Why do Stores Drive Online Sales? Evidence of Underlying Mechanisms from a Multichannel Retailer

Anuj Kumar
Warrington College of Business
University of Florida
akumar1@ufl.edu

Amit Mehra
Naveen Jindal School of Management
University of Texas Dallas
Amit.Mehra@utdallas.edu

Subodha Kumar
Fox School of Business
Temple University
subodha@temple.edu

This Version
June 2018

Abstract

We utilize the event of store opening by a large apparel retailer and use customer-level data to estimate the effect of store presence on the online purchase behavior of its existing customers. We find that the retailer's store openings resulted in an increase in online purchases from such customers. Drawing on the Theory of Planned Behavior and Prospect Theory, we propose two mechanisms to explain this complementary effect of store presence on online purchases by existing customers. These mechanisms are the *store engagement effect* – customers making higher online purchases due to higher engagement from store interactions, and the *store return effect* - reduced risk of online purchase due to the option of store returns. We provide direct empirical evidence of these mechanisms on customer-level data. We further show that these effects increase as customers' distances from the retailer's store reduce due to the store openings. Our findings have significant implications for multichannel retailers.

Keywords: E-Commerce, Multichannel customer behavior. Omni-channel retail, Matching estimators, Average treatment effect.

1 Introduction

Retailers that have traditionally relied on the store channel are increasingly dependent on their online channels to deliver sales growth. For example, Walmart online sales increased by 30 percent compared to its overall sales growth of one percent in the first quarter of 2013.[1] Gap's online sales increased by 13 percent while their store sales declined in the last quarter of 2013 leading to a meager growth rate of 1.2 percent in the overall sales.[2] This trend has led many retailers to scale down their store presence while investing more in their online channels. Thus, Gap has closed down 20 percent of its stores over the years (McIntyre and Hess 2014) and Nordstrom is considering increased investments in its online channel (Cook 2014).

Therefore, it appears that retailers consider stores as a losing proposition and are placing their bets on the online channel. However, such an approach should not ignore the possibility that absence of a store may affect sales at the online channel. Indeed, recent work by Bell et al. (2015) and Avery et al. (2012) found that the physical channel complements sales on the online channel for multi-channel retailers.

A store may complement sales on the online channel due to two possible reasons. First, a retailer may acquire new customers due to its store presence in a geographical area. If some of these new customers purchase on the retailer's online channel as well, then the store will have a complementary effect on online sales. Second, the presence of a store may facilitate online transactions for existing customers, and thus, complementarity between store and online sales. While existing research by Avery et al. (2012) and Bell et al. (2015) point out that the store complements online sales, they cannot identify whether this effect is only due to purchases made by new customers, or also due to additional purchases made by existing customers.[3]

[1] http://news.walmart.com/news-archive/investors/walmart-reports-a-46-percent-increase-for-q1-eps-of-114-us-businesses-forecast-positive-comp-sales-for-q2-1820850
[2] http://www.internetretailer.com/2014/05/22/e-commerce-accounts-all-q1-growth-gap
[3] Avery et al. (2012) find that the number of existing customers purchasing from the online channel increases due to store presence, but they do not analyze sales from these customers.

These papers could not differentiate between the two effects because they analyzed sales data at the ZIP code level whereas one needs customer-level purchase data to separate these two factors.

Therefore, *our first research objective* focuses on identifying whether store presence increases online purchases made by existing customers of a multichannel retailer selling products rich in non-digital attributes (e.g. apparel). If this effect were economically significant, then it would be useful to understand the mechanisms that drive this store-facilitation effect for existing customers. An understanding of these mechanisms can help retailers design appropriate marketing policies to improve their online and overall sales.

The first possible mechanism through which a store may facilitate online sales is the *store engagement effect*. This effect arises because when customers visit the store, they may explore and evaluate the products and become interested in the retailer's offerings. This increased engagement with the retailer may translate into customers purchasing more from the store. However, customers may sometimes not purchase the products they liked during a store visit due to a variety of reasons. For example, they may not get their desired SKU (e.g., the size, style, or color of an apparel) in store or they may like to explore products from another retailer before making their purchase decision. We describe such possibilities in detail in Section 2 of the paper. In such cases, customers may utilize the convenience of retailer's online channel to purchase products they are interested in later, instead of making another store visit.

The second possible mechanism through which a store may facilitate online sales is the *store return effect.* This effect arises because customers cannot properly evaluate products with non-digital attributes on the online channel. Therefore, they incur a higher risk in purchasing such products if they conducted product evaluation solely on the online channel. However, if they do not like the products they purchased online, they can easily return those products at a nearby store. Thus, presence of a store could reduce customers' risk of online purchases, which, in turn, could lead to higher online purchases.

To our knowledge, no prior research has examined the underlying mechanisms through which the store channel of a retailer complements its online sales. An understanding of which of these mechanisms are operative can help managers take appropriate actions to leverage these mechanisms. For example, if the complementary effect of stores is rooted in the store engagement effect, then retailers should design strategies to increase store foot traffic, such as organizing special events in stores. In view of this, *our second research objective* is to identify the effectiveness of the mechanisms that could lead to a complementary effect of the store channel on online sales.

Finally, note that customers located at a relatively smaller distance from the store may visit the store more frequently, resulting in higher store engagement effect that may drive higher online purchases by such customers. Similarly, the store return effect for such customers may be higher because easy accessibility to stores may prompt these customers to purchase more online. Thus, evidence of higher store engagement and return effects for customers located relatively nearer to the store would provide additional support for the existence of these effects.

1.1 Experiment Design, Analysis, and Results

To examine the effect of store openings by a retailer on its online sales, we designed an experiment around the six new store openings in 2003 by a large retailer of fashion apparel, accessories, and home products in the U.S. Due to such store openings, while the distance from the nearest store significantly reduced for the retailer's existing customers living near the newly opened stores (*affected customers*), it remained unchanged for its customers living in other parts of the U.S. (*unaffected customers*). We selected the samples of affected and unaffected customers in such a way that the influence of other factors, such as sales tax incidence, price and non-price promotions, and retailer's shipping and return shipping policies on their purchase propensities were similar. We collected purchase and return data on the online and store channels over four years for the selected sample of customers and used a difference-in-difference experimental design around the event of retailer's store openings to estimate the treatment effect of store openings on the online purchase behavior of its existing customers. We further accounted for the differences between the samples of affected and unaffected customers with a fixed effect and a propensity score matched panel data regression specifications. In the matching specifications, we matched the two categories of customers on their individual-level purchase data in pre-store opening period, their individual socioeconomic factors, and their aggregate zip code-level socioeconomic and demographic factors. Besides the parametric propensity score matching, we also estimated our treatment effect from the nonparametric coarsened exact matching estimator to show the robustness of our findings.

Our treatment effect estimates reveal that the retailer's store openings resulted in a complementary effect on the online purchases of its existing customers. Specifically, for our sample of customers, we find that: (1) online purchase probability increased by 0.029 (36 percent) over the mean value of 0.08; (2) number of items purchased per year on the online channel increased by 0.056 (14 percent) over the mean value of 0.4; and (3) annual online purchase revenue increased by US $6.59 (29 percent) over the mean value of US $22.4.

We further provide empirical evidence of the mechanisms for complementary effect of stores on customers' online purchases. We find an increase in percentage of customers making their online purchases with store interactions due to store openings, which indicates the existence of store engagement effect. Specifically, the probability of store interactions for customers making online purchases increases by 0.11 (25 percent) due to store opening over the pre-store opening period mean value of 0.44. We also find that customers return higher percentage of their online purchases in store after the store opening, which indicates the existence of store return effect. Specifically, customers return an additional 2.76 percent (113 percent) of their total online purchases in store after the store opening over the pre-store opening period mean value of 2.45 percent. We further find higher overall complementary, store engagement, and store return effects of store opening for those customers who experience greater reduction in their store distances, which indicates that the reduction in customers' store distances is the driver for these effects.

1.2 Contributions

This study makes several contributions to the literature. First, it combines the power of customer-level data with matching estimators in a quasi-experimental setting to make causal estimate of the effect of easier store access on the online purchase behavior of existing customers. Previous studies either used experimental setup with zip code-level aggregate data (Avery et al. 2012, Bell et al. 2015) or used customer-level data without an experimental setup (Kushwaha and Shankar 2013, Venkatesan et al. 2007).

Second, this is the first paper that uses customer-level purchase data to provide direct empirical evidence of the two mechanisms through which easier store access can increase the online purchases of existing customers. Prior studies could only conjecture that the retailer's store opening could increase customers' awareness about their products on aggregate zip code-level data (Avery et al. 2012, Bell et al. 2015). Similarly, prior studies have only examined how the option to return on a channel would affect the sales on that channel (Anderson et al. 2009, Petersen and Kumar 2009). This study demonstrates why facilitating returns on one channel (easier access to store channel) affects the demand on another channel (online channel). To our knowledge, this is the first study that provide evidence of cross-channel interactions between product returns and purchases.

The rest of the paper is organized as follows. We present the related literature and theory development in Section 2; describe our field setup, sample selection, and empirical design in Section 3; present our data and econometric specifications in Section 4; discuss our results and robustness checks in Section 5; and conclude with managerial implications and opportunities for future extensions in Section 6.

2. Related Literature and Theory Development

Prior literature suggests that the competition between store and online retailers are affected by three factors: relative prices between these retailers (Brynjolfsson and Smith 2000, Nelson et al 2012), offline transportation costs (Forman et al. 2009), and differences in offered product assortment between these retailers (Brynjolfsson et al. 2009, Choi and Bell 2011). Out of these factors, the customers' transportation costs to the store are reduced when a new store opens in their area. Thus, all else being equal, such store openings should decrease the sales of competing online retailers in that area. Forman et al. (2009) show that store entry decreases consumers' sensitivity to online product discount and lowers sales of the online retailer in the area of store entry. Brynjolfsson et al. (2009) show that with increase in the number of traditional store retailers of apparel in an area, sales of popular products significantly decrease for an internet apparel retailer but sales of its niche products remain unaffected. These works provide evidence of the substitutive effect of a retailer's store opening on the sales of its online competitors, which suggests that a multichannel retailer can expect a reduction in its online sales in an area where it opens a new store.

However, recent literature has provided evidence of a complementary effect on online sales due to a store opening. For example, Avery et al. (2012) show that the online sales of an apparel retailer increases in the areas where it opens its physical stores. Similarly, Bell et al. (2015) show that the online sales of an eyeglass retailer increases in the areas where it opens its inventory showrooms. Tang et al. (2016) show the possibility of both showrooming (complementary) and competing (substitutive) effects on the sales of an online footwear retailer in an area from the store openings of its competing retailers in that area. Therefore, we conclude from previous literature that a store opening may either increase or decrease online sales.

In the present study, we are interested in examining whether store openings of a multichannel retailer selling products rich in non-digital attributes, such as apparel, has a complementary effect on its online sales. Further, if the store has a complementary effect on online sales, this could be merely at the level of aggregate sales because of additional online and store purchases made by newly acquired customers, or this effect could be at the level of an individual customer. We wish to establish whether the complementary effect, if it exists, occurs in aggregate only, or at the individual customer level. Finally, if the complementary effect exists at the individual customer level, we want to examine the reasons why store openings may complement online sales.

A complementary effect of a store on the online sales for an individual customer may arise if customers plan future purchases at stores and then purchase later at the online channel. There is considerable evidence that consumers plan and research their apparel purchases (see footnote[4] and Doupnik 2017). Therefore, an application of the Theory of Planned Behavior (TPB, henceforth) in the context of multiple retail channels seems appropriate to investigate this issue of complementarity. Further, the information that customers gather during their research phase may be imperfect, and when customers consider entering any transaction with such information, they may incur a loss. Prospect Theory provides us with a lens to investigate consumer decision making when their decisions can lead to a loss. Hence, we apply both TPB and Prospect Theory to analyze how stores can potentially drive online sales.

TPB posits that a customer's behavioral intention for making a purchase depends upon her attitude towards making that behavior. As per the expectancy-value model, this attitude is shaped by the belief that this purchase will be favorable for her. Clearly, the process of attitude formation is critical, and this is where channel capabilities that help the customer evaluate a product assume importance. For products rich in non-digital attributes, customer needs physical inspection to gather information on such attributes (Lal and Sarvary 1999). For example, a women's dress may be available in several options of cuts, styles, fabric, colors, and sizes. After physically trying out various options, the customer may be able to form an attitude supportive of the intention to purchase particular products if those products enhance her beliefs regarding the favorability of purchasing the product along multiple dimensions of cut, style etc. In absence of the retailer's store near her, a customer may be able to physically evaluate the retailer's products in a distant store less frequently and thus may have an unfavorable attitude towards purchasing the retailer's products. However, once the retailer opens a store near the customer, she is able to physically evaluate the retailer's products and this change may make her attitude towards the retailer's products more favorable. Therefore, all else being equal, a customer's intention to purchase the retailer's products as compared to that of other competing retailers may increase after the retailer's store opens near her.[5] Naturally, such increased purchase intentions for the retailer's products during store interactions should increase customer's store

[4] https://www.statista.com/statistics/813954/planned-back-to-school-apparel-purchase-of-consumers-us/
[5] In line with findings in prior research (Brynjolfsson et al. 2009, Forman et al. 2009), after the retailer's store opening, customers may substitute some of their purchases from other competing retailers with the retailer's products. This naturally happens when customers divert their attention from competing retailers' products to the retailer's products.

purchases. However, there may be several scenarios where the customer may not purchase the evaluated products in store but purchase those products later from the retailer's online channel. We outline these possibilities below:

a) There may be situations where a customer may not be able to find a product of the desired size, color, cut, or style of her chosen apparel in store, as only limited number of options can be stocked in the limited space available. It is possible that a customer evaluates the non-digital attributes (e.g. cut, fabric and style) of a product in the store, and forms an unfavorable attitude towards purchase because of unavailability of some desired digital attribute (e.g. color). In such cases, she can later form a positive attitude supportive of the purchase intention once she finds the product in the online channel with the fabric and style she likes (identified in store) and the color she wants (identified in the online channel). This is plausible because typically bigger assortments are available online due to lower stocking costs (Brynjolfsson et al. 2006).

b) The TPB further posits that a customer's purchase intention is influenced by the strength of social norms. The social norms in the present case would be determined by customer's perception about the relevant others' opinions about her purchase. While at the store, a customer may not be able to determine the weight of the social norms because opinions of family or friends may not be easily available during the store visit[6], but she can learn about their opinion later on (e.g. by showing them pictures of the item or showing the items on the retailer's website). If the weight of the social norms strengthens after a store visit, the intent to purchase may become higher. At this stage, the customer can make a purchase using the online channel without needing to take the time and effort to revisit the store.

c) The Prospect Theory of mental accounting (Thaler 2008) posits that a customer attaches higher weights to losses than to gains. If a customer believes that the price of the product in the retailer's

[6] Customer can seek the opinion of friends/family on how she looks in the apparel by sending her picture from the store, but it may be only possible if friends/family are free to respond at that time.

store is higher than its reference point, she would perceive a loss (negative utility) from purchasing the product in store and would therefore not purchase it in store even if it fits her expectations. As the reference point for a product is based on the average price of the product at competing stores (Rajendran and Tellis 1994), the customer may first want to explore other options available at competing retailers (Cachon et al. 2005). Once the customer has determined competing prices, her transaction utility from purchase of item with the retailer may become positive (if the price at the retailer is the lowest compared to other stores). At this point, she can make a purchase using the online channel without needing to revisit the store.

Therefore, in situations (a), (b), and (c), the origination of the customers' demand for product occurs at the store channel but the customers may purchase the product later using the convenience of the online channel without visiting the store again. The possibility of customers using store channel for product evaluation and online channel for product purchase is consistent with prior research, which indicates that customers may use different channels during different stages of their purchase process (Verhoef et al. 2007, Galleno and Moreno 2014). *The increased frequency (probability) of store interactions of customers after a store opening would increase the frequency of occurrence of scenarios described in (a), (b), and (c) above*, and this increase could lead to a complementary effect on their online purchases from the retailer. We call this possible effect of higher store visits on increased online purchases as the *store engagement effect*.

The TPB plays an important role in influencing customers' purchase behavior on the online channel in yet another way. Product evaluation and purchase on the online channel may not always result in a successful transaction because of product misfit, and the possibility that wrong, or damaged products may be shipped. Thus, the perceived behavioral control for online purchasing is low leading to weak purchase intent. However, if it is possible for customers to easily return the online purchases they made, thereby mitigating any mistakes made in the purchase transaction from the online channel, then the obstacle to a successful online purchase is reduced. The easier accessibility of a store due to a new store opening, and the provision to return the online purchases at the store increases the effectiveness of perceived behavioral control and consequently increases the intent to purchase online. This prediction is in line with prior research that shows that easier return policies could lead to higher customer purchases (Wood 2001). Further, Anderson et al. (2009) show that the option to return on catalog channel increases catalog sales. In our setting, we extend this idea in the context of store and online channels for products with non-digital attributes, where the perceived behavioral control of customers on online purchase is much lower. Clearly, availability of an option to return products in store could make the customers more likely to purchase products with non-digital attributes online even if they are uncertain about its fit/looks. We refer to this effect of easy store returns on customers' online purchases behavior as the *store return effect*.

Overall, the net effect of store openings by a multichannel retailer on its online sales is an open empirical question. If the store openings have a complementary effect on its online sales, it could be caused by two mechanisms, store engagement effect and store return effect. If customers' online purchases increase *due to store engagement effect, it should be associated with their higher store interactions*. Further, if customers' online purchases increase *due to store return effect, it should be associated with higher return of their online purchased quantities in store*. Furthermore, since the *driver of the two effects is reduction in customers' store distances, these effects should be higher for those customers who experience greater reduction in their store distances due to store openings*. We utilize these insights to provide empirical evidence of these effects as well as the driver for these effects in our field data.

To highlight our contributions, we succinctly compare our paper with other papers closest to our work using in Table A1 in Appendix A.

3. Research Setting, Sample Selection, and Experimental Strategy

3.1 Field Setup

We conduct our study on a large fashion apparel, accessories, and home products retailer in the U.S.[7] The retailer mainly sells products through physical stores and a website.[8] The retailer has an annual revenue of about US $10.0 billion.[9] In the present study, we utilize the event of store openings by the retailer to examine the effect of reduction in store access costs of *existing customers* on their purchase and return transactions on the store and online channels. We chose to examine the behavior of existing customers because store opening can only result in additional sales from the newly acquired customers, as these customers were not purchasing from the retailer prior to store opening. In contrast, store opening can affect the online purchases of existing customers in either direction, and hence it is academically interesting to examine.

We obtained data on purchase and return transactions on the physical store and the online channel of a 10% random sample (approximately 1.5 million customer) of the total population of customers who made at least one purchase with the retailer in the period from July 1999 to June 2006. Customers of the retailer means customer households that purchase apparel/home goods with the retailer. The retailer identifies its customers through multiple measures such as phone number, loyalty card, credit card, e-mail address, physical address, and other personal information about customers.[10] Because of this extensive identification process, the retailer is able to relate a transaction (in scanner panel data as well as in online sales data) to a specific customer.[11]

To examine the long-term effect of store openings on the online purchase behavior of existing customers, we require reasonable period of data before and after the event of store opening. This requirement dictated our choice of existing customers and the store opening events. We chose six store openings by the retailer in 2003 for our present analysis so that we have about three-year data on customers' post-store opening period purchases. We chose 656,949 customers who made at least one transaction with

[7] The identity of the retailer is not disclosed due to a non-disclosure agreement.
[8] Store and online sales account for roughly 95% of the total sales of the retailer.
[9] The exact annual revenue of the retailer is not disclosed to keep its identity confidential.
[10] https://econsultancy.com/blog/68559-how-to-identify-retail-customers-in-store/
[11] The process of identifying customers and relating it to the scanner panel data is quite prevalent among most of the big apparel and home goods retailers, such as Macy's, JC Penny, Nordstrom's, and Dillard's.

the retailer in the period from July 1999 to 1st January 2001 as existing customers for our analysis so that we have at least two years purchase data for the existing customers in the pre-store opening period.

The retailer opened approximately 30 new stores from July 1999 to June 2006.[12] We utilized the great-circle distance formula to compute the distance between the zip codes of customers and their nearest store (called *store distance*). As a result of these store openings the distance from their nearest store reduced for 90,326 existing customers (called *affected customers*) but remained the same for the remaining 566,623 existing customers (called *unaffected customers*).

3.2 Sample Selection

Since we wish to examine the effect of 2003 store openings, out of 90,326 affected customers we selected those customers, whose distance from the nearest store was affected by 2003 store openings only. The store distance for affected customers may change multiple times by store openings in the pre-2003 and/or post-2003 period besides in 2003 period. This way, we avoid picking up the effect of pre-2003 or post-2003 store openings on customers' purchase behavior.

The online purchases in a state were only liable to sales tax if the retailer has a physical store in that state (Anderson et al. 2010). Therefore, the sales tax liability on online purchase and hence online purchase propensity of affected customers may change if the store opened in 2003 is the retailer's first store in the state. We checked and found that all six store openings of the retailer in 2003 were in the states where it already had stores.[13] Therefore, the sales tax liability on the online purchases of the affected customers remained unchanged. We finally got a sample of 17,277 affected customers from 1,200 unique zip codes in the U.S., which constitutes our treatment group. We report the distribution of store distance before and after the store openings for these affected customers in Table 1, which indicates a substantial reduction in store distance for affected customers due to store openings – the change in store distances for the median affected

[12] The exact number of retailer's stores are not given due to non-disclosure agreement.
[13] The retailer closed only two stores in the study period and these stores were closed in the states where it had other stores. So, these store closing did not change the sales tax liability on online purchases.

customer was 33.9 miles and the store distances for median affected customer before and after store opening were 154.1 miles and 22.5 miles, respectively.

Type of Customers	Store distance in miles	Mean	Std. Dev.	Percentile Values				
				0	25	50	75	100
17,277 Affected customers	before store opening	148.0	127.4	0.9	22.9	154.1	234.9	473.0
	after store opening	78.0	97.9	0.0	7.9	22.5	144.1	412.0
	change in store distance	70.0	78.9	0.0	4.7	33.9	114.7	233.5
201,096 Unaffected customers	before/after store opening	53.0	73.8	0	6.1	24.9	68.2	794.5

Table 1: Customers store distance

Out of the total 566,623 unaffected customers, we selected a subsample in such a way that we have sufficient number of unaffected customers whose store distances match with the store distances of affected customers in all quartiles of distribution prior to store opening as shown in Table 1. Out of the total 566,623 unaffected customers, only 22,403 customers had the store distance of more than 154 miles. Therefore, we retained all of these 22,403 unaffected customers. Out of the remaining 544,220 unaffected customers with store distance below 154 miles, we randomly selected 200,000 customers.[14] From this sample of 200,000 + 22,403 = 222,403 unaffected customers, we dropped all those customers whose sales tax liability on online purchases changed due to opening of a first store in their state anytime during our analysis period.[15] As a result, we finally got a sample of 201,096 unaffected customers from 12,859 unique zip codes in the U.S., which constitutes our control group.

3.3 Experimental Design

To identify the treatment effect of store openings on online purchase behavior, we compare the change in online purchases of our selected sample of affected and unaffected customers around the store opening period. A major challenge in making this comparison is that in addition to the treatment of store opening,

[14] We intentionally kept larger sample size of unaffected customer to find sufficient number of control customers with similar purchase behavior as treated customers prior to store opening for our matching estimator.

[15] For unaffected customers living at the boundary of a state, the nearest retailer's store was in the adjoining state. Thus, by opening of a retailer's store in their state, the sales tax liability on their online purchases changed but their distance from the nearest store remained unaffected.

certain other factors may also differentially change for the affected and unaffected customers and hence may confound the treatment effect. For example, a retailer may offer preferential promotions in the area where it opens a store to attract new customers. Such preferential treatment could additionally influence the online purchases of affected customers, and thus, the difference of change in purchases behavior of affected customers from that of unaffected customers cannot be attributed to store opening alone.

To assuage this concern, we describe the promotion, shipping, and return policies of the retailer which could potentially be different for affected and unaffected customers.

Price promotions - The retailer offers sales at different times in a calendar year on different categories of items. These sales are pre-decided and announced well in advance. These sales are uniformly applicable on all stores in the U.S. and on the online channel of the retailer. Therefore, at a given time, the price of an item is same across all stores in the U.S. and on its online channel. The retailer did not send price coupons/discounts to its customers through either mail, or email, or catalog in the study period. It also did not offer any price discounts in conjunction with a new store opening. Therefore, both the affected and unaffected customers face the same price for an item on all channels irrespective of store opening.

Shipping policy – The retailer's charged shipping rates based on the value of purchased items and it applied uniformly over the whole U.S. So, customers living in different parts of the U.S. faced the same shipping charges on their online purchases of the same value. During the study period, the retailer made changes in its shipping rates in late 2005, but this change applied uniformly over the whole U.S. Therefore, the affected and unaffected customers face the same shipping rates and thus the same final price for their online purchases of same value of items purchased.

Return policy – Any product purchased online or in store can be returned within 90 days of date of purchase. Products can be returned either by mail free of charge with no minimum purchase threshold, or in store. This return policy of the retailer was uniformly applicable over the whole U.S. Therefore, the impact of these policies is expected to be similar on the affected and unaffected customers.

Non-price promotions – The retailer offered non-price promotions by periodically sending catalog and emails to randomly selected customers to inform them about its products. It also served limited ads on the national media (TV, radio, and newspapers). Thus, the exposure of the retailer's ads to the affected and unaffected customers are likely to be similar. Normally, the retailer did not advertise its products in the local and regional media, but it publicized its store opening in the local media. These promotions were done for a short period around the time of store opening. Still, such promotions may result in higher awareness about the retailer in the area of new store opening and thus higher purchase preferences of affected customers as compared to the unaffected customers. One way to account for the impact of non-price promotions would have been to include the magnitude of such promotions as a variable in our econometric specification. Unfortunately, we did not have data on such promotions, and so we excluded the six-month period around the store opening date– called as store opening period – from our analysis to avoid picking up the effect of increased awareness about the retailer around the store opening date.

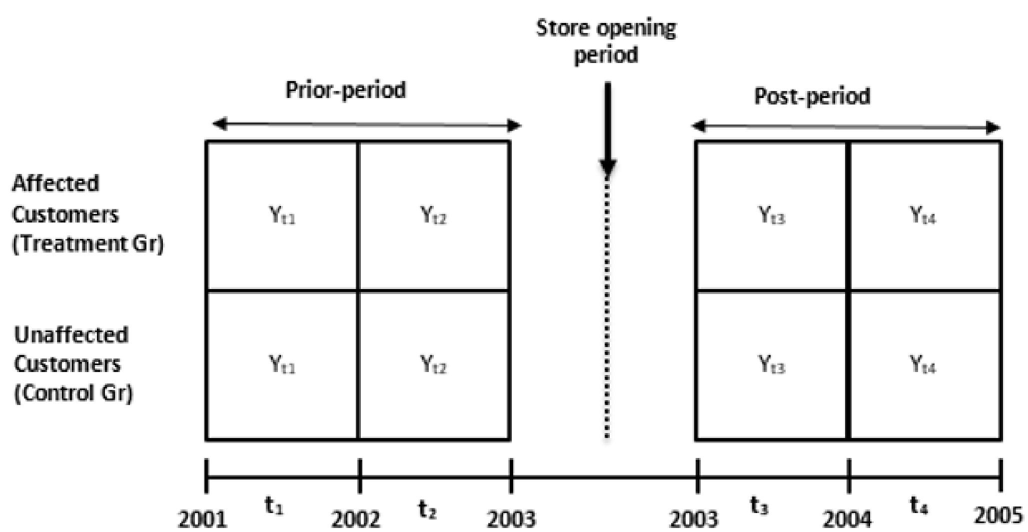

Figure 1: Experimental Setup

We use a diff-in-diff experimental design as shown in Figure (1) to estimate the effect of store opening on customers' online purchase behavior. In this design, we use customers' purchase data (denoted as Y) for four annual periods (denoted by t_1 to t_4); two annual periods prior to the beginning of store opening period and for two annual periods after the end of store opening period. We chose to conduct our analysis on customers' purchase data at the annual level for two reasons. First, the objective of the present study is to find the long-term effect of physical stores on customers' online purchase behavior and not to examine the short term dynamics in customers' demand with store openings. Second, the median customer in our data does not make any purchase with the retailer in a year and thus analysis on monthly- or quarterly-level customer purchase data would not be meaningful. The exact dates for four annual periods for different store openings may vary based on the date of store opening. For example, for a store opening on 15th August 2003, store opening period will be 15th May 2003-15th November 2003, period 1 will be 15th May 2001-14th May 2002, period 2 will be 15th May 2002-14th May 2003, period 3 will be 15th Nov 2003-14th Nov 2004, and period 4 will be 15th Nov 2004-15^{4h} Nov 2005. However, all annual periods constructed around different store opening dates will cover a whole year. For example, for store openings of 15th August 2003, period 1 will be will be 15th May 2001-15th May 2002 and for store opening of 15th May 2003, period 1 will be will be 15th February 2001-15th February 2002. Both of these periods cover the full year.

4 Data and Econometric Specifications

4.1 Data Description

We obtained detailed customer-level transaction data from the retailer. In this data, for each transaction by a customer, we have precise information on the customer identity (unique customer number assigned by the retailer), date of transaction, channel (online/physical store) of transaction, and data (description, category, price, quantity, and discount) on each product purchased/returned in the transaction. We aggregated this customer–level transaction data to obtain the number of transactions, number of items purchased, and purchase revenue separately on the online and physical store channel for each customer in the sample of 17277 affected and 201096 unaffected customers in four annual time periods in Figure 1. The summary statistics of this data is reported in Table 2.

The retailer also collects data on the age of head of customer household and household annual income through surveys. The retailer categorizes its customers into six income categories based on their annual household income [1 → < $50K, 2 → $50-$75K, 3 → $75-$100K, 4 → $100-$150K, 5 → >$150K, and 6 → unknown] and seven age categories based on the head of customer household [1 → < 25 years, 2→ 25-34 years, 3 → 35-44 years, 4 → 45-54 years, 5 → 55-64 years, 6 → >65 years, and 7 → unknown]. When customers do not reveal their income or age in their surveys, they are assigned to the unknown category. We report the breakup of number (%) of affected and unaffected customers in different age and income categories in Table 3, which shows similar distribution of age and income categories in the two groups of customers.

Variables	17277 Affected customers				201096 Unaffected customers			
	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max
Pre-store opening period 1								
No. of online purchase transactions	0.14	0.74	0	32	0.10	0.57	0	46
No. of online purchased items	0.36	2.65	0	186	0.24	1.93	0	254
Online purchase revenue (US $)	20.26	169.29	0	15480.05	13.45	105.81	0	11910.8
No. of store purchase transactions	1.21	3.25	0	75	2.26	4.95	0	117
No. of store purchased items	4.18	13.62	0	592	7.85	21.20	0	1887
Store purchase revenue (US $)	221.54	1127.02	0	111337.3	373.84	1172.07	0	141621.4
Pre-store opening period 2								
No. of online purchase transactions	0.17	0.91	0	58	0.13	0.73	0	80
No. of online purchased items	0.44	3.28	0	287	0.38	14.48	0	6390
Online purchase revenue (US $)	24.53	224.56	0	23937.6	19.99	585.18	0	254961
No. of store purchase transactions	1.16	3.17	0	61	2.21	5.06	0	136
No. of store purchased items	4.28	14.30	0	643	8.01	22.61	0	1888
Store purchase revenue (US $)	231.71	1442.70	0	159120.7	383.60	1226.59	0	101560.3
Post-store opening period 3								
No. of online purchase transactions	0.24	1.00	0	45	0.17	0.87	0	83
No. of online purchased items	0.58	3.36	0	207	0.43	3.19	0	406
Online purchase revenue (US $)	36.60	195.52	0	11112.64	27.15	204.16	0	27221
No. of store purchase transactions	1.90	4.33	0	76	2.29	5.29	0	136
No. of store purchased items	6.72	20.52	0	1244	8.22	22.73	0	993
Store purchase revenue (US $)	390.56	2880.43	0	340892	413.62	1347.07	0	131423.5
Post-store opening period 4								
No. of online purchase transactions	0.29	1.21	0	55	0.20	1.04	0	138
No. of online purchased items	0.72	4.14	0	264	0.53	4.33	0	1081
Online purchase revenue (US $)	46.67	262.47	0	12609.01	34.61	286.21	0	50050.28
No. of store purchase transactions	2.09	4.69	0	91	2.47	5.53	0	136
No. of store purchased items	7.15	21.79	0	1525	8.71	23.38	0	921
Store purchase revenue (US $)	421.39	2957.66	0	352297.8	457.36	1451.03	0	124808.1

Table 2: Summary statistics of online purchase transactions

Age Category	Affected Customers		Unaffected Customers		Annual Income Category	Affected Customers		Unaffected Customers	
	No.	%	No.	%		No.	%	No.	%
<25 Years	988	5.7%	12036	6.0%	<50K	6109	35.4%	63201	31.4%
25-34 Years	1701	9.8%	18751	9.3%	50-75K	3294	19.1%	42072	20.9%

35-44 Years	2434	14.1%	29520	14.7%	75-100K	2158	12.5%	27657	13.8%
45-54 Years	3436	19.9%	40380	20.1%	100-150K	2275	13.2%	27351	13.6%
55-64 Years	3376	19.5%	40175	20.0%	>150K	1892	11.0%	24767	12.3%
>65 Years	3550	20.5%	38637	19.2%	Unknown	1549	9.0%	16048	8.0%
Unknown	1792	10.4%	21597	10.7%	Total	17277		201096	
Total	17277		201096						

Table 3: Distribution of customers in different age and income categories

In addition to the customer-level data, we also collected the zip-code level socioeconomic and demographic data for the year 2003 from Sourcebook and ESRI datasets. We provide details of these data in Section 4.2.2. The basic purpose for collecting these data is to account for zip code-level differences that could differentially affect the purchase behavior of affected and unaffected customers in the post-store opening period. Our selection of zip code-level variables were based on the previous literature (Avery et al 2012, Bell et al. 2015).

4.2 Econometric Specifications

The treatment effect of store opening on the online purchase behavior of affected customers in our experimental setup is identified by inferring their counterfactual behavior in post-store opening period from the purchase behavior of unaffected customers in that period. Thus, for clean identification of treatment effect, the effect of all observed and unobserved factors, other than store opening, on the purchase behavior of affected and unaffected customers should be similar. In the previous sections, we discuss that the retailer's policies on promotions, product returns, and shipping equally influence the affected and unaffected customers. Since we compare the purchase behavior of affected and unaffected customers in the same calendar year period in our experimental setup, the effect of time related factors on their purchase behaviors is likely to be similar. But, the retailer selects an area to open its store, and this may be based on the differences in the purchase propensities of the affected and unaffected customers. We run two alternative specifications to account for these difference in the two categories of customers.

4.2.1 Fixed-Effect Estimator

The most straightforward way to implement our experimental design in Figure 1 is the following fixed effect specifications:

$$Y_{it} = \beta_i + \beta_1 \, Post_t + \beta_2 \, (Treat \times Post)_{it} + \varepsilon_{it} \quad ; \qquad \text{--- (1)}$$

where $t = t_1$, t_2, t_3, and t_4 denote four time periods in Figure 1; i denotes customers; $Treat_i$ is an indicator variable equal to one if i is an affected customer and zero otherwise; and $Post_t$ is an indicator variable equal to one for $t = t_3$ or t_4, and zero otherwise. The left hand side variable Y_{it} denotes the online purchase variable for customer i in period t. Coefficient β_i captures customer fixed effects that accounts for time invariant unobserved differences across customers. Thus, the fixed effect accounts for time invariant unobserved differences between affected and unaffected customers. The coefficient of $Treat \times Post$ (i.e., β_2) is of interest, and it captures the treatment effect of store opening on the online purchase behavior of affected customers. The retailer may select an area to open its store based on the increasing purchase trends (preferences for apparel) in that area, and thus we may expect a higher pre-existing purchase trends for affected customers compared to that of unaffected customers prior to store opening. The diff-in-diff design specification (1) can spuriously identify the pre-existing higher purchase trends of affected customers as the treatment effect of store opening. In Appendix B, we check and find similar (parallel) purchase trends of our sample of affected and unaffected customers in the pre-store opening period.

4.2.2 Matching Estimator

The treatment effect in fixed effect specifications (1) is identified by the difference in average online purchase behaviors of the affected and unaffected customers after accounting for differences in scales of purchases across customers. However, a more precise estimate of treatment effect is the difference of online purchase behavior of affected customers from the average purchase behavior of their matched sample of unaffected customers from whom their counterfactual purchase behavior after store openings could be inferred. We chose the following four categories of variables (as reported in Table 4) to match the affected and unaffected customers.

1. Unaffected customers, who exhibited similar purchase behavior as the affected customer prior to store opening, are likely to closely approximate the counterfactual purchase behavior of affected customers after store openings. Thus, matching the two categories of customers on their pre-store opening period purchase behavior is the most obvious choice in a matching estimator. Marketing literature indicates that the purchase behavior of customers can be captured by RFM variables describing the recency, frequency, and monetary values of their purchases/returns. Thus, we use the observed customer-level purchase/return variables capturing the recency, frequency, and monetary values of their purchases prior to store openings as described in Table 4.
2. We also utilize customer-level socio-demographic variables, such as age of customer (i.e. head of customer household), customers' household income, and customers' distance from the nearest store to match the affected and unaffected customers. Customers of similar age and income are likely to have similar purchase preferences. Also, customers with similar distance from the nearest store may similarly choose between online and store channels of the retailer.
3. Although, we account for the actual purchases/returns of affected and unaffected customers, we suspect that any zip code-level factors prior to store opening may result in differential trends in purchase behavior of affected customers. For example, the retailer may open its store in an area experiencing economic growth, which may result in increasing trends in purchase propensities of fashion apparels of residents of that area (affected customers). Comparing the purchase of such customers in post store opening period from that of customers in other areas (who made similar volumes of purchase previously but at a constant rate) could spuriously capture the increasing trends of their purchase as a positive treatment effect of store opening. To account for such possibilities, we additionally match the change in customer-level annual purchase on the online and store channels for the affected and unaffected customers prior to store openings.
4. It is also plausible that some of the changes in the purchase propensities of residents of an area due to area specific factors prior to store opening may occur in the post-store opening period. If this is

the case, the counterfactual behavior of affected customers in the post-store opening period may differ from that of their matched unaffected customers even if they exhibited similar levels and trends in purchase behavior in the pre-store opening period. To account for this possibility, we also collected data on zip code-level socioeconomic-demographic factors for the year 2003 from Sourcebook and ESRI datasets to match the affected and unaffected customers. Specifically, we match the zip codes of affected and unaffected customers on socioeconomic factors, such as population, median income, and median age. A retailer may choose to open its store in an area based on these factors. A retailer may also decide to open a store in an area based on the internet penetration, extent of computer usage, and the preference for apparel in that area. Accordingly, we use these aggregate variables at the zip code-level to match the two categories of customers.

(A) - Observed customer-level cumulative purchase/return variables prior to store opening
Recency measures - Time between last product purchase and store opening date in weeks Frequency measures - Number of annual purchase transactions in store and online channels - Number of annual return transactions in store and online channels Monetary value measures - Number of products purchased in store and online channels - Number of products returned in store and online channels - Purchase revenue in store and online channels in US $ - Return revenue in store and online channels in US $
(B) - Customer-level socio-economic-demographic variables
- Distance from the nearest store in miles - Income category of the customer household - Age category of the head of the customer household
(C) – Customer-level purchase trends prior to store opening
- Change in annual purchase transactions from 2001-02 to 2002-03 on store channel. - Change in annual purchase transactions from 2001-02 to 2002-03 on online channel.
(D) - Zip code-level aggregate variables in 2003
- Population - Median age of population - Median income - Internet usage index - Computer usage index - Apparel preference index

Table 4: Variables used for matching affected and unaffected customers

We also report the summary statistics of these variables used for matching in Table C1 in Appendix C. Form Table C1, we find significantly different distributions of these variables for our sample of affected and unaffected customers, which necessitates the use of matching methods to account for these differences. We use two types of matching estimators– a parametric inverse propensity score weighted doubly robust estimator and a nonparametric coarsened exact matching estimator.

4.2.2.1 Inverse Propensity Score Weighted Matching Estimator

We compute the propensity score $p\ (X_i)$ for customer i, which is simply the logit probability that customer i is an affected customer given her values of matching variables (X_i). We compute the inverse probability weights as $Weight_i = [\{Treat_i / p\ (X_i)\} + \{(1-Treat_i) / (1- p\ (X_i))\}]$, where $Treat_i$ is an indicator variable equal to one if customer i is an affected customer and zero otherwise. Using such computed weights, we estimate the following weighted OLS specification:

$$Y_{it} = \beta_0 + \beta_1\ Treat_i + \ \beta_2\ Post_t + \ \beta_3\ (Treat\ \times Post)_{it} + \varepsilon_{it} \quad ; \qquad \text{-------- (2)}$$

where, the variables in specification (2) have the same meaning as in specification (1). Specification (2) with inverse probability weighting on propensity scores is referred to as the Doubly Robust Estimators and is widely used to estimate treatment effect in economics literature (Hirano and Imbens 2001). The coefficient β_3 in this specification captures the treatment effect of store opening on the online purchase behavior of the affected customers after accounting for the purchase behavior of their matched unaffected customers. In Appendix D, we check and find that after propensity score weighting the two groups of customers were well balanced on all matching variables. We also plotted the histogram of propensity scores for the two groups of customers and found that there is sufficient overlap between the two, which satisfies the overlap assumption required in estimation of average treatment effect.[16]

4.2.2.2 Coarsened Exact Matching (CEM) Estimator

We additionally deploy Coarsened Exact Matching (CEM) estimators to estimate the *ATE*. Unlike propensity score matching method that uses maximum likelihood estimators to control for the differences in pretreatment variables across the affected and unaffected customer groups, CEM is a nonparametric

[16] The propensity score histograms are omitted for brevity. But, they are available on demand from the authors.

matching method that allows researchers to ex-ante bound the imbalance (both on individual variables and jointly) between the affected and unaffected customer groups by manually coarsening the pretreatment variables into different sized bins and thereby ex-ante control the error in *ATE* estimates (see Iacus et al. 2009 for details).

In CEM estimators, we try different coarsening of bin sizes for each variables X_i such that both the affected and their matched unaffected customers fall under the same bin for each variable. Then, we compute the difference in change in purchase behavior of each affected customer from that of the average of corresponding values for their matched sample of unaffected customers. The average treatment effect of store opening is the average of these computed differences over all affected customers and it can be mathematically represented as:

$$ATT = \frac{1}{N_{af}} \sum_{af} \left[\left(Y_{post}^{af} - Y_{prior}^{af} \right) - \frac{1}{N_{uaf_{af}}} \sum_{uaf_{af}} \left(Y_{post}^{uaf_{af}} - Y_{prior}^{uaf_{af}} \right) \right] \; ; \qquad \text{-- (3)}$$

where, $Y_{prior} = (Y_{period1} + Y_{period2})/2$, and $Y_{post} = (Y_{period3} + Y_{period4})/2$, respectively, indicates the average purchase variables for customers before and after the store openings; *af* denotes affected customers; and uaf_{af} denotes matched unaffected customers for affected customer *af*. N_{af} and $N_{uaf_{af}}$, respectively, denote the number of affected customers and the number of matched unaffected customers for an affected customer.

5 Results and Discussions

5.1 Model Free Evidence on Data

In this section, we compare the number of customers making online / store interactions, and the mean values of their online purchases for the samples of affected and unaffected customers during the four annual periods of our experimental setup. The purpose of these comparisons is to show preliminary suggestive evidence of the complementary effect, the store engagement effect, and the store return effect due to store opening on our data.[17]

5.1.1 Complementary Effect of Store Opening on Online Purchases

In Table 5, we compare the number of affected and unaffected customers making store and online interactions with the retailer in the prior- and post-store opening periods. We observe that after the retailer's store openings (in 2003), the percentage of affected customers who do store interactions increases from 28.7% to 38.7% while the percentage of unaffected customers who do so marginally reduces from 40.9% to 39.8%. This change is expected, as affected customers would interact more at stores once they get easier access to these stores due to reduction in their store distances. But, we also find a greater increase in the percentage of affected customers making online purchase interactions after the retailer's store openings compared to the unaffected customers (8.5 % to 12.1% compared to 7.2% to 8.3%). We also observe a greater increase in online purchase variables for the sample of affected customers than that of unaffected customers, e.g., increase in online purchase interactions from 0.17 to 0.24 versus from 0.13 to 0.17 and increase in online purchase quantity from 0.44 to 0.58 versus from 0.38 to 0.43.[18]

Mean value of Variables	Period 1	Period 2	Period 3	Period 4	Period 1	Period 2	Period 3	Period 4
	Affected Customers (17277)				Unaffected Customers (201096)			
No. (%)of customers doing store interactions	5422 (31.4)	4951 (28.7)	6644 (38.5)	6950 40.2)	89089 (44.3)	82269 (40.9)	80029 (39.8)	85315 (42.4)
No. (%)of customers doing online interactions	1302 (7.5)	1461 (8.5)	2090 (12.1)	2286 (13.2)	11337 (5.6)	14382 (7.2)	16786 (8.3)	19659 (9.8)
Online purchase interactions	0.14	0.17	0.24	0.29	0.09	0.13	0.17	0.20
Online purchase quantities	0.36	0.44	0.58	0.72	0.24	0.38	0.43	0.53

[17] As the objective is not to show whether the differences in mean values of variables are statistically significant, the sample standard deviations for variables are not reported. This also helps reduce the clutter in tables.
[18] The mean values of online purchase variables are low, as less than 10% of the total customers in the sample make online interactions with the retailer. However, this is in line with the industry averages of 5-10 percent for multichannel apparel retailers in the study period (e.g., see https://www.internetretailer.com/2009/02/27/web-sales-at-macy-s-grow-29-in-2008-while-total-sales-sink-7).

Online purchase revenue in US $	20.26	24.53	36.60	46.67	13.45	19.99	27.15	34.61

The values in parenthesis are percentage of total customers. Online purchase interactions, quantities, and revenue are, respectively, the annual number of purchase interactions, total quantity purchased, and total purchase revenue (in US $) on the online channel.

Table 5: Model free evidence of complementary effect

We also show this graphically in Figure 2. Overall, we find a complementary effect of store openings on online purchases of affected customers. To understand this complementary effect of store openings on existing customers' online purchase behavior, we specifically examine the purchase behavior of those customers who make online purchases, hereafter called *online customers*.

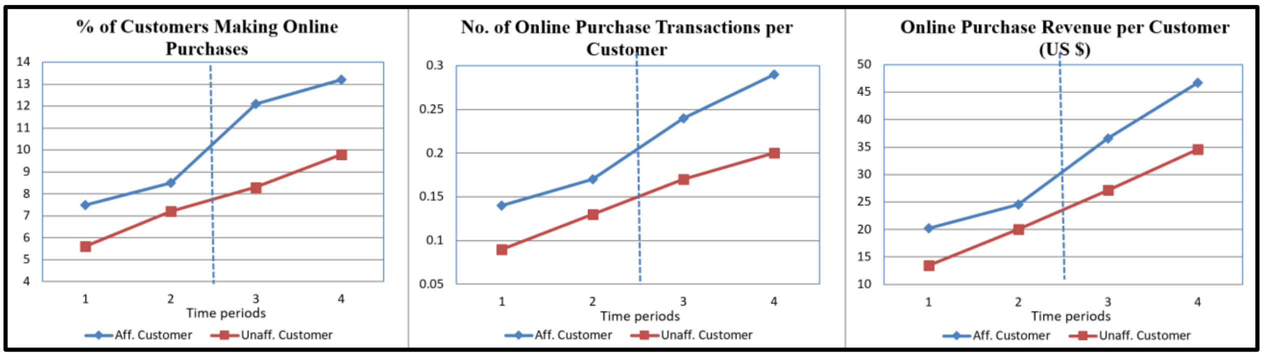

Figure 2: Model free evidence of complementary effect

5.1.2 Model Free Evidence of Store Engagement Effect

The store engagement effect suggests that the store opening near affected customers improves their engagement with the retailer by increased store interactions that may, in turn, lead to increase in their online purchases. Accordingly, to provide evidence of store engagement effect, we show that after store opening, the extent of online transactions by affected customers increases with their store interactions. To check out this fact, we provide the breakup of the affected and unaffected customers making only online purchases and both online and store purchases in Table 6. We find a higher increase in percentage of affected customers making both online and store purchases after store opening as compared to the corresponding increase for unaffected customers (44.8% to 66.8% versus 61.6% to 66.6%). This suggests a higher association between online and store purchases for affected customers after store opening near them. We graphically show this evidence in left part of Figure 3.

To better understand a higher association between online and store interactions for affected customers after store opening, we examine the intensity of their store purchases. The intensity of customers' store purchases is captured by the mean values of their number of store purchase interactions, store purchase quantities, and store purchase revenue in Table 6. We find a significantly greater increase in the store purchase variables after store opening for affected online customers than that for unaffected online customers (e.g., annual store purchase revenue increased from 408.4 to 947.1 versus 739.0 to 946.9). This significant (more than 100 percent) increase in store purchases of affected online customers indicates that: (1) they make higher online purchases with more intense store purchases, and (2) store opening primarily increases their store purchases. Both of these facts support the existence of our proposed store engagement effect.

Customers who make online purchases	Affected customers				Unaffected customers			
	Period 1	Period 2	Period 3	Period 4	Period 1	Period 2	Period 3	Period 4
No. of customers	1302	1461	2090	2286	11337	14382	16786	19659
No. of customers who make only online purchases	743	806	694	757	4410	5525	5609	5813
No. of customers who make both store and online purchases	559	655	1396	1529	6927	8857	11177	13846
% of customers with store purchases	42.9	44.8	66.8	66.9	61.1	61.6	66.6	70.4
Store purchase variables for customers who make online purchases								
Store purchase interactions	1.91	2.01	4.03	4.34	3.75	3.92	4.74	5.40
Store purchase quantity	7.23	7.96	14.49	15.31	14.33	15.66	18.37	20.68
Store purchase revenue	369.6	408.4	947.1	875.5	662.1	739.0	946.9	1129.7

Table 6: Model free evidence of store engagement effect

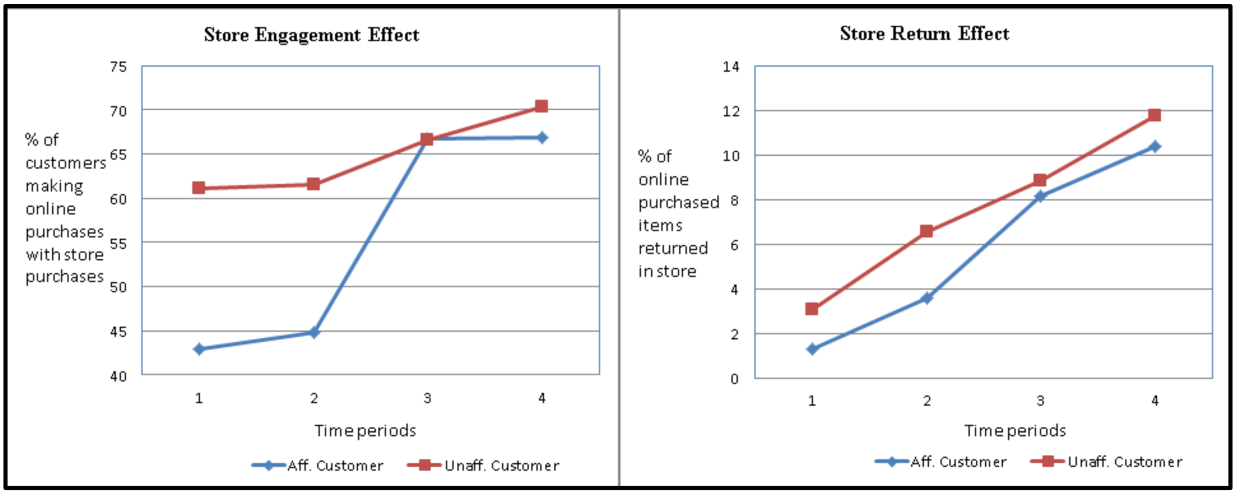

Figure 3: Model free evidence of store engagement and return effects

5.1.3 Model Free Evidence of Store Return Effect

To provide empirical evidence for the store return effect, we show that a greater number of affected customers, who make online purchases, make store returns, and they return higher percentage of their online purchases in store after store opening than that of the corresponding unaffected customers. Table 7 reports the comparison of the percentage of customers making store returns and online purchases, and the mean values of online purchase quantity, online return quantity, online purchased quantity returned in store/online for the two categories of customers making online purchases in different periods.[19]

Customers with online purchase interactions	Period 1	Period 2	Period 3	Period 4	Period 1	Period 2	Period 3	Period 4
	Affected customers				Unaffected customers			
No. of Customers	1302	1461	2090	2286	11337	14382	16786	19659
% of customers making store returns	13.1	15.6	29.2	29.7	24.2	26.6	30.6	35.0
Online purchase quantity	4.72	5.19	4.79	5.44	4.28	5.32	5.17	5.41
Online return quantity	0.39	0.93	0.64	0.76	0.44	0.81	0.77	0.78
Online purchase quantity return in store	0.06	0.19	0.39	0.57	0.13	0.35	0.46	0.64
% of online purchase quantity returned in store	1.3	3.6	8.2	10.4	3.1	6.6	8.9	11.8
% of online purchase quantity returned online	8.3	17.9	13.4	14.0	10.3	15.2	14.9	14.4

[19] The retailer identified its products by their SKU numbers. If the SKU number of an item purchased online by a customer appears again within 90 days from purchase date as item returned in store (online) by the same customer, then we consider this item as purchased online and returned in store (online) by the customer.

Table 7: Model free evidence of store return effect

We find a higher increase in percentage of affected customers making online purchases and store returns after store opening as compared to that of corresponding unaffected customers (15.6% to 29.2% versus 26.6% to 30.6%). We also find a higher increase in percentage of online purchased quantity returned in store after store opening for such affected customers than that for the corresponding value for unaffected customers (3.6% to 8.2% versus 6.6% to 8.9%). Moreover, we find a higher decrease in percentage of online purchased quantity returned online after store opening for the affected customers than that for the unaffected customers (17.9% to 13.4% versus 15.2% to 14.9%). These findings indicate that affected customers utilize the store more relative to unaffected customers to return their online purchases after store opening, which suggests the existence of store return effect in our data. Moreover, we graphically show this evidence in the right half of Figure 3.

5.2 Results of Econometric Specifications

The model free evidences provided in the previous section were merely suggestive and lacked on two counts. First, it did not show whether the observed changes in online purchase behavior of affected customers after store openings were statistically different from that of unaffected customers. Second, it did not account for possible systematic differences between the affected and unaffected customers. In this section, we propose econometric specifications to address these deficiencies and provide rigorous estimates of the treatment effect of store opening on customers' online purchase behavior.

5.2.1 Complementary Effect of Store Opening on Online Purchases

In order to show the complementary effect of store opening on online purchases, we run the fixed effect specifications (1) and the matching estimator specification (2) on our sample of affected and unaffected customers. For the dependent variable, we use three online purchase variables: (1) online purchase probability – indicator variable equal to one if customer *i* does online purchase interaction in period *t*, and zero otherwise; (2) online purchase quantity – number of items customer *i* purchases online in period *t*; and

(3) online purchase revenue – value (in US $) of the online purchases for customer *i* in period *t*. The coefficient estimates are reported in Table 8.

We find similar magnitudes, signs, and significances for the coefficient estimates of online purchase related dependent variables in all specifications, which indicates the robustness of our results. We note that coefficient estimates for *Treat*Post* variable for all online purchase related dependent variables are positive and significant, which indicates an increase in online purchases for affected customers due to store opening. Specifically, the coefficient estimates of *Treat*Post* in the matching estimator indicate the following effect of store opening on affected customers: (1) online purchase probability increases by 0.029, which is a 36 percent increase over their mean online purchase probability value of 0.08 in the prior period; [20] (2) online purchase quantities increase by 0.056, which is a 14 percent increase over their mean online purchase quantities of 0.40 prior to store opening; and (3) online purchase revenues increase by US $6.59, which is a 29 percent increase over their mean online purchase revenue of US $22.4 in the prior period.

Variables	Online purchase probability		Online purchase quantity		Online purchase revenue	
	Coeff. Est.	Std. Err	Coeff. Est.	Std. Err	Coeff. Est.	Std. Err
Fixed Effect Specification (1)						
Post	0.027^{***}	0.001	0.17^{***}	0.02	14.16^{***}	0.77
*Treat*post*	0.020^{***}	0.002	0.08^{***}	0.03	5.07^{***}	1.53
Intercept	0.065^{***}	0.000	0.32^{***}	0.01	17.17^{***}	0.36
N (No. of Groups)	873492 (218373)					
R Square value	0.56		0.61		0.65	
Matching Estimator Specification (2)						
Treat	-0.013	0.010	-0.116	0.078	-6.04	4.32
Post	0.027^{***}	0.003	0.222^{***}	0.036	15.94^{***}	1.42
*Treat*post*	0.029^{**}	0.014	0.056^{**}	0.025	6.59^{**}	3.03
Intercept	0.072^{***}	0.002	0.344^{***}	0.022	18.34^{***}	0.82
N	873492					

$^{***, **, *}$ = statistically significant at $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.10$ levels (two-sided test), respectively. Standard errors are cluster corrected at customer level. 873492 observations are for four time periods observations for 218373 customers → 17277 affected and 201096 unaffected customers.

Table 8: Treatment effect estimates

[20] The mean online purchase variables for affected customers prior to store opening are the average of their mean values in 2001-02 and 2003-03 period from Table 2. For example, mean online purchase probability = (7.5% + 8.5%)/2= 8 % or 0.08, mean online purchase quantity = (0.36+0.44)/2=0.40, and mean online purchase revenue = (20.26+24.53)/2 = 22.4.

5.2.2 Evidence of Store Engagement Effect

In model free evidence of store engagement effect in Section 5.1.2, we have shown that the customers affected make higher online purchases with higher store interactions after the store opening near them. To rigorously test this fact, we compare the probability of store interactions for those customers who make online purchases for the two categories of customers with our proposed econometric specifications. If we show a higher increase in probability of store interactions given online purchase after store opening for affected customers than that of corresponding unaffected customers, then we provide the direct evidence of store engagement effect. Accordingly, we run specifications (1) and (2) on the unbalanced panel of the two categories of customers who make online purchases in different periods of our experiment (details provided below Table 10). The dependent variable Y_{it} (probability of store interactions given online purchase) in this analysis is one if customer *i* (who makes online purchases) makes store purchase in period *t* and zero otherwise. The resulting coefficient estimates are reported in the second and third columns in Table 9.[21]

Variables	Store Engagement Effect: Prob. (store purchase |online purchase)		Store Return Effect: % Online purchase Qty returned in store |online purchase	
	Coeff. Est.	Std. Err.	Coeff. Est.	Std. Err.
Fixed effect Specification (1)				
Post	-0.002	0.004	2.65^{***}	0.22
*Treat*post*	0.179^{***}	0.016	3.28^{***}	0.63
Intercept	0.638^{***}	0.002	5.89^{***}	0.12
N (No. of groups)	69303 (43613)		69303 (43613)	
R squared value	0.37		0.29	
Matching Estimator Specification (2)				
Treat	0.046	0.06	-1.81^{***}	0.27
Post	0.048^{**}	0.02	3.30^{***}	0.26
*Treat*post*	0.110^{**}	0.05	2.76^{**}	1.34
Intercept	0.611^{***}	0.02	4.53^{***}	0.20
N	69303		69303	

***, **, * = statistically significant at $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.10$ levels (two-sided test), respectively. Standard errors are cluster corrected at customer level. Unbalanced panel of 60303 customer-period observations for 43613 unique customers who make online purchases. There are 7139 customer-period observations for 4347 affected customers and 62164 customer-observations for unaffected customers.

[21] We also estimated fixed effect logit specification and found qualitatively similar results. We report the coefficient estimates from the fixed effect OLS in the paper due to its easy interpretability.

Table 9: Evidence of store engagement and return effects

We find a positive and significant coefficient estimate for *Treat*Post* indicating an increase in probability of store purchases for affected customers who make online purchases due to store opening. The coefficient value of 0.11 indicates that the probability of store purchase conditional on making online purchases increase by 0.11 due to store opening, a 25 percent increase over mean probability value of 0.44 for affected customers in prior period.[22]

5.2.3 Evidence of Store Return Effect

For evidence of store return effect, we show a higher increase in percentage of online purchases returned in store after store opening for affected customers than that of corresponding unaffected customers. To implement this, we use Y_{it} – percentage of online purchases returned in store by customer *i* in time period *t* – as the dependent variable in specifications (1) and (2). The resulting coefficient estimates are reported in the fourth and fifth columns in Table 9. We find a positive and significant estimate for coefficient of *Treat*post* variable, which indicates that affected customers return higher percentage of their purchases in store after store opening. A coefficient value of 2.76 indicates that the percentage of online purchases returned in store for affected customers increased by 2.76 due to store opening, a 113 percent increase over their mean value of 2.45 prior to store opening.[23] This result shows that affected customers who increase their online purchases after store opening also return higher percentage of their store purchases in store.

5.2.4 Evidence of Separate Existence of the Two Effects

If an affected customers make more online purchases and then return those items which do not fit their expectation in store and purchase replacement item in store. This would result in higher probability of online purchase with store purchase (measure of store engagement effect) and also higher percentage of

[22] The mean probability (store purchase | online purchase) for affected customers prior to store opening is the average of their mean values in 2001-02 and 2003-03 period from Table 3, i.e., (42.9% + 44.8%)/2= 44 % or 0.44.
[23] The mean value of percent of online purchases returned in store for affected customers prior to store opening is the average of their mean values in 2001-02 and 2003-03 period from Table 4, i.e., = (1.3% + 3.6%)/2= 2.45%.

online purchased items return in store (measure of store return effect). On the one hand, it is higher customer engagement with the retailer that causes higher online purchases and thus returns in store – due to store engagement and not return effect. On the other hand, it could be the ability to return easily in store that causes customer to purchase more online in the first place – due to store return and not engagement effect. Therefore, it is important to examine whether both effects separately exist in our data.

We resolve this issue by recognizing that not all customers return their online purchases in store. If we consider only these customers, the store return effect is likely less important for these customers because they did not return their online purchases in store. If, for this sample of affected customers, we can show that our measure of store engagement effect is positive and significant, then we would allay the doubts about the existence of store engagement effect. In Table 10, we provide the breakdown of number of customers (both affected and unaffected) who make online purchases but do not return any of their online purchases in store. We note that there is a significantly higher increase in number of affected customers who make online purchases with store purchases but do not return such online purchases in store after store opening as compared to the corresponding value for unaffected customers (580 to 1155 versus 7216 to 8864 or 42% to 63% versus 57% to 62%). This shows evidence of store engagement effect without (or with little) store return effect in our data.

	Affected customers				Unaffected customers			
	Period 1	Period 2	Period 3	Period 4	Period 1	Period 2	Period 3	Period 4
Customers who do not return any online purchased items in store								
No. of customers	1270	1375	1828	1914	10662	12611	14348	16237
No. of customers who make only online purchases	742	795	673	739	4374	5395	5484	5648
No. of customers who make both store and online purchases	528	580	1155	1175	6288	7216	8864	10589
% of customers with store purchases	41.6%	42.2	63.2	61.4%	59.0%	57.2	61.8	65.2%

Table 10: Purchase behavior of customers who do not return online purchases in store

Next, we run our fixed effect and matching specifications on the dataset of online customers who do not return their online purchase in store. The resulting estimates are reported in Table 11, where the positive and significant coefficient of the interaction term (*Treat*Post*) indicates the store engagement effect. Thus, we show that our store engagement effect is not driven by customers who purchase replacement items while returning their online purchases in store.

Store Engagement Effect	Prob. (store purchase |online purchase)			
	Fixed Effects (Specification 1)		Matching estimator (Specification 2)	
	Coeff. Est.	Std. Err.	Coeff. Est.	Std. Err.
Treat			0.036	0.06
Post	-0.004	0.004	0.032**	0.016
*Treat*post*	0.18***	0.02	0.13**	0.05
Intercept	0.595***	0.002	0.421***	0.02
N (No. of groups)	60245 (39827)		60245	

Table 11: Store engagement effect for customers who do not return online purchases in store

We further estimate the gains in online purchase revenue separately for customers, who do not return their online purchases in store. This will provide us an estimate of how much of the complementary effect of store is from the store engagement effect alone. For this analysis, we use data on all affected and unaffected customers, who either do not make any online purchase or if they make online purchases then they do not return it in store. We report the resulting estimates for the fixed effect and matching estimators for this dataset along with the corresponding results for all customers in Table 12. We find that the majority of complementary effect is coming due to store engagement effect.

Online purchase revenue	All customers		Customers with no store return of their online purchase	
	Coeff. Est.	Std. Err	Coeff. Est.	Std. Err
Fixed Effect Specification (1)				
Post	14.16***	0.77	8.77***	0.71
*Treat*post*	5.07***	1.53	4.99***	1.29
Intercept	17.17***	0.36	14.51***	0.33
N (No. of Groups)	873492 (218373)		837669 (217886)	
R Square value	0.65		0.61	
Matching Estimator Specification (2)				
Treat	-6.04	4.32	-5.92	4.83
Post	15.94***	1.42	13.26***	2.13
*Treat*post*	6.59**	3.03	6.01**	3.02
Intercept	18.34***	0.82	21.32***	1.27

***, **, * = statistically significant at $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.10$ levels (two-sided test), respectively.
Standard errors are cluster corrected at customer level.

Table 12: Complementary effect of store for customers who do not return online purchases in store

A customer may not return her online purchases in store because it fits her expectations, but she may still make higher online purchases due to reduction in cost to return it in nearby store after store opening. While we can show evidence of store return effect with a higher increase in percentage of online purchased items return in store by the affected customers, store return effect may exist with availability of nearby store without any increase in store returns of online purchases. Thus, it is not possible to precisely estimate the extent of complementary effect of store on online purchases due to store engagement and return effects separately in our present setup.

5.3 Robustness Checks

We performed several checks to ensure that out treatment effect estimates are robust. First, we estimated treatment effect of store opening by CEM estimator with several different coarsening of bin sizes of matching variables to match the affected and unaffected customers. In Appendix E, we show two such example CEM estimates with manual coarsening along with the doubly robust estimators. We find that our *ATE* estimates from the two options of manual coarsening remain qualitatively similar to that obtained from the propensity score matching, which suggests the robustness of our causal estimates.

If the complementary effect of store opening is due to ease in customers' store access, the higher the reduction in store distance, the higher should be the effect of store opening. We examine how the effects of store openings on customers' online purchase behaviors varies with the change in their store distances in Appendix F. We find that the complementary effect, store engagement effect, and store return effect increases with the increase in reduction of customers' store distance. This further provides support that our estimated effects are caused by the ease in customers' store access.

We further varied the length of pre-treatment and post-treatment periods in our analysis and find qualitatively similar treatment effect. Specifically, we used two alternatives: Case 1- one pre-treatment period with two post-treatment periods and Case 2- two pre-treatment periods with one post-treatment period for this analysis.[24] Similarly, we find qualitatively similar estimates for the treatment effect with removal of three and nine month period around the store opening date (i.e. varied the store opening period), which indicates that our treatment effect is not picking up the effect of increase awareness about the retailer during store opening period.[25] We further find a higher complementary effect of store opening by estimating our specifications without excluding 6 months data around the date of store opening in Appendix G, which further suggests a higher awareness about retailer around the data of store opening.

5.4 Net Benefit of Store Opening from Existing Customers

Out of the sample of 1.5 million customers, the store distance for 90,326 existing customers were affected by the six store openings of the retailer in 2003 (see Section 3.1). Thus, one store opening by the retailer may affect the store distance of approximately 15,000 existing customers in this sample. Since the 1.5 million customers sample is ten percent of the entire customer population of the retailer, a total of approximately 150,000 existing customers would be affected by one store opening of the retailer. Our estimates indicate that customers make an additional online purchases of US $6.59 due to store opening, but after accounting for the increase in the returns (both online and in store) of their online purchases after store opening, the net additional annual online purchase revenue was US $4.00. Similarly, the store opening results in the net increase of $48.35 in customers' store purchases.[26] Thus, one store opening of retailer results in an additional annual online revenue of 150,000 *4.00 = US $0.6 million and an additional online revenue of 150,000*48.35 = US $ 7.25 million from its existing customers. Note that the store opening would also result in additional sales from newly acquired customers.

6. Conclusions and Future Research Directions

[24] These results are available on demand from the authors

[25] These results are available on demand from the authors.

[26] This is obtained by estimating our specifications on the net store purchases (store purchase - store return) by our samples of affected and unaffected customers. The detailed results are available on request from the authors.

We designed a quasi-natural experiment and used customer-level data to estimate the causal effect of store openings by a fashion retailer in the U.S. on the purchase and return behavior of its existing customers. We employed several different analysis approaches (fixed effect, parametric propensity score matching, and nonparametric coarsened exact matching estimator) to obtain a range of estimates for the effect of store on online sales. We find a complementary effect of store opening on online sales – specifically the annual online purchase revenue per customer increased in the range US $5.07 – $7.47 (an increase of 23% - 33% over the mean online purchase revenue of $22.4 prior to store opening).[27] These estimates are not comparable with that obtained in earlier studies (Bell et al. 2015 and Avery et al. 2012) because of differences in product category and the unit of analysis – our study is on customer level data whereas the previous studies are on Zip code level data. Yet, these findings highlight our first contribution compared to these earlier papers. While these papers also found evidence of the complementary effect of stores on online sales, they did not identify whether this complementarity came from the acquisition of new customers in the zip code where a new store opened, or it came due to additional purchases made at the online channel by existing customers. We are able to point out that complementary effect of stores on online sales exists on the level of an individual consumer because of the availability of consumer level data.

In addition to identifying the complementary effect of stores on online sales, we use theory of planned behavior and prospect theory to identify two underlying mechanisms – store engagement and return effects - that could explain the complementary effect of store on online sales. We provide empirical evidence of these mechanisms on our data. We show existence of the store engagement effect with the increase in probability of store interactions for customers making online purchases after the store opening - this probability value increased in the range 0.11 – 0.18 (an increase of 25% – 41% over the average value of 0.44 prior to store opening). Similarly, we show existence of store return effect with the increase in percentage of online purchased quantity returned in store after the availability of store - this value increased in the range 2.49 – 3.30 (an increase of 102% – 135% over the mean value of 2.45 prior to store opening).

[27].

We further show that these effects are higher for customers who face a higher reduction in their store distance from store opening, which is line with theory and thus provides support to the definitions of the two effects. The identification of these two mechanisms is our second and the most important contribution to literature as no extant work has identified these mechanisms earlier.

The valence of these two effects may vary based on the category of products. For example, for commodity products where customers are certain about the product’s quality (brand) and the quantity they need e.g. diapers, neither the store engagement effect nor the return effect can be expected to be operative. But for commodity products where customers know the brands they prefer but are uncertain about the quantity they need, e.g. typical food items, the return effect is expected to be more salient (because extra items may need to be returned). For experience products where customers look for variety e.g. gourmet chocolates/ wines. Customers need to taste such products to decide whether they like it or not. However, once the packaging is open, it cannot be returned. Hence, the store engagement is important to showcase items that customers may like to try out, but return effect is not operative. For products with non-digital attributes and where personal tastes are important (e.g. apparel) we expect both store engagement and store return effect to be operative. As per TPB, the store visit is an important component of forming a positive attitude towards purchase of such products. Therefore, we expect stronger store engagement effect for products with non-digital attributes. Similarly, customers may care more about the social norms for highly visible products, such as apparel as compared to less visible products, such as book. Therefore, we expect stronger store engagement effect for highly visible products.

The managerial prescriptions to obtain the complementary effect of store vary based on the relative valence of the two effects. If the store engagement effect is predominant in driving the complementary effect of store, then the managers should: (1) Design appropriate worker incentive schemes to optimize its total (store + online) sales –the store workers may be compensated not only for the sales at that store, but also for online sales that occur from customers residing in the catchment area of that store. With this kind of compensation policy, the store workers will be incentivized to provide information about products that the customer may not purchase in store, but may purchase in future using the online channel, (2) offer promotions and host events at the store to increase store foot traffic because customer store visits enhance their engagement with the retailer. If the store return effect is predominant then managers should: (1) Design liberal omnichannel return policies. (2) Locate product returns counter in such a way that customers are exposed to retailer's products so they can easily find replacement products while returning products in store. (3) Examine the possibility of creating low cost return centers (if costs of opening a full-fledged store are exorbitant) where customers can easily return their online purchases.

Our findings also help managers to correctly compute the viability of an existing store and conduct the cost-benefit analysis of opening a new store. Our results indicate that availability of existing stores of the retailer selling non-digital goods results in about 25-30% increase in online sales from its existing customers. The retailer should account for the loss of this online sales besides the loss of store sales while analyzing the impact of closing its existing store. Similarly, the retailer should factor in the gain in online sales from the store while doing the cost-benefit analysis of opening a new store.

Our paper has some limitations that offers opportunities for future research. First, we assumed that the counterfactual purchase and return behavior of the affected customers can be estimated by the purchase and return behavior of the unaffected customers matched on observable demographic characteristics and past purchase and return behavior. Although, it is a reasonable assumption in our context, yet there is an outside possibility that the two groups of customers are systematically different on unobserved characteristics that may differentially influence their purchase and return behaviors. Second, the results in this paper are obtained for non-digital products, such as fashion apparel, accessories and home products and they may not be generalizable to other products. For instance, the store engagement and return effects may vary in magnitude based on the extent of non-digital attributes in the product under study. Third, although we show the evidence for existence of both mechanisms, we could not precisely compute how much complementary effect on online sales is attributable to individual mechanism. Fourth, the complementary effect of stores on the online sales of a multichannel retailer is dependent on the extent of competition it faces in a geography. In case of fierce price competition, the showrooming effect (Mehra et al. 2017) of the multichannel retailer may result in cannibalization in its store sales to the online sales of a pure online retailer or of other multichannel retailers. The estimated effects in the present paper could also be due to higher differentiation or branding effect of the retailer and may not be generalizable in other competitive settings. Further research is required to study the impact of store openings by retailers in different product categories and different competitive settings on the online sales of the multichannel retailers and the pure-play online retailers. Fifth, this paper examines the causal effect of facilitating customers' access on store channel on their purchase and return behaviors on store and online channels. It would be interesting to examine the effect in opposite direction, i.e., the effect of facilitating customers' access to the online channel, such as by reducing shipping fees, on their store and online purchase behavior. Finally, in line with research by Mithas and Rust (2016), one could examine how the market value of multi-channel retailers is explained by their emphasis on lowering cost (promotion of their online channel) and emphasis on boosting revenues (utilizing stores to do so).

References

Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes* 50, 179-211.

Anderson E. T., K. Hansen, D. Simester (2009). The Option Value of Return: Theory and Empirical Evidence. *Marketing Science* 28(3), 405-423.

Anderson, E. T., N. M. Fong., D. Simester, C. E. Tucker (2010). How Sales Taxes Affect Customer and Firm Behavior: The Role of Search on the Internet. *Journal of Marketing Research* 47(2), 229–239.

Ansari, A., C. F. Mela, S. A. Neslin (2008). Customer Channel Migration. *Journal of Marketing Research* 45(1), 60-76.

Avery, J., T. J. Steenburgh, J. Deighton, M. Caravella (2012). Adding Bricks to Clicks: Predicting the Patterns of Cross-Channel Elasticities over Time. *Journal of Marketing* 76(3), 96-111.

Bell, D., S. Gallino, A. Moreno (2015). Offline Showrooms and Customer Migration in Omni-Channel Retail. *Working Paper* (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2370535).

Bharadwaj, A., O. A. El Sawy, P. A. Pavlou, N. Venkatraman (2013). Digital Business Strategy: Toward a Next Generation of Insights. *MIS Quarterly* 37(2), 471-482.

Brynjolfsson, E., M. D. Smith (2000). "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," *Management Science* 46 (4), 563-85.

Brynjolfsson, E., Y. Hu, M. S. Rahman (2009). Battle of the Retail Channels: How Product Selection and Geography Drive Cross-Channel Competition. *Management Science* 55(11), 1755-1765.

Brynjolfsson, E., Y. Hu, M. D. Smith (2006). From Niches to Riches: The Anatomy of the Long Tail. *MIT Sloan Management Review* 47(4), 67-71.

Cachon, G. P., C. Terwiesch, Y. Xu (2005). Retail Assortment Planning in the Presence of Consumer Search. *Manufacturing and Services Operations Management* 7(4), 330-346.

Choi, J., D. R. Bell (2011). Preference Minorities and the Internet. *Journal of Marketing Research* 48, 670-682.

Cook, A. (2014). With Sales Falling at Department Stores, Nordstrom Plans Bigger Online Push. *Geekwire* February 22, 2014 (http://www.geekwire.com/2014/sales-falling-department-stores-nordstrom-plans-bigger-online-push/).

Danaher, P. J., I. W. Wilson, R. A. Davis (2003). A Comparison of Online and Offline Consumer Brand Loyalty. *Marketing Science* 22(4), 461-476.

Dennis, S. (2017). Unsustainable Customer Acquisition Costs Make Much of Ecommerce Profit Proof. *Forbes*. August 31.

Doupnik, E. (2017). Customers Exhaustively Research, Plan Purchases. *WWD*. October 20.

Forman, C., A. Ghose, A. Goldfarb (2009). Competition between Local and Electronic Markets: How the Benefit of Buying Online Depends on Where You Live. *Management Science* 55(1), 47-57.

Gallino, S., A. Moreno (2014). Integration of Online and Offline Channels in Retail: The Impact of Sharing Reliable Inventory Availability Information. *Management Science* 60(6), 1434-1451.

Geykens, I., K. Gielens, M. G. Dekimpe (2002). The Market Valuation of Internet Channel Additions. *Journal of Marketing* 66(2), 102-119.

Hirano, K., G. W. Imbens (2001). Estimation of Causal Effect using Propensity Score Weighting: An Application to Data on Right Heart Catheterization. *Health Services & Outcomes research Methodology* 2, 259-278.

Kushwaha, T., V. Shankar (2013). Are Multichannel Customers Really More Valuable? The Moderating Role of Product Category Characteristics. *Journal of Marketing* 77(4), 67-85.

Lal, R., M. Sarvary (1999). When and How is the Internet Likely to Decrease Price Competition? *Marketing Science* 18(4), 485–503.

McIntyre, D. A., A. E. M. Hess (2014). Nine Retailers Closing the Most Stores. *24/7 Wall Street* March 12, 2014 (http://www.usatoday.com/story/money/business/2014/03/12/retailers-store-closings/6333865/).

Mehra, A., S. Kumar, J. S. Raju (2017). Competitive Strategies for Brick-and-Mortar Stores to Counter "Showrooming." *Management Science* Forthcoming.

Nelson, G., A. Gupta, R. Kauffman (2012). Online and Offline Demand and Price Elasticities: Evidence from the Air Travel Industry. *Information Systems Research* 23(1), 164-181.

Petersen, J. A., V. Kumar (2009). Are Product Returns a Necessary Evil? Antecedents and Consequences. *Journal of Marketing* 73, 35-51.

Rajendran, K. N., G. J. Tellis (1994). Contextual and Temporal Components of Reference Price. *Journal of Marketing* 58(1), 22-34.

Tang, Q., M. Lin, Y. Kim (2016). Showrooming vs. Competing: The Role of Product Assortment and Price. *Working paper* (http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2732948).

Thaler, R. H. (2008). Mental Accounting and Consumer Choice. *Marketing Science* 27(1), 15-25.

Venkatesan, R., V. Kumar, N. Ravishanker (2007). Multichannel Shopping: Causes and Consequences. *Journal of Marketing* 71, 114-132.

Verhoef, P. C., S. A. Neslin, B. Vroomen (2007). Multichannel Customer Management: Understanding the Research-Shopper Phenomenon. *International Journal of Research in Marketing* 24,129-148.

Wood, S. L. (2001). Remote Purchase Environments: The Influence of Return Policy Leniency on Two-Stage Decision Processes. *Journal of Marketing Research* 38(2), 157-169.

APPENDIX

Appendix A: Comparison with Related Papers in Literature

Papers examining the effect of new store openings				
	Bell et al. 2015	Forman et al. 2009	Avery et al. 2009	Our paper
Customer-level data	No	No	No	Yes
Same retailer with both physical and online channel	Yes, they find a complementary effect of inventory showroom (not store) on online channel	No, they find that opening of physical store cannibalizes sales from an online retailer	Yes, they find complementary effect of physical store on online channel	Yes, we find a complementary effect of physical store on online channel
Store opening as a quasi-experiment	Yes	No	Yes	Yes
Show underlying mechanisms for effect of store opening on online channel sales	No	No	No	Yes, we show evidence of two mechanisms – store engagement effect and store return effect
Show varying effect of store opening with variation in customers' store distance	No	No	No	Yes, we show that higher reduction in customers' store distance results in higher complementary effect of store opening on customers' online purchases as well as higher store engagement and return effects

Papers examining the effect of product returns			
	Anderson et al. 2009	Peterson and Kumar 2009	Our paper
Data with and without easier option to return	No. They use customer-level transaction data to build and estimate a structural model to estimate option value of return on	No. They use customer-level transaction data to analyze how customers' current product return behavior affects their future purchase	Yes. We design a quasi-experiment with data on both with and without nearby store to

	customers' purchase behavior.	behavior and firms future marketing strategy.	estimate the causal effect of easier product returns at store on customers' online purchase behavior.
Estimate the effect of easier return on one channel on demand on another channel	No. They estimate the option to return on catalog channel on the demand on the same channel	No. They estimate the effect of overall product returns on the overall future demand. They don't specifically show the effect of product returns on one channel on purchase on another channel.	Yes. We estimate the effect of easier option to return in store on the demand on the online channel

Table A1: Comparison with related papers

Appendix B: Checking for Parallel Trends Assumption

It may be possible that the affected customers online purchase behavior may have a higher upward trend than that of unaffected customers. In this case, the treatment effect identified in the DD estimation will be due to the differences in their pre-existing trends in online purchases prior to store opening and not because of the store opening. Thus, we need to check that the trends in online purchase variables are similar for the affected and unaffected customers prior to store openings. We run the following fixed-effect specifications in the pre-store opening period to test for the parallel trends between affected and unaffected customers

$$Y_{it} = \beta_i + \beta_1 \, Post1_t + \beta_2 \, (Treat \times Post1)_{it} + \varepsilon_{it} \quad ; \qquad \text{--- (A1)}$$

where t = t_1, and t_2 denote period 1 and period 2 in Figure 1; i denotes customers; $Treat_i$ is an indicator variable equal to one if i is an affected customer and zero otherwise; and $Post1_t$ is an indicator variable equal to one for t = t_2 (period 2) and zero otherwise. The left hand side variable Y_{it} denotes the three online purchase variables for showing complementary effect of store opening for customer i in period t – online purchase probability, online purchase quantity, and online purchase revenue. Coefficient β_i captures customer fixed effects that accounts for time invariant unobserved differences across customers. The coefficient of *Treat* × *Post1* (i.e., β_2) is of interest, and it captures the differential trends in online purchase behavior of affected customers as compared to unaffected customers. An insignificant coefficient estimate of β_2 would indicate a statistically similar online purchase behavior for affected and unaffected customers. Table B1 report the resulting coefficient estimates. We find an insignificant coefficient estimates for *Treat* × *Post1* for all online purchase variables.

Pre-store opening period analysis	Coeff	St. Err	Coeff	St. Err	Coeff	St. Err
	Online purchase probability		Online purchase quantity		Online purchase revenue	
Post1	0.015^{***}	0.001	0.14^{***}	0.03	6.55^{***}	1.30
Treat*post1	-0.006	0.004	-0.06	0.04	-2.27	2.20
Intercept	0.058^{***}	0.000	0.25^{***}	0.01	13.98^{***}	0.60
N (groups)	436746 (218373)		436746 (218373)		436746 (218373)	

***, **, * = statistically significant at α = 0.01, α = 0.05, and α = 0.10 levels (two-sided test), respectively. Standard errors cluster corrected at customer level. Two periods observations for 218373 customers -17277 affected +201096 unaffected.

Table B1: Parallel trends for online purchase variables

We also estimate a similar model to ensure that the affected and unaffected customers have similar trends of dependent variables for store engagement effect (probability of store purchase | online purchase) and for store return effect (% of online purchase quantity returned in store) prior to store openings. Since the store engagement and return effects are shown from the purchase and return behavior of customers who make online transactions, for this analysis we use data on only those customers who make online transactions with the retailer. Table B2 reports the resulting coefficient estimates, which shows similar trends of the two categories of customers in the period prior to store openings.

Pre-store opening period analysis	Coeff	St. Err	Coeff	St. Err
	Prob (store purchase |online purchase)		% Online purchase quantity returned in store	
Post1	-0.012^{*}	0.007	3.24^{***}	0.299
Treat*post1	0.010	0.022	-1.24	0.964
Intercept	0.603^{***}	0.004	3.45^{***}	0.154
N (groups)	28482 (23159)		28482 (23159)	

***, **, * = statistically significant at α = 0.01, α = 0.05, and α = 0.10 levels (two-sided test), respectively. Standard errors cluster corrected at customer level. Unbalanced panel of 28482 customer-period observations for 23159 unique customers.

Table B2: Parallel trends of variables used in store engagement and return effect

Appendix C: Summary Statistics of Variables Used for Matching

We provide the distribution of matching variables at different quantile values in Table C1 to check the balance between the full sample of affected and unaffected customers.

Variables	Customer Group	Percentile Values							Mean	St. Dev.
		0	25	50	75	90	95	100		
Customer-level cumulative RFM variable prior to store opening										
Time between last purchase and store opening date in weeks	Unaff	0.03	2.37	14.20	36.00	44.53	46.43	49.57	19.32	17.06
	Aff	0.01	7.27	21.90	36.97	44.70	46.67	49.91	22.53	16.01
No. of online purchase interactions	Unaff	0	0	0	0	1	2	160	0.37	1.61
	Aff	0	0	0	0	1	2	76	0.48	1.91
No. of store purchase interactions	Unaff	0	1	3	11	27	42	473	10.03	18.27
	Aff	0	0	1	5	14	23	254	5.12	10.85
No. of online purchased items	Unaff	0	0	0	0	2	5	639	0.96	15.23
	Aff	0	0	0	0	2	6	370	1.21	6.31
No. of store purchased items	Unaff	0	2	9	33	88	147	6520	33.63	76.44
	Aff	0	0	3	15	47	82	2602	17.40	46.49
Online purchase revenue (US $)	Unaff	0	0	0	0	98	259.9	25496	52.95	651.75
	Aff	0	0	0	0	151.6	341.2	29058	68.13	393.12
Store purchase revenue (US $)	Unaff	0	75	399.9	1507.2	4128.2	7123.3	487040.8	1631.7	4209.7
	Aff	0	0	153.35	801.6	2443.5	4357.4	509364.3	959.6	4640.7
No. of online return interactions	Unaff	0	0	0	0	0	0	107	0.07	0.58
	Aff	0	0	0	0	0	1	39	0.09	0.63
No. of store return interactions	Unaff	0	0	0	1	4	7	231	1.54	4.95
	Aff	0	0	0	0	2	4	83	0.75	2.76
No. of online returned items	Unaff	0	0	0	0	0	0	352	0.13	1.41
	Aff	0	0	0	0	0	1	150	0.17	1.71
No. of store returned items	Unaff	0	0	0	2	6	13	880	2.90	12.60
	Aff	0	0	0	0	3	7	571	1.42	8.38
Online return revenue (US $)	Unaff	0	0	0	0	0	0	14652.4	8.57	86.75
	Aff	0	0	0	0	0	38.1	11887.4	11.54	122.17
Store return revenue (US $)	Unaff	0	0	0	98	451	929.57	69345.4	204.76	957.59
	Aff	0	0	0	0	235	530	89031	119.30	1019.11
Customer-level socio-demographic variables										
Store distance in miles before store opening	Unaff	0	6.1	24.9	68.2	156.7	204.9	794.5	53.00	73.80
	Aff	0.9	22.9	154.1	234.9	338.7	411.3	473.0	148.01	127.41
Income category	Unaff	1	1	2	4	5	6	6	2.38	1.43
	Aff	1	1	2	4	5	6	6	2.37	1.46
Age category	Unaff	1	3	5	6	7	7	7	4.88	1.53
	Aff	1	3	5	6	7	7	7	4.80	1.52
Customer-level purchase trends prior to store opening										
Change in no. of items purchased online from 01-02 to 02-03	Unaff	-158	0	0	0	0	1	639	0.14	14.41
	Aff	-131	0	0	0	0	1	279	0.08	3.25
Change in no. of items purchased in store from 01-02 to 02-03	Unaff	-409	-1	0	0	8	16	1782	0.16	13.86
	Aff	-210	0	0	0	5	11	197	0.10	9.81
Zip code-level aggregate variables in 2003										
Population in "000"	Unaff	8.21	14.88	25.43	37.27	48.97	56.58	113.65	27.16	16.56
	Aff	6.32	16.42	28.81	39.81	55.00	77.30	116.85	31.35	21.13
Median Age	Unaff	12.5	34.9	38.1	41	44	46.5	78.5	38.09	5.45
	Aff	19.2	33.6	36.5	39.8	44.3	47.3	65.8	37.09	5.95
Median Income in "000"	Unaff	8.13	45.21	58.14	75.66	95.73	111.46	234.79	63.05	25.06
	Aff	6.98	40.46	54.28	70.58	90.39	104.38	162.17	57.75	22.75
Internet usage index	Unaff	0.49	0.63	0.65	0.68	0.70	0.71	0.76	0.66	0.03
	Aff	0.55	0.62	0.65	0.68	0.70	0.72	0.75	0.65	0.04
Computer usage index	Unaff	0	92	122	163	219	272	668	137.65	67.98
	Aff	0	84	118	151	212	234	352	126.86	58.40
Apparel preference index	Unaff	0	90	114	146	193	235	561	126.93	54.73
	Aff	0	87	111	139	185	207	318	120.29	46.04

“Unaff” and “Aff”, respectively, denote unaffected and affected customers

Table C1: Summary statistics of variables used for matching

From Table C1, it is apparent that the full sample of affected and unaffected customers significantly differ in distribution of matching variables, and we need to use matching methods to account for these differences.

Appendix D: Balancing Treated and Control Groups with Propensity Score Weighting

The basic purpose of inverse propensity score weighting of treated and control observations is to balance them on various covariates used in their matching (Hirano and Imbens 2002). Accordingly, we examine the balance between the two groups of customers on different covariates with and without weighting in this appendix. In Table D1: columns two and three report the means of different covariates; column four report the t-statistics for difference in these means; column five and six report the weighted means of the same covariates; and the last column report the t-statistics for difference in these weighted means.

Mean values of variables	Treated	Control	*t-stats*	Treated (Weighted)	Control (Weighted)	*t-stats* (Weighted)
Customer-level purchase/return variables prior to store opening						
Time between last purchase and store opening date in weeks	22.53	19.32	25.13	21.06	20.06	0.85
No. of online purchase interactions	0.48	0.37	7.27	0.46	0.40	0.57
No. of store purchase interactions	5.12	10.03	-53.33	7.84	8.17	-0.99
No. of online purchased items	1.21	0.96	4.11	1.04	1.10	-0.31
No. of store purchased items	17.40	33.63	-41.33	19.76	22.61	-1.58
Online purchase revenue (US $)	68.13	52.95	4.57	60.50	57.87	0.26
Store purchase revenue (US $)	959.59	1631.71	-18.40	1047.62	1103.24	-0.74
No. of online return interactions	0.09	0.07	4.13	0.08	0.07	1.01
No. of store return interactions	0.75	1.54	-33.31	0.86	0.91	-1.77
No. of online returned items	0.17	0.13	3.05	0.17	0.13	3.05
No. of store returned items	1.42	2.90	-21.19	1.81	1.90	-0.90
Online return revenue (US $)	11.54	8.57	3.13	9.39	8.95	0.56
Store return revenue (US $)	119.30	204.76	-10.63	159.30	172.76	-1.67
Customer-level socio-demographic variables						
Store distance in miles	148.01	53.00	96.63	77.40	72.25	1.10
Income category	2.37	2.38	-1.63	2.32	2.40	-0.53
Age category	4.80	4.88	-6.07	4.82	4.86	-0.09
Customer-level purchase trends prior to store opening						
Change in no. of items purchased online from 01-02 to 02-03	0.10	0.16	-0.78	0.11	0.15	-0.49
Change in no. of items purchased in store from 01-02 to 02-03	0.08	0.14	-1.37	0.09	0.12	-0.83
Zip code -level aggregate variables in 2003						
Population in '000'	31.35	27.16	25.43	31.17	30.93	0.93
Median age	37.09	38.09	-21.24	41.04	42.16	-1.51
Median household income in '000'	57.75	63.05	-29.12	61.94	64.02	-0.13
Internet usage index	0.65	0.66	-15.57	0.74	0.77	-0.39
Cable TV penetration index	126.86	137.65	-22.99	121.10	131.13	-1.47
Apparel preference index	120.29	126.93	-17.91	123.99	125.75	-0.27

Table D1: Balance between affected and unaffected customers

We find significant t-statistics for differences in means of all covariates, which indicates that the two groups of customers differ significantly on almost all variables. However, after inverse propensity score weighting, the means of most of the covariates become closer, and we find insignificant t-statistics for differences in weighted means of almost all covariates. Thus, weighting by propensity scores have balanced the samples of affected and unaffected customers extremely well.

Appendix E: Results of Coarsened Exact Matching Estimators

We used a non-parametric exact matching method, Coarsened Exact Matching (CEM) estimators, to estimate the treatment effect of store opening (*ATE*). Based on the distribution of each variable, we manually create different sized bins so as to capture similar purchase behavior of customers inside the range of values of variables in a bin. We tried several different manual coarsening of bin sizes for different variables in our data and show two example bins that we created in Table E1.

Variables	Manual Coarsening	Cutoff points for manual bins
Customer-level cumulative RFM variable prior to store opening		
Time between last purchase and store opening date in weeks	1	(0, 2, 10, 20, 50)
	2	(0, 2, 5, 10, 20, 30, 50)
No. of online purchase interactions	1	(0, 5, 25, 50, 200)
	2	(0, 2, 5, 25, 50, 100, 200)
No. of store purchase interactions	1	(0, 5, 10, 25, 50, 500)
	2	(0, 2, 5, 10, 25, 50, 100, 500)
No. of online purchased items	1	(0, 5, 25, 50, 100, 1000)
	2	(0, 2, 5, 10, 25, 50, 100, 500, 1000)
No. of store purchased items	1	(0, 5, 25, 50, 100, 1000, 10000)
	2	(0, 2, 5, 10, 25, 50, 100, 500, 1000, 10000)
Online purchase revenue (US $)	1	(0, 25, 100, 1000, 30000)
	2	(0, 25, 50, 100, 500, 1000, 10000, 30000)
Store purchase revenue (US $)	1	(0, 100, 500, 1000, 10000, 100000)
	2	(0, 50, 100, 500, 1000, 5000, 10000, 100000)
No. of online return interactions	1	(0, 1, 10, 50, 100)
	2	(0, 1, 5, 10, 25, 50, 100)
No. of store return interactions	1	(0, 5, 10, 50, 250)
	2	(0, 2, 5, 10, 25, 50, 250)
No. of online returned items	1	(0, 5, 50, 500)
	2	(0, 1, 5, 10, 50, 500)

No. of store returned items	1	(0, 10, 50, 250, 1000)
	2	(0, 2, 5, 10, 25, 50, 250, 1000)
Online return revenue (US $)	1	(0, 50, 1000, 15000)
	2	(0, 50, 250, 1000, 15000)
Store return revenue (US $)	1	(0, 50, 1000, 100000)
	2	(0, 50, 250, 1000, 10000, 100000)
Customer-level socio-demographic variables		
Store distance in miles before store opening	1	(0, 10, 50, 100, 250, 1000)
	2	(0, 5, 10, 25, 50, 100, 250, 1000)
Income category	1	(0, 1.1, 2.1, 3.1, 4.1, 5.1, 6.1)
	2	
Age category	1	(0, 1.1, 2.1, 3.1, 4.1, 5.1, 6.1, 7.1)
	2	
Customer-level purchase trends prior to store opening		
Change in no. of items purchased online from 01-02 to 02-03	1	(-200, 0, 1, 50, 1000)
	2	(-200, 0, 1, 50, 100, 1000)
Change in no. of items purchased in store from 01-02 to 02-03	1	(-500, -1, 0, 10, 25, 2000)
	2	(-500, -1, 0, 10, 25, 100, 500, 2000)
Zip code-level aggregate variables in 2003		
Population in "000"	1	(0, 10, 25, 50, 100)
	2	
Median Age	1	(0, 20, 30, 40, 50, 60, 80)
	2	
Median Income in "000"	1	(0, 25, 50, 75, 100, 150, 250)
	2	
Internet usage index	1	(0, 0.5, 0.6, 0.7, 0.8)
	2	
Computer usage index	1	(0, 100, 125, 150, 250, 700)
	2	
Apparel preference index	1	(0, 100, 150, 200, 300, 600)
	2	

Table E1- Bin cut-off values for different variables from two manual coarsening examples

In Table E2, we report the treatment effect estimates for two examples of manual coarsening and propensity score based doubly robust estimators. We find that the two manual coarsening results in different numbers of matched affected and unaffected customers and accordingly different imbalances ($\mathcal{L}_1$ statistic values). But, we still find similar sign and significance of the average treatment effect estimates for the two manual coarsening, which suggests the robustness of our treatment effect estimates. We further find that our treatment effect estimates from coarsened exact matching are similar in sign, significance, and magnitudes to that from propensity score matching, which further shows the robustness of our estimates.

	Propensity score matching		Manual Coarsening 1		Manual Coarsening 2	
Number of matched (total) affected customers	Not applicable		13995 (17277)		9150 (17277)	
Number of matched (total) unaffected customers	Not applicable		153323 (201096)		95782 (201096)	
Overall imbalance ($\mathcal{L}_1$ statistics)	Not applicable		0.765		0.889	
Average Treatment Effect	Coeff.	St. Err.	Coeff.	St. Err.	Coeff.	St. Err.
Complementary effect of store opening on online purchases						
Online purchase probability	0.029^{**}	0.014	0.056^{***}	0.005	0.058^{***}	0.005
Online purchase quantity	0.056^{**}	0.025	0.096^{***}	0.018	0.118^{***}	0.015
Online purchase revenue	6.59^{***}	3.03	6.85^{***}	1.12	7.47^{***}	1.040
Store engagement effect						
Prob (store purchase | online purchase)	0.11^{**}	0.05	0.13^{***}	0.015	0.127^{***}	0.016
Store return effect						
% Online purchase quantity returned in store	2.76^{**}	1.34	2.49^{**}	1.06	3.30^{**}	1.45

***, **, * = statistically significant at α = 0.01, α = 0.05, and α = 0.10 levels (two-sided test), respectively.
Robust standard errors used.

Table E2: Comparison of CEM and Propensity Score matching estimators

Appendix F: Variations in Store Opening Effect with Change in Store Distance

The extent of complementary effect of store opening depends on how much the store distance (store access costs) reduces for a customer. The higher the reduction in store distance, the higher will be the number of store visits for purchase and return. To test this intuition, we examine how the effects of store openings on customers' online purchase behaviors are moderated by the change in their store distances. For this analysis, we use the reduction in a customer's store distances in miles due store opening (denoted as *Chdist*) as the treatment variable. In Table 1, we found a wide variation in the *Chdist* values across the sample of affected customers (mean value = 70 miles, standard deviation = 79 miles, and the median value = 34 miles), which indicates the wide variations in treatment of store opening across our sample of affected customers. We estimate fixed effect specifications (1a) & (1b), and the matching specification (2), with *Chdist* as the treatment variable in place of indicator variable *Treat*, on our data and report the coefficient estimates in Table 10.

The coefficient of *Post*Chdist* is of interest, as it captures the change in a customer's online purchase behavior due to the reduction in her store distance by one mile. We find a positive and significant estimate for *Post*Chdist* for the three online purchase variables from all specifications, which indicates that customers' online purchase probabilities / quantities / revenue increase with the magnitude of reduction in their store distances. The *Post*Chdist* estimate of 0.091 for online purchase revenue indicates an increase in customer's annual online purchase revenue by US $ 0.091 with reduction of her store distance by one mile, which translates into an increase in annual online purchase revenue of US $6.37 for the mean reduction in store distance of 70 miles for affected customers (Table 2).

Variables	Online purchase probability		Online purchase quantity		Online purchase revenue	
	Coeff. Est.	Std. Err	Coeff. Est.	Std. Err	Coeff. Est.	Std. Err
Fixed effects specification (1)						
Post	0.027^{***}	0.001	0.170^{***}	0.017	14.09^{***}	0.74
*Post*Chdist*	0.0003^{***}	0.000	0.001^{**}	0.000	0.08^{***}	0.01
Intercept	0.065^{***}	0.000	0.318^{***}	0.008	17.17^{***}	0.36
N (groups)	873492 (218373)					
R sq. value	0.60		0.68		0.71	
Matching estimator specification (2)						
Chdist	0.0002^{***}	0.000	0.0009^{***}	0.000	0.052^{***}	0.010
Post	0.0430^{***}	0.017	0.2519^{***}	0.061	19.47^{***}	4.94
*Chdist*post*	0.0001^{***}	0.000	0.0005^{***}	0.000	0.091^{***}	0.019
Intercept	0.0630^{***}	0.006	0.2682^{***}	0.017	14.35^{***}	0.79
N	873492					

***, **, * = statistically significant at $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.10$ levels (two-sided test), respectively.

Table F1: Variations in store opening effect with change in store distance

We further examine whether the extent of store engagement and return effects on customers' online purchases are also moderated by the magnitude of their reduction in store distances. For this analyses, we replicate the analyses in Sections 5.2.1, 5.2.2, and 5.2.3 by replacing the indicator treatment variable *Treat* with continuous treatment variable *Chdist*. The resulting coefficient estimate are reported in Table 11. We find a positive and significant coefficient estimate for *Post*Chdist*, the coefficient of interest, in all specifications for variables pertaining to both store engagement and return effects. Specifically, we find a higher probability of association between store and online purchases for customers who face a higher reduction in their store distances after store opening, i.e., a positive association between store engagement effect and reduction in store distance. Moreover, we find a higher percentage of online purchases returned in store for affected online customers with higher reduction in their store distances.

Variables	Prob. (store purchase | online purchase)		% Online purchase Qty returned in store | online purchase	
	Coeff. Est.	Std. Err.	Coeff. Est.	Std. Err.
Fixed effects specification (1)				
Post	0.004	0.004	2.82***	0.21
*Post*Chdist*	0.001***	0.000	0.02***	0.005
Intercept	0.638***	0.002	5.89***	0.12
N (No. of groups)	69303 (43613)			
R Sq. value	0.41		0.32	
Matching estimator specification (2)				
Chdist	-0.002***	0.000	-0.023***	0.003
Post	0.101**	0.051	3.398***	0.179
*Chdist*post*	0.001***	0.000	0.015***	0.004
Intercept	0.661***	0.035	3.873***	0.142
N	69303			

***, **, * = statistically significant at $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.10$ levels (two-sided test), respectively.

Table F2: Variations in store engagement and return effects with change in store distance

Appendix G: Store Opening Effect with inclusion of store opening period

We estimate our specifications without excluding 6 months data around the date of store opening. The results are reported in Table G1. We find a higher magnitude of treatment effect estimates of store opening on online purchase probability, quantity, and revenue, which suggests a higher awareness about retailer around the data of store opening.

Variables	Online purchase probability		Online purchase quantity		Online purchase revenue	
	Coeff. Est.	Std. Err	Coeff. Est.	Std. Err	Coeff. Est.	Std. Err
Fixed Effect Specification (1)						
Post	0.021***	0.001	0.19***	0.04	17.21***	0.82
*Treat*post*	0.026***	0.002	0.13***	0.05	8.13***	1.72
Intercept	0.074***	0.000	0.62***	0.11	19.28***	0.96
N (No. of Groups)	873492 (218373)					
R Square value	0.54		0.64		0.63	
Matching Estimator Specification (2)						
Treat	-0.011	0.010	-0.13	0.078	-7.12	5.18
Post	0.034***	0.003	0.26***	0.036	18.93***	3.12
*Treat*post*	0.033**	0.014	0.062**	0.025	7.12**	3.51
Intercept	0.078***	0.002	0.42***	0.022	20.61***	1.47
N	873492					

Table G1