeneity. None of the coefficients for gender, age, monetary value, recency, or frequency is statistically significant at the 5% level. Thus, commonly available customer descriptors do not identify which demand components are most responsive to the promotion. This result illustrates a limitation of relying exclusively on demographic and aggregate behavioral measures: customers with similar observed characteristics may nevertheless differ substantially in the demand routines through which they respond.

In contrast, pre-treatment demand-component spending reveals a pronounced component-specific pattern. All seven statistically significant coefficients among the component-spending moderators lie on the diagonal of the matrix. Customers who previously spent more in Intensive Hydration (C1), Routine Mask Care (C2), Household Personal Care (C4), Beauty & Wellness (C6), Skincare Maintenance (C7), Appearance Renewal (C9), or Complete Moisture Care (C11) exhibit significantly larger promotional effects within the corresponding component. The diagonal coefficients range from 0.70 for Household Personal Care to 2.37 for Appearance Renewal. Corrective Treatment (C5) displays a similar positive own-component relationship, although it is significant only at the 10% level.

To verify that the estimated heterogeneity is not mechanically generated by the modeling framework, we repeat the analysis after randomly reassigning treatment. Web Appendix E shows that the resulting placebo estimates exhibit no systematic heterogeneity across observed customer characteristics or pre-treatment component spending.

The absence of corresponding off-diagonal relationships indicates that this heterogeneity is primarily routine-specific rather than broadly transferable across components. For example, prior spending in Appearance Renewal predicts a stronger response within Appearance Renewal, but not systematically within the other purchasing routines. Overall, these findings suggest that the promotion primarily intensifies demand within customers' established consumption routines rather than generating uniform responses based on general customer characteristics or shifting demand indiscriminately across routines.

The preceding analysis demonstrates that coupon responses are concentrated within customers' existing demand components. An important remaining question is whether differences in customers' aggregate responsiveness reflect uniformly stronger effects across all demand components or instead arise from disproportionately large responses within a limited set of purchasing routines. To address

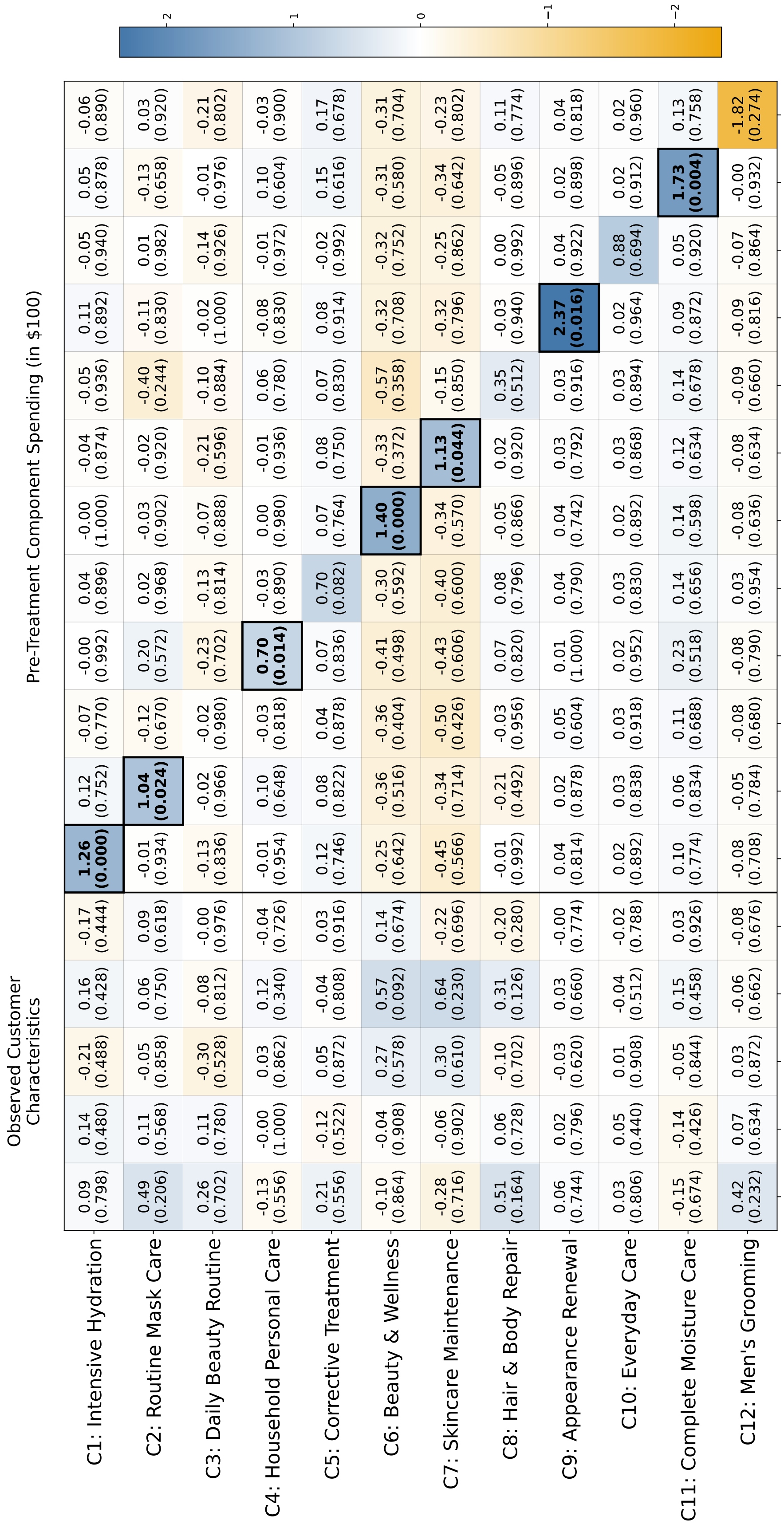

Figure 5: Component-Level Treatment-Effect Heterogeneity

this question, we decompose predicted aggregate treatment effects jointly across customers and demand components.

8.4 Concentration of Treatment Effects Across Customers and Demand Components

To better understand the behavioral sources of customer heterogeneity, we conduct an additional analysis. We rank customers according to their predicted aggregate treatment effect[9] and divide them into ten equally sized groups, ranging from the lowest-response group ($G1$) to the highest-response group ($G10$).

Figure 6 reports the posterior mean treatment effect for each demand component within each group. The figure shows that customers with larger aggregate treatment effects are not uniformly more responsive across the assortment. Instead, their additional response is concentrated primarily in Beauty & Wellness (C6) and Skincare Maintenance (C7). The average effect for Beauty & Wellness increases from -0.03 in $G1$ to 1.07 in $G10$, while the effect for Skincare Maintenance rises from -0.01 to 1.06, reaching 1.15 in $G9$. Complete Moisture Care (C11) provides a secondary source of heterogeneity, increasing from 0.04 in $G1$ to 0.50 in $G10$. Intensive Hydration (C1) and Routine Mask Care (C2) exhibit smaller but similarly increasing gradients.

Other components contribute little to the separation between low- and high-response customers. Effects for Appearance Renewal (C9) and Everyday Care (C10) remain close to zero across most groups. Daily Beauty Routine (C3) exhibits the opposite pattern: its effect is largest among lower-response customers and declines from 0.24 in $G1$ to 0.05 in $G10$. Thus, customers with stronger aggregate responses are not uniformly more responsive across all components. Their incremental response is concentrated in a specific set of skincare, wellness, and moisture-care routines.

The figure further reveals that promotional responses are concentrated along two dimensions: across customers and across demand components. The top 40% of customers (groups $G7$–$G10$) account for approximately 60.5% of the total treatment effect. Within these high-response groups, Beauty & Wellness and Skincare Maintenance account for 25.9% and 34.5% of the treatment effect, respectively, jointly explaining approximately 60.4% of the response generated by these customers. Thus, a disproportionate share of the aggregate lift is generated by a relatively small subset of customers and transmitted through a limited set of demand components.

9 From Demand-Components to High-Response Products

Our proposed framework provides a direct basis for product-level management decisions during broad-scope promotions. Decomposing the aggregate treatment effect across latent demand components makes it possible to approximate the incremental demand back to individual products, allowing retailers to prioritize the products most responsible for the promotional lift.

[9] The additivity of the decomposition implies that customer i's aggregate treatment effect is obtained by summing the predicted treatment effects across demand components: $\widehat{\tau}^{a}_{1,i} = \sum_{k=1}^{K} \widehat{\tau}^{c}_{1,i,k}$.

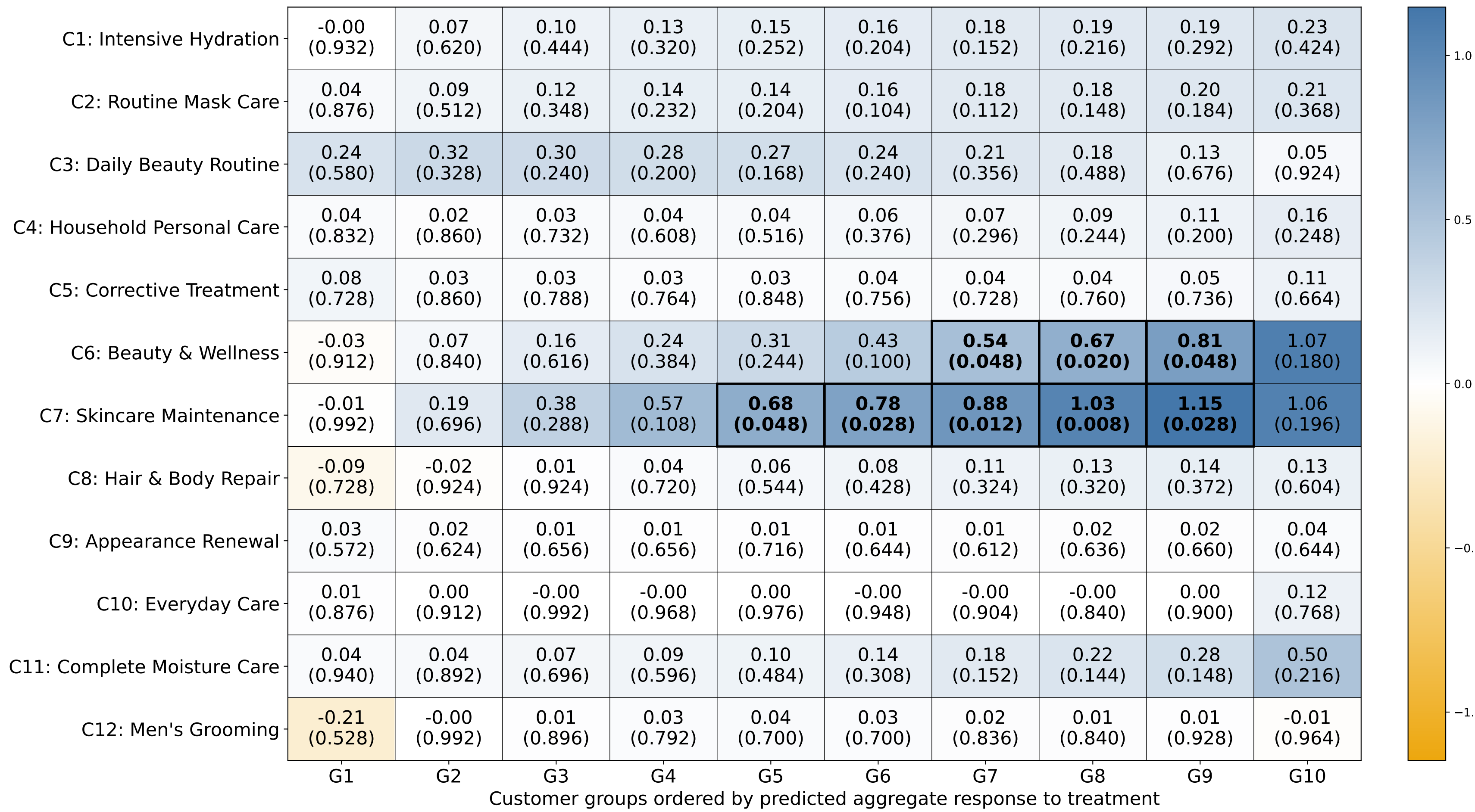

Figure 6: Treatment Effects Across Customer Groups Ordered by Predicted Aggregate Response

9.1 Product-Level Attribution of Incremental Demand

We assume that customer i's treatment effect for demand component k is allocated across products in proportion to the component's posterior product-association probabilities. Recall that our approach yields demand components as probability vectors over products $\boldsymbol{\phi}_k$, and consumer level treatment effects $\tau_{1,i,k}^{c,\text{het}}$. The model-implied treatment effect for consumer i for product j can be written as

$$\widehat{\tau}_{1,i,j}^{c} = \sum_{k=1}^{K} \mathbb{E}_d \left[\phi_{kj}^{(d)} \tau_{1,i,k}^{c,\text{het},(d)} \right], \tag{18}$$

where $\mathbb{E}_d$ denotes the expectation over the post-burn-in MCMC draws. This formulation allows a product to inherit incremental demand from multiple components, with larger allocations assigned to products that are strongly associated with each demand component among the retailer's customers. To illustrate, consider the limiting case in which demand component k assigns weight one to product j and zero to all other products. The component-level treatment effect for customer i is then attributed entirely to product j. Conversely, if product j receives zero weight in component k, none of that component's treatment effect is attributed to the product.

The implied average treatment effect for product j is then obtained by averaging the customer-specific product-level treatment effects:

$$\bar{\tau}_{1,j}^{c} = \frac{1}{I} \sum_{i=1}^{I} \widehat{\tau}_{1,i,j}^{c}. \tag{19}$$

This decomposition indicates which underlying customer demand patterns generate that effect. Figure 7 decomposes the allocated incremental demand of the ten highest-ranked products in the Skin Care and Makeup categories. Each stacked bar shows the contribution $I^{-1} \sum_i \mathbb{E}_d \left[\phi_{kj}^{(d)} \tau_{1,i,k}^{c,\text{het},(d)} \right]$ of component k to the total allocation for product j. Products within the same retailer category can receive similar total allocations for different reasons: some are primarily associated with a single highly responsive component, whereas others draw smaller contributions from several components.

Consequently, the decomposition can be used to quantify each product's contribution to the aggregate treatment effect. In particular, product j's implied share of the aggregate treatment effect is

$$\Delta_j^c = \frac{\bar{\tau}_{1,j}^{c}}{\hat{\tau}^a},$$

where $\hat{\tau}^a$ is the estimated aggregate treatment effect. As a contrast, one can also employ the category-level results to allocate treatment effects at the product-level during broad-scope promotions. This approach assigns each product the same share of the category-level effect as it does not distinguish among products within a category.

Thus, the two approaches allocate the same aggregate promotional response across products and differ only in how that response is distributed across the assortment. Figure 8 illustrates this comparison. A category-level approach produces a piecewise-constant share in which all products

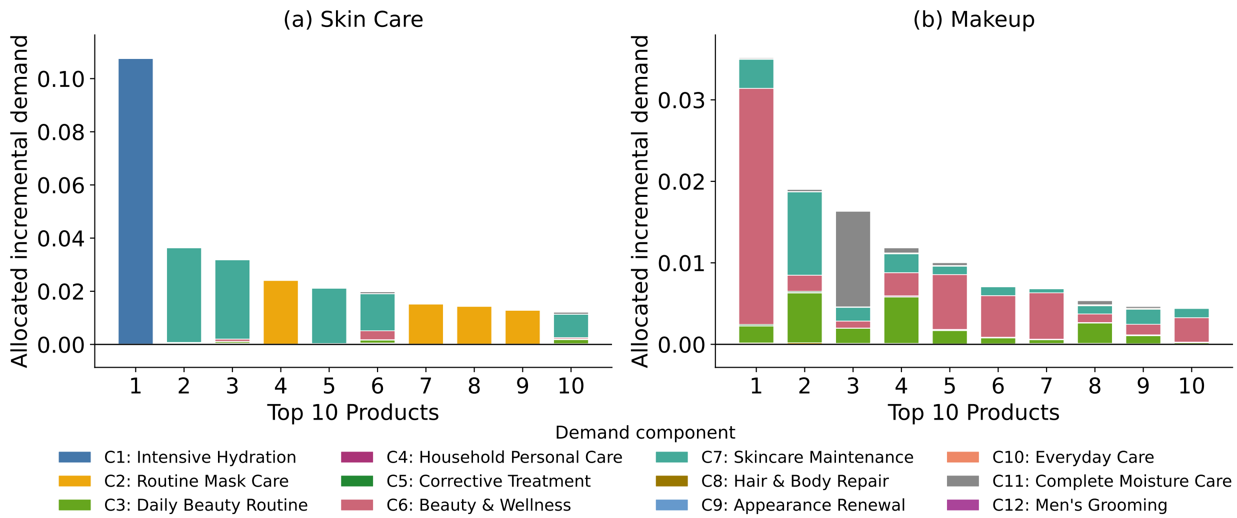

Figure 7: Implied Component Contributions for Top-Ranked Skin Care and Makeup Products

within a category receive the same treatment effect. For example, all products in the Skin Care category contribute 0.09% of the aggregate treatment effect. In contrast, our approach generates substantial within-category variation by incorporating each product's association with the demand components. Although it preserves the aggregate treatment effect, it produces a more concentrated product-level allocation. For example, the two highest-ranked products in the Skin Care category are implied to account jointly for more than 5% of the aggregate treatment effect.

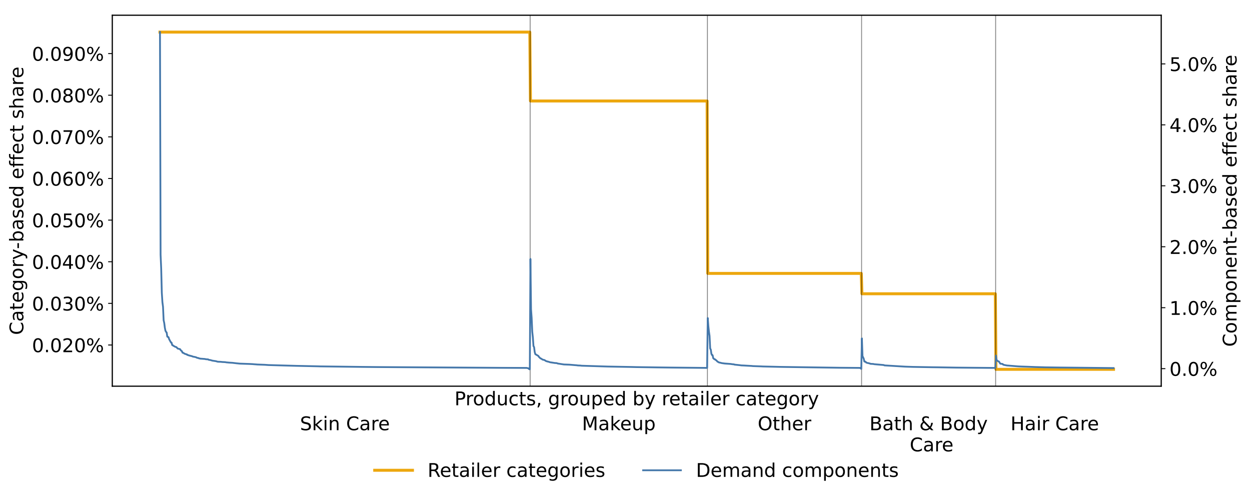

Figure 8: Product-Level Effect Allocation: Retailer-Category versus Demand-Component Approaches

9.2 Identifying High-Response Products

We conduct an assortment-removal exercise to test whether the approach identifies sets of products that account for a large share of the observed treatment effect.

We compare three product-removal strategies. The demand-component strategy ranks products by their component-implied treatment effects. The retailer-category strategy ranks categories by their estimated treatment effects and randomly orders products within each category. The random-removal benchmark averages outcomes across randomly generated product orderings. Note that we ignore consumer substitution that may follow an actual assortment change (Goli and Chintagunta, 2021). Instead, the analysis evaluates how effectively each ranking identifies products that account for the observed promotional response.

Figure 9 reports the *actual* aggregate treatment effect estimated on the remaining data as progressively larger portions of the assortment are removed. The benchmark is the treatment effect estimated using the full assortment, equal to $\$1.94$, as reported in Table 1. The vertical axis reports the remaining treatment effect in dollars, and the error bars represent 95% confidence intervals. A lower remaining treatment effect indicates that the corresponding product removal strategy is more effective in identifying the products responsible for the promotional gains.

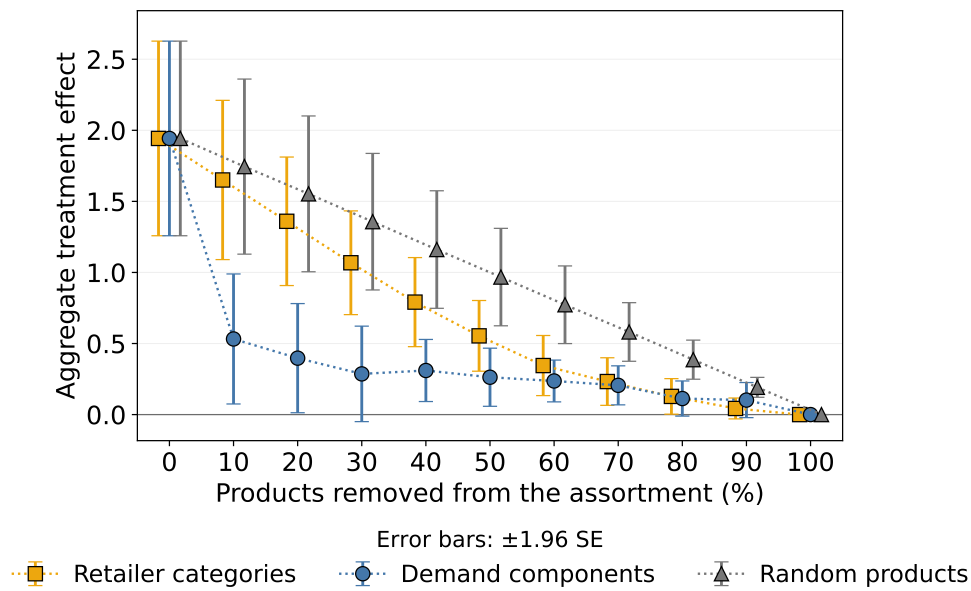

Figure 9: Aggregate Treatment Effect Remaining After Product Removal

The results reveal substantial differences in the ability of the three strategies to identify products that account for the promotional gains. Removing the top 10% of products identified by the demand-component analysis reduces the estimated aggregate treatment effect from $\$1.94$ to $\$0.53$, leaving only 27.4% of the original effect. In contrast, the retailer-category strategy leaves an effect of $\$1.65$, or 85.0% of the original effect, while random removal leaves $\$1.74$, or 89.8%. Thus, at a 10% removal rate, the retailer-category strategy leaves 57.6 percentage points more of the original promotional effect than the demand-component strategy. Although the retailer-category ranking performs modestly better than random removal, it fails to identify many of the individual products most responsible for the incremental response.

The advantage of the demand-component ranking remains economically meaningful over a broad range of removal rates. After removing 20%, 30%, and 40% of the assortment, the demand-component strategy leaves estimated treatment effects of $\$0.40$, $\$0.29$, and $\$0.31$, corresponding to 20.4%, 14.7%, and 16.0% of the original effect. The corresponding estimates under the retailer-category strategy are $\$1.36$, $\$1.07$, and $\$0.79$, or 70.0%, 55.0%, and 40.7% of the original effect. These differences amount to 49.6, 40.2, and 24.8 percentage points, respectively. Random removal performs considerably worse, leaving estimated effects of $\$1.55$, $\$1.36$, and $\$1.16$, corresponding to 79.9%, 69.9%, and 59.8% of the original effect. As increasingly large portions of the assortment are removed, the strategies converge because little incremental demand remains in the retained assortment.[10]

Overall, the results suggest that retailer-defined categories are too coarse to guide product-level decisions following a broad-scope promotion. The demand-component analysis identifies a relatively small subset of products that accounts for a large share of the observed promotional effect. The component-based approach additionally links these products to interpretable customer purchasing routines and ranks them by their implied incremental contribution. This provides a behaviorally informed and experimentally grounded basis for coordinating replenishment decisions for products that may span conventional retailer categories.

Table 6: Top 10% Component-Ranked Products versus Full Assortment

Metric	Top 10%	Full Assortment
Retailer-Category Distribution		
Skin Care (%)	61.78	38.83
Makeup (%)	19.11	18.58
Other (%)	10.83	16.16
Bath & Body Care (%)	4.46	14.05
Hair Care (%)	3.82	12.38
Demand-Component Distribution		
C1: Intensive Hydration (%)	2.86	2.15
C2: Routine Mask Care (%)	7.90	3.46
C3: Daily Beauty Routine (%)	13.33	6.34
C4: Household Personal Care (%)	5.42	7.53
C5: Corrective Treatment (%)	6.23	7.72
C6: Beauty & Wellness (%)	14.89	9.50
C7: Skincare Maintenance (%)	22.61	7.99
C8: Hair & Body Repair (%)	6.17	9.50
C9: Appearance Renewal (%)	3.95	11.87
C10: Everyday Care (%)	4.02	12.65
C11: Complete Moisture Care (%)	8.29	9.51
C12: Men's Grooming (%)	4.31	11.79
Pre-Treatment Product Characteristics		
Average product price	26.78	26.66
Nb. Customers who purchased per product	99.6	20.3
Purchase records per buyer	3.29	2.39
Repeat-purchase rate (%)	38.9	31.3
Share of pre-treatment sales (%)	57.7	100.0

[10] Small reversals at high removal rates arise because some products have negative estimated effects. Removing such products can mechanically increase the treatment effect remaining among the retained products.

9.3 Characteristics of High-Response Products

Table 6 characterizes the products with the top 10% largest component-implied treatment effects. These products are disproportionately concentrated in Skin Care, which accounts for 61.8% of the top decile compared with 38.8% of the full assortment. By contrast, Bath & Body Care and Hair Care are substantially underrepresented.

The component distribution provides a more granular account of this concentration.[11] In particular, Skincare Maintenance (C7) accounts for 22.6% of the component distribution among the top 10% products, compared with 8.0% in the full assortment. Daily Beauty Routine (C3), Beauty & Wellness (C6), and Routine Mask Care (C2) are also overrepresented. In contrast, more diffuse routines such as Appearance Renewal (C9), Everyday Care (C10), and Men's Grooming (C12) receive considerably smaller shares among the targeted products. Thus, the approach identifies a narrow set of products associated with recurring skincare and beauty-maintenance routines that is not fully captured by the retailer's broader category structure.

Furthermore, the differences are not explained by product price: average prices are nearly identical for the top decile and the full assortment, at $\$26.78$ and $\$26.66$, respectively. Instead, the targeted products differ primarily in the breadth and intensity of their pre-treatment demand. On average, each high impact product was purchased by more customers (99), compared with customers for a product in the full assortment (20). Buyers of these products also purchased them more frequently and were more likely to make repeat purchases. Although the targeted set contains only 10% of the assortment, it accounts for 57.7% of pre-treatment sales. Thus, these patterns indicate that the component approach identifies products embedded in recurring consumption routines.

10 Conclusion

Broad-scope promotions have become an increasingly common tool for stimulating consumer spending while preserving purchasing flexibility. Yet understanding these promotions requires more than determining whether they increase sales. It also requires understanding where within a large assortment those gains originate. As assortments continue to expand, this becomes not only a high-dimensional estimation problem but also a problem of organizing product-level responses into behaviorally meaningful patterns.

In this paper, we develop a Bayesian framework that decomposes demand into a small number of latent demand components and leverages randomized experimental variation to estimate promotional effects at a behaviorally meaningful level of resolution. Applying this framework to a large-scale field experiment reveals that broad-scope promotions produce responses that are substantially more concentrated than what conventional category-level analyses suggest. While aggregate- and category-level analyses imply that promotional gains are broadly distributed across the assortment,

[11] For each posterior draw, we first normalize the component–product probabilities across components within each product, $\widetilde{\phi}_{kj}^{(d)} = \phi_{kj}^{(d)} / \sum_{\ell=1}^{K} \phi_{\ell j}^{(d)}$, so that $\sum_k \widetilde{\phi}_{kj}^{(d)} = 1$. The reported share of component k is $J^{-1} \sum_j \mathbb{E}_d[\widetilde{\phi}_{kj}^{(d)}]$. We calculate this quantity separately for the top 10% products and for the full assortment; each resulting component distribution sums to one.

our demand decomposition shows that these gains are concentrated within a relatively small number of demand components and consistent with reinforcement of established routines.

From a managerial perspective, our findings suggest that retailers should rethink how broad-scope promotions are executed rather than narrowing their scope. Broad-scope promotions remain attractive because they are simple to communicate, easy for consumers to understand, and preserve purchasing flexibility. However, the execution of these promotions can be considerably more targeted. Because a relatively small number of demand components accounts for a disproportionate share of the promotional response, retailers can prioritize these underlying purchasing routines in merchandising, digital communications, product recommendations, and inventory planning without restricting the promotion itself. In other words, the promotion can remain broad while its execution becomes increasingly precise.

From a methods perspective, our results showcase that the level of aggregation at which causal effects are measured is not simply a matter of analytical convenience. Instead, it fundamentally shapes the conclusions that researchers and managers draw about how marketing interventions influence consumer behavior. Behaviorally derived demand components can reveal patterns that remain obscured under conventional product classifications, providing more information about where treatment effects arise.

Although our application focuses on broad-scope promotions, the underlying contribution extends more broadly. Many contemporary marketing interventions, including loyalty programs, personalized offers, and platform incentives, operate in environments characterized by large assortments and complex purchasing behavior. In these contexts, managerial conclusions may depend critically on the behavioral resolution at which treatment effects are measured. We hope this work encourages future research on causal inference in high-dimensional marketing environments and provides a foundation for studying a broader class of interventions in which understanding the distribution of treatment effects is as important as estimating their aggregate magnitude.

Disclosure: The authors declare that they have no financial or non-financial conflicts of interest related to the subject matter or materials presented in this manuscript. No external funding was received for this work. The authors used ChatGPT to assist with code development, theme-label refinement, and language editing. All empirical analyses were conducted and verified by the authors.

References

Ailawadi, Kusum L., Bari A. Harlam, Jacques César, and David Trounce (2007), "Practice Prize Report—Quantifying and Improving Promotion Effectiveness at CVS," *Marketing Science*, 26 (4), 566–575.

Anderson, Eric T. and Duncan I. Simester (2004), "Long-Run Effects of Promotion Depth on New versus Established Customers: Three Field Studies," *Marketing Science*, 23 (1), 4–20.

Angrist, Joshua D and Jörn-Steffen Pischke (2009), *Mostly Harmless Econometrics: An Empiricist's Companion* Princeton University Press.

Blattberg, Robert C, Richard Briesch, and Edward J Fox (1995), "How Promotions Work," *Marketing science*, 14 (3_supplement), G122–G132.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003), "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993–1022.

Boughanmi, Khaled and Asim Ansari (2021), "Dynamics of Musical Success: A Machine Learning Approach to Multimedia Data Fusion," *Journal of Marketing Research*, 58 (6), 1034–1057.

Broniarczyk, Susan M., Wayne D. Hoyer, and Leigh McAlister (1998), "Consumers' Perceptions of the Assortment Offered in a Grocery Category: The Impact of Item Reduction," *Journal of Marketing Research*, 35 (2), 166–176.

Gabel, Sebastian and Artem Timoshenko (2022), "Product Choice with Large Assortments: A Scalable Deep-Learning Model," *Management Science*, 68 (3), 1808–1827.

Gelman, Andrew, John B Carlin, Hal S Stern, and Donald B Rubin (1995), *Bayesian Data Analysis* Chapman and Hall/CRC.

Goli, Ali and Pradeep K. Chintagunta (2021), "What Happens When a Retailer Drops a Product Category? Investigating the Consequences of Ending Tobacco Sales," *Marketing Science*, 40 (6), 1169–1198.

Gopalakrishnan, Arun and Young-Hoon Park (2021), "The Impact of Coupons on the Visit-to-Purchase Funnel," *Marketing Science*, 40 (1), 48–61.

Heinrich, Gregor (2011), "Infinite LDA Implementing the HDP with Minimum Code Complexity," *Technical note, Feb*, 170.

Hui, Sam K. and Eric T. Bradlow (2012), "The Bayesian Customer-Level Multi-Category Latent Factor Model for Grocery Store Purchasing," *Quantitative Marketing and Economics*, 10 (2), 123–153.

Imbens, Guido W. and Donald B. Rubin (2015), *Causal Inference in Statistics, Social, and Biomedical Sciences* Cambridge University Press.

Jacobs, Bruno, Dennis Fok, and Bas Donkers (2021), "Understanding Large-Scale Dynamic Purchase Behavior," *Marketing Science*, 40 (5), 844–870.

Jacobs, Bruno JD, Bas Donkers, and Dennis Fok (2016), "Model-Based Purchase Predictions for Large Assortments," *Marketing Science*, 35 (3), 389–404.

Jiang, Zhaohui (Zoey), Jun Li, and Dennis Zhang (2025), "A High-Dimensional Choice Model for Online Retailing," *Management Science*, 71 (4), 3320–3339.

Kalyanam, Kirthi, Sharad Borle, and Peter Boatwright (2007), "Deconstructing Each Item's Category Contribution," *Marketing Science*, 26 (3), 327–341.

Kim, Mingyung, Eric T. Bradlow, and Raghuram Iyengar (2026), "A Bayesian Dual Clustering Approach for Selecting Data and Parameter Granularities," *Marketing Science*, 45 (3), 556–575.

Leeflang, Peter S. H. and Josefa Parreño-Selva (2012), "Cross-Category Demand Effects of Price Promotions," *Journal of the Academy of Marketing Science*, 40 (4), 572–586.

Murphy, Kevin M. and Robert H. Topel (2002), "Estimation and Inference in Two-Step Econometric Models," *Journal of Business & Economic Statistics*, 20 (1), 88–97.

Nijs, Vincent R., Marnik G. Dekimpe, Jan-Benedict E. M. Steenkamp, and Dominique M. Hanssens (2001), "The Category-Demand Effects of Price Promotions," *Marketing Science*, 20 (1), 1–22.

Plummer, Martyn (2015), "Cuts in Bayesian Graphical Models," *Statistics and Computing*, 25 (1), 37–43.

Ringel, Daniel M and Bernd Skiera (2016), "Visualizing Asymmetric Competition Among More Than 1,000 Products Using Big Search Data," *Marketing Science*, 35 (3), 511–534.

Rooderkerk, Robert P. and Donald R. Lehmann (2021), "Incorporating Consumer Product Categorizations into Shelf Layout Design," *Journal of Marketing Research*, 58 (1), 50–73.

Rosenbaum, Paul R. (1984), "The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment," *Journal of the Royal Statistical Society: Series A (General)*, 147 (5), 656–666.

Ruiz, Francisco J. R., Susan Athey, and David M. Blei (2020), "SHOPPER: A Probabilistic Model of Consumer Choice with Substitutes and Complements," *The Annals of Applied Statistics*, 14 (1), 1–27.

Sethuraman, Jayaram (1994), "A Constructive Definition of Dirichlet Priors," *Statistica sinica*, pages 639–650.

Shi, Haolun and Guosheng Yin (2021), "Reconnecting p-Value and Posterior Probability Under One- and Two-Sided Tests," *The American Statistician*, 75 (3), 265–275.

Stephens, Matthew (2000), "Dealing with Label Switching in Mixture Models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62 (4), 795–809.

Teh, Yee Whye, Michael I. Jordan, Matthew J. Beal, and David M. Blei (2006), "Hierarchical Dirichlet processes," *Journal of the American Statistical Association*, 101 (476), 1566–1581.

Teh, YW, MI Jordan, MJ Beal, and DM Blei (2004), "Hierarchical Dirichlet Processes (Technical Report 653)," *UC Berkeley Statistics*.

Urban, Glen L. and John R. Hauser (2004), ""Listening In" to Find and Explore New Combinations of Customer Needs," *Journal of Marketing*, 68 (2), 72–87.

Van Heerde, Harald J., Sachin Gupta, and Dick R. Wittink (2003), "Is 75% of the Sales Promotion Bump Due to Brand Switching? No, Only 33% Is," *Journal of Marketing Research*, 40 (4), 481–491.

Van Heerde, Harald J., Peter S. H. Leeflang, and Dick R. Wittink (2004), "Decomposing the Sales Promotion Bump with Store Data," *Marketing Science*, 23 (3), 317–334.

Web Appendix

These materials have been supplied by the authors to aid in the understanding of their paper. The AMA is sharing these materials at the request of the authors.

CONTENTS

A Web Appendix A: Estimation Settings & Convergence Diagnostics

We provide details on the priors used in the estimation, along with convergence diagnostics for the MCMC chains.

A.1 Priors and Hyperparameters

For the HDP, we placed independent $\mathrm{Gamma}(0.1, 1)$ priors on the customer- and population-level concentration parameters, α and γ, and fixed the product-distribution concentration parameter at $\eta = 1$. For the causal regressions, coefficient vectors received diffuse zero-mean Normal priors with variance 5, and the error variances received inverse-Gamma priors with shape and scale parameters equal to 1.

A.2 Convergence Diagnostics for HDP

Figure A.1 reports the MCMC trace of the number of active HDP components over 10,000 iterations. After 4,000 iterations, the sampler stabilizes and explores representations with roughly 12-15 demand components. After 8,000 iterations, we retain $K = 12$ components (the mode of the sampled number of components). The trace is therefore intentionally flat as it reflects the freeze step used to stabilize the treatment-effect decomposition.[12]

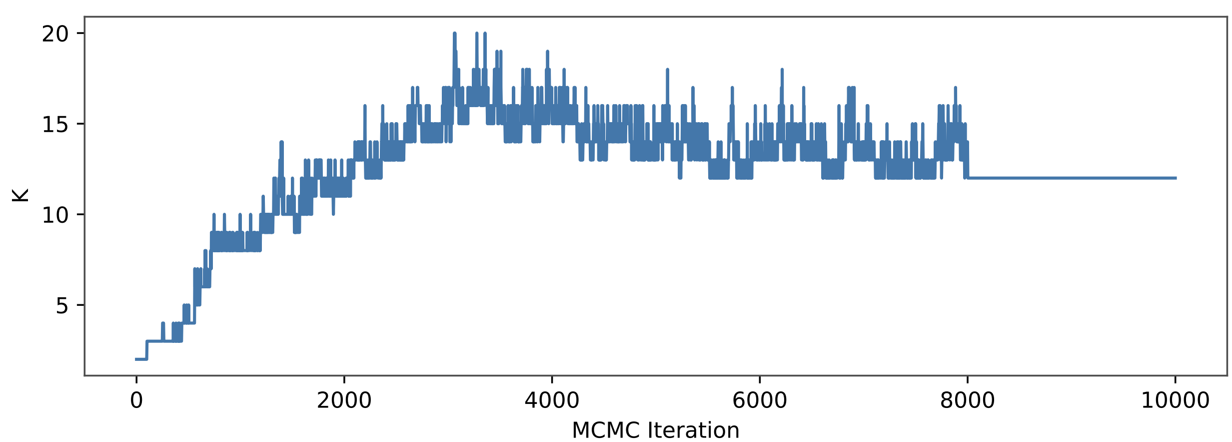

Figure A.1: Trace of the number of active components across MCMC iterations

Figures A.2 and A.3 display the MCMC draws for the concentration parameters γ and α, respectively. In both cases, the chains converge and stabilize around persistent posterior regions, with no visible trends or systematic drift, after 4,000 iterations. These diagnostics provide evidence of satisfactory

[12]The concentration parameters α and γ are also fixed at this stage. The remaining HDP quantities, including customer-component mixtures and component-product distributions, continue to be sampled so that uncertainty is propagated into the causal estimates.

convergence. After 8,000 iterations, these parameters are fixed at their posterior modes to stabilize the treatment-effect decomposition.

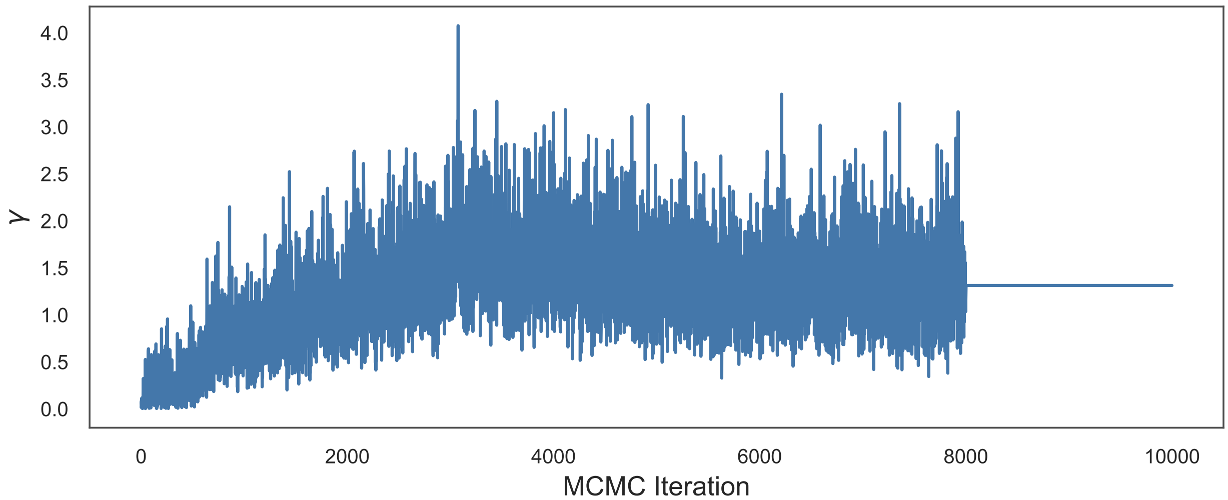

Figure A.2: MCMC Draws for Population Level Concentration γ Parameter

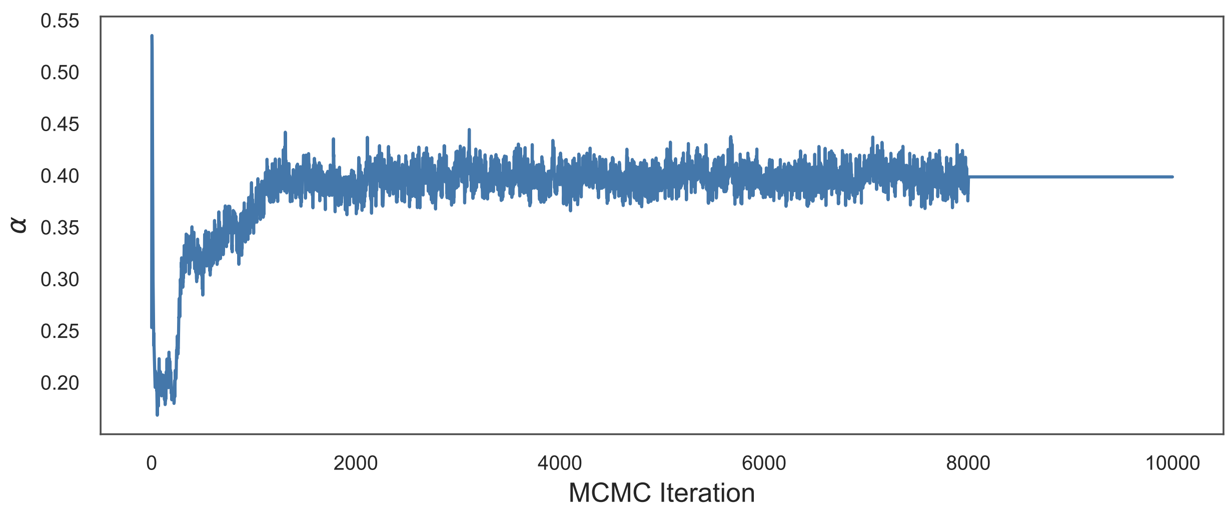

Figure A.3: MCMC Draws for Population Level Concentration α Parameter

A.3 Convergence Diagnostics for Regressions

We estimate the treatment-effect regressions using 2,000 HDP posterior draws. We discard the first 1,000 as burn-in and retain the remaining 1,000 for causal-effect reporting. Estimating the regressions for each retained draw propagates uncertainty in the latent demand components into the treatment-effect estimates.

We next examine convergence for parameters central to the empirical analysis. Figure A.4 reports the MCMC traces for the component-level ATEs $\{\tau_k\}$ across the post burn-in 1,000 iterations. Each trajectory fluctuates around a stable center with no persistent drifts or signs of nonstationarity. The marginal variance appears roughly constant throughout the chain, and visual inspection reveals no stickiness or slow-moving segments indicative of poor mixing. Successive draws show no evidence of severe autocorrelation that would impede effective sampling.

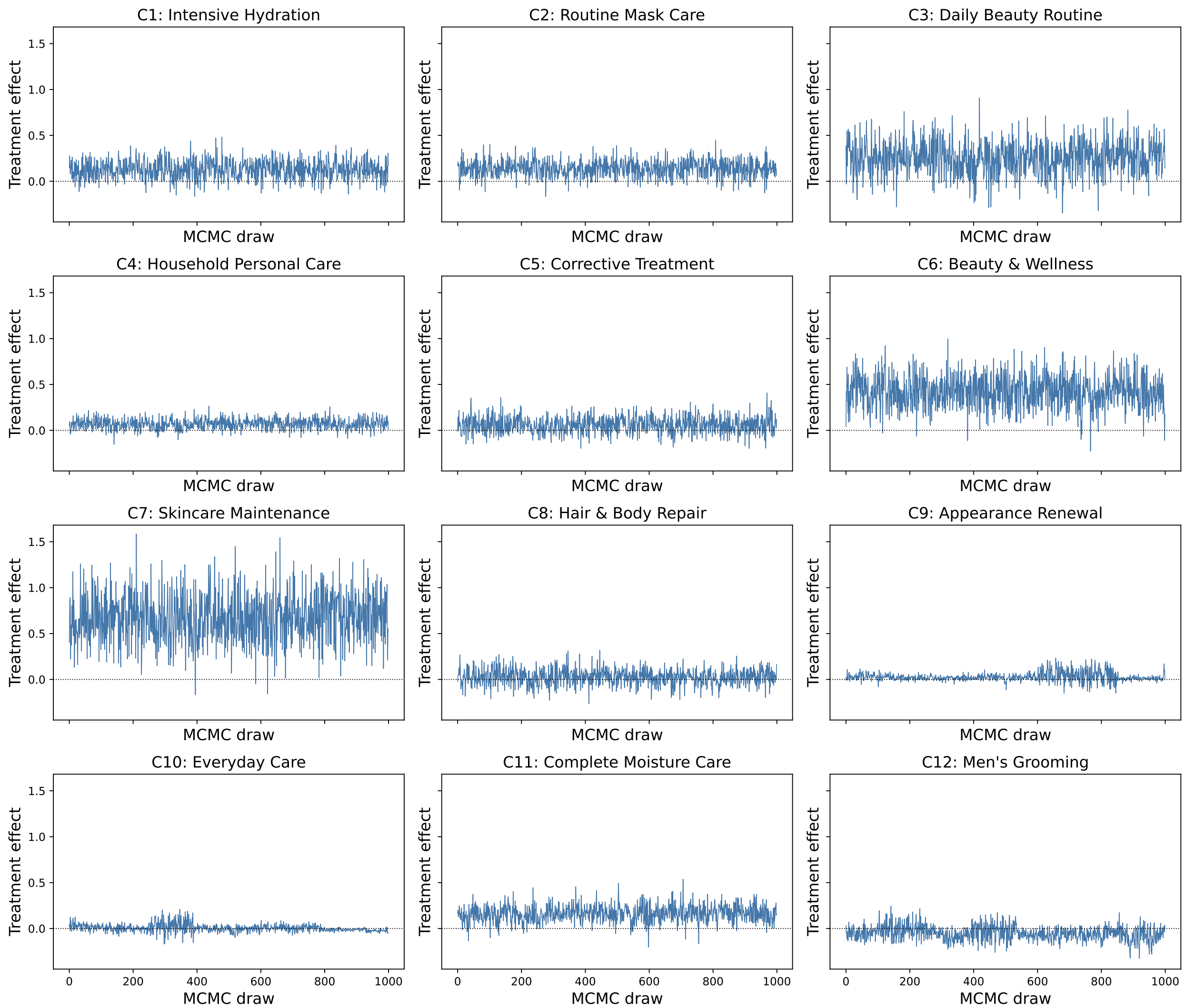

Figure A.4: MCMC Draws for Component-Level Treatment Effects

B Web Appendix B: Customer Heterogeneity over Demand Components

Figure B.1 illustrates consumers' distinct distributions ($\boldsymbol{\theta}_i$) recovered by the HDP for three random customers. Each customer places substantial weight on only a small subset of demand components, but the dominant components differ markedly across individuals. Consumer 1 is primarily associated with Daily Beauty Routine (C3), Consumer 2 with Beauty & Wellness (C6), and Consumer 3 with a combination of Daily Beauty Routine (C3) and Skincare Maintenance (C7). The figure therefore shows that customers are not assigned to a single segment; instead, each is represented by a distinct mixture of shared purchasing routines.

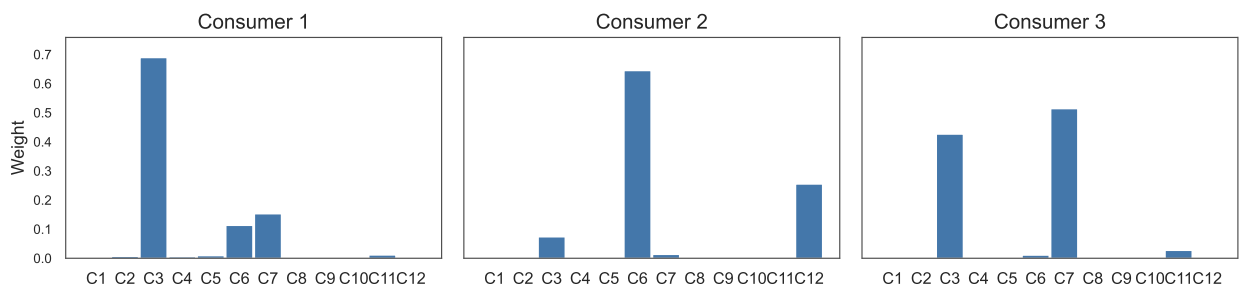

Figure B.1: Customer Specific Component Weights

C Web Appendix C: Benchmark Models Details

We provide details over the benchmark models used in our validation section.

C.1 Naive Uniform Model

The naive uniform model assumes that every product in the assortment is equally likely to be purchased by every customer. Let $v_{in} \in \{1, \ldots, J\}$ denote the product associated with purchase instance n for customer i. The model specifies

$$\Pr(v_{in} = j) = \frac{1}{J}, \qquad j = 1, \ldots, J. \tag{20}$$

The model contains no customer-specific parameters and does not use the training data to estimate product preferences. Its predictive distribution is therefore determined entirely by the number of products in the assortment. Model performance is evaluated by applying this fixed probability distribution to the held-out purchase instances in the test sample.

C.2 Global Product-Popularity Model

The global product-popularity model assumes that all customers share the same product-choice distribution. Let $v_{in} \in \{1, \ldots, J\}$ denote the product associated with purchase instance n for customer i. The model specifies

$$\Pr(v_{in} = j) = p_j, \qquad j = 1, \ldots, J, \tag{21}$$

where p_j denotes the population-level probability of purchasing product j. These probabilities are estimated from the training sample using smoothed product frequencies,

$$\widehat{p}_j = \frac{n_j + \lambda}{\sum_{\ell=1}^{J} n_\ell + \lambda J}, \tag{22}$$

where n_j is the number of training-sample purchase instances associated with product j and $\lambda > 0$ is a symmetric Dirichlet smoothing parameter. The resulting predictive distribution is identical for all customers and ranks products according to their estimated aggregate popularity.

C.3 Observed-Category Popularity Model

The observed-category popularity model assumes that customers differ in their preferences across the retailer's predefined product categories, while product choice within each category is governed by aggregate product popularity. Let $c_j \in \{1, \ldots, C\}$ denote the observed category of product j. The model specifies

$$\Pr(v_{in} = j \mid i) = \Pr(c_j \mid i) \Pr(j \mid c_j). \tag{23}$$

Customer-specific category probabilities are estimated from training purchases using shrinkage toward the population-level category distribution,

$$\widehat{\Pr}(c \mid i) = \frac{n_{ic} + \lambda \widehat{\pi}_c}{n_i + \lambda}, \tag{24}$$

where n_{ic} is the number of customer i's training purchases in category c, $n_i = \sum_{c=1}^{C} n_{ic}$, $\lambda > 0$ controls the strength of shrinkage, and $\widehat{\pi}_c$ is the smoothed global probability of category c. Product probabilities within category are estimated as

$$\widehat{\Pr}(j \mid c_j) = \frac{n_j + \eta}{n_{c_j} + \eta J_{c_j}}, \tag{25}$$

where n_j is the number of training purchases of product j, n_{c_j} is the total number of training purchases in category c_j, J_{c_j} is the number of products assigned to that category, and $\eta > 0$ is a symmetric smoothing parameter. The resulting predictive distribution is customer-specific across categories but common across customers within each category.

D Web Appendix D: Estimating the HDP Using Post-Treatment Purchases

Our identification strategy requires the latent demand components to be estimated exclusively from pre-treatment purchase histories. Because treatment assignment occurs after the latent representation is constructed, the resulting component memberships are predetermined with respect to the intervention and can be interpreted as valid pre-treatment covariates in the subsequent causal analysis.

To illustrate the importance of this design choice, we re-estimate the HDP using the combined pre- and post-treatment purchase histories. This specification allows the treatment to influence the construction of the latent demand components themselves, thereby violating the temporal ordering required for causal inference. The latent representation becomes a post-treatment variable that may partially encode the effects of the promotion, introducing post-treatment bias into subsequent estimates of component-level treatment effects.

Table D.1: Pre-Treatment Balance Between Control and Treatment Groups

Component	Control	Treatment	Difference	SE	p
C1	0.085	0.075	-0.009	0.009	0.308
C2	0.083	0.092	0.009	0.008	0.304
C3	0.057	0.055	-0.002	0.007	0.774
C4	0.252	0.261	0.010	0.012	0.444
C5	0.270	0.274	0.004	0.013	0.740
C6	0.112	0.109	-0.003	0.009	0.750
C7	0.023	0.029	0.007	0.004	0.107
C8	0.036	0.029	-0.007	0.005	0.214
C9	**0.058**	**0.041**	**-0.016**	**0.007**	**0.015**
C10	0.001	0.001	-0.001	0.001	0.457
C11	0.003	0.004	0.001	0.001	0.298
C12	0.000	0.001	0.001	0.000	0.077
C13	0.004	0.004	0.000	0.002	0.909
C14	0.002	0.002	-0.001	0.001	0.583
C15	**0.016**	**0.023**	**0.007**	**0.003**	**0.028**
C16	0.000	0.000	-0.000	0.000	0.810
C17	0.000	0.000	0.000	0.000	0.117

Table D.1 demonstrates the consequences of this specification. Relative to the main analysis, the estimated latent structure changes substantially, including an increase in the number of inferred demand components from 12 to 17. More importantly, the posterior component memberships are no longer balanced across the randomized groups. The treatment group has a significantly lower membership in C9 than the control group, with a difference of -0.016 ($p = 0.015$), and a significantly higher membership in C15, with a difference of 0.007 ($p = 0.028$). C12 also exhibits a positive imbalance of 0.001 with marginal significance ($p = 0.077$). Because these components are estimated using post-treatment purchases, these differences reflect treatment-induced changes in the latent representation rather than pre-existing differences between customers. This contamination makes the component memberships endogenous to treatment and invalidates their use in a causal decomposition.

These results demonstrate that estimating the HDP on post-treatment purchases contaminates the latent representation with treatment-induced behavior, breaking the causal validity of the subsequent decomposition. For this reason, the main analysis estimates the HDP exclusively on pre-treatment purchase histories before estimating treatment effects.

Table D.2: Treatment Effects with the HDP Estimated on Pre- and Post-Treatment Purchases

Component	Post. Mean	Post. SD	p_{Bayes}
C1	0.107	0.086	0.220
C2	0.131	0.091	0.126
C3	0.059	0.046	0.190
C4	0.252	0.173	0.148
C5	1.035	0.301	<0.001
C6	0.210	0.142	0.138
C7	0.037	0.062	0.520
C8	-0.056	0.083	0.496
C9	0.038	0.092	0.644
C10	0.002	0.006	0.676
C11	0.022	0.053	0.592
C12	0.000	0.011	0.920
C13	0.025	0.040	0.468
C14	0.000	0.004	0.838
C15	0.052	0.066	0.368
C16	0.000	0.003	0.910
C17	0.000	0.002	0.984

Table D.2 reports treatment effects when the HDP is estimated using both pre- and post-treatment purchases. Although the aggregate treatment effect remains essentially unchanged, the component-level decomposition changes markedly. The treatment effect becomes concentrated almost exclusively in Component 5 (1.035, $p_{\text{Bayes}} < 0.001$). In contrast, the previously significant effects identified in multiple demand components largely disappear, with the remaining components exhibiting small and statistically insignificant treatment effects.

This change is a direct consequence of estimating the latent representation using post-treatment outcomes. Because the HDP is allowed to learn from treatment-induced purchasing behavior, the latent components themselves become functions of the treatment assignment. Consequently, part of the promotional effect is absorbed into the construction of the components rather than being estimated as a causal effect conditional on fixed pre-treatment demand. The resulting decomposition no longer admits a causal interpretation: it describes a latent structure that has already been influenced by the intervention. These findings underscore the importance of estimating the HDP exclusively on pre-treatment purchase histories before conducting the causal analysis.

E Web Appendix E: Placebo Test Using Random Treatment Assignment

To assess whether the estimation procedure mechanically produces aggregate, component-level, or heterogeneous treatment effects, we conduct a placebo test based on a randomly generated treatment assignment. Specifically, we randomly divide customers into placebo treatment and control groups while leaving their observed post-treatment purchases unchanged. We then repeat the full treatment-effect analysis using this placebo assignment. Because the assignment is unrelated to customers' actual exposure to the promotion, the resulting estimates should be centered around zero and should not exhibit systematic heterogeneity.

Table E.1 reports the aggregate and component-level placebo estimates. The aggregate placebo effect is 0.494, with a posterior standard deviation of 0.457 and $p_{\text{Bayes}} = 0.276$. The component-level estimates are similarly small relative to their posterior uncertainty. None of the twelve components provides evidence of a placebo treatment effect, with all component-level Bayesian p-values exceeding 0.49. These results indicate that the component-level effects documented in the main analysis do not arise mechanically from the estimation procedure.

Table E.1: Placebo Treatment Effects Using Random Treatment Assignment

Component	Post. Mean	Post. SD	p_{Bayes}
Aggregate	0.494	0.457	0.276
C1: Intensive Hydration	0.028	0.105	0.784
C2: Routine Mask Care	0.030	0.091	0.742
C3: Daily Beauty Routine	0.085	0.183	0.614
C4: Household Personal Care	0.025	0.059	0.642
C5: Corrective Treatment	0.040	0.090	0.650
C6: Beauty & Wellness	0.020	0.186	0.944
C7: Skincare Maintenance	0.128	0.264	0.634
C8: Hair & Body Repair	0.055	0.085	0.494
C9: Appearance Renewal	0.016	0.049	0.698
C10: Everyday Care	0.004	0.042	0.950
C11: Complete Moisture Care	-0.043	0.096	0.666
C12: Men's Grooming	0.025	0.076	0.772

We also examine whether the placebo assignment generates spurious heterogeneity in component-level treatment effects. Figure E.1 reports the estimated interactions between the placebo treatment indicator and the observed customer characteristics used in the main analysis, as well as customers' pre-treatment spending across the twelve demand components. Each cell reports the posterior mean interaction coefficient, with its Bayesian p-value in parentheses.

The placebo heterogeneity estimates provide no evidence of systematic moderation. None of the interaction coefficients is statistically distinguishable from zero at conventional levels. Although a small number of cells have Bayesian p-values between 0.07 and 0.09, these estimates are isolated

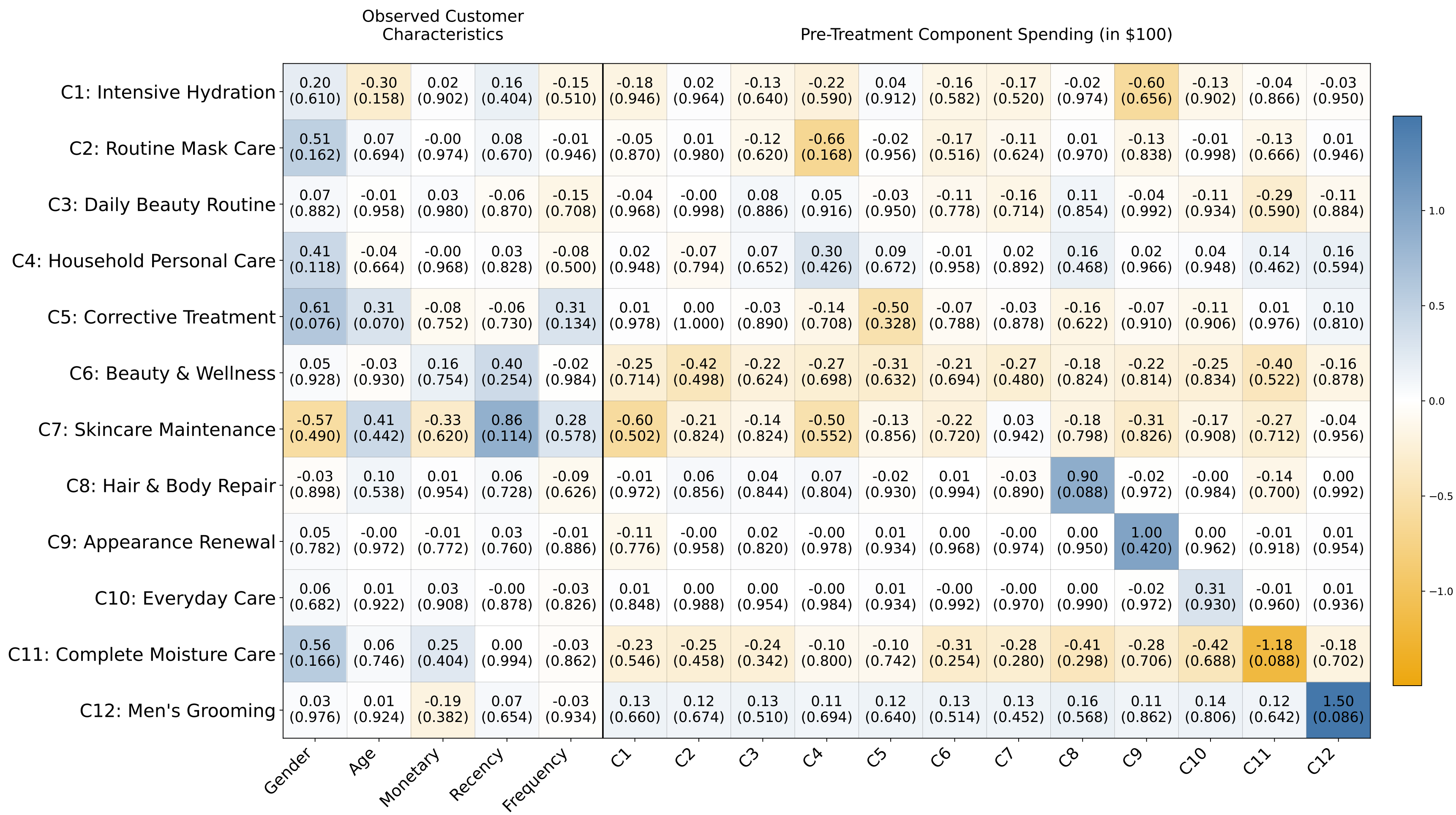

Figure E.1: Placebo Component-Level Treatment-Effect Heterogeneity across a large set of interactions and do not reveal a coherent pattern across customer characteristics or pre-treatment component spending.

Overall, the aggregate, component-level, and heterogeneity placebo results support the interpretation that the effects reported in the main analysis reflect the randomized promotion rather than artifacts of the demand-component decomposition or treatment-effect estimation procedure.