Does Integration of Bricks With Clicks Affect Online-Offline Price Dispersion?

Does Integration of Bricks with Clicks Affect Online-Offline Price Dispersion?*

Siddhant Agrawal, Harvard University†

March 29, 2019

Abstract

The technological advancements in mobile computing and augmented reality have given rise to omnichannel retailing. This paper studies the effect of increasing omnichannelness on online-offline price dispersion. I construct a measure of omnichannelness — omnichannel index — using detailed omnichannel data from Digital Commerce 360's omnichannel reports. Using the omnichannel index, I find a negative association between omnichannelness of retailers and online-offline price dispersion: more omnichannel retailers present lower price differences across channels. This finding is robust across three measures of price difference — binary indicator for price difference, relative price difference, and absolute relative price difference. Among the omnichannel variables, integration features and in-store features have a negative relationship with price dispersion, while geographical variables are positively associated with price dispersion. I also find that an increase in omnichannelness is associated with increased likelihood of having synchronized sales across the channels.

Keywords: Geographical Features, In-Store Features, Integration Features, Omnichannel, Omnichannel Index, Pricing Strategies, Price Dispersion, Reputational Risk, Synchronized Sale

*I would like to thank my supervisor Antonio Moreno for his guidance, wonderful support, and kind advice as I worked on my senior thesis. I am indebted to Alberto F. Cavallo for disclosing the names of retailers in BPP price data and sharing other information for my thesis. I am very thankful to Alex Albright for continuous feedback and support throughout the process. I am also thankful to the Billion Price Project team at MIT for making the price data publicly available through Harvard/MIT dataverse.

†Email: siddhantagrawal@harvard.edu

1 Introduction

The technological advancements in mobile computing and augmented reality have led to a new revolution in retail: omnichannel strategy (Brynjolfsson, Hu and Rahman, 2013). Rather than being purely online or brick-and-mortar players, retailers are integrating the online and offline channels to provide consumers with a seamless shopping experience regardless of the channel used (Bell, Gallino and Moreno, 2014). In addition to "traditional" online channels like websites, many retailers are engaging with customers through mobile apps and social media (Piotrowicz and Cuthbertson, 2014). As retailers make strategic decisions to remain competitive in the age of omnichannel retailing, an important factor they must consider is their pricing strategy across channels. Do retailers charge the same price across channels for an identical product as they adopt omnichannel strategies?

This paper studies the effect of the omnichannel revolution on online-offline price dispersion. Retailers often employ price discrimination whereby they charge customers different prices depending on their perceived ability to pay and other market forces. This pricing strategy has become increasingly complex because of the abundance of data available on customers' preferences and development of sophisticated algorithms to predict future demand. One form of price discrimination is channel-based price discrimination which involves charging different prices across channels for an identical product. While well-implemented channel-based price discrimination can increase profits (Grewal et al., 2010), retailers risk potential customer irritation that could result from difference in prices across channels. Retailers are concerned about their brand name and desire to convey a sense of fairness. This reputational cost has further increased due to the ubiquity of smartphones, which allow customers to instantly compare prices across channels. Likewise, menu costs associated with changing prices increase as retailers become more omnichannel (Stamatopoulos, Bassamboo and Moreno, 2017). These factors present an important challenge for omnichannel retailers in terms of strategic management and coordination of online and offline pricing. I hypothesize that as retailers become more omnichannel, the price dispersion of their products across the channels should decrease. This follows from the fact that an increase in omnichannelness raises the costs, but does not substantially affect the profits derived from channel-based price discrimination.

I empirically test the association between omnichannelness and online-offline price dispersion in this paper. I use data from MIT's Billion Prices Project (BPP), which features both online and offline prices for a set of products (Cavallo and Rigobon, 2016). The BPP team carried out the first large-scale collection of prices across channels simultaneously. I merge the BPP price dataset with omnichannel data obtained from Digital Commerce 360's omnichannel reports, which quantitatively describe the omnichannel developments in US retail (Digital Commerce 360, 2016-2017). I construct a measure of omnichannelness — omnichannel index — using detailed data on omnichannel variables from the reports. Using the index, I study the association between omnichannelness and three measures of price dispersion: binary price difference, relative price difference and absolute relative price difference[1]. To my knowledge, this is the first study about how omnichannel strategies affect cross-channel price dispersion.

I also divide the omnichannel variables into three categories — integration features, in-store features, and geographical features — based on their level of relevance to online-offline price dispersion. Then I study the association between the three categories of omnichannel variables and online-offline price dispersion to better understand the mechanism behind how omnichannelness might affect price dispersion. Two variables — "Buy Online Pick Up in Store" and "Online Price Matching" — are especially interesting for price dispersion, so I consider them individually. Finally, I examine if there is any relationship between omnichannelness and features like synchronized sale (discount) that would be expected of omnichannel retailers.

I find a statistically significant and negative relationship between omnichannelness of retailers and online-offline price dispersion of their products. This finding is robust across the three different measures of price difference. Among the three categories of omnichannel variables, the results show that integration and in-store omnichannel features have negative relationship with price dispersion, while geographical features have positive association. As per the individual omnichannel variables, "Buy Online Pick Up in Store" is negatively associated with price dispersion, but "Online Price Matching" does not have a consistent relationship with price dispersion. Finally, an increase in omnichannelness is associated with increased likelihood of having synchronized sale (discount) across the channels. Since this study is the first of its kind and historical omnichannel data is not readily available, the correlations reported in this paper should provide useful insights on how omnichannelness and online-offline price dispersion are related.

This paper proceeds in six sections: I do a literature review in section 2 and provide

[1] Absolute relative price difference is the absolute value of relative price difference. This measure of price difference ignores the direction of the price difference.

a theoretical framework in section 3. The BPP price data and Digital Commerce 360s omnichannel data are described in section 4. In section 5, I discuss construction of the omnichannel index and division of omnichannel variables into three categories. Section 6 provides empirical models and findings related to the relationship between omnichannelness and online-offline price dispersion. Section 7 concludes and discusses the limitations of this study.

2 Literature Review

The majority of the prior studies on price dispersion has focused on dispersion within a channel— either online or offline — across the retailers. Following the Internet boom in 1990s, it was expected that the market on the Internet would be efficient where customers are fully informed of prices and product offerings. Studies emerged hypothesizing that the unique features of the Internet would lead to a decrease in price dispersion across retailers on the Internet. Baye et al. (2006) survey a large literature and find that the web has not reduced price dispersion across different retailers. There are few studies which compare prices of products across channels (brick-and-mortar vs online) within a retailer. This is partly due to the inherent difficulty in simultaneously collecting online and offline prices of the products. This data collection is further complicated when shipping costs, taxes, and sale are considered.

Brynjolfsson and Smith (2000) examined the prices of compact disks (CDs) and books, both of which are homogeneous goods, across conventional and Internet retailers. They found that prices on the Internet for these two goods were 9-16% less than prices of identical products in the physical stores. Likewise, Clay et al. (2002) compared prices for 107 books on 14 online and 2 physical stores. They found similar average prices across channels, but substantial price dispersion within the online channel. These two studies, even though they focus on a narrow category of homogeneous products, have contradicting findings about price dispersion across channels. Clay et al. (2002) argue that the disparate findings can be accounted for by differences in the samples of books and stores, the weighting of samples, and the treatment of shipping costs and the opportunity cost of time.

When considering a broader categories of products, the literature is divided about existence of price dispersion across channels in multichannel retail. It must be emphasized that multichannel retail is different from omnichannel retail in that being multichannel does not necessitate the integration of channels. It simply means that there are multiple channels that customers can use to interact with a retailer, but those channels jointly may not provide customers with a seamless shopping experience. Wolk and Ebling (2010) find that multichannel retailers do engage in channel-based price differentiation, and that big companies with market power are more likely to engage in such behavior. The findings of this paper deviate from prior studies that had failed to find any evidence of channel-based price discrimination (Ancarani and Shankar, 2004).

Cavallo (2017) carried out the first large-scale comparison of offline and online prices in retail. The Billion Prices Project team at MIT (Cavallo and Rigobon, 2016) simultaneously collected both offline and online prices of randomly sampled goods from 56 largest retailers in 10 countries: Argentina, Australia, Brazil, Canada, China, Germany, Japan, South Africa, United Kingdom, and the United States. Using this novel dataset, Cavallo finds that online and offline price levels are identical 72% of the time, with significant heterogeneity at the country, sector and retailer level. Cavallo's study provides evidence that there is price dispersion across channels, but it is small.

Because the omnichannel revolution in retail is very recent, there are very few studies related to pricing strategies or coordination of other information (like inventory) across channels in omnichannel retail. Identifying the challenges associated with inventory management and pricing in omnichannel retail, Harsha, Subramanian and Uichanco (2016) develop an optimization model which jointly optimizes cross-channel fulfillment inventories and lifecycle channel prices. Gao and Su (2016) use a theoretical framework to study how retailers can effectively deliver online and offline information to omnichannel consumers who strategically choose whether to gather information online or offline and whether to buy products online or offline. Likewise, (Bell, Gallino and Moreno, 2014) analyze the impact of the implementation of a "Buy Online Pick Up in Store (BOPS)" and find that implementation of this project is associated with a reduction in online sales and an increase in store sales and traffic. I did not find any academic paper that investigates how an increase in omnichannelness affects online-offline price dispersion. Thus, I use the price data collected by the BPP team to study price dispersion in omnichannel retail. In the theoretical framework on price dispersion in omnichannel retail described in the next section, I hypothesize that online-offline price dispersion should decrease as retailers become more omnichannel.

3 Theoretical Framework

In this section, I present a simple theoretical model to illustrate how a change in omnichannelness of a retailer affects online-offline price dispersion of its products. Cavallo (2017) identifies that (unsynchronized) sales/discounts tend to create some discrepancies between prices across channels. An unsychronized sale is a condition where a sale is provided on one channel, but not the other. For a sale to be synchronized, it has to be present on both the channels or on neither channel[2]. Cavallo also mentions that some of the price differences are sector-level behavior. He finds that a drug retailer is more likely to have an online-offline price difference than a clothing retailer because customers would be willing to pay a premium for immediate access to a drug in a physical store. The two reasons for price dispersion identified by Cavallo — unsynchronized sale and sector-level behavior — are different forms of channel-based price discrimination. Below, I consider the incentives associated with channel-based price discrimination and how they are affected by change in omnichannelness.

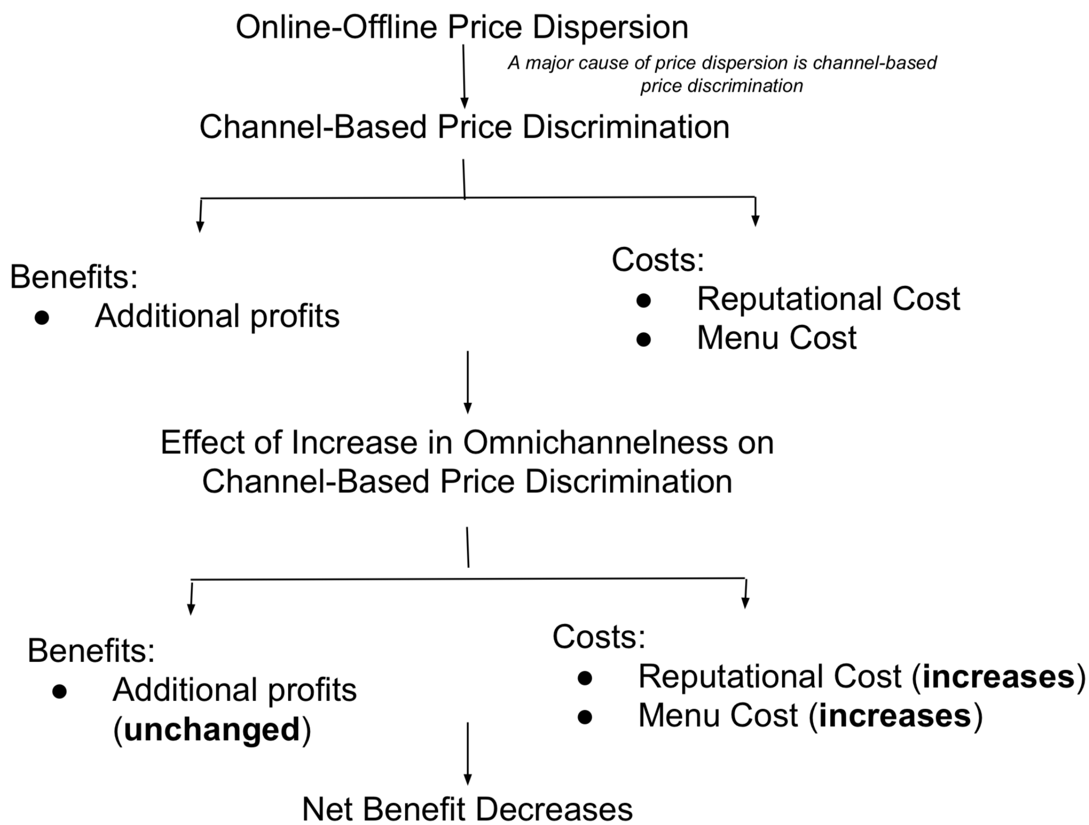
The literature on channel-based price discrimination identifies the benefits as well as the costs of such pricing strategy. It has been shown that retailers can increase their profits by applying channel-based price discrimination (Wolk and Ebling, 2010). Channel-based price discrimination has two main costs: (*i*) reputational cost and (*ii*) menu cost. Reputational cost is the risk of alienating the customers by charging different prices across channels. Price dispersion can evoke a sense of unfairness among customers, which might lead to a decline in their trust level and loyalty toward the retailer (Vogel and Paul, 2015). Likewise, channel-based price discrimination requires coordination of prices across channels. Menu cost is the expense associated with changing and coordinating prices across channels (Stamatopoulos, Bassamboo and Moreno, 2017).

As a retailer becomes more omnichannel, its net gain from channel-based price discrimination decreases – the costs of channel-based price discrimination increase while the benefits remain unchanged. Because of increased integration of channels, customers can further easily discern difference in prices across channels, which results in a higher reputational cost for retailers. Likewise, better integration would also mean an increase in coordination of prices across channels resulting in increased menu costs. An increase in omnichannelness does not affect retailer's ability to profit from channel-based price discrimination. It must be emphasized that an increase in omnichannelness can increase profits, but that is usually

[2] Synchronized sales need not eliminate price dispersion. For example, there would be a dispersion if the amount of sale differs across the channels.

a consequence of increase in customer expenditure following better engagement with the retailer (Digital Commerce 360, 2016-2017). On the whole, the net benefit from channel-based price discrimination decreases as omnichannelness increases. Therefore, I hypothesize that online-offline dispersion decreases as omnichannleness of retailers increases. Figure 1 summarizes the theoretical framework discussed above.

Figure 1: Theoretical Framework for Effect of Omnichannelness on Price Dispesion



4 Data

I combine two datasets to generate a unique dataset for analyzing the relationship between omnichannelness and online-offline price difference. The first dataset contains online and offline prices of randomly sampled products collected through MIT's Billion Prices Project (BPP) (Cavallo and Rigobon, 2016). The dataset is an unbalanced panel at the product-level containing prices collected between December 2015 and March 2016 for 17 US retailers. I merge the BPP price data with retailer-year level omnichannel data reported in Digital Commerce 360's omnichannel reports for 2015 and 2016.[3] The omnichannel reports contain data on 12 of the 17 retailers in the BPP dataset. Thus, the final merged dataset contains price observations between 2015 and 2016 for 12 US retailers at the product-level. I describe the individual datasets and the merged dataset in detail in subsections below.

4.1 BPP Price Data

The BPP price data publicly available through Harvard/MIT Dataverse contains 45,253 observations, 20,193 of which are for US retailers. The BPP team collected online and offline prices simultaneously from the 56 largest retailers in 10 counties: Argentina, Australia, Brazil, Canada, China, Germany, Japan, South Africa, United Kingdom, and the US. In this paper, I limit my analysis to US retailers because Digital Commerce 360's omnichannel reports only contain metrics on US retailers (see Section 3.2). The dataset also contains barcode of products (product id), date of offline price collection, location of physical store from which offline price was collected, an indicator for whether online and offline prices were collected on the same day, a binary for whether there was a sale online, and two other variables containing comments about whether there was a sale offline.

As described in detail in Cavallo (2017), the retailers for BPP price data were selected based on their market share and whether the product barcode could be matched across samples. Cavallo's team employed a combination of crowd sourcing platforms (Amazon Mechanical Turk, Elance and UpWork), a customized mobile phone app, and web scraping methods to collect prices. The workers hired through the crowd sourcing platforms were instructed to collect prices for randomly selected 10-50 products in physical stores using a special application developed to simplify and standardize the data collection process — workers scanned product barcodes, manually entered the price, took a photo of the price tag, and sent all the information via e-mail to BPP servers. The server automatically processed the data, and then a scraping software used the bar code find the same product on retailer's

[3]https://https://www.digitalcommerce360.com/product/omnichannel-report/

website and collected the online price within a period of seven days. Table 1 presents the retailers in BPP price data.

Table 1: Retailers in Billion Prices Project Data

Country	Retailer
Argentina	Carrefour, Coto, Easy, Sodimac, Walmart
Australia	Coles, Masters, Target, Woolworths
Brazil	Droga Raia, Extra, Magazine Luiza, Pao de Azucar, Renner
Canada	Canadian Tire, Home Depot, The Source, Toys R Us, Walmart
China	Auchan Drive, Sams Club
Germany	Galeria Kaufhof, Obi, Real, Rewe, Saturn
Japan	Bic Camera, K's Denki, Lawson, Yamada
South Africa	Clicks, Dis-Chem Pharmacy, Mr. Price, Pick n Pay, Woolworths
United Kingdom	Asda, Marks and Spencer, Sainsburys, Tesco
United States	Walmart, Target, Safeway, Stop&Shop, Best Buy, Home Depot, Lowe's, CVS, Macy's, Banana Republic, Forever 21, GAP, Nike, Urban Outfitters, Old Navy, Staples, Office Depot

Source: Cavallo (2017)

4.1.1 Summary Statistics for Online-Offline Price Difference

There are several ways of measuring the online-offline price difference. The simplest would be a binary measure of price difference which would equal 1 if the online and offline prices are different and 0 otherwise. This metric hides the information on the magnitude as well the direction of the price difference. A next reasonable measure would be to simply take the difference between online and offline price. Using this measure of price difference, I exclude 18 observations in BPP price data with unreasonable online-offline price difference. For 17 of these observations, the product ID was equal to offline price, which points to an error in data collection or data entry. The remaining single observation also had an unreasonable price difference though its ID did not equal offline price. These observations are provided in Appendix A1.

The difference between online and offline price, however, is not ideal for studying the association between omnichannelness and price difference. This is because products with larger prices are more likely to have large price differences. Therefore, I use relative price difference — the ratio of nominal online-offline price difference to the online price — which normalizes this difference between expensive and cheaper products. A challenge associated with using relative price difference is that it is sensitive to the base. For instance, a product with online price of 5 and offline price of 50 will have a relative price difference of 0.9 with offline price as the base, whereas the relative price difference would be 9 with online price as base. To account for this sensitivity, I remove 60 observations with excessive relative price difference, which I define as having relative price difference higher than 5 with either online or offline price as base. These observations are provided in Appendix A2. There are 20,115 remaining observations in the BPP price dataset. For the rest of the analysis, I use online price as the base for the relative price difference as shown below. In the equation below, p indexes product and t indexes time.

$$Relative\ Price\ Diff_{p,t} = \frac{Online\ Price_{p,t} - Offline\ Price_{p,t}}{Online\ Price_{p,t}}$$

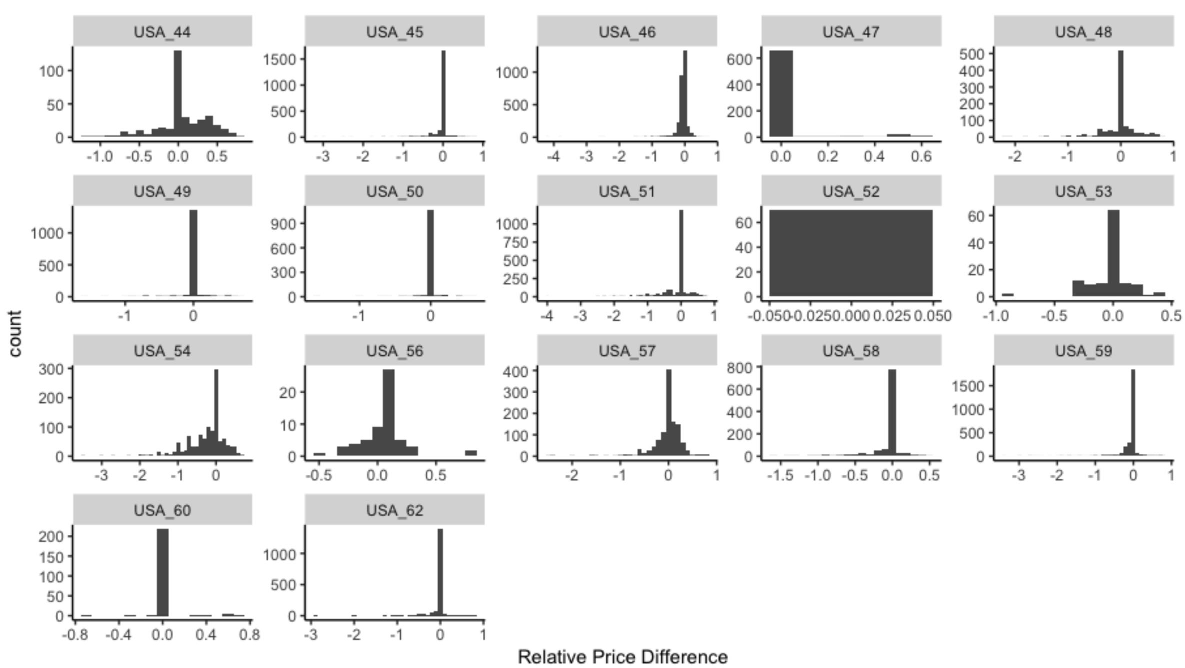
Table 2 presents the summary statistics for binary price difference and relative price difference. I keep the names of the retailers anonymous like in Cavallo (2017). The pricing strategy of individual retailers in the dataset is not the primary interest of this paper. The second column contains the number of observations for each retailer, which ranges from 59 for USA56 to 2962 for USA46. At the retailer level, the number of observations in the BPP price data is less than ideal. The third column contains the number of unique products in the sample for each retailer. The count of unique products is fewer than the count of total observations for all retailers except USA52. This confirms that prices of some of the products were collected more than once for all retailers except USA52. For USA52, prices for all the products were collected just once.

Table 2: Summary Statistics for Online-Offline Price Difference for US Retailers

Retailer	Obs Count	Unique Product Count	% Same Price	% Higher Online	% Higher Offline	Min Rel Price Diff	Max Rel Price Diff	Mean Rel Price Diff	Median Rel Price Diff	Std Rel Price Diff
USA_44	355	247	33.80	42.54	23.66	-1.18	0.77	0.05	0.00	0.33
USA_45	2179	1256	72.65	8.12	19.23	-3.17	0.81	-0.04	0.00	0.21
USA_46	2962	1647	23.84	12.05	64.11	-4.17	0.70	-0.07	-0.04	0.20
USA_47	692	588	94.65	5.35	0.00	0.00	0.59	0.02	0.00	0.11
USA_48	972	435	51.03	25.82	23.15	-2.22	0.76	-0.02	0.00	0.32
USA_49	1515	715	87.79	6.07	6.14	-1.61	0.68	-0.00	0.00	0.12
USA_50	1130	401	92.30	3.81	3.89	-1.60	0.55	-0.00	0.00	0.08
USA_51	2062	1287	55.53	15.71	28.76	-4.00	0.78	-0.11	0.00	0.40
USA_52	70	70	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
USA_53	120	46	35.00	21.67	43.33	-0.95	0.40	-0.04	0.00	0.19
USA_54	1138	461	22.67	21.62	55.71	-3.49	0.74	-0.24	-0.11	0.50
USA_56	59	57	6.78	66.10	27.12	-0.45	0.81	0.06	0.09	0.21
USA_57	1172	543	24.15	40.87	34.98	-2.50	0.81	-0.02	0.00	0.30
USA_58	1052	272	69.96	7.41	22.62	-1.55	0.46	-0.05	0.00	0.19
USA_59	2700	1551	64.48	7.85	27.67	-3.37	0.80	-0.06	0.00	0.26
USA_60	228	112	95.61	3.07	1.32	-0.70	0.67	0.01	0.00	0.11
USA_62	1709	726	76.36	7.96	15.68	-2.87	0.83	-0.02	0.00	0.18

The next three columns provide summary statistics for binary measure of price difference.59% of the observations for US retailers have the same online and offline price at the aggregate level. As seen in the table, however, there is substantial heterogeneity at the retailer level: The percentage of observations with same price across channels ranges from a low of 6.7% for USA56 to a maximum of 100% for USA52. Likewise, only 6 retailers (USA44, USA47, USA48, USA56, USA57, USA60) have more observations with higher online price than observations with lower online price. This suggests that on average it is more common to see higher prices offline than online given there exists a difference. The last five columns provide the summary statistics for relative price difference defined earlier. The mean relative price difference is negative for all but 4 retailers (USA44, USA47, USA56, USA60). This would imply that prices are higher offline than online on average for the retailers with the negative mean. Interestingly, all 4 retailers with positive mean relative price difference have more observations with higher online price than observations with lower online price. The mean relative price difference is exactly 0 only for USA52. In fact for USA52, the online and offline prices are the same for all the observations. The maximum relative price difference for the US retailers never exceeds 1; however, the minimum goes up to -4.17. Lastly, USA44, USA51 and USA54 have relatively high standard deviation in relative price difference. This is also reflected in the figure below. Figure 2 plots histogram of relative price difference for each retailer in the BPP price data. The retailers have a tall bar around 0 suggesting a median relative price difference of approximately zero. USA52 has no variation in the price difference, while USA44 and USA54 (subsidiaries of the same firm) have a large spread.

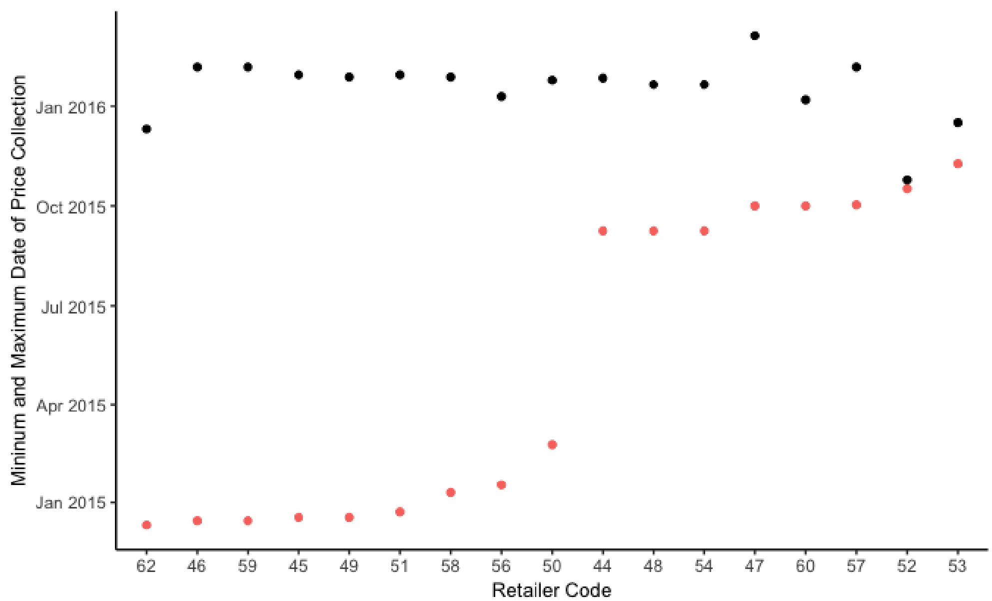
Figure 2: Histogram of Relative Price Difference for US Retailers



4.1.2 Period and Frequency of Price Collection

The online and offline prices for US retailers were collected by the BPP team between December 2015 and March 2016. Though this was the overall period of price collection for US retailers, the prices were sampled from the US retailers in varying time intervals. Figure 3 shows the minimum and maximum date of price collection. Two groups of retailers become apparent in the figure. For 9 retailers on the left side of the plot (up to retailer 50), the period of price collection spanned most of 2015 and parts of 2016 (except for retailer 62). The prices were collected between 2015Q4 and 2016Q1 for the remaining 8 retailers on the right (retailer 44 onwards). The period of price collection was especially short for retailer 62.

Figure 3: Period of Price Collection for US Retailers in BPP Price Data



Note: The red and black points indicate the beginning and the end of the price collection period respectively.

It is also interesting to consider the frequency of price collection of products. We inferred from Table 2 that prices were collected more than once for some of the products for all retailers except USA52. As mentioned earlier, the price data is an unbalanced panel i.e. the BPP team did not collect prices on same dates for all the products. The frequency of overtime price observations of products ranges from single to over five. Table 3 shows the number of unique products for which prices were collected once, twice, thrice, four times, five times and over five times. Except for USA53, over 50% of the products just have single observation in the BPP price dataset. Because of substantial differences in the period of price collection as well as the frequency of price collection, it is hard to exploit overtime variation in prices to study the effect of omnnichannelness on price difference.

Table 3: Count of Unique Products with Overtime Prices

Retailer	One Obs	Two Obs	Three Obs	Four Obs	Five Obs	More Than Five Obs	% of Prods With Single Obs
USA_44	173	49	18	5	2	0	0.70
USA_45	915	128	91	27	26	72	0.73
USA_46	1346	130	22	15	25	110	0.82
USA_47	521	50	7	2	6	2	0.89
USA_48	244	87	33	22	14	35	0.56
USA_49	580	18	5	6	12	94	0.81
USA_50	314	6	1	2	3	77	0.78
USA_51	1086	90	21	11	7	72	0.84
USA_52	70	0	0	0	0	0	1.00
USA_53	8	9	22	7	0	0	0.17
USA_54	258	72	29	29	22	51	0.56
USA_56	55	2	0	0	0	0	0.96
USA_57	410	45	7	10	4	67	0.76
USA_58	153	12	12	7	5	83	0.56
USA_59	1296	86	27	20	19	107	0.83
USA_60	75	17	3	1	3	13	0.67
USA_62	567	41	18	6	5	90	0.78

Note: This table presents the number of unique products for which prices were collected just once, twice, thrice, four time, five times and move than five times.

4.1.3 Synchronized Sale

The BPP price data also contains data on presence of sale on online as well as offline channel. Presence of sale is indicated using a binary variable for each channel; it does not tell us the magnitude of the sale. Cavallo (2017) mentions that sales tend to create some discrepancies between online and offline prices. Say, for example, price of a product on both online and offline channel is the same to begin with. If the retailer only introduces sale on the online channel (perhaps because it is easier to do so), a price dispersion between the channels would emerge. Cavallo, however, argues that the impact of sales on his aggregate results is small because the number of sale observations in BPP price data is relatively small.

In Table 4, I show the percentage of observations with sale either online or offline or on both channels. From columns (3) and (4), it can be inferred that sales are more common online than offline. This is likely due to lower menu costs associated with changing prices or indicating sale on the online channel. The exceptions to this are USA47, USA50, USA52, USA56, USA58, and USA60. Among these six retailers, all but USA50 have no observations with online sale. Since sales account for some of the price dispersion across channels, they must be controlled for while studying the effect of omnichannelness on price dispersion. A reasonable control would to synchronize sales across the channels. If there was a sale on both channels or if there was no sale on either channel, then the sales are said to be synchronized. It must be emphasized that this is not a perfect control: Even with sale on both channels, prices could be different if the magnitude of sale differ. Because of limitations of data, I cannot use this control. In the last column of Table 4, I present the percentage of observations with synchronized sale. Except for retailers USA44, USA51 and USA54, most of the sales tend to be synchronized. Even though the percentage of observations with unsynchronized sale is small, controlling for it leads to more robust results.

Table 4: Percentage of Observations with Sale/Discount

Retailer	Obs Count	% Sale Online	% Sale Offline	% Sale Both Channels	% Sale Neither Channels	% Sale Only Online	% Sale Only Offline	% Sync Sale
USA_44	355	49.01	48.17	32.39	35.21	16.62	15.77	67.61
USA_45	2179	19.05	10.10	6.98	77.83	12.07	3.12	84.81
USA_46	2962	16.21	3.85	2.43	82.38	13.77	1.42	84.81
USA_47	692	0.00	5.20	0.00	94.80	0.00	5.20	94.80
USA_48	972	36.73	32.30	25.41	56.38	11.32	6.89	81.79
USA_49	1515	0.79	0.20	0.07	99.08	0.73	0.13	99.14
USA_50	1130	1.24	1.86	0.62	97.52	0.62	1.24	98.14
USA_51	2062	51.75	17.85	10.33	40.74	41.42	7.52	51.07
USA_52	70	0.00	0.00	0.00	100.00	0.00	0.00	100.00
USA_53	120	2.50	2.50	2.50	97.50	0.00	0.00	100.00
USA_54	1138	71.09	38.14	29.44	20.21	41.65	8.70	49.65
USA_56	59	0.00	1.69	0.00	98.31	0.00	1.69	98.31
USA_57	1172	8.19	7.34	3.92	88.40	4.27	3.41	92.32
USA_58	1052	0.00	3.23	0.00	96.77	0.00	3.23	96.77
USA_59	2700	4.70	4.59	1.44	92.15	3.26	3.15	93.59
USA_60	228	0.00	2.19	0.00	97.81	0.00	2.19	97.81
USA_62	1709	13.11	1.11	0.64	86.42	12.46	0.47	87.07

Note: This table presents the breakdown of the observations by sale on online and offline channels. The % refers to the percentage of observations out of total observations.

4.2 Digital Commerce 360's Omnichannel Report

Digital Commerce 360 ranks the leading 1000 e-retailers selling in the US and Canada. The firm studies website traffic data and key trends in the e-commerce industry to identify all the significant e-retail websites selling to North American consumers. Their team has been publishing annual omnichannel report for US retail since 2016. The reports describe emerging omnichannel technologies in retail and also document individual performance of top omnichannel retailers in the US.[4] This is the best dataset available on omnichannel performance of US retailers. In the 2016 and 2017 omnichannel reports, they include data on 17 omnichannel features for top 30 omnichannel players in the US. All of the 17 omnichannel variables are binary. I use the 2016 and 2017 report to get omnichannel data for 2015 and 2016 respectively. There are two limitations of using omnichannel data from these reports. First, I would be limiting the analysis to US retailers since the report only contains omnichannel data on US retailers. Second, the price data collected in December 2014 cannot be used since we do not have omnichannel data for 2014. However, this does not affect the analysis in this paper substantially since we only have 143 observations from 2014. Lastly, the omnichannel reports do not contain information for subsidiaries separately. For instance, both Banana Republic and Old Navy (subsidiaries of GAP) do not have individual observations. In such cases, I use the metrics of the parent company for all of its subsidiaries. I further discuss omnichannel variables, their division into three categories based on their relevance to online-offline price difference, and construction of omnichannel index in section 4.

4.3 Merged Dataset

I merge BPP price data with Digital Commerce 360's omnichannel data. Retailer 52 and 56 only have 70 and 59 observations in BPP price data. I drop the observations for two of these retailers since we cannot make substantial conclusions using few observations of these retailers. We also do not have omnichannel data for these two retailers. This leaves us with 15 US retailers. The omnichannel report does not contain data on the other 4 US retailers (USA46, USA47 and USA58) represented in the BPP price data. Retailer 58 is a non-US based company for which Digital Commerce 360 does not collect data. No other sources had rich omnichannel data for retailers 46 and 47; thus, I drop the observations of these two retailers as well. This leaves 15,148 matched observations for 12 retailers in the merged dataset.

[4] https://www.digitalcommerce360.com/product/omnichannel-report/

5 Omnichannel Index

In order to study the effect of omnichannelness on online-offline price difference, we need a measure of omnichannelness of retailers. While being omnichannel broadly implies integration of online and offline channels, there are several fronts on which the integration can take place. For example, both "buy online pick up in store" and "ship from store" are omnichannel features. Therefore, I create an omnichannel index using individual omnichannel features to capture a retailer's overall engagement with various omnichannel fronts. Because omnichannel revolution is a recent phenomenon in retail, new omnichannel technologies continue to emerge. This makes it difficult to pin down the exact features of a successful omnichannel retailer.

Prior studies on omnichannel retail have focused on specific omnichannel attributes like "buy online pickup in store" (Gallino and Moreno, 2014). I did not find any work that measured the overall omnichannelness of a retailer. Since I cannot base the omnichannel index on prior literature, I use detailed omnichannel data on US retailers available through Digital Commerce 360's omnichannel reports from 2016 and 2017 to create the index (Digital Commerce 360, 2016-2017).[5] The reports contain data on 17 omnichannel variables, which are presented in Table 5. The omnichannel variables in the dataset are binary, and for each variable a value of 1 implies being more omnichannel than a value of 0. For instance, a retailer offering "Barcode on Products" (value of 1) is more omnichannel than another retailer which does not offer this feature (value of 0) with all the other omnichannel features being the same.

Creating an index to capture overall omnichannel behavior using these individual features is not obvious. Omnichannel variables like "Online Price Matching" seem to be more relevant for determining online-offline price difference than, say, "Directions". Ideally, assigning weights to individual omnichannel variables based on their relevance to online-offline price difference would make the index more robust. However, there is no objective way of assigning the weights given the limitations of the data. Thus, I create the omnichannel index by taking an equally weighted average of all then 17 omnichannel variables. Since all the variables are binary, the index ranges between 0 and 1. An omnichannel index value of 1 would imply the highest level of omnichannelness. Though this is not an ideal measure of omnichannelness, it still captures the overall omnichannel performance of the retailers.

[5] The 2016 and 2017 reports were created using data collected in 2015 and 2016 respectively.

Table 5: Omnichannel Variables in Digital Commerce 360's Omnichannel Report

Retailer	Year	44	45	48	49	50	51	53	54	57	59	60	62
Barcode on Products	2015	1	1	1	1	1	1	1	1	1	1	1	1
	2016	0	1	0	1	1	1	1	0	1	1	1	1
Beacons	2015	0	1	0	1	0	1	0	0	0	1	1	1
	2016	0	1	0	1	1	1	0	0	0	1	1	1
Buy Online Pick Up in Store	2015	0	1	0	1	1	1	1	0	1	1	1	1
	2016	1	1	1	1	1	1	1	1	1	1	1	1
Directions	2015	1	1	1	1	1	1	1	1	1	1	1	1
	2016	1	1	1	1	1	0	1	1	1	1	1	1
Geofencing	2015	0	1	0	1	1	1	1	0	0	1	1	1
	2016	0	1	0	0	0	1	1	0	0	0	1	1
In-store Kiosks	2015	0	0	0	1	0	1	0	0	1	0	0	0
	2016	0	0	0	0	1	0	0	0	0	1	0	1
In-store Mobile Promotions	2015	0	0	0	0	0	0	0	0	0	1	1	0
	2016	0	0	0	0	0	0	0	0	0	1	0	0
In-store Stock Counts	2015	0	1	0	1	1	1	1	0	1	1	1	0
	2016	0	1	0	1	1	1	1	0	1	1	0	1
Location-Based Offers	2015	0	0	0	1	0	0	0	0	1	1	0	0
	2016	1	0	1	0	0	1	0	1	0	1	1	1
Mobile Devices for Store Associates	2015	0	1	0	0	1	1	0	0	0	1	1	0
	2016	1	1	1	0	0	1	0	1	0	1	0	0
Online Price Matching	2015	0	1	0	0	0	1	1	0	0	1	0	0
	2016	0	1	0	1	0	0	1	0	1	0	0	1
Registry on App	2015	0	1	0	0	0	1	0	0	0	1	0	1
	2016	0	0	0	0	0	1	0	0	0	0	0	1
Reserve Online Pick Up in Store	2015	1	1	1	0	0	0	1	1	0	0	0	0
	2016	1	1	1	0	0	0	1	1	0	0	0	1
Return to Store	2015	1	1	1	1	1	1	1	1	1	1	1	1
	2016	1	1	1	1	1	1	1	1	1	1	1	1
Ship from Store	2015	0	1	0	1	1	1	0	0	1	1	1	1
	2016	1	1	1	1	0	1	1	1	1	1	0	1
Store Locator	2015	1	1	1	1	1	1	1	1	1	1	1	1
	2016	1	1	1	1	1	1	1	1	1	1	1	1
Wayfinding	2015	0	0	0	1	0	0	0	0	0	1	0	0
	2016	0	0	0	1	1	1	0	0	0	1	0	1

Source: Digital Commerce 360's Omnichannel Report 2016 and 2017.

Note: This table contains data on 17 binary omnichannel variables for 12 US retailers.

5.1 Three Categories of Omnichannel Variables

As mentioned earlier, some omnichannel variables are more likely to influence price dispersion than others. Splitting the omnichannel variables into categories based on their relevance to price dispersion would allow us to better understand the mechanism by which omnichannelness affects price dispersion. I divide the 17 omnichannel variables into three categories and create an index (just like the omnichannel index) for each category by taking the average of the variables in the category. The three categories are: (*i*) Integration features, (*ii*) In-Store Features, and (*iii*) Geographical Features. The variables included in each of the three categories are shown in Table 6.

Integration features include "fundamental" omnichannel variables that help retailers integrate their channels. These variables are "fundamental" because both the online and offline component of shopping experience are equally relevant when using these features. In-store features also allow retailers to integrate the online and offline channels, but they are store-based technologies that primarily affect customers visiting the physical store. Lastly, geographical features contains variables that are centered around the physical location of the customers. They too help integrate the channels, but to a lesser degree than both integration features and in-store features.

Table 6: Three Categories of Omnichannel Variables

Integration Features	In-Store Features	Geographical Features
Barcode on Products	Beacons	Directions
Buy Online Pick Up in Store	In-Store Kiosks	Geofencing
Online Price Matching	In-Store Mobile Promotions	Store Locator
Registry on App	In-Store Stock Counts	Location Based Offers
Reserve Online Pick Up in Store	Mobile Devices for Store Associates	Wayfinding
Return to Store		
Ship from Store		

Table 7 presents the values of omnichannel index, integration index, in-store index and geographical index for US retailers in 2015 and 2016. The omnichannel index varies between 0.29 and 0.88 and the integration index varies between 0.43 and 1. Likewise, the in-store index varies between 0 and 1. The geographical index has the value of either 0.4, 0.6, 0.8, or 1.0. Though it may seem that retailers would increase their omnichannel efforts in 2016 compared to 2015, the omnichannel index decreased from 2015 to 2016 for 6 of the retailers. A potential reason for this could be that some of the omnichannel strategies were not necessarily effective and were simply adding to retailers' costs. Hence the retailers might have chosen to focus on fewer omnichannel fronts.

Table 7: Indexes Constructed Using Omnichannel Variables

Retailer	Year	Omnichannel Index	Integration Index	In-store Index	Geographical Index
USA_44	2015	0.29	0.43	0.00	0.40
USA_44	2016	0.47	0.57	0.20	0.60
USA_45	2015	0.76	1.00	0.60	0.60
USA_45	2016	0.71	0.86	0.60	0.60
USA_48	2015	0.29	0.43	0.00	0.40
USA_48	2016	0.47	0.57	0.20	0.60
USA_49	2015	0.71	0.57	0.60	1.00
USA_49	2016	0.59	0.71	0.40	0.60
USA_50	2015	0.53	0.57	0.40	0.60
USA_50	2016	0.53	0.43	0.60	0.60
USA_51	2015	0.76	0.86	0.80	0.60
USA_51	2016	0.71	0.71	0.60	0.80
USA_53	2015	0.53	0.71	0.20	0.60
USA_53	2016	0.59	0.86	0.20	0.60
USA_54	2015	0.29	0.43	0.00	0.40
USA_54	2016	0.47	0.57	0.20	0.60
USA_57	2015	0.53	0.57	0.40	0.60
USA_57	2016	0.47	0.71	0.20	0.40
USA_59	2015	0.88	0.86	0.80	1.00
USA_59	2016	0.76	0.57	1.00	0.80
USA_60	2015	0.65	0.57	0.80	0.60
USA_60	2016	0.47	0.43	0.20	0.80
USA_62	2015	0.53	0.71	0.20	0.60
USA_62	2016	0.88	1.00	0.60	1.00

6 Empirical Methodology and Results

I study the association between omnichannelness and online-offline price dispersion using binary as well as continuous measures of price difference. First, I consider a binary indicator of price dispersion: The variable equals 0 if price of a product is same online and offline, otherwise it equals 1. Second, I use an absolute measure of relative price difference, which ignores the direction of price difference but captures its magnitude unlike the binary measure. To construct this measure, I take the absolute of relative price difference as defined in Section 3. Third, I consider relative price difference to examine both the magnitude and the direction of price dispersion. Next, I analyze the relationship between price dispersion and three categories of omnichannel variables outlined in Section 5. This allows us to identify the group of omnichannel variables that have stronger association with price dispersion and gain insight into the mechanism by which omnichannelness affects price dispersion. I also study the effect of two individual omnichannel variables — "Buy Online Pick Up in Store" and "Online Price Matching" — since these variables have interesting implications for price dispersion. Lastly, I study the association between omnichannelness and synchronized sale. The subsections below outline the methodology and show the results.

6.1 Binary Measure of Price Dispersion

It is helpful to study the association between omnichannelness and price dispersion using a binary measure before analyzing the same relationship with more complicated measures. Since the dependent variable is binary, I use logistic as well as ordinary least squares (OLS) regression models. I perform regression analysis at the product-level since product-level determinants like synchronized sale and imputed are important to control.

$$Binary\ Price\ Diff_{p,r,y,m} = \beta_0 + \beta_1 Omni\ Index_{r,y} + \beta_2 Sync\ Sale_{p,r,y,m} + \beta_3 Imputed_{p,r,y,m} + \gamma_s + \delta_y + \lambda_m \quad (i)$$

In specification (i) above, p , r, y, and m index product, retailer, year and month respectively. The variable of interest is *Omni Index*, which is the average of 17 omnichannel variables reported in Digital Commerce 360's omnichannel reports. *Sync Sale* is a binary variable which equals 1 if the sale (discount) was synchronized across the channels, otherwise it equals 0. As discussed in section 3, Cavallo (2017) mentions that (unsynchronized) sales tend to create some discrepancies between online and offline prices. Therefore, controlling for synchronized sale is important as some of the online-offline price dispersion may simply exist due, for example, to the presence of a sale on one channel but not on the other. Likewise, *Imputed* equals 1 if online and offline prices for a product were not collected on the same day. Because price dispersion is more likely if the prices were not collected on the same day, it is important to control for this variable. γ, δ, λ denote sector, year, and month fixed effects respectively. Sector fixed effects are important since online-offline price dispersion vary by sector. As shown in Cavallo (2017), a drug retailer is likely to have a larger online-offline price difference than a clothing retailer because customers would be willing to pay a premium for immediate access to a drug in a physical store. Month fixed effects control for seasonality trends in retail and year-fixed effects control for yearly trends. I cluster the standard errors at the product level because it is the smallest level of observation and the analysis is being performed at the product level.

Table 8: Binary Online-Offline Price Difference and Omnichannel Index

VARIABLES	(1) Binary Price Diff: Logit Coeff	(2) Binary Price Diff: Logit Coeff	(3) Binary Price Diff: OLS Coeff	(4) Binary Price Diff: OLS Coeff
Omnichannel Index	-1.506***	-0.696***	-0.298***	-0.132***
	(0.111)	(0.244)	(0.0217)	(0.0461)
Synchronized Sale	-1.985***	-1.985***	-0.442***	-0.428***
	(0.049)	(0.0881)	(0.00918)	(0.0162)
Imputed	0.465***	0.197***	0.0926***	0.0345***
	(0.039)	(0.0505)	(0.00762)	(0.00874)
Constant	1.791***	1.439***	0.876***	0.805***
	(0.091)	(0.205)	(0.0167)	(0.0368)
Observations	15,148	15,148	15,148	15,148
R-squared			0.164	0.255
Sector FE	No	Yes	No	Yes
Year FE	No	Yes	No	Yes
Month FE	No	Yes	No	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The dependent variable is binary measure of price difference which equals 1 if the price of a product differs across the channels.

The results of specification (i) are shown in Table 8. Columns (1) and (2) present the results of logistic model, while columns (3) and (4) present the results of OLS model outlined in specification (i). In all four columns, the coefficient on the omnichannel index is negative and statistically significant. This suggests a negative relationship between omnichannelness and presence of online-offline price difference on average. The magnitude of the coefficient on omnichannel index decreases in both logistic and OLS models when fixed effects are added. Model (2) says that for every unit increase in omnichannel index, the log odds of having an online-offline price difference decreases by 0.7. Likewise, model (4) can be interpreted as for every unit increase in omnichannel index, the binary price difference decreases by 0.13 on average. Besides the omnichannel index, the coefficient on synchronized sale is negative in all four columns as expected. If there is a synchronized sale, the price difference (or log odds of having a price difference in logistic model) decreases on average. This suggests that some of the price difference may be due to unsynchronized sale across the two channels. Lastly, the coefficient on imputed is positive. This implies that if the online and offline prices were not collected on the same day (imputed equals 1), the binary price difference increases on average. This make sense since prices can change the very next day and hence difference in time of price collection across channels can result in a price difference.

6.2 Continuous Measure of Price Dispersion

I consider two continuous measures of price dispersion: absolute relative price difference and relative price difference. Relative price difference is the ratio of online-offline price difference to online price of the product. Absolute relative price difference is the absolute value of the relative price difference and ignores the direction of the price difference. Its value is either zero or positive. The econometric specification with absolute relative price difference follows below. Specification (*ii*) is similar to specification (*i*) except for the dependent variable: The new dependent variable is the absolute of relative price difference.

$$\begin{aligned} Absolute\ Relative\ Price\ Diff_{p,r,y,m} = \beta_0 + \beta_1 Omni\ Index_{r,y} + \beta_2 Sync\ Sale_{p,r,y,m} + \\ \beta_3 Imputed_{p,r,y,m} + \gamma_s + \delta_y + \lambda_m \quad (ii) \end{aligned}$$

Relative price difference can take positive (online $>$ offline), zero (online $=$ offline), or negative (online $<$ offline) value. The fact that relative price difference can be positive, zero, or negative makes it harder to study its association with online-offline price difference. This is because the coefficient on the omnichannel index would vary based on the direction of price dispersion. For instance, let's assume the association between price dispersion and relative price difference is negative. Then the coefficient on the omnichannel index would have to be negative for observations with a positive price difference and positive for observations with a negative price difference. To account for this, I construct an indicator for whether the price difference was positive and interact it with the omnichannel index. Likewise, I construct an indicator for positive or zero price difference and interact it with the omnichannel index. The econometric specifications are given below. In specification (*iii*) and (*iv*), I use relative price difference as the response variable. I include $Pos\ Price\ Diff$ an indicator for whether price difference was positive and an its interaction with the omnichannel index in specification (*iii*). Specification (*iv*) is similar to specification (*iii*) except that the indicator $Pos/Zero\ Price\ Diff$ equals 1 for observations with positive or no price difference.

$$\begin{aligned} Rel\ Price\ Diff_{p,r,y,m} = \beta_0 + \beta_1 Omni\ Index_{r,y} + \beta_2 Pos\ Price\ Diff_{p,r,y,m} * Omni\ Index_{r,y} + \\ \beta_3 Pos\ Price\ Diff_{p,r,y,m} + \beta_4 Sync\ Sale_{p,r,y,m} + \beta_5 Imputed_{p,r,y,m} + \\ \gamma_s + \delta_y + \lambda_m \quad (iii) \end{aligned}$$

$$\begin{aligned} Rel\ Price\ Diff_{p,r,y,m} = \beta_0 + \beta_1 Omni\ Index_{r,y} + \beta_2 Pos/Zero\ Price\ Diff_{p,r,y,m} * Omni\ Index_{r,y} \\ \beta_3 Pos/Zero\ Price + \beta_4 Sync\ Sale_{p,r,y,m} + \beta_5 Imputed_{p,r,y,m} + \\ \gamma_s + \delta_y + \lambda_m \quad (iv) \end{aligned}$$

Table 9 presents the results of regressions with continuous measures of price dispersion. All the columns in Table 9 use OLS model with fixed effects. Column (1) presents the results of specification (*ii*). The omnichannel coefficient is negative and statistically significant, which implies a negative relationship between omnichannelness and absolute relative price difference. If the omnichannel index increases by 1, the absolute relative price difference decreases by 5% on average. The decrease in price dispersion with increase in omnichannelness is the same for both observations with positive and those with negative price dispersion. The coefficient on synchronized sale and imputed are significant and have the same direction as they did in Table 8 with binary measure of price dispersion. Their interpretation remains unchanged. The coefficient on omnichannel index in column (2) is not significant. This might suggest a lack of relationship between omnichannelness and relative price difference. The lack of significance on omnichannel coefficient, however, is likely due to the difference in the direction of omnichannel coefficient based on the direction of price difference as explained earlier.

Columns (3) and (4) show the results of specifications (*iii*) and (*iv*) respectively. The use of indicator and interaction terms in these two specifications makes it possible to separately

Table 9: Continuous Measure of Price Dispersion and Omnichannel Index

VARIABLES	(1) Abs Rel Price Diff	(2) Relative Price Diff	(3) Relative Price Diff	(4) Relative Price Diff
Omnichannel Index	-0.052**	0.007	0.067**	0.216***
	(0.025)	(0.028)	(0.028)	(0.052)
Pos Price Diff * Omni Index			-0.163***	
			(0.044)	
Positive Price Diff			0.505***	
			(0.028)	
Pos/Zero Price Diff * Omni Index				-0.329***
				(0.050)
Positive/Zero Price Diff				0.637***
				(0.032)
Synchronized Sale	-0.219***	0.160***	0.188***	0.002
	(0.011)	(0.013)	(0.011)	(0.010)
Imputed	0.020***	-0.020***	-0.021***	-0.005
	(0.005)	(0.006)	(0.005)	(0.005)
Constant	0.375***	-0.184***	-0.321***	-0.534***
	(0.018)	(0.021)	(0.020)	(0.031)
Observations	15,148	15,148	15,148	15,148
R-squared	0.175	0.070	0.283	0.410
Sector FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

examine the omnichannel coefficients for observations with positive, negative, positive/zero, and negative/zero price difference. The results of columns (3) and (4) need to be jointly considered to infer the omnichannel coefficients summarized in Table 10 below. Let's consider column (3) first. The positive omnichannel coefficient 0.067 is for observations with negative/zero price difference as the interaction term is for observations with strictly positive price difference. The positive coefficient suggests that an increase in omnichannelness is associated with a decrease in price difference (move toward zero) on average for observations with negative/zero price difference. Hence the omnichannel effect on price dispersion is negative. This negative association also holds for observations with strictly positive price difference. The coefficient on the interaction term in column (3) is negative, and its magnitude is greater than the coefficient on omnichannel index. Since the sum of omnichannel and interaction coefficient is negative, the omnichannel effect for observations with strictly positive price difference is also negative. Based on the values in column (3), the omnichannel effect for observations with positive price difference would be 0.067 - 0.163 = -0.096. The same reasoning holds for results in column (4). The omnichannel coefficient 0.216 is for observations with strictly negative price difference as the interaction is for observations with positive/zero price difference. For observations with positive/zero price difference, the omnichannel coefficient is negative (0.216 - 0.329 = -0.113).

Table 10: Omnichannel Coefficient Based on Direction of Price Difference

Direction of Price Difference	Omnichannel Coefficient
Positive	-0.096
Negative	0.216
Positive/Zero	-0.113
Negative/Zero	0.067

6.3 Three Omnichannel Categories and Price Dispersion

In this section, I consider the association between three categories of omnichannel variables and online-offline price dispersion. As explained in section 5.1, the three categories of omnichannel variables are: (*i*) Integration features, (*ii*) In-store Features, and (*iii*) Geographical Features. I construct an index for each category by taking the average of the omnichannel variables in the category. The econometric specification (*v*) outlines the model. Instead of having a single omnichannel index like in earlier specifications, the three indexes are included separately. I run the econometric models with binary and absolute relative price difference as the response variable. I do not include a model with relative price difference because it would require two interaction terms and identification of coefficients separately for the observations with positive and negative price dispersion. This would make the interpretation very complicated.

$$(Binary/Absolute\ Relative)\ Price\ Diff_{p,r,y,m} = \beta_0 + \beta_1 Integration\ Index_{r,y} + \beta_2 InStore\ Index_{r,y} + \beta_3 Geog\ Index_{r,y} + \beta_4 Sync\ Sale_{p,r,y,m} + \beta_5 Imputed_{p,r,y,m} + \gamma_s + \delta_y + \lambda_m \quad (v)$$

The results of specification (*v*) are presented in Table 11. The coefficients on all 3 indexes in all the columns are statistically significant. The coefficients on integration features index and in-store features index are negative. This implies that there is a negative association between these two groups of variables and online-offline price dispersion. Additionally, the coefficient on integration features index is bigger in magnitude than the coefficient on in-store

Table 11: Online-Offline Price Difference and Three Categories of Omnichannel Variables

VARIABLES	(1) Binary Price Diff: Logit	(2) Binary Price Diff: OLS	(3) Absolute Rel Price Diff: OLS
Integration Features Index	-2.259***	-0.349***	-0.139***
	(0.371)	(0.062)	(0.038)
In-Store Features Index	-0.657***	-0.128***	-0.036**
	(0.165)	(0.031)	(0.018)
Geographical Features Index	2.431***	0.367***	0.115***
	(0.332)	(0.048)	(0.024)
Synchronized Sale	-2.142***	-0.441***	-0.223***
	(0.090)	(0.016)	(0.011)
Imputed	0.154***	0.022**	0.015***
	(0.051)	(0.009)	(0.005)
Constant	1.739***	0.836***	0.399***
	(0.296)	(0.050)	(0.027)
Observations	15,148	15,148	15,148
R-squared		0.266	0.179
Sector FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

features index. This suggests that the association between price dispersion and integration features index is stronger than that between price dispersion and in-store features index. This is line with the earlier explanation that the integration features are the “fundamental” omnichannel variables that place equal emphasis on both online and offline channels, whereas the in-store features are relevant only for the customers visiting the store.

The coefficient on geographical features index is positive, which implies that an increase in geographical omnichannel variables is associated with an increase in online-offline price dispersion. This positive direction of association is opposite of the association between price dispersion and the two indexes. As mentioned in section 5.1, the geographical omnichannel variables are not strong omnichannel variables. In the later editions of Digital Commerce 360’s omnichannel reports, geographical omnichannel variables are not reported. It can be argued that geographical omnichannel variables like “Location Based offers” would allow retailers to price discriminate based on the location of the customers. This increase in the ease of price discrimination across channels would lead to an increase in the price dispersion.

6.4 Individual Omnichannel Variables

Though omnichannel index captures overall omnichannel behavior of a retailer, the interpretation of the index is not obvious. For instance, an increase of 0.1 in the omnichannel index can result from several combinations of changes in the individual omnichannel variables. In other words, we do not know which omnichannel features a retailer changes when its omnichannel index changes. Dividing the omnichannel variables into three categories provided additional insights on how the three groups of omnichannel variables differ in terms of their association with online-offline price dispersion. But the indexes for the three categories also suffer from the same interpretation issue as the omnichannel index. As such, it would be interesting to study the association between individual omnichannel variables and price dispersion for a meaningful interpretation. Because of strong correlations between individual omnichannel variables, a model with all the individual omnichannel variables would not provide accurate results. Among the 17 omnichannel variables in Digital Commerce 360's omnichannel reports, the variables "Buy Online Pick Up in Store" and "Online Price Matching" are interesting for price dispersion. In the subsections below, I examine the how these two variables are related to online-offline price dispersion.

6.4.1 Buy Online Pick Up in Store

Buy Online Pick Up in Store (BOPS) is one of the most popular omnichannel features. Customers who buy a product online and pick it up in store can easily check the in-store price of the product. Therefore it would not make much sense for a retailer to have different prices across channels if they offer BOPS. Because the retailer increases its reputational cost by offering BOPS, the theoretical framework suggests a negative association between online-offline price dispersion and BOPS. I test this hypothesis using all three measures of price dispersion — binary, absolute relative and relative — and present the results in Table 12. The econometric specifications are similar to those used in previous sections.

The models with binary and absolute relative price difference support our hypothesis. As seen in the first three columns, the association between BOPS and price dispersion is negative and statistically significant. The last two columns must be jointly considered. The BOPS coefficient is 0.105-0.083 = 0.022 for observations with strictly positive price difference and 0.129 for those with strictly negative price difference. The association is not negative for observations with positive price difference, which is a departure from the hypothesis. However, if we consider the observations with positive/zero price difference, the association is negative (0.129-0.185 = -0.056).

Table 12: Online-Offline Price Difference and Buy Online Pick Up in Store (BOPS)

VARIABLES	(1) Binary: Logit Coeff	(2) Binary: OLS Coeff	(3) Abs Rel Price Diff	(4) Rel Price Diff	(5) Rel Price Diff
Buy Online Pick Up in Store	-0.892***	-0.167***	-0.064***	0.105***	0.129***
	(0.120)	(0.0233)	(0.017)	(0.019)	(0.027)
Pos Price Diff * BOPS				-0.083***	
				(0.023)	
Positive Price Diff				0.481***	
				(0.021)	
Pos/Zero Price Diff * BOPS					-0.185***
					(0.025)
Positive/Zero Price Diff					0.589***
					(0.023)
Synchronized Sale	-2.016***	-0.428***	-0.219***	0.190***	-0.004
	(0.087)	(0.0160)	(0.011)	(0.011)	(0.010)
Imputed	0.131**	0.0224**	0.015***	-0.012**	-0.003
	(0.052)	(0.00895)	(0.005)	(0.005)	(0.005)
Constant	1.796***	0.862***	0.396***	-0.368***	-0.504***
	(0.172)	(0.0285)	(0.017)	(0.019)	(0.023)
Observations	15,148	15,148	15,148	15,148	15,148
R-squared		0.261	0.177	0.288	0.411
Sector FE	No	No	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
Month FE	No	No	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6.4.2 Online Price Matching

While it may seem that offering online price matching (OPM) would eliminate any difference in online and offline prices, this need not be the case. There are two reasons for this. A retailer lowers its reputational risk by providing customers OPM and places the burden of "correcting" price differentials on the customers. The retailer might leave it to customers to act against any difference in prices across the channels and hence continue to engage in channel-based price discrimination to increase profits. Second, the BPP team collected offline prices as labeled on products. The offline prices in the dataset do not reflect what the prices would have been had the customers asked the retailers to match the online price. Thus, there are two opposing forces which makes the association between price dispersion and online price matching uncertain. I test this hypothesis using all three measures of price difference — binary, absolute relative and relative. The results are presented in Table 13.

Though the coefficient on OPM is negative in the first three columns, it is insignificant at 5% level (standard significance level) in columns (1) and (3) suggesting a lack of relationship between online-offline price dispersion and OPM. Likewise, in columns (4) and (5), the coefficient on OPM is statistically insignificant even at 10% significance level. This again suggests a lack of association between price dispersion and OPM for observations with negative and negative/zero price difference. The OPM coefficient is statistically significant and negative for observations with positive (-0.081-0 = -0.081) as well as positive/zero (-0.074 - 0 = -0.074) price difference, which is a negative association. All in all, some of the models show a negative relationship between OPM and price difference, while others suggest no relationship. Therefore the association between the two is not obvious as hypothesized.

Table 13: Online-Offline Price Difference and Online Price Matching

VARIABLES	(1) Binary: Logit Coeff	(2) Binary: OLS Coeff	(3) Abs Rel Price Diff	(4) Rel Price Diff	(5) Rel Price Diff
Online Price Matching (OPM)	-0.174*	-0.0379**	-0.016*	0.005	0.031
	(0.092)	(0.0171)	(0.009)	(0.009)	(0.021)
Pos Price Diff * OPM				-0.081***	
				(0.016)	
Positive Price Diff				0.439***	
				(0.013)	
Pos/Zero Price Diff * OPM					-0.074***
					(0.021)
Positive/Zero Price Diff					0.466***
					(0.014)
Synchronized Sale	-1.985***	-0.428***	-0.219***	0.189***	0.007
	(0.088)	(0.0163)	(0.011)	(0.011)	(0.010)
Imputed	0.233***	0.0393***	0.021***	-0.024***	-0.009*
	(0.052)	(0.00908)	(0.006)	(0.006)	(0.005)
Constant	1.117***	0.745***	0.352***	-0.284***	-0.420***
	(0.150)	(0.0256)	(0.014)	(0.015)	(0.016)
Observations	15,148	15,148	15,148	15,148	15,148
R-squared		0.255	0.175	0.283	0.406
Sector FE	No	No	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
Month FE	No	No	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6.5 Synchronized Sale and Omnichannelness

As discussed in section 3, an unsynchronized sale across channels can result in price dispersion. Unsynchronized sale is a form of channel-based price discrimination where sale is used as an instrument for changing prices. The results of the earlier models show that price dispersion decreases as omnichannelness increases. As such, retailers are also more likely to have synchronized sale across channels as they become more omnichannel. Synchronized sale is a property, at least to some degree, of an omnichannel retailer. In this section, I examine if there exists a positive association between omnichannelness and synchronization of sales. Specification (*vi*) outlines the econometric model.

$$Synchronized\ Sale_{p,r,y,m} = \beta_0 + \beta_1 Omni\ Index_{r,y} + \beta_2 Imputed_{p,r,y,m} + \delta_y + \lambda_m \quad (vi)$$

Ideally, a synchronized sale would require synchronization of both the presence as well as the magnitude of the sale, but the BPP price data only contains information on whether or not a sale was present. Therefore, the dependent variable *Synchronized Sale* in the specification above is binary. The variable equals 1 if there is a sale on both the channels or on neither channel. It equals 0 if a sale is present on one channel, but not the other. Just like for binary measure of price difference, I run logistic as well as OLS regressions. The results are presented in Table 14.

Table 14: Synchronized Sale and Omnichannel Index

VARIABLES	(1) Logit Coeff	(2) Logit Coeff	(3) OLS Coeff	(4) OLS Coeff
Omnichannel Index	1.103***	0.734***	0.169***	0.121***
	(0.129)	(0.188)	(0.0199)	(0.0302)
Imputed	-0.415***	-0.280***	-0.0586***	-0.0404***
	(0.046)	(0.060)	(0.00632)	(0.00815)
Constant	1.078***	1.025***	0.745***	0.737***
	(0.092)	(0.165)	(0.0144)	(0.0260)
Observations	15,148	15,148	15,148	15,148
R-squared			0.015	0.044
Sector FE	No	No	No	No
Year FE	No	Yes	No	Yes
Month FE	No	Yes	No	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The first two columns present the results of logistic model, while the last two columns present the results of OLS model. For both OLS and logistic models, I start with no fixed effects and then add time (year as well as month) fixed effects. The time fixed effects are important here because sales vary by the time of the year; they are more likely, for example, during holidays. However, sales do not really vary across sectors due to which I do not include sector fixed effects. In all four columns, the omnichannel coefficient is positive and statistically significant. The results suggest that an increase in omnichannelness is associated with an increase in likelihood of having a synchronized sale across channels on average. This is reassuring as it implies that omnichannel retailers make effort to synchronize sales across the channels and thus reduce price dispersion. The coefficient on imputed in all the columns is negative and statistically significant. This is because if the online and offline prices (and hence the information about sale) were not collected on the same day, we are less likely to see a synchronized sale on average. This is again consistent since presence of sales can change over time. All in all, the results in Table 14 suggest that omnichannel retailers are more likely to have a synchronized sale across channels.

7 Conclusion and Discussion

In the age of omnichannel retailing, retailers face an important question about their pricing strategy: Do they charge the same price across channels for an identical product? Though the question about price-dispersion across channels has been studied before, it was done in the context of multichannel retail (Wolk and Ebling, 2010). The optimal pricing strategy in omnichannel retail, where online and offline channels are integrated into one, is further complicated. This paper studies the relationship between omnichannelness and online-offline price dispersion.

I extend the research by Cavallo (2017) on large-scale comparison of online-offline prices by considering omnichannelness as a factor that explains the price dispersion. I construct a measure of omnichannelness — omnichannel index — from detailed omnichannel data in Digital Commerce 360's omnichannel reports. Using the omnichannel index, I examine the association between omnichannelness of retailers and online-offline price dispersion. I find a statistically significant and negative relationship between omnichannelness of retailers and the online-offline price dispersion of their products. This finding is robust across three different measures of price dispersion — binary indicator for price difference, relative price difference, and absolute relative price difference. To my knowledge, this is the first study about the relationship between omnichannelness and price dispersion.

I also divide the omnichannel variables into three categories based on their level of relevance to online-offline price dispersion. The integration features and in-store features have a negative association with price dispersion, while the geographical features have a positive association. The geographical features, unlike the other two categories, do not substantially affect the integration of two channels and hence they do not lead to a decrease in price dispersion. I also consider two individual omnichannel variables: "Buy Online Pick Up in Store" and "Online Price Matching". The results show that "Buy Online Pick Up in Store" is negatively associated with price dispersion, while "Online Price Matching" does not have a consistent relationship with price dispersion. "Buy Online Pick Up in Store" lowers the price dispersion because it increases the reputational cost as customers can compare prices in the store before picking up the online order. "Online Price Matching", on the other hand, decreases the reputational risk for the retailers and places the burden of "correcting" price dispersion on customers. Finally, I find that omnichannel retailers are more likely to have a synchronized sale across their channels.

There are several limitations of this study. The number of product-level observations for retailers in BPP price data is relatively small. In addition to more observations per retailer, price data consistent over time would allow researchers to exploit variation over time. The retailers in this dataset are some of the biggest retailers in the US. Smaller omnichannel retailers might pursue different pricing strategies due to the lack of market power. Future studies could also improve the measurement of omnichannelness. Since there is no prior literature on measurement of omnichannelness, I had to rely on omnichannel data from Digital Commerce 360's omnichannel reports to construct the index. The current functional form of the index, where I take an equally weighted average of the omnichannel variables, could be improved by weighting the variables based on their relevance to price dispersion. This can be done with the help of a rigorous study of what it means to be omnichannel and how to accurately measure the overall omnichannel performance of retailers.

References

Ancarani, Fabio, and Venkatesh Shankar. 2004. “Price levels and price dispersion within and across multiple retailer types: Further evidence and extension.” *Journal of the academy of marketing Science*, 32(2): 176.

Baye, Michael R, John Morgan, Patrick Scholten, et al. 2006. “Information, search, and price dispersion.” *Handbook on economics and information systems*, 1: 323–375.

Bell, David R, Santiago Gallino, and Antonio Moreno. 2014. “How to win in an omnichannel world.” *MIT Sloan Management Review*, 56(1): 45.

Brynjolfsson, Erik, and Michael D Smith. 2000. “Frictionless commerce? A comparison of Internet and conventional retailers.” *Management science*, 46(4): 563–585.

Brynjolfsson, Erik, Yu Jeffrey Hu, and Mohammad S Rahman. 2013. “Competing in the age of omnichannel retailing.”

Cavallo, Alberto. 2017. “Are online and offline prices similar? evidence from large multi-channel retailers.” *American Economic Review*, 107(1): 283–303.

Cavallo, Alberto, and Roberto Rigobon. 2016. “The billion prices project: Using online prices for measurement and research.” *Journal of Economic Perspectives*, 30(2): 151–78.

Clay, Karen, Ramayya Krishnan, Eric Wolff, and Danny Fernandes. 2002. “Retail strategies on the web: Price and non–price competition in the online book industry.” *The Journal of Industrial Economics*, 50(3): 351–367.

Digital Commerce 360. 2016-2017. “Digital Commerce 360’s Omnichannel Reports.” *http://www.digitalcommerce360.com/topic/omnichannel/*.

Gallino, Santiago, and Antonio Moreno. 2014. “Integration of online and offline channels in retail: The impact of sharing reliable inventory availability information.” *Management Science*, 60(6): 1434–1451.

Gao, Fei, and Xuanming Su. 2016. “Online and offline information for omnichannel retailing.” *Manufacturing & Service Operations Management*, 19(1): 84–98.

Grewal, Dhruv, Ramkumar Janakiraman, Kirthi Kalyanam, PK Kannan, Brian Ratchford, Reo Song, and Stephen Tolerico. 2010. “Strategic online and offline retail pricing: a review and research agenda.” *Journal of Interactive Marketing*, 24(2): 138–154.

Harsha, Pavithra, Shivaram Subramanian, and Joline Uichanco. 2016. "Omni-channel revenue management through integrated pricing and fulfillment planning." *Ross School of Business, University of Michigan.*

Piotrowicz, Wojciech, and Richard Cuthbertson. 2014. "Introduction to the special issue information technology in retail: Toward omnichannel retailing." *International Journal of Electronic Commerce*, 18(4): 5–16.

Stamatopoulos, Ioannis, Achal Bassamboo, and Antonio Moreno. 2017. "The Effects of Menu Costs on Retail Performance: Evidence from Adoption of the Electronic Shelf Label Technology."

Vogel, Julia, and Michael Paul. 2015. "One firm, one product, two prices: Channel-based price differentiation and customer retention." *Journal of Retailing and Consumer Services*, 27: 126–139.

Wolk, Agnieszka, and Christine Ebling. 2010. "Multi-channel price differentiation: An empirical investigation of existence and causes." *International Journal of Research in Marketing*, 27(2): 142–150.

Appendix A1

Table 15: Observations With Unreasonable Price Difference in BPP Price Data

	Retailer	ID	Offline Price	Online Price	Price Diff	Rel Price Diff (Online Base)	Rel Price Diff (Offline Base)
1	USA_46	974540	974540.00	9.49	-974530.51	102690.25	1.00
2	USA_48	112636	112636.00	38.99	-112597.01	2887.84	1.00
3	USA_48	721303	721303.00	29.95	-721273.05	24082.57	1.00
4	USA_50	146514	146514.00	19.98	-146494.02	7332.03	1.00
5	USA_50	475494	475494.00	10.98	-475483.02	43304.46	1.00
6	USA_50	116480	116480.00	34.98	-116445.02	3328.90	1.00
7	USA_50	92074	92074.00	10.98	-92063.02	8384.61	1.00
8	USA_50	16925	16925.00	24.98	-16900.02	676.54	1.00
9	USA_50	593660	593660.00	24.98	-593635.02	23764.41	1.00
10	USA_54	714473	714473.00	9.00	-714464.00	79384.89	1.00
11	USA_54	597769	597769.00	34.94	-597734.06	17107.44	1.00
12	USA_54	814026	814026.00	59.94	-813966.06	13579.68	1.00
13	USA_54	716792	716792.00	45.00	-716747.00	15927.71	1.00
14	USA_54	712642	712642.00	32.00	-712610.00	22269.06	1.00
15	USA_54	753005	753005.00	69.94	-752935.06	10765.44	1.00
16	USA_59	490240806874	24100175.00	6.99	-24100168.01	3447806.58	1.00
17	USA_59	060021128	60021128.00	7.99	-60021120.01	7512030.04	1.00
18	USA_62	553490984	553490984.00	2.96	-553490981.04	186990196.30	1.00

Note: For 17 of these observations, the product ID equals offline price, which points to an error in data collection or data entry. The 16th observation also has unreasonable price difference though its ID does not equal offline price.

Appendix A2

Table 16: Observations with Large Relative Price Difference

	Retailer	Price Offline	Price Online	Price Diff	Rel Price Diff Online Base	Rel Price Diff Offline Base
1	USA_44	28.00	228.00	200.00	0.88	7.14
2	USA_45	116.99	8.43	-108.56	-12.88	-0.93
3	USA_45	44.99	359.99	315.00	0.88	7.00
4	USA_45	24.99	2.99	-22.00	-7.36	-0.88
5	USA_45	399.99	19.86	-380.13	-19.14	-0.95
6	USA_45	399.99	19.86	-380.13	-19.14	-0.95
7	USA_46	79.00	0.77	-78.23	-101.60	-0.99
8	USA_46	1.87	87.99	86.12	0.98	46.05
9	USA_49	1.97	89.97	88.00	0.98	44.67
10	USA_49	36.97	299.00	262.03	0.88	7.09
11	USA_49	4.28	40.98	36.70	0.90	8.57
12	USA_49	4.28	40.98	36.70	0.90	8.57
13	USA_49	5.99	167.00	161.01	0.96	26.88
14	USA_49	15.98	164.99	149.01	0.90	9.32
15	USA_49	7.97	423.24	415.27	0.98	52.10
16	USA_50	49.98	499.00	449.02	0.90	8.98
17	USA_50	19.98	3.18	-16.80	-5.28	-0.84
18	USA_50	1.88	198.79	196.91	0.99	104.74
19	USA_50	4.98	378.00	373.02	0.99	74.90
20	USA_50	9.98	87.03	77.05	0.89	7.72
21	USA_50	7.98	0.40	-7.58	-18.95	-0.95
22	USA_57	1.50	36.49	34.99	0.96	23.33
23	USA_57	1.50	36.49	34.99	0.96	23.33
24	USA_57	1.50	36.49	34.99	0.96	23.33
25	USA_57	1.50	36.49	34.99	0.96	23.33
26	USA_57	1.53	36.49	34.96	0.96	22.85
27	USA_57	1.49	36.49	35.00	0.96	23.49
28	USA_57	1.49	36.49	35.00	0.96	23.49
29	USA_57	1.49	36.49	35.00	0.96	23.49
30	USA_57	1.49	36.49	35.00	0.96	23.49
31	USA_59	3.00	39.99	36.99	0.92	12.33
32	USA_59	6.49	119.99	113.50	0.95	17.49
33	USA_59	6.99	139.99	133.00	0.95	19.03
34	USA_59	5.99	45.99	40.00	0.87	6.68
35	USA_59	17.79	2.49	-15.30	-6.14	-0.86
36	USA_59	54.99	4.99	-50.00	-10.02	-0.91
37	USA_59	54.99	4.99	-50.00	-10.02	-0.91
38	USA_59	15.99	2.49	-13.50	-5.42	-0.84
39	USA_59	56.99	3.99	-53.00	-13.28	-0.93
40	USA_59	99.99	11.99	-88.00	-7.34	-0.88
41	USA_62	4.98	42.32	37.34	0.88	7.50
42	USA_62	3.98	88.05	84.07	0.95	21.12
43	USA_62	2.48	0.03	-2.45	-81.67	-0.99
44	USA_62	0.75	12.45	11.70	0.94	15.60
45	USA_62	3.00	35.76	32.76	0.92	10.92
46	USA_62	3.00	70.81	67.81	0.96	22.60
47	USA_62	5.58	34.90	29.32	0.84	5.25
48	USA_62	0.98	27.76	26.78	0.96	27.33
49	USA_62	3.78	45.88	42.10	0.92	11.14
50	USA_62	3.78	49.06	45.28	0.92	11.98
51	USA_62	5.98	68.21	62.23	0.91	10.41
52	USA_62	3.64	95.94	92.30	0.96	25.36
53	USA_62	1.47	21.88	20.41	0.93	13.88
54	USA_62	1.47	13.91	12.44	0.89	8.46
55	USA_62	2.97	105.59	102.62	0.97	34.55
56	USA_62	1.00	7.47	6.47	0.87	6.47
57	USA_62	2.84	251.99	249.15	0.99	87.73
58	USA_62	1.98	60.21	58.23	0.97	29.41
59	USA_62	2.97	33.99	31.02	0.91	10.44
60	USA_62	1.77	11.99	10.22	0.85	5.77

Note: These observations have absolute relative price difference greater than 5.