How Do Product Recommendations Help Consumers Search Products? Evidence of Underlying Mechanisms from a Field Experiment

Xiang (Shawn) Wan,[a] Anuj Kumar,[a] Xitong Li [b]
[a] Warrington College of Business, University of Florida; [b] HEC Paris
xiang.wan@ufl.edu; akumar1@ufl.edu; lix@hec.fr

Abstract

Although conventional wisdom suggests that product recommendations can help consumers find products that are lower priced and fit their tastes better, the literature lacks empirical evidence of these mechanisms. We separately estimate the benefits of recommendations to consumers due to these mechanisms through a randomized field experiment on an apparel retailer's website in the US. We collect unique data on the affinity scores computed by the recommendation algorithm to estimate the effect of product recommendations in helping consumers discover larger-value products, which are lower-priced, fit their tastes better, or both. The discovery of higher-value products results in a higher consumers' purchase probability (lower likelihood of failed search efforts) and purchase of lower-priced and/or better-fit products. We provide additional evidence of these mechanisms by showing a larger benefit of product recommendations to consumers in product categories with higher average prices, larger relative price dispersions, and higher heterogeneity in consumers' tastes. Finally, we find that consumers substitute other search tools on the website with product recommendations when available. Our findings have implications for the design and deployment of algorithmic product recommendations on digital platforms.

Keywords: Product recommendations; benefits of recommendations; search cost; consumer search; field experiment.

1. Introduction

Algorithmic product recommendations have emerged as one of the most widely used product promotion tools on e-commerce websites. Although product recommendations are designed to assist consumers in their search process, the literature on product recommendations mainly examines their economic benefits to firms, such as their positive effect on product sales and how they can maximize product sales (De et al. 2010, Lee and Hosanagar 2019). A few studies also examine how recommendations affect the aggregate demand in the market, such as their effect on sales diversity (Hosanagar et al. 2014, Lee and Hosanagar 2019). However, little research examines why and how much product recommendations could benefit consumers, especially based on real-life field data. Answers to these questions could help design recommendation systems that result in higher consumer satisfaction and retention.

The recommendation systems infer consumers' preferences based on the product pages they browse and recommend a few related products on those product pages. Such recommendations allow consumers to directly navigate from one product's page to another, thereby allowing them to view a higher number of product pages in the same total number of webpage views. Moreover, recommending products related to their revealed preferences should help consumers efficiently search among the products aligned to their preferences out of a large number of available options on the website. With such a search process, consumers could find products that they derive larger net value (value/utility minus price), where a product's value to consumers is its quality (consumer preferences on vertical product attributes) and its fit to their tastes (consumer preferences on horizontal product attributes). Thus, product recommendations could benefit consumers by helping them find (i) a lower-priced product, (ii) a product that fits their tastes better, or both.

While it is easy to observe the price of purchased products, measuring how well they match consumers' preferences is challenging, especially based on the field data (Ratchford and Srinivasan 1993). For this reason, extant research has largely examined the effect of consumer search on the purchase price

(Ratchford and Srinivasan 1993, Seiler and Pinna 2017, Ursu et al. 2020), but not on the net value or taste fit to the consumers. We utilize unique data on affinity scores of the recommended products computed by the recommendation algorithm to estimate the effect of recommendations in helping consumers search and purchase larger net value products. We further separately estimate if this larger net value is attributable to the lower product price, the fit of the product to consumers' tastes, or both.

Besides product recommendations, e-commerce websites commonly offer two search tools to consumers: keyword-based and product category-based searches. Because consumers choose their search tools endogenously, measuring how they substitute other search tools with product recommendations is non-trivial. For this reason, prior literature on consumer search and product recommendations offer little empirical evidence on the substitution among the different search tools on e-commerce websites. Understanding how consumers substitute one search tool with another could allow us to design better search tools for e-commerce websites. In this paper, we aim to fill the abovementioned research gaps.

Specifically, we answer the following research questions: (1) Do product recommendations help consumers search for larger value products? (2) How do product recommendations influence consumers' purchase probability, and the price and fit of purchased products? (3) How do product recommendations affect consumers' usage of other search tools on e-commerce websites?

There is demand- and supply-side challenges in answering the above research questions. On the demand side, consumers choose their search efforts (product recommendation usage) endogenously. Many unobserved confounders (such as consumers' price sensitivity) may simultaneously drive consumers' product recommendation usage and search process outcomes, e.g., the price or fit of searched products. We address this challenge by exogenously manipulating the availability of product recommendations to consumers in a field experiment on a US apparel and home goods retailer's website. In this experiment, we randomly assigned visitors to view one of the two versions of the website. The website's treated version recommended four products that are most co-viewed and co-purchased with, and in the product category of the focal product (hereafter FP) on the FP's page. These four products are called recommended products (hereafter RP).[1] The control version of the website hid the RPs on the FP's page. On the supply side, retailers may strategically recommend products to optimize profits or reduce inventory.[2] In such cases, the retailer-level unobserved factors may confound the estimate of the benefits of recommendations to visitors. Fortunately, the retailer in our research context did not strategically recommend products on its website.

The recommendation algorithm computes the weighted sum of co-views and co-purchases of the FP and its RP by past visitors – called *affinity scores*. We collected the data on affinity scores of each FP-RP pair. Among the four RPs shown on the FP's page, we find that visitors are more likely to view and buy the RP of a higher affinity score. Moreover, among the various products viewed during the search process, we find that visitors are more likely to buy products of higher affinity scores. Since visitors view and buy products that provide larger net value to them, the affinity score of a product could be used to measure its net value to the visitors. We further decompose the affinity score of a searched product into two parts: one correlated to and the other uncorrelated to the product's price. We argue that the part that is uncorrelated to product price could measure the extent to which the product's attributes fit with visitors' tastes (see details about the measures of a product's net value and fit to visitors' taste in Section 5).

We measure a visitor's recommendation usage with (1) an indicator variable for the availability of recommendations and (2) the total number of RP page views during a search process. Although the availability of recommendations to consumers is random, visitors choose their number of RP page views endogenously. We use the exogenous availability of recommendations as an instrumental variable to account for the endogeneity in the number of RP page views (search effort under recommendations). Then,

[1] Recommendation algorithms based on such co-view (co-purchase) relationships between products belongs to item-based collaborative filtering recommendation algorithms. Details of the recommendation engine are presened in Section 3.2.

[2] For example, few recommendation systems are designed to maximize retailers' profits besides providing relevant recommendations to the users (Abdollahpouri et al. 2020). Some recommendation systems balance these two objectives by recommending the high margin or high-value products from the pool of relevant products.

we estimate the effect of recommendation usage by comparing the search and purchase behavior of treated visitors with control visitors.

We find that treated visitors browse higher affinity score, lower-priced, and better fit products, compared to control visitors. The higher number of RP page views due to recommendations on average increases affinity scores and taste fit of browsed products. Moreover, additional RP page views due to recommendations reduce the minimum, average, and median values, but not the maximum value, of price distribution of searched products.

If recommendations help visitors find larger net value products, we should expect a higher purchase probability (lower likelihood of failed search efforts) for treated visitors than control visitors. In line with this conjecture, we find that product recommendations increase purchase probability by two percent. We further find that the higher likelihood of product purchase results from the fact that treated visitors view higher net value products (as measured by affinity scores), perhaps because they find lower-priced products, better fit products, or both. For the converted visitors, we further estimate the effect of recommendations on purchased products. We find that the treated visitors purchase higher affinity-score, lower-priced, and better-fit products than control visitors. Economically, our estimates suggest that the net value of purchased products (as measured by the affinity scores) on average increases by 22% due to product recommendations.

As we pointed out above, the higher net value could result from lower product price, better fit to visitors' tastes, or both. Our additional analyses provide further support to these underlying mechanisms. Specifically, we find that the recommendations result in the purchase of lower-priced products in product categories with higher average prices and larger relative price dispersions. These results support the mechanism of finding lower-priced products. We also find evidence for the mechanism of finding better-fit products by showing that the recommendations result in a higher fit of purchased products in product categories with higher heterogeneity in consumers' tastes (such as products in the women's apparel category).

Finally, we find evidence that visitors substitute other search tools with product recommendations, plausibly because product recommendations help them search for larger-value products better than other alternatives. Specifically, our estimates indicate that an additional RP page view under recommendations results in a decrease of 0.14 search page views and 0.63 product category page views, respectively.

This paper contributes to the literature on product recommendations and online consumer search. First, our study is perhaps the first to provide empirical evidence that product recommendations help consumers find and buy larger net value products in a real-life business setting, whereas very few prior studies document this finding perhaps due to the unavailability of data that measures product value or fit. Second, while conventional wisdom suggests that product recommendations could benefit consumers due to two underlying mechanisms – find lower-priced and better-fit products, little research could tease out the benefits of these mechanisms separately. We are among the first to provide evidence for the impacts of these mechanisms on consumers' search process, purchase probability, and finally, purchased products. Lastly, we quantify how consumers substitute for the existing search tools with product recommendations when it is exogenously made available to them. These estimates highlight the efficacy of product recommendations as a search tool compared to other search tools on websites. Our findings have implications for the design of product recommendations and their deployment vis-à-vis other search tools on digital platforms.

2. Related Literature

Our research draws from two literature streams: the literature on product recommendations and the literature on consumer search costs and returns to search.

2.1. Related Literature on Product Recommendations

Our research is closely related to the literature on production recommendations. Many prior studies in this literature stream estimate the positive impact of product recommendations on product sales. De, et al. (2010)

conducted one of the earliest studies to show that using a recommender system could increase both promoted and non-promoted products' sales. Lee and Hosanagar (2019) find the positive impact of two types of collaborating filtering – purchase-based and view-based collaborative filtering – on sales volume, while the effect of purchase-based collaborative filtering is more pronounced. More recently, Lee and Hosanagar (2021) examined the heterogeneity in the positive effects of recommendations on purchase conversion rate and found that the effects are larger for hedonic goods than for utilitarian goods, but the effects are not significantly different between experience goods and search goods.

Many product recommendations are developed based on co-view or co-purchase relationships between products (Goldenberg et al. 2012, Thorat et al. 2015). There is emerging literature that focuses on examining the effects of product recommendations from this perspective. Oestreicher-Singer and Sundararajan (2012) find that the visibility of a co-purchase relationship can lead to up to an average threefold amplification of the influence that complementary products have on each other's demands. Kumar and Tan (2015) document that recommending products on the focal products' webpages can not only increase the sales of the focal products (direct effect) but also increase the sales of complementary products (spillover effect). While the jointly displayed products are largely complementary in the study of Kumar and Tan (2015) (e.g., apparel and accessories), Kumar and Hosanagar (2019) examine the situation where the focal product and the recommended products are substitutes. They find that a recommendation link increases the daily focal product page views by seven percent, reduces focal product sales conditional on the page views by 8.5 percent, and increases recommended products' sales by 24.5 percent. Overall, recommendations have led to an increase in the total sales of the focal and recommended products. Lin et al. (2017) examine how the product recommendation network's diversity could influence the effect of recommendations on product sales. They find that a one percent increase in the category diversity of the incoming (outgoing) co-purchase network of a product is associated with an increase (decrease) in product sales.

Prior studies in this literature largely focus on examining the economic benefits that product recommendations bring to firms and online retailers. These studies explore the effects of product recommendations on product sales and purchase conversion rate and how these effects differ with product characteristics, such as hedonic versus utilitarian goods (Kumar and Hosanagar 2019, Lee and Hosanagar 2021) or the category diversity of the product recommendation network (e.g., Lin et al. 2017). To the best of our knowledge, no prior study has examined either the direct benefits of product recommendations to consumers or the moderating effect of product characteristics on product recommendations' benefits to consumers. For example, how do product recommendations affect consumers' purchase probability? Do the benefits of product recommendations to consumers result from finding lower-priced or better-fit products? Our research answers these questions to fill the gaps in the literature on product recommendations.

2.2. Related Literature on Search Cost and Return to Consumer Search

Our research is also related to the literature on consumers' search cost. Consumers' search for product information is nontrivial and often costly. In this paper, we assume consumers perform a sequential search (Reinganum 1982, Weitzman 1979), i.e., consumers continue searching until the marginal cost exceeds the expected marginal benefit of an extra search. Most studies on consumer search in an online environment assume a sequential search by consumers (Chapelle and Zhang 2009, Chen and Yao 2017, Ghose et al. 2019). Accordingly, in the equilibrium, consumers stop searching when the expected marginal benefit of search equals the marginal search cost, thereby allowing researchers to measure the return to consumer search.

While it is challenging to determine the quality and direct utility of consumer search, Ratchford and Srinivasan (1993) have conducted one of the earliest studies to estimate the monetary return to consumer search by measuring lower price paid as the outcome of consumer search. Specifically, they assumed that consumers stop searching when the marginal return to search becomes equal to the marginal search cost. They found that the median consumer could save \$17.76 by spending an additional hour searching for a lower car price. In a later study, Honka (2014) has quantified the range of search costs in the US auto insurance industry from $35 to $170. She further found that search cost was the most important driver of consumer retention than switching cost and customer satisfaction. Accordingly, the elimination of search costs could be the main lever to increase consumer welfare in the auto insurance industry. Interestingly, Ngwe et al. (2019) have found that under certain conditions, deliberately increasing search cost by varying website navigation elements, especially those associated with accessing discounted items, could increase online retailers' average selling prices and overall expected purchase probability.

It is nontrivial to estimate the monetary return to consumer search because consumers' search efforts are endogenous. For example, more price-sensitive consumers exert more effort to search for lower-priced products. Ursu, et al. (2020) account for this endogeneity by explicitly modelling a consumer's decision on how much time she chooses to spend searching and estimate an average search cost of $0.07 per minute. Ursu, et al. (2020) further show that consumers are more likely to purchase from restaurants that they spend more time searching online. More closely related to our research, Seiler and Pinna (2017) estimate the causal return to consumers' search efforts using "path-tracking" data obtained from shopping carts equipped with radio-frequency identification (RFID) tags in a physical store environment. Like Ratchford and Srinivasan (1993), Seiler and Pinna (2017) estimate the monetary return to consumer search in terms of lower prices paid for purchased products. They find an additional minute of search results in a price saving of $2.10. Seiler and Pinna (2017) further show that there is little heterogeneity in the return to consumer search across product categories (e.g., category size, average price, price dispersion), but considerable heterogeneity across consumer types (e.g., shopping frequency) and product locations in the physical store.

In sum, prior research in this literature stream has estimated the monetary return to consumer search as the lower price paid (Ratchford and Srinivasan 1993, Seiler and Pinna 2017). However, to our knowledge, no prior study has examined how consumers choose among the different search tools available to them on the retailing website and how consumers find the lower-priced products from their search. Our research contributes to this literature by (1) estimating how consumers substitute the usage of other search tools with product recommendations and (2) whether they find larger value, lower-priced or better-fit products during the purchase process. Besides, while Seiler and Pinna (2017) do not find heterogeneity in returns to consumer search across product categories, we find substantial heterogeneity depending on relative price dispersions across product categories. Moreover, to our knowledge, no prior research compares consumers' search behaviors when product recommendations are available versus when they are unavailable, with the only exception of Dellaert and Häubl (2012). Dellaert and Häubl (2012) conduct laboratory experiments to investigate the effect of recommendations on consumer decision processes in product search. Unlike their study, our research estimates the impact of recommendations on consumers' search from a large-scale randomized field experiment. Our analysis compares consumers' product searches in the two scenarios by leveraging the exogenous shock to their search cost due to the random availability of product recommendations. We show how consumers' usage of product recommendations affects the outcomes of consumer product search, such as finding larger net value and better fit products.

3. Field Setup

3.1. Website Description

We conducted a field experiment on a mid-size online retailer's website in the US. The retailer's annual sales are over $400 million, and the online channel accounts for 10 percent of the total sales. The retailer offers over 35,000 products for sale on its website. These products are organized into eight main categories, including women's clothing, men's clothing, etc. Each of these main categories is further subdivided into subcategories. For example, the main category of women's clothing consists of several subcategories, such as women's dress, top, shorts, and skirts.

The contents on the website are organized into four levels: home page, product category pages, product subcategory pages, and product description page (hereafter referred to as "product page").

Webpages on a higher level have hyperlinks to navigate visitors to the subsequent lower-level webpages. For example, a visitor can click on the "women" hyperlink on the home page to go to the main page of the women's product category. Then, the visitor can click on the "skirts" hyperlink on the women's category page to go to the main page of the women's skirts product subcategory.

Figure 1 shows an example of a product page for a women's top (called *focal product* or *FP*) on the website. The product page has a large FP image with its information, such as regular and discounted retail prices, available colors and sizes, product descriptions, etc. For visitors presented with product recommendations, the FP's product page also displays four recommended products or RPs, as the smaller images of four women's tops under the heading "MORE OPTIONS" on the right in Figure 1. A visitor can reach the product page of an RP by clicking its image on the FP's page. After the visitor clicks to view the RP's product page, the RP becomes a new FP and its product page will display another set of recommended products.

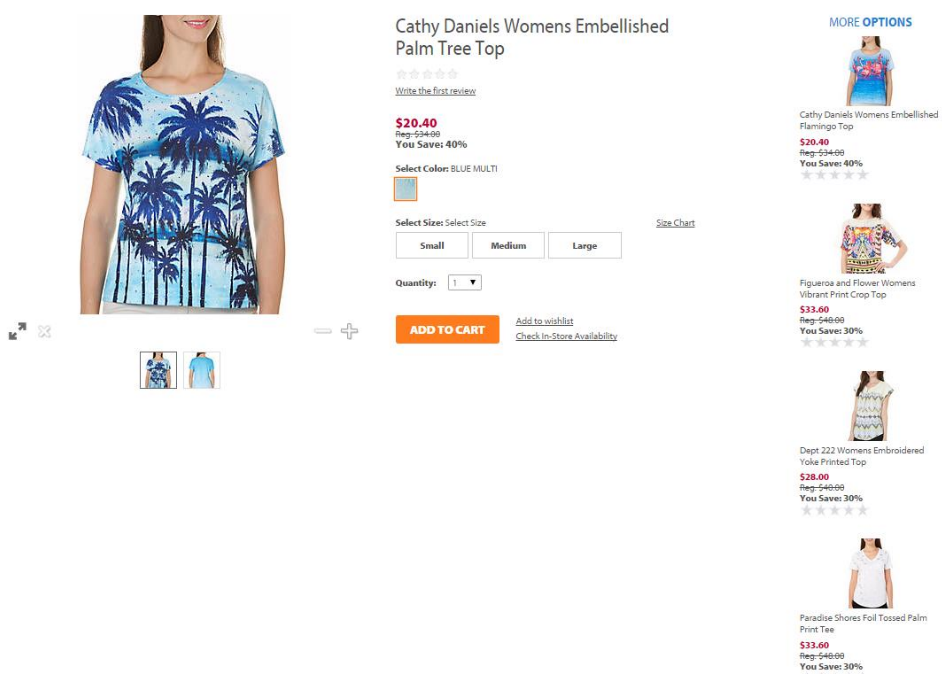

Figure 1. Product Page of an FP with four RPs

3.2. Collaborative Filtering Product Recommendations

The retailer uses IBM's Coremetrics digital recommendation engine to recommend the RPs on an FP's page. The recommendation algorithm selects the RPs based on two rules. First, the recommendation algorithm computes the affinity score between every two products on the retailer's website based on their co-views and co-purchases over the last 30 days. The affinity score consists of the following four component scores: (1) view-to-view score that counts the number of times a visitor view both the FP and RP in the same session; (2) view-to-buy score that counts the number of times a visitor view the FP but buy the RP in the same session; (3) buy-to-buy score that counts the number of times a visitor buy both the FP and RP but not necessarily in the same session; (4) abandon-to-buy score that counts the number of times a visitor abandons the carted FP to buy the RP in the same session. These four component scores are aggregated by the following weights: 70 (view-to-view score), 20 (view-to-buy score), 5 (buy-to-buy score), and 5 (abandon-to-buy score). The recommendation engine computes the affinity scores for each pair of products on the retailer's website at 4 AM daily.

Thus, the recommendation engine infers the relationship between two products from the co-views and co-purchases of visitors on the retailer's website, an item-based collaborative filtering recommendation algorithm (Linden et al. 2003, Sarwar et al. 2001). Item-based collaborative filtering algorithms are among the most widely used recommendation algorithms in the current business practice (Lee and Hosanagar 2021). For example, "Amazon Personalize" provides such similar item-based recommendation service – "People who purchased this item also purchased…"

Besides the affinity score, the recommendation algorithm only selects RPs from FP's product subcategory. Thus, the RPs are largely substitutes to the FP in our research context. The algorithm stores the top 15 products with the highest affinity scores and the same product subcategory as the FP. Finally, the top four products with the FP's highest affinity scores are recommended on the FP's page.

3.3. Product Category Search and Keyword Search

As described in Section 3.1, the retailer's website organizes products in different product categories and then under various subcategories within each product category. This hierarchical organization of products helps visitors conduct *product category-based searches* on the website. For example, a visitor interested in searching for women's tops can select the women's category on the home page and then choose the women's top subcategory on the women's category page to browse different women's tops.

Besides the hierarchical organization of products, the retailer's website also offers a search text box on each webpage in which visitors can enter their search keywords. The search algorithm identifies the products that match the search keywords and displays them on the search result page. For example, if a visitor enters the search keywords "women's floral tops" in the search text box, the search result page would show a number of women's floral tops available on the website.

3.4. Experimental Setup

In practice, visitors' usage of product recommendations is correlated to their unobserved characteristics. For example, a price-sensitive visitor would choose to view recommendations more intensively, and she also has a stronger intention to find lower-priced products. Therefore, we need an exogenous variation in the availability of product recommendations across visitors to estimate its unbiased effects on visitors' search behavior. We accomplish this exogenous shock by conducting a randomized field experiment on the retailer's website. Specifically, the retailer created two versions of product pages on its website. The FP's page displays four RPs identified by the recommendation system in the treated version of product pages, as illustrated in Figure 1. In contrast, the four RPs are suppressed (not displayed) on the FP's page in the control version of product pages.

We ran the field experiment for nine weeks, from 8th April 2015 to 9th June 2015. In the experiment period, we randomly assigned half of the visitors on the website to the treated version (treated visitors) and the remaining half to the control version (control visitors). The recommendation engine can identify the visitors based on different information, such as cookies and IP addresses, and consistently assigned visitors to the same experimental version in their all visits.

3.5. Effect of Recommendations

Product recommendations allow consumers to conduct efficient product search on the website. Once a visitor reveals her preferences by selecting an FP on the website, the recommendation system displays the related RPs on the FP's page. Thus, recommendations enable visitors to directly navigate from the FP's page to a related RP that is likely aligned to the visitors' preferences. In this way, recommendations could help visitors view a higher number of products that potentially fit their tastes (preferences on a product's horizontal attributes) from a large number of available alternatives on the retailer's website.

If product recommendations are effective, the recommendations should help visitors find products that offer them larger net value. A product's net value to consumers is the difference between its value and price, that is, (Value – Price). Prior literature indicates that the value derived from a product can be decomposed into two orthogonal dimensions: the value from the vertical quality (consumer preferences on vertical product attributes) and horizontal fit (consumer preferences on horizontal product attributes) (Kwark et al. 2017, Sutton 1986). Vertical product attributes, such as a product's raw material and craftsmanship, determine its vertical quality that correlates to its price. On average, consumers have a higher preference for products with higher vertical quality. In contrast, how well the horizontal product attributes match a consumer's tastes determines its horizontal fit. For example, different consumers may prefer different colors of the same product, even though the product's material, design, and craftsmanship are the same. By definition, the horizontal fit measure is uncorrelated to the product's price.[3]

[3] Sometimes, retailers may strategically charge higher prices for products in popular colors and styles that most consumers prefer to increase sales revenue. In that case, such popular colors and styles commanding higher prices may become part of the product's vertical quality.

4. Data Description and Summary Statistics

4.1. Purchase Funnels and Balance Checks

A visitor searches for desired products in a product category by exploring various relevant web pages, such as category/subcategory product page, product detail pages, and search result pages on a website. If the visitor finds her wanted product, she would purchase it; otherwise, she ends her search. In line with the prior literature, we call a visitor's sequence of web page views during her product search in a product subcategory as a "purchase funnel" (Bleier and Eisenbeiss 2015).[4] A visitor's purchase funnel begins from a session when she starts searching for products in a product subcategory and ends when she purchases a product in that product subcategory or the end of our data period, whichever is earlier. After purchasing products in a subcategory, if a visitor again explores products in the same subcategory, it is considered as a new purchase funnel. Thus, visitors can have multiple purchase funnels in a product subcategory in our data. A visitor may be in multiple purchase funnels during a web session if she views web pages related to multiple product subcategories.

We conduct our analyses on purchase funnels for product subcategories under five main product categories. The product subcategories under these main categories are well defined and only contain substitute products.[5] For example, women's skirts contain only women's skirts of all types. During the experiment period, visitors searched products in 707,777 purchase funnels. We dropped 135,530 funnels in which visitors do not view a product page because we do not have information on those visitors' product search. Ultimately, we have data of 572,247 purchase funnels across 68 product subcategories by 429,212

[4] Visitors may go through four stages of the purchase process while purchasing a product in a product category. They search for product information, form a consideration set of products, evaluate products in the consideration set, and finally purchase the chosen product.

[5] Since consumers search, evaluate, and choose a product among its substitute products in a purchase funnel, we dropped those product subcategories that also contain complementary or unrelated products in the same product subcategory. For example, under the product main category "Home", the product subcategory of "Home Décor" includes furniture, rugs and mats, sheets, and decorative pillows. It is not appropriate to consider the furniture, rugs, sheets, and pillows browsed by a visitor in her multiple sessions to be one purchase funnel. Accordingly, we dropped three main product categories.

unique visitors. Out of the total purchase funnels, 286,653 (50.09 percent) funnels were for treated visitors (called *treated purchase funnels*), and the remaining were for control visitors (called *control purchase funnels*).

We check whether the random assignment of recommendations is valid in the final sample of visitors and purchase funnels. We conduct balance checks at both the visitor and purchase funnel levels. The results support that visitor characteristics between the treated and control conditions are statistically balanced. The results are presented in Appendix A.

4.2. Page Views and Purchase Probability

Table 1 presents the summary statistics of purchase probability and different types of page views for treated and control purchase funnels. We find that visitors purchase products in 6.2 percent of the treated purchase funnels, and 6.0 percent of the control purchase funnels. The difference in purchase probability is statistically significant, providing preliminary evidence that displaying product recommendations on FP pages can increase visitors' likelihood to buy.

Table 1. Summary Statistics of Page Views and Purchas Probability

	Control Purchase Funnels		Treated Purchase Funnels		Diff in Means
	Obs.	Mean (S.D.)	Obs.	Mean (S.D.)	(*t*-stats)
Purchase probability	285,594	0.060 (0.24)	286,653	0.062 (0.24)	0.002* (2.07)
# of webpage views	285,594	4.69 (7.35)	286,653	4.73 (7.24)	0.04* (2.12)
# of FP page views	285,594	1.53 (1.60)	286,653	1.65 (1.74)	0.13*** (28.54)
# of RP page views	285,594	0.18 (0.66)	286,653	0.32 (0.89)	0.14*** (67.84)
FP page views / Total webpage views	285,594	0.66 (0.36)	286,653	0.67 (0.35)	0.01*** (8.99)
# of category page views	285,594	1.79 (5.12)	286,653	1.71 (4.95)	-0.08*** (-6.13)
# of search result page views	285,594	0.71 (2.40)	286,653	0.69 (2.37)	-0.02** (-3.01)

Notes: Diff. in means = Treated - Control. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

In addition to the purchase probability, we compute several variables related to visitors' search behaviors, such as the number of webpage views, FP page views, RP page views, category page views, and search result page views. In our operationalization, each product's page view in a purchase funnel is counted as an FP page view. If a product viewed in the purchase funnel appears as an RP on one of the earlier viewed products' pages, this page view is also counted as an RP page view. We observe several interesting findings.

First, we find a statistically higher average number of webpage views in treated purchase funnels than control purchase funnels (4.73 versus 4.69). Second, we find a statistically higher average number of FP page views (1.65 versus 1.53) and RP page views (0.32 versus 0.18) in treated purchase funnels than control purchase funnels. These results reveal that displaying related RPs on the FP's pages plausibly drives more FP page views and more RP page views in treated purchase funnels because treated visitors can directly navigate to an RP's page from the FP's page. Third and more interestingly, we find a statistically higher ratio of FP page views over the total number of webpage views in treated purchase funnels than control purchase funnels (0.67 versus 0.66). The higher ratio indicates that visitors with recommendations can view more product pages out of the same total number of webpage views, suggesting a more efficient search under recommendations. Finally, we find a statistically lower average number of category page views (1.71 versus 1.79) and search result page views (0.69 versus 0.71) in treated purchase funnels than control purchase funnels. These results indicate that treated visitors substitute their usage of alternative navigation tools (product category-based and keyword-based search) due to the availability of recommendations on FP's pages.

4.3. Model-Free Evidence for Lower Price

Table 2 reports the summary statistics of the price distribution of products browsed in treated and control purchase funnels. We find a significantly lower mean value of average prices for products browsed in treated purchase funnels than control purchase funnels (\$39.16 versus \$39.35). We further find significantly lower mean values of the minimum prices, the prices at the 25th percentile, and the median prices of browsed products in treated purchase funnels. However, we do not find a significant difference in the mean values of maximum prices or the prices at the 75th percentile of browsed products between the two types of purchase funnels. These results provide model-free evidence that treated visitors could find lower-priced products during the product search than control visitors, while they view similar high prices with and without recommendations.

Table 2. Summary Statistics of Price Distribution of Products Searched

	Control Purchase Funnels		Treated Purchase Funnels		Diff in Means
	Obs.	Mean (S.D.)	Obs.	Mean (S.D.)	(t-stat)
Average	285,594	39.35 (24.52)	286,653	39.16 (24.64)	-0.19** (-2.89)
Minimum	285,594	37.93 (24.55)	286,653	37.56 (24.65)	-0.37*** (-5.71)
25th percentile	285,594	38.21 (24.49)	286,653	37.89 (24.59)	-0.32*** (-4.90)
Median	285,594	39.26 (24.55)	286,653	39.06 (24.67)	-0.20** (-3.04)
75th percentile	285,594	40.46 (25.06)	286,653	40.40 (25.27)	-0.06 (-0.87)
Maximum	285,594	41.01 (25.45)	286,653	41.03 (25.73)	0.02 (0.31)

Notes: Diff. in means = Treated - Control. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5. Measures of Net Value and Horizontal Fit

5.1. Net Value Measure

In our research, we collected unique data on daily affinity scores between each pair of products computed by the recommendation algorithm on the retailer's website. Below, we show that a product's affinity scores could be a measure of its net value to an average visitor.

The affinity score between two products is the weighted sum of their number of co-views and co-purchases by past visitors on the website. For a visitor who views Product A's page, if the affinity score between Products B and A is higher than between Products C and A, then the visitor is more likely to choose to view Product B than Product C by the definition of affinity scores. But visitors would view products that offer them high net perceived value (utility), taking into account all the product attributes, such as the fit to their preferences and product price. Therefore, if visitors who view Product A prefer to view Product B which has a higher affinity score (with Product A) than Product C, Product B should offer on average a higher net value to the visitors. In other words, the affinity score of Product B (with Product A) is positively correlated with Product B's net value to those visitors who view Product A. To this end, we believe the affinity score of a product (with other products) can be used as a measure of its net value to visitors. Below, we conduct two sets of empirical tests on the products viewed in purchase funnels to show that a product's affinity score with other products in the funnel measures its net value for an average visitor.

Our first empirical test shows that an RP with a higher affinity score with the FP is more likely to be viewed (purchased) in the purchase funnel. Visitors view a few products (FPs) in a purchase funnel. While four RPs are explicitly displayed as recommendations on the FP's page in the purchase funnel for treated visitors, these RPs are hidden on the FP's page for control visitors. We estimate the likelihood of each RP's view (purchase) on an FP's page in the purchase funnel with the following Specification

$$Y_{frt} = \alpha_f + \alpha_t + \beta \times AffScore_{frt} + \delta \times Control_{frt} + \varepsilon_{frt}, \tag{1}$$

where (f), (r), and (t), respectively, denote the FP, RP, and days. The dependent variables are the indicator variable for RP view and purchase. We control for the RP's position on FP's page (Position=1,2,3, or 4), whether the RP is explicitly displayed (in treated purchase funnels) or not (in control purchase funnels), FP fixed effects, and day fixed effects. We use the log transformation of affinity scores to account for its highly skewed distribution. Table 3 reports the estimation results.

Columns (1)-(3) in Table 3 report estimates for the likelihood of an RP view. We first note that the signs of the estimated coefficients of the RP's position and whether the RP is explicitly displayed are consistent with our expectations and thus provide face validity of the empirical results. Specifically, a lower RP's placement on the FP's page (higher value of the RP's position) is associated with a lower likelihood of the RP view, and the explicit display of the RP on the FP's page is associated with a higher probability of the RP view. Notably, we find positive and significant coefficients for affinity scores in all specifications, indicating that the probability of the RP view increases with its affinity score with the FP. We find qualitatively similar results for the likelihood of the RP purchase in Columns (4)-(6) in Table 3. Taken together, the results in Table 3 provide consistent evidence that an RP with a higher affinity score with the FP is more likely to be viewed and purchased.

Table 3. The Effect of RP's Affinity Score on RP View and Purchase

	DV: Whether RP View			DV: Whether RP Purchase		
	(1) Treated	(2) Control	(3) Pooled	(4) Treated	(5) Control	(6) Pooled
AffScore	0.0022^{***} (0.0001)	0.0011^{***} (0.0001)	0.0017^{***} (0.0001)	0.0002^{***} (0.0000)	0.0001^{*} (0.0000)	0.0001^{***} (0.0000)
RP's Position	-0.0154^{***} (0.0002)	-0.0101^{***} (0.0001)	-0.0129^{***} (0.0001)	-0.0013^{***} (0.0000)	-0.0011^{***} (0.0000)	-0.0012^{***} (0.0000)
Rec. Indicator			0.0266^{***} (0.0002)			0.0010^{***} (0.0001)
Constant	0.0821^{***} (0.0018)	0.0538^{***} (0.0015)	0.0548^{***} (0.0012)	0.0060^{***} (0.0005)	0.0056^{***} (0.0004)	0.0054^{***} (0.0003)
FP fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Day fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.0404	0.0414	0.0379	0.0132	0.0151	0.0091
Obs.	1,868,121	1,716,835	3,584,956	1,868,121	1,716,835	3,584,956

Notes: + p < .10, * p < .05, ** p < .01, *** p < .001

A visitor starts the product search on the website by selecting a product that is aligned to her preference and offers her high perceived net value. Therefore, subsequent products (FPs) in the purchase funnel with high affinity scores with the first FP would also be aligned to visitors' preferences and provide a similar high net value. Accordingly, our second empirical tests show that products with higher affinity scores with the first FP in the purchase funnel are more likely to be purchased. We compute the affinity scores for all subsequent products with the first FP in the purchase funnel and estimate the effect of these affinity scores on their purchase probability with the following Specification.[6]

$$PurchProb_{fit} = \alpha_i + \alpha_t + \beta \times AffScore_{it} + \delta \times Control_{fit} + \varepsilon_{fit} , \quad (2)$$

where (f), (i), and (t), respectively, denote products (FPs), purchase funnels, and days. The dependent variable (*PurchProb*) is the probability of an FP purchase in the funnel. We further control for the product's order (or position) in the purchase funnel, whether it is the treated purchase funnel, product funnel fixed effects, and day fixed effects. We use the log transformation of the affinity scores to account for its highly skewed distribution. Table 4 reports the estimation results.

Table 4. The Effect of FP's Funnel-level Affinity Score on FP Purchase

	DV: Whether FP Purchase		
	(1) Treated	(2) Control	(3) Pooled
AffScore with first FP	0.0012*** (0.0001)	0.0012*** (0.0001)	0.0012*** (0.0001)
FP's Position	0.0039*** (0.0002)	0.0030*** (0.0002)	0.0035*** (0.0001)
Rec. Indicator			Dropped
Constant	0.0314*** (0.0007)	0.0427*** (0.0007)	0.0366*** (0.0005)
Funnel fixed effect	Yes	Yes	Yes
Day fixed effect	Yes	Yes	Yes
R^2	0.5025	0.5251	0.5135
Obs.	435,224	472,950	908,174

Notes: + p < .10, * p < .05, ** p < .01, *** p < .001

We first note that the signs of the estimated coefficients of the FP's position (order) are significantly positive, suggesting that products viewed later in the purchase funnel have a higher likelihood of purchase than those viewed earlier. This result is consistent with the prior findings that consumers are more likely to

[6] For the first FP in a purchase funnel, we define it as the average of its affinity score with all other products in the same product subcategory, because a product with a higher average affinity with all other products in the same product subcategory is more likely to be viewed first. We also define it as zero or leave it as missing, and find that our results are qualitatively similar.

purchase products viewed in the later stages of the purchase funnel (Bleier and Eisenbeiss 2015, Lambrecht and Tucker 2013), thus providing face validity of the empirical results. Notably, the estimated coefficients of affinity scores are all positive and statistically significant, supporting that an FP with a higher affinity score with the first FP in a purchase funnel is more likely to be purchased after controlling for the funnel fixed effects and day fixed effects.

Put together, the results in Tables 3 and 4 provide consistent evidence that visitors are more likely to view and purchase higher affinity score products. Since visitors would choose to view and purchase products that offer them high net value (after considering all the product attributes, including price), these findings indicate that the affinity score measures the product's net value to visitors.

5.2. Horizontal Fit Measure

As described in Section 3.5, the value derived from a product can be decomposed into two orthogonal dimensions: the value from the vertical quality (denoted by VQ) and horizontal fit (denoted by HFit) (Kwark, et al. 2017, Sutton 1986). Thus, the net value of a product = (VQ+HFit) – Price = (VQ–Price) + HFit. While the component (VQ–Price) is correlated to its price, the component HFit is not.

We utilize the residual approach to measure the product's horizontal (taste) fit in our context. The labor economics and financial accounting literature have widely used the residual approach to decompose certain theoretical constructs into a component related to a factor and another component unrelated to the factor (Ali and Zhang 2015, Basu et al. 2006, Brynjolfsson et al. 2021, Dou et al. 2013, Jones 1991, Jorgenson and Stiroh 1999, Kim et al. 2014, Solow 1956, Solow 1957). As discussed in Section 5.1, the affinity score of the viewed product f in purchase funnel i measures the net value of product f, where Net Value = (VQ–Price) + HFit, and the first component (VQ–Price) could be explained by $Price_f$. We estimate the following regression:

$$AffinityScore_{if} = \beta_i + \beta_1 \times Price_f + \varepsilon_{if}, \tag{3}$$

where β_i indicates purchase funnel fixed effects. The residual of Specification (3) captures the variation in Affinity Scores that is uncorrelated to price. As a product's affinity score measures its net value, the residual from Specification (3) measures the component of net value uncorrelated to the product's price. Thus, the residual measures the horizontal fit of the viewed product f in purchase funnel i, denoted by $HFit_{if}$. We also compute the horizontal fit-over-price ratio of product *f*, as $Ratio_{if} = HFit_{if}/Price_i$ to capture the tradeoff between horizontal fit and product price.

Since our estimation of the effects of product recommendations is on the purchase funnel level, we aggregate $AffinityScore_{if}$, $HFit_{if}$ and $Ratio_{if}$ across all viewed products in purchase funnel *i* by taking their averages across all viewed products, denoted by $AveAffScore_i$, $AveHFit_i$ and $AveRatio_i$, respectively. Table 5 reports their summary statistics for treated and control purchase funnels. These results show that the products browsed in treated purchase funnels have a higher average affinity score, a higher value of horizontal fit, and a larger fit-over-price ratio than control purchase funnels. Thus, the results in Table 5 provide preliminary evidence that product recommendations, on average, help visitors find and view products with higher net value, better horizontal fit with their tastes, and a larger fit-over-price ratio.

Table 5. Summary Statistics of AveAffScore, AveHFit, and AveRatio

	Control Purchase Funnels		Treated Purchase Funnels		Diff in Means
	Obs.	Mean(S.D.)	Obs.	Mean(S.D.)	(t-stat)
AveAffScore	285,594	1.234 (2.259)	286,653	1.632 (2.683)	0.398*** (60.69)
AveHFit	285,586	0.256 (1.849)	286,646	0.464 (1.973)	0.208*** (41.10)
AveRatio	285,580	0.005 (0.085)	286,641	0.013 (0.091)	0.007*** (32.03)

Notes: Diff. in means = Treated - Control. + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001.

6. Results

6.1. Effect of Recommendations on Consumers' Search Behavior

We examine the effect of recommendations on visitors' search behavior on the website with the following specifications:

$$Y_i = \beta_c + \beta_1 \times Rec_i + \varepsilon_i, \quad (4.1)$$

$$Y_i = \beta_c + \beta_1 \times NRPView_i + \varepsilon_i, \quad (4.2)$$

where (i) denotes purchase funnels, and (c) denotes product categories. *Y(i)* denotes visitors' search behavior in terms of (1) average price, average affinity score, average horizontal fit, and the average fit-over-price ratio of the products browsed in a purchase funnel, and (2) different price-related variables, including average price; minimum price; prices at 25th, 50th, and 75th percentiles; and maximum price of price distribution in purchase funnel (i). Rec_i denotes the indicator variable for the treatment of recommendations. Coefficient β_1 estimates the treatment - *intent to treat* – effect of recommendations. As discussed in Section 4 and shown in Appendix A, statistically similar visitor characteristics between the treated and control purchase funnels assure us that variable Rec_i is exogenous in Specification (4.1). Parameter β_c denotes the product category fixed effects that account for differences in the effect of unobserved product category-level factors on the dependent variable.

As discussed in Section 3.5, recommendations help visitors find products of higher net value by enabling them to view a higher number of RP page views. Thus, we estimate the effect of the number of RP page views in purchase funnel *i* (called $NRPView_i$) on visitors' search behavior in Specification (4.2). Although recommendations are randomly assigned, visitors endogenously choose their recommendation usage in purchase funnels. Therefore, $NRPView_i$ is endogenous in Specification (4.2) because customer-level unobserved characteristics may be correlated to the number of product pages views and the product price. For example, more price-sensitive visitors may be more likely to view RPs on the FP's page, and such customers can also search for lower-priced products. We account for endogeneity by instrumenting the number of RP page views with the recommendation indicator (*Rec*) and estimate Specification (4.2) using a two-stage least square regression (2SLS). Variable *Rec* satisfies the exogeneity condition for instrument variable (IV), as recommendations are randomly assigned across visitors. Moreover, *Rec* should significantly affect the number of RP page views to satisfy the relevance condition. As expected, we find an *F-value* of 142.9 for the exclusion of *Rec* in the first stage regression of *NRPView* on *Rec*. This value is significantly higher than the threshold value of 10 for the weak instrument, which indicates that *Rec* satisfies the relevance condition for the IV (Bound et al. 1995, Dinkelman 2011). The second stage of regression estimates the effect of the number of RP page views due to recommendations on visitors' price search behavior. Table 6 reports the estimated coefficients from Specifications (4.1) and (4.2).

We first estimate recommendations' impact on visitors' search behavior on the website using Specification (4.1). We capture visitors' search behavior in terms of average price, average affinity score, average horizontal fit, and the average fit-over-price ratio of the products browsed in a purchase funnel. Panel A of Table 6 reports the effects of our treatment variable (*Rec*. Indicator) on the different outcomes. The results in Column (1) show that recommendations help visitors find higher affinity score products that offer them larger net value. The results in Columns (2) and (3) suggest that recommendations help visitors find lower-priced products and also products that fit their tastes better (better fit products). Hence, the higher average net values of products browsed under recommendations are due to their lower average prices and a better fit of their attributes with visitors' tastes. Results in Column (4) indicate that visitors browse higher fit-over-price ratio products under recommendations. This finding further suggests that visitors can find products with either a better taste fit with similar prices, lower prices with similar taste fit, or both.

Panel B of Table 6 reports the effect of recommendations on the price distribution of the products browsed in a purchase funnel from Specification (4.2). Interestingly, recommendations have a negative effect on the minimum price, the price at the 25 percentile, and the median price. However, the effect of recommendations on the price at the 75 percentile and the maximum price is insignificant. These results further reveal that recommendations help visitors explore a larger number of lower-priced products but a similar number of higher-priced products in the purchase funnels.

Table 6. Recommendations and Consumers' Search Behavior

Panel A. Consumer Search

	(1) AveAffScore	(2) AvePrice	(3) AveHFit	(4) AveRatio
	Fixed Effect Specification (4.1)			
Rec. Indicator	0.392***	-0.122*	0.208***	0.007***
	(0.017)	(0.048)	(0.016)	(0.001)
Constant	1.237***	39.314***	0.256***	0.005***
	(0.009)	(0.024)	(0.008)	(0.000)
R^2	0.130	0.457	0.320	0.236
	Fixed Effect 2SLS Specification (4.2)			
No. of RP Page View	2.803***	-0.869**	1.487***	0.053***
	(0.166)	(0.320)	(0.040)	(0.004)
Constant	0.726***	39.473***	-0.015*	-0.005***
	(0.042)	(0.081)	(0.010)	(0.001)
Obs.	572,245	572,245	572,245	572,245

Notes: + p < .10, * p < .05, ** p < .01, *** p < .001. Standard errors cluster corrected at product subcategory level in parentheses.

Panel B. Price Search

	(1) Minimum Price	(2) Price at 25th Percentile	(3) Median Price	(4) Price at 75th Percentile	(5) Maximum Price
	Fixed Effect Specification (4.1)				
Rec. Indicator	-0.304***	-0.251***	-0.131**	0.009	0.085
	(0.051)	(0.047)	(0.047)	(0.054)	(0.054)
Constant	37.900***	38.172***	39.222***	40.426***	40.979***
	(0.025)	(0.023)	(0.023)	(0.027)	(0.027)
R^2	0.451	0.453	0.455	0.442	0.431
	Fixed Effect 2SLS Specification (4.2)				
No. of RP Page View	-2.176***	-1.797***	-0.937**	0.064	0.608
	(0.354)	(0.328)	(0.315)	(0.388)	(0.407)
Constant	38.297***	38.500***	39.393***	40.414***	40.868***
	(0.090)	(0.083)	(0.079)	(0.098)	(0.103)
Obs.	572,245	572,245	572,245	572,245	572,245

Notes: + p < .10, * p < .05, ** p < .01, *** p < .001. Standard errors cluster corrected at product subcategory level in parentheses.

6.2. Benefits of Recommendations to Consumers

Analyses in the previous section reveal that recommendations help visitors find products that offer them higher net value due to lower prices and better fit for their tastes. We expect that discovering higher-value products should affect visitors' purchase behaviors with recommendations. We first examine the effects of recommendations on visitors' purchase probability and then on the outcomes related to the purchased products.

6.2.1. Effect of Recommendations on Purchase Probability

We estimate the following specifications to evaluate the effect of recommendations on the probability of product purchase:

$$ProbPurchase_i = \beta_c + \beta_1 \times Rec_i + \varepsilon_i, \tag{5.1}$$

$$ProbPurchase_i = \beta_c + \beta_1 \times X_i + \varepsilon_i, \tag{5.2}$$

where $ProbPurchase_i$ denotes the probability of product purchase in purchase funnel *i*. X_i denotes average affinity score, average fit-over-price ratio, average horizontal fit, and the variables capturing price distribution of products browsed in purchase funnels, including the minimum price, the price at the 25th percentile, and the median price of products browsed in the purchase funnel *i*. All other variables have the same meanings as in the previous specifications.

We first use the OLS Specification (5.1) to estimate the effect of the intent to treat (Rec. Indicator) on visitors' purchase probability; the Logistic regression produces qualitatively similar results. Column (1) of Table 7 indicates that showing product recommendations on average results in a higher purchase probability. The coefficient value of 0.001 translates into a two percent increase in purchase probability over the mean value of 0.06. The increase in the visitors' purchase likelihood plausibly results from that they can find higher net value products with the help of recommendations.

To empirically show this fact, we examine the effect of average affinity score, average fit-over-price ratio, average horizontal fit, and the variables capturing price distribution of products browsed in purchase funnels on the purchase probability using Specification (5.2). Since these variables related to visitors' product search are endogenous, we use randomized treatment indicator variable as an instrumental variable and use the two-stage least square (2SLS) regression to estimate their effects. Since recommendations affect purchase probability by helping visitors find higher net value (affinity score) products, the *Rec* indicator satisfies the exclusion restrictions for the average affinity score in Specification (5.2). Column (2) in Table 7 reports the results from the 2SLS regressions. We find a positive and significant coefficient for the average affinity score, indicating that higher affinity scores of browsed products under recommendations indeed lead to a higher purchase probability.

Similarly, we find a positive and significant coefficient for the average fit-over-price ratio in Column (3), indicating a higher fit-over-price ratio of the browsed products under recommendations leads to a higher purchase probability. Note that a higher fit-over-price ratio could be due to a better fit with the same price, a similar fit with a lower price, or both. Thus, the recommendations can affect visitors' purchase probabilities through either of the two mechanisms – helping visitors find higher HFit and lower-priced products.

We further estimate the effect of better-fit and lower-priced products due to recommendations on purchase probability from Specification (5.2). We note that the Rec indicator does not satisfy the exclusion restriction for horizontal fit and price variables individually in Specification (5.2). However, we present these estimates merely as suggestive evidence.

We find a positive and significant estimate for HFit in Column (4), indicating that a higher taste fit under recommendations is associated with increased purchase probability. The results in Columns (5)-(8) of Table 7 shows that while the change in the average price of the browsed products does not have a significant effect on purchase probability, a decrease in the minimum price, the price at the 25 percentile, and the median price of the browsed products results in a higher purchase probability.

Table 7. Effects of Recommendations on Purchase Probability

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS	(7) 2SLS	(8) 2SLS
Rec. Indicator	0.001^{*} (0.001)							
AveAffScore		0.003^{*} (0.001)						
AveRatio			0.167^{*} (0.078)					
AveHFit				0.006^{*} (0.003)				
AvePrice					-0.010 (0.006)			
Minimum Price						-0.004^{*} (0.002)		
Price at 25th Percentile							-0.005^{*} (0.002)	
Median Price								-0.009^{+} (0.005)
Constant	0.060^{***} (0.000)	0.056^{***} (0.002)	0.059^{***} (0.001)	0.059^{***} (0.001)	0.456^{+} (0.243)	0.212^{**} (0.073)	0.246^{**} (0.092)	0.426^{*} (0.215)
R^2	0.0064							
Obs.	572,247	572,247	572,247	572,247	572,247	572,247	572,247	572,247

Notes: + p < .10, * p < .05, ** p < .01, *** p < .001. Standard errors cluster corrected at product subcategory level in parentheses.

Overall, the results suggest that recommendations lead to a higher purchase probability and thus a lower likelihood of failed product search without purchase. More importantly, the higher chance of product purchase is driven by the fact that visitors with recommendations can find and view products of larger net value, which are either lower-priced, better-fit products, or both.

6.2.2. Effect of Recommendations on Purchased Products

We estimate the effect of recommendations on the purchased products in this section using Specification (6). Specifically, we estimate the impact of recommendations on the affinity score, price, horizontal fit, and fit-over-price ratio of the purchased products.

$$Y_i = \beta_c + \beta_1 \times Rec_i + \varepsilon_i, \tag{6}$$

where the dependent variable Y_i includes the affinity score (*PurAffScore*), price (*PurPrcie*), horizontal fit (*PurHFit*), and fit-over-price ratio (*PurRatio*) of the purchased products. All other variables have the same meanings as in previous specifications.

Table 8 reports the results from Specification (6). We find that recommendations help visitors purchase products of larger net value (as measured by the affinity score), lower price, and better fit to their tastes. We also find recommendations help visitors buy products with a higher fit-over-price ratio.

Table 8. Effects of Recommendations on Purchased Product

	(1) PurAffScore	(2) PurPrice	(3) PurHFit	(4) PurRatio
Rec. Indicator	0.506^{***} (0.048)	-0.396^{***} (0.105)	0.414^{***} (0.032)	0.016^{***} (0.001)
Constant	2.270^{***} (0.024)	32.858^{***} (0.053)	0.510^{***} (0.016)	0.016^{***} (0.001)
R^2	0.037	0.453	0.090	0.070
Obs.	34,824	34,824	34,824	34,824

Notes: + p < .10, * p < .05, ** p < .01, *** p < .001. Standard errors cluster corrected at product subcategory level in parentheses.

Taken together, the results in Tables 7 and 8 show that visitors benefit from recommendations in two ways. First, they have a lower likelihood of failed search effort due to a higher probability of product purchase under recommendations. Second, visitors with recommendations can search and purchase products that offer them higher net value due to lower product prices, better fit to their tastes, or both.

6.2.3. Additional Evidence for Higher Taste Fit and Lower Price with Recommendations

Next, we utilize the variation in product characteristics across product categories to provide additional evidence of better fit and lower prices of purchased products under recommendations. If recommendations help visitors find lower-priced products in a product category, this effect should be more prominent in the product categories with higher average prices and higher price dispersions. Similarly, if recommendations help visitors find better fit products, this effect should be more pronounced in product categories with highly heterogeneous visitors' tastes, such as women's apparel.

We examine the moderating effect of product categories on the impact of recommendations on the characteristics of purchased products with the following specifications (7).

$$Y_i = \beta_c + \beta_1 \times Rec_i + \beta_2 \times (Rec_i \times PCat_c) + \varepsilon_i, \quad (7)$$

where the dependent variable Y_i is the price of purchased products. $PCat_i$ denotes the product category-level moderating factors in product subcategory c. We use the average product price and relative price dispersion (measured by the ratio of the price variance over the average price) in a product subcategory to show the differential effects of recommendations in finding lower-priced products. All other variables have the same meanings as in the previous specifications.

Columns (1) and (2) in Table 9 reports the results for the price of purchased products as DV in Specification (7). We find negative and significant coefficients for the interaction term of *Rec. Indicator* with moderating factors related to price distribution. First, we find a significant negative coefficient for the interaction of *Rec. Indicator* and average product price, suggesting that visitors pay a lower purchase price by using recommendations for product categories with a higher average price. It is understandable because product categories with higher prices have a larger room for searching lower prices. Second, we find a significant negative coefficient for the interaction of *Rec. Indicator* and relative price dispersion, suggesting that visitors pay a lower purchase price by using recommendations in product categories with a larger relative price dispersion. It is perhaps because such product subcategories offer visitors a greater opportunity to search for lower-priced products from a larger price variation around the average price.

Next, we examine the moderating effect of women’s product category in the impact of recommendations in finding products that fit visitors' tastes better. We first estimate Specification (7) on products in women’s and men's apparel categories, separately. The dependent variable in the analyses is the horizontal fit of purchased products, and *PCat* denotes the indicator variable for women’s apparel in Specification (7).

Table 9. Moderating Effects of Category-level Average Price and Price Dispersion

DV: PurPrice	(1)	(2)
Rec. Indicator	0.335	0.112
	(0.220)	(0.153)
Rec. Indicator × Average Price	-0.019***	
	(0.006)	
Rec. Indicator × Relative Price Dispersion		-0.072***
		(0.015)
Constant	32.858***	32.859***
	(0.045)	(0.045)
R^2	0.453	0.453
Obs.	34,824	34,824

Notes: + p < .10, * p < .05, ** p < .01, *** p < .001. Standard errors cluster corrected at product subcategory level in parentheses.

Table 10. Moderating Effects of Category-level Taste Heterogeneity

DV: HFit	(1) Women	(2) Men	(3) Pooled
Rec. Indicator	0.453***	0.339***	0.339***
	(0.041)	(0.050)	(0.048)
Rec. Indicator × Women			0.114+
			(0.063)
Constant	0.512***	0.114**	0.423***
	(0.021)	(0.025)	(0.017)
R^2	0.054	0.066	0.059
Obs.	14,322	4,091	18,413

Notes: + p < .10, * p < .05, ** p < .01, *** p < .001. Standard errors cluster corrected at product subcategory level in parentheses.

Columns (1) and (2) of Table 10 report the results on horizontal fit of purchased products for women's and men's apparel categories, separately. We find positive and significant coefficients of *Rec. Indicator* for both product categories, while the magnitude of the estimated effect for women's apparel categories is larger. Moreover, Column (3) of Table 10 suggests a positive and significant coefficient for the interaction term of *Rec. Indicator* with Women product category indicator. Therefore, the results in Table 10 consistently support that the effect of recommendations on increasing horizontal fit of purchased products is greater for women's apparel categories than for men's apparel categories. Perhaps because of the higher heterogeneity in visitors' tastes for products in women's compared to men's product categories, there is a larger scope for recommendations to help find products that better fit visitors' tastes in the women's product categories.

6.3. Substitution of Existing Search Tools with Recommendations

The retailer's website offers visitors two additional search tools besides product recommendations: (1) product category-based search through the hierarchical organization of products on the website and (2) keyword-based search. We measure visitors' product category-based and keyword-based searches with the number of their product category/subcategory page views and search result page views in a purchase funnel, respectively. Visitors endogenously decide which and how much search tool to use. In our experiment, we additionally provide the recommendation tool to some randomly selected visitors. Thus, our experimental setup allows us to examine the effect of exogenous availability of recommendations on visitors' endogenous choice of search tools on the retailer's website.

We estimate the effect of the recommendation availability and usage (measured by the number of RP page views) on existing search tools use on the website with the following specifications:

$$Y_i = \beta_c + \beta_1 \times Rec_i + \varepsilon_i, \tag{8.1}$$

$$Y_i = \beta_c + \beta_1 \times NRPView_i + \varepsilon_i, \tag{8.2}$$

where dependent variable Y_i denotes the usage of two existing search tools on the website, i.e., the number of search result page views and product category/subcategory page views in the purchase funnel i. All other variables have the same meanings as in the previous specifications.

Table 11 reports the estimated coefficients from Specifications (8.1) and (8.2). In Columns (1) and (4), we find negative and significant coefficients for the recommendation indicator variable in Specification (8.1). These estimates indicate that visitors reduce their keyword-based and product-category-based searches with the availability of recommendations. Specifically, visitors with recommendations view 0.02 fewer search result pages and 0.09 fewer product subcategory pages in their purchase funnels.

Table 11. Effects of Recommendations on Usages of Other Search Tools

DV:	*No. of Search Result Page Views*			*No. of Product Category Page Views*		
	(1)	(2)	(3)	(4)	(5)	(6)
Rec. Indicator	-0.02^{**}			-0.09^{***}		
	(0.01)			(0.02)		
No. of RP Page Views		0.48^{***}	-0.14^{***}		1.61^{***}	-0.63^{***}
		(0.04)	(0.04)		(0.25)	(0.09)
Constant	0.71^{***}	0.58^{***}	0.74^{***}	1.79^{***}	1.34^{***}	1.91^{***}
	(0.00)	(0.01)	(0.01)	(0.01)	(0.06)	(0.02)
Instrumental Variable		No	Rec. Indicator		No	Rec. Indicator
Obs.	572,247	572,247	572,247	572,247	572,247	572,247

Notes: + p < .10, * p < .05, ** p < .01, *** p < .001. Standard errors cluster corrected at product subcategory level in parentheses.

We suspect that the number of RP page views (*NRPView*) in Specification (8.2) would be endogenous, as visitor-level unobserved characteristics may determine both their usage of recommendations and other search tools on the retailer's website. Columns (2) and (5) of Table 11 show a positive and significant correlation of the number of RP page views with the number of the search result page and product category page views, respectively, from a fixed effect OLS estimation of Specification (8.2). The positive correlation between keyword search and recommendations suggests complementarity between the two search tools. However, the estimate of positive correlation is biased and misleading due to the omission of visitor-level factors. For example, while more price-sensitive visitors may use product category-based and keyword-based searches more intensively, they are also likely to click and view recommended product pages.

To address the endogeneity of *NRPView* in Specification (8.2), we instrument it with the indicator variable of recommendations. In the previous sections, we have already shown that the recommendation indicator satisfies the conditions of exogeneity and relevance for instrumental variables. Thus, we find the effect of the number of RP page views attributed to the availability of recommendations on the number of search result page views. Columns (3) and (6) in Table 11 report the estimates from the 2SLS regressions. The results show negative and significant coefficients for the number of RP page views from Specification (8.2), which indicates substitution between the RP page views and the number of the search result page (and also product subcategory page views). Specifically, the coefficient estimates indicate that an additional RP page view due to the recommendations decreases in 0.14 search page views and 0.63 product category page views, respectively. Thus, the availability of recommendations results in visitors substituting the existing search tools with recommendations.

7. Conclusion

We conduct a field experiment on a US-based mid-sized retailer website to investigate the effect of recommendations on a visitor's search behavior during her purchase process. We use the unique data of affinity scores computed by the recommendation algorithm as a measure of a product's net value to the visitor. We further measure a product's horizontal fit with visitors' tastes with the part of the affinity score uncorrelated to its price. In this way, we separately estimate the two underlying mechanisms through which recommendations benefit visitors – helping them find lower-priced products and better fit products.

We examine the effect of these two mechanisms of recommendations on visitors' search behavior, purchase probability, and finally, characteristics of purchased products. We find that visitors could search for lower-priced products and also better fit products, thereby providing them higher net value with the help of recommendations. The ability to find higher value products with recommendations results in increased purchase probability and ultimately the purchase of lower-priced products, better fit products, or both. Our estimates of the benefits of recommendations to consumers due to lower price and better fit offer insights into their relative values and thus have implications for recommendation algorithm design. Furthermore, our results of higher benefit of finding lower-priced (better taste fit) products in product categories with heterogeneous prices (consumer tastes) offer practical guidance for adjusting recommendation algorithms to the product category characteristics.

We also find that visitors substitute the existing search tools with recommendations after it is made available. Our estimates indicate that the availability of recommendations results in more than two times higher substitution in category-based searches compared to keyword-based searches. These estimates provide cross elasticities of different search tools and have implications for deploying different search tools on eCommerce platforms. For example, our findings suggest that eCommerce websites should provide better instead of more search tools to the consumers.

Our research has several limitations that provide opportunities for future research on product recommendations. Our estimates are for a widely used generic collaborative filtering-based recommendation system in the context of apparel and accessories. Our estimates may not be applicable for recommendation algorithms that are designed to optimize retailers' revenue (or profit). Future research should estimate the benefit of such recommendation systems to consumers. Moreover, our estimates for consumers' taste fit is relevant for the apparel product categories, but may not relevant for other product categories with low horizontal differentiation across products. Further studies can seek to understand the contributions of the two mechanisms of recommendations across products in different product categories. Another limitation of our experiment is the possible imprecise identification of visitors. A treated visitor may be misclassified as a control visitor when she uses a different device. However, this is a common limitation of most online experiments. Moreover, such overlap of visitors to different treatment conditions only makes our results more conservative. Future research should devise better methods of visitor identification.

References

Abdollahpouri H, Adomavicius G, Burke R, Guy I, Jannach D, Kamishima T, Krasnodebski J, Pizzato L (2020) Multistakeholder recommendation: Survey and research directions. *User Modeling and User-Adapted Interaction*. 30(1) 127-158.

Ali A, Zhang W (2015) CEO tenure and earnings management. *Journal of Accounting and Economics*. 59(1) 60-79.

Basu S, Fernald J G, Kimball M S (2006) Are technology improvements contractionary? *American Economic Review*. 96(5) 1418-1448.

Bleier A, Eisenbeiss M (2015) Personalized online advertising effectiveness: The interplay of what, when, and where. *Marketing Science*. 34(5) 669-688.

Bound J, Jaeger D A, Baker R M (1995) Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*. 90(430) 443-450.

Brynjolfsson E, Rock D, Syverson C (2021) The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics*. 13(1) 333-372.

Chapelle O, Zhang Y (2009) A dynamic bayesian network click model for web search ranking. *Proceedings of the 18th International Conference on World Wide Web* 1-10.

Chen N, Li A, Talluri K (2021) Reviews and self-selection bias with operational implications. *Management Science*.

Chen Y, Yao S (2017) Sequential search with refinement: Model and application with click-stream data. *Management Science*. 63(12) 4345-4365.

De P, Hu Y, Rahman M S (2010) Technology usage and online sales: An empirical study. *Management Science*. 56(11) 1930-1945.

Dellaert B G, Häubl G (2012) Searching in choice mode: consumer decision processes in product search with recommendations. *Journal of Marketing Research*. 49(2) 277-288.

Dinkelman T (2011) The effects of rural electrification on employment: New evidence from South Africa. *American Economic Review*. 101(7) 3078-3108.

Dou Y, Hope O-K, Thomas W B (2013) Relationship-specificity, contract enforceability, and income smoothing. *The Accounting Review*. 88(5) 1629-1656.

Ghose A, Ipeirotis P G, Li B (2019) Modeling consumer footprints on search engines: An interplay with social media. *Management Science*. 65(3) 1363-1385.

Goldenberg J, Oestreicher-Singer G, Reichman S (2012) The quest for content: How user-generated links can facilitate online exploration. *Journal of Marketing Research*. 49(4) 452-468.

Honka E (2014) Quantifying search and switching costs in the US auto insurance industry. *The RAND Journal of Economics*. 45(4) 847-884.

Hosanagar K, Fleder D, Lee D, Buja A (2014) Will the global village fracture into tribes? Recommender systems and their effects on consumer fragmentation. *Management Science*. 60(4) 805-823.

Hu N, Pavlou P A, Zhang J J (2017) On self-selection biases in online product reviews. *MIS Quartely*. 41(2) 449-471.

Jones J J (1991) Earnings management during import relief investigations. *Journal of Accounting Research*. 29(2) 193-228.

Jorgenson D W, Stiroh K J (1999) Information technology and growth. *American Economic Review*. 89(2) 109-115.

Kim K, Mauldin E, Patro S (2014) Outside directors and board advising and monitoring performance. *Journal of Accounting and Economics*. 57(2-3) 110-131.

Kumar A, Hosanagar K (2019) Measuring the Value of Recommendation Links on Product Demand. *Information Systems Research*. 30(3) 819-838.

Kumar A, Tan Y (2015) The demand effects of joint product advertising in online videos. *Management Science*. 61(8) 1921-1937.

Kwark Y, Chen J, Raghunathan S (2017) Platform or wholesale? A strategic tool for online retailers to benefit from third-party information. *MIS Quarterly*. 41(3) 763-785.

Lambrecht A, Tucker C (2013) When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research*. 50(5) 561-576.

Lee D, Hosanagar K (2019) How do recommender systems affect sales diversity? A cross-category investigation via randomized field experiment. *Information Systems Research*. 30(1) 239-259.

Lee D, Hosanagar K (2021) How Do Product Attributes Moderate the Impact of Recommender Systems? *Management Science*. 67(1) 524-546.

Li X, Hitt L M (2008) Self-selection and information role of online product reviews. *Information Systems Research*. 19(4) 456-474.

Lin Z, Goh K Y, Heng C S (2017) The demand effects of product recommendation networks: an empirical analysis of network diversity and stability. *MIS Quarterly*. 41(2) 397-426.

Linden G, Smith B, York J (2003) Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*. 7(1) 76-80.

Ngwe D, Ferreira K J, Teixeira T (2019) The impact of increasing search frictions on online shopping behavior: Evidence from a field experiment. *Journal of Marketing Research*. 56(6) 944-959.

Oestreicher-Singer G, Sundararajan A (2012) The visible hand? Demand effects of recommendation networks in electronic markets. *Management Science*. 58(11) 1963-1981.

Ratchford B T, Srinivasan N (1993) An empirical investigation of returns to search. *Marketing Science*. 12(1) 73-87.

Reinganum J F (1982) Strategic search theory. *International Economic Review*. 23(1) 1-17.

Sarwar B M, Karypis G, Konstan J A, Riedl J (2001) Item-based collaborative filtering recommendation algorithms. *In Proceedings of the 10th international conference on World Wide Web.* 285-295.

Seiler S, Pinna F (2017) Estimating search benefits from path-tracking data: measurement and determinants. *Marketing Science*. 36(4) 565-589.

Solow R M (1956) A contribution to the theory of economic growth. *The Quarterly Journal of Economics*. 70(1) 65-94.

Solow R M (1957) Technical change and the aggregate production function. *The Review of Economics and Statistics*. 39(3) 312-320.

Sutton J (1986) Vertical product differentiation: some basic themes. *The American Economic Review*. 76(2) 393-398.

Thorat P B, Goudar R, Barve S (2015) Survey on collaborative filtering, content-based filtering and hybrid recommendation system. *International Journal of Computer Applications*. 110(4) 31-36.

Ursu R M, Wang Q, Chintagunta P K (2020) Search duration. *Marketing Science*. forthcoming.

Weitzman M L (1979) Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*. 47(3) 641-654.

Appendix A. Balance Check

To validate the randomness of our assignment, we check whether visitor characteristics of treated and control visitors (purchase funnels) are similar. Specifically, we examine whether the following variables related to visitors are statistically similar for treated and control visitors (purchase funnels): (a) new (vs. repeat) visitors; (b) frequent (vs. infrequent) visitors, indicating whether the total visitor session number is above the 75th percentile value in the distribution of the number of sessions by a visitor; (c) the way through which a visitor comes to a session, such as paid search, and organic search, and others. Table A1 reports the results, showing that these variables are statistically indistinguishable between the treated and control visitors (purchase funnels).

Table A1. Balance Check at the Visitor and Purchase Funnel Level

		Control		Treated		Diff. in means
		Obs.	Mean (Std. Dev.)	Obs.	Mean (Std. Dev.)	(*t*-stats)
Visitor level						
New visitor		214,737	0.8778 (0.33)	214,475	0.8767 (0.33)	-0.0010 (-1.03)
Frequent visitor		214,737	0.0113 (0.11)	214,475	0.0112 (0.11)	-0.0001 (-0.16)
	Paid search	214,737	0.3228 (0.47)	214,475	0.3215 (0.47)	-0.0013 (-0.95)
Loading	Organic search	214,737	0.1109 (0.31)	214,475	0.1115 (0.31)	0.0006 (0.65)
	Others	214,737	0.5663 (0.50)	214,475	0.5670 (0.50)	0.0007 (0.48)
Purchase funnel level						
New visitor		285,594	0.8451 (0.36)	286,653	0.8453 (0.36)	0.0002 (0.20)
Frequent visitor		285,594	0.0343 (0.18)	286,653	0.0349 (0.18)	0.0006 (1.20)
	Paid search	285,594	0.3156 (0.46)	286,653	0.3148 (0.46)	-0.0008 (-0.66)
Loading	Organic search	285,594	0.1141 (0.32)	286,653	0.1144 (0.32)	0.0003 (0.42)
	Others	285,594	0.5703(0.50)	286,653	0.5707 (0.49)	0.0004 (0.35)

Notes: Diff. in means = Treated - Control. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.