ABSTRACT

Title of Dissertation:	STRATEGIC MERCHANT COMPETITION AND LARGE ASSORTMENT MANAGEMENT IN ONLINE MARKETPLACES
	Min Kim, Doctor of Philosophy in Business and Management, 2020
Dissertation directed by:	Jie Zhang
	Professor of Marketing and the Harvey Sanders Fellow of Retail Management
	R. H. Smith School of Business

In this dissertation research, I study strategic merchant competitions on a retail deal platform and propose a new modeling approach for discovering consumer preference structures in online shopping environments with large and frequently changing assortments. In the first essay, I examine the impact of alternative platform policies and changes in market primitives on payoffs of players on a retail deal platform where merchants make frequent strategic decisions based on competitive

considerations in a two-sided market. To this end, I construct a dynamic game model and conduct counterfactual simulations. My analyses reveal novel insights regarding the interplay between platform policies and merchants' profits and relationships among competing merchants, and identify several mutually beneficial growth opportunities for all parties involved via policy changes. They also provide a comprehensive assessment of the net contribution of each merchant's action to the platform eco-system, which sheds new light on how to identify and foster "star brands" in such marketplaces. In the second essay, I develop a scalable stock-keeping-unit-level modeling framework of discovering consumers' preference structures in large and frequently changing assortments at the store/marketplace level. My proposed model identifies the underlying "topics of interest" that drive online browsing and purchase activities concerning the entire store assortment. It overcomes several limitations of standard Latent Dirichlet Allocation models for handling large assortments and enables demand forecast for products not in the existing assortments. I apply the proposed framework to investigate "topics" driving browsing and purchase activities in an online marketplace of fashion products, which reveals distinct topics driving these two types of shopping activities and their time-varying patterns of relevance. Finally, I propose a personalized product recommendation system that is based on individuals' preference structures inferred from the model and powered by the Bayesian optimization algorithm. Hold-out comparisons show that our approach substantially outperforms benchmark recommendation algorithms commonly used in practice (such as user-, item-, and content-based collaborative filters), and it is particularly powerful in prescribing recommendations for products that do not exist in the previous assortment. This dissertation research offers a range of policy recommendations and valuable managerial insights to marketing researchers and E-commerce practitioners.

STRATEGIC MERCHANT COMPETITION AND LARGE ASSORTMENT MANAGEMENT ON A DEAL PLATFORM

by

Min Kim

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2020

Advisory Committee:

Professor Jie Zhang, Chair

Professor Michel Wedel

Associate Professor Yogesh Joshi

Assistant Professor Bruno Jacobs

Professor Andrew Sweeting, Dean's Representative

Acknowledgements

I am wholeheartedly thankful to my advisor, Dr. Jie Zhang, whose caring guidance and full support enabled me to complete this dissertation and my Ph.D. She is the role model that I respect and want to follow in my academic career down the road. I would like to thank Dr. Michel Wedel, who introduced me to the joy of Bayesian modeling. I am also grateful for Dr. Yogesh Joshi, who has helped me to sharpen critical thinking through our discussions. I would like to thank Dr. Bruno Jacobs, whose keen insights and valuable suggestions have contributed greatly to the completion of this dissertation. I am also in debt to Dr. Andrew Sweeting, who first introduced me to the beauty of the marketplace mechanism and kindly served on my defense committee. I am also grateful to Professor Lily Du of Xi'dian University for her help with acquiring the data. My deepest gratitude goes to my parents, Daeyoung Kim and Sunja Lee, my younger brother, Minho Kim, and his wife, Eunji Song, for their unconditional support throughout my life. This dissertation is impossible without them.

Table of Contents

INTRODUCTION

This dissertation examines behaviors, decisions, and interactions of consumers, merchants, and platform managers in a platform ecosystem. It involves two essays, each of which studies distinct aspects and managerial issues in the context of a platform economy. The platform presents a two-sided market, which may alter the conventional competitive relationships among merchants due to indirect network effects. A merchant's contribution to the ecosystem is no longer just a matter of its own market share and direct impact on its competitors' sales, but would also depend on its indirect impact on the market size and competitive actions. As a consequence, it requires a more comprehensive approach to assess the net contribution of a merchant to a platform. In addition, many success stories from today's marketplace suggest that market primitives and platform policies may be crucial factors to determine their success. In order to quantify their impact, the first essay focuses on how market primitives and different platform policies affect the payoffs of various players in the ecosystem.

Furthermore, many retailers and e-commerce operators (including online platforms) sell products from large and frequently changing assortments, such as sellers of clothes, footwear, bags, fashion accessories, video games, music, movies, books, electronics, and gadgets. For marketers, inferring consumer preference structures in such context is important for making a variety of marketing actions such as making merchandise assortment decisions and offering personalized merchandise recommendations, yet it is a very challenging task due to the large and varying nature of the merchandise assortment, and traditional modeling approaches developed in the marketing literature are not well equipped to handle it. In the second essay of this dissertation, I take up this challenge and develop a scalable modeling framework of inferring consumer preferences in the context of large and frequently changing assortments.

In Essay I, I examine the impact of alternative platform policies and changes in market primitives on payoffs of players on a retail deal platform where merchants make frequent strategic decisions based on competitive considerations in a two-sided market. To this end, I construct a dynamic game model in which each merchant decides the number of deals offered and depth of discount based on its expectations of other merchants' decisions and future profit streams, and a consumer demand model that endogenizes the effects of merchant decisions on the market size. I estimate the models using data from a retail deal platform specializing in fashion products, and conduct counterfactual simulations to examine the impact of alternative commission policies, imposing requirements of a minimum number of deals, and changes in a merchant's cost structure or brand equity on the payoffs of all players in the platform eco-system. My analyses reveal novel insights regarding the interplay between platform policies and merchants' profits and relationships among competing merchants, and identify several mutually beneficial growth opportunities for all parties involved via policy changes. They also provide a comprehensive assessment of the net contribution of each merchant's action to the platform eco-system, which sheds new light on how to identify and foster "star brands" in such marketplaces.

In Essay II, I develop a scalable stock-keeping-unit-level modeling framework of discovering consumers' preference structures in large and frequently changing assortments at the store/marketplace level. My proposed model identifies the underlying "topics of interest" that drive online browsing and purchase activities concerning the entire store assortment. It overcomes several limitations of standard Latent Dirichlet Allocation models for handling large assortments and enables demand forecast for products not in the existing assortments. I apply the proposed framework to investigate "topics" driving browsing and purchase activities in an online marketplace of fashion products, which reveals distinct topics driving these two types of shopping activities and their time-varying patterns of relevance. Finally, I propose a personalized product recommendation system that is based on individuals' preference structures inferred from the model and powered by the Bayesian optimization algorithm. Hold-out comparisons show that our approach substantially outperforms benchmark recommendation algorithms commonly used in practice (such as user-, item-, and content-based collaborative filters), and it is particularly powerful in prescribing recommendations for products that do not exist in the previous assortment.

In both essays, I offer a range of policy recommendations and valuable managerial insights to marketing researchers and E-commerce practitioners.

ESSAY I

Strategic Merchant Competitions on a Retail Deal Platform: Implications for Platform Policies and Management

INTRODUCTION

Retail deal platforms are online marketplaces where merchants sell directly to consumers of products which are usually at deep discounts and available for only a limited time. They provide a cost-efficient way of selling products with limited-time deals, especially for small- and medium-sized merchants which are often constrained by marketing budgets and resources to pay upfront costs associated with conventional promotions (Cao, Hui, and Xu 2015). In addition, they could be utilized as a sales channel similar to outlet stores that sells older and less desirable merchandise to capture price-sensitive consumers while at the same time prevents cannibalization of regular store revenues (Ngwe 2017). From the demand side, consumers benefit from shopping on these deal platforms because they decrease the search costs of finding deals and comparing attributes across different merchants' websites. These advantages contribute to the rapid growth of retail deal platforms in recent years around the world (e.g., Vipshop in China, Cdiscount in France, Snapdeal in India, Gmarket in Korea, and Museco in Japan). For example, Vipshop, founded in 2008, operates the largest retail deal platform and is the third largest E-commerce site in China. Its annual sales have grown from $1.7 billion in 2013 to $11.21 billion in 2017 (Amigobulls 2018). Cdiscount, offering appliances and fashion products since 1998, has recorded gross merchandise value growth from €1.6 billion in 2012 to €3.4 billion in 2017 (Statista 2017, Cnova 2018).

Retail deal platforms share some similarities with conventional retail marketplaces (e.g., Amazon marketplace, eBay marketplace, Spartoo marketplace, Tmall, and Rakuten) and deal distribution platforms (e.g., Groupon, LivingSocial, Scoopon, and Cudo), but also have unique features. Like the former, they provide a marketplace for merchants to sell directly to consumers and in general leave the specific merchandising decisions to sellers, but products sold on retail deal platforms usually are offered with (deep) discounts and available for only a limited time, and thus merchants need to make frequent decisions on how many products to sell and at what discounts. Their deal orientation is a similarity shared with Groupon type of deal distribution platforms. But unlike the latter, which distribute one-time discount coupons to be redeemed later at the merchants' stores, a retail deal platform offers a marketplace for merchants and consumers to transact directly on a repeated basis. The basic business model behind retail deal platforms is also different from that of online off-price retailers (e.g., Overstock.com, Gilt.com, Zuily.com, RueLaLa.com, and Woot.com). Off-price retailers source from various vendors, make merchandising decisions for individual products, and own the merchandise. Retail deal platforms, on the other hand, do not own the merchandise and usually leave the specific merchandising decisions to individual merchants. Their main source of revenue is commissions paid by merchants based on sales generated on their platforms.

Prior research on deal platforms has mainly focused on Groupon type of deal distribution platforms where consumers can find one-time deal offers from a large number of merchants, and the pricing and assortment decisions are usually influenced if not controlled by the platforms. Some studies have looked into consumers' purchase behaviors on these platforms (e.g., Cao et al. 2015). Others have investigated merchants' decisions on whether or not to use a given deal platform (e.g., Kumar and Rajan 2012, Wu, Liang, and Chen 2015, Edelman, Jaffe, and Kominers 2016), and the impact of bargaining power between platforms and merchants (Zhang and Chung 2016). The focus of our study is retail deal platforms, which have several distinct features different from deal distribution platforms and other online retail formats as outlined above. Despite their growing popularity in E-commerce practices, many issues regarding how to manage such platforms remain unanswered. For example, should the platform charge the same commission rate to all merchants or differential rates? If the latter, should they offer preferential treatment to market share leaders or smaller brands, or could they milk more commissions out of the top brands, and how would a change in the commission policy affect the competition among merchants and the sales consequences of all parties?

As more E-commerce firms entering the retail deal platform business, a pressing issue is how to grow the customer base of their platforms. A policy that platform operators could consider is to impose a requirement of the minimum number of deals/products by each merchant. This could force merchants to offer more deals than they otherwise would have, and quickly attract more consumers to a platform through the joint actions of all merchants, which in turn could benefit all parties. On the other hand, if the requirement is considered too stringent, merchants may choose to withdraw from the platform instead of complying to the minimum requirement, which would then backfire on the platform. Is imposing a minimum number of deals an effective way to grow a platform, and how should the minimum number be determined? Answers to these questions have direct managerial implications and require a comprehensive examination of the interplay between the platform, merchants, and consumers.

Another issue facing platform operators is how to evaluate the net contribution of actions taken by each merchant selling on its marketplace and to identify and foster “star brands”. Are they necessarily the market share leaders or most premium brands? How would one merchant’s cost reduction effort affect not only its own sales or profit but also those of other merchants and the platform? How would changes in a merchant’s brand equity affect the well-being of all parties in the eco-system?

We address these managerial issues by conducting an empirical investigation utilizing data from a retail deal platform. Several important features of retail deal platforms need to be taken into consideration in this research endeavor. First, on these platforms, merchants tend to have frequent presence in certain product categories, even though the specific products/deals that they offer are usually of a limited time. Therefore, a merchant needs to make frequent and repeated decisions on the number of deals offered and the depth of discounts on a given platform. Second, each merchant's decisions are likely to be influenced by those of other merchants. They face multiple competitors and the competitive environment changes rapidly because most deals are short-lived, and thus each merchant needs to evaluate the relative competiveness of their offerings and adjust its number of deals and depth of discounts accordingly. Moreover, like in other types of marketplaces, the behaviors of agents (merchants and consumers) on a retail deal platform well represent the nature of a two-sided market with indirect network effects, where the actions of merchants affect the attractiveness of the platform, which in turn affects the participation and purchase behavior of consumers, and vice versa (Rochet and Tirole 2006). From the consumers' perspective, an increase in the number of deals and/or depth of discounts on a platform is likely to enhance the attractiveness of the platform and thus increase the number of active users and sales on the platform. For merchants, more active users would lead to a larger potential market size, and thus expectations of higher future profit streams. Therefore, merchants would take into account how their merchandise decisions in each time period affect their future profitability, which means that their decisions are likely to be forward-looking in nature. To account for these features, we construct a dynamic game model where merchants maximize their expected total profits by strategically choosing the number of deals and depth of discounts based on their expectations of the other merchants' decisions and anticipated future profit streams, and a consumer demand model that incorporates the effects of merchant decisions on the market size and thus merchants' future profit potentials. We estimate the models using data from a retail deal platform specializing in fashion products. The dataset includes detailed records of merchants' deal offerings on the platform as well as individual consumers' browsing and purchase history information. The model estimation results reveal how strategic interactions among merchants may drive their merchandise decisions as well as consumer participation and purchase behavior on the platform, and enable us to conduct various counterfactual analyses to address the aforementioned managerial questions.

We intend to make the following contributions in this research. Our primary contribution is a substantive one. We investigate several important issues regarding how to manage the relatively new and fast-growing E-commerce format --- retail deal platforms. Our analyses provide novel insights on how alternative policies may impact individual merchants and the platform, and reveal several mutually beneficial growth opportunities for merchants and the platform via policy changes. For example, we find that lowering the commission rate does not necessarily lead to higher profits for all merchants due to intensified competitions among them, but the platform could charge differential commission rates and deliver mutual benefits to all merchants via profit sharing. We also find that favorable changes in the market primitive of one merchant (such as increase in brand equity or reduction in fixed costs) can benefit its competitors, and that the merchant that brings about the greatest benefit to the ecosystem is not the market share leader or the most premium brand. Our research also contributes to the deal platform literature by broadening the scope of contexts investigated to retail deal platforms, which have unique features and managerial problems that require different modeling frameworks and research approaches. In addition, we contribute to the dynamic game model literature by endogenizing the impact of agents' decisions on the market size. Our analyses show that this additional feature to the classic game models plays a central role in driving several key results, which may seem counterintuitive otherwise, in our context.

The rest of the paper proceeds as follows. Section 2 reviews the relevant literature. Section 3 describes the data and institutional background of the retail deal platform studied here. Section 4 describes the model formulation and discusses the estimation strategy. Section 5 presents estimation results. In Section 6, we conduct counterfactual analyses and discuss results. Finally, we summarize key findings and discuss their implications in Section 7.

LITERATURE REVIEW

Our research builds upon two streams of prior literature. In terms of the substantive topic, it is related to the research area on deal distribution platforms. On the methodology front, it is built upon the dynamic oligopoly game model originally developed in the economics literature, which has also been applied by marketing academics. In the following subsections, we will provide a brief review of each stream of research and discuss the relations of our study to the prior literature.

Deal Platforms

Prior research on deal platforms has focused on Groupon type of deal distribution platforms and the specific topics examined can be categorized into three aspects: consumers, merchants, and platforms. Several studies empirically examine consumers' purchase behaviors, with an emphasis on the effectiveness of the depth of discounts. For example, Jeon et al. (2014) examine the impact of the discount rates on deal demand using data from a large online deal distribution site in Asia, and find that increases in discount rates are likely to enhance consumers' attention, thereby generating additional sales. Also using data from Groupon, Cao et al. (2015) find that higher depth of discounts can lead to the unintended consequence of consumers' quality concerns.

Some studies have investigated whether selling products via deal distribution platforms are profitable from the merchants' perspective. Based on analytical modeling frameworks, Kumar and Rajan (2012) show that offering discounts via deal platforms may not be effective to increase revenues or to attract repeat customers for those businesses offering services that are purchased infrequently, and Edelman et al. (2016) find that offering deals tends to be more profitable for merchants relatively unknown or with low marginal costs. Adopting a static game structural model in which platforms and merchants strategically set commission rates and depth of discounts, Wu et al. (2015) find that, while offering deals on deal distribution platforms can incur a loss in the current period due to deep discounts, it can increase merchants' future profits. This study highlights the importance of considering future profit implications when studying merchants' or platforms' decisions. A key difference of our study with those discussed above is that we focus on merchant decisions in the type of platforms where they need to make repeated merchandising decisions (as opposed to one-time or very infrequent decisions), and thus we need to adopt a dynamic game model with repeated merchant decisions.

Retail deal platforms are a classic example of two-sided markets, and prior research has demonstrated that it is important to take into account this feature when studying buyers' and sellers' decisions on platforms. Using data from an online marketplace, Chu and Manchanda (2016) empirically show the existence of positive indirect network effects from the buyers' and sellers' sides and identify factors that affect growth on both sides, such as the installed bases, product variety, buyer quality, and product prices. Adopting a structural model of bilateral Nash bargaining, Zhang and Chung (2016) study the impact of relative bargaining power between merchants and two deal distribution platforms, Groupon and Living Social. They show that capturing indirect network externalities is important in modeling pricing decisions on these platforms. Their results suggest that merchants can be worse off if platforms have higher price bargaining power.

Despite the growing literature on platform economies, there has been little empirical research on the relatively new and fast growing E-commerce format --- retail deal platforms. Many pressing issues regarding how to manage such platforms and set effective platform policies remain unanswered. Building upon valuable insights from the prior research, we examine a variety of platform policies and changes in market primitives to identify mutually beneficial growth opportunities via a comprehensive modeling framework that simultaneously takes into account strategic interactions between merchants, forward-looking considerations of each merchant, and indirect-network effects between merchants and consumers.

Dynamic Game Models and Estimation

In terms of the modeling approach, our research builds on the development of dynamic game models. As well known in the literature, solving a dynamic game model with closed-form analytical results is infeasible even with just a few state variables. The seminal work by Ericson and Pakes (1995) proposes a model of infinite-horizon dynamic oligopoly competition in a differentiated-products market and provides a tool to numerically solve the model with multiple state variables. A key feature of the model is that actions taken in a given period affect both current profits and future strategic interactions among players by influencing a set of commonly observed state variables (Bajari et al. 2007). There have been extensive applications of dynamic game models in the economics and marketing literature to understand firms' strategic behaviors over time (e.g., Dube et al. 2005, Yao and Mela 2011, Liu et al. 2016, Zhou 2017). For examples, Dube et al. (2005) examine optimal advertising scheduling by taking into account competitors' decisions, and Liu et al. (2016) investigate detailing schedules of firms selling competing drugs.

Two issues often complicate the estimation of a dynamic game model. First, the classic nested fixed point algorithm proposed by Rust (1987) requires solving for an equilibrium at each trial of the parameter vector by iteratively evaluating the value function, and thus the computational burden increases exponentially with state spaces, especially in the case of estimating a dynamic game (Bajari et al. 2007, Yao and Mela 2011). Second, the possibility of multiple equilibria requires a modeler to consider all possible equilibria in the game and to impose equilibrium selection rules *a priori*. A recent development of estimation method, the BBL estimator proposed by Bajari et al. (2007), mitigates the problems by relying on a two-step approach, which avoids the computational burden of evaluating the value function for each set of parameters. Yao and Mela (2011) and Liu et al. (2016) are examples in the marketing literature that have applied the BBL estimator to recover structure parameters in their dynamic game models, and have shown that this approach works very well empirically.

Among prior works employing dynamic game models, Yao and Mela (2011) and Zhou (2017) explicitly analyze two-sided markets. Yao and Mela (2011) study the equilibrium dynamic bidding strategies in sponsored search advertising for music management software by incorporating the behaviors of searchers, advertisers, and the search engine. Zhou (2017) studies video game consoles as platforms where consumers and video game sellers interact, and focuses on incumbent video game sellers' dynamic pricing decisions and potential entrants' entry decisions in the software market. While sharing some similarities in the modeling approach, the substantive context and managerial focus of our research are very different. We study merchants' repeated merchandising decisions on a retail deal platform, and incorporate the dynamics between strategic merchant decisions, size of the market, and consumer demand, which will serve as the basis to examine policy implications for merchants, consumers, and platform operators.

INSTITUTIONAL BACKGROUND AND DATA

Institutional Background

The dataset for our empirical analysis was provided by a mobile deal platform operator based in China. The platform was launched in April 2013 as a marketplace for merchants to sell fashion merchandise at discounted prices for limited time periods. The platform processes payments for merchants who pay the platform commissions at a fixed percentage of sales from each successful transaction. A customer must first join the platform by installing its app on a mobile device and registering as a user. After joining the platform, they can view available deals and detailed product information, such as the product picture, initial price, sales price, and payment methods, and place orders via the app. This platform operates exclusively through its mobile app and does not have a web presence.

Overall Data Pattern

Deals

Our dataset contains individual customers' browsing and purchase records and information on all deals (i.e., products) offered on the platform from August 2013 to June 2014, spanning a total of 46 weeks. Each observation of a deal contains information such as descriptions of the product, brand name, initial price[1], and sales price. Our empirical analyses are carried out using data from the largest merchandise category as classified by the platform – Female Tops, which consists of women's T-shirts, blouses, jackets, sweaters, and coats. Most merchants of this broad category offer deals in all relevant product categories and thus directly compete with each other. Focusing on examining one broad merchandise category not only keeps computations of our complex models feasible, but also is consistent with the premise of our research context in which merchants are close competitors in the same consumer market. We focus on the top five merchants in this merchandise category, which accounted for 72.8% of the total number of deals and 71.5% of the total sales revenues in the category. Table 1 presents descriptive statistics of deals offered by the top five merchants. It shows large variations in the initial prices, sales prices, depth of discounts, and the number of deals offered, not only across merchants but also within each merchant.

Table 1. Descriptive Statistics of Deals by the Top Five Merchants

Merchant	Variable	Mean	SD	Min	Max
A	Initial price (yuan)	770.40	369.21	370.58	1,540.15
	Sales price (yuan)	353.61	144.71	175.90	601.12
	Depth of discount (%)	51.76	7.76	36.80	60.73
	Number of deals	38.54	11.70	10.00	56.00
B	Initial price (yuan)	798.90	291.83	403.17	1,259.62
	Sales price (yuan)	601.59	213.52	280.00	1,013.94
	Depth of discount (%)	23.76	9.78	12.01	46.56
	Number of deals	29.15	10.60	8.00	60.00
C	Initial price (yuan)	1199.35	422.76	498	1,883.33
	Sales price (yuan)	688.27	262.7	225	1,199.75
	Depth of discount (%)	36.43	10.65	17.82	66.05
	Number of deals	13.17	4.47	1.00	21.00
D	Initial price (yuan)	940.59	378.50	356.50	1,635.00
	Sales price (yuan)	462.31	156.20	178.50	713.00
	Depth of discount (%)	48.26	7.16	35.44	57.71
	Number of deals	33.26	13.11	4.00	54.00
E	Initial price (yuan)	791.65	284.90	417.75	1,405.86

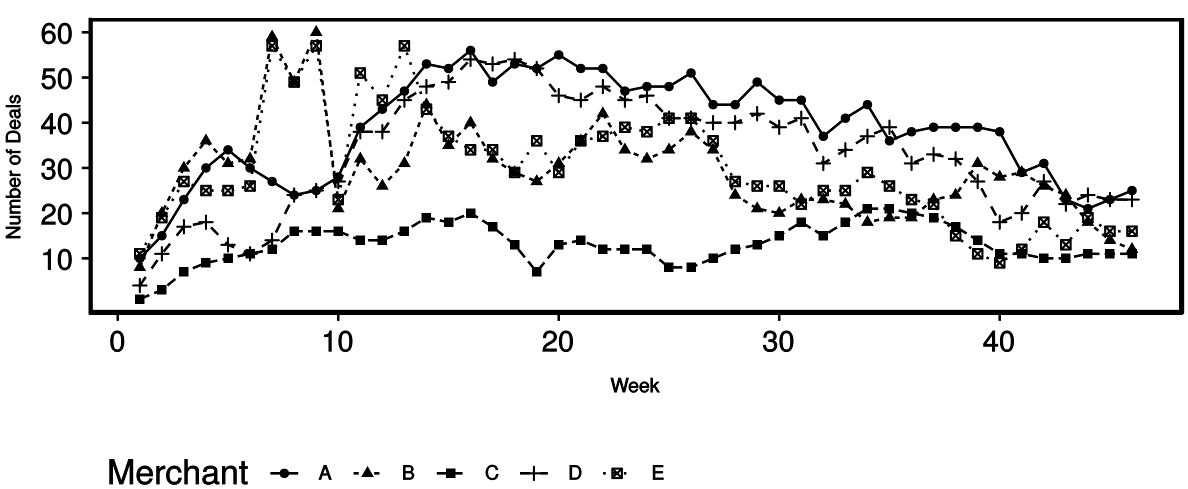
[1] The initial price is the price printed on the product label and is set before the product is sold in any stores or on any platforms. The focal retail deal platform presents both the initial price and sales price for each product.

Sales price (yuan)	588.06	174.40	282.64	1,004.21
Depth of discount (%)	24.04	9.20	12.01	40.48
Number of deals	29.61	12.64	9.00	57.00

Note: One yuan is approximately 0.15 US dollars.

Figure 1 shows the time-varying pattern of the number of deals offered by each merchant. Variations in the number of deals across merchants suggest that strategic interactions may exist among merchants. For example, a merchant's incremental number of deals on the platform appears to have a negative correlation (-0.17, p-value = 0.011) with the total number of deals offered by the other merchants in the previous period, suggesting that merchants might reduce their own number of deals if they saw the number of deals by its competitors was higher, possibly due to concerns that consumers' attention to a particular merchant's products may decrease as the total number of deals available on the platform goes up. In addition, one possible explanation for variations in the number of deals within merchants is that the costs incurred to the merchants associated with the number of deals may drive their decisions on the number of deals. The two varying patterns within and across merchants suggest that the model of merchants' decisions may need to capture both strategic forces and cost consideration.

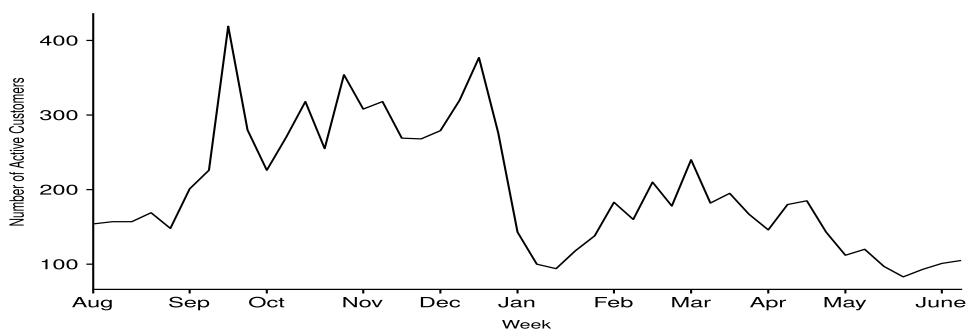
Figure 1. Time-Varying Pattern of the Number of Deals Offered by the Top Five Merchants



Market Size: The Number of Active Customers

We define the number of active customers as the number of users who browsed the Female Tops category on the platform in a given week. Figure 2 presents time series of the number of active customers by week, which shows cyclical patterns. We speculate that, in addition to seasonal variations, changes in the number of active customers may be influenced by merchants' decisions in each time period. For example, there is a high correlation (0.62, p-value = 0.001) between the number of active customers of the category on the platform in a given week and the total number of deals offered by the top five merchants, indicating that a higher number of deals available on the platform may attract more customer participations.

Figure 2. The Number of Active Customers



MODEL FORMULATION AND ESTIMATION

We focus on two of the most important decisions by merchants on a retail deal platform: the number of deals offered and the depth of discount. To maintain the tractability of the model, we focus on each merchant's decisions at the category level. The number of deals offered is measured as the number of stock-keeping units (SKUs) in the category sold by a merchant on the platform in a given week, and it is allowed to be zero to accommodate a merchant's decision to (temporarily) exit the platform. The depth of discount is operationalized as the average percentage discount off the initial price across the SKUs sold by a merchant in the focal category in a given week. Consistent with the set-up of dynamic oligopoly game models in the literature, we assume that merchants first make decisions on the total number of deals to be offered in the next period, and that, given the number of deals by all merchants in the current period, the average depth of discount is determined based on a static Nash-equilibrium in each period. The assumption of a static game regarding the depth of discount may not be very restrictive because most deals lasted for only a limited time on the platform.

Let j denote a merchant. Time in our model is discrete and each period is a week, $t \in \{1, \dots, T\}$. For the focal category, we assume that there is a fixed J number of merchants (i.e., no new entry) over the sample period, $j \in \{1, \dots, J\}$. This assumption is made because our empirical analyses focus on the top five merchants in the category and they were all present from the beginning of the data. Let $n_{jt} \in \{0,1, \dots, N\}$ represent the number of deals of merchant j at time t. The state vector $\boldsymbol{n_t} \equiv \{n_{jt}: j = 1,2, \dots, J\}$ represent the number of deals offered by each merchant at t.

The Timing of Decisions in the Game

Within each week t, the sequence of merchant decisions is modeled as follows:

1. At the beginning of the week, each merchant j observes i) the current states $\boldsymbol{n_t}$ which represent each merchant's number of deals at t; ii) exogenous variables that can affect profits, denoted by the vector $\boldsymbol{x_t} = (\boldsymbol{x_{1t}}, \dots, \boldsymbol{x_{jt}}, \dots, \boldsymbol{x_{Jt}})$, where $\boldsymbol{x_{jt}} \equiv \{x_{jct}: c = 1,2, \dots, C\}$, and C is the number of exogenous merchant characteristics, and iii) the number of active customers on the platform, A_t;

2. Each merchant incurs the internal administrative costs (*IC*) for maintaining their existing deals on the platform, which include costs associated with inventory management for products sold via the platform;
3. Each merchant observes the private information shock, ϵ_{jt}, to its decision on the change in the number of deals to offer in the next period (a_{jt}), and makes its choice for the next period, $a_{jt} \in \{-\overline{N}, \dots, -1, 0, 1, , \dots, \overline{N}\}$; [2]
4. Each merchant competes on the average depth of discount, earns the sales of the period, incurs the external administrative costs (*EC*) which is associated with a_{jt}, and receives realized shock to its profit, $\epsilon(a_{jt})$, which is a function of the decision a_{jt}, and earns the profit of the period;
5. $\boldsymbol{n_t}$, $\boldsymbol{x_t}$, and A_t evolve to the state in the next week according to merchants' decisions on the change in the number of deals (a_{jt}) and stochastic evolutions of exogenous merchant characteristics.

In the following sections, we will describe the models of market size and consumer demand, the merchants' profit functions, the equilibrium depth of discount and number of deals, and the estimation strategy.

The Market Size Model: Number of Active Customers

Retail deal platforms are two-sided markets with network externalities, where decisions of merchants affect the attractiveness of the platform, which in turn affects decisions of consumers, and vice versa (Rochet and Tirole, 2006). For consumers, an increase in the number of deals offered or the depth of discounts on a platform could raise their expectations of finding good

[2] We empirically determine the proper value of $\overline{N}$ and will address this issue in a later section.

deals and thus is likely to increase the number of active customers on the platform. The extant literature shows empirical evidence on the existence of such indirect network externalities on C2C online platforms (e.g., Chu and Manchanda 2016). We capture this phenomenon by explicitly modeling the number of active customers in a given week, A_t, as a function of:

$$A_t = \rho_0 + \rho_1 \sum_{j=1}^{J} n_{jt-1} + \rho_2 \bar{d}_{t-1} + e_t, \tag{1}$$

where $\sum_{j=1}^{J} n_{jt-1}$ is the total number of deals offered by the merchants and $\bar{d}_{t-1} = 1/J \sum_{j=1}^{J} d_{jt-1}$ is the average depth of discount across the merchants at time $t-1$, and e_t is the error term. Parameters ρ_1 and ρ_2 capture the indirect network effects of the total number of deals and average depth of discounts on the number of active customers. We use the lagged explanatory variables to avoid potential endogeneity. The number of active customers positively affects a merchant's profit through the market size. The empirical estimation result of Equation (1) will also shed light on which factor(s), being the number of deals and/or depth of discount, may be the driving force for forward-looking considerations by merchants (instead of making repeated static optimization decisions).

The Consumer Demand Model

We model S_{jt}, the share of customers who choose merchant j in week t. We assume that only one product is bought if a consumer makes a purchase from the focal category in week t.[3] Since each merchant sells only one brand in our data, we will use the terms "brand" and "merchant" interchangeable in the following exposition. We derive the shares from an individual brand-

[3] In our data, most consumers who made a purchase from the focal category in a given week bought only one product. Our demand model can be extended to accommodate the cases when consumers bought multiple products from the focal category in a given week, by treating these cases as different purchase incidences of the same consumer. See detailed discussion on this issue by Villas-Boas and Zhao (2005). Allowing the possibility of multiple purchases would severely increase the computation complexity. Since the main thrust of our modeling system is dynamics in merchants' decisions on the optimal number of deals, we keep the demand model fairly simple in order to maintain computational tractability.

choice demand model, in which a consumer i chooses one of J available brands indexed by j or the outside good (denoted by $j = 0$) in each week t.

The indirect utility of customer i choosing brand j at time t is as follows:

$$u_{ijt} = \alpha_i d_{jt} + o_i n_{jt} + \boldsymbol{x_{jt}} \boldsymbol{\beta_i} + \xi_{jt} + \zeta_{igt} + (1 - \lambda)\nu_{ijt}, \tag{2}$$

where d_{jt} is the average depth of discount, n_{jt} is the number of deals, $\boldsymbol{x_{jt}}$ is a vector of observed brand characteristics, including the average initial price of brand j and brand intercepts, to capture heterogeneity in the brand attractiveness that is fixed over time. α_i and o_i are individual-specific parameters reflecting consumers' responsiveness to the number of deals and average depth of discount offered by a merchant, and $\boldsymbol{\beta_i}$ is a vector of individual-specific parameters capturing effects of observed brand characteristics. Note that the initial prices are prices printed on product labels and are determined before a product is sold in any stores, include on the focal retail deal platform. Therefore, we treat them as exogenous. We use a Hierarchical Bayes formulation to account for unobserved heterogeneity in these parameters across consumers, and assume $\boldsymbol{\Lambda}_i = (\alpha_i, o_i, \boldsymbol{\beta_i}) \sim MVN(\overline{\boldsymbol{\Lambda}}, \boldsymbol{\Sigma_\Lambda})$.

The term $(\zeta_{igt} + (1 - \lambda)\nu_{ijt})$ captures the random shock to the utility function and is assumed to follow an extreme value distribution, where ζ_{igt} is common to all brands in group g, and has a distribution function that depends on λ, and ν_{ijt} follows an i.i.d. extreme value distribution (Berry 1994). Each j belongs to either the outside group ($j = 0$) or the inside group ($j \in \{1,2, \dots, J\}$). This specification leads to the nested logit formulation which allows the degree of substitutability among brands in the category to be greater than that between any brand and the outside option (Villas-Boas and Zhao 2005). $\lambda \in [0,1)$ is the inclusive-value parameter and represents the degree of similarity among the inside brands. When $\lambda = 0$, the model reduces to a standard multinomial logit formulation (Berry 1994). We use a logit transformation in the estimation to ensure that $\lambda \in [0,1)$. The indirect utility of choosing the outside option $j = 0$ is set to $u_{i0t} = 0 + \nu_{i0t}$ for identification purposes.

$\boldsymbol{\xi_t}$ is a vector of ξ_{jt}'s, unobserved merchant/brand-specific shocks (to the researcher) that vary over time, due to factors such as display order of deals on the platform pages, and $\boldsymbol{\xi_t}$ is assumed to follow $\boldsymbol{\xi_t} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma_d})$. By the set-up of decision timings in the game, the number of deals at t is determined at t-1, and thus not correlated with the common demand shock, ξ_{jt}, while the average depth of discount is determined at t and thus may be correlated with ξ_{jt}. To address the potential endogeneity of the average depth of discount, d_{jt}, we use each merchant's average depth of discounts in other categories as an instrument for the depth of discounts. Our justification for using this instrument is that it is correlated with the depth of discounts of the focal category due to a merchant's common marginal cost structure across categories, which can affect the depth of discounts, yet it is unlikely to be correlated with idiosyncratic demand shocks of the focal category in a given week because consumers' needs and benefits sought outside the broad Women's Top merchandise category are likely to be different. We estimate the model via a limited information MCMC approach, following Yang, Chen, and Allenby (2003, p.272), and use diffuse priors in the estimation procedures.

Given the above utility specification, the probability of consumer i choosing brand j in week t is (see Villas-Boas and Zhao 2005, p. 86):

$$s_{ijt} = \frac{exp\left[\frac{V_{ijt}}{1-\lambda}\right]}{\left(\Sigma_{j\in g} exp\left[\frac{V_{ijt}}{1-\lambda}\right]\right)^{\lambda}\left[\Sigma_g\left(\Sigma_{j\in g} exp\left[\frac{V_{ijt}}{1-\lambda}\right]\right)^{(1-\lambda)}\right]}, \tag{3}$$

where $V_{ijt} = \alpha_i d_{jt} + o_i n_{jt} + \boldsymbol{x_{jt}\beta_i} + \xi_{jt}$. The aggregated share for each brand is $S_{jt} = 1/A_t \cdot \sum_i s_{ijt}$, where A_t is the number of active customers who browsed the category in week t.

Merchants' Profit Functions

Each merchant, j, makes decisions to maximize the net present value of its profit stream over an infinite time horizon, $E_t\left(\Sigma_{l=0}^{\infty}\delta^l \Pi_{jt+l}\right)$, where $\delta \in (0,1)$ represents the discount factor. In each week t, merchants compete on the depth of discount given the states and earn the profits of the week. The net profit of a merchant j in week t, Π_{jt}, is specified as follows:

$$\Pi_{jt} = GP_{jt}\left(d_{jt}, \boldsymbol{n}_t, \boldsymbol{x_t}, A_t\right) - IC\left(n_{jt}\right) - EC\left(a_{jt}\right) + \epsilon_{jt}\left(a_{jt}\right), \tag{4}$$

where GP_{jt} is the gross profit computed as:

$$GP_{jt} = \left[(1-\phi)\cdot p_{jt}\cdot(1-d_{jt}) - mc_{jt}\right]A_t S_{jt}. \tag{5}$$

In Equation (5), ϕ is the commission rate charged by the platform, p_{jt} is the average initial price across deals offered by merchant j, d_{jt} is the average depth of discount across the offered deals, and mc_{jt} is merchant j's marginal costs in week t. A_t is the number of active customers on the platform in week t, and S_{jt} is the share of customers who purchase from merchant j in week t, which is determined by a demand model to be described in the previous section.[4] The marginal costs, mc_{jt}, involve costs of merchandise and direct product costs associated with shipping and handling. Following Berry et al. (1995), we assume that the log of a merchant's marginal costs is a linear function of the observed component (by the researcher) of deals offered by a given merchant ($\boldsymbol{W_{jt}}$) and an unobserved component (ω_{jt}):

$$\ln\,(mc_{jt}) = \boldsymbol{x_{jt}\gamma} + \omega_{jt}, \tag{6}$$

where $\boldsymbol{x_{jt}}$ are the observed deal characteristics. In the net profit function, Equation (4), $IC_{jt}(.)$ represents the internal administrative costs if merchant j offers any deals on the platform in week t, and includes warehousing and book-keeping costs incurred for those products. It is specified

[4] We assume that each purchaser in a given week buys one product from the focal category. Our data showed that this was indeed the case for most consumers.

as a function of the number of deals on the platform, $IC(n_{jt}) = \theta_1 \cdot n_{jt}$, where θ_1 is the internal administrative cost parameter.

In each week t, a merchant decides the number of deals to be offered in the next week and must work with the platform if it chooses to increase its number of deals prior to $t + 1$. Let $a_{jt} \in \{-\overline{N}, \dots, 0, 1, 2, \dots, \overline{N}\}$ denote the change in the number of deals by merchant j from t to $t + 1$. We denote the merchants' decisions by the vector $\boldsymbol{a_t} \equiv \{a_{jt}: j = 1, 2, \dots, N\}$. By definition, $n_{jt+1} = n_{jt} + a_{jt}$, which is the total number of deals offered by merchant j in week $t + 1$. We specify $EC(a_{jt})$, the costs of increasing the number of deals, as follows:

$$EC(a_{jt}) = \begin{cases} \theta_2 \cdot a_{jt}, & if\ a_{jt} > 0 \\ 0, & otherwise \end{cases} \tag{7}$$

where θ_2 represents the marginal cost of increasing the number of deals[5]. $EC(a_{jt})$ can be interpreted as the external administrative costs incurred by a merchant, including labor and logistics costs involved in working with the platform operator when increasing its number of deals. The cost parameters θ_1 and θ_2 will be estimated based on estimation results of other model components (more details later).

Lastly, $\epsilon_{jt}(a_{jt})$ is the private information shock to a merchant's profit from each possible change in the number of deals, which is assumed to follow independent and identical type-I extreme value distributions over merchants and time.[6] It captures the impact of all factors that may affect a merchant's action but are not included in the model, including unobserved heterogeneity in merchants' actions.

[5] To capture the possibility of diminishing marginal costs in internal and external administrative costs, we have tried the natural log specification of the cost terms and found it to be slightly inferior to the linear specification. In addition, θ_1 and θ_2 may vary across merchants. But due to data limitations, we assume and estimate common cost functions for all merchants.

[6] This assumption guarantees that the dynamic game has at least one equilibrium in pure strategies, although it does not rule out the possibility of multiple equilibria (Doraszelski and Satterthwaite 2010).

The net present value of a merchant's profit stream over an infinite time horizon affects how it strategically increases or decreases the number of deals. By increasing the number of deals, a merchant may contribute to the attractiveness of the platform and increase the number of active customers, and thus could increase the expected demand and profits in subsequent periods, which needs to be balanced with the consideration of incurring higher external administrative costs through $\theta_2 \cdot a_{jt}$ in the current period and higher internal administrative costs through $\theta_1 \cdot n_{jt}$ in future periods. Our model takes into account such strategic trade-offs.

Equilibrium Depth of Discount

As shown by the empirical estimation result of Equation (1) in the next section, the number of active customers is significantly affected by the number of deals on the platform but not the depth of discount. Therefore, we model merchants' decisions on the latter as outcomes of static optimizations, which is consistent with the conventional treatment of pricing decisions in the dynamic game model literature. In each period, a merchant chooses the depth of discount for the deals offered and competes for demand from consumers of mass A_t to maximize its (static) gross profit. To keep the model tractable, we model the average depth of discount across all deals (SKU's) offered by a merchant on the platform in week t as the decision variable. Merchant j at week t solves

$$\text{Max}_{d_{jt}} GP_{jt} = \text{Max}_{d_{jt}} \left[(1-\phi) \cdot p_{jt} \cdot (1-d_{jt}) - mc_{jt} \right] A_t S_{jt}. \tag{8}$$

Pierson et al. (2013) proves that under fairly general conditions, a unique Bertrand-Nash equilibrium exists under mixed multinomial logit demand functions. Under the unique pure-strategy Bertrand-Nash equilibrium, the equilibrium depth of discount in our setting must satisfy the following first-order condition (e.g., Nevo 2000):

$$-(1-\phi)\cdot p_{jt}\cdot S_{jt}+\left[(1-\phi)\cdot p_{jt}\cdot(1-d_{jt})-mc_{jt}\right]\frac{\partial S_{jt}}{\partial d_{jt}}=0. \tag{9}$$

Numerically solving for d_{jt} yields the equilibrium depth of discount, which will be used to calculate the single-period profit for each merchant.

Equilibrium Number of Deals

In our dynamic game model, merchants maximize the net present value (NPV) of discounted profits over an infinite time horizon given expectations about the optimal decisions made by the other merchants. At equilibrium, those expectations are fully consistent with the optimal decisions of all merchants in a category. In other words, the optimal strategy on the number of deals offered by each merchant takes into account the implication of their current decision on future reactions of competitors, which in turn can affect the future profits of merchants. We assume that merchants' strategies depend only on payoff-relevant state variables, which is the standard Markov Perfect Equilibrium (MPE) assumption made in the dynamic oligopoly game framework (Ericson and Pakes 1995). Following the literature on estimating dynamic discrete games (e.g., Aguirregabiria and Mira 2007, Bajari et al. 2007), we assume that $\boldsymbol{n_t}$, $\boldsymbol{x_t}$, and A_t are payoff-relevant state variables and common knowledge of all merchants, and ϵ_{jt} is private information to each merchant. For notation simplicity, let $\boldsymbol{s_t}$ be the vector of all observable states, $\boldsymbol{n_t}$, $\boldsymbol{x_t}$, and A_t.

Denote $\boldsymbol{\sigma}\equiv\{\sigma_j(\boldsymbol{s},\epsilon_j): j=1,2,\dots,J\}$, a vector of strategy functions for each merchant. An MPE for the game is defined as a vector of strategy functions $\boldsymbol{\sigma}$ such that each merchant's strategy maximizes the net present value of the merchant for each possible state $\{\boldsymbol{s},\epsilon_j\}$ given other merchants' optimal strategies over time. Let $V_j(\boldsymbol{s_t},\epsilon_{jt};\boldsymbol{\sigma})$ represent the value function for merchant j that uses its best response strategy against the other merchants who take an action based on their respective strategies in $\boldsymbol{\sigma}$. By Bellman's optimality principle (Bellman 1957),

$$V_j(\boldsymbol{s_t}, \epsilon_{jt}; \boldsymbol{\sigma}) = \max_{a_{jt}}\{\Pi_j(a_{jt}, \boldsymbol{s_t}, \epsilon_{jt}) + \delta E[V_j(\boldsymbol{s_{t+1}}, \epsilon_{jt+1}; \boldsymbol{\sigma})|a_{jt}, \boldsymbol{s_t}]\}. \tag{10}$$

For every merchant j and every state $\{\boldsymbol{s_t}, \epsilon_{jt}\}$, the set of strategies $\boldsymbol{\sigma}$ is an MPE if the following is satisfied:

$$\sigma_j(\boldsymbol{s_t}, \epsilon_{jt}) = \underset{a_{jt}}{argmax}\{\Pi_j(a_{jt}, \boldsymbol{s_t}, \epsilon_{jt}) + \delta E[V_j(\boldsymbol{s_{t+1}}, \epsilon_{jt+1}; \boldsymbol{\sigma})|a_{jt}, \boldsymbol{s_t}]\}. \tag{11}$$

That is, in equilibrium, each merchant maximizes its expected NPV, given expectations about the optimal decisions made by other merchants. An equilibrium specified in the model describes the dynamics of changes in the number of deals and, combined with the static equilibrium in the depth of discounts, the dynamics of the depth of discounts.

Estimation Strategy

In this section, we provide a brief description of the model estimation procedures. As described previously, the BBL estimator proposed by Bajari et al. (2007) is a recent development in the estimation method that mitigates the severe computational burden of estimating dynamic game models, and it relies on a two-step approach. Following prior examples by marketing academics (e.g., Yao and Mela 2011, Liu et al. 2016), we apply the BBL estimator to estimate our dynamic game model.

Our estimation procedure consists of the following two stages. In the first stage, we estimate all relevant parameters in the models except the internal and external administrative cost parameters. We start with estimating parameters in the consumer demand model in Equation (2). Next, we recover the marginal costs using the first order conditions of the equilibrium depth of discount in Equation (9). Based on the estimated marginal costs, we estimate parameters in the marginal cost function, Equation (6).

We also estimate the transition functions that describe how the state vector evolves, and the merchants' policy functions that map the state vectors to their actions, which together characterize the observed policy and will be used to simulate alternative policies in the next stage. One of the transition functions to be estimated is the market size model, Equation (1). The other is for the exogenous variable, the average initial price across deals for each merchant, which we use an auto-regression (AR1) estimator proposed by Arellano and Bond (1991).

We now describe the policy functions. Let $\boldsymbol{P} = \{P_j(a_j|\boldsymbol{s};\boldsymbol{\sigma}): \forall j, \boldsymbol{s} \in S\}$ be the vector of conditional choice probabilities (CCPs) associated with a vector of strategy functions $\boldsymbol{\sigma}$, where $P_j(a_j|\boldsymbol{s};\boldsymbol{\sigma})$ is the probability of merchant j changing the number of deals by a_j in the next period given states of $\boldsymbol{s}$:

$$P_j(a_j|\boldsymbol{s};\boldsymbol{\sigma}) = Pr(\sigma_j(\boldsymbol{s_j};\boldsymbol{\sigma}) = a_j|\boldsymbol{s}). \tag{12}$$

Since the change in the number of deals, $a_{jt} \in A$, takes on discrete values, we employ an ordered logit model to formulate its policy function. A merchant's choice is determined by:

$$a_{jt} = \begin{cases} a_1 & if\ a_{jt}^* \leq c_1, \\ a_2 & if\ c_1 \leq a_{jt}^* \leq c_2, \\ \ldots & \ldots \\ a_A & if\ c_{A-1} < a_{jt}^*, \end{cases} \tag{13}$$

where $a_1 < a_2 < \cdots < a_A$, and $c_1, c_2, \ldots, c_{A-1}$ are cutoff values to be estimated. The latent tendency to increase the number of deals, a_{jt}^*, is specified as:

$$a_{jt}^* = \psi_1 n_{jt} + \psi_2 \sum_{i\neq j}^{J} n_{it} + \psi_3 A_t + \overline{p}_{t,-j}\psi_4 + Merchant_j + \epsilon_{jt}, \tag{14}$$

where n_{jt} is the number of deals offered by merchant j, A_t is the number of active customers, $\overline{p}_{t,-j}$ is the average values of exogenous merchant characteristics across all merchants except the focal merchant j, $\boldsymbol{\psi} = \{\psi_1, \psi_2, \psi_3, \psi_4\}$ are their parameters, and $Merchant_j$'s are merchant-specific fixed effects to capture the differences in their baseline tendencies. In addition, we use a summary measure of merchant j's competitive states, $\sum_{i \neq j}^{J} n_{it}$, which approximates the other merchants' states in a limited information nature and reflects the notion that merchants are more likely to monitor the overall states across its competitors instead of each individual competitor's activities, because the latter would involve much more efforts and higher costs and may be infeasible. Previous studies that estimated dynamic oligopoly games (e.g., Yao and Mela 2011, Liu et al. 2016) used similar aggregated measures of competitive states. Finally, ϵ_{jt} is an independent and logistically distributed error term. These estimates in the first stage enable us to compute the equilibrium gross profits for each merchant given any state vectors in week t.

In the second stage, we recover the internal and external administrative costs parameters (θ_1 and θ_2), by assuming that the observed policy of the merchants (estimated in the first stage) is the most profitable one over (a large number of) other simulated alternative policies. The parameters are estimated by minimizing the average squared difference of the predicted profit under between the observed policy and simulated policies. The intuition of this procedure is to find parameter values that best rationalize the estimated policy functions given the states, such that the internal and external administrative cost parameters yield the highest expected profits compared to under other simulated policies of a focal merchant that deviate from the estimated policy while assuming that the other merchants use the estimated policies. Details of the procedure are described in Appendix A.

Identification

We now discuss identification of the parameters in our models. The linear parameters in the demand model are identified through cross-sectional and intertemporal variations in consumers' brand choice decisions in response to the depth of discounts, the number of deals, and the other observed merchant characteristics. The diagonal elements of the variance-covariance matrix, which represent the extent of unobserved consumer heterogeneity, are identified through variations in consumers' responses to changes in the inter-temporal merchant characteristics. The non-diagonal elements of the variance-covariance matrix, which capture how responses to merchant characteristics are correlated pair-wise, are identified through correlations in consumer utilities derived from the corresponding pair of merchant characteristics. The nested logit inclusive value parameter is identified through variations in consumers' choice decisions among merchants of the inside group as well as the share of the outside good in response to changes in merchant characteristics. Parameters in the state transition of the exogenous merchant characteristics are identified through auto-regression (AR1) processes. Parameters in the merchants' policy functions, except the merchant fixed effects, are identified through correlations between changes in the number of deals and the state variables. The merchant fixed effects are identified through variations in changes in the number of deals across the merchants. Similarly, parameters in the model of the number of active customer are identified through changes in inter-temporal variations in the number of active customers, the total number of deals, and the depth of discount. The identification of the dynamic structural parameters, those of the internal and external administrative costs, relies on the inter-temporal variations in the number of deals and the optimality condition of the equilibrium specified in the previous section.

ESTIMATION RESULTS

The Market Size Model

The parameter estimates of the number of active customer model are presented in Table 2. We find that the total number of deals at the previous time period has a significant and positive effect on the number of active customers, while the effect of the average depth of discount at time $t-1$ is not significant. The result shows the existence of indirect network effects of merchants' decisions on consumer participations on the platform. It implies that the number of deals on the platform affects not only merchants' current-period profits but also their expected future profit streams via its impact on the market size, and thus should be treated as a dynamic optimization decision as we have formulated. The empirical result also supports our modeling approach that treats the depth of discount as a static decision variable.

Table 2. Estimates of the Market Size Model (Number of Active Customers)

Variable/Parameter	***Estimate***	***Standard Error***
Constant	-17.048	55.802
$\Sigma_{j=1}^{J} n_{jt-1}$	**1.179**	0.257
$\bar{d}_{t-1}$	128.527	115.834

Note: The bold font indicates that p-value < 0.05.

The Demand Model

In order to identify the individual-level parameters, we chose to include only those customers who made at least two purchases in the Female-Top category over the data period, which yields 324 customers in the estimation data. They made purchases from the five brands in 1,398 of the 3,416 platform visits. Conditional on category purchases, purchase incidence shares for brands A, B, C, D, and E are 19%, 28%, 11%, 29%, and 13%, respectively. Convergence of all parameters is achieved within 300,000 iterations. We use the first 50,000 as burn-in, with a thinning of one in every 50 draws to minimize auto-correlation. We present the posterior means and standard deviations of the (hyper-) parameters in the nested logit demand model in Table 3.

Table 3. Demand Model Estimates

Variable/Parameter	***Posterior Mean [a]***	***Posterior SD***
Merchant A	**-1.860**	0.632

Merchant B	**-1.435**	0.411
Merchant C	**-3.759**	0.766
Merchant D	**-1.067**	0.476
Merchant E	**-2.820**	0.393
Initial price (in 1,000 yuan)	**-0.689**	0.375
(Instrumented) Average depth of discount	**1.712**	0.979
Number of deals (in 10 SKUs)	0.027	0.095
Inclusive-value, λ	**0.141**	0.025
Number of observations	20,496	

[a]: Bold font indicates that the 95% credible interval does not include zero.

As expected, we find a negative and "significant"[7] effect of the average initial price on the probability of choosing a brand/merchant. The instrumented depth of discount shows a positive and "significant" effect on the choice probability, while the effect of the number of deals is positive (as expected) but not "significant". The posterior mean of the inclusive-value parameter λ is 0.141, and its 95% credible interval does not contain zero. This indicates that the degree of similarity among the merchants included in the model is greater than that between a merchant and the outside option, which supports the nested logit formulation over the standard logit model in our application. The covariance matrix of the distribution of heterogeneity indicates that there was a fairly substantial degree of heterogeneity in the brand preference and responsiveness to merchant actions among consumers. (The variance-covariance matrix of the heterogeneity distribution is presented in Appendix B.)

Estimates of Marginal Costs and Initial Price Transition Parameters

Based on the demand model parameter estimates and the Bertrand-Nash equilibrium condition for the depth of discount specified in the previous section, we can impute estimates of marginal costs, which are then used as the dependent variable in the marginal cost function (Equation 6) to estimate the parameters. The estimate results are presented in Table 4. As expected, the marginal

[7] "Significant" means that the 95% credible interval of the parameter estimate does not include zero.

costs increase with the average initial price. We also find that the number of deals offered by a merchant has a positive effect on its marginal costs, suggesting decreasing return to scale. A plausible reason for this is that, as a merchant offers more deals on the platform, congestions in the fullfillment process may occur in handling orders and customer inquiries.

For the transition function of the average initial price over all deals for each merchant, we employ the auto-regression estimator including merchant fixed effects proposed by Arellano and Bond (1991). The AR1 estimate for the average initial price is 0.936 with a standard error of 0.02.

Table 4. Marginal Cost Function Estimates

Variable/Parameter	***Estimate***	***Standard Error***
Constant	**-1.518**	0.010
Merchant A	**-0.606**	0.019
Merchant B	-0.015	0.011
Merchants C	**-0.012**	0.003
Merchant D	**-0.421**	0.020
Average initial price (in 1,000 yuan)	**0.852**	0.007
Number of deals (in 10 SKUs)	**0.070**	0.001

Note: The bold font indicates p-value < 0.05.

Policy Function Estimates

Our data showed that most changes in a merchants' number of deals from $t-1$ to t were within five deals.[8] Thus, we set the range of the merchants' decision options in the policy function (Equation 13) to $A = \{-9 \text{ or lower}, -8, -7, \dots, 0, \dots, 7, 8, 9 \text{ or higher}\}$. Table 5 presents the parameter estimates of the policy function, with the cut-off value estimates omitted to save space.[9]

Table 5. Policy Function Estimates

[8] Table A3 in Appendix B presents the changes in the number of deals by each merchant.

[9] We tried an alternative specification using B-splines, which allows for a more flexible approximation to the relationship between merchants' observed actions and states. The model with 5 knots shows the best fit. The result shows that the linear specification only slightly underperforms the specification using splines. Therefore, we use parameter estimates of the linear specification in the subsequent estimation procedures for their ease of interpretation.

Variable/Parameter	***Estimate*** [a]	***Standard Error***
Merchant A	**1.105**	0.415
Merchant B	0.148	0.395
Merchant C	**-0.720**	0.444
Merchant D	**0.796**	0.388
n_{jt}	**-0.912**	0.147
$\Sigma_{i \neq j}^{J} n_{it}$	**-0.174**	0.064
A_t	**0.008**	0.002
$\overline{p}_{t,-j}$	**1.609**	0.599
(Cut-off values omitted)		
Log Likelihood	-602.568	
BIC	1,422.510	

[a]: The bold font indicates p-value < 0.05.

The estimation result of the policy function can well explain the observed data pattern of merchants' number of deals as depicted in Figure 1. As shown in Table 5, Merchant A and D have the highest brand-specific intercepts (1.105 and 0.796), which are also the closed to each other in magnitude. Merchant B and E have similar brand-specific intercepts (0.148 and 0) which are not significantly different from each other. Brand C has the lowest brand-specific intercept (-0.720), which is also the farthest away from the others in magnitude. The parameter estimates for brand-specific intercepts not only explain the relative ranking in terms of the number of deals across brands, but also are consistent with the observed co-movement patterns among them (as shown in Figure 1), in particularly, Merchants A's and D's decisions were similar to each other's, Merchant B and E tended to move in the same way, and Merchant C had the lowest number of deals and moved in opposite fashions as the other merchants.

The policy function estimates suggest a few intuitions about the observed strategies across merchants. First, the incremental number of deals offered in the next period is negatively related to a merchant's current number of deals, which is likely due to the concern that the internal administrative costs in the subsequent period would increase with the number of deals offered in the current period. Second, the total number of deals offered by the other merchants negatively affects the incremental number of deals offered in the next period. A plausible reason for this is that consumers' attention to a particular merchant's deals may diminish when the total number of competing products on the platform is higher, and thus a merchant may be better off by offering fewer deals when its competitors are offering more. Third, the number of active customers on the platform positively affects a merchant's incremental number of deals to be offered in the next period, likely because higher customer traffic to the platform signals a greater profit potential in the future. Together with the finding of the positive effect of the number of deals on customer traffic, this confirms the existence of indirect network externalities on the platform from both consumers' and merchants' sides. Lastly, the incremental number of deals offered in the next period is positively associated with the average initial price of the other merchants, which suggests that merchants may see higher initial prices by competitors as an opportunity to introduce more deals as consumers may be more likely to switch away from higher-priced competing offerings.

Estimates of Internal and External Administrative Cost Parameters

As outlined in Section 4.7, we use estimation results in the first stage to conduct forward simulations in order to estimate the internal and external administrative cost parameters in Equation (7). Following Erdem and Keane (1996), we set the discount factor, δ, to 0.995 to compute the net present value, which has been shown to be reasonable for weekly data like ours (e.g., Sun 2005). We derive standard errors for the estimates using bootstrapping procedures as described in Bajari et al. (2007). Our C++ codes use parallel computing for forward simulations and yield substantial savings of the CPU time.

The estimation result is presented in Table 6. The parameter estimates are measured in thousands of Chinese yuan. We find that, for each additional deal placed on the platform, the incremental internal administration cost is 0.696 (approximately $108), while the incremental external administration cost is 1.034 (approximately $161). The estimates indicate that both sources of costs significantly affect merchants' profits.

Table 6. Estimates of the Internal and External Administrative Costs

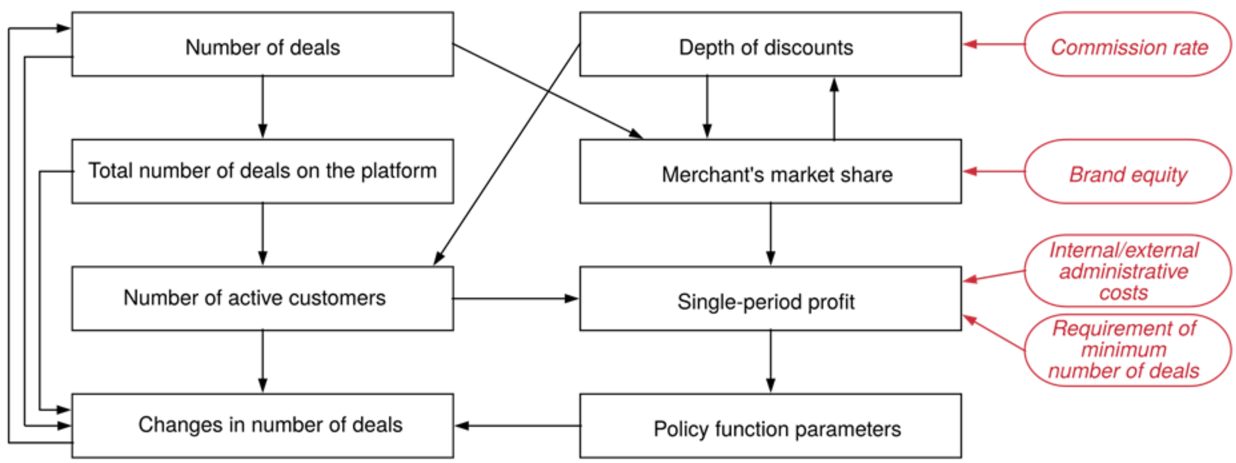
P***arameter***	***Estimate***	***Standard Error***
Internal administrative costs (θ_1)	**0.696**	0.160
External administrative costs (θ_2)	**1.034**	0.033

Note: The bold font indicates p-value < 0.05.

COUNTERFACTUAL SIMULATION ANALYSES

In this section, we investigate the impact on the merchants' and platform's payoffs due to changes in platform policies and market primitives through counterfactual simulations. These analyses enable us to address the managerial questions outlined previously, such as identifying mutually beneficial growth opportunities for merchants and the platform. We consider four sets of counterfactual scenarios. In the first set, we examine alternative commission policies. In the second set, we look at whether imposing a requirement of the minimum number of deals from each merchant facilitates or hinders the growth of the platform. In the third set, we examine the impact of merchants' efforts to reduce their internal and external administrative costs. And in the last set, we investigate changes in a merchant's brand equity, possibly due to marketing efforts or shifting consumer preferences. Each source of changes would lead to a series of reactions in the platform ecosystem due to competitions among merchants and indirect network effects between merchants and consumers. They will affect merchants' decisions on the depth of discount and number of deals to offer, the number of active customers on the platform, and demand for each merchant's deals, and thus eventually impact each merchant's profit as well as the platform's revenue. These complex dynamics require a comprehensive modeling framework like the one we propose in this study. Figure 3 provides a conceptual diagram of the chain reactions in the platform ecosystem that would be induced by each type of changes, which are captured by our models and counterfactual analyses.

Figure 3. **Chain Reactions in the Ecosystem Induced by Four Types of Changes**



For each scenario, we simulate merchants' new equilibrium policy on the number of deals to offer and the depth of discounts for 52 weeks, starting from the states at the first week in our data.[10] We compute the expected weekly profit for each merchant, expected weekly platform revenue, and the ecosystem payoff which is measured as the sum of platform revenue and total merchant profits.

Alternative Commission Policies

We examine two types of changes in the commission rate. First, the platform would change the current uniform commission rate charged to all merchants to a lower (3%) or higher (7%) level. Second, the platform would set commission rates based on merchants' market shares on the platform, to charge them with either higher or lower commission rates. Table 7 summarizes the

[10] Directly solving the equilibria is computationally difficult due to the many state variables and decision choices in our model. Instead, we adopt an approximate dynamic programming approach to deriving the equilibrium outcomes under each scenario, which has been applied by prior research in the literature (e.g., Yao and Mela 2011). The detailed procedures are described in Appendix C.

results of commission-related policy simulations, in which we report the incremental values of the outcome measures under each scenario relative to those under the current policy of 5% commission.

Table 7. Average Weekly Payoffs under Alternative Commission Policies

Commission Policy	# Active Customers	Total # of Deals	Merchant Profit (¥)					Total Merchant Profit (¥)	Platform Revenue (¥)	Eco-system Payoff (¥)
			A	B	C	D	E			
Current Practice (55555)	657.6	480.4	2766.8	3743	920.3	6085.5	627.7	14143.2	1404.1	15547.3
33333	-2.2	-1.7	+108.5	+24.4	+2.8	-255.2	+14.8	-104.8	-581.5	-686.3
77777	-31.4	-26.1	-305.9	-163.8	-61.8	-576.4	-21.3	-1129	-661.5	-1790.5
57575	-14.8	-12.7	+309.7	-158.1	-48.6	-570.7	-30.1	-497.8	+292.9	-204.8
53535	+34.6	+29.5	+278	+2838.3	-12.4	-515.4	-156.3	+2432.2	-261.8	+2170.4
55353	+28.2	+24.5	-14	+812.7	+379.5	-150.9	+14.1	+1041.5	+57.3	+1098.8

Notes: Each number in the first column represents commissions charged for each merchant. For example, 53535 represents 5%, 3%, 5%, 3%, and 5% commissions for merchant A, B, C, D, and E, respectively.

The scenario of lowering the uniform commission rate to 3% reveals several surprising and interesting results. We find that the total merchant profit, the platform revenue, and thus the ecosystem payoff all decrease, and the reason is that the lower commission rate would lead to more intense competition among merchants as shown by our detailed analyses. This result implies that, although lowering the commission rate might be a well-intended strategy for a platform to incentivize merchant participation at the expanse of its own commission revenue, it would actually intensify competition among merchants, which could drastically offset the direct benefit to merchants from the lower commission rate. It indicates that it is critically important to take into consideration of competition among merchants when designing and/or evaluating platform commission policies. On the other hand, the platform may intend to extract more revenues directly from all merchants by slightly raising the uniform commission rate, for example to 7% (“77777”). Our counterfactual simulation results indicate, however, that this would make all parties involved worse off, including the platform. We find that the total merchant profit and platform revenue would decrease and so would the ecosystem payoff. More detailed analysis shows that all merchants would reduce their respective number of deals and average depth of discounts, which would in turn reduce the number of active customers, total revenues across merchants, and thus result in lower platform revenues. These findings speak to the importance of taking into account indirect network effects when designing and/or evaluating platform commission policies.

The platform may consider charging merchants of higher market shares with higher commission rate to extract more revenue from them. To test the implication of this commission policy, we examine a scenario in which the commission rate of only the top two selling merchants (B and D) is raised to 7% (“57575”), and find that with the higher commission rate, merchants B and D would lower their total number of deals and the depth of discounts on the platform, which in turn would reduce the number of active customers on the platform. Overall, the ecosystem payoff would decrease, which is primarily driven by profit losses of the two top market share merchants. Among the five commission rate based policies examined, this one generates the greatest amount of gain in the platform’s direct revenue, but it would hurt merchants’ incentives, which in turn would hurt the attractiveness of the platform and customer participation. The key reason is that higher share merchants exercise stronger indirect network effect. When their number of deals or depth of discount drops due to higher commission rate, it would bring a greater negative impact to the whole ecosystem.

For the scenario in which the commission rate for the top two selling merchants (B and D) is lowered to 3% and that for the other merchants (A, C, and E) remains at 5% (“53535”), we find that it would increase the total merchant profits (primarily from Merchant B) by an amount that outweighs the loss in the platform’s direct revenues and thus increase the total ecosystem payoff. It is interesting to note that the two merchants with the preferential commission rates do not necessarily benefit from this commission structure. Our results indicate that Merchant B's profit would increase while merchant D's profit would decrease. A key reason is that the lower commission rate intensifies the competition between merchants B and D, and merchant B is a stronger competitor between the two in terms of offering more deals under the reduced commission rates. Among the five alternative commission policies examined, this one generates the highest gain in the ecosystem payoff. This suggests an opportunity to achieve mutual benefits for all the merchants and the platform through a profit-sharing arrangement, such that every party's payoff would be higher than that under the current practice. Another differential commission structure is to offer preferential rates to merchants with smaller market shares, in which we lowered the rate to 3% for the two smallest merchants (C and E) among the five ("55353"). We find that this policy can be another effective way to increase the total ecosystem payoff and make all merchants and the platform better off via a profit-sharing agreement, although the incremental benefit to the system is lower than the previous scenario.

Requirement of Minimum Number of Deals

A potential strategy for retail deal platforms like ours to increase the amount of merchandise available on their platform is to impose a requirement on the minimum number of deals offered by each merchant. Such requirement could effectively increase the number of deals offered and thus stimulate the growth of customer base, especially in the early stage of a platform. On the other hand, if the requirement is too stringent, it may force some merchants to opt out of the platform and thus result in the opposite effect. Our context is an ideal setting to test the impact of adopting such strategy because our dataset starts at an early period of the platform. We examine three counterfactual scenarios in which the platform requires merchants to offer a minimum number of deals in any given week, which we set to 10, 20 and 30. Under each scenario, merchants′ decisions on the number of deals in the next period are subject to the minimum number constraint. A merchant would offer the optimal number of deals dictated by the policy function if the constraint is satisfied. Otherwise, it chooses to either opt out from the platform or offer the required minimum number of deals in the next period, by comparing future profit streams under these two options. In Table 8, we report the simulation results of those scenarios over 52 weeks to evaluate the impact of these policies at the early stage of the platform.[11] Note that, given our empirical focus on the Women's Top category, the results reported here should be interpreted as impact on the growth of the category on the platform.

Table 8. Average Weekly Payoffs with Requirements of Minimum Number of Deals

Scenario	# Active Customers	Total # of Deals	Merchant Profit (¥)					Total Merchant Profit (¥)	Platform Revenue (¥)	Eco-system Payoff (¥)
			A	B	C	D	E			
Current Practice	657.6	480.4	2766.8	3743	920.3	6085.5	627.7	14143.2	1404.1	15547.3
Minimum 10	+133.9	+112.9	+1628.2	+2204.8	+341.4	+3165.2	-190.7	+7148.9	+1115.1	+8264
Minimum 20	+135.7	+115.9	+1624.6	+2300.3	+340.8	+3202.3	-195.9	+7272.2	+1135.4	+8407.6
Minimum 30	+135	+116.5	+1631.5	+2281	+332.1	+3162.5	-197.4	+7209.7	+1126.9	+8336.7

Our results reveal two interesting findings. First, imposing a required minimum number of deals could be an effective policy to benefit both merchants and the platform compared to the commission-based policy changes examined above. In all three cases involving a requirement of the minimum number of deals, we see an increase in the total number of available deals, number of active customers, total merchant profits, the platform's revenue, and thus the ecosystem payoff. Our detailed analysis indicates that this is primarily driven by the indirect network effect, especially during the early weeks of the platform's development. As the requirement would

[11] We also conducted the counterfactual simulations over 156 weeks and the results are very similar. The detailed results are available upon request.

increase the total number of deals on the platform during these weeks, more consumers would visit the platform during those periods, which in turn leads to a substantially larger customer base sustained into later periods. Even though the required minimum number of deals is sub-optimal for some merchants, they would all benefit from the much larger market size which is brought about by all merchants collectively complying with the minimum number of deals requirement. We find that, while merchants affected by the minimum number requirement tend to opt out in the very early weeks, they would choose to stay on the platform and comply with the requirement as time goes by. The requirement would no longer be binding after week 10 for all merchants. Second, the marginal increment of merchants' profits goes down as a more stringent minimum requirement is imposed. This is due to a tradeoff between merchants' decision on whether to comply with the requirement or opt out of the platform in a given week. As the requirement becomes more stringent, merchants are more likely to opt out of the platform instead of meeting the minimum requirement. In summary, imposing a required minimum number of deals can create win-win opportunities for both the platform and merchants, but managers should conduct careful analyses to determine the specific level of the minimum number requirement suitable for their platforms.

Reductions in the Internal and External Administrative Costs

The internal administrative costs borne by a merchant include warehousing and book-keeping for products sold on the platform. The external administrative costs include the labor and logistics costs involved in working with the platform when additional deals are offered there. We consider two sets of counterfactual scenarios of reducing the internal and external administrative costs, respectively, which correspond to merchants' efforts to improve efficiency in their internal inventory management and joint efforts with the platform to streamline the coordination processes. We operationalize the cost reduction efforts by reducing the internal administrative cost parameter (θ_1) or the external administrative cost parameter (θ_2) to half, respectively, for one merchant at a time. Table 9 summarizes the results of the counterfactual simulations, in which we report the incremental weekly average number of active customers, total number of deals on the platform, profit of each merchant, total profit across merchants, platform revenue, and ecosystem payoff under each scenario, relative to those under the current practice.

Our results indicate that the cost reduction efforts by each merchant would lead to a greater number of deals on the platform, more active customers, higher platform revenue and ecosystem payoff, and would increase profits for almost all merchants. In general, with the exception of Merchant E, all parties involved in the ecosystem (consumers, merchants, and the platform) would benefit from a given merchant's cost reduction efforts. In addition, reducing external costs would yield greater gains in the total merchant profits and ecosystem payoffs than reducing internal costs. This is because the marginal external cost associated with an additional deal is greater than that for the marginal internal cost. Merchant E presents an interesting case. It would suffer a profit loss of about 30% when the internal administrative cost parameter is halved by one of the other four merchants, and the profit loss is only slightly lower when its own internal administrative cost parameter is halved. Our in-depth analysis shows that, when a given merchant reduces its internal costs, it would increase its number of deals, which would then lead to all other merchants to increase their respective number of deals. With the increased number of deals by all merchants, all except Merchant E would also increase their depth of discount, while Merchant E would decrease its depth of discount due to competition according to Equation (9), which would then reduce its sales and profit. In contrast, Merchant E would benefit from reduction of its own external administrative costs because, as shown by our follow-up analysis, the greater cost savings achieved via this route would allow it to overcome the competitive pressure and to offer a larger number of deals than in the case of internal administrative cost reduction, which would be high enough to boost its sales and profit.

We also find that the platform can benefit from any merchant's cost reduction efforts, which collectively would increase the number of deals on the platform and attract more customers. It suggests that a platform operator should help merchants engage in cost reduction initiatives, such as providing technical supports and streamlining the administration processes for working with merchants. In addition, the results shed light on how the platform may choose to work with individual merchants depending on the differential benefits each merchant may generate for the platform and/or the ecosystem. For example, in our analyses, helping Merchant A or E reduce its external administrative costs would lead to the greatest gains in the platform's revenue and the ecosystem payoff, and the platform may wish to give priority to these two merchants when streamlining the administration processes. Interestingly, Merchant A ranks the lowest among the five examined in terms of the average initial or sales prices, while Merchant E ranks the lowest in terms of market share on the platform. Contrary to conventional wisdom, the "start brands" that a platform should foster for a given initiative may not be the market share leaders or the most premium ones.

Table 9. Average Weekly Payoffs under Half Marginal Internal or External Administrative Cost

Scenario	# Active Customers	Total # of Deals	Merchant Profit (¥)					Total Merchant Profit (¥)	Platform Revenue (¥)	Eco-system Payoff (¥)
			A	B	C	D	E			
Current Practice	657.6	480.4	2766.8	3743	920.3	6085.5	627.7	14143.2	1404.1	15547.3
Half internal										
by A	+134.1	+112.7	+1649	+2290.2	+340.8	+3106.6	-193.1	+7193.6	+1112.1	+8305.7
by B	+129.6	+108.9	+1559.1	+2178.1	+335.1	+3016.1	-201.7	+6886.7	+1065.5	+7952.3
by C	+132.4	+111.3	+1616.7	+2233.7	+351.8	+3049.2	-193.4	+7058	+1090.5	+8148.5
by D	+133.4	+112.1	+1590.8	+2245.5	+341	+3206.3	-197.9	+7185.6	+1121.4	+8307
by E	+134.1	+112.7	+1626.7	+2224.6	+339.8	+3185	-179.9	+7196.2	+1121	+8317.1
Half external										

by A	+134.3	+113	+2352.2	+2261.4	+342.8	+3160.6	-193.2	+7923.8	+1120.1	+9043.9
by B	+123	+103.3	+1539.3	+2724.7	+313.3	+2826.8	-208.6	+7195.5	+1020.1	+8215.7
by C	+131.1	+110.2	+1563.3	+2208.5	+830.7	+3069.5	-197.4	+7474.7	+1084.5	+8559.2
by D	+130.5	+109.7	+1570.5	+2231.5	+327.5	+3721.4	-195.9	+7655	+1082.5	+8737.5
by E	+133.6	+112.3	+1612.2	+2261.4	+340.2	+3138.1	+347	+7699	+1112.8	+8811.8

Changes in the Brand Equity of a Given Merchant

A merchant's brand equity affects consumers' demand for products offered by the merchant and thus influences merchants' strategic decisions and their profits and platform revenue. Following Goldfarb, Lu, and Moorthy (2009), we view merchants' brand equity as driven by distinctive branding elements such as user and usage imagery, emotional and status benefits, and brand personality. This equity is captured in the merchant (brand) intercepts in the demand model specified in Equation (2). We follow the methodology developed by Goldfarb et al. (2009) to evaluate the impact due to changes in brand equity at the merchant level by taking into account of competitors' reactions, and focus on two counterfactuals involving the top (D) and bottom (C) selling merchants in our estimation data. For the first counterfactual, we increase the posterior mean of merchant C's brand constant ($\beta_{i,merchant\ C}$) to that of merchant D's brand constant ($\beta_{i,merchant\ D}$) in Equation (2). For the second counterfactual, we decrease the posterior mean of merchant D's brand constant ($\beta_{i,merchant\ D}$) to that of merchant C's brand constant ($\beta_{i,merchant\ C}$). For both counterfactuals, estimates of the other merchants' constants remain the same as reported in Table 3.

Table 10 summarizes the results of the two counterfactual simulations. Our analyses indicate that, if merchant C's brand equity is increased to the level of merchant D's, it would enjoy a substantial gain in profit. More interestingly, all merchants would increase their numbers of deals offered, which would in turn increase the number of active customers on the platform. In addition, the platform's revenue and the total ecosystem payoff would also be increased substantially. In contrast, if merchant D's equity is reduced to the level of merchant C's, all parties involved in the ecosystem would be worse off, including the other four merchants who are direct competitors of merchant D. The key reason is that, due to merchant D's prominent position in the category, the assumed drop in its brand equity would lead to a greater reduction in its number of deals than the initial increase in the number of deals by the other merchants, leading to a decrease in the total number of deals on the platform and thus other subsequent negative consequences for all merchants and the platform. This result is consistent with the finding by Goldfarb et al. (2009) that changes in brand equity not only affect market shares of the brands involved but also the category sales in a static equilibrium framework. In our case, the negative impact on merchant and platform payoffs due to the lowered category share is amplified over time periods because of the indirect network effects in a dynamic game framework.

Our findings from these two counterfactual simulations suggest that merchant equity plays a key role in their competitive position by affecting customers' demand for a merchant's deals and their strategic decisions, which in turn impact the payoffs of the other merchants and the platform. It implies that the platform may wish to actively engage in merchants' brand equity management for the benefits of all parties involved in the platform ecosystem. For example, the platform could monitor the quality of the products listed, and customer services and communications offered by each merchant.

Table 10. Average Weekly Payoffs Resulting from Brand Equity Changes

Scenario	# Active Customers	Total # of Deals	Merchant Profit (¥)					Total Merchant Profit (¥)	Platform Revenue (¥)	Eco-system Payoff (¥)
			A	B	C	D	E			
Current Practice	657.6	480.4	2766.8	3743	920.3	6085.5	627.7	14143.2	1404.1	15547.3
C→D	+116.1	+101.4	+466.8	+108.3	+8017.2	+1140.5	-948.1	+8784.7	+1201.8	+9986.4
D→C	-457.6	-383.1	-2208.8	-2234.8	-533.2	-6159.6	-405.3	-11541.7	-1327.9	-12869.6

SUMMARY AND DISCUSSIONS

This study expands the literature on platform marketing to a new type of context, retail deal platforms, where merchants make repeated merchandising decisions and consumers can purchase deals on a regular basis. We focus on merchants' decisions on the number of deals and depth of discount by considering their strategic interactions in a competitive environment and the two-sided market nature of the platform. We propose a comprehensive model framework that incorporates merchants' forward-looking considerations, strategic interactions among each other, and indirect network effects between merchants and consumers. We estimate the models using a unique dataset from a mobile shopping platform specializing in fashion products.

Our model estimation results suggest that merchants and platform managers need to take into account two key forces when designing and/or evaluating their decisions or platform policies: indirect network effects and strategic considerations among merchants. First, merchants' decisions on the number of deals and depth of discounts affect consumer demand and the platform traffic, which in turn affect merchants' decisions on what to offer on the platform. These findings together indicate that indirect network effects are a key force on a retail deal platform. Second, merchants strategically consider each other's decisions when making their own decisions on the platform. We find that a merchant would reduce its number of deals if the number of deals by competitors was higher, and vice versa. This is another important force that drives the dynamics in the platform ecosystem.

The proposed modeling framework enables us to examine a variety of policy changes via counterfactual simulations, which incorporate the complex chain reactions in the ecosystem brought about by changes from each type of players (consumers, merchants, or the platform). We have examined the profit/revenue implications for merchants and the platform resulting from alternative commission policies, imposing a requirement on the minimum number of deals, and merchants' cost reductions and changes in brand equities. Our analyses have identified multiple opportunities to improve the platform's commission policies that would benefit all parties involved. For example, offering a preferential commission rate to top selling merchants appears to be the most beneficial commission-based policy among those examined in increasing the total ecosystem payoff, and thus create a win-win opportunity via profit-sharing arrangements. We believe that profit-sharing arrangements are implementable because the platform processes payments for and keeps transaction records of each merchant. Furthermore, imposing a requirement of the minimum number of deals is an effective way to grow a platform and enables merchants to benefit from each other's actions via indirect network effects, which in turn would benefit consumers and the platform operator. In addition, we find that the requirement of minimum number of deals appears to be more effective than commission-based policies in bringing mutual benefits to the ecosystem, in the early stage of a platform's development.

Our analyses reveal several novel insights which would not have been possible without considering the competition among merchants and network effects due to the two-sided market. We find that cost reduction efforts or brand equity enhancement by individual merchants not only benefit the focal merchant but can generate substantial gain to the ecosystem and their competitors. Due to the indirect network effects, instead of being "enemies", merchants selling similar products on a platform could bring mutual benefits even though they directly compete with each other. Interestingly, the merchants whose cost reduction effort brings about the greatest gain to the ecosystem are smaller brands in terms of market share or price premium. Another intriguing finding is that not all merchants may benefit from platform policies which intend to help them, due to the strategic interactions among merchants. For example, one may expect all merchants to enjoy profit gains in the scenario where the platform reduces the commission rate to 3%, yet our analysis shows that the lower commission rate could intensify competitions among merchants and lead to lower total merchant profit and platform revenue and a smaller number of deals available to consumers.

Our study also sheds light for retail platform operators on potential trade-offs between shorter term revenues and the well-being of the ecosystem in the longer term. For instance, if a platform aims at increasing its direct commission revenues, milking merchants of higher market shares with a higher commission rate would be the most effective commission policy, although it comes at the expense of lowering total merchant profit and ecosystem payoffs. This strategy would not be desirable for the platform in the long run. In contrast, offering preferential commission rates to high-market-share merchants could substantially increase the total merchant profit. Although the weaker merchants would incur a direct profit loss, the platform could facilitate a profit-sharing arrangement among merchants to make all of them better off, which in term should boost merchant loyalty and benefit the platform in the longer term. To conclude, we hope our study will stimulate more research about retail deal platforms in particular and the platform economy in general.

ESSAY II

Discovering Online Shopping Preference Structures in Large and Frequently Changing Assortments

INTRODUCTION

Many retailers carry large and frequently changing assortments. With the development of E-commerce, the assortment size of a typical online store or marketplace is even larger and growing steadily. Unlike for consumer-packaged goods, consumers hardly buy the same item multiple times in many types of merchandise, such as clothes, footwear, bags, fashion accessories, video games, music, movies, books, electronics, and gadgets, and retailers selling such merchandise need to frequently update their assortments. In addition to the large assortment size, this feature makes it even more difficult to predict consumers' demand for specific products and to prescribe effective marketing actions, such as making merchandise assortment decisions and offering personalized merchandise recommendations.

The general approach proposed in the marketing literature to study consumers' preference structures in fairly large assortments is to translate individual stock-keeping-units (SKUs) into relevant product attributes and then model purchase decisions with regard to those product attributes (e.g., Fader and Hardie 1996, Ho and Chong 2003, Inman et al. 2008). As is well-articulated in the seminal work by Fader and Hardie (1996), an advantage of this approach is a parsimonious model specification. Another advantage is to be able to make demand forecasts for products that have yet to be included in the assortment. In fact, conjoint analysis, which shares a similar feature in that it decomposes the overall product preference into attribute level part-worths, also offers such benefit and has played an important role for new product developments in marketing research and practices. Nonetheless, prior SKU-level models and their applications were almost exclusively constructed and estimated at the product category level, and the size of assortment that these models could handle was still quite limited, usually in the range of dozens to up to a few hundreds of SKUs.

Recent developments in the fields of natural language processing and machine learning, especially in the domain of topic models, have offered tools for a break-through in handling the scale of retail assortments. A natural analogy can be drawn between words and documents studied by natural language processing to products (or SKUs) purchased and a consumer's purchase history record in retailing data. Based on this analogy, Jacobs et al. (2016) adopted the Latent Dirichlet Allocation (LDA) topic model to study consumers' purchase behaviors concerning the entire assortment of an online grocery store, and investigated the underlying topics (which they called "motivations") that drove the purchase decisions. The scalability of their model is substantially greater than prior SKU-level models, with an empirical application involving about 400 "products" defined as category-brand combinations. Another advantage of their approach is to provide rich insights on consumers' preference structures in terms of plausible motivating factors. In the context of grocery products, their analyses revealed interesting underlying "motivations" such as preference for eco-friendly products, low-fat products, or products for sensitive skin.

While showing impressive improvement over prior SKU-level models in dealing with large assortments, the standard LDA models have several limitations which hinder their ability to assist managerial decisions facing retailers with large and frequently changing assortments. First, the meaning of an identified "topic"/"motivation" needs to be inferred from the estimated probability distribution over a large number of products and is based on the commonality of products that are loaded on the same topic, and thus it is often subject to interpretation and not always clearly defined.[12] Second, they cannot incorporate the effects of time-varying product

[12] As an illustration, suppose that a topic's probability distribution is loaded on baked potato chips, diet coke, low-fat ice cream, and fruit platter, it could be interpreted as preference for "party food", "healthy eating", or "sweet tooth".

attributes, such as marketing mix variables (e.g., price and discounts), which are important for many marketing and merchandising decisions. Third, they cannot provide demand forecast for products that do not exist in the calibration data of existing assortments (i.e., "out-of-vocabulary" products), yet such insights are critically important for managerial decisions by retailers/marketplaces with frequently changing assortments (such as fast fashion retailers). Fourth, the scalability of these models still needs further improvement to truly tackle purchase decisions at the SKU-level, because directly estimating Dirichlet distributions over a large number of products for each topic is a computationally challenging task and requires many rounds of tuning. Fifth, products in the assortment are treated as entirely different from each other and similarities among those sharing certain common attributes are ignored.

To address these limitations and provide a modeling approach better suited for frequently changing assortments, we build on the LDA framework to tackle the entire store assortment, while incorporating the idea of decomposing individual products (SKUs) into managerially relevant product attributes, as in the aforementioned SKU-level models. The basic idea is to construct a model with the underlying topics defined by probability distributions on product attributes instead of individual products. We adopt the terminology "topic" from the natural language processing literature in this study. A "topic" is similar to the underlying interest, need, objective, or goal that drive consumers' shopping behaviors. Our approach offers several advantages over previous methods. First, by translating individual products (SKUs) into a product attribute space of much lower dimensionality, it enables us to handle a much larger number of products and substantially improves the scalability of the model, as shown by our empirical application. Second, the topics revealed by our model are clearly defined by the relevant product attributes and offer directly actionable managerial insights. Third, it can easily incorporate the impact of time-varying marketing mix variables on shopping/purchase activities, because each of them can be simply treated as a product attribute in the model. Fourth, it enables us to make demand forecast for "out-of-vocabulary" products at the individual consumer level, which can then assist in a variety of important managerial decisions such as making personalized product recommendations.

We apply the proposed modeling framework to study the topics that drive consumers' online browsing and purchase activities concerning the entire store assortment. To this end, we construct a hierarchical LDA topic model which relates the topics driving the two levels of activities using the approach proposed by Liu and Toubia (2018). In addition, to enhance the managerial value of the model, we take into account the impact of time-varying situational and consumer characteristics on the relevance of each topic at each stage of the shopping process (browsing or purchase). We apply the proposed model to data from an online retail marketplace for fashion products. The dataset includes detailed records of products offered in the marketplace as well as individual consumers' browsing and purchase history information. Our model will reveal easy-to-interpret and directly actionable "fashion topics of interest" that drive shoppers' browsing vs. purchase activities and the time-varying patterns of their relevance to each individual, which can be utilized to make various merchandising decisions and product recommendations by online retailers or marketplace operators.

Personalized product recommendations are playing an increasingly important role in the context of large and frequently changing assortments in today's retail environment. For example, Stitch Fix, an online subscription-based personal styling service that delivers fashion products individually curated for its customers, generated $1.2 billion in net revenue in its fiscal year 2018 with earnings of $45 million (Smiley 2019). Understanding consumers' preference structures in large and frequently changing assortments is a key pillar for making personalized merchandise recommendations. As another contribution of this research, we propose a personalized product recommendation system based on individual consumers' preference structures concerning the entire store assortment, as revealed by our proposed model. We utilize the Bayesian optimization algorithm to address the challenge imposed by having to simultaneously search over a large space of discrete as well as continuous product attributes. As shown by our empirical analyses, the proposed method substantially outperforms benchmark algorithms commonly used in practice, such as user-based, item-based, and content-based collaborative filters, and is particularly powerful in making "out-of-vocabulary" product recommendations beyond the existing assortment. These recommendations offer valuable insights for assortment planning, new product development, and personalization.

To summarize, we intend to make the following contributions in this study. First, we develop a highly scalable stock-keeping-unit-level Latent Dirichlet Allocation model for inferring consumers' preference structures in large and frequently changing assortments at the store/marketplace level that overcomes several limitations of the standard LDA models. Second, we propose a personalized product recommendation system based on consumers' preference structures inferred from the entire store/marketplace assortment, which is capable of making recommendations of products not included in the calibration data of existing assortment and substantially outperforms personalized recommendation systems commonly used in practice. Third, our empirical analyses reveal novel insights on the underlying factors driving consumers' browsing and purchase activities in an online marketplace for fashion products and how they evolve with consumers' shopping experiences, which provide a range of managerial insights for e-commerce practitioners.

RELEVANT LITERATURE

Our research builds upon several streams of prior literature. On the methodology front, we build on the literature in 1) assortment and stock-keeping-unit-level (SKU) modeling, and 2) topic models in machine learning developed for natural language processing, which have also been applied by marketing academics. In terms of the substantive topic, it is related to the literature on understanding online shopping and purchase processes and personalized product recommendations. We will provide a brief review of each stream of research and outline how our study contributes to the prior literature in this section.

Assortment and Stock-Keeping-Unit-Level Modeling

To deal with sales response or consumer demand/choice at the SKU level, models proposed in the marketing literature have essentially adopted two approaches. The first approach explicitly contains SKU-specific intercept terms to capture preference for a SKU (e.g., Guadagni and Little 1983). A key limitation of this approach is that the number of parameters to be estimated grows proportionally to the number of SKUs, and thus it cannot well accommodate even a fairly small assortment, especially for consumer segment mixture models. The second approach is to define and parameterize the preference for a SKU as the sum of preference for each of its product attributes. A seminal work taking this approach is Fader and Hardie (1996). They propose a multinomial logit model of product choice based on a set of product attributes (e.g., brand name, package size, type, price) that characterizes a large number of SKUs in a category in a parsimonious manner. Compared to the first approach, their model prevents explosion in the number of parameters. They apply the model to data of fabric softeners and show that it outperforms models based on SKU-specific intercept terms. Moreover, this approach enables demand forecast for new products which are new combinations of product attributes included in the model, which provides valuable managerial insights on identifying line extension opportunities. A good number of studies have followed this basic approach and extended to incorporate other aspects of consumer decisions. For example, the model proposed by Ho and Chong (2003) captures the impact of consumers' attribute-level experience and product-specific experience. Inman et al. (2008) develop a model that identifies product attributes driving switching vs. repeat purchase behaviors, respectively. Rooderkerk et al. (2013) accommodate substitution patterns between SKUs by incorporating attribute-level similarity between SKUs in their model. Nonetheless, these models are best suited to handle the size of assortment at the product category level, because their scalability is usually limited to several dozens of SKUs. As a result, they are unable to capture co-occurrences of purchases across multiple product categories, let alone at the level of the entire store assortment. In this study, we develop a SKU-level model to study preference structures at the store/marketplace level, which offers substantially greater scalability (our empirical application of the model involves over 7,600 SKUs).

Latent Dirichlet Allocation (LDA) Topic Models

As mentioned previously, the LDA topic models originally developed for natural language processing offer substantial improvements in the scalability and have been applied to handle large assortment sizes (e.g., Jacobs et al. 2016). The LDA model is an unsupervised Bayesian learning algorithm developed in the computer science literature mainly used to extract latent "topics" by relating the words under a topic (see Blei et al. 2003 for details). There have been a growing number of applications of the LDA modeling framework in the marketing literature, primarily adopted to understand latent dimensions extracted from rich text data (e.g., Tirunillai and Tellis 2014, Puranam et al. 2017, Liu and Toubia 2018, Ansari et al. 2018, and Toubia et al. 2019). For examples, Tirunillai and Tellis (2014) extract latent dimensions of consumer satisfaction about brand qualities from consumer reviews (e.g., instability, portability, receptivity, compatibility, discomfort, secondary features), which provide insights on how brands compete on this multidimensional space of quality over time.

The topic models have been mainly applied to analyzing text data in the marketing literature, such as user reviews. An exception is Jacobs et al. (2016). They apply the LDA framework to study purchase decisions concerning the entire store assortment and to identify the underlying "motivations" that drive the purchases. The scalability is much enhanced over previous SKU-level models for category-level purchase decisions. In addition, their model outperforms several benchmark methods such as collaborative filtering and Dirichlet-Multinomial mixtures in its predictive capability. Despite the impressive improvement, the models applied in their study share the limitations of standard LDA models when it comes to studying consumers' preference structures in large assortments, as detailed in our Introduction section. In this study, we propose a new LDA model that overcomes these limitations. A prominent distinction of our proposed model with those adopted by Jacobs et al. (2016) is that our model is particularly suited to study consumer preferences in frequently changing assortments, while theirs cannot deal with changing assortments and are best suited for the context of fairly stable assortments (such as grocery and consumer packaged goods). In addition, we study the underlying "topics" for both browsing and purchase activities in online stores, and our analyses reveal the similarities and differences of the underlying product interests driving these two types of shopping activities and how they evolve with consumers' shopping experience.

Table 1 provides a comparison of the proposed model and our study to selected studies in the literature that model purchase decisions at the SKU level. In addition to further improved scalability, our framework reveals easy-to-interpret “topics” that are based on product attributes and incorporates the impact of time-varying marketing mix variables on shopping activities. Moreover, our model allows consumers’ preferences to change over time and captures the dynamics in topic relevance due to consumers’ shopping experience. Finally, in terms of the substantive insights, our study looks into the driving factors for both browsing and purchase activities.

Table 1. Overview of Selected Studies of SKU-level Models for Large Assortments.

Study	Assortment level	Number of SKUs	Represent SKUs on product attributes	Purchase Decision	Time-varying preference structures?	Demand forecast for new products?	Empirical context
Fader and Hardie (1996)	Category	56	Yes	Choice	No	Yes	Offline, fabric softeners
Ho and Chong (2003)	Category	38-421	Yes	Choice	Yes	Yes	Offline, 15 grocery categories
Inman et al. (2008)	Category	121	Yes	Choice	Yes	Yes	Offline, salad dressing
Jacobs et al. (2016)	Store	394 category*brand combinations	No	Purchase incidence	No	No	Online, grocery
This study	**Store**	**7658**	**Yes**	**Browsing & purchase incidence**	**Yes**	**Yes**	**Online, Fashion**

Online Shopping and Purchase Behaviors

Our research also contributes to the literature on online shopping and purchase behaviors. Prior research has extensively examined online shopping and purchase processes and/or decision outcomes (e.g., Bucklin and Sismeiro 2003, Danaher et al. 2003, Montgomery et al. 2004, Shi and Zhang 2014, just to name a few). For example, Bucklin and Sismeiro (2003) study online users' browsing behaviors on a website and found that visitors' propensity to continue browsing is affected by the depth of a given site visit and the number of repeat visits. Montgomery et al. (2004) focus on examining the paths of online users' browsing behaviors and found that a user's past path information is crucial in accurately predicting a subsequent path. Shi and Zhang (2014) investigate how online shoppers' usage experience with various decision aids influence the evolution of their store visit and purchase spending decisions over time.

In contrast, the substantive focus of our study is to identify the *underlying product interests* (which we call "topics") that drive online browsing and purchase activities concerning individual products in large and frequently changing assortments. These underlying factors are expressed in terms of probability distributions over clearly defined product attributes and can be interpreted as "topics" of fashion interests. They serve as a powerful diagnostic tool which can be used to make a variety of managerial decisions for (online) retailers, such as assortment decisions for a new season, fine-tuning cross-selling pitches for individual consumers, and as the basis for making personalized merchandise recommendations.

Personalized Product Recommendation Systems

Our study builds on and contributes to the stream of research on personalized product recommendations. A variety of personalization systems have been proposed for different contexts in the marketing literature. For example, Ansari et al. (2000) propose a recommendation system based on a hierarchical Bayesian approach that captures unobserved heterogeneity in consumer preferences and product attributes. They apply the proposed system to movie recommendations. Ansari and Mela (2003) develop a model for customizing the design and content of e-mails to increase web site traffic, which incorporates viewers' responses to email design attributes, content descriptors, and user characteristics. Zhang and Krishnamurthi (2004) propose a decision-support system that provides recommendations on when to promote how much (e.g., price cut) to whom, primarily for use in online stores. Their model accounts for competitive prices and promotions, as well as consumers' internal factors such as brand preferences and variety seeking/inertia tendency. Lu et al. (2016) construct a video-based automated recommender system (VAR) for apparel using real time in-store videos. While most of these previous personalized recommendation systems may not be easily scalable for large assortments due to computational burdens, our proposed framework is highly scalable and particularly suited for retailers with frequently changing assortments, because it is capable of making recommendations of products not included in the calibration data of existing assortments. A notable exception is the adaptive personalization approach for music or news articles proposed by Chung et al. (2009, 2016), which is also capable of making recommendations out of a large repertoire. Chung et al. (2016) further utilizes customer's social network structures for better personalization. Their model adaptively learns preferences based on implicit user behaviors (such as how long have songs been listened to or articles been read) rather than user ratings of those songs or articles. Similar to their approach, our model does not require explicit ratings of products by consumers either, and it only needs clickstream data on browsing and purchase behaviors. Also, both their approach and ours utilize information on product attributes (of songs or physical goods). Nonetheless, our approach differs from theirs in several ways. First, in terms of utilization of the attribute information, they adopt the Bayesian variable selection on attributes to pin down those that are most relevant for an individual. Our LDA approach focuses on discovering latent preferences over attributes. One advantage of our approach of inferring latent structures in preference is that our approach can directly generate dash-board insights on consumer preferences that marketers and platform managers can easily understand. Second, the scalability is much improved with our approach (e.g., 400 unique songs from 86 individuals vs. 7658 unique SKUS from 1738 individuals), which is possible largely due to the difference in estimation methods (MCMC vs. variation inference). Third, while the adaptive personalization approach by Chung and colleagues assumes fixed underlying preference by an individual and their algorithm adaptively learns the relevance of specific attributes, our approach not only incorporates a comprehensive depiction of individual-specific preference structures but also allows them to change over time, which can be affected by time-varying factors such as seasonality and previous shopping experience. Thus, the application of our approach is more general, and it is particularly useful for contexts such as apparels, footwear, bags, jewelry, cosmetics, and other fashion products, where consumer preferences tend to change over time.

MODEL FORMULATION AND ESTIMATION

Our model is built on the Latent Dirichlet Allocation (LDA) modeling framework, originally developed by Blei et al. (2003). The key idea is that consumers browse or purchase certain products in a shopping session (i.e., a store visit) at the focal online store as driven by a mixture of latent "topics" of interest, and the probability distributions over these topics represent their preference structures regarding browsing and purchase activities, respectively. Our goal is to discover consumers' preference structures based on each individual's browsing and purchase history data.

Let i denote a consumer, $i \in \{1, \ldots, I\}$, and j denote a product (i.e., a SKU) in the store assortment, $j \in \{1, \ldots, J\}$. Each visit made by a consumer to the online store defines a session, in which he/she browses certain products and possibly purchases some or even all of them. For each consumer i, we observe S_i sessions in total. The set of products browsed by consumer i in session s, $s = 1, \ldots, S_i$, is given by $\boldsymbol{d}_{is}^{[b]} = \{\boldsymbol{d}_{is1}^{[b]}, \ldots, j, \ldots, \boldsymbol{d}_{isN_{is}^{[b]}}^{[b]}\}$, where the superscript $[b]$ refers to browsing activities and $N_{is}^{[b]}$ is the number of unique products browsed in session s. Similarly, the set of products purchased by consumer i in session s is given by $\boldsymbol{d}_{is}^{[p]} = \{\boldsymbol{d}_{is1}^{[p]}, \ldots, j, \ldots, \boldsymbol{d}_{isN_{is}^{[p]}}^{[p]}\}$, where the superscript $[p]$ now refers to purchase activities and $N_{is}^{[p]}$ is the number of unique products purchased in session s.[13]

An innovation of our model is that each product is first translated into a representation on a product attribute space, and then topics are inferred as distributions over product attributes, instead of treating topics as probability distributions over a large number of products directly as in standard LDA models.

In the rest of this section, we first describe the product attribute space and how topics are defined on this space, followed by the formulation through which topics drive shopping activities. We then present the link between the browsing activities component and purchase activities component, and describe our estimation techniques.

Product Attribute Space

To capture consumers' preference over product attributes, we represent each product/SKU as a combination of a set of product attributes. For example, attributes of a fashion product can be the product's category, brand, gender of the target user (which we call "gender" for simplicity), color, initial price, and depth of discount. These attributes are captured in a vector $\boldsymbol{d}_j = (\boldsymbol{c}_j, \boldsymbol{y}_j)$. Here, $\boldsymbol{c}_j = [c_{j1}, \ldots, c_{ja}, \ldots, c_{jA}]$ contains discrete attributes of product j such as category, brand,

[13] In case no purchases are made in the session, $N_{is}^{[p]} = 0$ and hence $\boldsymbol{d}_{is}^{[p]} = \emptyset$. Note that by definition the purchased products in session s are a subset of the browsed products, i.e. $\boldsymbol{d}_{is}^{[p]} \subseteq \boldsymbol{d}_{is}^{[b]}$and it holds that $N_{is}^{[p]} \leq N_{is}^{[b]}$.

“gender”, and color, where a denotes a given attribute, the number of discrete attributes is A, and the number of levels for attribute a is denoted by $|a|$. The vector $\boldsymbol{y}_j$ contains continuous attributes of product j, such as initial price and depth of discount, and the number of continuous attributes is denoted by $|\boldsymbol{y}|$. Some of the product attributes can be time-varying, for example, the depth of discount in our empirical application.

Topics: Probability Distributions over Product Attributes

We assume that consumers’ browsing or purchase activities in a session are driven by a mixture of topics and each topic is meant to partially capture preferences for product attributes. As stated earlier, a novelty of our model is that topics are characterized by probability distributions over product attributes (instead of products as in standard LDA models), which substantially improves the interpretability of topics and scalability of the model. Each topic $m \in \{1,\ 2,\ \dots,\ M\}$, where M is the total number of topics, is characterized by a joint distribution involving $A + 1$ individual probability distributions: A probability mass functions, each corresponding to one of the A discrete product attributes, and a joint density over continuous product attributes such as initial price and the depth of discount. We assume that the probability vector for attribute levels of a given discrete product attribute a, denoted by $\boldsymbol{\phi}_m^{[a]}$, follows a Dirichlet distribution:

$$\boldsymbol{\phi}_m^{[a]} \sim \text{Dirichlet}_{|a|},$$

which reflects the relevance of each level of attribute a to topic m. For example, a topic that reflects preference for men’s layered outfits of “Blazers” and “Overcoats” by “Zara” and “H&M” would place high probabilities on “Male” for the “gender” attribute, “Blazers” and “Overcoats” for the “category” attribute, and “Zara” and “H&M” for the “brand” attribute.

Consumers’ preference for continuous product attributes under topic m is captured by a joint density function and assumed to follow a multivariate Normal distribution with mean $\boldsymbol{\mu}_m$

and covariance $\boldsymbol{\Sigma}_m$: $\boldsymbol{\phi}_m^{[y]} \sim \mathrm{Normal}_{|\boldsymbol{y}|}(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$. In our empirical application, the continuous attributes include the initial price and percentage discount of each product, both in the logarithm transformation.

Note that a consumer's preference structure in a given session is not captured by a single topic, but rather a mixture of topics, as described in the next sub-section.

How Topics Drive Shopping Activities

The salience of topics in driving the browsing activity of consumer *i* in session *s* is captured by the probability vector $\boldsymbol{\theta}_{is}^{[b]} = (\theta_{is1}^{[b]}, \dots, \theta_{ism}^{[b]}, \dots\ \theta_{isM}^{[b]})$, with each element $\theta_{ism}^{[b]}$ represents the relevance of topic m, and is assumed to follow a Dirichlet distribution. $\boldsymbol{\theta}_{is}^{[b]}$ reflects the mixture of topics that are most relevant in a given browsing session.

In the conventional specification of an LDA topic model (e.g., Blei et al. 2003), $\boldsymbol{\theta}_{is}^{[b]}$ is drawn from a straightforward Dirichlet distribution: $\boldsymbol{\theta}_{is}^{[b]} \sim \mathrm{Dirichlet}_M(\boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ is a vector of scalars to be estimated directly. However, this specification has two restrictions: 1) lack of systematic correlations among topics[14], and 2) no scalable inclusion of additional independent variables at the consumer-session level. To relax these two restrictions, we follow Jacobs et al. (2017) and use the following formulation:

$$\boldsymbol{\theta}_{is}^{[b]}(\boldsymbol{\alpha}_{is}) = \left(\frac{\exp(\alpha_{is1})}{\sum_{m=1}^{M} \exp(\alpha_{ism})}, \dots, \frac{\exp(\alpha_{ism})}{\sum_{m=1}^{M} \exp(\alpha_{ism})}, \dots, \frac{\exp(\alpha_{isM})}{\sum_{m=1}^{M} \exp(\alpha_{ism})}\right), \qquad (1)$$

where $\boldsymbol{\alpha}_{is}$ is a vector that describes consumer i's inclination toward the topics in session s. This formulation is an extension of the correlated topic model, originally developed by Blei and Lafferty (2007). The logit specification guarantees a valid parameter vector for a categorical

[14] Correlations between all elements in a Dirichlet distribution are innately negative because these elements sum to 1. However, the larger the number of topics *M* in the model, the lower the correlations between all elements of a Dirichlet.

distribution (i.e., the vector of non-negative elements sum up to 1). A larger value of α_{ism} leads to larger expected value of θ_{ism}, indicating a higher probability of topic m in session s. α_{ism} is modeled as the following function: $\alpha_{ism} = \kappa_{ism} + \boldsymbol{\beta}'_m \boldsymbol{X}_{is} + \epsilon_{ism}$. Here κ_{ism} is an individual-, session-, and topic-specific intercept; ϵ_{ism} is the error term which is assumed to follow a Normal distribution with mean zero and standard deviation σ_m. $\boldsymbol{X}_{is}$ is a vector containing predictor variables specific to a session s of consumer i that may shift the relevance of topics at the consumer and/or session level. For example, quarterly dummies could be used to capture potential seasonality in the preference for fashion products. In addition, we are particularly interested in whether and how a consumer's preference structure may evolve over time as he/she accumulates more shopping experience in the focal online store. Therefore, we include in $\boldsymbol{X}_{is}$ a few covariates that represent a consumer's cumulative shopping experience prior to the current session, such as the total number of brands browsed or purchased, and the total cumulative spending in the past.

Similar to Jacobs et al. (2017), we allow for correlations between topics at the session-level by formulating that $\boldsymbol{\kappa}_{is} = (\kappa_{is1}, \ \dots, \ \kappa_{isM})$ follows a multivariate Normal distribution with mean $\boldsymbol{\mu}_\kappa$ and covariance matrix $\boldsymbol{\Sigma}_\kappa$. This formulation allows us to capture richer relationships between topics and therefore between product attributes. For example, browsing and purchases across categories may exist such as pairs of pants and tops. Being able to identify such relationships in the data would offer additional insight for designing cross-selling recommendations.

Likelihood Functions

Let $z_{isn}^{[b]}$ denote the topic that drives the browsing of a product in a given session s by consumer i. The relevance of each topic $m \in \{1, \dots, \ M\}$ to the observed browsing outcome is captured by its assignment probability $Pr\left(z_{isn}^{[b]} = m\right)$, which is driven by the topic distribution probabilities $\boldsymbol{\theta}_{is}^{[b]}$. As the topics are latent, we account for all possible topic assignments to obtain $L(\cdot)$, the marginal likelihood that consumer i browses product $\boldsymbol{d}_j = (\boldsymbol{c}_j, \boldsymbol{y}_j)$ in session s, resulting in:

$$\begin{aligned} &L\left(\boldsymbol{d}_{isn}^{[b]} = (\boldsymbol{c}_{isn}^{[b]} = \boldsymbol{c},\ \boldsymbol{y}_{isn}^{[b]} = \boldsymbol{y}) \middle| \left\{\boldsymbol{\phi}_m^{[a]}\right\}_{m\in\{1,\dots,M\}}^{a\in\{1,\dots,A\}}, \{\boldsymbol{\mu}_m\}_{m=1}^{M},\ \{\boldsymbol{\Sigma}_m\}_{m=1}^{M}, \boldsymbol{\theta}_{is}^{[b]}\right) \\ &= \sum_{m=1}^{M} \left[\Pr\left(\boldsymbol{c}_{isn}^{[b]} = \boldsymbol{c} \middle| \{\boldsymbol{\phi}_m^{[a]}\}_{m\in\{1,\dots,M\}}^{a\in\{1,\dots,A\}}, z_{isn}^{[b]} = m\right) \cdot L(\ \boldsymbol{y}_{isn}^{[b]} = \boldsymbol{y} |\ \boldsymbol{\mu}_m,\ \boldsymbol{\Sigma}_m,\ z_{isn}^{[b]} = m)\ \right] \Pr(z_{isn}^{[b]} = m | \boldsymbol{\theta}_{is}^{[b]}) \\ &= \sum_{m=1}^{M} \left[\prod_{a=1}^{A} \phi_{m,c_a}^{[a]} \times \frac{\exp\left(-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}_m)\boldsymbol{\Sigma}_m^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_m)\right)}{\sqrt{(2\pi)^{|\boldsymbol{y}|}\det(\boldsymbol{\Sigma}_m)}}\right] \theta_{ism}^{[b]}, \end{aligned} \tag{2}$$

where c_a denotes the ath element in $\boldsymbol{c}$. The likelihood of observing the set of products browsed in session s by consumer i, $\boldsymbol{d}_{is}^{[b]} = \{\boldsymbol{d}_{is1}^{[b]}, \dots, \boldsymbol{d}_{isn}^{[b]}, \dots, \boldsymbol{d}_{isN_{is}^{[b]}}^{[b]}\}$, is the multiplication across products browsed in the session:

$$L\left(\boldsymbol{d}_{is}^{[b]}\right) = \prod_{n=1}^{N_{is}^{[b]}} L\left(\boldsymbol{d}_{isn}^{[b]} = (\boldsymbol{c}_{isn}^{[b]}, \boldsymbol{y}_{isn}^{[b]}) \middle| \left\{\boldsymbol{\phi}_m^{[a]}\right\}_{m\in 1,\dots,M}^{a\in 1,\dots,A}, \{\boldsymbol{\mu}_m\}_{m=1}^{M}, \{\boldsymbol{\Sigma}_m\}_{m=1}^{M}, \boldsymbol{\theta}_{is}^{[b]}\right).$$

Note that similarities of products browsed in the same session are captured by the common set of topics driving the session, $\boldsymbol{\theta}_{is}^{[b]}$.

Linking Browsing and Purchase Activities

The underlying preference structure that drives a consumer's browsing and purchase activities may not be the same. For example, her browsing activities may be driven by preference for products with deep discounts while her purchase activities may be driven by preference for particular brands. In order to capture the similarities and differences in the preference structure that drives these two types of shopping activities, we adopt the Hierarchically Dual Latent Dirichlet Allocation (HDLDA) modeling structure developed by Liu and Toubia (2018). Our application of this formulation involves two premises. First, purchase and browsing activities share the same set of M possible topics but are subject to different probability distributions over these topics even within the same session, which means that the topics that actually drive browsing and purchase activities in a session can be different. Second, the topic assignment probabilities for purchase activities are specified as functions of the topic assignment probabilities for browsing activities in a given session, which explicitly captures the notation that the relevance of topics for the purchase activities can be reinforced or weakened by the relevance of topics for browsing activities in the session.

Formally, $\boldsymbol{\theta}_{is}^{[p]}$, the topic distribution probabilities corresponding to purchase outcomes, are specified to be functions of $\boldsymbol{\theta}_{is}^{[b]}$. We capture the mapping between the two types of shopping activities in a matrix $\boldsymbol{R}$, where each element $r_{mm'}$ in matrix $\boldsymbol{R}$ indicates how the relevance of topic m for browsing activities affects that of topic m' for purchase activities, in the following parameterization of a Dirichlet distribution (see Liu and Toubia 2018):

$$\boldsymbol{\theta}_{is}^{[p]} \sim \text{Dirichlet}_M\left(\exp\left(\boldsymbol{R}_1'\boldsymbol{\theta}_{is}^{[b]}\right), \ldots, \exp\left(\boldsymbol{R}_M'\boldsymbol{\theta}_{is}^{[b]}\right)\right). \tag{3}$$

The exponential of the multiplication between $\boldsymbol{R}_m$ (the mth column of $\boldsymbol{R}$) and $\boldsymbol{\theta}_{is}^{[b]}$ is proportional to $E\left[\boldsymbol{\theta}_{ism}^{[p]}\right]$, the expected relevance of topic m for purchase activities in the session. To illustrate how this formulation captures the relationship between browsing and purchase activities, suppose $M = 2$, then the relevance of topics for purchase activities is as follows:

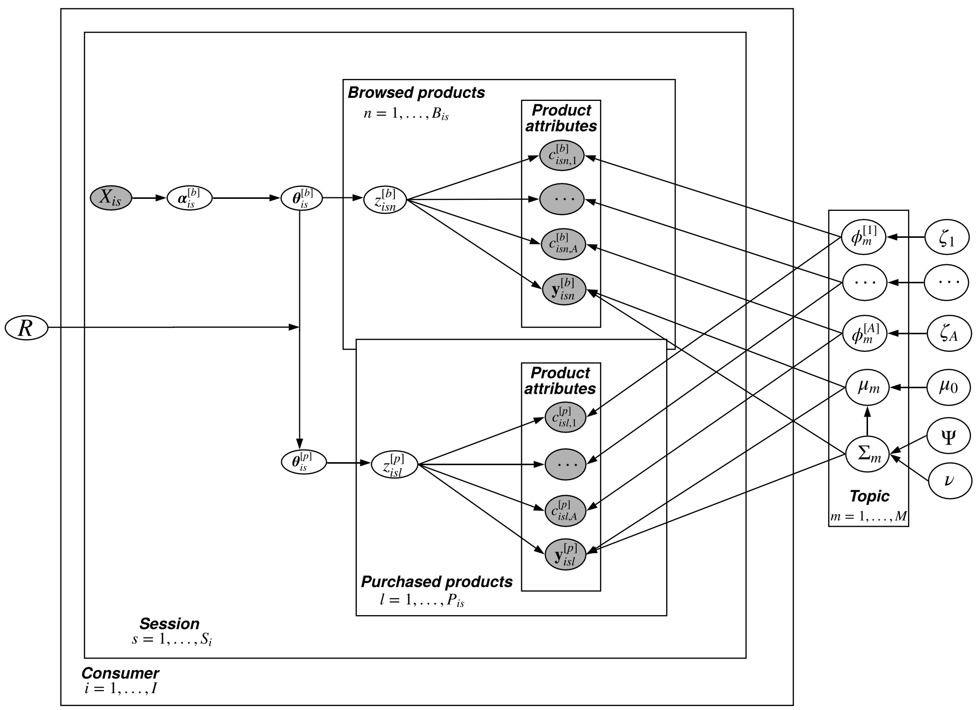
$$\left(\theta_{is1}^{[p]},\ \theta_{is2}^{[p]}\right) \sim \text{Dirichlet}_2\left(\exp\left(r_{11}\theta_{is1}^{[b]} + r_{21}\theta_{is2}^{[b]}\right),\ \exp\left(r_{12}\theta_{is1}^{[b]} + r_{22}\ \theta_{is2}^{[b]}\right)\right), \tag{4}$$

where r_{11} indicates the effect of the relevance of topic 1 for browsing activities on the relevance of topic 1 for purchase activities, and r_{21} indicates the effect of the relevance of topic 2 for browsing activities on the relevance of topic 1 for purchase activities in the same session. Positive (negative) value of $r_{mm'}$ indicates that the relevance of topic m for browsing activities is positively (negatively) associated with the relevance of topic m' for purchase activities in a consumer's preference structure.

The likelihood function of observing consumer i's purchased products in session s, $\boldsymbol{d}_{is}^{[p]}$, is similar to that for browsing activities, except that the distribution over the M topics are drawn as $\boldsymbol{\theta}_{is}^{[p]} \sim \text{Dirichlet}_M\left(\exp\left(\boldsymbol{R}_1'\boldsymbol{\theta}_{is}^{[b]}\right), \dots, \exp\left(\boldsymbol{R}_M'\boldsymbol{\theta}_{is}^{[b]}\right)\right)$.

Figure 1 presents the directed acyclic graph that summarizes the set of relations between the data and the model parameters, where shaded circles indicate observed variables while unshaded circles indicate hidden ones (to the researcher).

Figure 1. The Graphical Model.



Estimation

The posterior distribution of parameters in the model, given the observed attributes of products browsed and possibly purchased in all sessions, can be expressed as follows:

$$P\left(\{\boldsymbol{\theta}_{is}^{[b]}, \boldsymbol{\theta}_{is}^{[p]}, \boldsymbol{\kappa}_{is}, \boldsymbol{Z}_{is}^{[b]}, \boldsymbol{Z}_{is}^{[p]}\}_{i=1,\dots,I}^{s=1,\dots,S_i}, \{\boldsymbol{\phi}^{[a]}\}_{a=1,\dots,A}, \{\boldsymbol{\beta}_m,\ \boldsymbol{\gamma}_m\}_{m=1,\dots,M}, \boldsymbol{\mu}_\kappa, \boldsymbol{\Sigma}_\kappa, \boldsymbol{R} \mid\ \{\boldsymbol{\alpha}_a\}_{a=1,\dots,A}, \boldsymbol{\Psi},\ \nu, \boldsymbol{\mu_0}\right) = P(\boldsymbol{\Omega}|\boldsymbol{Y}).$$

Due to the complexity of the model structure and the scale of data, it is computationally challenging to infer the posterior distribution using inference methods based on sampling such as

MCMC, which is not easily scalable especially when the dataset is large and the probability model is complex (Kucukelbir et al. 2017). Instead, we use the variational inference (VI) estimation technique to infer the posteriors, which was developed in the machine learning literature for fast approximate inference (Jordan et al. 1999). It has been adopted by marketing academics recently and shown to offer great advantages in computational efficiency (e.g., Ansari et al. 2018). The basic idea of VI is to approximate the posterior distribution $P(\boldsymbol{\Omega}|\boldsymbol{Y})$ with a variational distribution denoted by $q(\boldsymbol{\Omega})$, and to minimize the Kullback-Leibler (KL) divergence from $q(\boldsymbol{\Omega})$ to $P(\boldsymbol{\Omega}|\boldsymbol{Y})$ which is defined as (see Blei et al. 2017, page 862):

$$KL[q(\boldsymbol{\Omega})||P(\boldsymbol{\Omega}|\boldsymbol{Y})\ ] \quad = E_q[\log q(\boldsymbol{\Omega})] - E_q[\log P(\boldsymbol{\Omega},\boldsymbol{Y})\ ] + \log P(\boldsymbol{Y}) \tag{5}$$

where E_q is the expected value under the variational distribution q. Intuitively, this corresponds to finding the variational distribution $q(\boldsymbol{\Omega})$ that best approximates the posterior $P(\boldsymbol{\Omega}|\ \boldsymbol{Y})$. However, directly evaluating $\log P(\boldsymbol{Y})$ is not feasible and thus one cannot compute the KL divergence. Instead, the so-called Evidence Lower bound (ELBO) is used: $ELBO[q(\boldsymbol{\Omega})] = E_q[\log P(\boldsymbol{\Omega},\boldsymbol{Y})\ ] - E_q[\log q(\boldsymbol{\Omega})]$, which provides a lower bound on $\log P(\boldsymbol{Y})$. It can be shown that the difference between the ELBO and $\log P(\boldsymbol{Y})$ is, in fact, equal to the KL divergence: $\log P(\boldsymbol{Y}) - ELBO[q(\boldsymbol{\Omega})] = KL[q(\boldsymbol{\Omega})||P(\boldsymbol{\Omega}|\boldsymbol{Y})]$. Since $\log P(\boldsymbol{Y})$ is constant with respect to the variational distribution q, maximizing the ELBO is equivalent to minimizing the KL divergence from $q(\boldsymbol{\Omega})$ to $P(\boldsymbol{\Omega}|\boldsymbol{Y})$ (Blei et al. 2017).

The complete data generating process assumed in our model is presented in Appendix A, and the prior distributions of parameters in the model are described in Appendix B. We use the ADVI (Kucukelbir et al. 2017) implementation provided in the software Stan (Stan Development Team 2018) to estimate our model. The convergence of the estimation procedure is determined by the convergence of the ELBO and carefully monitoring the running average of ELBO changes

(Yao et al. 2018). In our empirical application, it takes about five days for each model to converge on a personal computer equipped with an Intel i5-6600 processor, which speaks to the necessity to utilize an efficient inference technique like the VI.

PERSONALIZED PRODUCT RECOMMENDATIONS

We propose a new personalized product recommendation system based on individual consumers' preference structures regarding the entire store assortment, as inferred from our model.

Personalized product recommendations (PPRs) play an increasingly important role in e-commerce practices. For example, almost 67% of watched movies in Netflix come from personalized recommendations, and personalized recommendation systems generate 38% more click-throughs for Google News and 35% more sales for Amazon (Ma et al. 2016). In order to implement PPRs, many online retailers use collaborative filtering (CF) algorithms that utilize consumer's past purchase behaviors to identify a neighborhood of consumers that exhibits similar choices, and then recommend new products based on the choices of the neighborhood (Karypis 2001). CF has been successfully applied in large scale recommendation systems (e.g., Amazon.com) (Linden et al. 2003). Nonetheless, there are several limitations when applying CF, especially for large and frequently changing assortments. First, the researcher has to decide *a priori* how to count the co-occurrence of products to find individuals with similar tastes (e.g., pairs, triplets, or even higher-order product combinations). Focusing on the co-occurrence of small sets of products to identify similar consumers in consumer base can result in information loss, which leads to less accurate predictions. On the other hand, co-occurrence counts tend to be sparse for large combinations of products, which could result in poor predictions and quality of recommendations (Jacobs et al. 2016). Second, incorporating consumer characteristics in collaborative filtering algorithms is difficult. Partitioning the consumer base based on the similarity between such characteristics may be possible (Adomavicius and Tuzhilin 2005, Ricci et al. 2011). However, as there are multiple such variables, small sample sizes per consumer subgroup may not be useful to make predictions and recommendations. Third, there is generally a "cold start" problem for making recommendations for both new users and new products that do not exist in the calibration dataset. In other words, if a new user has not browsed or purchased any product yet, collaborative filtering cannot recommend any product. Moreover, both user-based and item-based collaborative filters cannot make recommendations for products that do not exist in the extant assortment and thus could not be browsed or purchased by anyone previously. The cold-start problem regarding new products is particularly relevant for firms managing frequently changing assortments, such as fashion retailers and marketplaces. For example, in our empirical application, many products in the hold-out dataset did not exist in the calibration dataset.

Content-based filtering can handle the cold start problem for new products by utilizing similarities of product attributes. Recommendations for new products are based on the similarities of attributes between a new product and those of previously purchased products. For example, the online music station Pandora recommends new songs that best match the attributes of previously listened songs (Howe 2009). Nonetheless, both collaborative and content-based filtering cannot necessarily utilize information on time-varying factors such as seasonality and consumer's shopping experiences (Ansari et al. 2000, Adomavicius and Tuzhilin 2005), which hinders the quality of predictions and recommendations. In addition, for online shopping behaviors, they cannot provide a holistic understanding of the interconnections between browsing and purchase activities and the differences in their underlying driving factors, which provides important insights for designing effective targeting strategies at the two different levels of the consumer journey. In contrast, our proposed personalization system can handle all three issues: the cold-start problem for new products, underutilization of information on time-varying factors, and an in-depth understanding of both browsing and purchase activities.

Deriving Personalized Recommendations: Optimizing Product Attributes

The basic idea behind our proposed recommendation system is to translate the prediction problem into a problem of finding the optimal set of product attributes that maximize a consumer's likelihood of browsing or purchase. We utilize the estimation results of our proposed attribute-based LDA model to derive the optimal product attributes for each consumer that maximize the likelihood that a consumer will browse or purchase a product in a given session. In the empirical application, we illustrate the approach by assuming that the focal firm would want to maximize the probability that a consumer browses the recommended product, but it can be easily modified to maximize the probability of making a purchase or other related objectives. The said problem translates to maximizing the following likelihood over the vector of product attributes $\boldsymbol{d}_{is} = (\boldsymbol{c}_{is}, \boldsymbol{y}_{is})$:

$$max_{\boldsymbol{c},\boldsymbol{y}}\, \boldsymbol{g} = max_{\boldsymbol{c},\boldsymbol{y}}\, L\left(\boldsymbol{d}_{is}^{[b]} = (\boldsymbol{c}, \boldsymbol{y}) \middle| \left\{\boldsymbol{\phi}_m^{[a]}\right\}_{m\in 1,\dots,M}^{a\in 1,\dots,A}, \{\boldsymbol{\mu}_m\}_{m=1}^{M}, \{\boldsymbol{\Sigma}_m\}_{m=1}^{M}, \boldsymbol{\theta}_{is}^{[b]}\right). \quad (6)$$

Solving this optimization problem involves two challenges. First, some product attributes take discrete values (e.g., category, brand, color, and "gender") while other product attributes take continuous values (e.g., initial price and depth of discount), which prevents us from adopting gradient-based optimization methods to maximize the objective function. Second, exploring the large combinatorial space of discrete product attributes using conventional methods such as grid or random search is a formidable task in terms of computing time and resources. To address these two challenges, we adopt the Bayesian optimization (BO) method to find optimal product attributes. BO is a derivative-free global optimization method. The basic idea of BO is to employ Bayesian updating of values of the objective function, which utilizes parameters obtained in the process of searching to update next candidate parameters (Frazier 2018). The details of our algorithm and application of the Bayesian optimization technique are presented in Appendix D.

EMPIRICAL APPLICATION

Institutional Background

We apply the proposed modeling framework to data from an online deal marketplace based in China which specializes in fashion products. It was launched in April 2013 as a marketplace for merchants to sell products at discounted prices for limited time periods. Consumers can view detailed information on available products, such as their picture, brand, product descriptions, in addition to the initial price and sales price, and place orders via the company's app. This marketplace operates exclusively through its mobile app and does not have a web presence.

This is an ideal context for the purpose of our research, because products sold in such venues usually are available for only a limited period of time and thus the assortment is frequently changing. We would like to note though that the general approach proposed in this research applies to assortment data by any retailers, and the specific model formulation that we construct can be applied to studying online browsing and purchase behaviors in any online stores or marketplaces.

Data Descriptions

Store-level assortment

Our dataset contains individual consumers' browsing and purchase records with detailed timestamps of their activities and information on all products offered on the marketplace from August 2013 to July 2014, spanning a total of 53 calendar weeks. Information on each product (SKU) includes its category, brand name, color, target users (such as men or women[15]), initial price, and sales price, etc. We eliminated a small number of SKUs that were browsed less than ten times in the entire data, which yielded 7,658 unique SKUs for model estimation and follow-up analyses. There are 22, 24, 2, 12 unique categories, brands, "genders" of the product, and colors, respectively, in the data. The brand names are recoded into 1, 2, …, 24 for confidentiality reasons. The depth of discount is measured as the percentage of a product's initial price. Table 2 provides more details of the product attributes used in our analyses. We use the logarithm of the initial price and depth of discount in the model to reduce the skewedness of the original variables.

Table 2. Description of Product Attributes.

Discrete Attribute	Number of levels	Details		
Category	22	Tees & Knits, Shirts & Blouses, Dresses, Pants, Jackets & Blazers, Sweaters, Overcoats, Jeans, Skirts, Lounge & Sleepwear, Underwear & Lingerie, Shorts, Loafers & Other Casual, Dress Shoes/Pumps, Boots/Booties, Jewel, Scarves & Wraps, Belts, Other accessories, Handbags, Backpacks, Other bags		
Brand	24	Coded as 1, 2, …, 24		
"Gender"	2	Female, Male		
Color	12	Black, Purple, White/Off-White, Red, Green/Deep Green, Blue/Navy, Yellow/Gold/Orange, Grey, Brown/Khaki, Pink, Mixed Colors, Other Colors		
Continuous Attribute	Mean	SD	Min	Max
Initial price	951.41	903.70	56.00	11,000
Depth of discount (%)	0.34	0.22	0.01	0.99

The average initial price and percentage depth of discount were 951.4 yuan and 34.2% respectively. The largest category was Tees & Knits (16.2%), followed by Loafers & Other Casual (11.8%), Shirts & Blouses (11.7%), Jackets & Blazers (9.0%), and Pants (8.2%). A brand

[15] A small number of SKUs were for children or unisex, which we eliminated in the analyses because of their very low total numbers of browses or purchases in the data.

spans across 12.3 categories on average. 63.9% of the SKUs were products for women. In terms of product colors, the largest shares are Black (21.5%) followed by Blue/Navy (19.1%), White/Off-white (12.3%), Red (8.4%), and Grey (8.0%).

Sessions and consumers' prior shopping experience

We define a session as a period of continuous browsing activities separated by no more than one hour of inactivity (similar to Schweidel and Moe 2016). We keep consumers who had at least three sessions in the estimation dataset for identification of consumer-specific parameters, which yielded browsing and purchase records of 1,738 consumers who made a total of 24,115 browsing sessions and 5,190 purchases. Table 3 presents descriptive statistics of the browsing and purchase activities of these 1,738 consumers. The mean number of browsed products and purchased products in a session with any purchase are 3.96 and 1.26 with standard deviations of 6.10 and 0.86. 17.1% of the sessions involved any purchases, while the rest were purely browsing sessions. The average number of browsing sessions and purchase occasions per consumer are 13.88 and 2.38, with standard deviations of 26.23 and 7.31, respectively. We use the number of brands browsed, the number of brands purchased, and the total spending amount before the current session as variables representing consumer's prior shopping experience.[16]

Table 3. Descriptive Statistics of Sessions and Consumer Characteristics.

Variable	Mean	SD	Min	Max
Number of browsed products per session	3.96	6.10	1	160
Number of purchased products in a session with any purchases	1.26	0.86	1	12
Made any purchases (yes=1/no=0) in a session	0.17	0.38	0	1
Number of sessions per consumer	13.88	26.23	3	303
Number of purchase occasions per consumer	2.38	7.31	0	186
Cumulative number of brands previously browsed	11.82	7.85	0	24
Cumulative number of brands previously purchased	1.80	3.73	0	23

[16] Due to high correlations between the cumulative number of brands and product categories browsed (or purchased), we only included the cumulative number of brands browed and that of brands purchased in the final model.

Cumulative store spending in the past (in 100 yuan)	24.29	97.68	0	1,402.93

Calibration and Hold-out Datasets

We estimate the model using data of the first 45 weeks and use data of the last eight weeks as the hold-out sample. The hold-out dataset will be used to assess the predictive performance of the proposed method and compare it with several benchmarks. Table 4 summarizes the two data subsets in terms of the numbers of consumers, unique products (SKUs), sessions, and purchase occasions. Of the 1,736 unique SKUs in the holdout dataset, 796 of them appeared in the calibration data and 940 of them did not exist in the calibration data, well reflecting the frequently-changing nature of the focal online marketplace's assortment. This allows us to examine the model's ability to perform demand forecast for not only SKUs in the existing assortment, but also those representing "out-of-vocabulary" products.

Table 4. Description of the Calibration and Hold-out Datasets.

Dataset	Number of consumers	Number of SKUs	Number of sessions	Number of purchase occasions
Full data	1,738	7,658	24,115	5,190
Calibration data	1,521	6,718	20,820	3,736
Hold-out data	561	1,736	3,295	1,454

ESTIMATION RESULTS

We estimate the proposed model using variational inference techniques. All results in this section are based on posterior means of 1,000 samples from the variational distributions of the parameters.

Topics

The number of topics M is a hyperparameter, which is set before the estimation. We need to balance the trade-off between the interpretability of the model results (which favors lower values of M) and the predictive power of the model (which favors higher values of M) (Liu and Toubia 2018). In this empirical application of the model, we set $M = 15$.[17]

Recall that each topic in the model is characterized by a set of distributions over the product attributes. The distribution over a discrete product attribute is represented by a Dirichlet distribution. The distribution over logarithms of the initial price and depth of discount (%) is formulated as a bivariate Normal distribution. Figure 2 presents a heat map of the topics characterized by their estimated probability distributions of discrete product attributes (category, brand, "gender", and color). Each row represents a topic, and each column represents a specific product attribute level. The numbers in the last two columns represent the mean of the initial price (in 100 yuan) and the depth of discount (%) corresponding to each topic.

Figure 2. Heat Map of Topics.

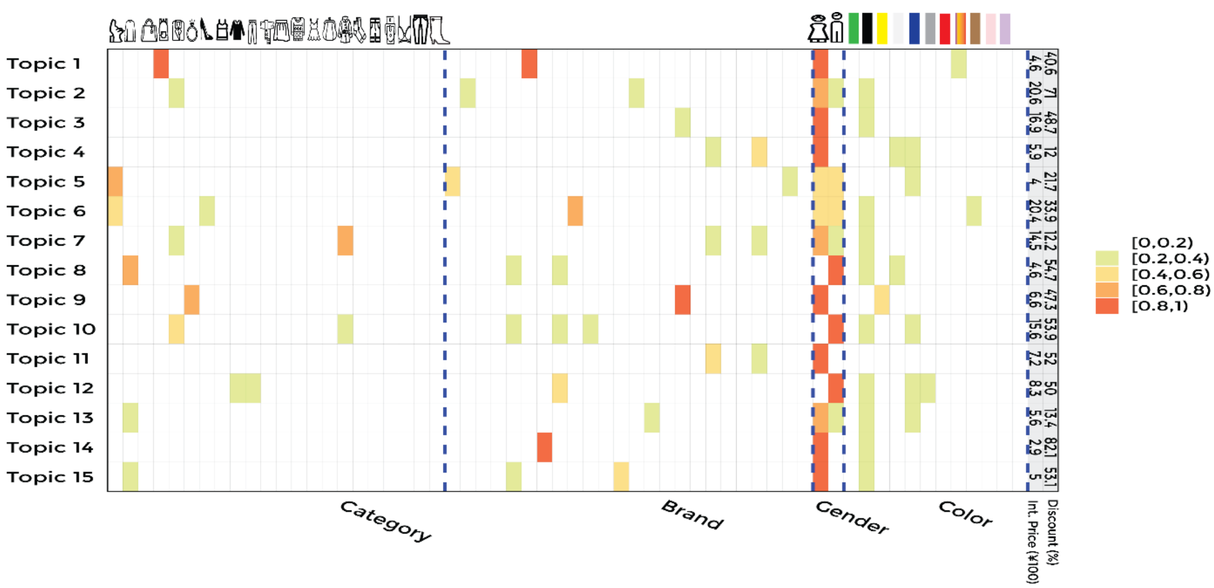


Figure 2 shows varied patterns in terms of how specific product attributes are related to the identified topics. Some topics are characterized by a single attribute level on most product

[17] The number of topics was chosen to sufficiently capture different topics yet is still manageable for practical usage. Jacobs et al. (2016) found the number of topics to be 11 to 15 in their analyses of assortment of an online grocery retailer.

attributes. For example, Topic 1 represents “brand 6, for women, backpacks, with the mean initial price of 460 yuan and 41% discounts”, and Topic 9 represents “brand 16, women’s yellow/gold/orange Jewel with the mean initial price of 660 yuan and 47% discounts”. Nonetheless, most other topics are characterized by attributes spanning across multiple categories, brands, “genders”, or colors. For example, Topic 2 shows “brands 2 and 13, women’s jackets & blazers, shirts & blouses, and sweaters, in black or blue/navy colors, with the mean initial price of 2,061 yuan and 71% discounts”.

As an illustration of how topics are related to specific products, we present the top five most likely products based on their likelihood for two sample topics in Table 5. The complete list of the top five most likely products for all 15 topics is reported in Appendix C.

Table 5. The Top Five Most Likely Products for Topics 2 and 12.

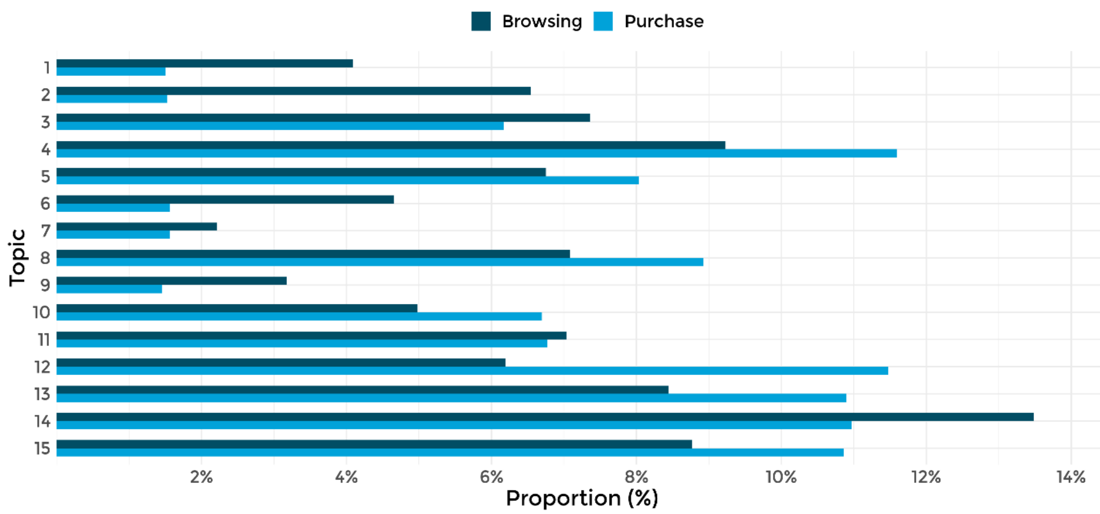
Topic 2						***Topic 12***					
Category	Brand	Gender	Color	Initial Price (¥)	Depth of Discount (%)	Category	Brand	Gender	Color	Initial Price (¥)	Depth of Discount (%)
Jackets & Blazers	2	Women	Black	800	70	Shirts & Blouses	8	Men	Blue/ Navy	839	50
Shirts & Blouses	13	Women	Black	290	70	Shirts & Blouses	8	Men	Black	739	50
Sweaters	2	Women	Black	800	70	Pants	8	Men	Black	839	50
Shirts & Blouses	13	Women	Black	990	70	Shirts & Blouses	8	Men	Blue/ Navy	929	50
Shirts & Blouses	2	Women	Black	100	70	Pants	8	Men	Black	839	60

The complementarity patterns among product categories and other attributes (such as brands and colors) as naturally revealed by the identified topics. Complementarity among product attributes is a common feature of fashion products. It not only lends support to the validity of our model but also offers valuable insights for cross-selling opportunities. For examples, based on Topic 2, a retailer could choose to recommend a women’s shirts & blouses of brand 13 to a consumer who has browsed a jacket/blazer of brand 2. Based on Topic 12, the recommendation could be a pair of men's black pants of brand 8 to a consumer who has browsed a men's navy shirt of the same brand.

Topic Relevance to Browsing versus Purchase Activities

A novel aspect of our study is to provide insights on the similarities and differences in the underlying topics that drive online browsing versus purchase activities. The relevance of a given topic in driving browsing or purchase activities is captured by its probability in the estimated Dirichlet distribution across topics. Figure 3 presents the average relevance of topics captured in $\boldsymbol{\theta}_{is}^{[b]}$ and $\boldsymbol{\theta}_{is}^{[p]}$ across all shopping sessions in the data. The length of each bar represents the average probability allocation of each topic to browsing or purchase across all shopping sessions.

Figure 3. The Relevance of Topics for Browsing and Purchase Activities.

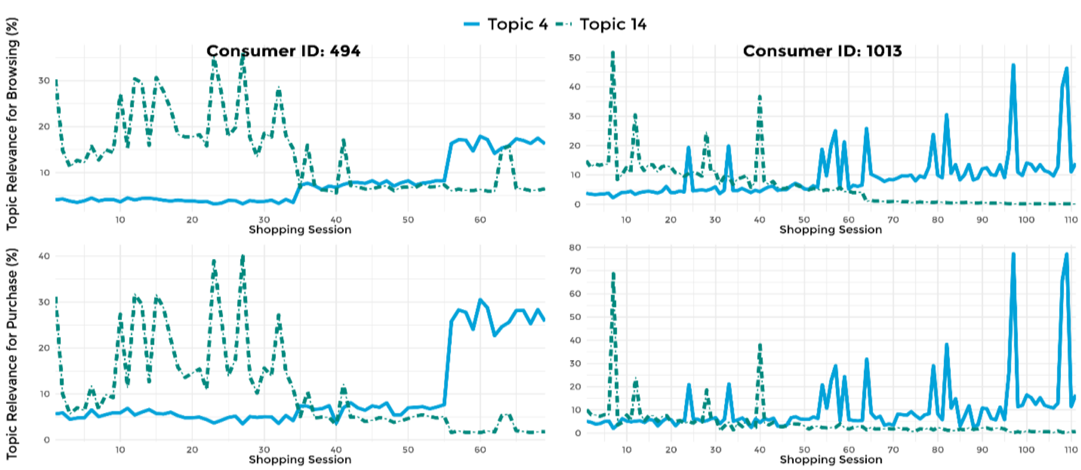


As shown in the figure, some topics play a prominent role in driving both browsing and purchase activities, while others influence one much more than the other. For example, across consumers and shopping sessions, Topic 14 (characterized by "Brand 7, for women, black color, mean initial price of 290 yuan and discount of 82%") plays an important role in driving both browsing and purchase activities, and Topic 12 (characterized by "Brand 8, for men, shirts and pants, black or navy/blue or grey color, mean initial price of 830 yuan and discount of 50%") is more relevant in driving purchase activities than browsing activities. In contrast, Topic 2 (characterized by "Brands 2 and 13, female's Jackets and Blazers, black color, mean initial price of 2,061 yuan and 71% discounts") is more relevant in driving browsing activities than purchase activities. These dashboard results provide useful insights to retail managers on what types of products tend to attract consumers to browse in an online store or marketplace and what types of products play a prominent role in converting browsing into purchases.

Consumer-Specific Evolution of Preference Structure over Shopping Sessions

Our model reveals how topic relevance evolves with a consumer's shopping experience in the focal marketplace, which reflects how consumers' preference structures change over time. As an example, Figure 4 shows how the relevance of Topic 4 and Topic 14 evolves across shopping sessions for two consumers.

Figure 4. Evolution of Preference Structure over Shopping Sessions for Two Sample Consumers.



The figure indicates that, for both consumers, the relevance of Topic 14 (characterized as "Brand 7, for women, black color, mean initial price of 290 yuan and discount of 82%") decreased over shopping sessions, while that of Topic 4 (characterized as "Brands 18 and 21, for women, white color, average initial price of 590 yuan and discount of 12%) increased over shopping sessions. In addition to changing preferences in brands and physical product attributes, these two consumers appeared to have been attracted by lower initial prices and very deep discounts in the early stage but gravitated towards products of higher initial prices and less discounts over time, indicating that the quality or brand appeal became more important to them than monetary savings in the focal deal marketplace. Retailers can use such insights to identify trends in consumers' preference evolutions, which would be valuable input for how to update their assortment, pricing, and promotions for future periods. In addition, they can be used to tailor merchandise offerings to individual consumers. For example, the focal marketplace could work with vendors of brands 18 and 21 to offer targeted promotions of these brands to the two customers in the above illustration. Such value-added services to vendors could create additional business opportunities for the marketplace.

Performance Comparison with Benchmark Recommendation Systems

We conduct performance comparisons among different recommendation systems. We use three popular recommendation systems in practice as benchmarks: 1) user-based collaborative filtering (cf. Ricci et al. 2011), 2) item-based collaborative filtering (cf. Karypis 2001), and 3) content-based filtering (cf. Ricci et al. 2011). For each method, we create consumer-specific prediction sets of products that are most likely to be browsed and then compare them against a consumer's actual browsing activities in the hold-out dataset.

We use prediction sets of different sizes to evaluate the recommendation algorithms for the different context of customization applications (Jacobs et al. 2016). A prediction set of size V contains the V highest ranked products for a consumer. For recommending a single product, the size of the prediction set is one. We examine V ranging from one to ten, because the number of recommended products in a single push hardly goes beyond ten in practical applications. For example, as of June 2019, “ralphlauren.com” (“You May Also Like”), “jcrew.com” (“Customer Also Love”), and “gap.com” (“Customers Also Viewed”) recommend six, five, and five products on their websites, respectively.

We assess the quality of a prediction set by matching its contents against products actually browsed by a consumer in the hold-out data. Denote these products by $\boldsymbol{y}_i$ for consumer i, and the number of these unique products in $\boldsymbol{y}_i$ by u_i. The vector $\boldsymbol{z}_i$ contains the V recommended products for consumer i in their ranking order, with the first element, z_{i1}, being the optimal product with the highest predicted probability of browsing based on our proposed method or with the highest predicted scores based on the benchmark methods.

The quality of a prediction set of size V can be measured by the number of products in the prediction set that overlap with the hold-out browsed products: $\sum_{v=1}^{V} I[z_{iv} \in \boldsymbol{y}_i]$, where $I[.]$ equals 1 if the argument inside the bracket is true, and 0 otherwise. This number should be best calculated relative to the maximum number of hits possible to evaluate a hit rate that is comparable across prediction sets of different sizes (Jacobs et al. 2016). This maximum is bounded by V, the size of the prediction set, and the number of unique hold-out browsed products u_i. Thus, the hit rate for consumer i is defined as: $\sum_{v=1}^{V} I[\mathrm{z}_{iv} \in \boldsymbol{y}_i] \,/\, min(V, u_i)$.

Along with the hit rate, we also consider a frequently used performance measure in machine learning literature, the F-measure, which combines precision and recall (Adomavicius and Tuzhilin 2005). Precision is the proportion of products in the recommendation list that match with products actually browsed in the hold-out data, and recall is the proportion of products actually browsed in the hold-out data that appear in the recommendation list. Since precision and recall of a given algorithm do not usually go in the same direction, the F-measure, a harmonic mean of precision and recall, is often used to compare the overall performance of recommender systems and is defined as: $\left((recall^{-1} + precision^{-1})/2\right)^{-1}$ (Adomavicius and Tuzhilin 2005). A higher F-measure indicates a better overall performance.

A limitation of performance measures such as the hit rate and the F-measure is their insensitivity to ordering in recommendation lists. However, it is likely that positions in a recommendation list affect consumer's reaction to the list (Smith and Brynjolfsson 2001, Xu and Kim 2008). To address the limitation, following Jacobs et al. (2016), we take into account weights in the positions within a recommendation list by weighing the hits according to their ranks in the list, which reflects importance in the recommendation list. For the vth ranked recommended product in a prediction set of size V, this weight is defined as $w(v, V) = 1 - (v - 1)/V$. And the weighted-ranking hit rate defined as:

$$h_i(z_i, V) = \frac{\sum_{v=1}^{V} I[z_{iv} \in y_i] \cdot w(v, V)}{\sum_{v=1}^{min(V,u_i)} w(v, V)}. \tag{7}$$

Results of Performance Comparisons

We examine the performance for the "in-vocabulary" products (i.e., those appeared in the calibration data) and "out-of-vocabulary" products (i.e., those appeared only in the hold-out data) separately to ensure fair comparisons for the recommendation systems, because collaborative-based filters are not capable of recommending "out-of-vocabulary" products.

Figure 5 and 6 present the average hit rate, ranking-weighted hit rate, and F-measure across all consumers that appeared in both the estimation dataset and the hold-out dataset for each recommendation method for "in-vocabulary" and "out-of-vocabulary" products, respectively.

Figure 5. Predictive Performance in the Hold-Out Dataset for "In-Vocabulary" Products.

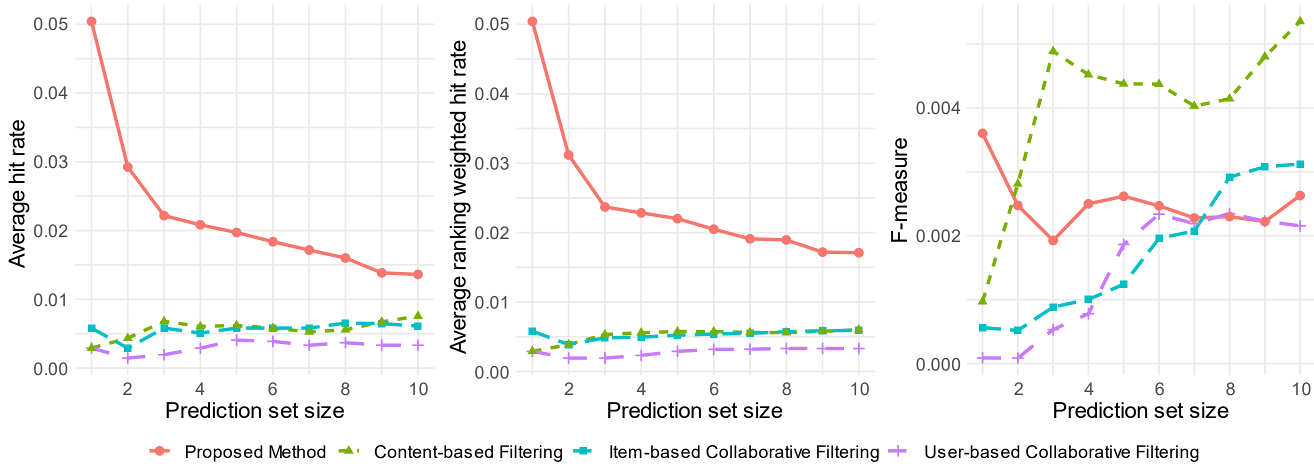


Figure 5 shows that our proposed recommendation system offers the best performance and substantially outweighs the benchmarks in terms of average hit rate and average ranking weighted hit rate for recommending “in-vocabulary”. Our method improves the weighted hit rate by almost 300% over the benchmarks. When it comes to the F-measure, the content-based filtering yields the best performance in general, although our proposed method is still the best when a single product is recommended. A closer examination indicates that our proposed method outperforms all benchmarks in terms of recall, while the content-based filtering generally performs the best in terms of precision among all four methods.

Figure 6. Predictive Performance in the Hold-Out Dataset for “out-of-vocabulary” Products.

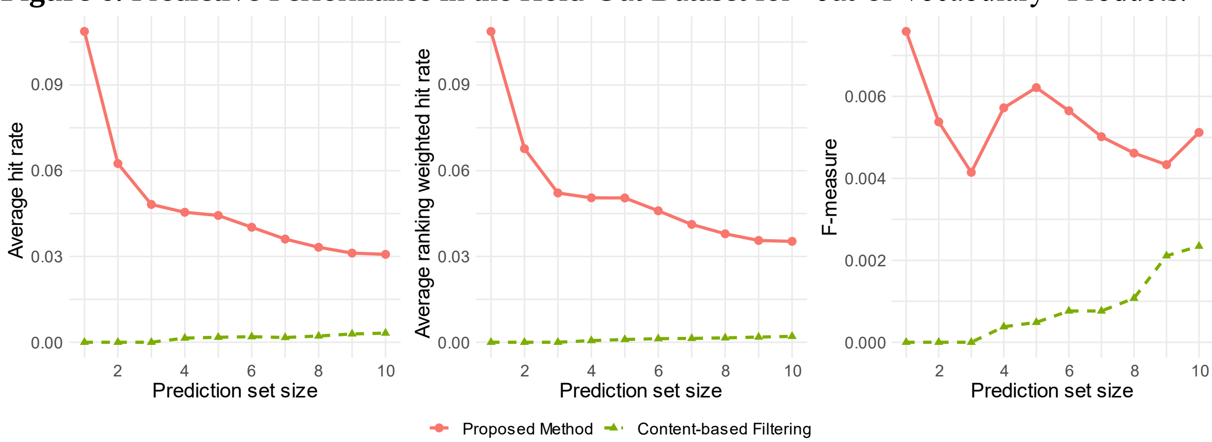


Figure 6 shows that our proposed method outperforms the content-based filtering in terms of average hit rate, average ranking weighted hit rate, and F-measure for "out-of-vocabulary" products, and it improves the weighted hit rate by almost 16 times. This comparison demonstrates that our approach is particularly powerful in making recommendations of products that did not exist in the previous assortment, and thus is especially valuable for firms managing frequently changing assortments.

Note that the average hit rate and average ranking weighted hit rate decline for our proposed algorithm. There are two reasons for this pattern. First, the declining pattern is a consequence of recommending products with sub-optimal attributes founded in the process of the optimization algorithm. This finding suggests that our proposed algorithm would best perform in applications where the number of personalized products is fairly small, such as in the context of fashion products or technology gadgets. In any case, managers may wish to limit their recommendation lists to a moderate number of products to avoid “choice overload” which could lead to consumers making no choice at all (Chernev et al. 2015). Second, recommended products with a combination of attributes that do not exist in both calibration and hold-out data are innately counted as misses even if these are novel to a consumer. It is likely, however, that consumers could choose some of those products with new combinations of attributes that suit their preferences, and if this is the case, the average weighted hit rate would be higher than the reported measures. Hence, the current performance measures reported can be interpreted as the lower bound of the performance of our proposed method.

As shown in Figure 6, one powerful advantage of our proposed recommendation system is its ability to recommend “out-of-vocabulary” products with much higher success rates than content-based filtering. To illustrate this application, Table 6 presents products with optimal attributes obtained from our algorithm for two sample consumers. The bold and italic characters indicate “out-of-vocabulary” products, not in the assortment of the estimation dataset. This advantage is particularly valuable for retailers that provide personal styling services (e.g., Stitch Fix, the Trunk Club by Nordstrom, Le Tote) to create new styles by recombining attributes from existing styles and slightly modifying them in a way that is most likely to suit a given consumer's tastes.

Table 6. Top Five Products That Are Most Likely to Be Browsed by Two Sample Consumers.

Rank	Category	Brand	Color	Gender identity	Initial price (¥)	Depth of discount
Consumer ID: 793						
1	***Scarves & Wraps***	***19***	***Grey***	***Women***	326	45
2	*Tees & Knits*	*24*	*Green/ Deep Green*	*Women*	432	54
3	*Shirts & Blouses*	*24*	*Green/Deep Green*	*Women*	368	46
4	***Loafers & Other Casual***	***20***	***Green/Deep Green***	***Women***	277	43
5	***Tees & Knits***	***24***	***Other Colors***	***Women***	705	56
Consumer ID: 1139						
1	***Shorts***	***15***	***Green/Deep Green***	***Men***	431	57
2	***Overcoats***	***19***	***Black***	***Men***	513	21
3	***Loafers & Other Casual***	***24***	***Green/Deep Green***	***Men***	282	34
4	Overcoats	24	Green/Deep Green	Men	301	32
5	***Belts***	***24***	***Purple***	***Men***	189	36

SUMMARY AND DISCUSSIONS

This paper develops a scalable modeling framework for discovering online consumers' preference structures in large and frequently changing assortments from their browsing and purchase activities. We build upon the latest developments in machine learning techniques and Bayesian estimation methods, and the proposed model overcomes several limitations of standard LDA topic models for handling large assortments. An advantage of our proposed attribute-based

LDA model is that the topics are now parsimoniously characterized by distributions over a much smaller number of dimensions (i.e., product attributes instead of SKUs), which provides straightforward and objective interpretation of the meaning of each identified topic and makes them directly actionable for making merchandising decisions. Another advantage of our approach is that it can be used to predict preferences for products that are not in the calibration data of existing store assortment and thus is particularly useful to retailers and marketplaces with frequently changing assortments. In addition, our approach incorporates the influence of product-specific variables that vary by time (such as price and promotion), which is an important improvement for assisting merchandising decisions. Moreover, we propose a comprehensive framework to uncover the topics driving online browsing versus purchase activities and how the relevance of these topics may evolve with consumers' shopping experience in a store. Our empirical analyses reveal novel insights on the topics of interest that drive online shopping behaviors. We summarize the key findings and their managerial implications below.

First, the proposed model offers an effective tool for retail managers to discover complementarity patterns across product categories (e.g., shirts and pants) or attributes (e.g., brands or colors) without imposing *a priori* structures, because such patterns are empirically revealed by those topics that span across multiple product attributes or multiple levels of a given attribute. These insights are particularly useful for retailers to devise cross-selling offerings, as illustrated in the previous section.

Second, our analyses show nuanced patterns of the relevance of different topics in driving online browsing versus purchase activities. Some topics play a prominent role in driving both types of activities, while others influence one much more than the other. These insights can help retail managers set priorities and fine-tune their assortments to intentionally encourage browsing or purchase conversion. Those topics that are important in driving both types of activities should be given priorities in merchandising decisions. In addition, managers can modify these plans by timeframes or other relevant considerations. For example, a new online retailer or marketplace may be more interested in attracting browsers and increasing its customer base, and it can adjust its assortment towards stocking more products that drive browsing. As time goes on, its priority may shift to purchase conversion, in which case the assortment can be adjusted accordingly to topics that are more relevant in driving purchase activities.

Third, our study provides detailed depictions of consumer-specific preference structures and how they may evolve over time with consumers' shopping experience in the store, which offer a deeper understanding of consumer heterogeneity and can serve as a basis for further improving personalization strategies. We find that, as the cumulative number of brands browsed or purchased by a consumer increases, certain topics became more prominent in driving their shopping activities while others became less relevant. Such evolution patterns of the topic relevance for individual consumers enable retail managers to modify their targeting objectives by tracking a shopper's prior activities in their store (for example, to focus on encouraging browsing or purchase conversion) and to further refine personalized targeting strategies to individual customers.

Finally, the empirical results well demonstrate advantages of our proposed personalized product recommendation system, which is based on in-depth understanding of individual consumer's preference structures gained from the proposed model. We find that it substantially outperforms benchmark product personalization algorithms commonly used in e-commerce practices, such as user- and item-based collaborative filters and content-based filtering. Hold-out comparisons show that, for "in-vocabulary" products, our proposed method improves the weighted hit rate over the benchmarks by at least 300% for recommendation lists of up to ten products, and that its relative advantage is even stronger for smaller recommendation lists. Moreover, our recommendation system is capable of prescribing "out-of-vocabulary" products beyond the existing assortment, which cannot be done by user- or item-based collaborative, and its performance is drastically better than the content-based filtering method, improving the weighted hit rate over the latter by almost 16 times. This shows that our proposed modeling framework and recommendation system are particularly valuable for retailers and marketplaces that manage large and frequently changing assortments. The content-based filtering relies on observed similarities of products previously browsed/purchased and the recommended new products, while our proposed method taps into consumers' underlying preference structures which are manifested on the product attribute space and could potentially span across seemingly dissimilar or unrelated product attributes. Their performance comparison indicates that our proposed model provides more in-depth and accurate depictions of consumer preferences, and that product recommendation systems based on such insights can bring about significant benefits to e-commerce practitioners.

Our empirical focus of the research is to study browsing and purchase activities in an online shopping environment regarding fashion products. Nonetheless, the modeling framework can be easily modified for other retail contexts. For example, offline retailers can simplify the model to focus on only purchase data, in which case the model does not need the hierarchical structure component. Retailers selling other types of merchandise can adjust the specific product attributes and their levels used in the model to fit the nature of their business. In this study, we choose to focus on common product attributes that are relevant to the entire assortment. Another way to expand the model is to incorporate additional attributes that are specific to individual merchandise departments. This would provide more in-depth knowledge within those merchandise departments, but a potential trade-off is that such attributes may be forced to load on only a very small number of topics and are not relevant to most consumers in the data.

In addition to modifying the attribute-based LDA model, our proposed personalized product recommendation system can also be adjusted to accommodate other managerial objectives or decision variables. For example, a retailer may wish to use a combination of browsing and purchase objectives in deriving personalized product recommendations or to adjust the objectives by individual consumers, especially by tracking their preference evolution patterns. In combination with adjusting products in the assortment, retail managers could also utilize personalized discounts as an additional vehicle to achieve their managerial objectives. Since our proposed model can well accommodate the effects of time-varying marketing mix variables such as promotion, decisions on personalized discounts are a natural outcome of the proposed methodology.

To conclude, we hope our study will stimulate more research on inferring consumers' preference structures from complex assortment data and on advancing personalized recommendation systems for e-commerce practices.

APPENDIX I

Appendix A: Details of Our Application of the BBL Estimator

We apply the BBL estimator (Bajari et al. 2007) in the second stage of the model estimation process to estimate the structural parameters θ_1 and θ_2, which represent the marginal internal and external administrative cost, respectively. The main intuition of the BBL procedure is to find a set of parameters that best rationalizes the estimated policy function given the states (revealed preference), such that the merchants' observed policy yields the highest profits on average compared to using other simulated policies that deviate from the observed policy, while assuming that other merchants play the observed policy.

To do this, we compare profits from two scenarios: (1) when all merchants follow the observed policy and (2) when all merchants except merchant j follow the observed policies, but merchant j follows a policy that is slightly perturbed from its observed policy. We can rewrite merchant j's period profit, Π_{jt}, as the product of two vectors:

$$\Pi_{jt} = \Psi\left(\boldsymbol{s_t}, IC(n_{jt}), EC(a_{jt}), \epsilon_{jt}(a_{jt})\right)\boldsymbol{\theta}, \tag{A1}$$

where $\Psi\left(\boldsymbol{s_t}, IC(n_{jt}), EC(a_{jt}), \epsilon_{jt}(a_{jt})\right) = \left[GP_{jt}(\boldsymbol{s_t}), n_{jt}, a_{jt}, \epsilon_{jt}(a_{jt})\right]$ and $\boldsymbol{\theta} = [1, \theta_1, \theta_2, 1]$.

Let $\sigma_j(\boldsymbol{s}, \epsilon_j)$ be merchant j's Markov strategy, $\boldsymbol{\sigma} \equiv \{\sigma_j(\boldsymbol{s}, \epsilon_j): j = 1,2, \dots, N\}$ be a vector of strategy functions for each merchant, $\boldsymbol{\sigma}_{-j} = \boldsymbol{\sigma}(\boldsymbol{s}, \boldsymbol{\epsilon}) \setminus \{\sigma_j(\boldsymbol{s}, \epsilon_j)\}$ be a vector of all merchants' Markov strategies except merchant j's, and $\delta \in (0,1)$ be a fixed discount factor common to all merchants. Merchant j's discounted sum of expected profits at state $\boldsymbol{s}$ under the strategy $\boldsymbol{\sigma}$ is the following:

$$\begin{aligned} V_j(\boldsymbol{s}; \boldsymbol{\sigma}) &= E\left[\Sigma_{l=0}^{\infty} \delta^l \Psi_j\left(\boldsymbol{s}_l, \sigma_j(\boldsymbol{s}_l, \epsilon_{jl}), \epsilon_{jt}(a_{jt})\right)\boldsymbol{\theta} | \boldsymbol{\sigma_{-j}}, \boldsymbol{s_0} = \boldsymbol{s}\right] \\ &= E\left[\Sigma_{l=0}^{\infty} \delta^l \Psi_j\left(\boldsymbol{s}_l, \sigma_j(\boldsymbol{s}_l, \epsilon_{jl}), \epsilon_{jt}(a_{jt})\right) | \boldsymbol{\sigma_{-j}}, \boldsymbol{s_0} = \boldsymbol{s}\right]\boldsymbol{\theta}. \end{aligned} \tag{A2}$$

By definition, a Markov-perfect equilibrium is defined as a vector of Markov strategies $\boldsymbol{\sigma}^*$ such that (Bajari et al. 2007)

$$V_j\left(\boldsymbol{s}; \sigma_j^*, \boldsymbol{\sigma}_{-\boldsymbol{j}}^*\right) \geq V_j\left(\boldsymbol{s}; \sigma_j', \boldsymbol{\sigma}_{-\boldsymbol{j}}^*\right) \text{ for all } j, \boldsymbol{s} \in S, \text{ and } \sigma_j'. \tag{A3}$$

Let $W_j(\boldsymbol{s}; \boldsymbol{\sigma}) = E\left[\Sigma_{l=0}^{\infty} \delta^l \Psi_j\left(\boldsymbol{s}_l, \sigma_j(\boldsymbol{s}_l, \epsilon_{jl}), \epsilon_{jt}(a_{jt})\right) | \boldsymbol{\sigma}_{-\boldsymbol{j}}, \boldsymbol{s}_{\boldsymbol{0}} = \boldsymbol{s}\right]$. The equilibrium condition can be written as:

$$\left[W_j\left(\boldsymbol{s}; \sigma_j^*, \boldsymbol{\sigma}_{-\boldsymbol{j}}^*\right) - W_j\left(\boldsymbol{s}; \sigma_j', \boldsymbol{\sigma}_{-\boldsymbol{j}}^*\right)\right]\boldsymbol{\theta} \geq 0 \text{ for all } j, \boldsymbol{s} \in S, \text{ and } \sigma_j'. \tag{A4}$$

To implement BBL, we consider N_P of alternative policies such that $\sigma_j^1, \sigma_j^k, \ldots, \sigma_j^{NP}$, and σ_j^0 denotes merchant j's observed policy. Using the policy functions and the transition functions obtained in the first stage, we simulate each merchant's decision on a_{jt} over T periods for N_S times. We save the values of merchant j's expected single-period profit given states, changes in a number of deals, and its cost shock. Then for each k^{th} alternative policy, we calculate $W_j\left(\boldsymbol{s}; \sigma_j^k, \boldsymbol{\sigma}_{-\boldsymbol{j}}^{\boldsymbol{0}}\right)$ as:

$$W_j\left(\boldsymbol{s}; \sigma_j^k, \boldsymbol{\sigma}_{-\boldsymbol{j}}^{\boldsymbol{0}}\right) = \frac{1}{N_S} \sum_{n=1}^{N_S} \sum_{l=0}^{\mathrm{T}} \delta^l \, \widetilde{\Pi_{jt}}^{k,n}. \tag{A5}$$

After calculating $W_j\left(\boldsymbol{s}; \sigma_j^k, \boldsymbol{\sigma}_{-\boldsymbol{j}}^{\boldsymbol{0}}\right)$ for each alternative policy k, we can then evaluate the magnitude of the differences in values between when following the observed policy and when deviating under the parameter vector, $\boldsymbol{\theta}$, using the following equation:

$$g_j^k(\boldsymbol{\theta}) = \left[W_j\left(\boldsymbol{s}; \sigma_j^0, \boldsymbol{\sigma}_{-\boldsymbol{j}}^{\boldsymbol{0}}\right) - W_j\left(\boldsymbol{s}; \sigma_j^k, \boldsymbol{\sigma}_{-\boldsymbol{j}}^{\boldsymbol{0}}\right)\right]\boldsymbol{\theta}. \tag{A6}$$

Then we define the following quadratic loss function:

$$\left(min\{g_j^k(\boldsymbol{\theta}), 0\}\right)^2. \tag{A7}$$

The minimum function is equal to zero when the profits from following the observed policy σ_j^0 are greater than the profits from following k^{th} alternative policy σ_j^k. Otherwise, it yields the squared differences in the values between using two different policies. We estimate the structural cost parameters by minimizing the following:[18]

$$\min_{\boldsymbol{\theta}} \frac{1}{J}\frac{1}{N_P}\sum_{j=1}^{J}\sum_{k=1}^{N_P}\left(\min\{g_j^k(\boldsymbol{\theta}), 0\}\right)^2. \tag{A8}$$

The detailed steps are as follows:

1. For each merchant j, simulate a series of state variables, $\boldsymbol{x_t}$ and A_t over T periods for N_S times by using the transition functions specified in the first stage procedure. Denote $\{\widetilde{\boldsymbol{X_l^n}}\}_{l=1}^{T}$ be its n^{th} series. We use the values of the first observation in data for the initial value $\widetilde{\boldsymbol{X_1^n}}$.

2. Generate merchant j's cost shocks $\epsilon_{jl}^{n}\left(a_{jl}^{k,n}\right)$ over T periods N_S times by drawing from Type I extreme value distribution whose mean is zero and whose variance is equal to $\pi^2/6$.

3. Generate merchant j's N_P alternative policies by perturbing the estimates of the policy function obtained in the first stage. To do this, first, generate vectors such that $(\kappa^1, \ldots, \kappa^k, \ldots, \kappa^{N_P})$ of *i.i.d.* random draws from the standard normal distribution. The length of κ^k is equal to the number of the parameters in the policy functions. Second, perturb the estimates of the observed policy function by increasing the parameters by 0.5% (i.e., multiplying (1+0.005κ^k) to the parameter estimates). Third, get the fitted values of the conditional choice probabilities.

4. For every k alternative policy, simulate merchant j's expected gross profits: $\{\widetilde{GP}_{jl}^{k,n}\}_{l=0}^{T}$, the number of deals $\{\tilde{n}_{jl}^{k,n}\}_{l=0}^{T}$, and changes in a number of deals: $\{\tilde{a}_{jl}^{k,n}\}_{l=0}^{T}$ by using its cost shocks $\{\tilde{\epsilon}_{jl}^{n}\}_{l=0}^{T}$ when merchant j follows σ_j^k while the other merchants follow the observed policy $\boldsymbol{\sigma_{-j}^0}$. The detailed steps are followings:

[18] We set $N_P = 100, N_S = 50, T = 100$ in the estimation procedure.

(4.a) For the initial value of $\tilde{\boldsymbol{n}}_1^{k,n}$, we use the values of the first observation in data.

(4.b) Simulate the decision of all merchants at the first period $\tilde{\boldsymbol{a}}_1^{k,n}$ by following σ_j^k with ϵ_{j1}^n for merchant j while assuming that the other merchants follow the observed policy $\boldsymbol{\sigma}^{\mathbf{0}}_{-\boldsymbol{j}}$.

(4.c) Update the state variables $\left(\boldsymbol{n}_{\mathbf{2}}^{\boldsymbol{k},\boldsymbol{n}}, \boldsymbol{X}_{\mathbf{2}}^{\boldsymbol{n}}\right) = \left(\boldsymbol{n}_{\mathbf{1}}^{\boldsymbol{k},\boldsymbol{n}} + \boldsymbol{a}_{\mathbf{1}}^{\boldsymbol{k},\boldsymbol{n}}, \boldsymbol{X}_{\mathbf{2}}^{\boldsymbol{n}}\right)$.

(4.d) Repeat (b)-(c) T times to simulate a series of state variables, $\{\tilde{\boldsymbol{n}}_l^{k,n}, \tilde{\boldsymbol{X}}_l^n\}_l^T$ over T periods.

(4.e) Calculate merchant j's expected gross profit $\widetilde{GP}_{jl}^{k,n}$ by solving the profit maximization problem given the states, $\tilde{\boldsymbol{n}}_t^{k,n}$ and $\tilde{\boldsymbol{X}}_t^n$.

(4.f) Calculate $\widetilde{W}_j\left(\sigma_j^k, \boldsymbol{\sigma}^{\mathbf{0}}_{-\boldsymbol{j}}\right)$.

Appendix B: Additional Estimation Results and Data Descriptions

Table A1. Posterior Mean (and Standard Deviation) of the Heterogeneity Covariance Matrix Σ_Λ

	Merchant A	***Merchant B***	***Merchant C***	***Merchant D***	***Merchant E***	***Average depth of discount***	***Regular price***	***Number of deals***
Merchant A	4.571 (1.140)							
Merchant B	-0.418 (0.271)	0.585 (0.168)						
Merchant C	0.185 (1.123)	0.404 (0.693)	12.886 (4.474)					
Merchant D	-0.784 (0.388)	-0.025 (0.195)	-3.299 (0.852)	1.186 (0.310)				
Merchant E	-0.858 (0.476)	0.423 (0.309)	4.741 (1.290)	-1.144 (0.347)	2.372 (0.653)			
Average depth of discount	-0.051 (0.231)	-0.083 (0.130)	-1.784 (0.529)	0.559 (0.168)	-0.709 (0.221)	0.600 (0.173)		
Regular price	0.339 (0.230)	-0.037 (0.127)	1.548 (0.420)	-0.438 (0.115)	0.468 (0.170)	-0.287 (0.112)	0.559 (0.149)	
Number of deals	-0.235 (0.090)	0.025 (0.043)	-0.190 (0.119)	0.081 (0.037)	-0.101 (0.063)	-0.004 (0.038)	-0.094 (0.042)	0.182 (0.022)

Table A2. Posterior Mean (and Standard Deviation) of the Covariance Matrix of Common Demand shocks, ξ, and Error Terms for the Depth of Discount η across Merchants

	ξ_A	ξ_B	ξ_C	ξ_D	ξ_E	η_A	η_B	η_C	η_D	η_E
ξ_A	2.515 (0.596)									
ξ_B	1.663 (0.456)	1.832 (0.455)								
ξ_C	0.301 (0.445)	0.329 (0.346)	1.462 (0.444)							
ξ_D	-0.226 (0.312)	-0.225 (0.241)	-0.489 (0.267)	1.878 (0.433)						
ξ_E	0.455 (0.621)	0.686 (0.459)	0.815 (0.438)	1.333 (0.492)	2.980 (0.877)					
η_A	-0.023 (0.065)	0.000 (0.051)	0.128 (0.027)	-0.066 (0.034)	0.062 (0.024)	0.214 (0.010)				
η_B	0.079 (0.091)	0.055 (0.064)	-0.223 (0.050)	-0.026 (0.047)	-0.264 (0.046)	-0.045 (0.008)	0.266 (0.012)			
η_C	0.154 (0.041)	0.098 (0.035)	0.103 (0.027)	-0.102 (0.033)	0.006 (0.034)	-0.03 (0.007)	-0.013 (0.007)	0.240 (0.011)		
η_D	0.158 (0.066)	0.14 (0.052)	0.018 (0.021)	0.021 (0.028)	0.059 (0.027)	-0.007 (0.008)	0.011 (0.008)	0.032 (0.008)	0.247 (0.014)	
η_E	-0.139 (0.068)	-0.07 (0.049)	-0.018 (0.022)	-0.011 (0.035)	0.006 (0.042)	0.029 (0.007)	-0.017 (0.008)	0.036 (0.007)	-0.022 (0.008)	0.247 (0.016)

Table A3. Changes in the Number of Deals by Each Merchant

Change in # Deals	Merchant A	Merchant B	Merchant C	Merchant D	Merchant E	Sum
(...,-9]	1	4	0	2	3	10
(-9,-7]	5	2	0	1	3	11
(-7,-5]	1	3	1	5	5	15
(-5,-3]	5	7	6	2	6	26
(-3,-1]	4	6	7	7	5	29
0	6	2	11	4	6	29
[1,3)	7	5	13	12	3	40
[3,5)	10	4	6	5	4	29
[5,7)	3	5	1	2	2	13
[7,9)	2	1	0	3	5	11
[9, ...)	1	6	0	2	3	12

Appendix C: Procedures for Solving the Dynamic Game in Counterfactual Simulations

The counterfactual analyses involve searching for the merchants' optimal strategies of offering the number of deals in response to the changes in counterfactual scenarios. Following prior research in the literature (e.g., Yao and Mela 2011), we adopt an approximate dynamic programming approach to deriving the equilibrium outcomes under each scenario. This iterated best response approach uses a parametric policy function and assumes that the parametric policy function is flexible enough to describe merchant strategies.

Denote the merchants' policy parameter for offering the number of deals as $\Omega_{\boldsymbol{J}} = \{\boldsymbol{\psi}, \boldsymbol{Merchants}\}$, where $\boldsymbol{\psi} = \{\psi_1, \psi_2, \psi_3, \psi_4\}$ is a vector of parameters as in Equation (14), which are common across merchants. $\boldsymbol{Merchants}$ is a vector of merchant-specific fixed effects. These policy function parameters are assumed to change under different counterfactual scenarios. Our goal is to find the set of optimal parameters $\boldsymbol{\psi}$ that maximize the expected future profit streams of merchants in response to different counterfactuals. Also denote $\Omega_{\boldsymbol{j}} = \{\boldsymbol{\psi}, Merchant_j\}$ and $\Omega_{-\boldsymbol{j}} = \{\boldsymbol{Merchants}_{-j}\}$, where j is a focal merchant.

The algorithm is as follows. First, we chose a merchant as the focal merchant, indexed as j. Second, we use the values of the first period observation in the observed data as the initial state variables $\boldsymbol{x_0}$ and A_0. Third, for all merchants j and period t we draw random shock ϵ_{jt}, for $T = 52$. We repeat this step $N_{PS} = 10$ times. Fourth, we search for parameters of the parametric policy function as in Equation (12) that maximize the expected future profit streams of merchants in a dynamic oligopoly game. Within the fourth step, parameters are chosen for the merchants and the detailed steps as follows.

1. Set $\Omega_{\boldsymbol{J}} = \{\boldsymbol{\psi}, \boldsymbol{Merchants}\}$ using the estimates of the policy function.

2. For a given set of parameters $\Omega_{\boldsymbol{j}} = \{\boldsymbol{\psi}, Merchant_j\}$, we calculate the following value of the expected future profit streams of merchants j:

$$\nu_j\left(\boldsymbol{x_0}, A_0; \sigma_j, \boldsymbol{\sigma_{-j}}, \boldsymbol{\epsilon};\ \ \Omega_{\boldsymbol{j}}, \Omega_{\boldsymbol{-j(old)}}\right)$$

$$= \Pi_j\left(\boldsymbol{x_0}, A_0; \sigma_j, \boldsymbol{\sigma_{-j}}, \epsilon_0;\ \ \Omega_{\boldsymbol{j}}, \Omega_{\boldsymbol{-j(old)}}\right) +$$

$$\Sigma_{t=1}^{T} \delta^t \Pi_j\left(\boldsymbol{x_t}, A_t; \sigma_j, \boldsymbol{\sigma_{-j}}, \epsilon_t;\ \ \Omega_{\boldsymbol{j}}, \Omega_{\boldsymbol{-j(old)}}\right) P(A_t | \boldsymbol{a_{t-1}}, A_{t-1}, \boldsymbol{x_{t-1}}).$$

3. Repeat step 2 above for N_{PS} times and average ν_j. This gives an approximation of the value function for merchant j.

$$\widehat{\nu_j}\left(\boldsymbol{x_0}, A_0; \sigma_j, \boldsymbol{\sigma_{-j}}, \boldsymbol{\epsilon};\ \ \Omega_{\boldsymbol{j}}, \Omega_{\boldsymbol{-j(old)}}\right) = \frac{1}{N_{PS}} \sum_{n=1}^{N_{PS}} \nu_j\left(\boldsymbol{x_0}, A_0; \sigma_j, \boldsymbol{\sigma_{-j}}, \boldsymbol{\epsilon^{(n)}};\ \ \Omega_{\boldsymbol{j}}, \Omega_{\boldsymbol{-j(old)}}\right)$$

4. Search for the $\Omega_{\boldsymbol{j(new)}} = \{\boldsymbol{\psi_{(new)}}, Merchant_{j(new)}\}$ that maximizes $\widehat{\nu_j}$ above

$$\Omega_{\boldsymbol{j(new)}} = argmax_{\Omega_{\boldsymbol{j}}} \widehat{\nu_j}\left(\boldsymbol{x_0}, A_0; \sigma_j, \boldsymbol{\sigma_{-j}}, \boldsymbol{\epsilon};\ \ \Omega_{\boldsymbol{j}}, \Omega_{\boldsymbol{-j(old)}}\right)$$

5. Iterate through all merchants one by one using the steps 2 and 4 to search for updated fixed effects $\boldsymbol{Merchants}_{-j}$, conditioned on the updated common parameters $\boldsymbol{\psi_{(new)}}$ for all merchants and the fixed effects for the other merchants.

6. Conditioned on new parameters $\boldsymbol{\psi_{(new)}}$ and $\boldsymbol{Merchants_{(new)}}$, calculate the updated value function for all merchants: $\widehat{\nu_j}\left(\boldsymbol{x_0}, A_0; \sigma_j, \boldsymbol{\sigma_{-j}}, \boldsymbol{\epsilon};\ \ \Omega_{\boldsymbol{j(new)}}, \Omega_{\boldsymbol{-j(new)}}\right) =$ $\frac{1}{N_{PS}} \sum_{n=1}^{N_{PS}} \nu_j\left(\boldsymbol{x_0}, A_0; \sigma_j, \boldsymbol{\sigma_{-j}}, \boldsymbol{\epsilon^{(n)}};\ \ \Omega_{\boldsymbol{j(new)}}, \Omega_{\boldsymbol{-j(new)}}\right), \forall j$

If the following condition is met, then stop the iteration. Otherwise, go back to 2.

$$\underset{j}{\text{Max}} \left| \widehat{\nu_j}\left(\boldsymbol{x_0}, A_0; \sigma_j, \boldsymbol{\sigma_{-j}}, \boldsymbol{\epsilon};\ \ \Omega_{\boldsymbol{j(new)}}, \Omega_{\boldsymbol{-j(new)}}\right) - \widehat{\nu_j}\left(\boldsymbol{x_0}, A_0; \sigma_j, \boldsymbol{\sigma_{-j}}, \boldsymbol{\epsilon};\ \ \Omega_{\boldsymbol{j(old)}}, \Omega_{\boldsymbol{-j(old)}}\right) \right| \leq 1.$$

APPENDIX II

APPENDIX A: DATA GENERATING PROCESS

Topics

- For each topic $m = 1, \dots, M$:
  - For each discrete product attribute $a = 1, \dots, A$:
    - Draw a probability vector $\boldsymbol{\phi}_m^{[a]}$ over the $|a|$ levels for attribute a:

$$\boldsymbol{\phi}_m^{[a]} \sim \text{Dirichlet}_{|a|}(\mathbf{1}\zeta_a)$$

  - For the $|\boldsymbol{y}|$ continuous product attributes jointly:
    - Draw a covariance matrix $\boldsymbol{\Sigma}_m$: $\boldsymbol{\Sigma}_m \sim \boldsymbol{\omega}_{|\boldsymbol{y}|}^{-1}(\boldsymbol{\Psi}, \nu)$, where $\boldsymbol{\omega}^{-1}$ is the inverse Wishart distribution.
    - Draw a mean vector $\boldsymbol{\mu}_m$:

$$\boldsymbol{\mu}_m \sim \text{Normal}_{|\boldsymbol{y}|}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_m)$$

Browsing activities

Subsequently, we model the nth browsed product $\boldsymbol{d}_{isn}^{[b]}$ in consumer i's sth session, by assigning $\boldsymbol{d}_{isn}^{[b]}$ to a particular topic. This assignment is denoted by $z_{isn}^{[b]} \in \{1, \dots, M\}$ and the corresponding assignment probabilities are given by $\boldsymbol{\theta}_{is}^{[b]}$. More formally, the generative process for browsing activities is as follows:

- For each consumer, $i = 1, \dots, I$:
  - For each of consumer i's session $s = 1, \dots, S_i$:
    - Draw a distribution over the M topics: $\boldsymbol{\theta}_{is}^{[b]}$
      - For each browsed product $n = 1, \dots, N_{is}^{[b]}$:
      - Draw the topic assignment for this browsed product:

$$z_{isn}^{[b]} \sim \text{Categorical}_M(\boldsymbol{\theta}_{is}^{[b]})$$

• For the vector of attributes $\boldsymbol{d}_{isn}^{[b]} = (\boldsymbol{c}_{isn}^{[b]},\ \boldsymbol{y}_{isn}^{[b]})$:

- For each discrete attribute $a = 1,\ \dots,\ A$:

  • Draw the level of the discrete attribute a from the probability vector corresponding to the topic assignment $z_{isn}^{[b]}$:

$$c_{isn,a}^{[b]} | z_{isn}^{[b]} \sim \text{Categorical}_{|a|}(\boldsymbol{\phi}_{z_{isn}^{[b]}}^{[a]})$$

- Combine the A discrete attributes into $\boldsymbol{c}_{isn}^{[b]} = (c_{isn,1}^{[b]}, \dots, c_{isn,A}^{[b]})$.

- For the continuous attributes $\boldsymbol{y}_{isn}^{[b]}$ jointly:

  • Draw the value of the continuous attributes given the multivariate Normal distribution corresponding to the topic assignment $z_{isn}^{[b]}$:

$$\boldsymbol{y}_{isn}^{[b]} | z_{isn}^{[b]} \sim \text{Normal}_{|\boldsymbol{y}|}(\boldsymbol{\mu}_{z_{isn}^{[b]}}, \boldsymbol{\Sigma}_{z_{isn}^{[b]}}).$$

- Combine the vector of discrete attributes $\boldsymbol{c}_{isn}^{[b]}$ and the vector of continuous attributes $\boldsymbol{y}_{isn}^{[b]}$ into $\boldsymbol{d}_{isn}^{[b]} = (\boldsymbol{c}_{isn}^{[b]},\ \boldsymbol{y}_{isn}^{[b]})$.

Purchase Activities

Given $\boldsymbol{\theta}_{is}^{[b]}$, the generative process for the observed purchases $\boldsymbol{d}_{is.}^{[p]}$ is similar to the generative process of browsing behavior $\boldsymbol{d}_{is.}^{[b]}$, except that the distribution over the M topics are drawn as $\boldsymbol{\theta}_{is}^{[p]} \sim \text{Dirichlet}_M\left(\exp\left(\boldsymbol{R}_1'\boldsymbol{\theta}_{is}^{[b]}\right),\ \dots,\ \exp\left(\boldsymbol{R}_M'\boldsymbol{\theta}_{is}^{[b]}\right) \right)$.

APPENDIX B: PRIOR SPECIFICATIONS IN THE PROPOSED MODEL

The parameters ζ_a control the sparseness of the Dirichlet distribution for $\boldsymbol{\phi}^{[a]}$. Smaller values for ζ_a lead to sparser distributions, while larger values lead to more diffuse distributions. Following Jacobs et al. (2017), we place a log-Normal prior distribution on ζ_a, with parameters $\mu = -\log(|a|)\frac{|a|+1}{|a|}$ and $\sigma^2 = 2\log(|a|)\frac{1}{|a|}$. The mean of the log-normal distribution on ζ_a is $1/|a|$, which favors sparse $\boldsymbol{\phi}^{[a]}$ vectors and makes interpretation of the topics easier. Following Das et al. (2015), we place conjugate priors on the vector of continuous product attributes: a Gaussian centered at zero ($\boldsymbol{\mu_0}$) and the covariance ($\boldsymbol{\Sigma}_m$) from an inverse Wishart distribution ($\boldsymbol{\omega}^{-1}$) with hyperparameters $\boldsymbol{\Psi} = 4 \cdot \boldsymbol{I}$, which is a positive definite matrix and $\nu = 5$. For all parameters in the specification of α_{ism}, we specify standard Normal priors. For the variance parameter σ_m, we choose half-Cauchy (0, 5) as a weak prior (Gelman et al. 2006). Lastly, $\boldsymbol{R}$ is considered to be a hyper-parameter that is estimated without a proper prior distribution (Liu and Toubia 2018).

APPENDIX C: TOP 5 MOST LIKELY PRODUCTS FOR EACH OF THE 15 TOPICS

Category	Brand	Gender	Color	Initial Price (¥)	Discount (%)
Topic 1					
Backpacks	6	Women	Mixed Colors	427	50
Backpacks	6	Women	Mixed Colors	497	50
Backpacks	6	Women	Mixed Colors	527	50
Backpacks	6	Women	Mixed Colors	397	50
Backpacks	6	Women	Mixed Colors	467	50
Topic 2					
Jackets & Blazers	2	Women	Black	1,800	70
Shirts & Blouses	13	Women	Black	2,290	70
Sweaters	2	Women	Black	2,800	70
Shirts & Blouses	13	Women	Black	2,990	70
Shirts & Blouses	2	Women	Black	2,100	70
Topic 3					
Dresses	16	Women	Blue/Navy	1,690	40
Dresses	16	Women	Blue/Navy	1,890	42

Dresses	16	Women	Blue/Navy	2,090	42
Sweaters	16	Women	Black	1,890	42
Handbags	16	Women	Black	1,890	42
Topic 4					
Dresses	21	Women	White/Off-White	599	12
Dresses	21	Women	White/Off-White	639	12
Dresses	21	Women	White/Off-White	569	12
Dresses	21	Women	White/Off-White	699	12
Dresses	18	Women	White/Off-White	639	12
Topic 5					
Loafers & Other Casual	1	Women	Blue/Navy	399	32
Loafers & Other Casual	1	Women	Blue/Navy	369	14
Loafers & Other Casual	1	Women	Blue/Navy	399	30
Loafers & Other Casual	1	Men	Blue/Navy	439	30
Loafers & Other Casual	1	Women	Red	399	32
Topic 6					
Loafers & Other Casual	9	Men	Black	1,899	37
Loafers & Other Casual	9	Women	Black	2,099	38
Loafers & Other Casual	9	Men	Black	1,999	20
Loafers & Other Casual	9	Men	Black	2,499	60
Loafers & Other Casual	9	Men	Black	1,999	15
Topic 7					
Overcoats	21	Women	Black	1,590	12
Overcoats	21	Women	Black	1,290	12
Overcoats	21	Women	Black	1,390	12
Overcoats	18	Women	Black	1,390	12
Overcoats	18	Women	Blue/Navy	1,190	12
Topic 8					
Tees & Knits	8	Men	White/Off-White	379	50
Tees & Knits	8	Men	Black	399	55
Tees & Knits	8	Men	Black	359	42
Tees & Knits	8	Men	Black	449	70
Tees & Knits	8	Men	Black	379	42
Topic 9					
Jewel	16	Women	Yellow/Gold/Orange	690	42
Jewel	16	Women	Yellow/Gold/Orange	590	42
Jewel	16	Women	Yellow/Gold/Orange	590	50
Jewel	16	Women	Yellow/Gold/Orange	650	42
Jewel	16	Women	Yellow/Gold/Orange	850	42
Topic 10					
Jackets & Blazers	5	Men	Black	1,599	52
Jackets & Blazers	5	Men	Black	1,499	52

Jackets & Blazers	5	Men	Black	2,199	55
Jackets & Blazers	5	Men	Black	1,490	52
Jackets & Blazers	5	Men	Black	1,090	52
Topic 11					
Tees & Knits	18	Women	White/Off-White	599	52
Tees & Knits	21	Women	White/Off-White	639	52
Tees & Knits	21	Women	White/Off-White	499	52
Tees & Knits	18	Women	White/Off-White	439	52
Tees & Knits	18	Women	Yellow/Gold/Orange	499	52
Topic 12					
Shirts & Blouses	8	Men	Blue/Navy	839	50
Shirts & Blouses	8	Men	Black	739	50
Pants	8	Men	Black	839	50
Shirts & Blouses	8	Men	Blue/Navy	929	50
Pants	8	Men	Black	839	60
Topic 13					
Tees & Knits	14	Women	Black	519	12
Tees & Knits	20	Women	Black	549	20
Tees & Knits	14	Men	Black	499	10
Tees & Knits	20	Women	Blue/Navy	499	20
Tees & Knits	20	Women	White/Off-White	599	20
Topic 14					
Scarves & Wraps	7	Women	Black	299	80
Handbags	7	Women	Black	339	71
Tees & Knits	7	Women	Black	299	84
Others	7	Women	Black	299	74
Handbags	7	Women	Black	299	80
Topic 15					
Tees & Knits	12	Women	Black	419	55
Tees & Knits	5	Women	Black	489	52
Tees & Knits	5	Women	White/Off-White	479	55
Tees & Knits	12	Women	Black	359	50
Tees & Knits	12	Women	White/Off-White	369	50

APPENDIX D: BAYESIAN OPTIMIZATION

Bayesian optimization has been used to solve the mixed-integer optimization problems such as finding hyperparameters for machine learning models, and it is proven to be very effective in terms of reducing computational burden compared to methods such as grid or random search

(Shahriari et al. 2016). Frazier (2018) provides an excellent introduction to the Bayesian optimization algorithm.

Our application of the Bayesian optimization algorithm consists of two steps. First, we fit a Gaussian process to a few observations of a pair of input parameters and a value of the objective function, which enables us to make a probabilistic inference on potential values of the objective function at input parameters that are not yet evaluated. Second, a set of parameters to explore next is chosen that maximizes a criteria function (also known as an acquisition function). The acquisition function balances the trade-off between exploitation and exploration so that it maximizes the efficiency of finding optimal values. We repeat the two steps until the predictive mean value of the Gaussian process for the objective function is maximized, or prespecified rounds are reached.

Using the Bayesian optimization algorithm, we find vectors of optimal attributes that jointly maximize the marginal likelihoods for each consumer. The steps are as follows:

1. Generate initial n samples of attribute vectors $\boldsymbol{D} = \{\boldsymbol{d}_1, \boldsymbol{d}_2, \dots, \boldsymbol{d}_n\}$ and associated objective function value $\boldsymbol{G} = \{\boldsymbol{g}_1, \boldsymbol{g}_2, \dots, \boldsymbol{g}_n\}$.
2. Set iteration index $c = 1$
3. If $c < C$ do Step 4-6, otherwise go to Step 7.
4. Maximize the acquisition function $acq(\boldsymbol{d})$ using the posterior mean and variance under the Gaussian process assumption:
$$\boldsymbol{d}_{n+1} = argmax_{\boldsymbol{d}}\, acq(\boldsymbol{d}).$$
5. Evaluate the objective function and set $\boldsymbol{g}_{n+1} = \boldsymbol{g}(\boldsymbol{d}_{n+1})$.
6. Augment the observations as

$\boldsymbol{D} = \boldsymbol{D} \cup \{\boldsymbol{d}_{n+1}\}$ and $\boldsymbol{G} = \boldsymbol{G} \cup \{\boldsymbol{g}_{n+1}\}$, set $c = c + 1$, and go back to Step 3.

7. Choose the attribute vectors $\boldsymbol{d}$ with the first vth maximum value of $\boldsymbol{g}$.

In our empirical application, we use $\beta_n = 2.5$ for the hyperparameter of the Gaussian process upper confidence bound, following Cox and John (1997). For the covariance function for the Gaussian process, we use the squared exponential kernel. We start with the initial random sample of 40 pairs of attributes. The number of iterations in Bayesian optimization is fixed to 30, after which pairs of attributes that give the first v highest-ranked likelihood is chosen for v recommendations.

REFERENCES

Adomavicius, Gediminas, and Alexander Tuzhilin (2005), "Toward the Next Generation of Recommender Systems: A survey of the State-of-the-art and Possible Extensions," *IEEE Trans. on Know. & Data Engine*, 17 (6), 734-749.

Aguirregabiria, Victor, and Pedro Mira (2007), "Sequential estimation of dynamic discrete games," *Econometrica*, 75 (1), 1-53.

Amigobulls (2018) Vipshop Revenue, Profits - VIPS Annual Income Statement. https://amigobulls.com/stocks/VIPS/income-statement/annual, accessed on July 29, 2018.

Ansari, Asim, Skander Essegaier, and Rajeev Kohli (2000), "Internet Recommendation Systems," *Journal of Marketing Research,* 37 (3), 363-375.

Ansari, Asim, Yang Li, and Jonathan Z. Zhang (2018), "Probabilistic Topic Model for Hybrid Recommender Systems: A Stochastic Variational Bayesian Approach," *Marketing Science*, 37 (6), 987-1008.

Ansari, Asim, and Carl F. Mela (2003), "E-Customization," *Journal of Marketing Res*earch, 40 (2), 131-145.

Arellano, Manuel, and Stephen Bond (1991), "Some tests of specification for panel data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58 (2), 277-297.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin (2007), "Estimating dynamic models of imperfect competition," *Econometrica*, 75 (5), 1331-1370.

Bellman R (1957) *Dynamic Programming* (Princeton University Press, Princeton, NJ).

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile prices in market equilibrium," *Econometrica*, 63 (4), 841–890.

Berry, Steven (1994), "Estimating discrete-choice models of product differentiation," *RAND Journal of Economics*, 25 (2), 242–262.

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

Blei, David M., Alp Kucukelbir, and Jon D. McAuliffe (2017), "Variational inference: A Review for Statisticians," *Journal of American Statistical Association*, 112 (518), 859-877.

Blei, David M., and John D. Lafferty (2007), "A Correlated Topic Model of Science," *The Annals of Applied Statistics*, 1 (1), 17-35.

Blevins, Jason, Ahmed Khwaja, and Nathan Yang (2018), "Firm expansion, size spillovers, and market dominance in retail chain dynamics," *Management Science*, 64 (9), 3971-4470.

Bucklin, Randolph E., and Catarina Sismeiro (2003), "A model of Web Site Browsing Behavior Estimated on Clickstream Data," *Journal of Marketing Research*, 40 (3), 249-267.

Cao, Zike, Kai-Lung Hui, and Hong Xu (2018), "When discounts hurt sales: The case of daily-deal markets," *Information Systems Research*, 29 (3), 525-777.

Chernev, Alexander, Ulf Böckenholt, and Joseph Goodman (2015), "Choice Overload: A Conceptual Review and Meta-analysis," *Journal of Consumer Psychology*, 25 (2), 333-358.

Chu Junhong, and Puneet Manchanda (2016), "Quantifying cross and direct network effects in online consumer-to-consumer platforms," *Marketing Science*, 35 (6), 870-893.

Chung, Tuck Siong, Roland T. Rust, and Michel Wedel (2009), "My Mobile Music: An Adaptive Personalization System for Digital Audio Players," *Marketing Science*, 28 (1), 52–68.

Chung, Tuck Siong, Michel Wedel, and Roland T. Rust (2016), "Adaptive Personalization Using Social Networks," *Journal of the Academy of Marketing Science*, 44 (1), 66–87.

Cnova (2018) CNOVA N.V. 2017 Financial Results. http://www.cnova.com/en/wp-content/uploads/sites/2/2018/02/2018-02-20-Cnova-FY17-Results-ENG-VF.pdf, accessed on July 29, 2018.
Cox, Dennis D., and Susan John (1997), "SDO: A Statistical Method for Global Optimization," *Multidisciplinary design optimization: State-of-the-Art. SIAM, Philadelphia, PA,* 315-329.
Danaher, Peter J., Isaac W. Wilson, and Robert A. Davis (2003), "A Comparison of Online and Offline Consumer Brand Loyalty," *Marketing Science*, 22 (4), 461–476.
Das, Rajarshi, Manzil Zaheer, and Chris Dyer (2015), "Gaussian LDA for Topic Models with Word Embeddings," *Proc. 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, 795-804.
Doraszelski, Ulrich, and Mark Satterthwaite (2010), "Computable Markov-perfect industry dynamics," *RAND Journal Economics*, 41 (2), 215-243.
Dubé, Jean-Pierre, Günter J. Hitsch, and Puneet Manchanda (2005), "An empirical model of advertising dynamics," *Quantitative Marketing and Economics*, 3 (2), 107-144.
Edelman, Benjamin, Sonia Jaffe, Scott Duke Kominers, (2016), "To groupon or not to groupon: The profitability of deep discounts," *Marketing Letters*, 27 (1), 39-53.
Erdem, Tülin, and Michael P. Keane (1996), "Decision making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets," *Marketing Science*, 15 (1), 1-20.
Ericson, Richard, and Ariel Pakes (1995), "Markov-perfect industry dynamics: A framework for empirical work," *Review of Economic Studies*, 62 (1), 53-82.
Evans, David S., and Richard Schmalensee (2016), *Matchmakers: The New Economics of Multisided Platforms*, Harvard Business Review Press, Boston, MA.
Fader, Peter S., and Bruce G. S. Hardie (1996), "Modeling Consumer Choice among SKUs," *Journal of Marketing Research*, 33 (4), 442-452.
Farias, Vivek, Denis Saure, and Gabriel Y. Weintraub (2012), "An approximate dynamic programming approach to solving dynamic oligopoly models," *RAND Journal of Economics*, *43* (2), 253-282.
Frazier, Peter I. (2018), "Bayesian Optimization," *INFORMS Tutorials in Operations Research*, 255-278.
Gelman, Andrew (2006), "Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article by Browne and Draper)," *Bayesian Analysis*, 1 (3), 515-534.
Goldfarb, Avi, Qiang Lu, and Sridhar Moorthy (2009), "Measuring brand value in an equilibrium framework," *Marketing Science*, 28 (1), 69-86.
Guadagni, Peter M., and John D. C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (3), 203–238.
Ho, Teck-Hua, and Juin-Kuan Chong (2003), "A Parsimonious Model of Stockkeeping-Unit-Choice", *Journal of Marketing Research*, 40 (3), 351-365.
Hotz, V. Joseph, and Robert A. Miller (1993), "Conditional choice probabilities and the estimation of dynamic models," *Review Economic Studies*, 60 (3), 497-529.
Howe, Michael (2009), "Pandora's Music Recommender", https://courses.cs.washington.edu/courses/csep521/07wi/prj/michael.pdf, accessed on May 23, 2019.

Inman, J. Jeffery, Joonwook Park, and Ashish Sinha (2008), "A Dynamic Choice Map Approach to Modeling Attribute-level Varied Behavior among Stock-Keeping Units," *Journal of Marketing Research*, 45 (1), 94-103.

Jacobs, Bruno J.D., Bas Donkers, and Dennis Fok (2016), "Model-Based Purchase Predictions for Large Assortments," *Marketing Science,* 35 (3), 341-537.

Jacobs, Bruno J.D., Bas Donkers, and Dennis Fok (2017), "Model-Based Marketing Insights from High-Dimensional Purchase Data", Working paper, Erasmus University, Rotterdam.

Jeon, Seongmin, Anindya Ghose, Jongmun Yoon, and Byungjoon Yoo (2014), "Attention economy in online daily deals: Demand estimation using structural models," *Proc. 35th International Conference on Information Systems Conf. 2014* (Association for Information Systems, Auckland, New Zealand), 1-8.

Jordan, Michael I., Zoubin Ghahramani, Tommi S. Jaakkola, and Lawrence K. Saul (1999) "An Introduction to Variational Methods for Graphical Models," *Machine Learning*. 37 (2), 183-233.

Karypis, George (2001), "Evaluation of Item-Based Top-N Recommendation Algorithms," *Proc. 10th International Conference on Information and Knowledge Management*, 247-254.

Kucukelbir, Alp, Dustin Tran, Rajesh Ranganath, Andrew Gelman, and David M. Blei (2017) "Automatic Differentiation Variational Inference," *Journal of Machine Learning Research*. 18 (14), 1-45.

Kumar, V., and Bharath Rajan (2012), "Social coupons as a marketing strategy: A multifaceted perspective," *Journal of the Academy of Marketing Science,* 40 (1), 120-136.

Linden, Greg, Brent Smith, and Jeremy York (2003), "Amazon.com Recommendations: Item-To-Item Collaborative Filtering," *IEEE Inter. Comp*., 76-80.

Liu, Jia, and Oliver Toubia (2018), "A Semantic Approach for Estimating Consumer Content Preferences from Online Search Queries," *Marketing Science*, 37 (6), 855-1052.

Liu, Qiang, Sachin Gupta, Sriram Venkataraman, and Liu Hongju (2016), "An empirical model of drug detailing: Dynamic competition and policy implications," *Management Science*, 62 (8), 2321-2340.

Lu, Shasha, Li Xiao, and Ming Ding (2016), "A Video-Based Automated Recommender (VAR) Systems for Garments," *Marketing Science,* 35(3), 484-510.

Ma, Wenping, Xiang Feng, Shanfeng Wang, and Maoguo Gong (2016), "Personalized Recommendation Based on Heat Bidirectional Transfer," *Physica A:Statistical Mechanics and its Applications*, 444, 713-721.

Montgomery, Alan L., Shibo Li, Kannan Srinivasan, and John C. Liechty (2004), "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science,* 23 (4), 579-595.

Nevo, Aviv (2000) "Mergers with differentiated products: the case of the ready-to-eat cereal industry," *RAND Journal of Economics*, 31 (3), 395-421.

Ngwe, Donald (2017), "Why outlet stores exist: averting cannibalization in product line extensions," *Marketing Science*, 36 (4), 523-541.

Pierson, Margaret, Gad Allon, and Awi Federgruen (2016) "Price competition under mixed multinomial logit demand functions," *Management Science*, 59 (8), 1817-1835.

Puranam, Dinesh, Vishal Narayan, and Vrinda Kadiyali (2017), "The Effect of Calorie Posting Regulation on Consumer Opinion: A Flexible Latent Dirichlet Allocation Model with Informative Priors," *Marketing Science*, 36 (5), 726-746.

Ricci, Francesco, Lior Rokach, and Bracha Shapira (2011), "Introduction to Recommender Systems Handbook," *Recommender Systems Handbook* (Springer, Boston, MA).
Rochet, Jean-Charles, and Jean Tirole (2006), "Two-sided markets: A progress report," *RAND Journal of Economics*, 37 (3), 645-667.
Rooderkerk, Robert P., Harald J. van Heerde, and Tammo H.A. Bijmolt (2013), "Optimizing Retail Assortments," *Marketing Science*, 32 (5), 699-715.
Rust, John (1987), "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica*, 55 (5), 999-1033.
Schweidel, David A., and Wendy W. Moe (2016), "Binge Watching and Advertising," *Journal of Marketing, 80 (5)*, 1-19.
Shahriari, Bobak, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas (2016), "Taking the Human Out of the Loop: A Review of Bayesian Optimization," *Proceedings of the IEEE*, 104 (1), 148-175.
Shi, Savannah Wei, and Jie Zhang (2014), "Usage Experience with Decision Aids and Evolution of Online Purchase Behavior," *Marketing Science*, 33 (6), 871-882.
Smiley (2019), "Stitch Fix's Radical Data-Driven Way to Sell Clothes–$1.2 Billion Last Year–is Reinventing Retail," https://www.fastcompany.com/90298900/stitch-fix-most-innovative-companies-2019, accessed on May 28, 2018.
Smith, Michael D., and Erik Brynjolfsson (2001), "Consumer Decision-Making at an Internet Shopbot: Brand Still Matters," *Journal of Industrial Economics*, 49 (4), 541-558.
Stan Development Team (2018), "Stan Modeling Language Users Guide and Reference Manual," *http://mc-stan.org*, 2018. Version 2.18.
Statista (2016) Gross merchandise value of Cdiscount.com in France from 2012 to 2016. https://www.statista.com/statistics/758156/gross-merchandise-value-cdiscount-france/, accessed on July 29, 2018.
Sun, Baohong (2005), "Promotion effect on endogenous consumption," *Marketing Science*, 24 (3), 430-443.
Suzuki, Junichi (2013), "Land use regulation as a barrier to entry: evidence from the Texas lodging industry," *International Economic Review*, 54 (2), 495–523.
Tirunillai, Seshadri, and Gerard J. Tellis (2014), "Mining Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation," *Journal of Marketing Research*, 51 (4), 463-479.
Toubia, Oliver, Garud Iyengar, Renèe Bunnell, and Alain Lemaire (2019), "Extracting Features of Entertainment Products: A Guided Latent Dirichlet Allocation Approach Informed by the Psychology of Media Consumption," *Journal of Marketing Research*, 56 (1), 18-36.
Verlinda, Jeremy (2005), "A Bayesian analysis of tree structure specification in nested logit models," *Economics Letters*, 87 (1), 67-73.
Villas-Boas, J. Miguel, and Ying Zhao (2005), "Retailer, manufacturers, and individual consumers: Modeling the supply side in the ketchup marketplace," *Journal of Marketing Research*, 42 (1), 83-95.
Wu, Chunhua, Yitian Liang, and Xinlei Chen (2015), "Is daily deal a good deal for merchants? An empirical analysis of economic value in the daily deal market," Working paper, University of British Columbia, Vancouver.
Xu, Yunjie (Calvin), and Hee-Woong Kim (2008), "Order Effect and Vendor Inspection in Online Comparison Shopping," *Journal of Retailing*, 84 (4), 477-486.
Yang, Sha, Yuxin Chen, and Greg M. Allenby (2003), "Bayesian analysis of simultaneous

demand and supply," *Quantitative Marketing and Economics*, 1 (3), 251-275.
Yao, Song, and Carl F. Mela (2011), "A dynamic model of sponsored search advertising," *Marketing Science*, 30 (3), 447-468.
Yao, Yuling, Aki Vehtari, Daniel Simpson, and Andrew Gelman (2018), "Yes, but Did It Work?: Evaluating Variational Inference," *Proc. 35th International Conference on Machine Learning*.
Zhang, Jie, and Lakshman Krishnamurthi (2004), "Customizing Promotions in Online Stores," *Marketing Science*, 23 (4), 561-578.
Zhang, Lingling, and Doug J. Chung (2016), "Strategic channel selection with online platforms: An empirical analysis of the Daily Deal Industry, Working paper, Harvard Business School, Boston.
Zhou, Yiyi (2017), "Bayesian estimation of a dynamic model of two-sided markets: Application to U.S. video games," *Management Science*, 63 (11), 3874-3894.