

How Do Recommender Systems Affect Sales Diversity? A Cross-Category Investigation Via Randomized Field Experiment

Dokyun Lee
Carnegie Mellon University

Kartik Hosanagar
The Wharton School*

Abstract

We investigate the impact of collaborative filtering recommender algorithms (e.g., Amazon’s “Customers who bought this item also bought”) commonly used in e-commerce on sales diversity. We use data from a randomized field experiment run on a top retailer in North America across 82,290 SKUs and 1,138,238 users. We report four main findings. First, we demonstrate across a wide range of product categories that the use of traditional collaborative filters (or CFs) is associated with a decrease in sales diversity relative to a world without product recommendations. Further, the design of the CF matters. CFs based on purchase data are associated with a greater effect size than those based on product views. Second, the decrease in aggregate sales diversity may not always be accompanied by a corresponding decrease in individual-level consumption diversity. In fact, it is even possible for individual consumption diversity to increase while aggregate sales diversity decreases. Third, co-purchase network analysis shows that recommenders can help individuals explore new products but similar users end up exploring the same kinds of products resulting in the concentration bias at the aggregate level. Fourth and finally, there is a difference between absolute and relative impact on niche items. Specifically, absolute sales and views for niche items in fact increase, but their gains are smaller compared to the gains in views and sales for popular items. Thus, while niche items gain in absolute terms, they lose out in terms of market shares.

Keywords: E-commerce, Personalization, Recommender systems, Sales diversity, Consumer purchase behavior, Collaborative filtering, Gini coefficient

*The authors gratefully acknowledge the financial support from the Jay H. Baker Retailing Center, the Wharton Risk Management and Decision Processes Center, the Mack Institute for Innovation Management, and the Fishman-Davidson Center for Service and Operations Management. We also thank Adit Bharat Sanghvi for research assistance. All errors are our own. This version: June, 2017.

1 Introduction

Recommender systems are widely used across a number of industries, ranging from online retail (e.g., Amazon’s product recommendations), music (e.g., Spotify’s weekly recommendations), video streaming (e.g., YouTube, Netflix), and news (e.g., Google News). These systems attempt to predict items of interest to users based on information about the users and items. The most common recommender design used in e-commerce is the Collaborative Filter, which is a recommender that finds other users whose product views or purchases are most similar to the focal user and recommends other items that they have viewed or purchased (Schafer et al., 1999). Examples include Amazon’s “people who bought X also bought Y” and Netflix’s “People like you liked X.”

Recommenders are an interesting example of a marketing technology that offers significant value to both consumers and firms. For consumers, recommenders help them learn about new products (Resnick and Varian, 1997) and sort through large choice sets (Häubl and Trifts, 2000). For firms, recommenders help convert browsers to buyers (Tam and Ho, 2005; Dias et al., 2008; De et al., 2010; Pathak et al., 2010; Oestreicher-Singer and Sundararajan, 2012), promote cross-selling (Pathak et al., 2010) and increase loyalty by providing a custom browsing experience (Schafer et al., 1999). As a result, most of the major online firms use recommenders. A recent survey indicated that 94% of the top e-commerce companies agree that personalization is critical to their success online¹. And they indicate that recommenders have a major impact on consumer choice—for example, 60% of media choices on Netflix (Thompson, 2008) and nearly 35% of sales at Amazon (Lamere and Green, 2008) originate from recommendations.

While we know that recommenders influence individual choice, there is growing interest in understanding how they affect the sales diversity, i.e., the overall distribution of sales. Some believe recommender systems will help consumers discover new products by lowering their search costs (Brynjolfsson et al., 2011; Anderson, 2008). For example, in the book *The Long Tail*, author Chris Anderson suggests that “the main effect of recommenders will be to help people move from the world of hits to the world of niches.” On the other hand, another viewpoint is that common recommender systems—in particular Collaborative Filters (CFs)—will decrease sales diversity (Mooney and Roy, 2000; Fleder and Hosanagar, 2009; Wu et al., 2011; Celma and Cano, 2008). Hosanagar and Fleder present an analytical model and simulation study to show that CFs only reinforce popularity of

¹<https://econsultancy.com/reports/the-realities-of-online-personalisation-report/>

already popular titles. They argue that this is because CFs rely on past purchase or view data and cannot recommend items with limited historical information.

The issue of sales diversity has interesting implications for the fields of marketing, information systems (IS), and operations management (OM). In marketing, there has been a lot of interest in understanding sales concentration and the long tail phenomenon and their implications (Borle et al., 2005; Jiang et al., 2011; Van Herpen and Pieters, 2002) for consumer surplus (Elberse, 2008). If recommender systems lower consumers’ search costs and help them find better product matches that they would have otherwise missed, then they are a highly appealing marketing technology. From an IS standpoint, we are interested in understanding whether the tacit design choices we make for decision support tools have a marked impact on consumer product consumption patterns. From an OM perspective, to the extent that these systems change sales diversity, there are implications for the firms’ product assortment (Jiang et al., 2014; Walter et al., 2012). In addition, recent research has shown that inventory levels & inventory costs increase with sales dispersion (Gallino et al., 2014; Walter et al., 2012; Kok et al., 2008) and, further, different levels of sales dispersion can call for different supply chain structures (Fisher, 2003). Thus, there can be significant operational adjustments needed in response to sales distribution changes caused by recommenders.

Despite the interest in the topic and its potential implications, there exists little to no empirical evidence to date that can help us reconcile these different views on the impact of recommenders on sales diversity. We believe a potential driver of this may be the lack of data that provides a contrast between users exposed and unexposed to recommendations. Without such a contrast, it is hard to empirically identify recommenders’ impact. Our empirical investigation attempts to fill this research gap and reconcile the different viewpoints. Specifically, we use data from a randomized field experiment run by a large online retailer to investigate how recommenders affect sales diversity across all the product categories sold on the website. Because much of the debate on recommenders’ impact on sales diversity is related to CFs, we focus only on CFs in the study. However, even among CFs, there can be designs based on purchases (“Customers who purchased this also purchased”) or views (“Customers who viewed this also viewed”). Accordingly, we also investigate how the design choice of CFs affects the answer by incorporating both view- and purchase-based designs.² Throughout the paper, we refer to view-based collaborative filtering as “VBCF” and purchase-based collaborative filtering as “PBCF”.

²<http://mahout.apache.org/>

We report four main findings. First, we demonstrate across a wide range of product categories that the use of traditional collaborative filters is associated with a decrease in sales diversity relative to a world without product recommendations. Second, we find that the decrease in aggregate sales diversity may not always be accompanied by a corresponding decrease in individual-level consumption diversity. In fact, it is even possible for individual consumption diversity to increase while aggregate sales diversity decreases. Third, our analysis of the co-purchase network shows that recommenders do help individuals explore new product categories but similar users end up exploring the same kinds of products. Thus, any increase in individual diversity does not show up at an aggregate level because of the correlated exploration across individuals. Fourth, we show that there is a difference between relative and absolute gains for niche items. Specifically, absolute sales and views for niche items in fact increase, but their gains are smaller compared to the gains in views and sales for popular items. Thus, while niche items gain in absolute terms, they lose out in terms of market shares. Together, these results present the most comprehensive empirical perspective to date on how traditional collaborative filters affect sales diversity.

These findings have implications for retailers, producers, and recommender designers. Consistent with prior literature, we find that retailers clearly benefit from increased sales under recommendations. However, for retailers whose strategy is to offer greater product variety —under the premise that consumers will find better-suited products — the use of traditional CF designs may not be fully consistent with that strategy. Our results suggest that retailers offering a broader product assortment might benefit from modifying off-the-shelf designs to allow discovery of relevant items with limited historical views and/or purchases. For producers of niche titles, one of the promises of the Internet is that such products can be efficiently matched to their customers by online tools, thereby allowing one to profitably produce and sell niche products. However, our results show that producers cannot passively rely on search tools like recommenders. Instead, such producers will need to continue exerting effort to ensure product discovery by consumers. Finally, from a design science standpoint, it is useful for system designers to recognize the inherent popularity bias in CFs. If there exist better product matches outside of the relatively popular titles, it might be useful to modify CF designs to ensure discovery of relevant items with limited views and purchases.

2 Prior Work

The vast majority of work on recommenders is focused on their design. A common taxonomy of recommenders broadly classifies them into content-based recommenders and Collaborative Filtering Algorithms (CFs) (Adomavicius and Tuzhilin, 2005). Content-based systems analyze product attributes to suggest products that are similar in this attribute space to products that a consumer has bought or liked in the past. A drawback of content-based designs is that they need rich metadata about the products, which can be expensive to collect. Furthermore, they do not work well when a retailer sells multiple categories of products. CFs were designed to address these limitations. CFs are unaware of product attributes and recommend products either purchased or liked by similar consumers, where similarity is measured by historical purchase (or like) data. Because they are easy to build and don't require detailed product attribute information, they are the most widely used class of recommenders in e-commerce. Hence, they are also the focus of this study. CFs can themselves be of many types, including those based on item similarity (item-based CFs such as in Sarwar et al. (2001)) and those based on user similarity (Breese et al., 1998). Even among item-based CFs, they can be based on product views ("People who viewed this also viewed") or purchases ("People who bought this also bought"). It is well documented that the demand effects of co-view and co-purchase recommendations differ (Lin et al., 2015). This may be because Purchase-based CFs are more likely to recommend complementary rather than substitute products (and vice-versa for View-based CFs) (Lin et al., 2015). Our study is based on a retailer's implementation of item-based CF. We consider both view-based and purchase-based designs.

While the vast majority of work in computer science has focused on design of recommender algorithms, an emerging stream of work is starting to explore their impact at the individual and market levels. Senecal and Nantel (2004) show experimentally that recommendations do influence choice and that online recommendations can be more influential than those by humans. Cooke et al. (2002) examine how purchase decisions under recommendations depend on the context and familiarity with the recommended items. De et al. (2010) and Hinz and Eckert (2010) show that this influence on consumer choice is not merely a product substitution effect but they help drive an increase in sales. While these studies ask how recommenders affect individual choice, our interest is the aggregate effect they have on product markets. In particular, we are interested in how recommenders affect sales diversity.

Brynjolfsson et al. (2006) and Anderson (2008) find that sales diversity can be higher on the Internet than in offline channels. They suggest supply-side causes such as the lower cost of offering a wide product assortment and several demand-side causes, such as active tools (search engines) and passive tools (recommender systems). However, they do not isolate the specific effect of recommenders from the other factors. Holding product supply fixed, Brynjolfsson et al. (2011) find that sales diversity for an apparel retailer’s Internet channel is greater than its catalog channel. Because their interest is also in channel differences, they also do not isolate the effect of recommenders relative to other factors such as search tools. Further, the retailer in their study did not use a CF. Other studies that suggest recommenders might help increase sales diversity include Hinz and Eckert (2010); Oestreicher-Singer and Sundararajan (2012); Zhou et al. (2010). In contrast, Fleder and Hosanagar (2009) use simulations to show that CFs are more likely to recommend items that have been viewed or purchased often, thereby reducing sales diversity. Similarly, Jannach et al. (2013); Wu et al. (2011); Celma and Cano (2008) also argue or show that the use of CFs will decrease aggregate sales diversity. Table 1 summarizes these academic papers and their main claims. In sum, there is no consensus among either popular or academic literature on how recommenders will affect sales diversity.

We believe this lack of consensus arises due to many reasons. First, different studies use different algorithms. While it appears as though content-based recommenders might increase sales diversity (see for example, Hosanagar et al. (2014)), the disagreement appears to be primarily about CFs. Thus, that is the design we study in this paper. Additionally, many studies are based on lab experiments or simulations calibrated to archival data, which makes generalization more difficult. Other studies measure non-purchase attributes like purchase-intentions, use-intentions, and satisfaction rather than the actual views or purchases. The few based on field data are constrained by the limitations of observational data including: (i) the inability to separate the impact of recommenders from other supply- and demand-side factors, and (ii) the lack of a contrast between users exposed to recommendations and an otherwise similar group of users who are unexposed to recommendations.

In this study, we isolate that effect of recommenders by varying their availability while also holding supply-side factors and other demand-side factors constant. We carry out a randomized field experiment on a large e-commerce website using multiple recommender algorithms. By doing so, we are able to offer compelling evidence that commonly used recommender designs affect sales concentration.

Study	Method & Data	Sales Diversity
Hinz and Eckert (2010)	MovieLens Data & Simulation	Increased niche product consumption leading to increase in aggregate sales diversity
Fleder and Hosanagar (2009)	Theoretical Models & Simulation	Decrease in aggregate sales diversity but increase in individual sales diversity
Hosanagar et al. (2014)	Archival Data & Econometrics	Content-based RS increase aggregate sales diversity and increase overlap/commonality in consumption
Oestreicher-Singer and Sundararajan (2012)	Crawled Amazon Data & Econometrics	Recommender shifts demand to niche item increasing aggregate sales diversity
Jannach et al. (2013)	MovieLens Data & Simulation	Different algorithms have different effects
Wu et al. (2011)	MovieLens Data & Simulation	Mixed result based on different algorithms. Collaborative filtering decreases aggregate diversity while content-based increases it
Celma and Cano (2008)	last.fm and Allmusic.com API data & Correlational Analysis	Collaborative filtering algorithm is linked to popularity bias suggesting decreased aggregate consumption diversity
Zhou et al. (2010)	Crawled YouTube Data & Correlational Analysis	Increases aggregate consumption diversity

Table 1: Literature on Impact of Recommender Systems and Claims

3 Problem Statement and Study Design

This section formally sets up research questions, experimental study designs, and empirical strategies.

3.1 Research Question

We are interested in studying the impact of recommenders on sales diversity. We measure the *sales diversity* of the products sold with a measure called the Gini coefficient. The Gini coefficient has been widely adopted in the long tail and the RS literature as a measure of sales diversity (Fleder and Hosanagar, 2009; Brynjolfsson et al., 2011). It is computed based on the Lorenz curve. Let $L(u)$ be the Lorenz curve denoting the percentage of the sales generated by the lowest $100u\%$ of items as shown in Figure 1. The Gini coefficient is defined as $G \equiv \frac{A}{A+B}$. It ranges from 0, representing the least amount of concentration or highest diversity, to 1, representing the highest amount of

concentration or lowest diversity. A Gini coefficient of 0 means that all products have equal sales, while values near 1 mean that a few broad-appeal blockbuster items account for most of the sales.

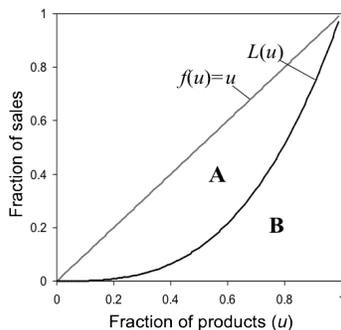


Figure 1: **Lorenz Curve**

We approach this problem with a field experiment in which consumers visiting a website are randomly assigned to a control or treatment group. The treatment group is shown a panel of different recommendations, much like Amazon’s “Customers who bought this item also bought” recommenders. The control group is shown nothing. For each group, we analyze the following variables of interest for sales diversity.

1. Aggregate firm-level **view and sales** diversity: This measures how the recommenders affect product view/sales diversity at the aggregate level (for each treatment group) and is measured by the Gini coefficient.
2. Individual average **view and purchase** diversity: This measures how the recommenders affect the diversity of products individuals view or purchase. Again, the Gini coefficient is used, but it is computed separately for each individual based only on their own purchases.

3.2 Treatment Conditions

We compare the sales diversity of users exposed to recommendations with that of another set of users unexposed to recommendations. Note that, in theory, it is always possible to design a recommender that increases concentration (recommend bestsellers) or diversity (recommend items with lowest sales). But our question is not whether there are designs that can achieve these effects. Instead, our empirical focus is on the impact of common designs. Accordingly, we use item-based CFs in this study. These CF designs have a long history (Sarwar et al., 2001) and continue to be very popular designs even today. There are three groups in our study:

1. Control (no recommendations)
2. View-based collaborative filtering (“People who viewed this item also viewed”) (henceforth “VBCF”)
3. Purchase-based collaborative filtering (“People who purchased this item also purchased”) (henceforth “PBCF”)

We have two different treatment groups corresponding to two different recommender algorithms, plus a control group that was not shown any recommendations. One treatment is based on views (“People who viewed this item also viewed”) and the other is based on purchases (“People who purchased this item also purchased”). We consider these two treatments because they are two of the most commonly used types of collaborative filtering algorithms. Further, while both are CF designs, they recommend very different kinds of products as discussed in Section 2. When we study the impact of recommenders, it is worth asking “relative to what.” We compare sales diversity under these designs against a group that receives no recommendations. This is how the problem has been framed in the literature (Hinz and Eckert, 2010; Jannach and Hegelich, 2009; Fleder and Hosanagar, 2009). One could alternatively study the impact of recommender systems relative to a random recommender or to a system that showcases the most popular items. These alternatives are unappealing for multiple reasons. From a practical perspective, when a consumer is on a specific product page, showing randomly generated recommendations or globally popular items will effectively show irrelevant items. As a result, it is not a format that is used by any retailer we know.

Our recommender system implements an item-based collaborative filtering algorithm using Apache Mahout (mahout.apache.org), an open-source, machine learning framework widely used in online retail. The item-based CF we implement in this experiment computes item-item similarity with pre-existing consumer purchase/view data. A critical input for the recommendation is the focal item (the product a user is viewing). However, the algorithm also considers the user’s past product views/purchases. The top N candidate products that are not yet purchased/viewed by the consumer are then recommended. We follow the details as described in Sarwar et al. (2001). Take, for example, VBCF. When a new user first visits a focal item page P with no prior history on the website, the recommendations are general and same for all brand-new users. As the user now views different items on the site, the user’s view vector changes. Now the view-based recommendations

(for the same focal item P) change to a new set of items based on newly updated view-history vector.

The algorithm uses the purchase/view data of the entire website 60 days prior to the start of the experimentation and the item-item similarity matrix is recomputed every three days. The number of products displayed is a function of the width of the user’s screen but the default is 6. We discuss more details of our treatment setup and data in Section 4. Next, we discuss our empirical strategies given data from this field experimental setup.

3.3 Study Design

Let g_i represent group i and let f represent a function that calculates an aggregate measure of interest, D_i , for the given group (e.g., the group-level purchase Gini coefficient). We define the following quantity of interest:

Aggregate Measure, f , of Group 1	$D_1 \equiv f(g_1)$
Aggregate Measure, f , of Group 2	$D_2 \equiv f(g_2)$
Difference in Aggregate Measures	$D \equiv D_1 - D_2$

The difference in the aggregate measure, D , shows how different Group 1 is from Group 2. Let $\mu \equiv \mathbb{E}[D]$, with the distribution of D unknown. All hypotheses testing in this paper takes the form:

Null Hypothesis H_0	$\mu \equiv \mathbb{E}[D] = 0$
Alternate Hypothesis H_a	$\mu \equiv \mathbb{E}[D] \neq 0$

Note that we carry out the hypotheses tests as two-sided tests (equal or not equal rather than greater than or less than) to remain conservative. Further, since we have one aggregate measure (or statistic) for each group, in order to produce a p-value, we utilize a permutation test technique (Good, 2005) that allows us to calculate a null distribution for a given aggregate measure. We choose to utilize permutation tests since 1) the permutation test does not make any distributional assumptions, 2) permutation tests can be utilized for any aggregate statistics, and 3) it is well-fitted to handle unbalanced designs, as in our case where there are more control users. The permutation test works as follows. Suppose the Null hypothesis is true and control and treated groups are similar. Then randomly relabeling some control users as treated or treated users as control users is unlikely to affect our test. A permutation test involves repeatedly and randomly relabeling individuals into Groups 1 and 2 (e.g., control and treated) to produce a null distribution for any test statistics. In

each iteration, we randomly take half of the sample from the control group and the other half from the treated group. By comparing statistics from null distributions to the actual test statistics from the real distribution and tallying how often null distribution statistics exceed the actual distribution statistic, we can determine the p-value. For more details, see Good (2005). In our study, we use 1,000 iterations to get an accurate p-value up to 0.001.

4 Data Description

Our dataset comes from a field experiment on the Canadian website of one of the top five retailers in North America. The experiment was conducted over two weeks between August 8, 2013 and August 22, 2013. The main dataset records item views and purchases of 1,138,238 unique users across all the product categories sold on the website (which spans 82,290 unique SKUs). The dataset has 2.8 million rows of individual-item-level data.

The field experiment was run by the retail company using a state-of-the-art A/B/n testing platform. The testing platform embeds snippets of code on the retailer’s website, much like Google Analytics, and controls all facets of A/B/n testing via a backend analytics dashboard. This A/B/n platform implements advanced cross-device customer identification strategies that combine IP addresses, cookies, log-in information, etc. with algorithms such as customer matching to assign a unique visitor ID to a customer.³ Then visitors’ behavior is tracked over the course of the experiment. This enables the website to track individuals’ viewing logs and purchases over many days. Users (both new and old) are randomly chosen to be in the control or one of the treatment groups, and their treatment status is fixed during the course of the experiment. When the company ran the field experiment, it wanted to test the recommenders with a small fraction of its visitors to reduce potential unwanted impact on the website. Therefore, it randomly allocated 10% of its visitors to each collaborative filtering treatment group. Upon clicking and viewing a particular item, the visitors are shown the appropriate recommender panel. Figure 2 shows a collaborative filtering recommender based on views (“People who viewed this item also viewed”) – which we call VBCF. Similarly, there is also a purchase-based collaborative filter – which we call PBCF. Users in the control group do not see this panel. At the end of the experiment, we have each consumer’s

³The A/B/n company outsources unique user identification to a specialized firm that utilizes a variety of data, such as IP, user-agent, log-in data, session ID, device, etc., to minimize duplicate and multi-device problems. However, it is always possible that a user on one device switches to another device and is not precisely identified. At the same time, almost every study based on randomized testing would be subject to this limitation of imperfect identification.

view logs and purchase logs at the item level. The algorithms were retrained every three days to propagate the influence of users' purchase history multiple times over the period of the experiment. About half of the users in our dataset were returning users and the remaining were new users on the website.

After the experiment was completed, we were concerned that the product information (e.g., review stars and numbers) displayed in the recommender panels might have influenced user click behavior, and eventually the impact of recommenders ⁴. Thus, we ran t-test on the average price, average review star, and average review number of all the viewed items in each treatments and control (user-number normalized) one against another and found that no comparisons were statistically significant.

People who viewed this item also viewed

The Dark Knight (Limited Edition)	Man Of Steel (Steelbook) (Blu-ray)	V For Vendetta (Limited Edition)	Batman Begins (Limited Edition)	Troy (Steelbook) (Bilingual)	I Am Legend (Limited Edition)
\$14.96	\$19.96	\$14.96	\$14.96	\$19.67	\$19.67
Add to cart	Add to cart	Add to cart	Add to cart	Add to cart	Add to cart

Figure 2: **Recommender Example:** Example of a recommender shown to a consumer. This consumer was in the treatment group of collaborative filtering based on views.

Figure 3 shows that most users in our dataset visit the retail website only once during our data collection. While many of the users view only one item, there are also many users who view multiple items in the retailer's catalog. Finally, the vast majority of buyers buy only one item during the two weeks of the experiment. There is also a reasonable tail of heavy users (15.1% of buyers buy 4 or more items in the 2-week period). Table 2 presents summary statistics by treatment group and Figure 4 visualizes the buying population and their activity levels. Together, they show that the recommenders help increase purchase activity by either driving an increase in the percentage of users who buy (Table 2, row 3) or the average number of purchases per buying user (Table 2, row

⁴We thank the anonymous reviewer for pointing this out.

4) or both.

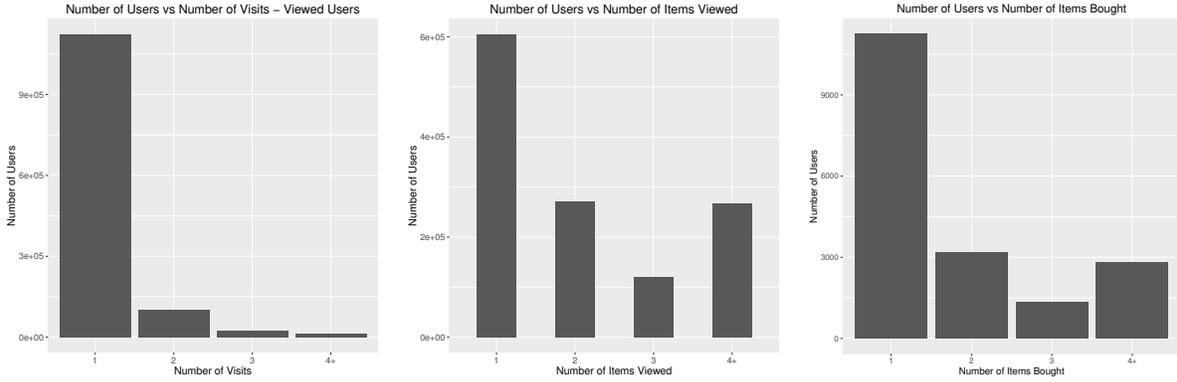


Figure 3: User Visit Frequency, Item # View Frequency, Item # Purchase Frequency

Description	Total	Control	Treated (PBCF)	Treated (VBCF)
Unique users who have viewed at least one item	1,138,238	876,301	137,167	124,770
Unique users who have made purchases	16,774	12,851	2,019	1,904
Percentage of users who buy	0.0147	0.0146	0.0147	0.0152
Number of purchases per buying user	2.460	2.444	2.560	2.461
Unique products viewed by users	82,290	76,820	45,611	44,899
Unique products purchased by users	12,976	10,910	2,825	2,646
Total number of item views	3,659,885	2,484,971	394,007	420,022
Total number of purchases	41,266	31,409	5,170	4,687
Total Purchase / Total Users	0.0362	0.0358	0.0376	0.0375

Table 2: Summary Statistics

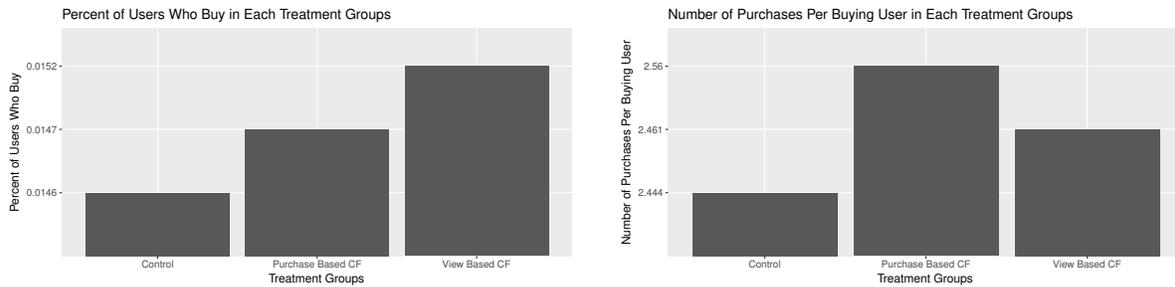


Figure 4: Purchase Activity of Users in Each Group

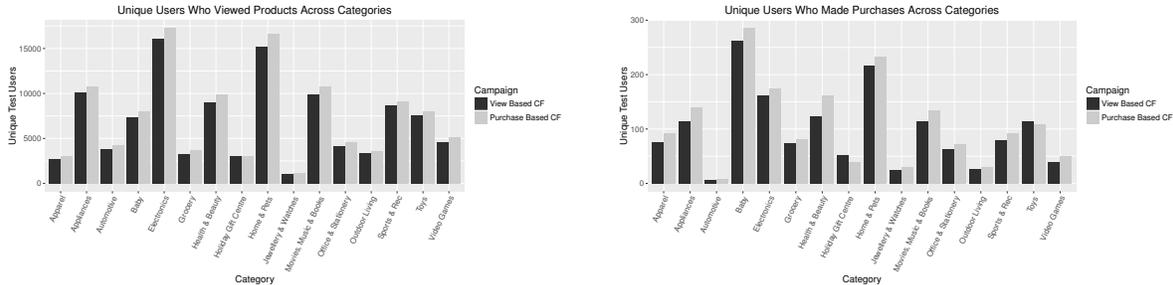


Figure 6: Number of Users Who Viewed or Purchased Products in Each Category.

curves by measuring sales at the item level, level-2 subcategory or at level-3 subcategory.⁵ Throughout our main analysis, we measure sales and compute Gini coefficients at the level-3 subcategory. Our results are qualitatively similar at other levels of analyses.⁶

4.2 Data Limitations

While our data are relatively clean, and causal effects are easier to extract compared to observational data, it has limitations – the most important being that we only implement two particular recommender system algorithms. We cannot fully control what the company is willing to implement and so we could not implement all well-known variants of recommender systems. To this end, we discuss how the results are generalizable. First, we used one of the most highly implemented item-based collaborative filters, as described in a seminal paper by Sarwar et al. (2001). We used both of the commonly used variants – view-based and purchase-based. Then, in implementing the algorithm, we also utilized a leading open-source platform, Apache Mahout. Additionally, our implementation of the CF used default parameter settings (e.g., how many days of historical data to use, how often to retrain the algorithm) that the A/B testing firm used across all clients. Finally, in talking to the A/B testing firm, we learned that out of several hundred firms that implement recommenders, only two utilized content-based recommenders. This translates to less than 1% of the firms. All of this suggests that our findings have wide applicability across many e-commerce companies.

An additional limitation is that we do not have purchase data beyond the experiment’s 2-week duration but it is possible that recommendations could have driven purchases after the experiment. Specifically, users may have viewed recommendations but taken some time to finalize their purchase

⁵Many items do not have a level 4 subcategory so we do not consider that possibility.

⁶The Gini coefficient changes are in fact more conservative at the subcategory level than at the item level, making the results more robust. The results are the same at the item level, with more group comparisons statistically significant. We present item-level analysis in Appendix B.

decisions. Given that we observe that recommendations drive an increase in purchases, it is possible that our estimate of that increase is conservative. On the other hand, it is also possible that the early response to recommendations might be enthusiastic but it might reduce over time. In this case, our estimates may be too aggressive. Nevertheless, our focus here is on sales diversity not volume. We do not expect later purchases to be systematically directed towards niche (or popular) products and we therefore believe that this limitation will not systematically bias our results in any one direction.

5 Results

In this section, we first analyze the effect of recommenders on aggregate and individual diversity on a pooled dataset. Next, we repeat our analysis by category for each of the major product categories on the retailer’s website. This is to test whether our results are qualitatively different based on the product category under consideration. Then, we conduct additional analysis to identify likely mechanisms that explain the changes in diversity that we observe. We conduct our analysis at both the item (SKU) level as well as the level-3 subcategory (not all items have a level-4 subcategory). In this section, we present our results at the subcategory level since it gives us more conservative estimates than the item level analyses. Appendix B shows the results at the item level, and our main findings are robust.

5.1 Pooled Aggregate and Individual Diversity Results

Figure 7 plots the Lorenz curve for aggregate sales at the firm. The unit of analysis is subcategory level 3 of the firm’s product catalog. For each subcategory, we calculate the totals views and sales associated with it. The Lorenz curve is then obtained based on these totals. The plots suggest a decrease in diversity of purchases for CF treatments relative to the control.

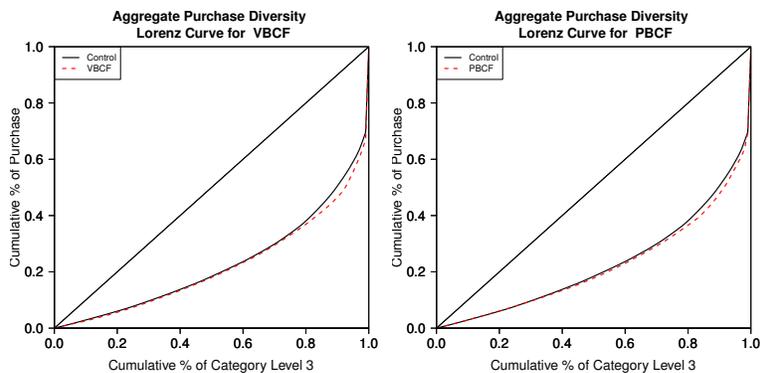


Figure 7: **Lorenz Curves for Subcategory Level 3 Purchased:** These Lorenz curves show that firm-level aggregate sales diversity is lower for both collaborative filtering algorithms relative to the control.

In Figure 8, we compute the Gini coefficients and test this more systematically for product views as well as purchases and at the aggregate and individual levels⁷. In the top panel, we plot the Gini coefficients for the control group and the two treatments groups. The figure on the left computes the Gini of views and the one on the right plots the Gini of purchases. Control group statistics are average of 1,000 random samplings to normalize for the number of user differences. Each test statistic bar of treated bars are also marked with the p-value stars if the difference from the control group is statistically significant. Stars are labeled as follows: '.' = p-value < 0.1, '*' = p-value < 0.05, '**' = p-value < 0.01, '***' = p-value < 0.001. For both views and purchases, the control groups' Gini coefficients are lower than that of either VBCF or PBCF. The differences between the control group and the treated groups are significant at p-value < 0.001. Both types of collaborative filtering are causing consumers to view and purchase less variety of products in aggregate. Finally, the PBCF's Gini coefficients were higher than those of the VBCF for both the views and purchases, suggesting that the concentration effect was stronger for PBCF. The bottom two graphs of the figure plot average view and purchase diversity at the individual level. Interestingly, we do not observe a concentration bias of the CFs at the individual level. In fact, directionally, the treated groups' Gini coefficients are lower than the control even though the differences were not statistically significant. Thus, it appears that while aggregate sales and view diversity decrease under CFs, individuals may or may not be exploring less.

⁷For individual Gini measure, we standardize the length of the purchase or view vector to include all genres/items in the site for all users. This ensures the measures are comparable across different groups.

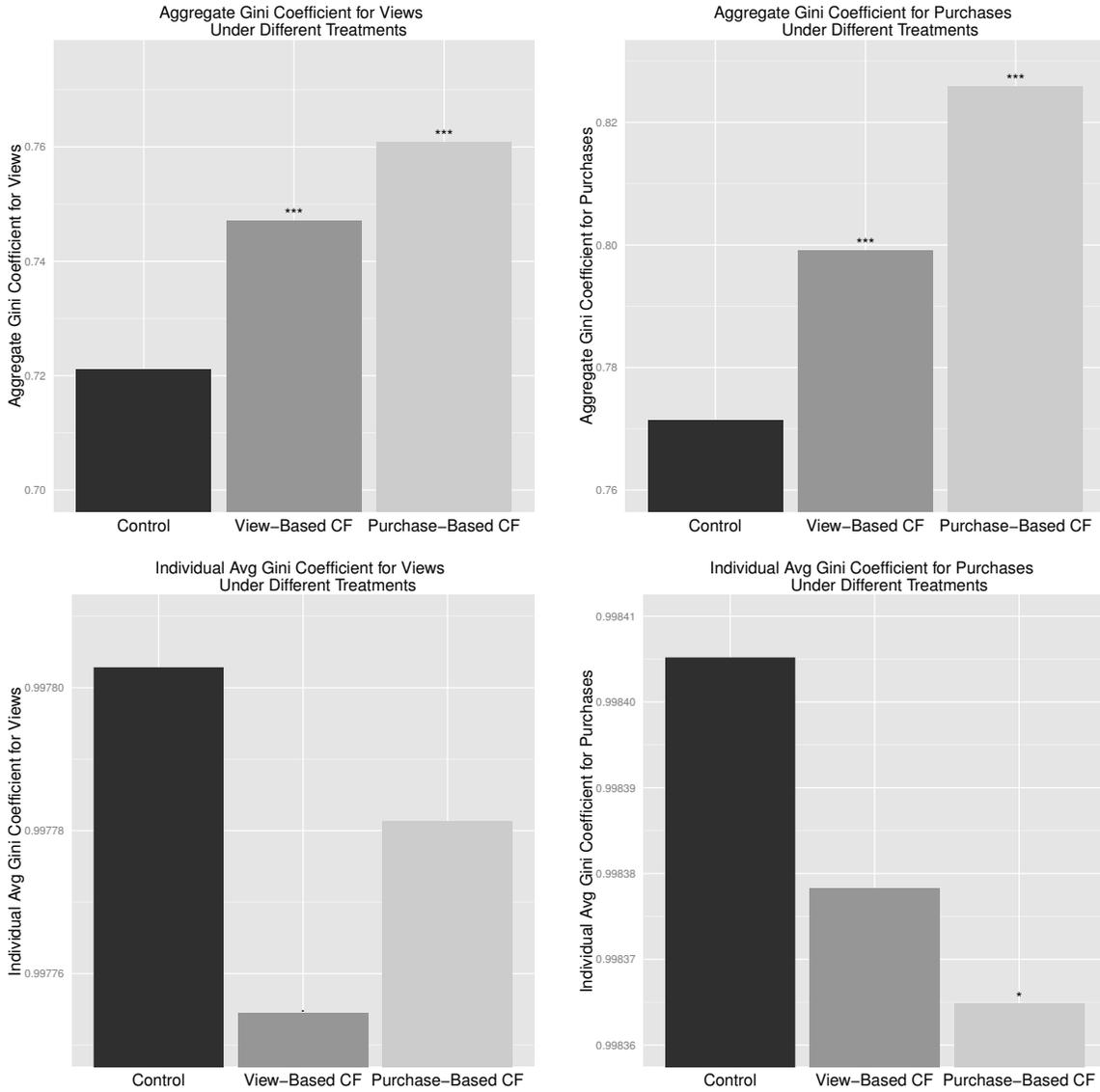


Figure 8: **Aggregate and Individual Gini Coefficient Differences Under Different Treatments.**

We next turn to analyze these results by product category to determine whether heterogeneity in effects across a diverse set of product categories might explain the findings.

5.2 Diversity By Product Category

Our partner firm has sixteen high-level categories as listed in Section 4.1. We now conduct our analysis separately for each of these sixteen categories. This allows us to investigate heterogeneity in results across product categories. Further, because recommended products are usually from the same Level-1 category as a focal product, there is additional merit in analyzing the results by

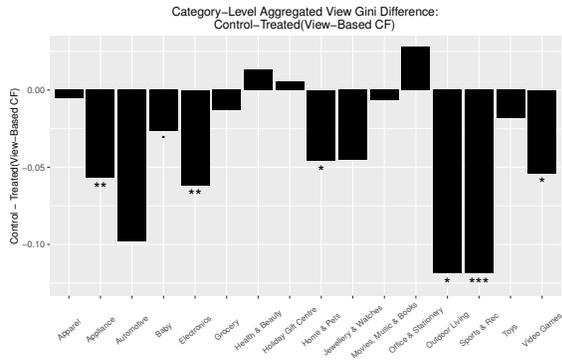
category. For space and clarity, we present our results in graphic formats here. Please see Appendix A for result numbers in table formats.

Figure 9 presents the results by category. The first three graphs show the change in aggregate view diversity for 1) control – view-based collaborative filtering (VBCF), 2) control – purchase-based collaborative filtering (PBCF), 3) VBCF-PBCF. The next three graphs show the results for change in aggregate purchase diversity in the same order. The Y-axis is the Gini difference. Each test statistic bar is denoted by the p-value stars associated with the value. Stars represent the statistical significance as before: '.'= p-value <0.1, '*'= p-value <0.05, '**'= p-value <0.01, '***'= p-value <0.001.

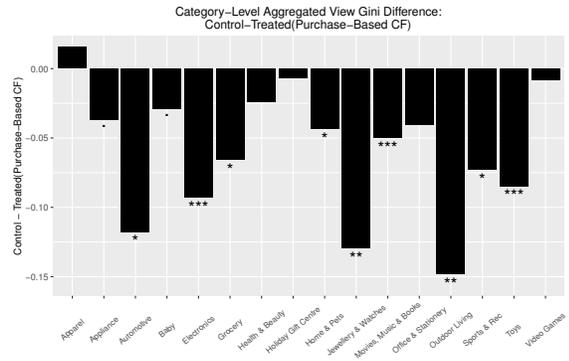
In both view diversity graphs at the top of Figure 9, the *control – treated* Gini differences are negative, if they are statistically significant. This shows that both recommenders have a concentration bias and that consumers, as a whole, are viewing and exploring fewer subcategories under the recommenders. We also repeated the analysis at the item level (these are presented in Appendix B). The results are even more pronounced at the item level, with Gini differences being negative and statistically significant for every category. Figure 9C compares the two different treatment groups (VBCF - PBCF). While the differences are usually negative, all the categories are not statistically significant at the p-value < 0.05 level.

The same pattern emerges for changes in aggregate purchase diversity in Figure 9D-F. When significant, the test statistics are always negative for the *control – treated* (for both VBCF and PBCF), which suggest that consumers are buying less variety of subcategories within all categories under the influence of recommenders. Once again, the analysis in Appendix B shows that the results are more pronounced at the item level and uniformly negative and statistically significant for all categories. In summary, the use of recommenders is once again associated with a decrease in the aggregate view and purchase diversity across multiple product categories.⁸

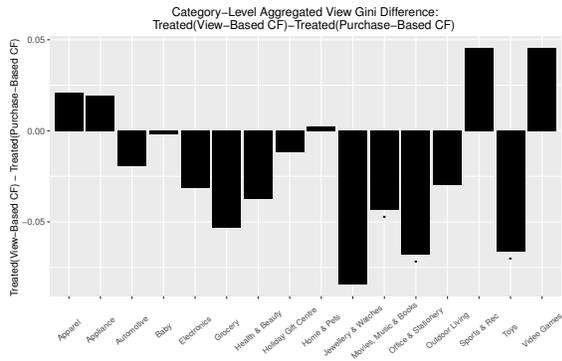
⁸One could ask whether the concentration bias is more pronounced in certain product categories. Magnitude of the concentration bias is indeed different across categories. However, a simple linear regression with Gini coefficient differences as the Y-variable and category level characteristics as the X-variable did not produce any statistically significant results. With sixteen categories at one firm, we do not have enough cross-sectional variation to explain whether the magnitude of the impact is systematically explained by category-specific attributes.



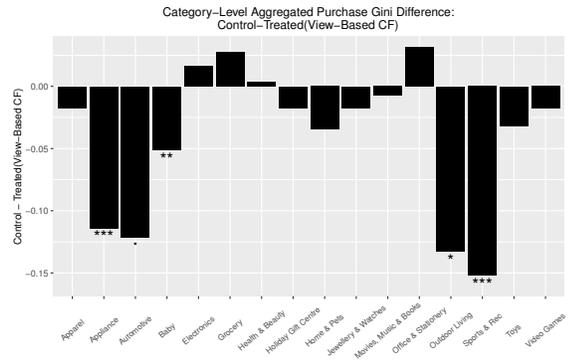
(a) Aggregate View Diversity Gini Difference: Control-VBCF



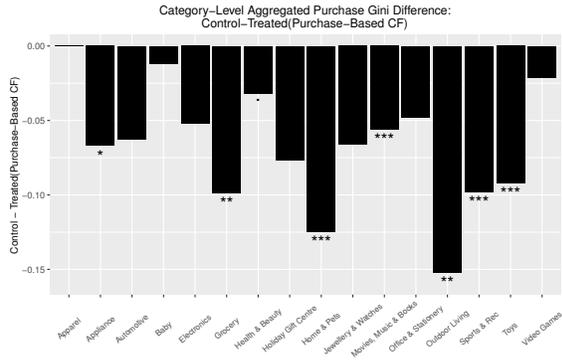
(b) Aggregate View Diversity Gini Difference: Control-PBCF



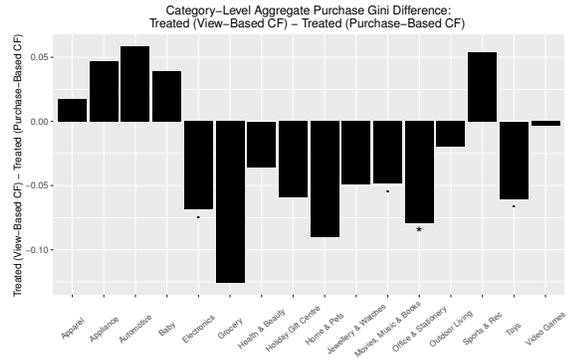
(c) Aggregate View Diversity Gini Difference: VBCF-PBCF



(d) Aggregate Purchase Diversity Gini Difference: Control-VBCF



(e) Aggregate Purchase Diversity Gini Difference: Control-PBCF



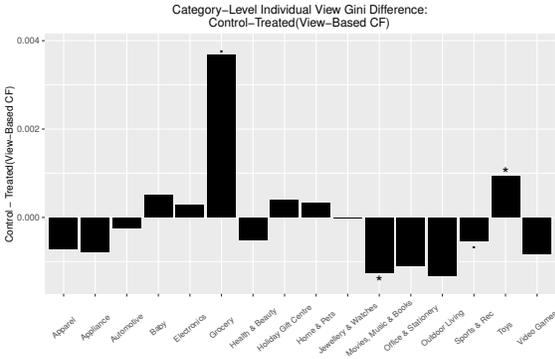
(f) Aggregate Purchase Diversity Gini Difference: VBCF-PBCF

Figure 9: Aggregate Diversity Results Across Cross Categories. Stars represent the statistical significance. ‘.’= p-value <0.1, ‘*’= p-value <0.05, ‘**’= p-value <0.01, ‘***’= p-value <0.001.

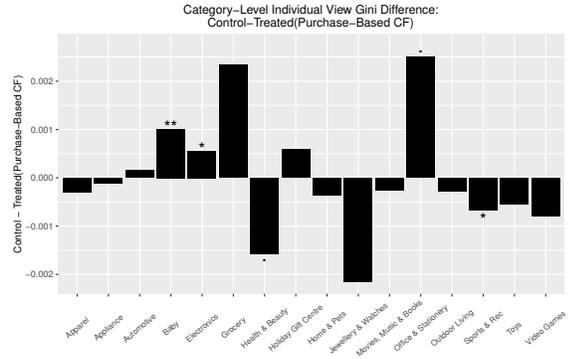
We next repeat the analyses at the individual level. This is done as we did in the previous section except that the analysis is done separately for each of the sixteen product categories. Figure 10 presents the results. While aggregate results show a clear concentration bias, there isn’t compelling

evidence of it at the individual level. Results are often not statistically significant and, when they are, the results are not directionally consistent. The lack of statistical significance could be due to the fact that many users in our dataset only purchase one product, leading to a lack of in-depth individual level data, unlike in aggregate cases where the impact was more pronounced and easier to measure. A simple solution would have been to run the experiment for a longer period of time, but this was not feasible due to practical constraints at the collaborating company.

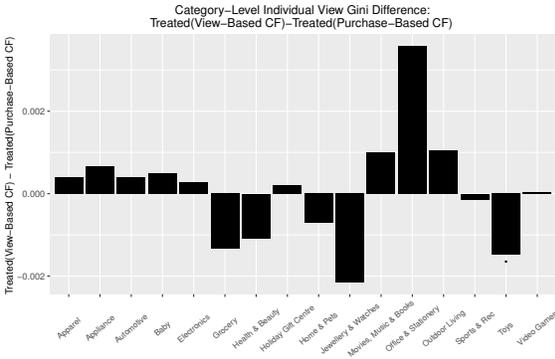
In summary, separately analyzing each of the main categories shows the same set of results that we found in our firm-level analysis. At the aggregate level, there is strong evidence of concentration bias. At the individual level, it is not clear whether user views or purchases are becoming less diverse. In the next two sections, we investigate product networks and analyze impact on absolute sales of niche items in order to explain the differences at the aggregate and individual levels.



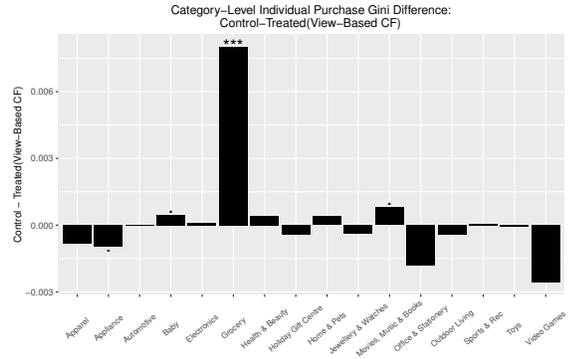
(a) Individual Average View Diversity Gini Difference: Control-VBCF



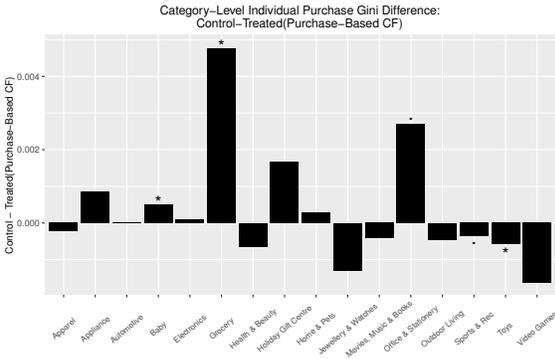
(b) Individual Average View Diversity Gini Difference: Control-PBCF



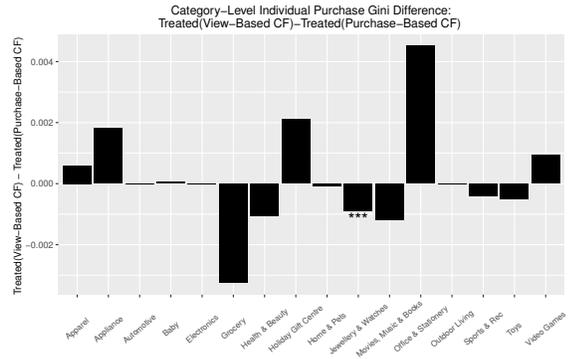
(c) Individual Average View Diversity Gini Difference: VBCF-PBCF



(d) Individual Average Purchase Diversity Gini Difference: Control-VBCF



(e) Individual Average Purchase Diversity Gini Difference: Control-PBCF



(f) Individual Average Purchase Diversity Gini Difference: VBCF-PBCF

Figure 10: Individual Average Diversity Results Across Cross Categories. Stars represent the statistical significance. ‘.’= p-value <0.1, ‘*’= p-value <0.05, ‘**’= p-value <0.01, ‘***’= p-value <0.001.

5.3 The Source of Diversity Shift: Co-Purchase Network Analyses

In this section, we visualize our results and explore how recommenders cause the shift in aggregate diversity. We use movies, one of the most-well known product types, to illustrate and visualize our

findings. Movies are a level-2 subcategory in our retailer’s catalog (within “Movies, Music & Books”). The main reason we choose movies is that the level-3 subcategory within it is movie genre which is easy to understand and interpret. Later in the section, we demonstrate that the observations we share here are not unique to movies.

In total, there were 71,122 users who viewed 15,064 unique movie-related items (total 151,709 views) and 993 users who actually ended up buying 1,478 (total 1,933) unique movie-related products, which include DVDs, Blu-Ray discs, etc. Our analysis of view and sales diversity for movies showed that aggregate diversity decreased under recommenders while individual increased (both results were statistically significant). These results are omitted due to space constraints.

We construct co-purchase networks for the control group and PBCF normalized by sample size (graphs and tables for VBCF are omitted due to space constraints, but the observations are qualitatively similar but sometimes less pronounced in magnitude). Each node in the graph represents a level-3 subcategory (movie genre), and the size of the node is proportional to the *percent* of overall sales that went to the genre. An edge between two nodes indicates that there was a user who purchased from these genres, and the thickness of the edge is proportional to the number of such users who exist. Thus, the relative sizes of the nodes convey the extent to which sales were (un)evenly distributed at the aggregate level, and the edges convey the extent to which individuals explored content in diverse genres. Figure 11 compares the network graphs for PBCF and control group side by side.

When visually comparing the two graphs, we note the following:

1. Relative size of the nodes shows that the majority of purchases by the control is distributed across a few genres: action, drama, and comedy. In the Purchase-based CF, however, comedy is much bigger than the rest, indicating that purchases were more concentrated in comedy. This might be because comedy titles were recommended more often by the algorithms or because consumers are more willing to explore and trust the recommender for comedy.
2. The Purchase-based CF graph is more well-connected (i.e., denser) than the control. This indicates that there are more users who are buying across genres in the Purchase-based CF group, or, more specifically, there is greater individual cross-buying behavior. The connectedness of the Purchase-based CF graph reflects the increase in individual diversity that we noted previously. Individual users may be exploring more genres while sales may be simultaneously

concentrated in a few genres at the aggregate level.

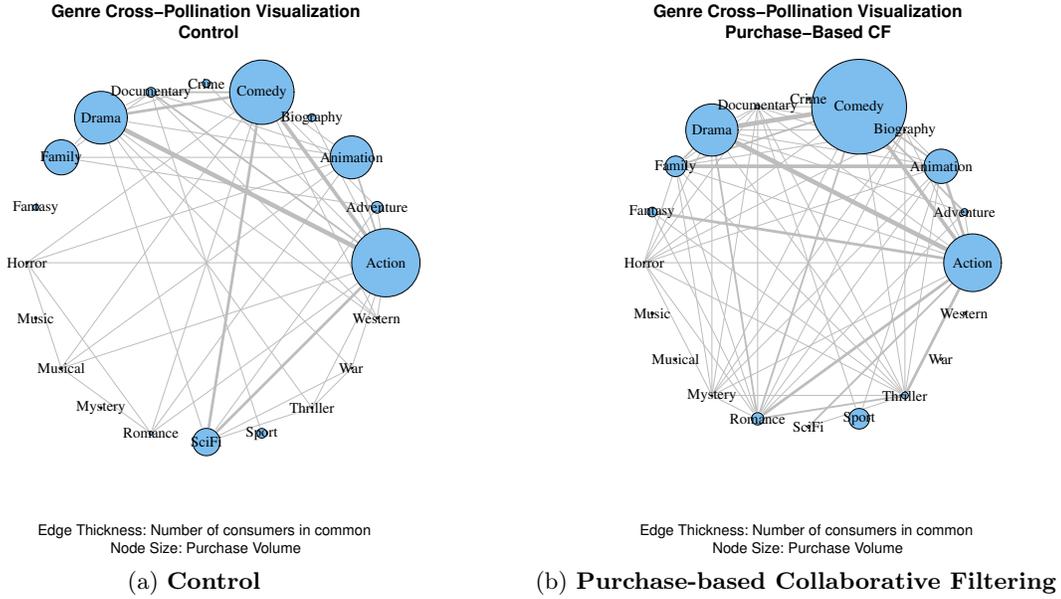


Figure 11: Co-Purchase Network Graphs of Genre Purchases Under Control and Purchase-Based Collaborative Filtering.

In summary, the Purchase-based collaborative filtering algorithm shifts users to buy a few top genres at the aggregate level while increasing individual diversity through a cross-buying behavior that is aided by a few “pathway” genres. To formally test these differences, we use a permutation test to evaluate the market share of the top genres in each graph (market share of top N nodes). Table 4 shows the difference between the market-share statistics for the CFs and control group as well as the corresponding p-values obtained via permutation tests. We replicated the analysis for top {1, 5, 10} genres. We see a clear shift to the top genres with the CFs. Under the Purchase-based CF, the top genre, comedy, took 11% more of the market share compared to the top genre in the control group, action. VBCF demonstrates a similar pattern.

	Top 1 Genre	Top 5 Genres	Top 10 Genres
	(CF Market Share - Control Market Share) P-value		
Purchase-based CF	0.11 0.04	0.10 0.04	0.08 0.004

Table 4: Permutation Test Results for Co-Purchase Network Comparisons: Purchase-Based CF vs. Control

To ensure that our observations with co-purchase networks generalize beyond movies, we repli-

cate the above analysis across all level-3 subcategories in the firm’s catalog. We again construct co-purchase network in which each node in the graph represents a level-3 subcategory in the catalog (e.g., for category Electronics, “Speakers” and “Desktop Computers” would be level-3 subcategories). Size of a node indicates the market share of the node. An edge between two nodes again indicates that there were consumers who purchased from these two sub-categories. We report two main measures associated with the networks. First, we compute the market share of the top N nodes in the network as we did above. In addition, we compute the average degree of top-N nodes. The degree of a node is simply the number of edges of that node. The average degree of top-N nodes is obtained by taking the degrees of each of the nodes that are among the Top-N nodes and averaging them. The market shares indicate aggregate diversity and the average degrees indicate the extent to which users cross-purchased across subcategories.

Table 5 presents the results for marketshare shift in the treated group vs the control group. The table presents the marketshare of top {1, 5, first quartile, second quartile} subcategories for the control group, PBCF group, *control – treated*, and p-value associated with the differences. For all rows, the recommender has increased the market share of the top nodes. The differences are statistically significant as well. This suggests that PBCF is increasing the concentration bias by increasing market share of the top subcategories. Table 6 presents the results for the average degree differences. The result suggests that PBCF is associated with a higher average degree for the top {1, 5, first quartile, second quartile} subcategories, meaning that even though users gravitate towards the top subcategories, they are also cross-purchasing across subcategories.

Taken together, the results in this section imply that consumers are cross-purchasing more, but at the same time their explorations are highly correlated due to the nature of the CF (recall that a CF uses data on other users’ views and purchases to recommend items). Due to this highly correlated exploration among users, the market share for the top selling products keeps increasing, creating a rich get richer concentration bias.

Top N	Control Stat	PBCF Stat	Control - Treat	P-value
1	0.022	0.056	-0.034	0.024
5	0.100	0.145	-0.045	0.016
First Quartile	0.815	0.840	-0.025	0.016
Second Quartile	0.961	0.968	-0.006	0.134

Table 5: Market Share of Top N Subcategory Permutation Test Results

Top N	Control Stat	PBCF Stat	Control - Treat	P-value
1	43	61	-18	<0.001
5	36.2	51.8	-15.6	<0.001
First Quartile	10.27	13.90	-3.63	0.160
Second Quartile	5.16	7.46	-2.3	<0.001

Table 6: Average Degree of Top N Subcategory Permutation Test Results

5.4 CFs' Impact on Absolute Sales of Niche Item

Our analyses thus far reveal that CFs cause marketshare concentration; i.e., popular items obtain increased marketshare while *niche items lose marketshare*. The market share and Gini coefficient analyses ultimately reflect relative gains for popular versus niche items in the seller's catalog. An important question is what happens at an absolute level. Table 7 suggests that recommenders increase total views and sales for the retailer. Specifically, the PBCF caused a 0.5% lift in views and a 5% lift in purchases and the VBCF showed an 11% lift in views and a 0.8% lift in purchases for our retailer (permutation tests reveal that that these changes are statistically significant for PBCF purchase and VBCF views)⁹. It is worth noting that VBCFs are more effective in increasing views while PBCFs are more effective in increasing purchases, relatively speaking. Given this increase in total volume, the question is whether niche items lose out in absolute terms as well. If so, these items are strictly worse off due to the use of CFs. On the other hand, if absolute views and purchases of niche items increase but their market share goes down, it reveals a more nuanced view of the impact of CFs. This would suggest that all kinds of products benefit from CFs but mainstream producers benefit more so than niche producers.

⁹Large effect sizes were observed for subcategories such as movies where PBCF cause a 25% lift in views and a 35% lift in the number of items purchased over the control group (no recommender). VBCF also showed a 3% lift in views and a 9% lift in the number of items purchased.

	Control Stat	PBCF Stat (P-value of Treat- Control)	VBCF Stat (P-value of Treat- Control)	PBCF Stat Increase % from Control	VBCF Stat Increase % from Control
Average # of Items Viewed	10.32	10.38 (0.778)	11.50 (<0.001)	0.5%	11%
Average # of Items Purchased	2.44	2.56 (0.034)	2.46 (0.210)	5%	0.8%

Table 7: **Sales Volume Effect of Recommenders Using Permutation Test:** The statistics represent average number of items purchased or viewed over 1000 iterations.

Figure 12 is a modified Lorenz curve which shows the absolute purchase count for all items (instead of the marketshare) with items rank ordered by sales volume. The charts are normalized by the number of users in the treated group to allow direct comparison of treated and control groups. X-axis orders the level-3 subcategories from the least popular to the most popular. Y-axis show cumulative absolute purchase counts. For both charts, the treated dotted lines are clearly above the control in all support. This suggests that under the influence of CFs, all items, regardless of popularity, obtain increased absolute sales.

To formalize this finding, we ran t-tests comparing treated and control cumulative purchase counts for bottom {25th, 50th, all} percentile ranked products. The alternative hypothesis is “true difference in means is greater than zero.” Rejecting the null would mean that the treated group’s cumulative purchase count is statistically significantly greater than that for the control group. We show the results in table format in Table 8. In all ranges, for both PBCF and VBCF, we reject the null and the t-test results corroborate intuition from the modified Lorenz curve in Figure 7.

Thus, although CFs show concentration bias in favor of popular items, they also manage to increase absolute sales for niche items. The increase in sales is more pronounced for the popular items resulting in market share concentration bias. We obtain qualitatively similar results for view counts (as opposed to purchase counts) and for an analysis at the item level (as opposed to level-3 subcategory).

Percentile	Test	T-Statistic	P-Value
0.25	PBCF vs Control	21.73560655	< 0.001
0.5	PBCF vs Control	11.32599092	< 0.001
1	PBCF vs Control	2.09396685	0.018
0.25	VBCF vs Control	20.74169952	< 0.001
0.5	VBCF vs Control	14.04737241	< 0.001
1	VBCF vs Control	1.69023882	0.045

Table 8: T-test Results Comparing Treated Group Cumulative Purchase Count is Greater than That of Control Group’s for Bottom {25th, 50th, all} Percentile.

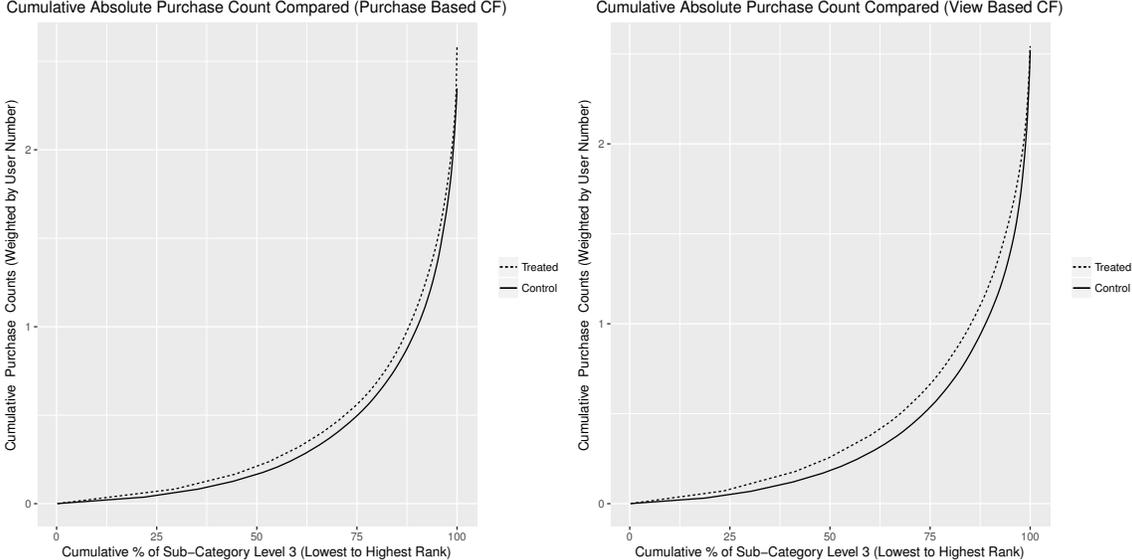


Figure 12: **Cumulative Absolute Purchase Counts Compared:** The absolute cumulative purchase counts are normalized by dividing by each group’s unique number of users.

5.5 Robustness Checks

We ran several different robustness checks to ensure that our results are not sensitive to certain choices we made for data analysis. They include the following:

- R1** Our analysis thus far is based on treatment groups smaller than control group sizes. We replicated the analysis by randomly sampling a fixed number of users in each group, ex-ante before the permutation test, so that each group has an equal number of consumers.
- R2** We removed consumers who are outliers in terms of number of views and purchased items. Specifically, we replicated the analysis excluding those users whose views or purchases exceeded three standard deviations from the mean.

R3 We replicated the analysis by using 1 or 0 (binary variable) for genre or items viewed (or purchased) instead of actual counts. This allows us to explore the notion of diversity from a perspective of number of unique items or genres viewed/purchased as opposed to proportion of sales.

R4 We replicated the analysis only for consumers who bought more than one item.

In all of these robustness checks, the findings are qualitatively similar to our main result.

6 Discussion and Conclusions

Recommenders and personalization technologies are fast taking over nearly every aspect of consumer interaction on the web. Their use spans the purchase of physical products (books, DVDs, clothing, electronics, etc.), digital media (movies, news), and even online services such as dating and peer-to-peer lending. Despite their ubiquity, we still have much to learn about how different recommender algorithms influence markets and society.

Our study contributes to emerging literature on the impact of personalization technologies by studying the impact of recommender algorithms on sales diversity with an expansive dataset from a field experiment run on the website of a top North American retailer. The dataset spans across 16 categories, 82,290 SKUs, and 1,138,238 users. Table 9 summarizes our results on the *sales diversity impact* of recommenders across a wide variety of different product categories. We have shown that collaborative filtering recommenders cause aggregate view and sales diversity to decrease, pushing consumers as a whole to explore and purchase less variety of products across all categories sold on this site. The result holds true whether diversity is analyzed at the item level or the subcategory level. It highlights a potential drawback of widely used collaborative filtering designs in terms of their ability to aid discovery of truly niche items. Interestingly, however, the individual average Gini coefficients were not significantly and consistently influenced by the recommender systems. We found that for some product categories, individual-level diversity may even increase. We investigated the source of diversity shift by analyzing co-purchase networks and comparing relative versus absolute purchase volume. Co-purchase analysis shows that individual users do explore new products and genres. However, their explorations are correlated, resulting in concentration at the aggregate level. Further, we find that the absolute volume of views and purchases of niche items, in fact, increase. But the increase in views and purchases is far greater for popular products such that market share

concentration increases. This suggests a nuanced view of the impact of recommenders. To the extent that scale helps, for example, by lowering per-unit costs or allowing producers to invest in better equipment, this shift helps niche producers. To the extent that market share helps, for example, by allowing producers to have greater consumer mindshare, this shift hurts niche producers.

Results
CF decreases aggregate product view and sales diversity across categories consistently
CF doesn't seem to influence individual view and purchase diversity consistently across product categories
CF enables greater exploration among individuals via cross-selling products. But the exploration is correlated across individuals
However, niche items do not necessarily lose because CF increases absolute sales volume for all type of items.

Table 9: **Result Summary**

These results have significant managerial relevance. As the amount of consumer data available to firms grows exponentially, many retailers have aggressively adopted data mining and personalization technologies without a deep understanding of how different designs may contribute toward (or deter) broader strategic goals. For example, a firm interested in exposing consumers to a broader assortment of products may prefer a different design from another simply interested in maximizing sales. To the extent that a firm is interested in pushing its “back catalog,” it may seek to augment traditional collaborative filtering algorithms so that it is possible to identify relevant products with limited historical data (past views/purchases) and/or increase diversity, serendipity, or novelty of the recommended products using techniques from the extant literature (e.g., Adamopoulos and Tuzhilin (2015); Adomavicius and Kwon (2014); Oh et al. (2011)).

For producers of niche products, it is important to recognize that co-purchase networks can be an important driver of consumer purchase decisions. As a result, producers might benefit from marketing strategies that seek to strategically place one’s product in the co-purchase network of other relevant products. For example, a publisher might initially discount its book (or target its advertising) for readers of another closely related, but popular, title. This can help ensure that the product enters the co-purchase network of the popular title. This is particularly relevant for producers of niche products. One of the main promises of the Internet is its ability to support niche

product markets. The Internet has reduced entry costs for producers in several markets such as books (Brynjolfsson and Smith, 2000), music (Graham et al., 2005), and movies (Zhu, 2001) and also made it feasible for retailers to carry niche products due to lower stocking costs (Brynjolfsson et al., 2006). On the demand side, search engines and recommenders lower consumer search costs. However, that alone might not be sufficient. Producers might need to exert additional marketing effort to enable product discovery by relevant customer segments.

We conclude by discussing some limitations of our study and opportunities for future work. First, our study focused on the two most commonly used collaborative filtering designs. It appears from our study that purchase-based collaborative filters have a greater impact than CFs based on views. This may be because: 1) the algorithm based on purchases might simply be better at delivering the best fit products, 2) consumers might be more influenced by the “purchased also purchased” signal than the “view also viewed” signal, 3) or both. However, our dataset lacks granular information on the actual products that were recommended and consumer response to each individual recommendation that would be needed to quantify which of these effects is in play. This would be a worthwhile extension of our work. Second, it is worth investigating the impact of other recommender designs, such as content-based and social-network-based recommenders. Third, a valuable extension to our work will be the studies that develop behavioral theories on how and why people react differently to different recommender systems and signals. Lab studies can be highly valuable in this regard. Finally, we found that the magnitude of the changes in diversity is different for different product categories. Future work that documents whether there are systematic patterns in terms of how the change in diversity depends on product category in consideration.

References

- Adamopoulos, P. and A. Tuzhilin: 2015, ‘On Unexpectedness in Recommender Systems : Or How to Better Expect the Unexpected’. *ACM Transactions on Intelligent Systems Technology* **5**(4), 54:1—54:32.
- Adomavicius, G. and Y. Kwon: 2014, ‘Optimization-Based Approaches For Maximizing Aggregate Recommendation Diversity’. *INFORMS Journal on Computing* **26**(2), 351–369.
- Adomavicius, G. and A. Tuzhilin: 2005, ‘Toward The Next Generation Of Recommender Systems: A Survey Of The State-Of-The-Art And Possible Extensions’.
- Anderson, C.: 2008, *The Long Tail: Why The Future Of Business Is Selling Less Of More*. Hyperion Books.

- Borle, S., P. Boatwright, J. B. Kadane, J. C. Nunes, and S. Galit: 2005, ‘The effect of product assortment changes on customer retention’. *Marketing science* **24**(4), 616–622.
- Breese, J. S., D. Heckerman, and C. Kadie: 1998, ‘Empirical analysis of predictive algorithms for collaborative filtering’. In: *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. pp. 43–52.
- Brynjolfsson, E., Y. J. Hu, and D. Simester: 2011, ‘Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales’. *Management Science* **57**(8), 1373–1386.
- Brynjolfsson, E., Y. J. Hu, and M. D. Smith: 2006, ‘From Niches to Riches : Anatomy of the Long Tail From Niches to Riches : Anatomy of the Long Tail’. *MIT Sloan Management Review* **47**(4), 67.
- Brynjolfsson, E. and M. D. Smith: 2000, ‘Frictionless commerce? A comparison of Internet and conventional retailers’. *Management science* **46**(4), 563–585.
- Celma, Ò. and P. Cano: 2008, ‘From Hits To Niches? Or How Popular Artists Can Bias Music Recommendation And Discovery’. *2nd Workshop on LargeScale Recommender Systems and the Netflix Prize Competition ACM KDD 1*(August 2015), 1–8.
- Cooke, A. D., H. Sujana, M. Sujana, and B. A. Weitz: 2002, ‘Marketing the unfamiliar: the role of context and item-specific information in electronic agent recommendations’. *Journal of marketing research* **39**(4), 488–497.
- De, P., Y. Hu, and M. S. Rahman: 2010, ‘Technology Usage and Online Sales: An Empirical Study’. *Management Science* **56**(11), 1930–1945.
- Dias, M. B., D. Locher, M. Li, W. El-Deredy, and P. J. G. Lisboa: 2008, ‘The Value of Personalised Recommender Systems to e-Business: A Case Study’. In: *Proceedings of the 2008 ACM Conference on Recommender Systems*. pp. 291–294.
- Elberse, A.: 2008, ‘Should you invest in the long tail?’. *Harvard business review* **86**(7/8), 88.
- Fisher, M. L.: 2003, ‘What is the right supply chain for your product’. *Operations management: critical perspectives on business and management* **4**, 73.
- Fleder, D. and K. Hosanagar: 2009, ‘Blockbuster Culture’s Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity’. *Management Science* **55**(5), 697–712.
- Gallino, S., A. Moreno, and I. Stamatopoulos: 2014, ‘Channel Integration, Sales Dispersion, and Inventory Management’. *Sales Dispersion, and Inventory Management (September 10, 2014)*.
- Good, P.: 2005, *Permutation, Parametric And Bootstrap Tests Of Hypotheses*. Springer.
- Graham, G., G. J. Lewis, G. Graham, and G. Hardaker: 2005, ‘Evaluating the impact of the internet on barriers to entry in the music industry’. *Supply Chain Management: An International Journal* **10**(5), 349–356.
- Häubl, G. and V. Trifts: 2000, ‘Consumer decision making in online shopping environments: The effects of interactive decision aids’. *Marketing science* **19**(1), 4–21.
- Hinz, O. and J. Eckert: 2010, ‘The Impact of Search and Recommendation Systems on Sales in Electronic Commerce’. *Bise* **2**(2), 67–77.
- Hosanagar, K., D. M. Fleder, D. Lee, and A. Buja: 2014, ‘Will the Global Village Fracture into Tribes: Recommender Systems and their Effects on Consumers’. *Management Science* **60**(4), 805–823.

- Jannach, D. and K. Hegelich: 2009, 'A Case Study On The Effectiveness Of Recommendations In The Mobile Internet'. In: *Proceedings of the third ACM conference on Recommender systems - RecSys '09*. p. 205.
- Jannach, D., L. Lerche, F. Gedikli, and G. Bonnin: 2013, 'What Recommenders Recommend - An Analysis Of Accuracy, Popularity And Sales Diversity Effects'. *User Modeling, Adaptation, and Personalization - Lecture Notes in Computer Science* **7899**, 25–37.
- Jiang, B., K. Jerath, and K. Srinivasan: 2011, 'Firm strategies in the mid tail of platform-based retailing'. *Marketing Science* **30**(5), 757–775.
- Jiang, H., X. Qi, and H. Sun: 2014, 'Choice-based recommender systems: a unified approach to achieving relevancy and diversity'. *Operations Research* **62**(5), 973–993.
- Kok, A. G., M. L. Fisher, and R. Vaidyanathan: 2008, 'Assortment planning: Review of literature and industry practice'. In: *Retail supply chain management*. Springer, pp. 99–153.
- Lamere, P. and S. Green: 2008, 'Project Aura: Recommendation For The Rest Of Us'.
- Lin, Z., K. Y. Goh, and C. S. Heng: 2015, 'The demand effects of product recommendation networks: an empirical analysis of network diversity and stability'.
- Mooney, R. J. and L. Roy: 2000, 'Content-based book recommending using learning for text categorization'. In: *Proceedings of the fifth ACM conference on Digital libraries*. pp. 195–204.
- Oestreicher-Singer, G. and A. Sundararajan: 2012, 'Recommendation Networks And The Long Tail Of Electronic Commerce'. *MIS Quarterly* **36**(1), 65–84.
- Oh, J., S. Park, H. Yu, M. Song, and S.-T. Park: 2011, 'Novel Recommendation Based on Personal Popularity Tendency'. In: *Data Mining (ICDM), 2011 IEEE 11th International Conference on*. pp. 507–516.
- Pathak, B., R. Garfinkel, R. D. Gopal, R. Venkatesan, and F. Yin: 2010, 'Empirical analysis of the impact of recommender systems on sales'. *Journal of Management Information Systems* **27**(2), 159–188.
- Resnick, P. and H. R. Varian: 1997, *Recommender Systems*, Vol. 40. Communications of the ACM.
- Sarwar, B., G. Karypis, J. Konstan, and J. Riedl: 2001, 'Item-based collaborative filtering recommendation algorithms'. In: *Proceedings of the 10th international conference on World Wide Web*. pp. 285–295.
- Schafer, J. B., J. Konstan, and J. Riedl: 1999, 'Recommender systems in e-commerce'. In: *Proceedings of the 1st ACM conference on Electronic commerce*. pp. 158–166.
- Senecal, S. and J. Nantel: 2004, 'The influence of online product recommendations on consumers online choices'. *Journal of Retailing* **80**(2), 159–169.
- Tam, K. Y. and S. Y. Ho: 2005, 'Web personalization as a persuasion strategy: An elaboration likelihood model perspective'. *Information Systems Research* **16**(3), 271–291.
- Thompson, C.: 2008, 'If You Liked This, You're Sure To Love That'.
- Van Herpen, E. and R. Pieters: 2002, 'The variety of an assortment: An extension to the attribute-based approach'. *Marketing Science* **21**(3), 331–341.
- Walter, F. E., S. Battiston, M. Yildirim, and F. Schweitzer: 2012, 'Moving recommender systems from on-line commerce to retail stores'. *Information Systems and E-Business Management* **10**(3), 367–393.

- Wu, L.-L., Y.-J. Joung, and T.-E. Chiang: 2011, 'Recommendation Systems and Sales Concentration: The Moderating Effects of Consumers' Product Awareness and Acceptance to Recommendations'. In: *System Sciences (HICSS), 2011 44th Hawaii International Conference on*. pp. 1–10.
- Zhou, R., S. Khemmarat, and L. Gao: 2010, 'The Impact Of Youtube Recommendation System On Video Views'. *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*.
- Zhu, K.: 2001, 'Internet-based distribution of digital videos: the economic impacts of digitization on the motion picture industry'. *Electronic Markets* **11**(4), 273–280.

A Result Table on Gini Differences By Category

In this appendix, we presents the table version of our cross-category results shown in the Results section. The columns mean the following:

- Category: Level 1 category name.
- Test Statistics: We show aggregate purchase diversity comparisons (agg_purchase), individual purchase diversity (ind_purchase), aggregate view diversity (agg_view), and individual view diversity (ind_view).
- Control, PBCF, VBCF: Respective group statistics value.
- Comparison: Group 1 minus Group 2. e.g., “control - PBCF” means we calculate the group difference by taking the control statistics minus the PBCF statistics.
- Difference: Group comparison value.
- p_value: P value associated with the difference via permutation test.

As presented in the Results section, the aggregate view and purchase diversity are consistently decreased for both VBCF and PBCF across categories and many are statistically significant. Individual diversity shift was not consistent however.

Category	Test Statistics	Control	PBCF	VBCF	Comparison	Difference	p_value
Apparel	agg_purchase	0.742492893	0.74244898	0.759854935	control-PBCF	4.39E-05	0.974
Appliances	agg_purchase	0.65093985	0.717904575	0.764912281	control-PBCF	-0.06696473	0.036
Automotive	agg_purchase	0.790322581	0.852941176	0.911764706	control-PBCF	-0.0626186	0.354
Baby	agg_purchase	0.782104769	0.794120062	0.833156361	control-PBCF	-0.01201529	0.522
Electronics	agg_purchase	0.723958862	0.776175549	0.707635107	control-PBCF	-0.05221669	0.922
Grocery	agg_purchase	0.55767739	0.656170704	0.530434783	control-PBCF	-0.09849331	0.008
Health & Beauty	agg_purchase	0.766843281	0.798559027	0.763032806	control-PBCF	-0.03171575	0.078
Holiday_Gift_Centre	agg_purchase	0.533446712	0.610169492	0.551020408	control-PBCF	-0.07672278	0.232
Home & Pets	agg_purchase	0.676595767	0.801456595	0.711047633	control-PBCF	-0.12486083	0
Jewellery & Watches	agg_purchase	0.736314363	0.802702703	0.753763441	control-PBCF	-0.06638834	0.202
Movies, Music & Books	agg_purchase	0.845687661	0.901813206	0.853298714	control-PBCF	-0.05612554	0
Office & Stationery	agg_purchase	0.749047619	0.796973317	0.717862166	control-PBCF	-0.0479257	0.198
Outdoor Living	agg_purchase	0.568	0.72	0.700606061	control-PBCF	-0.152	0.004
Sports & Rec	agg_purchase	0.633957592	0.732082077	0.786093217	control-PBCF	-0.09812449	0
Toys	agg_purchase	0.713961615	0.806133444	0.745844875	control-PBCF	-0.09217183	0
Video Games	agg_purchase	0.805786445	0.826844262	0.823369565	control-PBCF	-0.02105782	0.218
Apparel	agg_purchase	0.742492893	0.74244898	0.759854935	control-VBCF	-0.01736204	0.514
Appliances	agg_purchase	0.65093985	0.717904575	0.764912281	control-VBCF	-0.11397243	0
Automotive	agg_purchase	0.790322581	0.852941176	0.911764706	control-VBCF	-0.12144213	0.086
Baby	agg_purchase	0.782104769	0.794120062	0.833156361	control-VBCF	-0.05105159	0.004
Electronics	agg_purchase	0.723958862	0.776175549	0.707635107	control-VBCF	0.016323756	0.956
Grocery	agg_purchase	0.55767739	0.656170704	0.530434783	control-VBCF	0.027242607	0.548
Health & Beauty	agg_purchase	0.766843281	0.798559027	0.763032806	control-VBCF	0.003810475	0.81
Holiday_Gift_Centre	agg_purchase	0.533446712	0.610169492	0.551020408	control-VBCF	-0.0175737	0.766
Home & Pets	agg_purchase	0.676595767	0.801456595	0.711047633	control-VBCF	-0.03445187	0.23
Jewellery & Watches	agg_purchase	0.736314363	0.802702703	0.753763441	control-VBCF	-0.01744908	0.7
Movies, Music & Books	agg_purchase	0.845687661	0.901813206	0.853298714	control-VBCF	-0.00761105	0.584
Office & Stationery	agg_purchase	0.749047619	0.796973317	0.717862166	control-VBCF	0.031185453	0.478
Outdoor Living	agg_purchase	0.568	0.72	0.700606061	control-VBCF	-0.13260606	0.016
Sports & Rec	agg_purchase	0.633957592	0.732082077	0.786093217	control-VBCF	-0.15213563	0
Toys	agg_purchase	0.713961615	0.806133444	0.745844875	control-VBCF	-0.03188326	0.136
Video Games	agg_purchase	0.805786445	0.826844262	0.823369565	control-VBCF	-0.01758312	0.318
Apparel	agg_purchase	0.742492893	0.74244898	0.759854935	VBCF-PBCF	0.017405956	0.626
Appliances	agg_purchase	0.65093985	0.717904575	0.764912281	VBCF-PBCF	0.047007706	0.358
Automotive	agg_purchase	0.790322581	0.852941176	0.911764706	VBCF-PBCF	0.058823529	0.416
Baby	agg_purchase	0.782104769	0.794120062	0.833156361	VBCF-PBCF	0.039036299	0.216
Electronics	agg_purchase	0.723958862	0.776175549	0.707635107	VBCF-PBCF	-0.06854044	0.062
Grocery	agg_purchase	0.55767739	0.656170704	0.530434783	VBCF-PBCF	-0.12573592	0.1
Health & Beauty	agg_purchase	0.766843281	0.798559027	0.763032806	VBCF-PBCF	-0.03552622	0.136
Holiday_Gift_Centre	agg_purchase	0.533446712	0.610169492	0.551020408	VBCF-PBCF	-0.05914908	0.428
Home & Pets	agg_purchase	0.676595767	0.801456595	0.711047633	VBCF-PBCF	-0.09040896	0.25
Jewellery & Watches	agg_purchase	0.736314363	0.802702703	0.753763441	VBCF-PBCF	-0.04893926	0.378
Movies, Music & Books	agg_purchase	0.845687661	0.901813206	0.853298714	VBCF-PBCF	-0.04851449	0.092
Office & Stationery	agg_purchase	0.749047619	0.796973317	0.717862166	VBCF-PBCF	-0.07911115	0.046
Outdoor Living	agg_purchase	0.568	0.72	0.700606061	VBCF-PBCF	-0.01939394	0.692
Sports & Rec	agg_purchase	0.633957592	0.732082077	0.786093217	VBCF-PBCF	0.05401114	0.168
Toys	agg_purchase	0.713961615	0.806133444	0.745844875	VBCF-PBCF	-0.06028857	0.088
Video Games	agg_purchase	0.805786445	0.826844262	0.823369565	VBCF-PBCF	-0.0034747	0.91

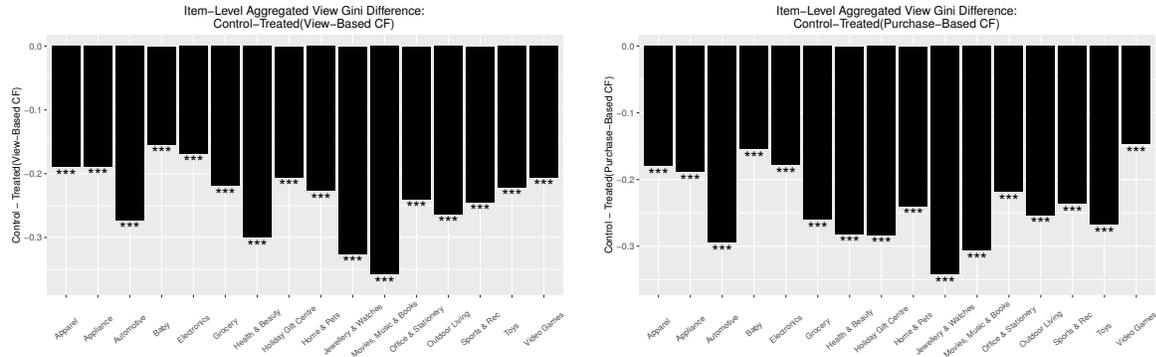
Apparel	agg_view	0.70470558	0.689177961	0.70997796	control-PBCF	0.015527619	0.512
Appliances	agg_view	0.628149274	0.665336985	0.684877851	control-PBCF	-0.03718771	0.074
Automotive	agg_view	0.693928908	0.81167691	0.792035398	control-PBCF	-0.117748	0.046
Baby	agg_view	0.660235279	0.68895574	0.686816562	control-PBCF	-0.02872046	0.056
Electronics	agg_view	0.614937002	0.707784622	0.676530478	control-PBCF	-0.09284762	0
Grocery	agg_view	0.573678533	0.639583333	0.586614853	control-PBCF	-0.0659048	0.034
Health & Beauty	agg_view	0.743683043	0.767950764	0.730693811	control-PBCF	-0.02426772	0.136
Holiday_Gift_Centre	agg_view	0.496135831	0.502623907	0.490822027	control-PBCF	-0.00648808	0.82
Home & Pets	agg_view	0.589562301	0.632848576	0.635187828	control-PBCF	-0.04328628	0.014
Jewellery & Watches	agg_view	0.664842454	0.794417863	0.71015873	control-PBCF	-0.12957541	0.004
Movies, Music & Books	agg_view	0.840538434	0.890596231	0.847106269	control-PBCF	-0.0500578	0
Office & Stationery	agg_view	0.672166728	0.712441783	0.644338624	control-PBCF	-0.04027506	0.16
Outdoor Living	agg_view	0.528659254	0.676872428	0.647239819	control-PBCF	-0.14821317	0.004
Sports & Rec	agg_view	0.657455156	0.730321309	0.776045184	control-PBCF	-0.07286615	0.014
Toys	agg_view	0.704591415	0.789210213	0.722713092	control-PBCF	-0.0846188	0
Video Games	agg_view	0.780719214	0.789093625	0.834558824	control-PBCF	-0.00837441	0.796
Apparel	agg_view	0.70470558	0.689177961	0.70997796	control-VBCF	-0.00527238	0.824
Appliances	agg_view	0.628149274	0.665336985	0.684877851	control-VBCF	-0.05672858	0.008
Automotive	agg_view	0.693928908	0.81167691	0.792035398	control-VBCF	-0.09810649	0.16
Baby	agg_view	0.660235279	0.68895574	0.686816562	control-VBCF	-0.02658128	0.064
Electronics	agg_view	0.614937002	0.707784622	0.676530478	control-VBCF	-0.06159348	0.002
Grocery	agg_view	0.573678533	0.639583333	0.586614853	control-VBCF	-0.01293632	0.708
Health & Beauty	agg_view	0.743683043	0.767950764	0.730693811	control-VBCF	0.012989232	0.396
Holiday_Gift_Centre	agg_view	0.496135831	0.502623907	0.490822027	control-VBCF	0.005313804	0.932
Home & Pets	agg_view	0.589562301	0.632848576	0.635187828	control-VBCF	-0.04562553	0.012
Jewellery & Watches	agg_view	0.664842454	0.794417863	0.71015873	control-VBCF	-0.04531628	0.424
Movies, Music & Books	agg_view	0.840538434	0.890596231	0.847106269	control-VBCF	-0.00656784	0.638
Office & Stationery	agg_view	0.672166728	0.712441783	0.644338624	control-VBCF	0.027828104	0.424
Outdoor Living	agg_view	0.528659254	0.676872428	0.647239819	control-VBCF	-0.11858057	0.034
Sports & Rec	agg_view	0.657455156	0.730321309	0.776045184	control-VBCF	-0.11859003	0
Toys	agg_view	0.704591415	0.789210213	0.722713092	control-VBCF	-0.01812168	0.388
Video Games	agg_view	0.780719214	0.789093625	0.834558824	control-VBCF	-0.05383961	0.022
Apparel	agg_view	0.70470558	0.689177961	0.70997796	VBCF-PBCF	0.020799999	0.652
Appliances	agg_view	0.628149274	0.665336985	0.684877851	VBCF-PBCF	0.019540866	0.608
Automotive	agg_view	0.693928908	0.81167691	0.792035398	VBCF-PBCF	-0.01964151	0.748
Baby	agg_view	0.660235279	0.68895574	0.686816562	VBCF-PBCF	-0.00213918	0.932
Electronics	agg_view	0.614937002	0.707784622	0.676530478	VBCF-PBCF	-0.03125414	0.264
Grocery	agg_view	0.573678533	0.639583333	0.586614853	VBCF-PBCF	-0.05296848	0.34
Health & Beauty	agg_view	0.743683043	0.767950764	0.730693811	VBCF-PBCF	-0.03725695	0.204
Holiday_Gift_Centre	agg_view	0.496135831	0.502623907	0.490822027	VBCF-PBCF	-0.01180188	0.782
Home & Pets	agg_view	0.589562301	0.632848576	0.635187828	VBCF-PBCF	0.002339252	0.91
Jewellery & Watches	agg_view	0.664842454	0.794417863	0.71015873	VBCF-PBCF	-0.08425913	0.168
Movies, Music & Books	agg_view	0.840538434	0.890596231	0.847106269	VBCF-PBCF	-0.04348996	0.076
Office & Stationery	agg_view	0.672166728	0.712441783	0.644338624	VBCF-PBCF	-0.06810316	0.07
Outdoor Living	agg_view	0.528659254	0.676872428	0.647239819	VBCF-PBCF	-0.02963261	0.652
Sports & Rec	agg_view	0.657455156	0.730321309	0.776045184	VBCF-PBCF	0.045723875	0.19
Toys	agg_view	0.704591415	0.789210213	0.722713092	VBCF-PBCF	-0.06649712	0.05
Video Games	agg_view	0.780719214	0.789093625	0.834558824	VBCF-PBCF	0.045465198	0.278
Apparel	ind_purchase	0.974435091	0.974662912	0.975250627	control-PBCF	-0.00022782	0.708

Appliances	ind_purchase	0.971514363	0.970655055	0.972479024	control-PBCF	0.000859308	0.17
Automotive	ind_purchase	0.970588235	0.970588235	0.970588235	control-PBCF	0	0
Baby	ind_purchase	0.987752319	0.987237665	0.987299745	control-PBCF	0.000514655	0.026
Electronics	ind_purchase	0.98808604	0.987983281	0.987989366	control-PBCF	0.000102759	0.516
Grocery	ind_purchase	0.955722393	0.950954727	0.947706922	control-PBCF	0.004767665	0.012
Health & Beauty	ind_purchase	0.969565279	0.970207683	0.969155629	control-PBCF	-0.0006424	0.434
Holiday_Gift_Centre	ind_purchase	0.856372322	0.854700855	0.85680593	control-PBCF	0.001671468	0.272
Home & Pets	ind_purchase	0.987336528	0.987039929	0.98694221	control-PBCF	0.000296599	0.354
Jewellery & Watches	ind_purchase	0.965380117	0.966666667	0.965777778	control-PBCF	-0.00128655	0.126
Movies, Music & Books	ind_purchase	0.97042314	0.970831404	0.969633484	control-PBCF	-0.00040826	0.356
Office & Stationery	ind_purchase	0.970491595	0.967792339	0.972304717	control-PBCF	0.002699256	0.096
Outdoor Living	ind_purchase	0.959553429	0.96	0.96	control-PBCF	-0.00044657	0.294
Sports & Rec	ind_purchase	0.988420304	0.988774666	0.988361849	control-PBCF	-0.00035436	0.054
Toys	ind_purchase	0.988352483	0.988926445	0.988405797	control-PBCF	-0.00057396	0.038
Video Games	ind_purchase	0.930120717	0.93175	0.932692308	control-PBCF	-0.00162928	0.496
Apparel	ind_purchase	0.974435091	0.974662912	0.975250627	control-VBCF	-0.00081554	0.114
Appliances	ind_purchase	0.971514363	0.970655055	0.972479024	control-VBCF	-0.00096466	0.082
Automotive	ind_purchase	0.970588235	0.970588235	0.970588235	control-VBCF	0	0
Baby	ind_purchase	0.987752319	0.987237665	0.987299745	control-VBCF	0.000452575	0.058
Electronics	ind_purchase	0.98808604	0.987983281	0.987989366	control-VBCF	9.67E-05	0.46
Grocery	ind_purchase	0.955722393	0.950954727	0.947706922	control-VBCF	0.008015471	0
Health & Beauty	ind_purchase	0.969565279	0.970207683	0.969155629	control-VBCF	0.00040965	0.614
Holiday_Gift_Centre	ind_purchase	0.856372322	0.854700855	0.85680593	control-VBCF	-0.00043361	0.474
Home & Pets	ind_purchase	0.987336528	0.987039929	0.98694221	control-VBCF	0.000394318	0.218
Jewellery & Watches	ind_purchase	0.965380117	0.966666667	0.965777778	control-VBCF	-0.00039766	0.79
Movies, Music & Books	ind_purchase	0.97042314	0.970831404	0.969633484	control-VBCF	0.000789656	0.096
Office & Stationery	ind_purchase	0.970491595	0.967792339	0.972304717	control-VBCF	-0.00181312	0.284
Outdoor Living	ind_purchase	0.959553429	0.96	0.96	control-VBCF	-0.00044657	0.248
Sports & Rec	ind_purchase	0.988420304	0.988774666	0.988361849	control-VBCF	5.85E-05	0.74
Toys	ind_purchase	0.988352483	0.988926445	0.988405797	control-VBCF	-5.33E-05	0.842
Video Games	ind_purchase	0.930120717	0.93175	0.932692308	control-VBCF	-0.00257159	0.312
Apparel	ind_purchase	0.974435091	0.974662912	0.975250627	VBCF-PBCF	0.000587715	0.416
Appliances	ind_purchase	0.971514363	0.970655055	0.972479024	VBCF-PBCF	0.001823969	0.168
Automotive	ind_purchase	0.970588235	0.970588235	0.970588235	VBCF-PBCF	0	0
Baby	ind_purchase	0.987752319	0.987237665	0.987299745	VBCF-PBCF	6.21E-05	0.89
Electronics	ind_purchase	0.98808604	0.987983281	0.987989366	VBCF-PBCF	6.08E-06	0.958
Grocery	ind_purchase	0.955722393	0.950954727	0.947706922	VBCF-PBCF	-0.00324781	0.536
Health & Beauty	ind_purchase	0.969565279	0.970207683	0.969155629	VBCF-PBCF	-0.00105205	0.576
Holiday_Gift_Centre	ind_purchase	0.856372322	0.854700855	0.85680593	VBCF-PBCF	0.002105075	0.506
Home & Pets	ind_purchase	0.987336528	0.987039929	0.98694221	VBCF-PBCF	-9.77E-05	0.954
Jewellery & Watches	ind_purchase	0.965380117	0.966666667	0.965777778	VBCF-PBCF	-0.00088889	0
Movies, Music & Books	ind_purchase	0.97042314	0.970831404	0.969633484	VBCF-PBCF	-0.00119792	0.282
Office & Stationery	ind_purchase	0.970491595	0.967792339	0.972304717	VBCF-PBCF	0.004512378	0.102
Outdoor Living	ind_purchase	0.959553429	0.96	0.96	VBCF-PBCF	0	0
Sports & Rec	ind_purchase	0.988420304	0.988774666	0.988361849	VBCF-PBCF	-0.00041282	0.29
Toys	ind_purchase	0.988352483	0.988926445	0.988405797	VBCF-PBCF	-0.00052065	0.21

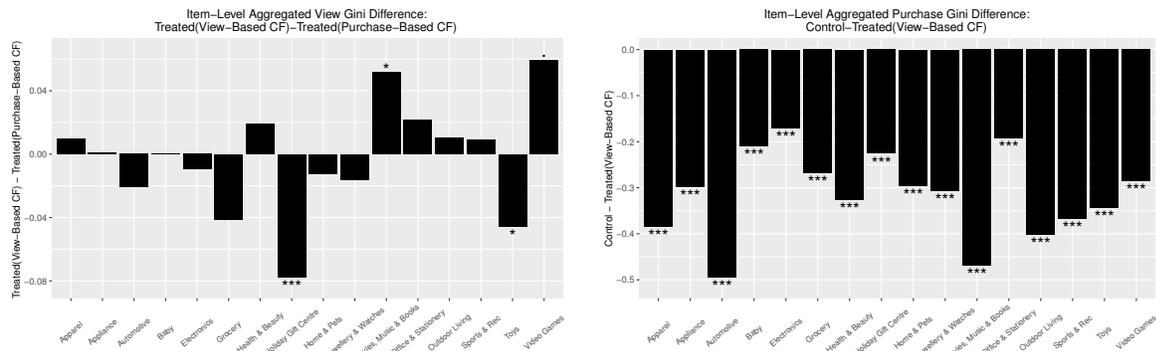
Video Games	ind_purchase	0.930120717	0.93175	0.932692308	VBCF-PBCF	0.000942308	0.806
Apparel	ind_view	0.971740369	0.97204259	0.972443421	control-PBCF	-0.00030222	0.642
Appliances	ind_view	0.970147614	0.970267002	0.970924991	control-PBCF	-0.00011939	0.788
Automotive	ind_view	0.969607309	0.969457014	0.969844728	control-PBCF	0.000150295	0.828
Baby	ind_view	0.98463077	0.983621496	0.984126414	control-PBCF	0.001009274	0.006
Electronics	ind_view	0.986452226	0.985894495	0.986181879	control-PBCF	0.000557731	0.014
Grocery	ind_view	0.952793865	0.950461722	0.949123587	control-PBCF	0.002332144	0.238
Health & Beauty	ind_view	0.968003496	0.969586503	0.968508087	control-PBCF	-0.00158301	0.064
Holiday_Gift_Centre	ind_view	0.853680122	0.853087671	0.853293018	control-PBCF	0.000592451	0.684
Home & Pets	ind_view	0.985599897	0.985964681	0.985265054	control-PBCF	-0.00036478	0.22
Jewellery & Watches	ind_view	0.959884468	0.96204059	0.959889601	control-PBCF	-0.00215612	0.28
Movies, Music & Books	ind_view	0.967642259	0.967891777	0.968901632	control-PBCF	-0.00024952	0.684
Office & Stationery	ind_view	0.969409982	0.96691202	0.9705058	control-PBCF	0.002497963	0.096
Outdoor Living	ind_view	0.957710723	0.957981265	0.959034014	control-PBCF	-0.00027054	0.81
Sports & Rec	ind_view	0.987029771	0.987699753	0.987555735	control-PBCF	-0.00066998	0.012
Toys	ind_view	0.986359163	0.986900867	0.985424714	control-PBCF	-0.0005417	0.164
Video Games	ind_view	0.927977465	0.928776341	0.928804243	control-PBCF	-0.00079888	0.72
Apparel	ind_view	0.971740369	0.97204259	0.972443421	control-VBCF	-0.00070305	0.37
Appliances	ind_view	0.970147614	0.970267002	0.970924991	control-VBCF	-0.00077738	0.108
Automotive	ind_view	0.969607309	0.969457014	0.969844728	control-VBCF	-0.00023742	0.766
Baby	ind_view	0.98463077	0.983621496	0.984126414	control-VBCF	0.000504355	0.174
Electronics	ind_view	0.986452226	0.985894495	0.986181879	control-VBCF	0.000270347	0.204
Grocery	ind_view	0.952793865	0.950461722	0.949123587	control-VBCF	0.003670278	0.054
Health & Beauty	ind_view	0.968003496	0.969586503	0.968508087	control-VBCF	-0.00050459	0.548
Holiday_Gift_Centre	ind_view	0.853680122	0.853087671	0.853293018	control-VBCF	0.000387104	0.814
Home & Pets	ind_view	0.985599897	0.985964681	0.985265054	control-VBCF	0.000334843	0.29
Jewellery & Watches	ind_view	0.959884468	0.96204059	0.959889601	control-VBCF	-5.13E-06	0.954
Movies, Music & Books	ind_view	0.967642259	0.967891777	0.968901632	control-VBCF	-0.00125937	0.024
Office & Stationery	ind_view	0.969409982	0.96691202	0.9705058	control-VBCF	-0.00109582	0.474
Outdoor Living	ind_view	0.957710723	0.957981265	0.959034014	control-VBCF	-0.00132329	0.17
Sports & Rec	ind_view	0.987029771	0.987699753	0.987555735	control-VBCF	-0.00052597	0.07
Toys	ind_view	0.986359163	0.986900867	0.985424714	control-VBCF	0.000934449	0.022
Video Games	ind_view	0.927977465	0.928776341	0.928804243	control-VBCF	-0.00082678	0.682
Apparel	ind_view	0.971740369	0.97204259	0.972443421	VBCF-PBCF	0.000400831	0.748
Appliances	ind_view	0.970147614	0.970267002	0.970924991	VBCF-PBCF	0.000657989	0.498
Automotive	ind_view	0.969607309	0.969457014	0.969844728	VBCF-PBCF	0.000387714	0.624
Baby	ind_view	0.98463077	0.983621496	0.984126414	VBCF-PBCF	0.000504919	0.448
Electronics	ind_view	0.986452226	0.985894495	0.986181879	VBCF-PBCF	0.000287384	0.534
Grocery	ind_view	0.952793865	0.950461722	0.949123587	VBCF-PBCF	-0.00133814	0.746
Health & Beauty	ind_view	0.968003496	0.969586503	0.968508087	VBCF-PBCF	-0.00107842	0.454
Holiday_Gift_Centre	ind_view	0.853680122	0.853087671	0.853293018	VBCF-PBCF	0.000205348	0.934
Home & Pets	ind_view	0.985599897	0.985964681	0.985265054	VBCF-PBCF	-0.00069963	0.23
Jewellery & Watches	ind_view	0.959884468	0.96204059	0.959889601	VBCF-PBCF	-0.00215099	0.514
Movies, Music & Books	ind_view	0.967642259	0.967891777	0.968901632	VBCF-PBCF	0.001009856	0.338
Office & Stationery	ind_view	0.969409982	0.96691202	0.9705058	VBCF-PBCF	0.003593781	0.134
Outdoor Living	ind_view	0.957710723	0.957981265	0.959034014	VBCF-PBCF	0.001052749	0.466
Sports & Rec	ind_view	0.987029771	0.987699753	0.987555735	VBCF-PBCF	-0.00014402	0.804
Toys	ind_view	0.986359163	0.986900867	0.985424714	VBCF-PBCF	-0.00147615	0.092
Video Games	ind_view	0.927977465	0.928776341	0.928804243	VBCF-PBCF	2.79E-05	0.998

B Aggregate & Individual Diversity Results at the Item Level

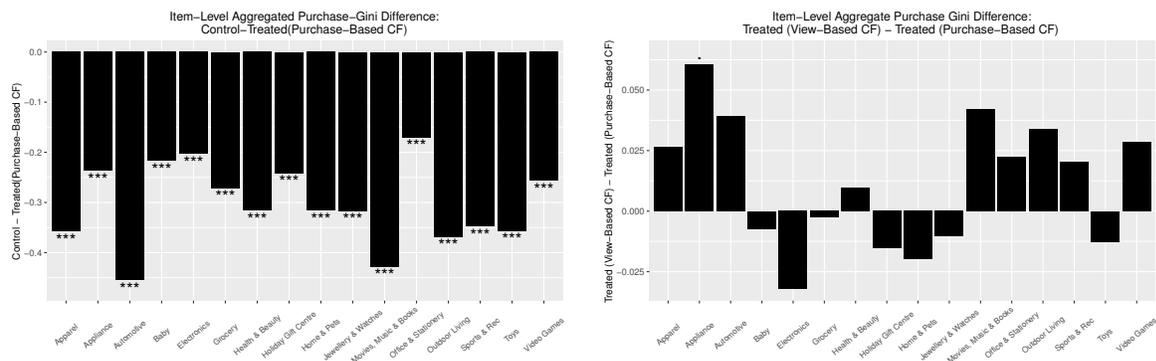
We provide the result calculated on the individual item-level in the same format as Section 5, in which we have provided the results for subcategory level.



(a) Aggregate View Diversity Gini Difference: Control-VBCF (b) Aggregate View Diversity Gini Difference: Control-PBCF

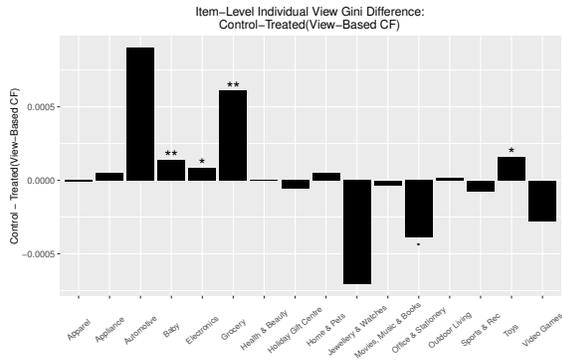


(c) Aggregate View Diversity Gini Difference: VBCF-PBCF (d) Aggregate Purchase Diversity Gini Difference: Control-VBCF

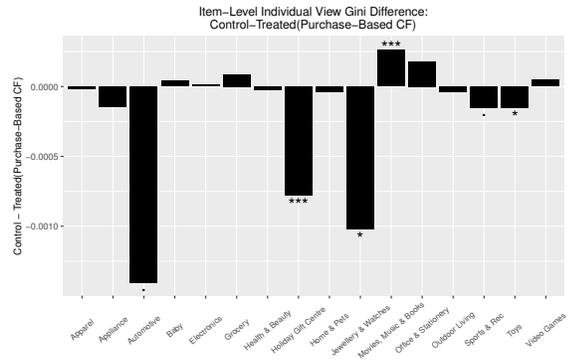


(e) Aggregate Purchase Diversity Gini Difference: Control-PBCF (f) Aggregate Purchase Diversity Gini Difference: VBCF-PBCF

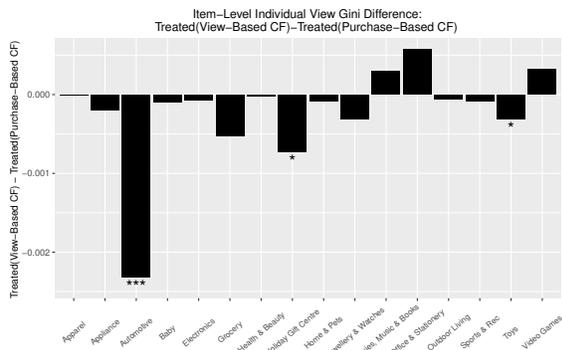
Figure 13: Aggregate Diversity Results At the Item Level. Stars represent the statistical significance. '.' = p-value < 0.1, '*' = p-value < 0.05, '**' = p-value < 0.01, '***' = p-value < 0.001.



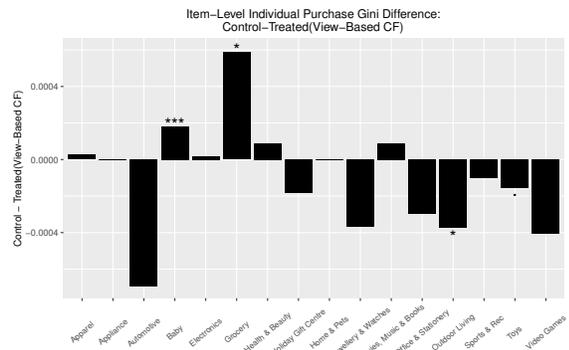
(a) Individual Average View Diversity Gini Difference: Control-VBCF



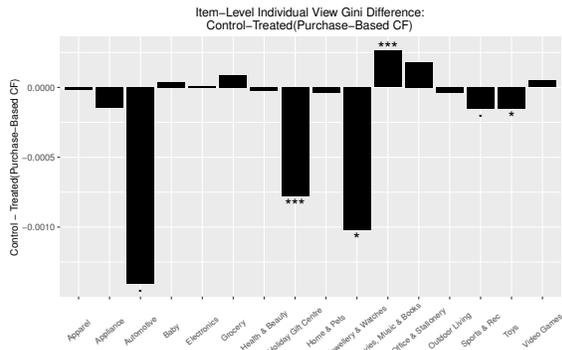
(b) Individual Average View Diversity Gini Difference: Control-PBCF



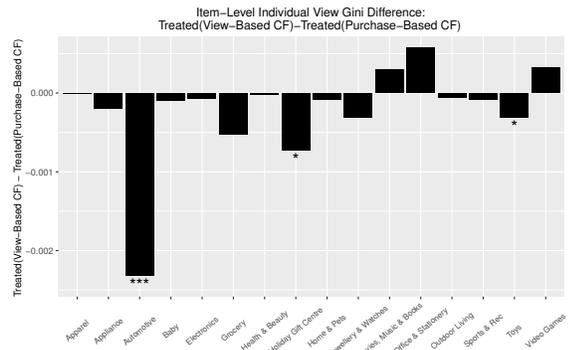
(c) Individual Average View Diversity Gini Difference: VBCF-PBCF



(d) Individual Average Purchase Diversity Gini Difference: Control-VBCF



(e) Individual Average Purchase Diversity Gini Difference: Control-PBCF



(f) Individual Average Purchase Diversity Gini Difference: VBCF-PBCF

Figure 14: Individual Average Diversity Results At the Item Level. Stars represent the statistical significance. ‘.’= p-value <0.1, ‘*’= p-value <0.05, ‘***’= p-value <0.01, ‘****’= p-value <0.001.