M-Commerce, Sales Concentration, and Inventory Management

Nitish Jain
Management Science and Operations, London Business School, njain@london.edu

Tom Fangyun Tan
Operations, Information and Decisions, Southern Methodist University, ttan@cox.smu.edu

The mobile commerce (m-commerce) channel is poised to be the future of online markets. It offers search features distinct from conventional personal-computer-based (PC) e-commerce channels. Its easy accessibility extends the time available for customers to search, although its shopping environment constraints (e.g., small screen size, single-tab browsing) may inflate search costs. Collectively, these competing features cause ambiguity about the mobile channel's true effect on sales concentration – a key criterion in managing retail operations. The focus of this study is to understand the net impact of the mobile channel on driving sales concentration. Our study extends the empirical literature focusing on online retailers' channel strategies and their implication on operations practice. It examines differences in primary online channels – mobile and PC – on shaping sales concentration across products and the cost of ignoring such a difference on inventory management, a core decision for operations managers. We collaborate with a large online apparel retailer to access customer-level transactional data. We identify the mobile channel's effect on sales concentration using a difference-in-differences strategy that leverages a quasi-experiment stemming from the retailer's decision to discontinue its PC sales channel. We find that the mobile channel increases the share of popular products by 6.4% as compared with the PC channel. We also identify scenarios where ignoring this significant sales concentration difference will yield sub-optimal inventory stocking by 4.2% to 12.9%. Our paper highlights that the mobile and PC channels have different sales concentrations because of different search features. Ignoring this difference affects inventory decisions, such as safety and cycle inventory levels. Therefore, it is imperative for managers to revise their status quo strategies, such as on inventory procurement, assortment planning, and product display, when integrating m-commerce with e-commerce.

Key words: online retail, m-commerce, e-commerce, mobile channel, sales concentration, inventory management, empirical operations

1. Introduction

Shopping on a mobile device (i.e., m-commerce) is a rapidly accelerating trend. Not only is online shopping overall growing at 25% per year – and predicted to reach $480 billion (US) in 2019 (Mulpuru et al. 2015) – but also the m-commerce share of that total is increasing. In 2014, m-commerce accounted for 11.6% of the $303 billion in US e-commerce; this share is estimated to reach 45%, and to generate $284 billion, by 2020 (Meola 2016). Likewise, m-commerce as a share of online retail is forecast to be 76.4% by 2020 in Asia Pacific countries. The growing importance of mobile devices to modern retailers has attracted the attention of scholars (Shankar et al. 2010, Ghose 2017) and practitioners (Ariely 2016) to understand how the mobile channel's shopping environment differs to that of the conventional PC channel in shaping key customer responses. In this study we focus on one such response that closely affects inventory management: sales concentration (Gallino et al. 2017).

Sales concentration measures the skewness in sales contribution of different products that, in turn, has important implications for inventory management. That is to say, a high sales concentration indicates that only a few products contribute to a disproportionately large share of sales. A high sales concentration can translate to lower inventory costs due to the benefits of economies of scale and more accurate demand forecasts to manage inventory procurement (Simchi-Levi 2010). In contrast, a low sales concentration may warrant higher inventory levels. Driven by these observations, scholars have studied the impact of new business models (Girotra and Netessine 2013), processes, and technology advancements on sales concentration. For instance, Gallino et al. (2017) examined the effect of a novel 'ship-to-store' cross-channel functionality on sales skewness across products and its concomitant impact on inventory levels.

Over the last decade, the evolution of sales concentration in online markets has been shaped primarily by: (i) the demand-side factors, such as lower search costs enabled by internet-based search technology; and (ii) the supply-side factors, such as expanded product variety because of lower supply costs on the internet. Anderson (2006) predicts that the combination of new online search technology and expanded product variety will help customers find products that better satisfy their respective preferences and increase the demand for niche products (i.e., reduce sales concentration). This prediction is referred to as the Long Tail effect. Previous empirical research has shown that the online channel indeed reduces sales concentration as compared with the bricks-and-mortar channel (e.g., Brynjolfsson et al. 2011). In the past, customers primarily used personal computers and laptops ('PC devices' hereafter) to access online markets. The future, however, belongs to the mobile channel.

Is the mobile channel a 'grand equalizer' that shrinks the sales gap between popular and niche products (i.e., reduces sales concentration)? Or does it actually make the popular products even more popular (i.e., increases sales concentration)? Or, does it simply mimic the conventional PC channel? A clearer understanding of these queries is instrumental for retail managers to make optimal inventory decisions.

Although an online retailer, when adopting the mobile channel, can well maintain the same supply-side operations (e.g., offering the same product variety as the PC channel), the mobile channel may materially affect customers' demand for different products due to its distinct search features. On the one hand, mobile devices are portable, thus allowing customers to shop virtually anywhere and anytime. In other words, customers may have more available time to search the breadth of a retailer's offered assortment. On the other hand, mobile devices have smaller screen sizes and less browsing flexibility than PC devices. These factors make it difficult to evaluate multiple products simultaneously and inflate search costs. Collectively, these competing factors make it unclear whether the mobile channel further reduces, maintains or even increases sales concentration, compared with the PC channel. In this paper, we empirically examine this question.

We collaborated with a large online apparel retailer in India for this study. We obtained detailed, transaction-level data for a period of 50 weeks (from 15 January through to 31 December 2015). During this period, the retailer primarily operated through two online channels: a mobile-based application ('the mobile channel') and web browsers installed on PC devices ('the PC channel'). The two channels offered the same product assortment and used the same order fulfillment process. In other words, they shared the same supply-side factors that may affect sales concentration. This facilitates identification of the mobile channel's search environment's (a demand-side factor) impact on sales concentration while controlling for the confounding supply-side effects.

We face the challenge of customers' 'self-selection' into different online channels in identifying the mobile channel's causal effect on sales concentration. For example, customers who shop on the mobile channel may have different fashion preferences from those shopping on the PC channel. Also, customers may endogenously choose a shopping channel to purchase products with specific characteristics (Wang et al. 2015). These sources of self-selection can induce endogeneity biases in estimating a channel's causal effect (Angrist and Pischke 2008). We overcome this challenge by matching similar customers and exploiting an exogenous shock to customers' channel usage decision. In the middle of our study period, the retailer shut down its PC channel. As a result, the customers had to shift all their purchase transactions to the mobile channel. We leverage this quasi-experimental setup and perform a customer-level difference-in-differences (DiD) estimation to identify the causal impact of the mobile channel's search environment on sales concentration.

We find that the mobile channel increases sales concentration, as compared with the PC channel. In particular, our results suggest that customers who are willing to shift purchases from the

PC channel to the mobile channel would exhibit an increase of 6.4% in the share of popular products. This finding suggests that the high search costs on the mobile channel dominate the expanded available search time in affecting sales concentration. This search cost explanation is further corroborated by the following evidence: we find that the share of top-displayed products is significantly higher in the mobile-channel purchases than in the PC channel purchases. Also, we find that the mobile channel's average impact on sales concentration is moderated by category-specific features such as the availability of complementary information about the products, and product feature variety. Next, combining the estimated difference in sales concentration between the two channels with our collaborating retailer's parameters on product categories' assortment sizes, we show that a manager may do suboptimal inventory stocking for a category by 4.2% to 12.9% if he/she ignorantly plans for the mobile channel as for the conventional PC channel. Specifically, the manager is prone to understock popular products by an average of 2.7% (median of 2.8%), and overstock niche products by an average of 10.1% (median of 9.7%). Finally, we find that our focal finding is robust to a wide variety of alternate analyses, including that with alternate outcome variables, product popularity ranking measures, and placebo tests.

Our study makes important contributions both to practice and to the literature on implications of innovative retail practices on operational decisions. First, for retail operations managers, it provides a strong case for not blindly treating the mobile channel as the conventional PC channel. It reinforces the call for better understanding of the differential impact of alternate online channels on customer responses. The nature and extent of such differences may require managers to adopt channel-specific strategies. Second, it offers rigorous empirical evidence to resolve the natural ambiguity around the mobile channel's impact on sales concentration. Our collaborating retailer's strategic decision to shutdown the PC channel creates a unique quasi-experiment that eliminates confounding channel selection issues. Finally, we find that, compared to the PC channel, the mobile channel actually reverses the trend of reduction in sales concentration that is commonly associated with the advent of online markets. This significant differential impact has considerable bearing for optimal inventory management. Failing to account for this significant differential impact, managers can suboptimally invest in inventory while co-managing m-commerce alongside e-commerce.

2. Theoretical Background and Related Literature

In this study, we examine whether the distinct search features of the mobile and PC channels lead to a significant difference in their respective impact on sales concentration. Any such difference, in turn, will have implications for inventory management. Below, we first establish a theoretical link between a channel's search environment and sales concentration. Then we discuss the contribution of our study to inventory management and studies of mobile-channel economics.

Online Channels: Search Environment and Sales Concentration Controlling for supply-side drivers (e.g., product variety), the search aspects of a shopping environment play an important role in determining the sales distribution across popular and niche products (Brynjolfsson et al. 2011). In theory, customers should search all the products and choose the one that best fits their preferences. Such an expansive search would reveal true demand preferences. In practice, however, customers either cannot or will not evaluate all the products. Evaluating an option requires time and effort, incurring a *search cost* (Stigler 1961, Brynjolfsson et al. 2011). Moreover, customers have a limited amount of time and effort that can be allocated to search (i.e., a limited *search budget*). Hence, customers tend to consider those products of which they are ex-ante cognizant or those that are easy to search for and evaluate. These products are typically the most popular products because: (i) they are generally heavily advertised and promoted; (ii) friends often discuss them in social life by word of mouth (Frank and Cook 2010); and/or (iii) such products satisfy customers' habitual needs (Fazio et al. 2000). Take apparel (our empirical setting) for instance. Customers typically have ex-ante knowledge about fashionable products because they read fashion magazines (and other media) and engage in social interaction. As a result, demand will be more concentrated on popular products than theory would suggest. The bottom line is that the combination of limited search budget and high search cost can dissuade customers from searching an entire product offering, thus increasing the sales concentration of popular products.

Although mobile and PC are both online channels, their respective shopping environments differ markedly (Shankar et al. 2010). The portability of mobile devices allows customers to access the internet virtually anywhere and anytime (e.g., while commuting). In effect, the mobile channel enables customers to spend more time searching and evaluating products than the PC channel. This larger search budget may increase the likelihood of discovering a niche product and reduce sales concentration.

Yet the mobile channel's portability advantage comes at the cost of its size: mobile screens are much smaller than PC monitors. The smaller screens on mobile devices cannot display as many products as a PC can show. Consumers therefore need to frequently swipe the screen to see the products. In addition, mobile devices often do not support the multi-tab functionality that PC browsers have. For these reasons, mobile customers incur *higher search costs* when evaluating a product alternative than PC customers do (Ghose 2017). In other words, for a given search budget, consumers are likely to evaluate fewer options on a mobile phone than on a PC, which may increase sales concentration.

In summary, the above discussion suggests two theoretical competing factors of the mobile channel: high search cost and large search budget. They cast ambiguity on the mobile channel's net impact on sales concentration compared to that of the PC channel. Thus, the question of the mobile channel's impact on sales concentration – both with regards to direction and extent – is best answered empirically.

Sales Concentration and Inventory Management It has been long recognized that sales concentration (or related customer preference dispersion) has far-reaching implications for inventory management. For example, if sales spread out to a larger number of niche products, this can negatively increase investment in inventory on account of poor demand forecasts and lower potential for economies of scale across such products (Simchi-Levi 2010). Similar qualitative insight of increase in product differentiation translating to an increase in inventory costs has been illustrated in studies of flexible manufacturing (Netessine et al. 2002), delayed product differentiation (Lee and Tang 1997), and benefits of a pooling strategy (Bimpikis and Markakis 2015). In addition to inventory level, sales concentration also affects product assortment (Gaur and Honhon 2006), delivery lead time (Alptekinoğlu and Corbett 2010), and production line breadth (Moreno and Terwiesch 2016). These studies all tended to focus on the implication of sales concentration for various inventory management decisions; little, however, has been done to understand the effect of the new and rapidly growing mobile channel in shaping sales concentration across products and its implications for inventory management.

We contribute to the literature by studying how the mobile channel changes sales concentration, relative to the PC channel, and its implication for one important operations decision – safety/cycle inventory level. We adapt the approach first proposed by Gallino et al. (2017) to analyze the impact of the skewness of sales on inventory management. They build on the conventional inventory models – Economic Order Quantity (EOQ) and newsvendor – and show that an increase in sales spread across products would lead to an increase in optimal safety and cycle inventories, even when product assortment is kept constant. Using this approach, in our context, we find that a manager may suboptimally stock inventory by 4.2% to 12.9% if he or she ignorantly treats the mobile-channel and the PC-channel sales concentrations the same.

Empirical Literature Our paper closely relates to the empirical literature on mobile economics (see Shankar et al. (2010) and Ghose (2017) for excellent reviews). This stream of work has primarily tended to focus on the effect of the mobile channel on total sales. For example, in a grocery retail context, Wang et al. (2015) find that adoption of a mobile app increases purchase frequency, especially for low-spenders, who end up ordering more often and also in larger amounts. Likewise, Xu et al. (2016) estimate that customer adoption of tablets, an alternative m-commerce channel, can significantly expand e-commerce sales. In contrast to these papers, our study focuses on the mobile channel's effect on a *purchase content* aspect, namely the shares of popular and niche products among all the products purchased by a customer.

Among the papers that focus on purchase content, our paper is closest to a recent working paper by Doosti et al. (2018). Using four weeks' data from a business-to-consumer (B2C) marketplace that primarily focuses on non-apparel products such as fishing poles and charger supplies, these authors study the difference in sales concentration between purchases through a mobile app (the mobile channel) and through browsers in desktops and laptops (the PC channel). Using both the aggregate sales model (such as Lorenz curves and Gini coefficients) and customer-level analyses, the authors find that sales in the mobile channel are less concentrated than in the PC channel. This attenuation, however, is moderated by the extent of cross-product searches (i.e., search across different products, as opposed to searching for different variants within the same product, which we call *within-product search*). In particular, the higher the cross-product searches, the lower the mobile channel's attenuation of sales concentration.

Our paper's empirical context complements Doosti et al. (2018) in two important ways. First, we focus on apparel products – a product category with steady increase in online sales worldwide. In 2018, worldwide revenue for apparel products was $418 billion and is expected to increase to $713 billion by 2022 (Orendorff 2019). Second, Doosti et al. (2018) exploit cross-sectional variation as 77.4% of customers in their data execute single transactions over the four-week study period. In comparison, we are able to exploit within-customer variation in channel usage and purchase content due to a longer 50-week study period, and a quasi-experiment that provides an exogenous shock to channel usage. A within-customer analysis enables control for time-invariant customer-level unobservables that confounds with channel selection and sales concentration.

Interestingly, in contrast to Doosti et al. (2018), we find that the mobile channel significantly increases sales concentration in purchases, compared to the PC channel. We find consistent evidence for our finding in both the aggregate sales model and customer-level analyses. Compared to product categories (such as fishing poles) studied in Doosti et al. (2018) that invoke both cross- and within-product searches, customer searches are primarily cross-products in the apparel category. As individuals' size information is relatively consistent over time, customers' do not typically engage in active within-product searches. This principal difference in the prevalence of cross-product searches in the studied categories seems to drive the two studies' observed contrasting findings. Despite these differences, both studies reinforce that the mobile channel not only offers a considerably different purchasing environment compared to the PC channel, but also has product category-dependent differences.

Finally, our paper also belongs to the growing empirical works in operations management on antecedents of retail inventory and its implications, including Olivares and Cachon (2009), Hendricks and Singhal (2009), Kesavan and Mani (2013), Gallino et al. (2017), and Cui et al. (2018). We extend this literature by studying the differential impact of primary online channels on sales concentration and its concomitant implications on inventory management.

3. Research Setting, Data, and Design of the Empirical Study

3.1. Research Setting

For this study, we collaborated with a large Indian fashion online retailer, and collected 50 weeks of its transaction data from 15 January through to 31 December 2015. During the first half of this period (January through mid-May), the retailer sold products through the mobile channel and the PC channel. On 15 May 2015, however, the retailer shut down its PC channel and thereafter operated only through the mobile channel. According to the retailer's senior management, the shutdown decision was strategically motivated: the company sought to devote its efforts on m-commerce, which it viewed as the future of online shopping. Essentially, the timing of terminating the PC channel is independent from any considerations towards sales concentration. As a result, this quasi-experimental setup provides an excellent opportunity for us to study the causal impact of a mobile search environment on sales concentration.

3.2. Data

We obtained detailed information on transactions, and product attributes. In particular, the transaction-level data include time stamp, price, and unique identifiers for channel, product, and customers. We drop transactions that occurred on tablets and through the mobile devices' web browser (approximately 0.6%). The product attributes data comprise a four-tier hierarchical classification – apparel type (e.g., topwear, bottomwear) $\rightarrow$ article (e.g., shirts, dresses) $\rightarrow$ style (e.g., crew-neck red T-shirts) $\rightarrow$ stock keeping unit (SKU, e.g. crew-neck red T-shirts of size M) of a product and its introduction date. Over the study period, we have information on 21.26 million transactions covering products across seven apparel types, and 99 articles.

The retailer's customer privacy policy allowed us to obtain detailed demographics data (such as city location, gender, and first purchase date) for only a random sample of customers, and not for all customers. We construct our sample by focusing on customers who joined the retailer before the start of our study period, have no missing demographics information, and transacted at least once during the pre-shutdown period. Collectively, this selection criteria enabled us to include customers who are active, and familiar with the retailer's platform. Among the customers who meet this criteria, we receive from the retailer a random sample of 20% customers. The resultant sample consists of 162,800 customers who cumulatively made 1.8 million transactions. Table A1 of the online appendix presents key statistics at the sample and population level. We find the drawn sample to be representative of the population with regards to the average customer transaction statistics, demographics, and products coverage. For instance, the sample (resp. population) consists of 57.40% (resp. 57.26%) of female customers, 76.80% (resp. 76.81%) of customers living in Tier 1 (urban) or Tier 2 (semi-urban) cities, and cover products across 95 (resp. 99) articles. The dropped four articles accounted for only 37 transactions in the population data.

3.3. Unit of Analysis and Variables

We set the unit of analysis at the customer(i)×time-unit(t) level. We define the time unit as two weeks to balance between measurement accuracy and statistical power (Gallino et al. 2017). A larger time-unit aggregation (e.g., a month or a quarter) would include more incidences of a product purchase, which bolsters accurate measurement of product popularity. However, longer time aggregation reduces the sample size for time-series regression, which in turn reduces statistical power. The two-week time unit choice allows for a reasonable number of observations in both the pre- and post-shutdown periods while capturing a sufficient number of product purchases per time unit. We have also conducted robustness checks of different time aggregations, which yield consistent results. Next, we define the analysis variables.

Product: we define product at the style level. The retailer defines styles at a level that it embeds color information but excludes the size information. It is the third tier in the retailer's four-tier classification system. The next granular level is SKU which captures both the style and size information. An example of the retailer's style definition is 'Women Blue Kurta' and that of a corresponding SKU description is 'Women Blue Kurta, Size: M'. Thus, in our context, the effect of styles on purchases represents customer fashion preferences and tastes, and the effect of SKUs on purchases is determined by customers' respective (and exogenous) body types. Since we are interested in the effect of the mobile channel's search environment on enabling customers to find products that fit their preferences, using the style (like Soysal and Zentner (2014)) to define a product seems more appropriate than the SKU. Our sample consists of 259,315 products.

Product Category: we define a product category at the article level (second classification tier) so as to balance the number of categories, and the number of products within a category.

Product Popularity Ranking: we follow previous studies on apparel demand to construct a channel-specific product popularity index , at the product-categories (e.g., shirts, dresses) level (Soysal and Zentner 2014). Customers typically decide on a particular category to browse through or purchase from before they start looking through the numerous styles within that category. Thus, category-specific popularity enables an accurate measure of the effect of search costs on purchase decisions. Accordingly, we create product popularity rankings using a product's sales value in each product category over each channel during the two-week period (a low rank indicates higher popularity). As a result, we have two popularity rankings for each time-unit in the pre-shutdown period, and one ranking for each time-unit in the post-shutdown period. We use all the 21.26 million sale transactions in the population data to construct these rankings. Similar to Soysal and Zentner (2014), we find subtle variation in products' popularity across the two channels. For instance, on average, 62% of the popular products (top 10% of products in a popularity rank) in the mobile channel remain popular in the PC channel, and 3% of these products rank in the bottom half of the PC channel's popularity rank (bottom 50%).

Share of Popular Products, S_{it}*:* we define S_{it} as the ratio of the number of popular products that customer i purchased in period t to the number of all products she/he purchased in that period (Brynjolfsson et al. 2011, Tan et al. 2017). This percentage-based classification allows us to account for varying product assortment sizes across product categories over time. We use the share of popular products to measure the sales concentration level in purchased products. A large S_{it} suggests a large sales concentration.

Relative Discount of Popular Products, $RDiscount_{it}$*:* we measure customer $i's$ relative discount choices between the popular and non-popular products as the ratio of the average 'discount level' of the popular products to the average discount of the non-popular products. The discount level is defined as 1 minus the ratio between price paid and the product's posted retail price.

Customer Channel Usage	Relationship Tenure (days) Average	25^{th}	50^{th}	75^{th}	Location Information† Tier 1	Tier 2	Tier 3	Tier 4
Mobile Only	380	95	259	655	38.4%	33.4%	25.9%	2.3%
PC Only	458	173	397	748	49.8%	30.3%	17.9%	2%
Omnichannel (Mobile + PC)	460	166	392	761	53.6%	32.2%	12.8%	1.4%

Notes. The period of observation is pre-shutdown of the PC channel. Relationship tenure measures days difference between the study period's start date and a customer's first purchase date. Location information captures economic status of a customer's residence.
† Tier 1: urban locations, Tier 2: semi-urban locations, Tier 3: semi-rural locations, and Tier 4: rural locations.

Table 1 Channel Usage and Customer Characteristics

3.4. Identification Strategy

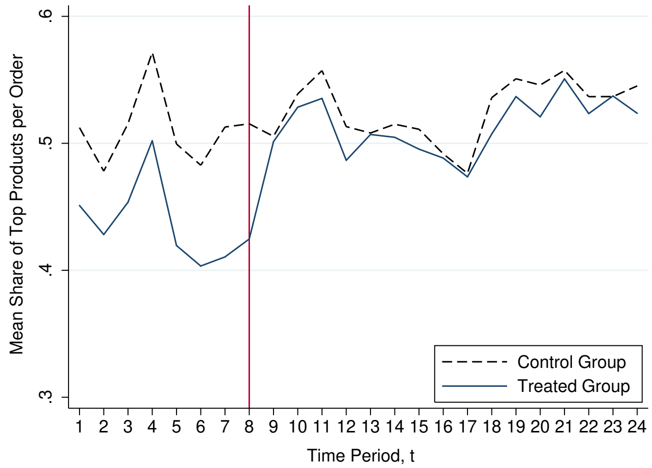
Two sources of selection bias compound the identification of the mobile channel's impact on sales concentration. First, customers using the mobile channel may be different to those using the PC channel in numerous demographic aspects, especially in an emerging country like India. The availability of affordable smartphone devices enables otherwise excluded customer segments living in semi-urban and rural locations to access online retail via the mobile channel (Srivastava 2016). These demographic aspects can influence customers' shopping preferences which, in turn, would affect sales concentration across products. For example, the purchase decisions of a handful of elite households in rural locations often strongly shape the local population's preferences, thus leading demand to be more homogeneous (Arnould 1989). In contrast, the rich cultural diversity of metropolitan areas means that the shopping decisions of urban customers may exhibit a wider variety, which translates into sales that are more dispersed. In our sample, customer characteristics are indeed different depending on their channel usage. Table 1 shows the differences among customers who made purchases – in the pre-shutdown period – using (i) only the mobile channel (hereafter, the 'mobile' customers, 50% of the sample); (ii) only the PC channel ('PC' customers, 36%); and (iii) both the mobile and the PC channel ('omnichannel' customers, 14%). In terms of relationship tenure, defined as the difference between the date of their historical first purchase and the start of our study period, PC customers are similar to omnichannel customers (e.g., their means are 458 days and 460 days, respectively). Their relationship tenure, however, is significantly longer than the mobile customers' (mean of 380 days). In terms of residence location, mobile customers have a relatively wider representation from Tier 1, 2, and 3 locations than PC and omnichannel customers.

The second source of bias may originate from the flexibility available to omnichannel customers who, unlike single-channel users, can endogenously select channels conditional on the purchase product's characteristics (i.e., self-sorting into the appropriate channels). For example, Wang et al. (2015) find that customers often prefer the mobile channel (to other online channels) when purchasing products that require a less thorough evaluation or have been purchased in the past (i.e., 'habitual' products).

A Difference-in-Differences Estimation Model To overcome these possible selection biases, we exploit the quasi-experiment created by the retailer's decision to shut down the PC channel in the middle of our study period. As a result, customers experienced shopping environment changes that varied depending on their respective pre-shutdown channel usage. Mobile-only customers, who remained in the post-shutdown period, continued to make all their purchase transactions over the mobile channel post-shutdown. In other words, they were not affected by the PC channel shutdown. In contrast, PC customers, who remained in the post-shutdown period, had to switch to a new mobile-based environment for post-shutdown shopping. Effectively, these customers received a treatment of channel switch in their purchasing environment. We therefore respectively classify the PC and mobile customers, who made at least one purchase in the post-shutdown period, as the treated and control group.

We perform a customer-level difference-in-differences (DiD) estimation to identify the causal effect of the mobile channel's shopping environment on sales concentration. The DiD framework exploits customers' change in search behavior when moving from the pre-shutdown to the post-shutdown period. For PC customers (the treatment group), this change is the composite effect of switching to a new shopping environment (the mobile environment) and different temporal conditions (e.g., changes in product assortment over time, or seasonality-driven sales concentration). For mobile customers (the control group), the change across the two periods depends only on the temporal conditions because the shutdown event itself does not alter their shopping environment. Using the control group's relative change, the DiD analysis 'differences out' the effect of temporal conditions from the relative change in the treated group, and so yields the mobile environment's causal impact on sales concentration. Given its potential to identify causal effects, DiD estimation is widely used across fields, including economics and operations management, to evaluate effects of policy or managerial intervention (e.g., Gallino et al. (2017), Lee et al. (2017), Ertekin and Agrawal (2019), Tan and Netessine (2020), Chen et al. (2020)).

In summary, the quasi-experiment enables control for selection bias in channel usage due to the difference in customer demographics, and our focus on single-channel customers eliminates the potential bias in estimates due to omnichannel customers' endogeneous channel selection, conditional on the purchased product's characteristics.



Notes. In the x-axis, a time-period t captures a two-week duration. The y-axis captures the average share of the top products in the matched sample customers' purchases. The PC channel shutdown event occurred in period 8.

Figure 1 Parallel Trend Assumption: Evidence from the Pre-Shutdown Period

Validation of DiD Assumptions The DiD framework relies on two critical assumptions. First, the control group exhibits the counterfactual change in the treated group had it *not* received the treatment (the 'parallel trends' assumption; see Angrist and Pischke 2008). There are several reasons why our data can be presumed to satisfy this assumption. We examine the evolution of sales concentration variables of the two groups during both the pre-shutdown and the post-shutdown periods. Figure 1 illustrates the temporal trends of the share of top products in period t for the treated group (solid line) and the control group (dashed line). The figure's x-axis (time) marks out two-week periods, and the vertical line corresponds to the shutdown event. The y-axis represents the average share of top products in t-period transactions by the customers in the matched sample. The treated and control groups exhibit very similar trends during our study period, although their gap shrinks immediately after the shutdown. This pattern supports the parallel assumption pre-shutdown. In addition, in Section A2 of the appendix we provide a formal test for the parallel trends assumption.

The second assumption is that conditional on a set of observed factors, individuals should be randomly allocated to the treated or control group (Angrist and Pischke 2008). In other words, the likelihood of an individual receiving treatment should not be contingent on unobservable factors that, in turn, also determines the outcome variable. One such unobservable factor can be customer engagement, which increases variety-seeking behavior (Simonson 1990). However, our setting and analysis alleviate such concern and should satisfy the second assumption of random treatment. In particular:

- In our setting, unconditional on any observables, the treated PC customers—in comparison to the PC customers who left the retailer after the PC channel shutdown—spent more money, purchased products with greater variety, spent relatively more money on the private brands (less known than public brands and, thus, require more search efforts), and purchased more during the office hours and weekdays (see Figure A1 in the appendix). Collectively, both higher engagement with the retailer's platform and exogenous factors that influence the PC channel access contribute to the treated PC customers' decision to remain after the PC channel shutdown. Besides, we find that the treated PC customers indeed exhibit significantly lower sales concentration (more variety-seeking) compared to the PC customers who left, most likely an artifact of relatively higher engagement of the treated PC customers. A formal test comparing difference in the average sales concentration between these two groups indicates significantly lower sales concentration in the treated PC customers at the p-value of 0.010. The higher engagement trait of the treated PC customers would enable them to continue their post-shutdown purchases in line with their pre-shutdown purchase patterns which, in turn, would bias against finding a treatment effect.
- In our analysis, we match the treated and control customers to minimize imbalance across them on covariates that may proxy for unobserved customer engagement and shopping preferences that can influence sales concentration. We estimate the DiD model on such a matched sample. In particular, we employ the Coarsened and Exact Matching (CEM) method of Iacus et al. (2012) to balance the following customer characteristics: (i) customer location; (ii) relationship tenure; (iii) gender; (iv) average order value of the pre-shutdown purchases; and (v) number of orders in the pre-shutdown period. We perform matching using CEM over the conventional Propensity Score Matching (PSM) methodology as it is more robust to the extent of pruned (unmatched) observations (King et al. 2016). In Section 5 Robustness Analysis, we present estimation results with PSM-matched samples.

  We note that matching effectiveness in eliminating all bias sources is contingent on relevance of the available observables in controlling for unobservables. In our context, we find that matching on observables diminishes, the above reported, significant difference in the average sales concentration between the treated PC customers and the PC customers who left. A formal test comparing the difference in the average sales concentration between the matched sample of these two groups fails to reject equality between them at the p-value of 0.914. Figure A2 show the difference in average sale concentration between these two groups across

	Unmatched Sample		Matched Sample	
Variable Description	PC Customers	Mobile Customers	PC Customers	Mobile Customers
Customer Share	32%	68%	50%	50%
Transaction Share	(29.1%, 29.0%)	(70.9%, 71%)	(48.6%, 46.7%)	(51.4%, 53.3%)
Popular Products Share (S)[1]	(43.8%, 51.5%)	(51.1%, 52.6%)	(43.8%, 51.5%)	(51.1%, 52.6%)
Rel. Discount (RDiscount)[1,2]	(1, 1.02)	(1.04, 1.04)	(1, 1.02)	(1.04, 1.04)
# Transactions per Customer[2]	(0.67, 0.73)	(0.80, 0.82)	(0.61, 0.71)	(0.64, 0.78)
Amount Spent per Period[1,2]	(1, 1.22)	(1.04, 1.23)	(0.96, 1.19)	(0.94, 1.19)
Female Customers Share[2]	1	0.9	1	0.9
Tier 1 & 2 Customers Share	8 2.0%	72.6%	81.9%	81.9%
Relationship Tenure	483	399	483	483
# Customers (i)	20,860	43,972	20,052	20,052
# Observations (i×t)[2]	(32.6K, 55.2K)	(78.7K, 135.0K)	(31.2K, 53.2K)	(33.9K, 61.1K)

Notes. Sample of the PC and mobile customers who made at least one purchase in the pre-and post-shutdown period. The matched sample is obtained using CEM matching on customers' location, relationship tenure, gender, and pre-shutdown average order value and number of orders.
[1] The first and second number in the bracket respectively denotes the pre-and post-shutdown statistics.
[2] As per the retailer's requirement, the PC customers numbers in the pre-shutdown period are normalized to 1.

Table 2 Summary Statistics

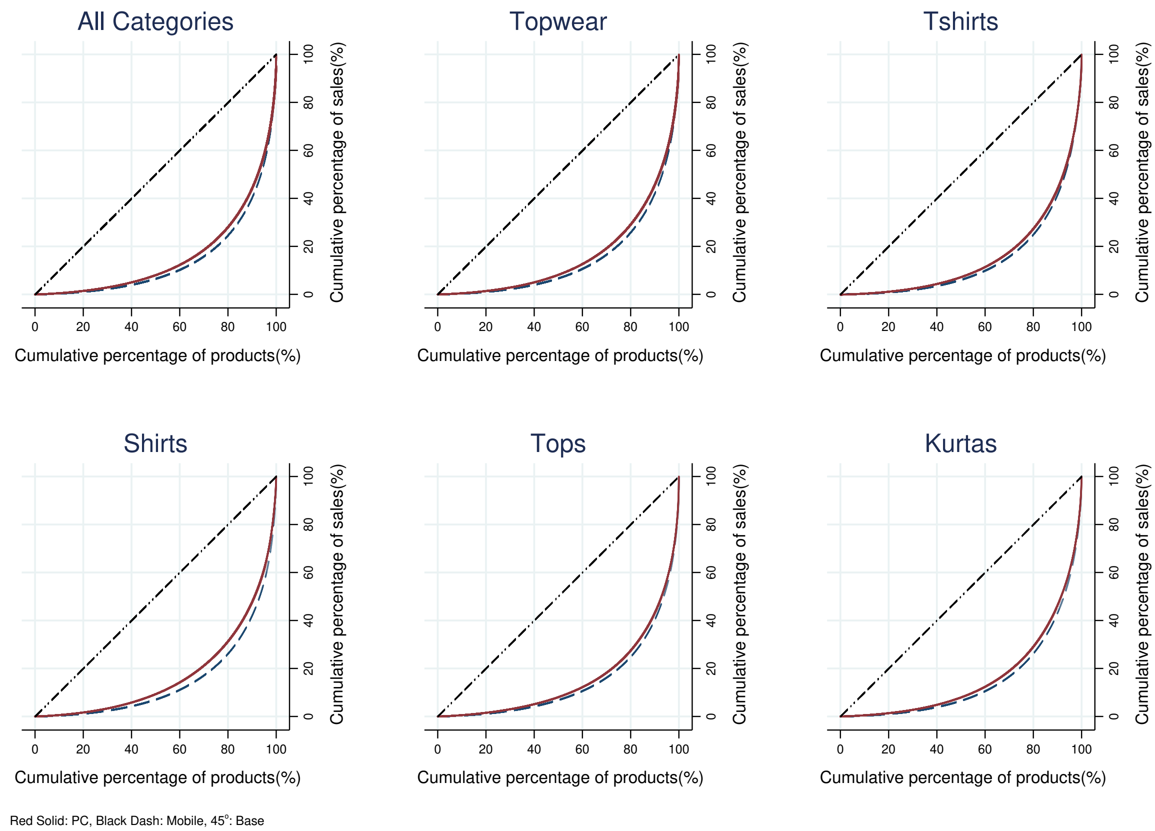
'all categories', in the top-selling apparel type (Topwear), and in the four top-selling product categories (T-shirts, Shirts, Tops, and Kurtas). Collectively, these figures indicate that the sales concentration levels of the two PC customer groups are closely intertwined: they are similar.

Model Specification We estimate the mobile channel's impact on sales concentration using the following customer-fixed effects DiD model:

$$S_{it} = \alpha_0 + \alpha_1 \times PC_i \times Post_t + \alpha_2 RDiscount_{it} + \mu_i + \tau_t + \varepsilon_{it}. \quad (1)$$

Here, PC_i is a dummy variable set equal to 1 for a treated PC customer or 0 for a mobile customer in the control group. $Post_{it}$ is a dummy variable set to 1 for the post-shutdown period and to 0 for the pre-shutdown period. Furthermore, customer-fixed effects μ_i control for unobserved customer heterogeneity. Time-fixed effects τ_t provide a nonparametric control for temporal factors (such as trends and seasonality), thus avoiding any functional assumption on evolution of temporal pattern during our study period. The coefficient α_1 captures the mobile channel's effect on sales concentration relative to the PC channel's effect. A positive (negative) α_1 indicates that the mobile channel will cause sales to be more (less) concentrated on popular products than the PC channel.

We estimate the above described DiD specification (eq(1)) on a matched sample with 179,449 customer×time-unit observations. Table 2 shows the summary statistics of customers in the treated



Notes. Lorenz curves reflect sales concentration in the two channels' pre-shutdown purchases. The solid and dash line respectively capture the PC and mobile channels' sales concentration levels.

Figure 2 Lorenz Curves for the Mobile and PC Channels

and the control groups. In the unmatched sample (columns 1 and 2), the control mobile customers account for a large percentage of transactions, exhibit higher purchase frequency, spent more per period, have a lower representation from Tier 1 & 2 locations, and started their relationship with the retailer relatively recently. In comparison, the two groups exhibit more balance across all these dimensions under the matched sample (columns 3 and 4).

4. Results

In this section, we first use all sale transactions of the population data to present preliminary evidence on the mobile channel's impact on sales concentration. Then, we show the customer-level DiD estimation results. Both sets of results are consistent.

4.1. Preliminary Model Free Evidence: Lorenz Curve

Figure 2 presents Lorenz curves – often used to understand distribution inequality in economics – for the mobile and PC channels as preliminary evidence of differences in sales concentration between them (Brynjolfsson et al. 2011). In particular, these curves demonstrate the sales concentration across all purchases made over a channel during the pre-shutdown period (in total 6 million transactions). Figure 2 plots sales information for: (i) all categories; (ii) the top-selling apparel type – Topwear (accounting for 65.6% of the total transactions); and (iii) the top four product categories (T-shirts, Shirts, Tops, and Kurtas). In all six graphs, the mobile channel's Lorenz curve (dashed line) sags below the PC channel's curve (solid line), suggesting that, compared with the PC channel, sales are more concentrated over the mobile channel. We formally test for this observed difference across the two channels using two aggregate level regression analyses: Gini Coefficient and Log-Linear sales model. Appendix A3 shows the results of these analyses. We find consistent support for higher sales concentration in the mobile-channel purchases compared to that in the PC channel.

4.2. Customer-Level DiD Model

Table 3 shows the estimation results of equation (1). A repeated transaction setting like ours is vulnerable to inaccurate inference due to biased standard errors on two accounts. First, because of unobservable customer-level factors (e.g., shopping preferences), the multiple transactions executed by a customer may exhibit clustered correlation in the corresponding error terms. Including customer-level fixed effects only partially corrects for the bias due to clustering (Petersen 2009). Second, our two-week time unit can be shorter than firm-level decision periodicity (e.g., seasonal replacement of product assortment) which may introduce serial correlation in the error terms; any such correlation will bias the standard errors (Wooldridge 2015). To mitigate these inaccurate inference concerns, we use robust and bootstrapped standard errors. In a seminal paper studying the effect of serial correlation on DiD coefficient–based inference, Bertrand et al. (2004) find that the nonparametric approach of bootstrapped errors outperforms the parametric approach of imposing a particular functional form (e.g., AR(1)) on serial correlation. The bootstrapped approach preserves the error terms' inherent correlation while empirically computing the estimates for standard errors.

Columns 1–3 of Table 3 show ordinary least-squares (OLS) estimation results with the unmatched sample. Column 1 presents results with robust standard errors that are corrected for the possibility of clustered correlation in error terms at the customer level, but without the discount control variable. Columns 2 and 3 respectively show the full specification results with robust and bootstrapped standard errors. Columns 4–6 report the corresponding estimation results obtained using the matched sample.

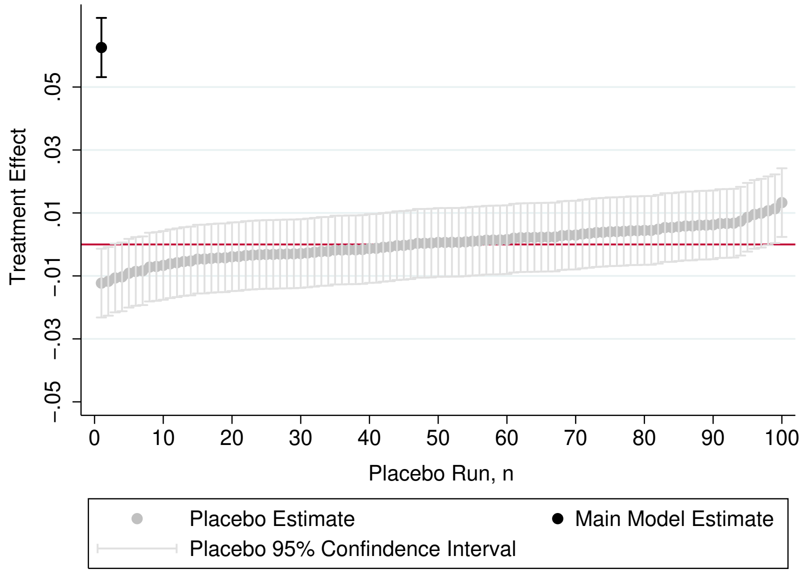
	(1)	(2)	(3)	(4)	(5)	(6)
	Share of Popular Products					
Estimation Description	Unmatched Sample			Matched Sample		
Standard Errors	Robust		Bootstrapped	Robust		Bootstrapped
PC×Post	0.0654***	0.0628***	0.0628***	0.0661***	0.0644***	0.0644***
	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.009)
RDiscount	-	0.1150	0.1150	-	0.0755	0.0755
	-	(0.093)	(0.083)	-	(0.110)	(0.096)
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Customer Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
# Observations	301,460	301,460	301,460	179,449	179,449	179,449
# Customers	64,832	64,832	64,832	40,104	40,104	40,104
Prob > Chi-sq	< 0.001%	< 0.001%	< 0.001%	<0.001%	<0.001%	<0.001%

Notes. Unit of analysis is customer×time-unit. Robust standard errors are clustered at the customer-level. Sample observations are less than #Customers×#Time-Periods as not all customers made purchases in all the periods.
*** p < 1%, ** < 5%, * < 10%

Table 3 Main Results

We find consistently significant and positive DiD coefficients (α_1; row 1 in Table 3) at a level of less than 1% p-value (0.0654, 0.0628, 0.0628, 0.0661, 0.0644, and 0.0644, respectively). This implies that, in comparison to the PC channel, the mobile channel increases sales concentration. In particular, interpreting the coefficient from column 6, we find that the customers, who are indifferent in channel preference when purchasing apparel, are likely to exhibit an increase in the share of popular products by 6.44% in their mobile-channel purchases compared to the PC-channel purchases. In other words, we find that the effect of the mobile channel's higher search costs dominates the effect of its larger search time budget.

Compared to the matched sample estimates, the unmatched sample estimates are relatively smaller. We note that the matching procedure results in exclusion of the PC customers' with atypical purchase behavior. In particular, the unmatched PC customers spent on average about three times more, and transacted on average five times more frequently than the matched PC customers during the pre-shutdown period (18.02 vs 3.31 purchases). This high frequency and spend pattern of the unmatched PC customers' is likely to be driven by factors that overshadow natural consumption of fashion products, including any channel considerations – for instance, a small business placing repeated orders at a high frequency for downstream sale to other businesses. Intuitively, one would expect these customers to largely continue with purchase patterns of the pre-shutdown period i.e., such customers would exhibit least change in the pre- and post-treatment sales con-



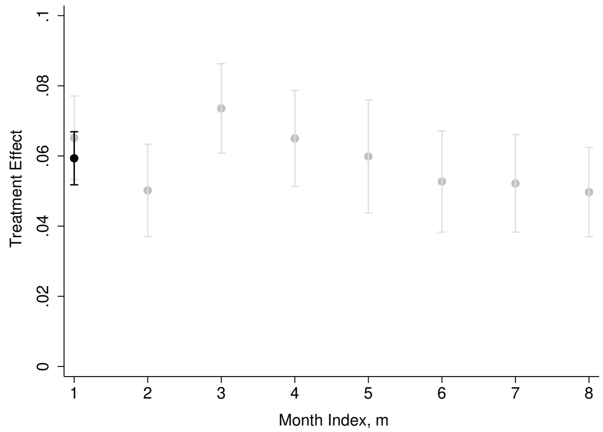
Notes. Placebo estimates are sorted in ascending order. The black dot and boundary line denote the main DiD treatment coefficient and its corresponding 90% confidence interval (CI). The grey dots and boundary lines reflect the DiD treatment effect and corresponding CIs obtained with pseudo-treated and control groups. The CIs are construed using bootstrapped standard errors.

Figure 3 Placebo Test: Pseudo-Treated Group

centration levels. Thus, we see a slight drop in the treatment effect with the unmatched sample compared with the matched sample.

We next examine the validity of our finding in two aspects. First, from the analysis standpoint, is the DiD estimation picking spurious effects that are driven by reasons other than the mobile channel effect? Second, from the customers' learning standpoint, is the observed finding an artifact of the PC customers' slow adjustment to the purchase environment change (the Hawthorne Effect Pierce et al. (2015)), and thus would attenuate over time, or of the mobile channel's innate search features, and thus would remain persistent over time? We examine these queries next.

4.2.1. Validating DiD Estimation: Placebo Test We conduct a placebo test to ascertain whether our DiD finding is spurious. If our estimation approach also finds that the control group customers were affected by the treatment event, then it strongly indicates that our finding is confounded with factors other than the treatment event. To examine this, we divide the control group customers into two pseudo groups of equal size – a pseudo-treated group and a control group. We then use equation (1) to replicate our DiD estimation. We repeat this placebo test 100 times. Figure 3 plots the estimated DiD coefficients (the grey points) for these tests, in the ascending order, along with their corresponding 90% confidence intervals (CIs, the gray bounding lines). The main analysis DiD coefficient and its 90% CI are marked by the black dot and bounding line around it. Only four of 100 tests have CIs completely disconnected from the horizontal line (zero). Therefore, we fail to



Notes. The *x*-axis shows month m after the shutdown event. The black dot and boundary line denote the main DiD treatment coefficient and its corresponding 90% confidence interval (CI). The grey dots and boundary lines denote the respective month's treatment effect and corresponding CIs. The CIs are construed using bootstrapped standard errors.

Figure 4 Monthly Effects of the Mobile Channel

reject the null hypothesis of an insignificant effect of the shutdown event on the pseudo treatment group's sales concentration level.

4.2.2. DiD Treatment Finding: Transient or Persistent? We aim to establish whether or not the treatment effect attenuates over time in our setting. For that purpose, we modify the DiD specification of equation (1) to estimate separate treatment effects for each of the eight months during the post-shutdown period (see Pierce et al. 2015, Tan and Netessine 2020). Figure 4 show the monthly treatment effects on the share of popular products. The black dot denotes the average DiD treatment effect across the months as estimated using the main specification. The grey dots represent the respective month m treatment effect. The associated bars are the 90% confidence intervals. The magnitudes of the monthly coefficients tend to be stable over time, with overlapping confidence intervals. In other words, we find support for the observed treatment effect to be caused by the mobile channel's innate search features, and not due to the PC customers' potential slow adjustment to the purchase environment change.

5. Robustness Analysis

In this section, we conduct a series of robustness tests by employing alternative dependent variables, matching criteria, product definition criteria, time aggregation methods, and unit of analysis. We also provide additional robustness checks in the Appendix. All of them yield congruent results.

Alternate Dependent Variables We first consider alternative definitions of a 'popular' product to ensure that our observed finding is not sensitive to the definition used in our main analysis (i.e., a product that ranks among the top 10% within time period t). In particular, we examine three other cutoffs: the top 20% for a popular product, the bottom 10% and bottom

20% for a niche product. Similarly, our main outcome variable measures sale concentration using the count of popular products. We examine whether our finding is sensitive to the alternate value-based view of sale concentration that can materially influence a firm's operational decisions like inventory investment. We define the value-adjusted share of popular products as $ws_{it} = \sum_{j \in \text{Popular products}(i,t)} sales_{jit} / \sum_{j \in \text{All Products(i,t)}} sales_{jit}$ where $sales_{jit}$ captures the sales value of product j purchased by customer i in period t, Popular Products(i,t) denotes the set of popular products purchased by the customer i in period t, and All Products(i,t) denotes the set of all products purchased by the customer i in period t. The regression results when using these alternative popularity definitions are shown in rows 2–5 of Table 4. Similar to our main result (row 1), the coefficients for popular products are consistently significant and positive (0.049, 0.025) while the coefficients for niche products are significant and negative (-0.0128, -0.079, -0.005).

Alternate Matching Criteria We perform three robustness tests about the matching choice in the main analysis. First, we implement a stricter matching criterion by including four additional variables that capture the shopping pattern differences of the mobile and PC customers. Towards this, we measure customers' relative share of purchase transactions over commuting hours (eight hours period: 06.00-10.00 and 18.00-22.00), working hours (eight hours period: 10.00-18.00), night hours (eight hours period: 22.00-06.00), and over the weekend. Collectively, these variables help control the potential confounding factors of shopping occasions that may influence a customer's search efforts. For instance, on weekends a customer may have more leisure time to engage in a more involved search for a product evaluation. In row 6, we present results based on this stricter matching approach. We note that this stricter matching reduces the sample size to 101,270 observations, approximately 56% of the original sample size. The smaller sample size makes it difficult to perform other complementary analyses of this study. We therefore use this stricter matching only as a robustness test. Next, we replicate the main analysis using two PSM-matched samples—nearest 1 and 3 neighbors— to examine whether the observed finding is sensitive to the matching methodology choice. Figure A3 of the appendix presents the bias reduction attained under the two PSM samples, and rows 7–8 report the estimates. Across all these three alternate sample construction approaches, the treatment coefficients are consistently positive and significant (0.0729, 0.0595, and 0.0612, respectively).

Alternate Product and Time Aggregation We examine whether our results are sensitive to the product, category and time aggregation choices made for the main analysis. To start, we use the

#		Test Description	Treatment Effect	# Obs
1	Main Model	Popular Products (Top 10%)	0.0644*** (0.009)	179,449
2	Alternate Dependent Variables	Popular Products (Top 20%)	0.0558*** (0.005)	179,449
3		Niche Products (Bottom 10%)	−0.0054*** (0.001)	179,449
4		Niche Products (Bottom 20%)	−0.0112*** (0.001)	179,449
5		Sales-Adjusted Popular Products (Top 10%)	0.0618*** (0.005)	179,403
6	Alternate Matching Definitions	Matching on Extended Variable List	0.0729*** (0.023)	101,270
7		PSM Matching (N-1)	0.0595*** (0.005)	182,242
8		PSM Matching (N-3)	0.0612*** (0.004)	226,178
9	Alternate Product and Time Aggregation	Product: SKU	0.0637*** (0.004)	179,449
10		Product Category: Apparel-Category	0.0642*** (0.004)	179,449
11		Time: Month	0.0563*** (0.004)	153,689
12	Alternate Unit of Analysis	Order-level Analysis	0.0486*** (0.005)	410,112
13		Before-After Analysis	0.0283** (0.011)	48

Notes. The unit of analysis in rows 1 to 11 is customer×time-unit. In rows 12 and 13 the unit of analysis is respectively customer-order and week. Standard errors in the parantheses are bootstrapped errors. In CEM extended matching, the additional matching variables are customers' relative share of purchase transactions over commuting hours, working hours, night hours, and weekend. In the PSM matching (rows 6 and 7), the maximum threshold is set to 0.01. The before-after analysis defines the product's popularity based on a weekly popularity ranking.
*** p < 1%, ** < 5%, * < 10%

Table 4 Robustness Results

SKU number as an alternative (to the style number) product definition. Next, we aggregate the products at the 'apparel type' level (the top tier of the apparel classification hierarchy) as an alternative to the product-category definition. Lastly, we analyze monthly aggregation as a robustness check of our current biweekly aggregation. Rows 9 through 11 show the results of these robustness checks. The coefficients are all significant and positive (0.0637, 0.0563, 0.0642, respectively).

Alternate Unit of Analysis Finally, we test the robustness of our results using two alternate unit-of-analysis: (a) customer-order, and (b) week. In the customer-order analysis, we replicate the DiD identification strategy at the order level. It offers the advantage of including a richer set of controls such as time of purchase, but is vulnerable to a confounded change in the PC customers' order-size and -frequency as they shift purchases to the mobile channel (Narang and Shankar 2016). In the weekly analysis, we implement a 'before-and-after' analysis that offers the advantage of examining a change in the firm-level sales concentration. Compared to the main analysis, this analysis is vulnerable to the selection issues discussed in Section 3.4. We measure the firm-level sale concentration in week w as the weighted average of the PC and mobile channel's share of popular products, where weights are set equal to respective channel's total sales value, including of omnichannel customers, in week w. Similar to the main analysis, we compute a channel-specific product popularity rank at week level. Rows 12 and 13 respectively report the results of the customer-order and before-and-after analysis. The treatment coefficients are significant and positive (0.0486, 0.0283 respectively).

To sum up, across the 12 robustness tests, we consistently find support for our finding: the mobile channel increases sales concentration, supporting our main results.

In Appendix A4, we present analysis using omnichannel customers where the identification rests on exploiting the customers' relative difference in the mobile channel exposure across product categories. We find qualitative support (in direction and significance) for our principal finding, identified using a much cleaner empirical setting of single-channel users. In Appendix A6, we show the robustness of our finding to an alternate outcome variable: the slope of the Pareto sales curve, a commonly studied sales dispersion measure (Brynjolfsson et al. 2011, Gallino et al. 2017). Finally, in Appendix A5, we provide additional evidence that our finding is driven by differential search features across the two channels. We find that more top display-ranked products, which require lower search costs, are purchased on the mobile channel than on the PC channel. Intuitively, high search costs make it difficult for customers to browse the products displayed at the bottom of the page, which results in more purchases of top display-ranked products

6. Impact of Mobile Channel on Sales Concentration: By Product Type

In this section, we explore the sensitivity of the mobile channel's impact on sales concentration to product type to understand the nuanced effects of search efforts. We test the following three product type characterizations that may affect customers' search behavior: value-based, information-based, and feature-variety based. In particular, we define a product to be of Value-High (resp. Value-Low) if its retail price is more (resp. less) than the median price of that product's category. Next, capturing the availability of complementary information sources, we classify a product as Information-Low (resp. Information-High) if it belongs to a private brand that is owned by the retail firm (resp.

external brand i.e. international/national brands like Gap, Peter England, etc.). Compared to the private brands, external brands typically publicize their products in multiple platforms such as weekly fashion magazines and social media, thus providing multiple sources of information for a product's features. Finally, reflecting the available feature variety, we classify products into Variety-High and Variety-Low types based on whether they belong to the 'Topwear' and 'Bottomwear' product category. Topwear products (such as tops and dresses) offer a much richer variety in patterns and colors than Bottomwear products (e.g., trousers, leggings, etc.). Indeed, our collaborator carries 2.5 times more variety in the Topwear category than in Bottomwear. Figure A4 in Appendix A1 shows a few examples of products in the Topwear and Bottomwear.

We obtain the moderating effects of the aforementioned product types by separately estimating the following model for the each product-type:

$$\begin{aligned} S_{imt} = \beta_0 + \beta_1 \times ProductType_m + \beta_2 PC_i \times Post_t + \beta_3 ProductType_m \times PC_i \times Post_t \\ + \beta_4 \times RDiscount_{imt} + \mu_i + \tau_t + \varepsilon_{imt}. \qquad (2) \end{aligned}$$

The dependent variable where S_{imt} denotes the share of popular products purchased by the customer i in period t of the product-type with the moderating feature $m \in \{$Value-High, Value-Low, Information-High, Information-Low, Variety-High, Variety-Low$\}$. We construct channel-specific product rankings for each of these moderating features. For example, we have separate PC-channel popularity rankings for the Value-High and Value-Low products. Thus, for a given product type, we have four rankings in each pre-shutdown period, and two rankings in each post-shutdown period. Across these ranks, we classify the top 10% of products as popular products. The dummy variable $ProductType_m$ is set to 1 if the observation's moderating-type classification is high as per the Table 5. The coefficients β_1 and β_3 respectively capture a product type's baseline effect, and its moderating effect on the mobile channel's impact on sales concentration.

Table 5 presents the results of the moderating effects of the three product type categorizations. First of all, consistent with the main analysis, we find that the mobile channel effects are significant and positive for all three types of product categorizations (column 2). In addition, the coefficient of High-value products is statistically significant and negative (-0.0504), while its interaction term is statistically insignificant (0.0027). These results suggest that customers are more invested in purchasing expensive products and, consequently search them more thoroughly than the cheaper ones, irrespective of the channel. As a result, sales of expensive products are more dispersed across different products.

(#)	Definition	Main Effect		Marginal Effect
		ProductType (β_2)	Mobile Channel (β_1)	Mobile Channel's on ProductType (β_3)
1	Value: High (1) vs Low (0)	-0.0504^{***} (0.002)	0.0565^{***} (0.006)	0.0027 (0.004)
2	Information: High (1) vs Low (0)	0.0422^{***} (0.003)	0.0483^{***} (0.006)	0.0090^{*} (0.005)
3	Variety: High (1) vs Low (0)	0.0212^{***} (0.003)	0.0526^{***} (0.006)	0.0124^{**} (0.005)

Notes. The unit of analysis is customer×moderating-feature×time-unit. Standard errors in the parantheses are bootstrapped errors. Rows 1, 2, and 3 respectively capture the moderating role of products' price levels, ease of information availability, and product-variety that effects cross-product searches.
*** p < 1%, ** < 5%, * < 10%

Table 5 **Varying Mobile Channel Effects on Sales Concentration by Product Type**

In the Information High-Low model (row 2), the coefficient of Information-High (external brands) and its interaction term with the mobile channel are both statistically significant and positive (0.0422 and 0.0090). These results imply that external brands show a higher sales concentration than private brands, and that this tendency exacerbates the sales concentration effect of the mobile channel. Customers have less information elsewhere about the private brands than the external brands, so they may search these brands more extensively than the external brands. As a result, they are more likely to discover the niche private brands that fit their heterogeneous needs, resulting in lower sales concentration.

Finally, in the Variety High-Low model (row 3), the coefficient of Variety-High (Topwear) and its interaction term with the mobile channel are both statistically significant and positive (0.0212 and 0.0124). These results suggest that the wider feature variety of topwear products may make customers search them less thoroughly than bottomwear, and consequently concentrate the sales more on the most popular topwear products than the bottomwear products. In other words, similar to (Doosti et al. 2018), we find that the mobile channel increases sales concentration in product categories that induce more cross-product searches. These results support Tan et al. (2017) who find that product variety increases sales concentration. These moderating effects provide further evidence that the search feature of the mobile channel is highly likely to affect sales concentration.

7. Implications of Mobile Channel Search on Inventory Management

Our empirical analysis shows that the mobile channel, compared to the conventional PC channel, results in a significant increase in sales concentration. Managers are vulnerable to do suboptimal inventory stocking if they fail to account for this key difference between the two online channels. Below, we quantify the extent of such suboptimal stocking under different scenarios and products.

Our analysis builds on the relationship between sales dispersion (opposite of sales concentration) and optimal inventory level first developed by Gallino et al. (2017). Although actual changes in inventory levels are co-determined by a gamut of supply-side, demand-side, and ordering policy factors, Gallino et al.'s approach provides an insightful lens to evaluate the implication of a change in sales concentration level on inventory investment. Next, we first expand on this approach and later combine it with our collaborating retailer's product parameters to quantify the extent of suboptimal inventory stocking.

Sales Concentration and Total Cycle Inventory We assume that the retailer offers N_c products in a product category c. For the product i in category c, the retailer expects mean demand μ_{ic} and incurs a category invariant fixed holding h_c , and ordering K_c cost. Furthermore, assume that the retailer orders cycle inventory as per the Economic Order Quantity (EOQ) formula. Collectively, these assumptions imply that the retailer's average total cycle inventory I_c for category c can be expressed as $I_c = \sum_{i=1}^{N_c} 1/2\sqrt{2K_c\mu_{ic}/h_c}$. Next, we substitute mean demand μ_{ic} by mean sales s_{ic}, which we assume follows the Pareto model. Without loss of generality, a product's index i is set equal to its popularity rank. Thus, mean sales of a product i in category c equals $s_{ic} = a_c i^{b_c}$ where parameters a_c and b_c respectively denote sales scale and dispersion factor. This reduces the average total cycle inventory expression to

$$I_c = \sqrt{\frac{K_c a_c}{2h_c}} \sum_{i=1}^{N_c} \sqrt{i_{ic}^{b_c}}. \tag{3}$$

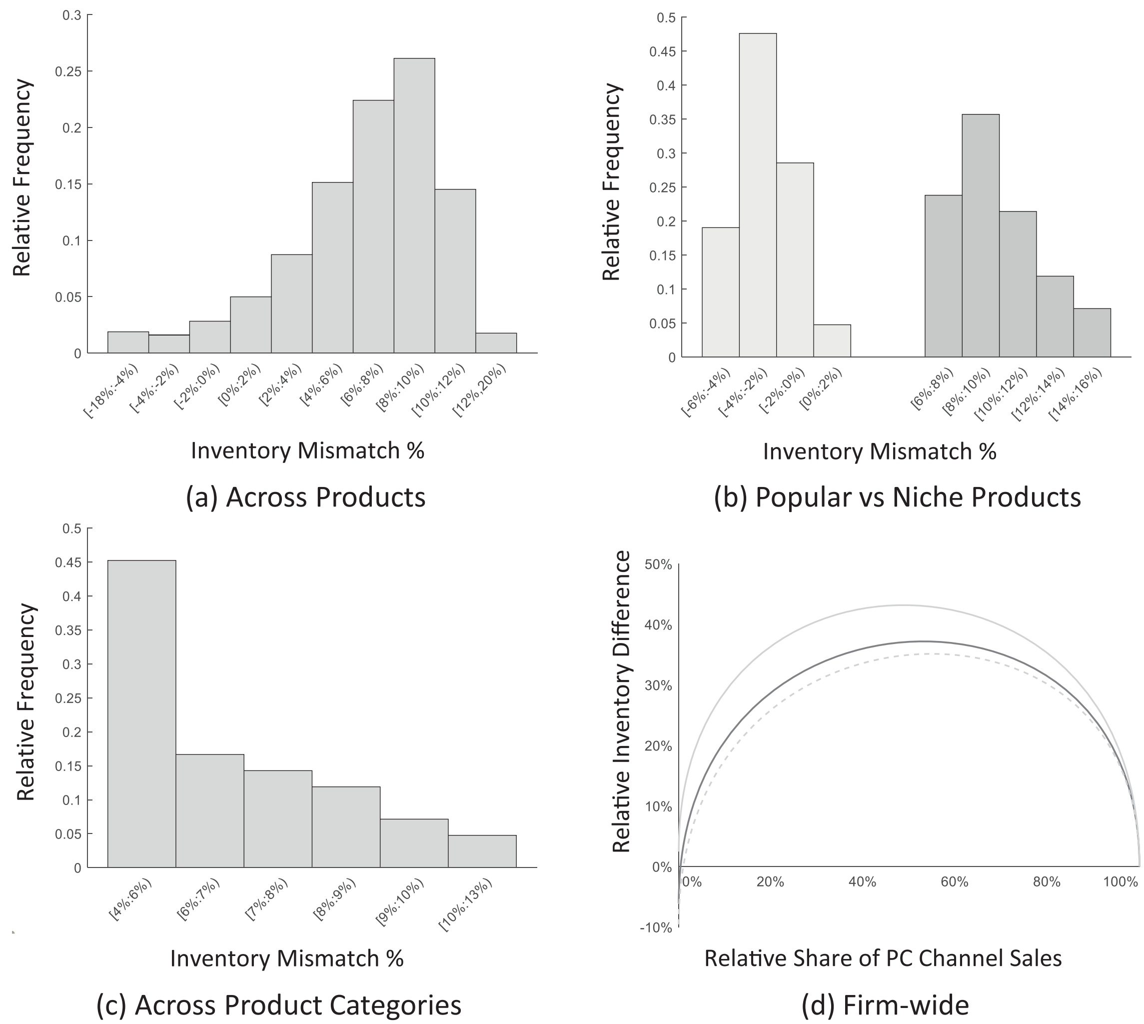
We use eq(3) to compute a change in inventory due to a change in sales dispersion level. More generally, consider that an introduction of a business practice (such as ship-to-store or sales via the mobile channel) results in a change of category c's total sales (from $s_c$to s'_c), and its sales dispersion (from $b_c$to b'_c). Let $s_{c'} = s_c(1+\gamma_c)$ where γ_c captures the extent of growth or contraction in total sales. Gallino et al. (2017) show that the associated change in category c's inventory level is given by

$$(\Delta I/I)(N_c, b_c, b'_c) = \sqrt{1+\gamma_c}\left(\frac{\sum_{i=1}^{N_c}\sqrt{i^{b'_c}}}{\sqrt{\sum_{i=1}^{N_c} i^{b'_c}}}\right)\left(\frac{\sum_{i=1}^{N_c}\sqrt{i^{b_c}}}{\sqrt{\sum_{i=1}^{N_c} i^{b_c}}}\right)^{-1} - 1. \tag{4}$$

We link the inventory change expression based on sales-dispersion factors (eq (4)) to the estimates of change in the share of popular and niche products in two steps. In the first step, we compute the share of product group $g(\in \{$ top 10%, top 20%, bottom 20%, bottom 10%$\})$ in purchases over channel $l\,(\in \{$PC, mobile$\})$ which we denote by $S_c^{g,l}$. We measure $S_c^{g,\mathrm{PC}}$as the share of group $g's$ products in the PC channel's pre-shutdown purchases. Next, we set $S_c^{g,\mathrm{mobile}} = S_c^{g,\mathrm{web}} + \alpha_1^g$ where α_1^g is the estimate of the mobile channel's impact on the share of product group g (as reported in Table 4). In the second step, we use these computed shares to infer the underlying sales dispersion factor for the channel l, denoted by b_c^l. In particular, we compute b_c^l as the average of group-specific dispersion factors $b_c^{g,l}$ inferred by equating group g's channel l share to the one derived from the Pareto model, i.e. $S_c^{g,l} = \sum_{i\in g} i^{b_c^{g,l}} / \sum_{i\in\{\text{All Products}\}} i^{b_c^{g,l}}$.

Implications of the Mobile Channel for Total Cycle Inventory Figure 5 shows the mismatch in inventory levels if a manager ignores that the mobile channel's impact on sales concentration significantly differs from that of the PC channel. We calibrate eq (4) using our collaborator's assortment size numbers (N_c for different product categories). Further, we assume that the manager makes independent stocking decisions for each of the product categories. Panel (a) shows that such an ignorance can lead to both understocking and overstocking of inventory. The inventory mismatch as a percentage of the optimal inventory varies from -17.9% to 19.2%. The mismatch direction is contingent on the product's popularity. Panel (b) shows that managers are prone to understock the popular products (top 10%) as the mobile channel's sales is more concentrated on the popular products compared to that in the PC channel. As a result, when the manager mistakenly treats the mobile channel the same as the PC channel, he/she is vulnerable to under account for the popular products' sales. Likewise, the manager may over account for the niche products' (bottom 10%) sales and, thus may overstock those products. Panel (c) shows the extent of inventory mismatch in our retailer's product categories, by equally weighting instances of a unit of inventory understock and overstock across all products within a category. Across all product categories, we find that instances of overstocking are higher than that of understocking, resulting in a net overinvestment in inventory (minimum: 4.2%, average: 6.6%, maximum: 12.9%).

From a firm-wide point of view, the extent of such an inventory mismatch naturally depends on the relative sales size of the two channels. In panel (d), the x-axis captures the share of PC-channel sales in total sales across the two channels. The y-axis captures the corresponding relative difference in firm-wide inventory levels between considering and neglecting the mobile channel's impact on sales concentration. We capture the mismatch in inventory levels of a product category with average number of products (1,386) across categories in our data. The solid black curve, solid grey cruve and dash grey curve respectively capture the average mismatch extent across all the



Notes. The figure illustrates the extent of inventory mismatch if a manager ignores the mobile channel's impact on sales concentration. Panel (a) shows the extent of inventory mismatch at the product level. Panel (b) splits inventory mismatch by product type: popular and niche. Panel (c) shows the mismatch at the product-category level. Panel (d) shows the firm-level mismatch as the relative share of the PC-channel sales.

Figure 5 Mobile Mimics PC: Overinvestment In Inventory Levels

products, popular, and niche products. Consistent with panel (c), we find that when the PC-channel share is 0% the manager is prone to overinvest in inventory when the manager plans for the firm-wide inventory based on the PC channel's sales concentration. Interestingly, we find that even with a moderate share of the PC channel, the manager may underinvest in inventory at the firm level. The mistaken application of the PC channel's lower sales concentration on firm-wide sales results in under accounting of a product's sales. This under accounting increases with the increase in the PC channel's share up to a threshold before reversing the trend; thus resulting in a U-curve relationship between the PC channel's sales share and the extent of inventory mismatch.

8. Conclusion

Our paper studies how the two primary online channels – mobile and PC – differ with respect to sales concentration, a metric that is highly relevant in making various operational decisions, including assortment planning and inventory investment (Gallino et al. 2017). Although the PC and mobile channels are both online channels, the mobile channel's search environment is quite different from of the PC channel's. The mobile channel's easy access increases the time available for product search and evaluation, which, in effect, increases the customer's search budget. However, the mobile channel's restrictive search features (e.g., small screen size, single-tab browsing) increases the search cost per searched item. The net effect of these two mobile channel search features determines the mobile channel's effect on sales concentration. Using a customer-level difference-in-differences estimation strategy, we find that the mobile channel increases sales concentration as compared with the PC channel, in the studied apparel context. Our study complements Doosti et al. (2018) by documenting that the mobile channel's impact on sales concentration is contingent on the product category's nature. In the apparel product categories, the mobile channel's 'higher search cost' effect seems to outweigh its 'larger search budget' effect, thus increasing sales concentration. Moreover, we find that the mobile channel's effect on a product category's sales concentration is moderated by category-specific features such as the availability of complementary information about the products, and product feature variety.

These findings make important contributions to research and practice. Our empirical results help fill an important gap in our knowledge of how the fast-growing mobile channel affects the purchase content aspect of customer behavior – more specifically, the sales distribution across popular and niche products. An accurate understanding of the mobile channel's effect on sales concentration is important for online retail managers too. The presence of a mobile channel can significantly alter the overall sales concentration for their businesses. We find that not accounting for the mobile channel's effect in increasing sales concentration can result in an inventory suboptimal stocking at the product category level between 4.2% and 12.9%. Managers therefore must carefully adopt innovative strategies for their online channels in order to maximize the value of internet-based information technology. For example, retailers can use assortment rotation or dynamic product display over the mobile channel to mitigate the negative implications of the channel's high search cost (Ferreira and Goh 2018).

Our study can be extended along different dimensions. First, our data set does not include details about the customers' browsing history or how the retailer ranks display of products. Therefore, we cannot disentangle the individual effects of our channel search features: search budget and search cost. Isolating these effects should be a fruitful avenue for future research. Second, our collaborating retailer had minimal presence in a growing online channel: tablets. Compared to the mobile channel, tablets offer limited freedom in accessibility but also features that lower search cost (such as larger screen size). Thus, it is not clear whether tablets would emulate the mobile or PC channel with respect to impact on sales concentration. This is a question that calls for empirical examination. Furthermore, we study a single retailer in a single industry in one country. This retailer is one of the largest apparel retailers in India, a country with widespread mobile phone adoption, so the observed finding can be considered as representative of the general effect of the mobile channel on purchases in the apparel category. Although data collection from multiple retailers is difficult, a deeper examination of the mobile channel's effect on sales concentration, especially across product categories, will require sampling across the mobile channels of more retailers. Finally, we only analyze safety/cycle inventory as an initial step to show the significance of the mobile channel for operations management. Abundant future research opportunities exist to examine the implications of increased sales concentration due to the mobile channel for other decisions, including assortment, fulfillment and pricing.

References

Alptekinoğlu A, Corbett CJ (2010) Leadtime-variety tradeoff in product differentiation. *Manufacturing & Service Operations Management* 12(4):569–582.

Anderson C (2006) *The Long Tail: Why the Future of Business Is Selling Less of More* (New York: Hyperion).

Angrist JD, Pischke J (2008) *Mostly harmless econometrics: An empiricist's companion* (Princeton university press).

Ariely D (2016) Time Pressure: Behavioral Science Considerations for Mobile Marketing. Accessed November 11, 2017, http://bit.ly/2Apm9PL.

Arnould EJ (1989) Toward a broadened theory of preference formation and the diffusion of innovations: Cases from zinder province, niger republic. *Journal of Consumer Research* 16(2):239–267.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *The Quarterly journal of economics* 119(1):249–275.

Bimpikis K, Markakis MG (2015) Inventory pooling under heavy-tailed demand. *Management Science* .

Brynjolfsson E, Hu Y, Simester D (2011) Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science* 57(8):1373–1386.

Chen C, Jain N, Yang SA (2020) The impact of trade credit provision on retail inventory: An empirical investigation using synthetic controls. *Available at SSRN 3375922* .

Cui R, Zhang DJ, Bassamboo A (2018) Learning from inventory availability information: Evidence from field experiments on amazon. *Management Science* 65(3):1216–1235.

Doosti S, Wang Y, Tan Y (2018) Do mobile applications bring longer tail? an empirical study of sales concentration in online channels. *Available at SSRN: https://ssrn.com/abstract=3255101* .

Ertekin N, Agrawal A (2019) How does a return period policy change affect multichannel retailer profitability? *Manufacturing & Service Operations Management* (forthcoming).

Fazio RH, Ledbetter JE, Towles-Schwen T (2000) On the costs of accessible attitudes: detecting that the attitude object has changed. *Journal of personality and social psychology* 78(2):197.

Ferreira K, Goh J (2018) Assortment rotation and the value of concealment. *Harvard Business School Technology & Operations Mgt. Unit Working Paper* (17-041).

Frank RH, Cook PJ (2010) *The winner-take-all society: Why the few at the top get so much more than the rest of us* (Random House).

Gallino S, Moreno A, Stamatopoulos I (2017) Channel integration, sales dispersion, and inventory management. *Management Science* 63(9):2813–2831.

Gaur V, Honhon D (2006) Assortment planning and inventory decisions under a locational choice model. *Management Science* 52(10):1528–1543.

Ghose A (2017) *Tap: Unlocking the Mobile Economy* (MIT Press).

Girotra K, Netessine S (2013) Om forum—business model innovation for sustainability. *Manufacturing & Service Operations Management* 15(4):537–544.

Hendricks KB, Singhal VR (2009) Demand-supply mismatches and stock market reaction: Evidence from excess inventory announcements. *Manufacturing & Service Operations Management* 11(3):509–524.

Iacus SM, King G, Porro G, Katz JN (2012) Causal inference without balance checking: Coarsened exact matching. *Political analysis* 1–24.

Kesavan S, Mani V (2013) The relationship between abnormal inventory growth and future earnings for us public retailers. *Manufacturing & Service Operations Management* 15(1):6–23.

King G, Nielsen R, et al. (2016) Why propensity scores should not be used for matching. *Copy at http://j. mp/1sexgVw Download Citation BibTex Tagged XML Download Paper* 378.

Lee HL, Tang CS (1997) Modelling the costs and benefits of delayed product differentiation. *Management science* 43(1):40–53.

Lee HS, Kesavan S, Deshpande V (2017) Understanding and managing customer-induced negative externalities in congested self-service environments. Technical report, Working paper.

Meola A (2016) The rise of m-commerce: Mobile shopping stats and trends. *Business Insider* .

Moreno A, Terwiesch C (2016) The effects of product line breadth: Evidence from the automotive industry. *Marketing Science* 36(2):254–271.

Mulpuru S, Boutan V, Johnson C, Wu S, Naparstek L (2015) Forrester Research Ecommerce Forecast, 2014 to 2019 (US). *Forrester Research, Inc., Cambridge, MA, USA* .

Narang U, Shankar V (2016) The effects of mobile apps on shopper purchases and product returns. *Working Paper* .

Netessine S, Dobson G, Shumsky RA (2002) Flexible service capacity: Optimal investment and the impact of demand correlation. *Operations Research* 50(2):375–388.

Olivares M, Cachon GP (2009) Competing retailers and inventory: An empirical investigation of general motors' dealerships in isolated us markets. *Management science* 55(9):1586–1604.

Orendorff A (2019) The state of the ecommerce fashion industry: Statistics, trends & strategy. Shopifyplus. Available at: http://bit.ly/2KY74hB. Accessed on February 20, 2019.

Petersen MA (2009) Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22(1):435–480.

Pierce L, Snow DC, McAfee A (2015) Cleaning house: The impact of information technology monitoring on employee theft and productivity. *Management Science* 61(10):2299–2319.

Shankar V, Venkatesh A, Hofacker C, Naik P (2010) Mobile marketing in the retailing environment: current insights and future research avenues. *Journal of interactive marketing* 24(2):111–120.

Simchi-Levi D (2010) *Operations rules: delivering customer value through flexible operations* (Mit Press).

Simonson I (1990) The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research* 27(2):150–162.

Soysal G, Zentner A (2014) Measuring e-commerce concentration effects when product popularity is channel-specific. *Working Paper* .

Srivastava M (2016) Mobile phones to dominate online sales medium in india: report. URL `http://bit.ly/2VFBLs3`, accessed on September 20, 2018.

Stigler G (1961) The economics of information. *The Journal of Political Economy* 69(3):213–225.

Tan FT, Netessine S (2020) At your service on the table: Impact of tabletop technology on restaurant performance. *Management Science* 66(10):4496–4515.

Tan FT, Netessine S, Hitt LM (2017) Is tom cruise threatenened? an empirical study of the impact of product variety on demand concentration. *Information Systems Research* 28(3):643–660.

Wang RJ, Malthouse EC, Krishnamurthi L (2015) On the go: How mobile shopping affects customer purchase behavior. *Journal of Retailing* 91(2):217–234.

Wooldridge JM (2015) *Introductory econometrics: A modern approach* (Nelson Education).

Xu K, Chan J, Ghose A, Han SP (2016) Battle of the channels: The impact of tablets on digital commerce. *Management Science* 63(5):1469–1492.

Appendices for "M-Commerce, Sales Concentration and Inventory Management"

A1. Additional Figures and Tables

No	Attribute Name	Population Count	Sample Count	No	Attribute Name	Population Count	Sample Count
1	Number of Transactions	9 million	1.8 million	7	Pre-Shutdown: App Share (Transactions)	37.57%	37.31%
2	Number of Customers	814,000	162,800	8	Pre-Shutdown: Order Size (amount)[4]	(1, 1.548, 2.439)	(1.001, 1.548, 2.440)
3	Apparel Product-Category[1]	7	7	9	Pre-Shutdown: Number of Orders per customer[4]	(3, 7, 14)	(3, 7, 14)
4	Article Type[2]	99	95[3]	10	Pre-Shutdown: Discount Availed[4]	(1, 1.644, 1.977)	(1, 1.644, 1.977)
5	Share of Tier 1 + Tier 2 cities	76.81%	76.80%	11	Pre-Shutdown: Average Share of Transactions During Office Hours	0.412	0.413
6	Share of Female Customers	57.26%	57.40%	12	Pre-Shutdown: Average Share of Transactions During Weekend	0.315	0.316

1: Describes the top-level apparel product-category; examples include Bottomwear, Topwear

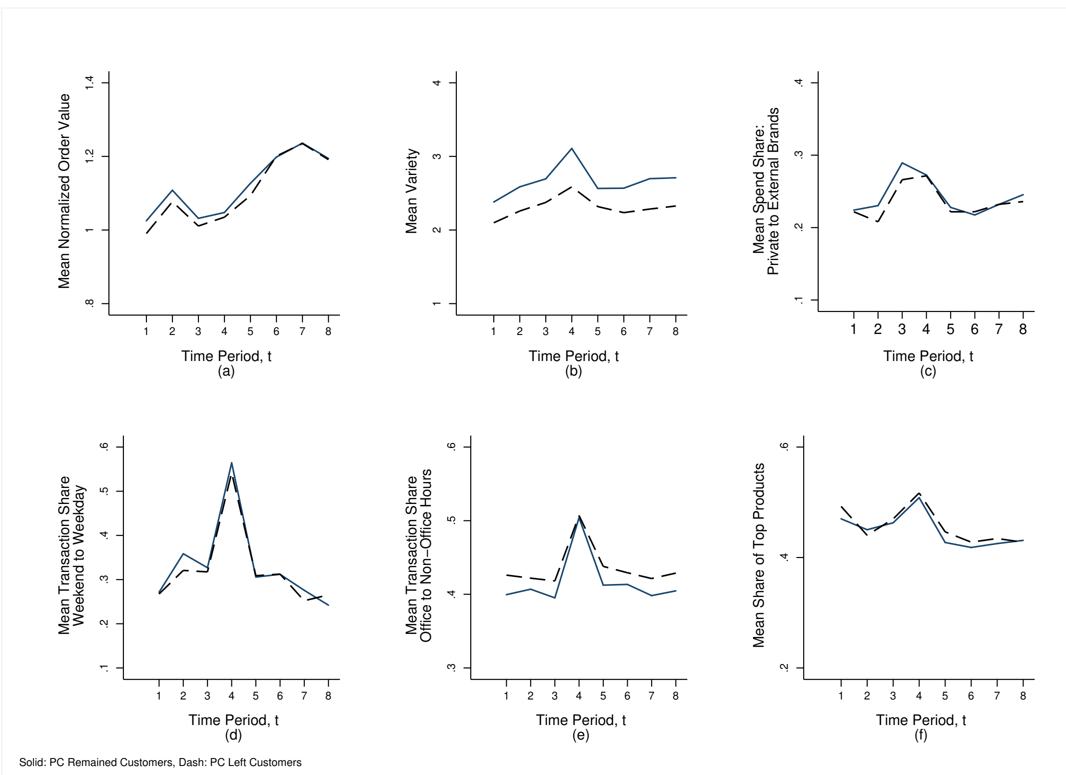
2: Describes the apparel sub product category; examples include Capris, Jeans, and Dhotis within the Bottomwear category, and Tshirts, Waistcoat, and Lehenga Choli within the Topwear Category

3: The omitted four article types collectively account for 37 transactions in the underlying population data

4: Numbers in brackets denote 25th, 50th and 75th percentile values. To accommodate the retailer's concern on sharing actual numbers, we normalized all numbers by the 25th percentile value.

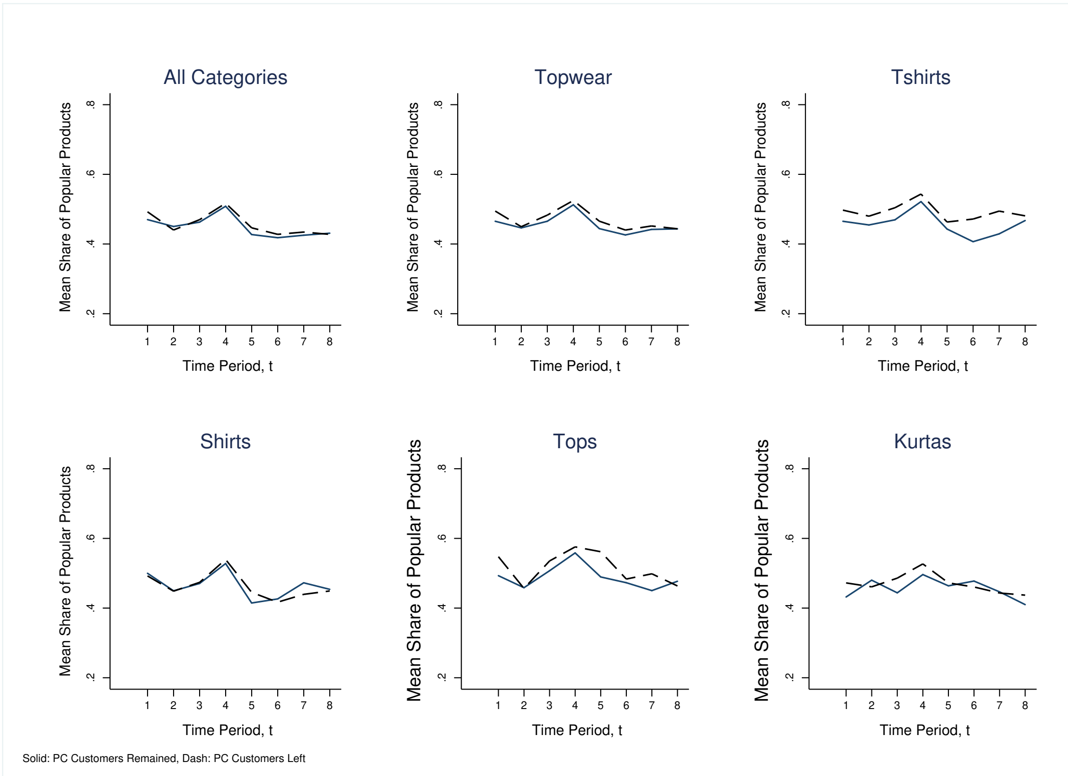
Notes. The population data captures all transactions between 15 January through to 31 December 2015. The 20% random sample consists of customers who satisfy the following three conditions:(i) transacted at least once in the pre-shutdown period, (ii) no missing information on demographics and the first purchase date, and (iii) joined the retailer before the start of the study period.

Table A1 Population vs Sample: Key Statistics



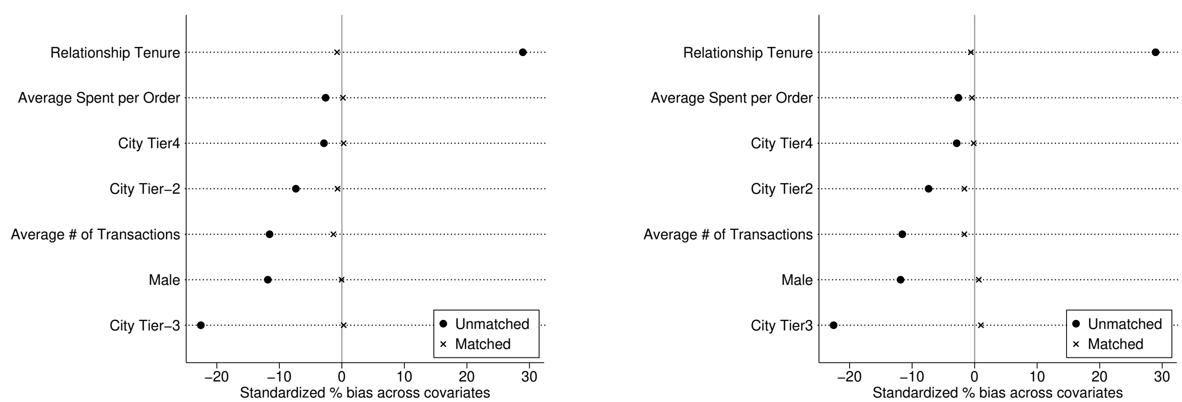
Notes. The x-axis shows the pre-shutdown period t. In panel(a), the y-axis shows the mean of order value, normalized by the mean order size. In panel(b), the y-axis shows the average number of purchased distinct styles. In panel(c), the y-axis shows the mean ratio of the amount spent on the private brands to the amount spent on the external brands. The y-axis in panel (d) (resp. panel (e)) captures the mean of the ratio of transactions completed during office hours (resp. weekend) to those completed during non-office hours (resp. weekdays). In panel(f), the y-axis captures the average share of the top products in the matched sample customers' purchases.

Figure A1 The PC Customers' Purchase Behavior: Treated vs Discountinued



Notes. The x-axis shows the pre-shutdown period t. The y-axis captures the average share of the top products in the matched sample customers' purchases.

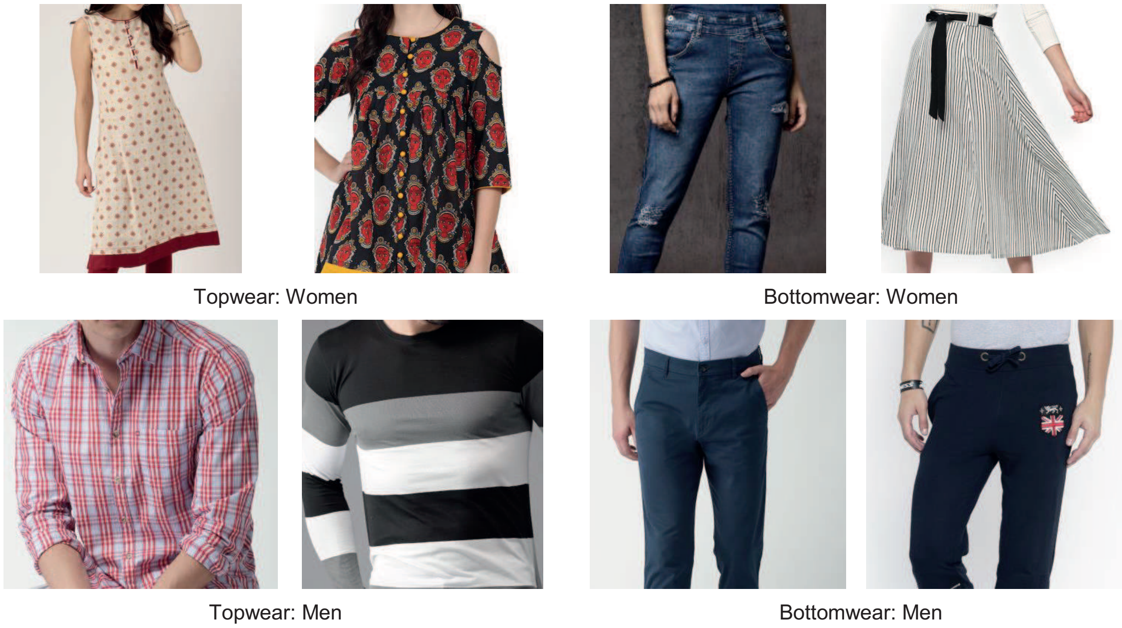
Figure A2 Sales Concentration Level in Purchases by Matched PC-channel customers: Treated vs Discontinued Customers



(a) Nearest Neighbor - 1 (b) Nearest Neighbor - 1

Notes. The figure shows standardized bias in the unmatched and matched sample under the two Propensity Score Matching (PSM) approaches. Panel (a) (resp. panel (b) shows bias in the N(1) (resp. N(3)) matched sample. In both these approaches we set the maximum threshold distance to 0.01.

Figure A3 PSM-Matched Samples: Covariates Balance



Notes. The figure shows example of product-variety in the Topwear and Bottomwear product-category.

Figure A4 Style Examples by Product Category

A2. Validating Difference-in-Differences (DiD) Estimation: Parallel Trend Assumption

Following Gallino et al. (2017), we formally test for the parallel trend assumption using the pre-shutdown period observations with the following specification:

$$S_{it} = \alpha'_0 + \alpha'_1 Discount + \alpha'_2 TimeUnit_{index} \times PC + \mu_i + \tau_t + \varepsilon_{it}, \quad \text{(A1)}$$

where TimeUnit$_{index} = 1 \ldots 8$ denotes the order of specific time unit in the pre-shutdown period and the coefficient α'_2 tests for any trend in difference between sales concentration level of the PC (treated group) and mobile (control group) customers. We find $\alpha_2 = -1.6 \times 10^{-3}$ to be insignificant at the p-value of 0.611. Thus, we fail to reject the null hypothesis of the two groups having a constant difference in the outcome variable (equivalently parallel trend) during the pre-shutdown period.

A3. Additional Aggregate Level Analysis: Gini Coefficient and Log-Linear Sales Model

We perform the Gini coefficient, and Log-Linear sales model estimation analysis using the pre-shutdown period sample. Table A2 shows estimates of the eight Gini coefficient analyses with a sample of all categories, and of top three categories that cumulatively account for 90.1% of transactions in the pre-shutdown period. A higher value of Gini coefficient indicates greater inequality in sales of different products or, equivalently, a higher sales concentration level. We present estimates of specifications with (see columns (1), (3), (5), and (7)) and without the control for the lagged Gini coefficient value (see columns (2), (4), (6), and (8)). Across all these eight tests, we find that the Gini coefficient value is significantly higher for the sales over the Mobile channel.

Table A3 shows estimates of the four Log-Linear model estimation analyses with a sample of all categories, and of top three categories. In all these four tests, we consistently find that the coefficient of Mobile Channel×Log(Sales Rank) is significant and negative in sign. This implies that the lower rank products accounts for lower sales on the mobile channel compared to that on the PC channel. Collectively, similar to the Lorenz Curves figures, we continue to find consistent support for significantly higher level of sales concentration in the mobile channel compared to that in the PC channel.

	Gini Coefficient							
	All Categories		Topwear		Bottomwear		Innerwear	
	Gini Value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mobile Channel	0.045***	0.037***	0.050***	0.039***	0.052***	0.046***	0.038***	0.030***
	(0.004)	(0.004)	(0.007)	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)
Gini Lagged Value		0.190***		0.129***		0.215***		0.187***
		(0.030)		(0.059)		(0.060)		(0.056)
Product Category Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	1388	1204	355	308	371	323	378	327
Adj-R sq	83.1%	82.8%	85.5%	85.3%	81.6%	81.0%	80.7%	81.3%

*p-values: *** < 1% ** < 5% * < 10%*

Notes. The unit of analysis is product-category × time-unit × channel. Product-category is defined at the article-level. The analysis use the pre-shutdown period transactions during which both the PC chand mobile channels were operational.

Table A2 Aggregate Level Analysis: Gini Coefficients Results

A4. Alternate Identification Using Omnichannel Customers

Using the omnichannel customers subsample, we identify the mobile channel's impact on sales concentration by exploiting such customers' varying mobile exposure across product categories. We measure a customer i's mobile channel exposure ME in product category c using his/her pre-shutdown purchase transactions. Formally, we set $ME_{ic} = N_{ic}^{\text{mobile}}/N_{ic}^{\text{all}}$ where N_{ic}^{mobile} captures the number of pre-shutdown purchases made over the mobile channel and N_{ic}^{all} the total number of pre-shutdown purchases. Leveraging this measure, we define treated and control groups at the customer×product category level (i,c). We re-estimate the main analysis DiD specification (eq (1)) with two alternate definitions of treatment. One, we define all (i,c) pairs with $ME = 0$ as

	Log-Linear Sales Model			
	All Categories	Topwear	Bottomwear	Innerwear
	Log(Sales)			
	(1)	(2)	(3)	(4)
Mobile Channel x Log(Sales Rank)	-0.032***	-0.044***	-0.045***	-0.031***
	(0.0004)	(0.0007)	(0.0011)	(0.0021)
Log(Sales Rank)	-0.949***	-0.961***	-0.969***	-0.908***
	(0.0004)	(0.0006)	(0.0010)	(0.0019)
Mobile Channel	0.580***	0.695***	0.662***	0.546***
	(0.0032)	(0.0055)	(0.0073)	(0.0113)
Product Category Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Sample Size	1,129,121	699,765	193,391	73,508
Adj-R sq	88.1%	87.0%	81.0%	84.5%

*p-values: *** < 1% ** < 5% * < 10%*

Notes. The unit of analysis is product $\times$ time-unit $\times$ channel. The analysis use the pre-shutdown period transactions during which both the PC chand mobile channels were operational. Products are ranked in descending order of their sales value, within their respective category that is defined at the article-type level i.e., product with highest sales is assigned rank 1.

Table A3 Aggregate Level Analysis: Log-Linear Sales Model Results

treated groups and with $ME = 1$ as the control groups. In the second definition, we define all (i, c) pairs with $ME \leq 25^{th}$percentile (ME) as treated groups and with $ME \geq 75^{th}$percentile as the control groups. Customers not belonging to either of the groups are dropped from the regression. Under both definitions, we continue to find a positive and significant estimate for the DiD coefficient α_1 with values of 0.0532^{***} (former definition, # Obs = 106,206, # Customers = 16,864, # Customer$\times$Product-Category = 51,854) and 0.0123^{***} (latter definition, # Obs = 256,082, # Customers = 16,864, # Customer$\times$Product-Category = 51,854). These results suggest that, despite having prior exposure to the mobile channel's shopping environment, a customer exhibits higher propensity to purchase popular products when shifted to the mobile channel due to the shutdown event. These results also support our focal finding that the mobile channel significantly increases the sales concentration. Such an increase is most likely due to the channel's embedded search features rather than lacking exposure to the mobile channel's shopping environment.

A5. Evidence Supporting the Effect of High Search Costs

When search costs increase, one would expect customers to shorten their search process and to select products that are displayed higher ($\equiv$ higher display ranking) on a search results page. Although we have no direct information on products' display rankings, we know from the channel

	(1)	(2)	(3)
		Age Threshold	
Estimation Description	$\log(\bar{A})$	$\tilde{A} = 50^{th}$ percentile	$\tilde{A} = 60$ days
PC×Post	-0.017^{**} (0.008)	0.015^{***} (0.005)	0.011^{***} (0.005)
Discount	0.179^{***} (0.003)	-0.100^{***} (0.002)	-0.102^{***} (0.002)
Time Fixed Effects	Yes	Yes	Yes
Customer Fixed Effects	Yes	Yes	Yes
Average Age/ Share of Recent Products	78 days	54%	44%
# Observations†	155,367	155,367	155,367
# Customers	39,750	39,750	39,750
Prob > Chi-sq	< 0.001%	<0.001%	<0.001%

Notes. The unit of analysis is customer×time-unit. Standard errors in the parantheses are bootstrapped errors.
*** p < 1%, ** < 5%, * < 10%

†Sample size is less than #Customers×#Periods since not all customers purchased in each period.

Table A4 Channel Effect on Age of Purchased Products

managers that recently introduced products are typically displayed higher in a search results page. Such pages are generated based on customers' applied filtering on user-specified attributes such as product category and price range. We exploit this observation to examine channels' search costs by studying the proportions of new products in customers' purchases. If purchases made over the mobile channel exhibit a higher share of newly introduced products (than do PC-channel purchases), then this tendency will corroborate the conjectured higher search cost mechanism (and thus increased sales concentration) in the mobile channel. In effect, the higher cost resulting from the mobile environment's embedded search features discourages customers from searching for 'back catalog' products.

To execute this strategy, we first measure the age (in days) of product p during period t (A_{pt}) as the difference between the product's introduction date and the last day in the period. Since a product can be introduced within a period, our age definition, based on the last day's date, ensures a non-negative value for age ($A_{pt} \geq 0$). We work with products that have, at most, six months of age to avoid concerns of products' cataloging recording error. Next, we replace the dependent variable in our DiD analysis, equation (1), with two types of dependent variables related to product age: (i) the log of average age of purchased products ($\log(\bar{A})_{it}$) made by customer i during period t ($\log(\bar{A})$ is normally distributed); and (ii) the share of recently introduced products ($\bar{R}_{it}$). Specifically, $\bar{A}_{it}$ and $\bar{R}_{it}$ are defined as $\bar{A}_{it} = \frac{\sum_{k \in \mathbb{N}_{it}} A_{kt}}{|\mathbb{N}_{it}|}$ and $\bar{R}_{it} = \frac{\sum_{k \in \mathbb{N}_{it}} \mathbb{I}(A_{kt} \leq \tilde{\mathtt{A}})}{|\mathbb{N}_{it}|}$, where $\mathbb{N}_{it}$ is the set of products purchased by customer i in period t; $|\cdot|$ is the set count operator, and $\mathbb{I}(\cdot)$ is an indicator function that returns 1 only if the age of product p in period t is not more than an age threshold $\tilde{\mathtt{A}}$ used to distinguish between recent and back-catalog products.

Table A4 reports the estimation results for these outcome variables. Column 1 presents results for the average aged- based outcome variable. We find that the average age of products purchased via the mobile channel is significantly lower than via the PC channel by $\approx 1.7\%$. Columns 2–3 of the table show the results with two alternate definitions of the share of recently introduced products outcome variable, R. In column 2, we measure R using a relative age definition by setting age threshold $\tilde{A}$ to 50^{th} percentile of the age distribution observed during the study period. In column 3, we test with an absolute age definition by setting $\tilde{A}$ to 60 days. Results for both these thresholds are significant and positive (with respective coefficients 0.015 and 0.011), suggesting that the share of relatively new products is significantly higher among purchases made over the mobile channel than over the PC channel.

A6. Additional Robustness Test: Pareto Slope as an Outcome Variable

In this section, we examine the impact of the mobile channel on sales concentration using the slope of the Pareto curve regression—a commonly used sales dispersion measure (Brynjolfsson et al. 2011, Gallino et al. 2017)—as the outcome variable. The outcome variable in the main analysis provides a measure of sales concentration using the share of a specific product group (popular products) in total sales. In contrast, the Pareto curve's slope (henceforth referred to as the Pareto measure) provides a sales dispersion measure that accounts for sales contribution of the full spectrum of products i.e., including the popular as well as niche products.

We estimate the Pareto measure by fitting a power law of the form $y = \alpha x^{\beta}$ to the observed sales in each product-class$\times$time$\times$channel unit. Here, x denotes the product-ranking with x=1 representing the highest-selling product, and y denotes the product $x's$ sales in period t and channel $c \in \{\text{PC, Mobile}\}$. We set product-class a at the article-type level. Further, in line with the main analysis, we set the product at the style level and time period at the biweekly level. The parameter β is the Pareto slope parameter that captures dispersion in sales of products within a product-class and α is the associated scale parameter. For each product-class$\times$time$\times$channel unit ($a \times t \times c$), we fit the Pareto curve on the observed sales to estimate the respective Pareto measure, $PARETO_{atc}$. We consider units with a minimum of 100 sales transactions so as to fit a Pareto curve.

	(1)	(2)
PCxPost	-0.004 (0.006)	-0.013** (0.006)
Product Class Fixed Effects	Yes	Yes
Channel Fixed Effects	Yes	Yes
Time Fixed Effects	Yes	Yes
Sample Size	769	2,106
Adj-R sq	75.5%	75%

*p-values: *** < 1% ** < 5% * < 10%*

Notes. The unit of analysis is product-class×time×channel. We define product-class at the article-type level, product at the style level, and time period at the biweekly level. For each observation unit, the outcome variable, the Pareto measure, is estimated by fitting a Pareto curve on the unit's observed sales. Columns (1) and (2) respectively show estimation results with the matched and unmatched sample.

Table A5 The Impact of Mobile Channel On Sales Dispersion: Pareto Slope

We estimate the mobile channel's impact on the Pareto measure using the following DiD specification: $PARETO_{atc} = \alpha_1 PC \times POST + \alpha_t + \alpha_a + \alpha_c + \varepsilon_{atc}$, where α_a, α_t, and α_c respectively captures the product-class, time, and channel fixed-effects. The coefficient of interest is α_1 that captures the treatment effect of the mobile channel's shutdown on the outcome variable. Table A5 shows the estimated treatment effects.

Columns 1 and 2 of Table A5 respectively show the results with the matched customers sample, and the unmatched customers (using all of the population data). We find that the sign of treatment effect (PC×POST) is negative under both analyses. This implies that the treatment resulted in a reduction of sales dispersion. In other words, the PC customers exhibit a higher sales concentration in their post-shutdown purchases than in pre-shutdown purchases. Under the matched customer sample analysis, with 729 observations, we find that the treatment effect is not significant. However, under the population data analysis, with a larger sample size of 2,106 observations, we find that the treatment effect is significant.