Channel Choice and Customer Value

Shirsho Biswas, University of Washington*
Hema Yoganarasimhan, University of Washington
Haonan Zhang, University of Washington

July 12, 2024

Abstract

This paper investigates the impact of omnichannel adoption on consumer spending behavior, focusing on two groups of consumers – offline-only consumers who became omnichannel organically (organic switchers) and offline-only consumers who became omnichannel due to the COVID-19 shock (covid switchers). Leveraging comprehensive data from a leading pet retailer in Brazil over 2019-2023 and a Differences-In-Differences analysis, we show that the two groups of consumers increase their spending by the same amount after becoming omnichannel. That is, conditional on switching to omnichannel, the reasons for switching have no significant impact on consumer revenue. From a managerial perspective, our findings are valuable since they suggest that retailers can use the Average Treatment Effect on Treated (ATT) of users who voluntarily become omnichannel (organic switchers) as a measure of the potential gains from converting other users to omnichannel. However, differences emerge in channel preferences, with externally induced switchers displaying a slower uptake of the online channel. We find that these differences are mostly driven by older covid switchers. Since the offline channel is more profitable in our setting, our results suggest that nudging consumers to become omnichannel can increase the profitability of these consumers even beyond that of customers who voluntarily become omnichannel.

Keywords: Retailing, COVID-19, Omnichannel, Channel choice, Customer value, Natural Experiment

*Authors contributed equally and names are listed in alphabetical order. E-mails: shirsho@uw.edu, hemay@uw.edu, hzhang96@uw.edu. We thank an anonymous pet supplies retailer for providing the data and for extensive thoughtful discussions on many aspects of the paper. We would like to thank Simha Mummanaleni for his detailed comments that have significantly improved the paper. We also thank the participants of the University of Washington Faculty Seminar Series and participants of the 2024 ISMS Marketing Science Conference for their helpful inputs. The opinions expressed in this paper are the authors' and do not reflect those of the data providers.

1 Introduction

Online shopping has been growing in popularity over the last two decades with the growing proliferation of computers and smartphones. In response, many retail stores like Walmart, Target, Costco, etc., which earlier only operated physical retail stores, now sell their products through both the physical stores and e-commerce channels as part of a digital transformation in retail. Given the size of the retail industry, and the growth of digital channels, quantifying the effect of using multiple channels on the revenue generated by a customer is important for managers and firms. As such, a large stream of research has tried to document the value of omnichannel[1] customers to retailers (Neslin and Shankar, 2009).

Broadly speaking, existing literature has documented two types of findings in this space. First, an early stream of literature provides *cross-sectional* descriptive findings (Kumar and Venkatesan, 2005; Ansari et al., 2008; Kushwaha and Shankar, 2013). These papers compare omnichannel customers to single-channel customers and find that the average multichannel customer buys more and is more valuable than the average single-channel customer. There are several potential reasons for why omnichannel consumers are found to be more valuable, and this early stream of literature based on cross-sectional comparisons does not address the issue of causality. For instance, Neslin and Shankar (2009) notes that, "the main question we need to resolve is whether the association between multichannel and purchase volume is causal. It may be that multichannel usage encourages customers to buy more due to convenience, or heavy volume customers naturally utilize more channels, or some third factor, e.g., brand loyalty, causes customers to be both high volume and multichannel."

A more recent stream of work focuses on *within-user* analysis Wang et al. (2015); Narang and Shankar (2019). These papers show that customers who adopt a retailer's mobile app (and become omnichannel) are more profitable (or spend more) than those who choose to remain offline only.[2] Nevertheless, because these papers compare users who become omnichannel to those who stayed offline-only, they cannot answer the question of whether those who switch to omnichannel due to external reasons (e.g., the COVID-19 pandemic) behave differently than those who switched to omnichannel organically.

In this paper, we seek to quantify the effect of becoming omnichannel on the revenue and profitability of different types of switchers. There are some fundamental challenges that we need to overcome to quantify the returns to becoming omnichannel. First, the gold standard approach to obtain the causal effect of becoming omnichannel is to run a randomized experiment where one group of consumers is randomly assigned to be omnichannel and another group to be offline-only. However, channel choice is a consumer-level decision (unlike other standard marketing interventions like advertising/promotions) and as such, this treatment (i.e.,

[1] While some previous literature (Verhoef et al., 2015) makes a distinction between the terms multichannel and omnichannel, in this paper we use the two interchangeably. So, an omnichannel customer is a customer who interacts with a retailer through both online and offline channels.

[2] There is a separate cross-channel literature, which, while not explicitly comparing multi-channel and single-channel consumers, has shown that online and offline channels are complementary in the long run and that adding a physical retail store can boost Avery et al. (2012); Wang and Goldfarb (2017); Bell et al. (2018); Li (2021) or cannibalize Forman et al. (2009); Shriver and Bollinger (2022) online channel sales and that different online mediums (smartphones vs. tablets vs. PC) can be substitutes or complements for each other Xu et al. (2017). See Zhang et al. (2010) and Cui et al. (2021) for a summary of the larger literature and challenges related to omnichannel retail strategy.

channel choice) cannot be randomly assigned by the firm. Thus, there is no direct way to measure the Average Treatment Effect (ATE) of being omnichannel.[3] Second, even if the firm could obtain the ATE of becoming omnichannel, this is not an actionable metric because the firm cannot force consumers to become omnichannel. Therefore, we focus on an alternative metric of interest, the Average Treatment Effect on Treated (ATT), which is a measure that captures the impact of the treatment on those who were treated (i.e., those who became omnichannel). In particular, a key question for the firm is this: do consumers who become omnichannel because of external incentives and/or macro shocks show the same incremental increase in revenue and profitability (post-switching) as those consumers who switched organically? If the answers to both these questions are yes, then the firm can set the appropriate expectations for returns to promotional campaigns that encourage switching to omnichannel. However, if the answer to either of these is no, then the firm should temper its expectations of the effects of such campaigns accordingly.

We use data from a large pet supplies retailer from Brazil that has a large network of brick-and-mortar stores as well as a strong online presence (website and mobile app) to answer these questions. We have access to data on consumer purchases across all channels for the period of 2019-2023. We take advantage of an exogenous shock that affected consumers' channel choice during this period – the COVID-19 pandemic. In March-April 2020, many consumers who were previously offline-only switched to online channels (i.e., became omnichannel) due to store closures, health concerns, or compliance with social distancing norms. We compare the behavior of these consumers (whom we refer to as covid switchers) with that of consumers who were also offline-only in 2019 and chose to become omnichannel in the period just before the onset of COVID-19 (we refer to these consumers as organic-switchers). Thus, unlike the earlier papers, we have two groups of switchers (or offline-only consumers who became omnichannel for different reasons), and we compare the pre- and post-switching behaviors of these two groups. This, in turn, allows us to answer the question of whether the ATT of the consumers who were incentivized to switch due to external forces (in this case COVID-19) is systematically different from that of organic switchers.

We construct consumer-level monthly panel data that quantifies the volume, frequency, and variety of a consumer's purchase behavior – the total spend, relative spend in offline and online channels, quantity bought, number of orders, number of items purchased, number of categories purchased, and so on. We find that, even though there are small differences in the demographics of organic- and covid switchers, the pre-switching purchase behavior of the two groups is quite similar. In contrast, both groups of switchers are systematically different from consumers who remain offline-only, on both demographics and purchase behavior. This is the first interesting finding for two reasons. First, it confirms that non-switchers or offline-only consumers are not the appropriate control group for switchers (of any kind) since they are systematically different from those who self-select to become omnichannel. Conversely, it also suggests that those who switch from offline-only to omnichannel – irrespective of the reason for switching – tend to have behaved similarly pre-switching, and as such the group of organic switchers can function as a control group for covid switchers (and vice-versa).

[3]Alternatively, the firm can run an experiment with the Intent-To-Treat (ITT) design, where some consumers are randomly provided incentives to switch to omnichannel while others are not. Nevertheless, even in these cases, compliance can be endogenous, and this design still does not provide a population-level ATE. See Angrist and Imbens (1995) and Mummalaneni et al. (2023) for details.

Next, we conduct a DID analysis and show that there is no significant difference in the post-switching spend of organic switchers and covid switchers, even though both groups spend more than offline-only customers. Together with the results on the pre-switching behavior, these findings suggest that, if the firm can convince consumers to become omnichannel, those who do so due to such external incentives are likely to provide the same incremental gains as those who became omnichannel organically. Managerially, this is a valuable finding since this allows firms to use the post-switching revenue of consumers who became omnichannel organically as a reasonable proxy of the expected revenue for consumers who may switch due to external incentives and/or encouragement by the firm. In other words, conditional on switching, the channel effect seems to dominate any potential self-selection effects.

Next, we examine if there are differences in how consumers' spending is distributed across different channels after they become omnichannel. We find that organic switchers spend 49.6% of their money offline, whereas covid switchers spend 56.7% of their money offline (on average). Further, we see that while both groups are slowly shifting more of their purchases to online channels, covid switchers are switching online at a slower rate. We investigate the age of covid switchers (on average, they are slightly older than organic switchers) as a potential explanation for this finding. We look into heterogeneity in the share of offline shopping by age among the two groups of switchers and find that the difference in the rate at which the share of offline shopping drops over time post-covid, is mainly being driven by older covid switchers who are likely less online-savvy (e.g. Yang and Ching, 2014) and therefore are less likely to switch most of their purchasing to online channels, after being nudged to start shopping online during covid.

These differences in channel choices between the two groups nevertheless have important implications for the profitability of the two groups. In our data, we see that both channels have largely similar costs, but the online channel has lower prices, i.e., offline channels tend to be more profitable, on average. Thus, even though both covid and organic switchers spend similar amounts after becoming omnichannel, they could be differentially profitable due to differences in their use of the two channels. To examine if this is indeed the case, we conduct a DID analysis on profitability and find that covid switchers are more profitable after conversion to omnichannel, compared to organic switchers. These results indicate that pushing customers to become omnichannel through either a macro/environmental shock or through marketing efforts of the firm, can indeed be profitable for the firm since such consumers may generate even higher profits than those who become omnichannel customers voluntarily.

In summary, our paper makes three main contributions to the literature on omnichannel retailing and consumer value. First, we show that consumers who become omnichannel increase their spending by the same amount after becoming omnichannel, irrespective of the reason for switching (either due to their own volition or external shocks). This is an important and distinct finding compared to the earlier research which provided only cross-sectional comparisons and/or compared organic switchers to non-switchers (offline-only consumers). As such, we show that conditional on switching, both groups exhibit similar incremental revenues. This finding suggests that firms can use the incremental revenue from organic switchers as a reasonable proxy for the incremental revenues from switchers who switch due to external incentives or shocks.

Second, we show that consumers who were incentivized to become omnichannel due to external incentives are slower in adopting the online channel, which in turn has profitability implications for the firm. We find that this is driven mainly by older consumers. To the extent that the offline channel is more profitable in our setting, we find that those consumers who were pushed to switch for external reasons (COVID-19) are more profitable than organic switchers. This finding cautions firms against naively interpreting the profitability of organic switchers as the profitability of users who switch due to macroeconomic shocks or firm-driven incentives. Finally, from a methodological perspective, we show how to leverage the external COVID-19 shock that affected consumers' retail shopping behavior to overcome some challenges with the problem formulation and analysis in this setting. COVID-19 and/or other macro-shocks combined with a similar DID approach can be used to quantify the effect of the retail environment on various constructs of user behavior, e.g., the impact of retail competition, offline-online product availability, etc. Our paper thus also contributes to the nascent literature that quantifies the impact of COVID-19 on consumers' purchase behavior (Oblander and McCarthy, 2022; Chen et al., 2021; Hwang et al., 2020).

2 Setting and Data

2.1 Setting

Our data come from one of the largest Brazilian pet supplies retailers, which sells a wide variety of products for pets, such as pet food and snacks, medicines/pharmacy items, hygiene and grooming products, and various pet toys and accessories. The firm has both offline and online channels. The retailer operates more than 200 physical stores in Brazil, which serve as the offline channel, and has a website as well as a mobile app, which serve as the online channel.

2.2 Data

Our data include transaction-level records of all individual customers from both online and offline channels during the 56-month period from January 1st, 2019 to August 31st, 2023. Each transaction record is uniquely identified by an order number with one or more purchased items. Each purchased item contains the unit price, purchased quantity, brand, purchased channel (online/offline), subcategory, and category (food, hygiene, etc.).

Customer purchases are tracked by the firm using a National Identification Number (similar to a Social Security Number), which is unique to each individual in Brazil. Customers who purchase online (either through website or app) must be logged in to their profile, and therefore all online purchases can be traced back to the National ID. Customers are also asked their National ID when making an in-store purchase, however this is not mandatory. Discussions with the retailer revealed that while most offline transactions are also linked with a customer profile, about a quarter of in-store transactions are not linked to a National ID. However, this number has remained fairly steady since 2019, and wasn't affected by Covid-19.

One concern with any channel analysis is that the results could be driven by assortment differences across the two channels. To ensure that this is not the case, we exclude data from four product categories that were not available in both channels for the entirety of the observation period; see Table A1 in Appendix A for a summary of these categories and their availability information. Together, sales from these categories account for only 0.63% of the total sales during the observation period; as such we do not expect these to play any significant role in the findings.

The total number of consumers who shopped at least once during our sample period is 5,517,802.[4] For each of these consumers, we aggregate the transaction-level data into customer-month-level panel data over the customer's entire purchase history within the observation period (January 2019 to August 2023), starting from his or her initial purchase month and ending with the last observed purchase month. For example, if a customer purchased for the first time in our data in August 2019 and the last observed purchase is in July 2020, then the panel for this customer spans 12 months, from August 2019 to July 2020. If there are intermediate months in which a customer did not make any purchase, the values for the monthly purchase-related variables are set to zero. Alternatively, we can let each consumer's panel run to the end of the observation period; doing so has no significant impact on the results (see the robustness checks in §4.4.4).

We now discuss the summary statistics of our data. In §2.2.1, we summarize the consumer-level demographic data, and in §2.2.2, we summarize the monthly purchase behaviors.

2.2.1 Time Invariant Consumer Attributes

Table 1: Summary Statistics of Customer-Level Categorical Demographic Variables

Gender	Count	Percentage (%)
Female	2,324,748	42.13
Male	1,507,303	27.32
No Information	1,685,751	30.55
Total	5,517,802	100.00

Table 2: Summary Statistics of Customer-Level Demographic Variables

Variable	Mean	SD	25th	50th	75th	(Min, Max)	Count	Missing (%)
Age	40.97	12.76	31	39	49	(7, 99)	4,264,631	22.71
HouseholdIncome	1861.12	1475.65	827.37	1361.94	2458.55	(8, 30826)	4,120,396	25.33
PanelLength	17.11	17.32	1	11	30	(1, 56)	5,517,802	0.00

We first summarize the time-invariant attributes of the consumers in the data. We have data on five consumer-specific variables:

- $Gender_i$: Categorical variable denoting $i's$ gender (Male, Female, or No Information).
- Age_i: Customer i's age.
- $HouseholdIncome_i$: Customer $i's$ average monthly household income.
- $PanelLength_i$: Period spanning the first and last purchases for customer i (in months).
- $Zipcode_i$: The zipcode of user i.

We present the summary statistics for demographic variables in Tables 1 and 2. Out of the 5,517,802 customers, 2,324,748 (42.1%) are Female, 1,507,303 (27.3%) are Male, and the rest (30.6%) do not report

[4] This is the set of consumers who purchased at least one item from one of the non-excluded categories.

gender information. The median consumer is 39 years old. In addition, the mean and median of monthly household incomes are 1861.12 and 1361.94 Brazilian Real, respectively, indicating a positively skewed distribution. Next, we see that the average and median panel lengths are 17.11 and 11 months respectively, which suggests that there is consumer entry as well as churn (and possibly longer purchase cycles) during the observation period. Finally, we observe that customers in our data sample come from 300,774 distinct zip codes across all of the 28 states of Brazil (though 19.5% of total customers do not provide the zip code information).

2.2.2 Customer Monthly Purchase Behaviors

Table 3: Customer-Month Summary Statistics

Variables	mean	std	min	25%	50%	75%	max	count
				Transactions across both Channels				
Spend	331.86	799.43	0	0	0	441.84	2,340,617.16	94,415,101
Quantity	2.75	11.25	0	0	0	3.00	44,158.00	94,415,101
Orders	0.68	1.05	0	0	0	1.00	430.00	94,415,101
UniqueItems	1.60	3.09	0	0	0	2.00	4,191.00	94,415,101
UniqueBrands	1.29	2.19	0	0	0	2.00	332.00	94,415,101
UniqueSubcategories	1.12	1.78	0	0	0	2.00	64.00	94,415,101
UniqueCategories	0.95	1.40	0	0	0	2.00	18.00	94,415,101
				Transactions in Offline Channel				
Spend	220.38	680.91	0	0	0	198.06	2,340,617.16	94,415,101
Quantity	2.05	10.04	0	0	0	2.00	44,158.00	94,415,101
Orders	0.49	0.90	0	0	0	1.00	430.00	94,415,101
UniqueItems	1.27	2.89	0	0	0	1.00	4,191.00	94,415,101
UniqueBrands	1.02	2.05	0	0	0	1.00	332.00	94,415,101
UniqueSubcategories	0.89	1.68	0	0	0	1.00	64.00	94,415,101
UniqueCategories	0.75	1.33	0	0	0	1.00	18.00	94,415,101
				Transactions in Online Channel				
Spend	111.48	424.95	0	0	0	0	276,399.66	94,415,101
Quantity	0.70	4.89	0	0	0	0	4,800.00	94,415,101
Orders	0.19	0.56	0	0	0	0	130.00	94,415,101
UniqueItems	0.34	1.17	0	0	0	0	141.00	94,415,101
UniqueBrands	0.28	0.91	0	0	0	0	44.00	94,415,101
UniqueSubcategories	0.26	0.77	0	0	0	0	22.00	94,415,101
UniqueCategories	0.23	0.65	0	0	0	0	13.00	94,415,101

We quantify consumers' purchase behavior on three dimensions – *Volume*, *Frequency*, and *Variety*.

- Volume of purchase is operationalized using two variables – (1) ($Spend_{it}$), which is the revenue generated from the customer i in month t, and (2) ($Quantity_{it}$), the total number of items purchased by customer i in month t.
- Frequency is measured by the number of unique orders per month ($Orders_{it}$). This metric focuses on purchase frequency without considering the value of purchases, which differs from volume metrics.
- Finally, we quantify the variety of consumers' purchases using the unique number of items ($UniqueItems_{it}$),

brands ($\textit{UniqueBrands}_{it}$), sub-categories ($\textit{UniqueSubcategories}_{it}$), and categories ($\textit{UniqueCategories}_{it}$) purchased in month t.

To preserve the confidentiality of the data provider, we mask the true spend value ($\textit{Spend}_{it}$) by multiplying it by an undisclosed scalar. All the descriptive and model estimation results are thus shown in terms of Masked Currency Units (MCU). The rest of the variables listed above are shown and used as is, without any masking.

Table 3 presents customer-month-level summary statistics of all variables discussed above for three cases: (1) Transactions across both channels, (2) Transactions in the offline channel, and (3) Transactions in the online channel. Table 3 shows the mean statistic of each variable, measured at the customer-month-channel level, which is calculated as $\frac{\sum_{i=1}^{N}\sum_{t=1}^{T_i} X_{itc}}{\sum_{i=1}^{N} T_i}$, where X_{itc} is the customer-month-channel level variable, T_i is the panel length of customer i, and N is the total number of customers, which is 5,517,802. $\sum_{i=1}^{N} T_i$ is the total number of observations, which is 94,415,101.

We find that, on average, consumers spend more offline than online. It is important to note that although the unconditional mean of monthly offline spend (220.38 MCU) is higher than online spend (111.48 MCU), we cannot conclude that, on average, customers spend more money offline than online. This is because some customers may only access one sales channel, or they strategically select the channel that benefits them the most or choose different channels for different types of purchases. We will discuss these ideas in detail later in the paper. Further, we find that all the variety and frequency variables also are higher for offline purchases compared to online purchases. All the purchase metrics have over 50% of observations as zero, which suggests that purchase frequencies are lower than once per month for a large fraction of consumers, though some consumers purchase very frequently and buy large varieties of products. We also report these summary statistics separately for the pre and post-Covid periods in Web Appendix B.

3 Descriptive Analysis

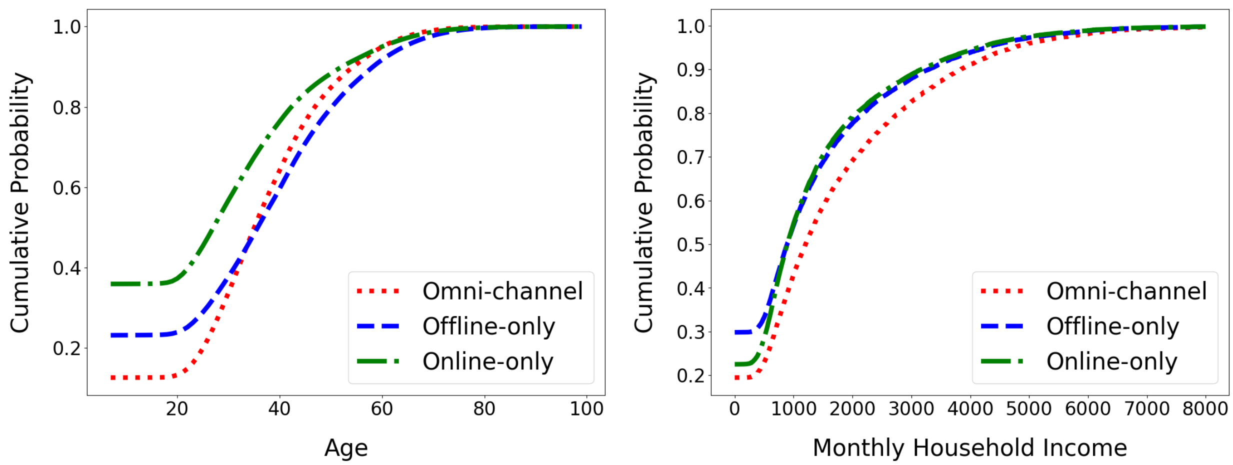
In this section, we provide some descriptive analyses summarizing the relationship between channel choice and consumers' purchase behaviors. In particular, we are interested in quantifying whether the value/revenue generated by a customer is systematically correlated with their channel choice. In §3.1, we first present some cross-sectional evidence linking channel choice to customer value. Next, in §3.2, we present a within-user analysis that can control for consumer-specific variables.

3.1 Cross-sectional Comparison

Table 4: Frequency distribution of gender - offline-only/online-only/omnichannel

Gender	Online-only (%)	Offline-only (%)	Omnichannel (%)	Overall (%)
No Information	35.32	33.47	22.01	30.56
Female	42.00	37.30	50.68	42.12
Male	22.69	29.24	27.30	27.32
Total	100	100	100	100

Figure 1: Cumulative distribution function (CDF) of Age/Monthly Household Income - offline-only, online-only/omnichannel



We start by categorizing consumers into three mutually exclusive groups:

- Omnichannel consumers – consumers who have purchased from both channels at least once during the observation period.
- Offline-only consumers – consumers who have exclusively purchased from offline stores during the observation period.
- Online-only consumers – consumers who have only purchased using the online channel during the observation period.

In total, our data consists of 1,585,442 (28.7%) omnichannel consumers, 2,781,473 (50.4%) offline-only consumers, and 1,150,887 (20.9%) online-only consumers.

We find that these three groups are systematically different from each other in terms of demographics; see Table 4 and Figure 1. Omnichannel consumers are more likely to be female compared to offline and online consumers. Online-only consumers tend to be younger than both offline and omnichannel consumers. Further, offline-only consumers tend to be lower income while omnichannel consumers tend to be the wealthiest. There are also significant differences in channel choice across geographic locations (see Figure A1 in Web Appendix C). Overall, these differences suggest that there is significant self-selection in channel choice based on consumer attributes.

Next, we summarize the purchase behaviors of each of these groups in Table 5[5]. The means shown in this table are calculated for each group g (where g is either omnichannel, offline-only, or online-only) as follows: $\frac{\sum_{i=1}^{N_g}\sum_{t=1}^{T_i} X_{it}}{\sum_{i=1}^{N_g} T_i}$, where i denotes a customer belonging to group g, t denotes the year-month index, T_i is the panel length of customer i, and N_g is the total number of consumers in group g. As we can see, omnichannel customers tend to spend more, and buy higher quantities and larger varieties of products compared to offline-only or online-only consumers. If we compare consumers who exclusively buy offline

[5]We also include versions of this split by pre- and post-Covid in Web Appendix C

Table 5: Customer-Month Summary Statistics By Customer Types - Omnichannel/Offline-only/Online-only

Variables	Mean	Std	Min	25%	50%	75%	Max	Count
				Omnichannel Customers				
Spend	424.70	817.98	0	0	0	594.97	301,134.12	44,190,696
Quantity	3.45	10.16	0	0	0	3.00	3,440.00	44,190,696
Orders	0.82	1.19	0	0	0	1.00	236.00	44,190,696
UniqueItems	1.90	3.32	0	0	0	3.00	355.00	44,190,696
UniqueBrands	1.52	2.40	0	0	0	2.00	127.00	44,190,696
UniqueSubcategories	1.31	1.93	0	0	0	2.00	40.00	44,190,696
UniqueCategories	1.10	1.50	0	0	0	2.00	16.00	44,190,696
				Offline-only Customers				
Spend	245.00	803.54	0	0	0	283.64	2,340,617.16	42,922,391
Quantity	2.22	12.77	0	0	0	2.00	44,158.00	42,922,391
Orders	0.56	0.93	0	0	0	1.00	430.00	42,922,391
UniqueItems	1.41	2.99	0	0	0	2.00	4,191.00	42,922,391
UniqueBrands	1.14	2.05	0	0	0	2.00	332.00	42,922,391
UniqueSubcategories	1.01	1.69	0	0	0	2.00	64.00	42,922,391
UniqueCategories	0.85	1.34	0	0	0	1.00	18.00	42,922,391
				Online-only Customers				
Spend	280.52	565.35	0	0	0	437.50	110,682.95	7,302,014
Quantity	1.61	6.96	0	0	0	1	4,800.00	7,302,014
Orders	0.50	0.69	0	0	0	1	130.00	7,302,014
UniqueItems	0.86	1.56	0	0	0	1	141.00	7,302,014
UniqueBrands	0.74	1.19	0	0	0	1	32.00	7,302,014
UniqueSubcategories	0.67	1.02	0	0	0	1	20.00	7,302,014
UniqueCategories	0.61	0.87	0	0	0	1	11.00	7,302,014

with those who exclusively buy online, we see that online-only consumers tend to spend more but purchase lower quantities and lower variety of items. This is mainly because online-only consumers tend to buy larger pack sizes of pet foods and more expensive products compared to offline-only consumers.

However, these patterns could be explained by the differences in observed consumer attributes across the three channels. To examine if this is the case, we estimate the following linear regression:

$$y_{it} = CustomerType_i + Gender_i + Age_i + HouseholdIncome_i + Zipcode_i + YearMonth_t + \epsilon_{it} \quad (1)$$

where y_{it} can denote one of the purchase metrics (e.g. Spend, Orders, UniqueCategories, etc.) of customer i in month t. The main explanatory variable of interest is $CustomerType_i$, which is a dummy variable that denotes the customer's channel choice: omnichannel, offline-only, or online-only. We specify online-only as the baseline category. In addition, we include other control variables related to customer demographics (*Gender*, *Age*, *Zipcode*, and *HouseholdIncome*) for the reasons discussed earlier, i.e., to control for individual characteristics that can affect purchase behavior. Finally, we also control for time-varying unobserved shocks in period t that can affect purchase behavior (such as seasonality and economic factors) using a period fixed effect ($YearMonth_t$).

Table 6: Cross-sectional Comparison of Customers' Monthly Spend

Dependent Variable:	Spend	
Model:	(1)	(2)
Variables		
Constant	280.522***	
	(0.515)	
Omnichannel	144.181***	127.761***
	(0.701)	(0.722)
Offline-only	-35.523***	-49.178***
	(0.624)	(0.684)
Control Variables		
Age		Yes
Gender		Yes
Zipcode		Yes
HouseholdIncome		Yes
Fixed-effects		
YearMonth		Yes
Fit statistics		
Observations	94,415,101	94,415,101
R^2	0.012	0.056
Within R^2		0.010

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

Table 6 presents the difference in total spend of the omnichannel and offline-only customers relative to the baseline category (i.e. online-only customers). In column (1), we use only customer type as the explanatory variable and find that spending patterns across customer types are the same as those discussed above. Next, in column (2), we estimate the same model with all the controls/fixed effects and find that the effects persist even after controlling for selection using the observed seasonality and demographic variables. Specifically, we find that, compared to offline-only consumers, omnichannel consumers spend about 178 MCU more per month. Interestingly, the magnitude of this effect is almost the same in both models (1) and (2), i.e., the addition of controls and fixed effects did not materially change this result. Further, we replicate the specification in column (2) of Table 6 with other purchase metrics in Table A4 in Web Appendix C, and find that the results are consistent with the observations from Table 5.

These findings align with previous literature such as Kumar and Venkatesan (2005); Ansari et al. (2008); Kushwaha and Shankar (2013), who find that omnichannel consumers are more valuable than consumers who only purchase via one channel. Given that offline-only consumers are the lowest spenders and omnichannel consumers are the heaviest spenders, a natural question for the firm is whether it can increase revenues by nudging offline-only consumers to also sign up for its online channels (and become omnichannel). However, the results from Table 6 could be a combination of the effect of having access to multiple channels (*channel effect*) as well as unobserved but systematic differences across consumers who use the two channels (*self-*

selection effect). For example, users who are loyal to the firm and/or heavy consumers are more likely to use both channels. As such, converting an offline-only consumer to become omnichannel may not necessarily lead to the same incremental gains.

3.2 Within-user Comparison

Table 7: Customers' Spend Before/After Switching to Omnichannel

Dependent Variable:	Total Spend	
Model:	(1)	(2)
Variables		
Constant	317.9^{***}	
	(0.622)	
Post_switch	219.4^{***}	115.1^{***}
	(0.899)	(0.841)
Fixed-effects		
Customer		Yes
YearMonth		Yes
Fit statistics		
Observations	27,239,397	27,239,397
No. of customers	871,158	871,158
R^2	0.01744	0.39461
Within R^2		0.00256

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

Given the problems with the cross-sectional analysis discussed above, we now perform a within-user comparison that allows us to control for some of the selection issues. We focus on consumers who were offline-only at the beginning of our observation period and became omnichannel sometime during the observation period. 871,158 consumers satisfy this criteria and we use their data for our analysis. To cleanly identify the difference in spending before and after switching to omnichannel, we exclude the observation of the month a customer decides to switch. We then run the following regression on these consumers' monthly spending:

$$y_{it} = \alpha_i + \beta_0 Post\_switch_{it} + YearMonth_t + \epsilon_{it}, \tag{2}$$

where y_{it} refers to spend or another outcome variables of interest, α_i is a consumer fixed-effect, $Post\_switch_{it}$ is an indicator for whether consumer i had switched to become omnichannel before period t, and $YearMonth$ is a timeperiod fixed effect as before. Table 7 presents the results from this regression. Model (1) shows the results without the customer and year-month fixed effects and model (2) shows the results with the fixed effects. We see that, on average, customers spend 115.1 MCU more after they switch to omnichannel. Note that this result is significantly smaller than the 178 MCU from the previous cross-sectional analysis. These results are consistent with earlier work that has shown that consumers who adopt digital channels tend to spend more (Venkatesan et al., 2007; Narang and Shankar, 2019). We also perform a similar analysis on other metrics of interest (Quantity, No. of orders, Unique items, Brands, Categories, and Subcategories) and find that all of them show a positive increase after switching to omnichannel (see Table A5 in Web Appendix C). However, as with Spend, the magnitude of these effects is much smaller than what we saw in the cross-sectional analysis.

In sum, while the within-consumer analysis controls for some self-selection problems, there is still a concern that consumers who choose to become omnichannel are still systematically different from those who choose not to, i.e., these findings could still be a combination of self-selection and channel effects. As such, the effect size here could still be an inflated measure of the incremental value from converting an offline-only consumer to an omnichannel consumer. We address this further in the next section.

4 Empirical Analysis

4.1 Firm's problem and Identification Strategy

We start by formalizing the firm's problem. At a high level, the firm's goal is to quantify the causal effect (if any) of a consumer being omnichannel compared to being offline-only. If there is a significant positive impact of being omnichannel, then the firm can consider incentives/strategies that encourage consumers to become omnichannel. However, quantifying the impact of channel choice on consumer behavior is not only challenging but also managerially irrelevant for several reasons.

First, the gold standard to obtain such a measurement would be a randomized experiment where one group of consumers is randomly assigned to be omnichannel and another group to be offline-only. However, since channel choice is a consumer-level decision, unlike advertising/promotions, this treatment (i.e., channel choice) cannot be randomly assigned by the firm. Thus, there is no direct way to measure the Average Treatment Effect (ATE) of being omnichannel. Alternatively, the firm can run an experiment with an intent-to-treat (ITT) design, where some consumers are randomly provided incentives (e.g., price reductions for in-app purchases or sending emails/promotions advertising the online channel) to switch to omnichannel while others are not. Nevertheless, even in this setting, the actual decision to become omnichannel is still a consumer choice and can be endogenous. To the extent that some consumers never switch online, it is again infeasible to measure ATE.[6] In other words, the Average Treatment Effect (ATE) is not only infeasible to obtain, but also not particularly actionable.

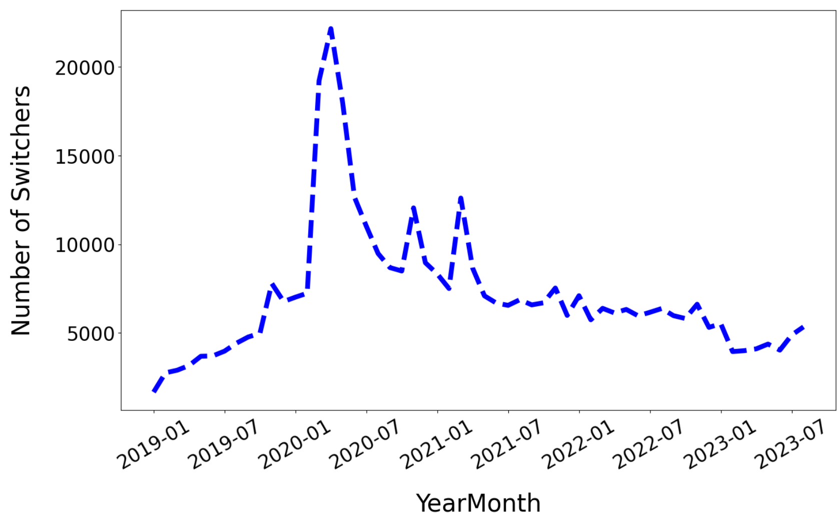
Second, even if the firm could somehow obtain the incremental impact of being omnichannel (compared to being offline-only), it cannot force consumers into certain channel choices. In other words, the Average Treatment Effect (ATE) is not only infeasible to obtain, but also not particularly actionable. Rather, from a firm's perspective, the main quantity of interest is the Average Treatment Effect on Treated (ATT). That is, if the firm could incentivize certain offline-only consumers to become omnichannel, what would the incremental revenue from the switchers be? Further, should we expect the incremental revenue from such switchers (i.e., those who switched due to external forces, e.g., other external shocks) to be systematically

[6] Indeed the measurement obtained from 2SLS estimators in these cases is not ATE, but the Average Causal Response, which is a weighted sum of individual treatment effects where the weights are compliance scores; see Angrist and Imbens (1995); Syrgkanis et al. (2019); Mummalaneni et al. (2023) for further discussions on the analysis of such experiments.

different compared to the consumers who organically switched into being omnichannel? That is, should we expect the incremental gains from §3.2 to reflect the ATT of nudging a consumer to become omnichannel?

To examine these questions, we take advantage of an exogenous shock that affected consumers' channel choice – the COVID-19 pandemic. In March-April 2020, many consumers who were previously offline-only switched to online channels (i.e., became omnichannel).[7] As Figure 2 shows, during March and April 2020, there was a large peak in the number of offline customers who first switched to online shopping; the number of switchers sharply increased from approx. 7000 in Jan and Feb 2020 to approx. 20,000 in March and April 2020.[8] This suggests that the onset of COVID-19 functioned as an exogenous shock that led many consumers to become omnichannel. As such, we can use this shock to quantify the ATT for users who switch due to external forces (in this case COVID-19) and compare it to the ATT for users who switch organically.

Figure 2: Total Number of Offline-only Customers First Switching to Omnichannel at each month



4.2 Sample Construction and Summary Statistics

To that end, we consider the cohort of customers who made their first offline purchase before 2019-12-31 and were offline-only at the end of 2019 and continued to be the firm's customers after the onset of COVID-19. Then, we categorize each consumer from this cohort into one of the following groups:

- Covid switchers: Customers who made the first online purchase between 2020-03-16 and 2020-04-30 [9]. These consumers became omnichannel during the onset of COVID-19.[10]

[7]This could be due to a combination of health concerns, compliance with Covid-19 social distancing norms as well as offline store closures.

[8]In Figure 2, we also see smaller peaks associated with the second and third waves of COVID-19.

[9]2020-03-16 was the date when the Brazilian government started implementing physical distancing and confinement measures in response to Covid-19 (Faria de Moura Villela et al., 2021).

[10]The set of people who switched during this period could also include consumers who would have organically switched even without the COVID-19 shock. We have no clear way of distinguishing such potential organic switchers from those who switched due to COVID-19. Nevertheless, based on Figure 2, it seems like approximately two-thirds of these consumers switched due to COVID-19. That is, a significant majority of these consumers switched due to COVID-19, and we call this cohort covid switchers

- Organic switchers: Customers who made the first online purchase between 2020-01-01 and 2020-02-29, i.e., the two months before the onset of COVID-19. These are the consumers who became omnichannel organically, in the period just before the onset of COVID-19.
- Offline-only: Customers who are active (i.e. at least one purchase before 2019-12-31 and after 2020-05-01), and never make online purchases during the observation period. These consumers did not become omnichannel anytime and continued as offline-only.

In total, we have 33,246 covid switchers, 12,799 organic switchers, and 573,934 offline-only customers. Note that a large fraction of customers were able to remain offline only even during COVID-19 because the focal retailer's stores were allowed to remain open as Pet Supplies retailing was deemed to be an essential service by the Brazilian government. We now discuss the descriptives of each of these groups below.

Demographics: We show the distribution of demographics of all three groups in Tables 8 and Table 9 Panel A. We see that all three groups are systematically different from each other on all the demographic variables, including gender, geographic location, age, and income. That said, the two groups of switchers are more similar than the offline-only customers on all demographic variables.

Table 8: Customer Demographics - Summary Statistics of Gender (Panel A) and Geographic State (Panel B)

Categories	Organic Switcher (%)	Covid Switcher (%)	Offline-only (%)
Panel A: Gender			
No Information	13.49	12.14	16.28
Female	57.36	59.55	46.21
Male	29.16	28.31	37.51
Total	100.00	100.00	100.00
Panel B: Geographic State			
São Paulo	49.68	58.78	63.70
Minas Gerais	12.13	8.58	4.81
Rio de Janeiro	6.95	6.69	4.30
Distrito Federal	5.48	6.21	4.26
Santa Catarina	3.46	3.33	2.90
Paraná	1.86	1.68	2.31
Rio Grande do Sul	2.77	3.21	2.20
Goiás	3.36	2.37	2.13
Mato Grosso do Sul	1.88	0.93	1.05
Espírito Santo	2.43	1.27	0.91
Bahia	3.31	1.74	0.69
Other States	2.41	1.14	0.39
Missing States	4.28	4.07	10.34
Total	100.00	100.00	100.00

Pre-switching purchase behavior: To summarize the pre-switching purchase behavior of each group of consumers, we define a series of consumer-level variables such as *AvgSpendPerMonth*, *AvgOrdersPerMonth*, *AvgUniqueCategoriesPerMonth*, and so on using data from 2019-01-01 to 2019-12-31.[11] Then we summarize

for simplicity.

[11] We use this observation period because this is the common period during which all three groups of customers purchase only

Table 9: Customer-Level Summary Statistics of Age and Monthly Household Income (Panel A), Pre-switching (Panel B) Purchase Behaviors (Panel B). Columns 1 - 3 describe statistics for offline-only, covid, and organic switchers respectively. Column 4 calculates the difference between covid and organic switchers, and columns 5 and 6 show the two-sample t-test results.

Variable	Offline-only mean	Covid Switchers mean	Organic Switchers mean	Diff	t-stats	p-value
Panel A: Numerical Demographic Variables						
Age	45.62	43.36	41.44	1.92	15.65	0.00
Monthly Household Income	2033.01	2432.11	2200.69	231.41	12.30	0.00
Panel B: Pre-switching Monthly Purchase Behaviors						
AvgSpendPerMonth	284.42	463.46	453.26	10.20	1.80	0.07
AvgQuantityPerMonth	2.96	4.98	4.43	0.55	6.15	0.00
AvgOrdersPerMonth	0.74	0.95	0.95	-0.00	-0.38	0.71
AvgUniqueItemsPerMonth	1.91	2.92	2.89	0.03	0.97	0.33
AvgUniqueBrandsPerMonth	1.55	2.32	2.34	-0.02	-0.92	0.36
AvgUniqueSubcategoriesPerMonth	1.39	2.02	2.04	-0.02	-1.10	0.27
AvgUniqueCategoriesPerMonth	1.17	1.65	1.66	-0.01	-0.62	0.54

these variables for each of the groups defined above. The means shown in Table 9 Panel B, for each group g (covid switchers, organic-switchers, and offline-only) are calculated as follows: $\frac{1}{N_g}\sum_{i=1}^{N_g}(\frac{1}{T_i^{pre}}\sum_{t=1}^{T_i^{pre}} X_{it})$, where i denotes a customer belonging to group g, t denotes the year-month index, T_i^{pre} denotes the length of the consumer i's panel pre-2020, and N_g is the number of customers in group g.

We find that organic-switchers and covid switchers are not systematically different in terms of their pre-switching purchase behaviors; the t-stats of the difference between all the variables for these two groups in panel B of Table A13 all fall below the 5% significance level (except for *AvgQuantityPerMonth* which shows a small but statistically significant difference). However, the purchase behavior of both these groups is quite different from that of offline-only customers; i.e., in the pre-switching period, the switchers (both organic and covid) spent more than offline-only customers, purchased more frequently, and bought a wider variety of brands and categories. Together, the main takeaway from these descriptives is as follows: even though the two groups of switchers show small differences in their demographics, their purchase behavior before switching is largely similar. This suggests that those who switch tend to be behaviorally similar, irrespective of the reason for switching.

Post-switching behavior: Table 10 shows the customer-level summary statistics of post-treatment purchase behavior for all the three groups.[12] [13] Interestingly, in the post-switching period, on average, offline-only customers spent less than what they did in the pre-period (207.43 MCU vs. 284.42 MCU). However, this is

through offline channels.

[12]The means for each group g (covid switchers, organic-switchers, and offline-only) are calculated as follows: $\frac{1}{N_g}\sum_{i=1}^{N_g}(\frac{1}{T_i^{post}}\sum_{t=1}^{T_i^{post}} X_{it})$, where i denotes a customer belonging to group g, t denotes the year-month index, T_i^{post} denotes the length of the consumer i's panel from May 2020, and N_g is the number of customers in group g.

[13]The counts of organic and covid switchers in Table 10 slightly differ from the raw count based on our cohorts' definition because 469 (3.66% of) organic switchers and 933 (2.81% of) covid switchers churned before 2020-05-01, and therefore not observed in the post-covid panel. Including these customers does not affect our identification of post-switching ATT, and the estimation results remain the same.

not the case for organic and covid switchers – both these groups increase their spending in the post-treatment period compared to the pre-period. Further, we see that COVID switchers and organic switchers largely behave similarly in the post-switching period, though there are some small differences in their purchase behaviors. In addition to spending, both groups of switchers also purchase higher quantities and a larger variety of products compared to the offline-only group in the post-treatment period. However, these raw patterns do not account for customer and time-period-specific differences. Therefore, we quantify these differences more carefully in §4.3.

Table 10: Customer-Level Summary Statistics for the Three Groups – Covid Switchers, Organic Switchers, and Offline-only for post-treatment period (i.e., after May 2020).

Variables	mean	std	min	25%	50%	75%	max	count
				Covid Switchers				
AvgSpendPerMonth	525.03	656.22	0.02	161.92	347.30	666.73	40153.38	32,313
AvgQuantityPerMonth	4.89	10.04	0.03	1.00	2.21	4.90	425.88	32,313
AvgOrdersPerMonth	0.90	0.80	0.03	0.38	0.70	1.18	17.81	32,313
AvgUniqueItemsPerMonth	2.29	2.39	0.03	0.80	1.59	2.96	58.62	32,313
AvgUniqueBrandsPerMonth	1.79	1.58	0.03	0.70	1.35	2.40	18.50	32,313
AvgUniqueSubcategoriesPerMonth	1.53	1.21	0.03	0.64	1.23	2.10	12.68	32,313
AvgUniqueCategoriesPerMonth	1.27	0.92	0.03	0.57	1.07	1.77	8.62	32,313
				Organic Switchers				
AvgSpendPerMonth	498.07	583.56	0.50	148.52	325.83	639.91	9751.80	12,330
AvgQuantityPerMonth	4.10	7.89	0.03	0.87	1.92	4.23	224.78	12,330
AvgOrdersPerMonth	0.90	0.82	0.03	0.35	0.68	1.20	10.95	12,330
AvgUniqueItemsPerMonth	2.05	2.11	0.03	0.69	1.41	2.69	30.65	12,330
AvgUniqueBrandsPerMonth	1.63	1.48	0.03	0.61	1.21	2.20	16.32	12,330
AvgUniqueSubcategoriesPerMonth	1.40	1.14	0.03	0.56	1.10	1.92	9.30	12,330
AvgUniqueCategoriesPerMonth	1.17	0.87	0.03	0.50	0.97	1.62	6.78	12,330
				Offline-only				
AvgSpendPerMonth	207.43	334.04	0.00	39.05	104.35	248.78	24573.87	573,934
AvgQuantityPerMonth	1.87	4.55	0.03	0.32	0.79	1.88	487.17	573,934
AvgOrdersPerMonth	0.46	0.52	0.03	0.13	0.29	0.60	40.11	573,934
AvgUniqueItemsPerMonth	1.14	1.54	0.03	0.28	0.65	1.41	123.22	573,934
AvgUniqueBrandsPerMonth	0.92	1.08	0.00	0.25	0.57	1.19	48.44	573,934
AvgUniqueSubcategoriesPerMonth	0.81	0.87	0.03	0.23	0.51	1.07	19.90	573,934
AvgUniqueCategoriesPerMonth	0.69	0.69	0.03	0.21	0.47	0.95	10.00	573,934

4.3 Differences-in-Differences Analysis

We use the customer-month-level panel data for our analysis and exclude the observations of customer purchases between 2020-01-01 and 2020-04-30, which is the period used to define the three groups.[14] Let y_{it} be an outcome variable that characterizes some aspect of user i's purchase behavior in month t.[15] To

[14]By excluding these observations, we avoid capturing the abnormal purchase behaviors used to define the switching behavior and the onset of COVID-19. As such, it allows us to cleanly identify the post-covid difference in customers' purchase behaviors.

[15]We mainly focus on spend in the main text, and present results on other variables (e.g., quantity, orders, unique brands) in the Web Appendix.

Table 11: Difference-in-Difference Results (All Three Groups) - Spend

Dependent Variable:	Monthly Spend	
Model:	(1)	(2)
Variables		
Offline_only	280.1***	
	(0.557)	
Organic_switcher	152.7***	
	(5.35)	
Covid_switcher	185.5***	
	(3.38)	
Post_covid	-51.4***	
	(0.512)	
Organic_Switcher * Post_covid	156.0***	141.6***
	(5.48)	(5.52)
Covid_Switcher * Post_covid	147.6***	141.2***
	(3.42)	(3.54)
Fixed-effects		
Customer		Yes
YearMonth		Yes
Fit statistics		
Observations	22,954,289	22,954,289
R^2	0.01889	0.38510
Within R^2		0.00097

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

formalize the incremental revenue from switching (i.e., the ATT of switching, for the two types of switchers), we now use a Diff-in-Diff specification as follows:

$$y_{it} = \alpha + \beta_0 \mathbb{I}\{Organic\_switcher_i\} + \beta_1 \mathbb{I}\{Covid\_switcher_i\} + \beta_2 \mathbb{I}\{Post\_covid_t\} \\ + \beta_3 \mathbb{I}\{Post\_covid_t\} \times \mathbb{I}\{Organic\_switcher_i\} + \beta_4 \mathbb{I}\{Post\_covid_t\} \times \mathbb{I}\{Covid\_switcher_i\} + \epsilon_{it}. \tag{3}$$

The intercept α denotes the pre-period average monthly spend of offline-only users. $\mathbb{1}\{Organic\_switcher_i\}$ and $\mathbb{1}\{Covid\_switcher_i\}$ denote the pre-period spend of organic-switchers and covid switchers (compared to offline-only customers), respectively. $\mathbb{1}\{Post\_covid_t\}$ is a binary variable that takes the value of 1 for the post-switch periods (i.e., from 2020-05-01 onwards) and 0 for pre-switching periods (i.e., from 2019-01-01 to 2019-12-31). Thus, β_2 indicates the average incremental revenue of offline-only customers post-May 2020, compared to the pre-2020 period. The coefficients of the interaction term, β_3 and β_4, represent the average incremental revenue per month from organic- and covid switchers in the post-switching period, respectively. We can refine this specification further to control for customer-specific time-invariant heterogeneity and time-variant common shocks by including customer fixed effects (α_i) and time-period fixed effects ($YearMonth_t$)

as follows:

$$\begin{aligned} y_{it} = \alpha_i + YearMonth_t + \beta_3 \mathbb{I}\{Post\_covid_t\} \times \mathbb{I}\{Organic\_switcher_i\} \\ + \beta_4 \mathbb{I}\{Post\_covid_t\} \times \mathbb{I}\{Covid\_switcher_i\} + \epsilon_{it} \end{aligned} \tag{4}$$

We present the results from these DID analyses in Table 11. The results show that, on average, both organic- and covid switchers spend more in the post-covid period, compared to offline-only customers. Organic switchers spend 141.6 MCU more than offline-only customers post-switching, and covid switchers spend 141.2 MCU more per month than offline-only customers, after controlling for customer fixed effects and year-month fixed effects. An interesting pattern here is the striking similarity in the ATT of the organic and covid switchers. That is, conditional on switching, both groups seem to spend very similarly.

To further investigate whether the difference between organic and covid switchers is statistically significant, we run another DID analysis where we set organic switchers as the control group and covid switchers as the treatment group, and estimate the following model:

$$y_{it} = \gamma + \delta_0 \mathbb{1}\{Covid\_switcher_i\} + \delta_1 \mathbb{1}\{Post\_covid_t\} + \delta_2 \mathbb{1}\{Covid\_switcher_i\} \times \mathbb{1}\{Post\_covid_t\} + \epsilon_{it}. \tag{5}$$

Here, the coefficient of the interaction term (δ_2) represents the difference in incremental revenue between covid and organic switchers in the post-switching period (i.e., after May 2020). The results from this model are shown in column (1) of Table 12. As we can see, there is no significant difference in the post-treatment behavior of these two groups. Further controlling for customer- and time-period fixed effects (in column (2)) confirms these findings.

In sum, these findings suggest, that irrespective of the reason for becoming omnichannel, consumers tend to behave similarly after switching (by increasing their overall spending with the firm), i.e., $ATT_{covid\_switchers} \equiv ATT_{organic\_switchers}$. This finding has important implications for omnichannel firms since it suggests that, if the firm can successfully convert users to become omnichannel due to external shocks/incentives, then their post-switching purchase behavior is likely to be similar to those of organic switchers. In other words, conditional on switching, the channel effect seems to dominate any potential self-selection effects.

Next, we examine if there are differences in *how* the spend is distributed across different channels and products post-switching for the two groups (covid and organic switchers). We run the DID analyses specified in Equation (5) but with *Offline spend* and *Fraction of offline spend* as the outcome variables. The results from this exercise are presented in columns (3)–(6) of Table 12. Interestingly, we see a difference between the two groups here – covid switchers spend more offline compared to offline switchers even though there is no significant difference in the spend itself. According to column (6), conditional on customers making some purchase in a month in the post-switching period (i.e., non-zero spend), covid switchers spend around 7% more in offline stores compared to organic switchers. While both groups spend 100% of their spend in offline stores before switching, after switching to omnichannel, organic switchers spend 49.6% of their money offline, whereas COVID switchers spend 56.7% of their money offline.[16]

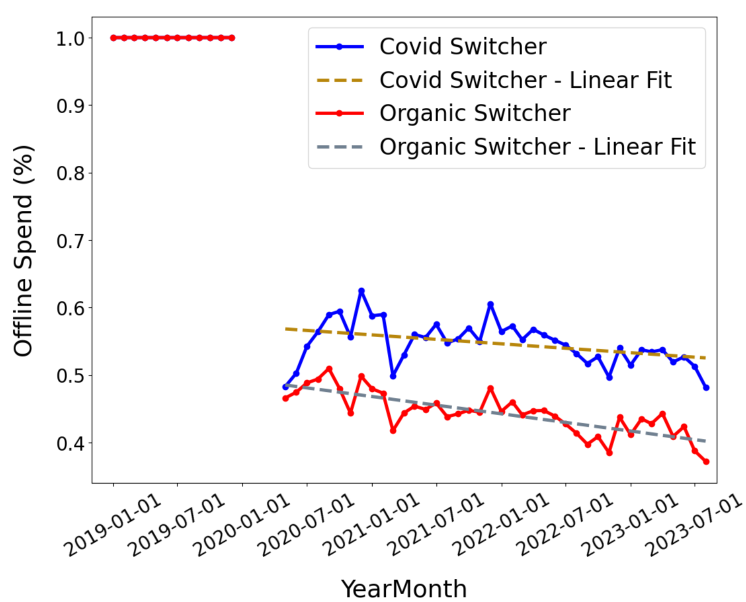
[16]Since covid switchers are more active in the offline channel, we similarly see a corresponding positive effect on the number of orders and variety of products purchased offline. See Table A10 in Web Appendix D).

Table 12: Difference-in-Difference Spend Results (Covid vs Organic Switchers)

Dependent Variables:	Total Spend		Offline Spend		Fraction of Offline Spend	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Organic_Switcher	432.8^{***}		432.8^{***}		1.00^{***}	
	(5.33)		(5.33)		(0.00)	
Covid_Switcher	32.8^{***}		32.8^{***}		0.00	
	(6.28)		(6.28)		(0.00)	
Post_covid	104.6^{***}		-194.0^{***}		-0.514^{***}	
	(5.46)		(4.51)		(0.003)	
Covid_Switcher * Post_covid	-8.41	-1.63	35.7^{***}	38.6^{***}	0.083^{***}	0.071^{***}
	(6.42)	(6.52)	(5.34)	(5.44)	(0.004)	(0.004)
Fixed-effects						
Customer		Yes		Yes		Yes
YearMonth		Yes		Yes		Yes
Fit statistics						
Observations	1,874,756	1,874,756	1,874,756	1,874,756	1,016,120	1,016,120
R^2	0.00187	0.39518	0.01035	0.36244	0.17300	0.50164
Within R^2		1.38×10^{-7}		0.00013		0.00136

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

Figure 3: Fraction of offline spend relative to total spend - Covid vs Organic Switchers



The above numbers are the averages for the entire post-switching period (from April 2020 to August 2023) and mask any trends. Therefore, next, we investigate whether there is a differential trend in the fraction of offline spending between two cohorts in the post-switching period. Figure 3 depicts the fraction of offline spend for organic and covid switchers over the entire observation period. We see that organic switchers persistently spend a lower percentage of their total money offline (red), compared to covid switchers (blue) in the post-switching periods. In addition, the downward trend is steeper for organic switchers, indicating that organic switchers shift their spending to online channels at a much faster than covid switchers.

To formally test whether these differential trends are systematic, we estimate the following model with linear time trend on the post-switch monthly panel data for both groups:

$$y_{gt} = \zeta + \rho_0 t + \rho_1 \mathbb{I}\{Covid\_switcher_g\} + \rho_2 \mathbb{I}\{Covid\_switcher_g\} \times t + \epsilon_{gt}. \tag{6}$$

Here, y_{gt} is the fraction of the offline spend relative to the total spend (ranging from 0 to 1) for group g in period t. t is the numerical time index, and we set the first post-switching month (i.e. 2020-05-01) as $t = 0$. The results from this regression are shown in Table A8 in Web Appendix §D. The estimate of the difference in trend (ρ_2) between these two cohorts is -0.001 (or 0.1%) and is statistically significant at 5%. More specifically, we see that both groups reduce their fraction of offline spend each month, but the group of organic switchers reduces by 0.21% each month, whereas covid switchers decrease by 0.11% every month, i.e., their rate of moving online is slower than that of organic switchers. Together, these findings suggest that even though the reasons for switching do not seem to have any impact on *how much* consumers spend with the firm, there are small differences in *where* they spend it.

We further investigate this finding by considering the attributes on which the two groups of switchers differ. Recall from Table 9 that covid and organic switchers show similar pre-switching purchase behavior but differ in demographics. In particular, covid switchers tend to be older than organic switchers. Therefore, one potential explanation for the above finding could be as follows: older users who switched online due to covid are more likely to continue using offline channels even after switching. That is, the differential trends in the fraction of offline purchases could be explained by the differences in the age distributions of the two groups. To test this hypothesis, we perform a median split of all the users based on age and plot the fraction of offline spending in the post-treatment period separately for senior and young switchers in Figure 4. As before, we see that, in both age groups, covid switchers tend to spend a higher fraction of their spend in offline channels. However, the trends are parallel for both types of switchers (organic and covid) for younger customers; in contrast, among older customers, covid switchers switch to online channels at a slower rate compared to organic switchers. We confirm these findings formally using linear time trend regressions in Table A9 in Web Appendix §D. Thus, the differential trends from Figure 3 are mainly driven by the fact that older covid switchers are more likely to continue to stick to offline shopping even after becoming omnichannel [17]. Overall, these findings suggest that while the firm can use external incentives to convince consumers to become omnichannel, older customers may exhibit a stronger preference for offline channels and less inclination to fully migrate to newer technologies (Yang and Ching, 2014).

Finally, to further explore the differences in the spending behavior of the two groups of consumers

[17]We also check whether differences in income drive the differences in offline shares, but do not find evidence in support of this hypothesis.

Figure 4: Heterogeneous Effect by Age - Fraction of offline spend relative to total spend

(a) Senior Switchers (Age $>=$ 41) (b) Young Switchers (Age $<$ 41)

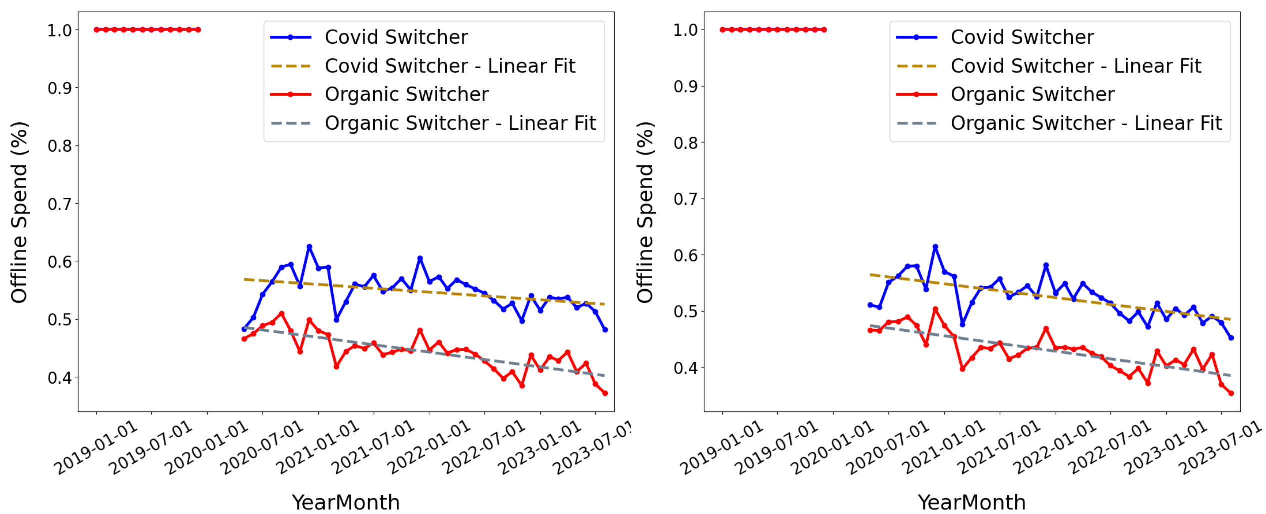


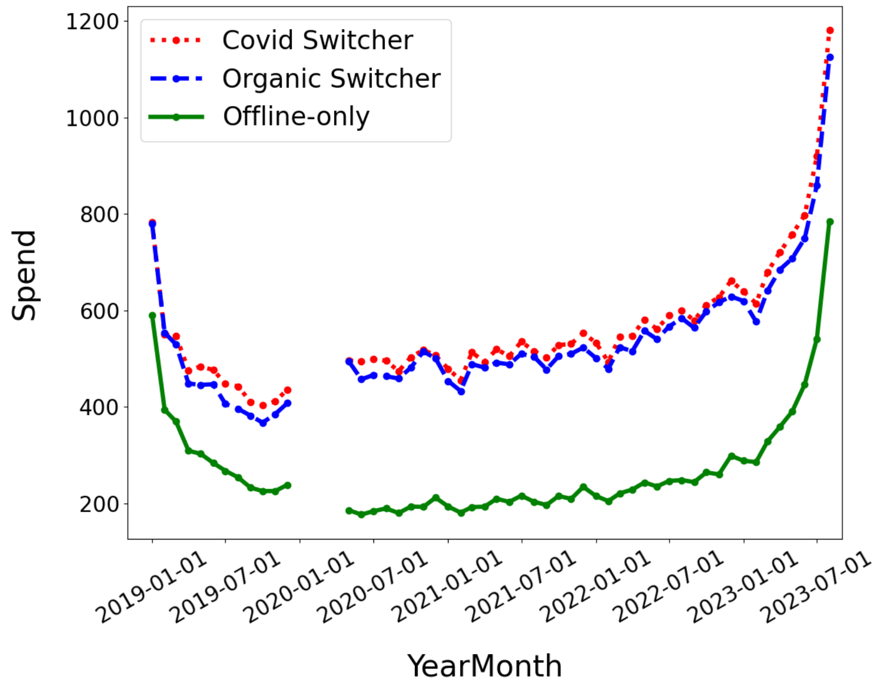
Table 13: Covid vs Organic Switcher Main Results – Other Metrics

DVs:	Quantities	Orders	Unique Items	Unique Brands	Unique Subcategories	Unique Categories
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Covid_Switcher * Post_covid	0.041	-0.027***	0.042	0.040**	0.041***	0.032***
	(0.098)	(0.010)	(0.028)	(0.020)	(0.016)	(0.012)
Fixed-effects						
Customer	Yes	Yes	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	1,874,756	1,874,756	1,874,756	1,874,756	1,874,756	1,874,756
R^2	0.49180	0.37024	0.36696	0.33318	0.31345	0.30243
Within R^2	5.16×10^{-7}	2×10^{-5}	5.39×10^{-6}	9.29×10^{-6}	1.47×10^{-5}	1.54×10^{-5}

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

post-switching, we estimate the same DID model (as shown in Equation (5)) on other metrics of interest (Quantity, Number of orders, Unique items, Brands, Categories, and Subcategories) and present the results in Table 13. Again, we find some small differences between the two groups. Covid switchers tend to place fewer orders but purchase a higher variety of items, brands, and categories, compared to organic switchers in total. This is consistent with their heavier use of the offline channel since consumers who use offline channels generally tend to buy more frequently and also buy a wider variety of products (Chintala et al., 2023).

Figure 5: Test Parallel Pre-trend Assumption: Spend



4.4 Validity and Robustness Checks

We now present a series of tests and robustness checks to establish the validity of our findings.

4.4.1 Parallel Trend Assumption

The validity of the DID approach relies on the parallel trends assumption, i.e., we need to confirm that all three groups followed the same time trend in their spending before 2020. To check if this is indeed the case, we plot the average monthly spend of all three groups in both the pre- and post-treatment periods in Figure 5. For each group, this average monthly spend is calculated by dividing the total spend of customers belonging to the group in a month by the number of customers who have not churned by the end of that month. As we can see, the spending of all three groups evolves parallelly during the pre-switching period from 2019-01-01 to 2019-12-31, and there are no differential trends in spend across the three groups.[18] In addition, we conduct a formal statistical test for the parallel pre-trend using an event study regression; the details of this analysis

[18]Each line in the parallel trend plots shows the average monthly spending for a given group of customers (organic, covid switchers, and offline-only) over the observation period. At any given period t, the averaging is done over the set of active consumers in that period. Both at the beginning and end of the observation period, we have a (relatively) small number of customers in all three groups because of customer acquisition and churn: (1) Not all customers in each group initiate their first purchase at the beginning of the time series (e.g., Jan and Feb 2019), i.e., many customers joined after early 2019, and (2) Customers in all three groups continue to churn in the post-treatment period, and many are no longer in our panel towards the end of the observation period (e.g., July and Aug 2023). Further, the consumers who were with the firm from early 2019 and/or those who did not churn by mid-2023 tend to be loyal consumers (who spend more on average). Since the early and late observations largely consist of such consumers, we observe higher average spending in these time periods. However, the existence of such spikes does not invalidate the parallel trends observed in the pre-treatment period.

and the results are shown in Web Appendix E.1. Finally, we also confirm that all DID analyses for other purchase metrics (as shown in Table 13) satisfy the parallel pre-treatment trend assumptions; see Figure A3 Web Appendix E.2.

4.4.2 Re-defining the Organic Switcher Cohorts

In the main analysis, we defined organic switchers as those offline-only customers who made their first purchase online between 2020-01-01 to 2020-02-29. This definition ensured that the set of customers across the two groups (organic and covid switchers) were largely similar since the time periods of switching were temporally close. Nevertheless, we now provide a robustness check by redefining organic switchers as those customers who were offline-only before 2019-10-31 and made their first online purchase between 2019-11-01 to 2020-02-29. The definition of covid switchers remains the same. Under this new definition, we continue to have 33,246 covid switchers, but also have a larger number of organic switchers (25,958). As before, we exclude the periods used for defining these two cohorts (2019-11-01 to 2020-04-30), and re-run the same set of DID models as in the main analysis, and present the results in Table A11 in Web Appendix E.3. We find that all the results are consistent with the main analysis, suggesting that consumers spend similarly after switching, irrespective of the reason for/timing of becoming omnichannel.

4.4.3 Propensity Score Matching and IPTW Models

In §4.2, we saw that there are some small differences in the demographics of organic and covid switchers. There may also be other pre-treatment confounders that can affect the propensity of an offline-only customer switch organically vs. during covid, such as (1) the length and extent of their experience with the firm, variables such as *Tenure (Days)* that denotes the number of days since the customer's first offline purchase, *CumulativeSpend*, and *CumulativeOrders* that denote the total spend/number of orders since the customer's first offline purchase; and (2) monthly engagement frequency metrics described in §4.2, e.g., *AvgQuantityPerMonth*.

To account for the differences in observables between organic and covid switchers, we estimate a propensity score model via a logistic regression. We label covid switchers as 1 and organic switchers as 0 and model the propensity score of customer i being a covid switcher as:

$$PropensityScore_i = f(X_i^T \theta), \tag{7}$$

where $f(\cdot)$ is the logit link function where $f(X_i^T\theta) = \frac{1}{1+\exp(-X_i^T\theta)}$. $X_i \in \mathcal{R}^{M\times 1}$ is a column vector with M covariates of customer i. These covariates include *Gender*$_i$, *Age*$_i$, *State*$_i$, *Tenure (Days)*$_i$, *CumulativeSpend*$_i$, *AverageSpendPerMonth*$_i$, etc. $\theta \in \mathcal{R}^{M\times 1}$ is a column vector of coefficients. The estimates from this model are shown in Table A12 in the Web Appendix E.4. We find that there are small differences in which types of customers switch organically vs. due to covid on both demographics and pre-switching purchase behaviors (see Figure A4 in the Web Appendix).

Next, we use these propensity scores to augment our DID analysis in two ways.

- First, we perform a customer-level propensity score matching and then re-run the DID analysis. Matching based on propensity score is much more efficient than matching based on a set of covariates (Rosenbaum and Rubin, 1985). Based on the predicted propensity score of each customer i, we use *MatchIt* package in R (Stuart et al., 2011) and perform one-to-one matching to find the nearest neighbor of each covid switcher i from the 12,799 organic switchers (with replacement). In total, this gives us 10,142 organic switchers, and we exclude 2657 organic switchers who were not chosen as the nearest neighbor for any of the covid switchers. We then perform a two-sample t-test on all the demographic and pre-treatment variables and confirm that the two samples are similar on all the variables after matching (see Table A13 in the Web Appendix E.4). We then run the DID analysis on the newly matched sample (see Table A14 in Web Appendix E.4). We find that the results from this exercise are consistent with the main findings.

- Next, we use a slightly different approach to control for any potential selection issues. Specifically, we use an inverse propensity of treatment weight-adjusted (IPTW) DID regression. Here, the sample is the same as the one from the main analysis, but each observation is inversely weighted by its propensity score. The results from this analysis are shown in Table A15 in Web Appendix E.4). Again, we find that the results are quite similar to those in the main section.

In summary, these findings confirm that while there are some differences in which types of consumers choose to switch organically vs. switching due to external forces (covid), conditional on switching, there are no significant differences in *how much* a consumer spends in total with the firm after switching, though those who switch organically are faster at shifting their purchases to the online channel.

4.4.4 Full Panel Analysis

In the main analysis, we constructed the customer-month panel data such that each consumer's panel started in the month of their first purchase and ended in the month of their last observed purchase. While this ensured that we did not augment the data with a lot of zeros (no purchase) for customers who had churned, this may also introduce some problems. For example, if switching to omnichannel, either organically or via shock, affects customers' churn rates, then curtailing the panel length of a customer at her/his last purchase can be problematic. Therefore, we now fix the end date of the panel for all the customers in the data to the end of our observation period. Thus, we extend all the consumers' panels to the end of the observation period (July 2023) and impute zeros for any inactive months between the last month of the customer's purchase and the end of the observational period. We then re-run DID models on this full panel data, and again find that the estimation results are consistent with the main results (see Table A16 in Web Appendix E.5).

4.5 Profitability Analysis

So far, we have shown that even though the reason for switching does not have a significant impact on the magnitude of post-switching spend, it affects the distribution of the spend across the offline-online channel. In particular, we found that covid switchers spend more offline after switching to omnichannel. If the firm's margins across the two channels are different, then even though the revenue from the two groups of consumers is the same, the differences in channel use can lead to a differential impact on profit. Therefore, we now directly examine whether the switch to omnichannel affected profitability for the two groups and how.

In our data, we observe item-level product margins for each purchase starting January 2020. This allows us to directly generate customer-month-profit panel data for each consumer i in the two groups (covid switchers and organic switchers). However, we lack data on the margins for 2019. Therefore, we obtain the margins of each item in a given channel by extrapolating the item-channel level margins from 2020–2023 to the 2019 timeframe based on Brazil's Extended National Consumer Price Index (IPCA). We refer interested readers to Web Appendix §F.1 for details.

We now provide some summary statistics of the differences in prices, costs, and margins across the two channels. First, to highlight the differences in these three variables across channels, we generate a panel of prices and costs at the item-channel-month level. For instance, in the case of price, an observation in this panel is the average price of an item j sold through a channel c in time period t and is calculated by the total sales of item j in channel c in month t divided by the total quantity of that item sold in that channel-month. Next, we calculate the offline-online differences at the item-month level to obtain item-specific offline-online differences in price, costs, and margin. Based on these three panels, Table 14 provides item-level summary statistics for the 22,803 items that were sold in both channels in our data (and were purchased at least once in the same month). For confidentiality, we multiply costs and margins by the same undisclosed multiplier used on price.

We now make a few important observations based on this table. First, on average, offline prices are higher than online prices, and this difference can be quite significant for a large portion of item-months. Second, differences in costs across the two channels are relatively small; the median difference is close to zero. Finally, margins are higher in the offline channel compared to the online channel. Overall, these descriptives suggest that offline prices tend to be higher than online prices, and costs are largely similar. Thus, consumers who spend more offline may be more profitable (assuming they buy similar assortments of products).

Table 14: Item-Level Offline Price, Cost and Margin Relative to Online

	mean	std	25%	50%	75%	(min, max)	count
Offline - Online Price	6.01	91.07	-0.84	4.61	17.71	(-7,598.46,1,430.91)	22,803
Offline - Online Cost	-1.57	24.08	-1.16	-0.07	0.02	(-2,030.71, 1,136.50)	22,803
Offline - Online Margin	5.37	78.49	-1.77	3.47	14.05	(-6,586.62, 2,163.31)	22,803

To that end, we now estimate the DID models in Equations (3)–(4), and present the results in Table 15.[19] As we can see, both groups became more profitable post-switching/post-covid, i.e., when consumers become omnichannel, they are more profitable irrespective of the reason for switching to omnichannel. However, we also see that covid switchers are more profitable than organic switchers post-switching, even after controlling for customer- and time-period fixed effects. We show that these results are robust even if we consider just those items that are available in both channels in a given month (see Table A17 in Web Appendix F.2).

This suggests that, when encouraging consumers to switch to omnichannel, retailers should consider not just the incremental revenue, but also incremental profitability. Further, this incremental profitability can be

[19]Figure A5 in Web Appendix F.2 shows that the parallel trends assumption holds for margins.

Table 15: Difference-in-Difference Margin Result (Covid vs. Organic Switchers)

Dependent Variable:	Total Margin	
Model:	(1)	(2)
Variables		
Constant	145.7^{***}	
	(1.70)	
Covid_Switcher	11.0^{***}	
	(2.01)	
Post_Covid	26.0^{***}	
	(1.81)	
Covid_Switcher * Post_covid	8.14^{***}	9.64^{***}
	(2.16)	(2.28)
Fixed-effects		
Customer		Yes
YearMonth		Yes
Fit statistics		
Observations	1,874,756	1,874,756
R^2	0.00178	0.30436
Within R^2		2.98×10^{-5}

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

driven by *which* channels the incremental spend comes from, and as such can affect the retailer's calculus on whether and how much to incentivize consumers to switch to omnichannel. An important managerial implication is that the firm should not naively assume that users who switch due to external incentives (in our case COVID-19, but also could be incentives provided by the firm) will be as profitable as organic switchers. Indeed, since users who are goaded to switch to omnichannel due to external reasons are slower to adopt online shopping, they can be more or less profitable depending on the profitability of the two channels.

5 Discussion and Conclusions

In conclusion, our study examines the impact of channel choice on customer value using data from a large pet supplies retailer in Brazil spanning the period of 2019-2023. We use the onset of COVID-19 as an exogenous shock that affected offline-only consumers' decision to become omnichannel by starting to (also) shop online.

Firstly, we establish that consumers who transition to omnichannel purchasing, whether by their own choice or due to external stimuli like the COVID-19 pandemic, exhibit similar increases in spending post-switching. This suggests that, regardless of the motivation behind the transition, the incremental revenue generated from omnichannel consumers remains consistent. Nevertheless, our analysis reveals differences in channel preferences between organic switchers and those prompted by external factors such as COVID-19. While both groups increase their spending after becoming omnichannel, the rate of adoption of online channels differs, with covid switchers displaying slower migration to the online channel. This divergence in channel utilization has implications for profitability, with covid switchers proving to be more profitable post-transition, due to their slower uptake of online channels (since the offline channel is more profitable in this specific case). An important managerial implication is that while firms may be able to use the incremental revenue from its organic switchers as a reasonable proxy for the expected revenue from users who switch due to external prodding, it should not naively assume that they would be equally profitable. Indeed, we see that the reasons for switching continue to have small but meaningful impact on channel choices after switching, which can affect profits. Thus, firms should take such persistent differences into account when designing incentives for consumers to become omnichannel.

Funding and Competing Interests

The authors were not paid by the data provider. A preliminary version of this manuscript was reviewed by the data provider to ensure that data was accurately depicted, but they did not review/comment on the content of the manuscript.

References

Joshua D Angrist and Guido W Imbens. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association*, 90(430):431–442, 1995.

Asim Ansari, Carl F Mela, and Scott A Neslin. Customer channel migration. *Journal of Marketing Research*, 45(1):60–76, 2008.

Jill Avery, Thomas J Steenburgh, John Deighton, and Mary Caravella. Adding bricks to clicks: Predicting the patterns of cross-channel elasticities over time. *Journal of Marketing*, 76(3):96–111, 2012.

David R Bell, Santiago Gallino, and Antonio Moreno. Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Science*, 64(4):1629–1651, 2018.

Haiqiang Chen, Wenlan Qian, and Qiang Wen. The impact of the covid-19 pandemic on consumption: Learning from high-frequency transaction data. In *AEA Papers and Proceedings*, volume 111, pages 307–311. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 2021.

Sai Chand Chintala, Jūra Liaukonytė, and Nathan Yang. Browsing the aisles or browsing the app? how online grocery shopping is changing what we buy. *Marketing Science*, 2023.

Tony Haitao Cui, Anindya Ghose, Hanna Halaburda, Raghuram Iyengar, Koen Pauwels, S Sriram, Catherine Tucker, and Sriraman Venkataraman. Informational challenges in omnichannel marketing: Remedies and future research. *Journal of Marketing*, 85(1):103–120, 2021.

Edlaine Faria de Moura Villela, Rossana Verónica Mendoza López, Ana Paula Sayuri Sato, Fábio Morato de Oliveira, Eliseu Alves Waldman, Rafael Van den Bergh, Joseph Nelson Siewe Fodjo, and Robert Colebunders. Covid-19 outbreak in brazil: adherence to national preventive measures and impact on people's lives, an online survey. *BMC Public Health*, 21(1):1–10, 2021.

Chris Forman, Anindya Ghose, and Avi Goldfarb. Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management science*, 55(1):47–57, 2009.

Elina H Hwang, Leela Nageswaran, and Soo-Haeng Cho. Impact of covid-19 on omnichannel retail: Drivers of online sales during pandemic. *Available at SSRN 3657827*, 2020.

Vipin Kumar and Rajkumar Venkatesan. Who are the multichannel shoppers and how do they perform?: Correlates of multichannel shopping behavior. *Journal of Interactive Marketing*, 19(2):44–62, 2005.

Tarun Kushwaha and Venkatesh Shankar. Are multichannel customers really more valuable? the moderating role of product category characteristics. *Journal of Marketing*, 77(4):67–85, 2013.

Hui Li. Are e-books a different channel? multichannel management of digital products. *Quantitative Marketing and Economics*, 19(2):179–225, 2021.

Simha Mummalaneni, Hema Yoganarasimhan, and Varad Pathak. How do content producers respond to engagement on social media platforms? 2023.

Unnati Narang and Venkatesh Shankar. Mobile app introduction and online and offline purchases and product returns. *Marketing Science*, 38(5):756–772, 2019.

Scott A Neslin and Venkatesh Shankar. Key issues in multichannel customer management: current knowledge and future directions. *Journal of Interactive Marketing*, 23(1):70–81, 2009.

Shin Oblander and Daniel McCarthy. Estimating the long-term impact of major events on consumption patterns: Evidence from covid-19. *Forthcoming at Marketing Science*, 2022.

Paul R Rosenbaum and Donald B Rubin. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1):33–38, 1985.

Scott K Shriver and Bryan Bollinger. Demand expansion and cannibalization effects from retail store entry: A structural analysis of multichannel demand. *Management Science*, 68(12):8829–8856, 2022.

Elizabeth A Stuart, Gary King, Kosuke Imai, and Daniel Ho. Matchit: nonparametric preprocessing for parametric causal inference. *Journal of statistical software*, 2011.

Vasilis Syrgkanis, Victor Lei, Miruna Oprescu, Maggie Hei, Keith Battocchi, and Greg Lewis. Machine learning estimation of heterogeneous treatment effects with instruments. *Advances in Neural Information Processing Systems*, 32, 2019.

Rajkumar Venkatesan, Vipin Kumar, and Nalini Ravishanker. Multichannel shopping: causes and consequences. *Journal of Marketing*, 71(2):114–132, 2007.

Peter C Verhoef, Pallassana K Kannan, and J Jeffrey Inman. From multi-channel retailing to omni-channel retailing: introduction to the special issue on multi-channel retailing. *Journal of retailing*, 91(2):174–181, 2015.

Kitty Wang and Avi Goldfarb. Can offline stores drive online sales? *Journal of Marketing Research*, 54(5): 706–719, 2017.

Rebecca Jen-Hui Wang, Edward C Malthouse, and Lakshman Krishnamurthi. On the go: How mobile shopping affects customer purchase behavior. *Journal of Retailing*, 91(2):217–234, 2015.

Kaiquan Xu, Jason Chan, Anindya Ghose, and Sang Pil Han. Battle of the channels: The impact of tablets on digital commerce. *Management Science*, 63(5):1469–1492, 2017.

Botao Yang and Andrew T Ching. Dynamics of consumer adoption of financial innovation: The case of atm cards. *Management Science*, 60(4):903–922, 2014.

Jie Zhang, Paul W Farris, John W Irvin, Tarun Kushwaha, Thomas J Steenburgh, and Barton A Weitz. Crafting integrated multichannel retailing strategies. *Journal of interactive marketing*, 24(2):168–180, 2010.

Web Appendix

A Data cleaning

Table A1: Item categories with substantially different availability time via online and offline channels (YY-MM-DD)

Item Category	First Offline Purchase Date	First Online Purchase Date
Large size food	2020-09-12	2020-02-28
Human food	2019-01-02	2021-02-12
Large size pharmacy	2021-02-03	2020-12-09
Pet Cleaning Supplies	2019-01-03	2022-03-13

There are four product categories with significant discrepancies in their initial availability time on the offline and online channels. For example, the product category *Large size food* (e.g. pack size greater than 10KG) was available online in February 2020 but only became available offline in September 2020. In addition, the product category Pet Cleaning Supplies was only available offline before COVID-19 but became available online in August 2022.

Including these food categories in our analysis can be problematic since it also reflects retailers' assortment and the channel-specific effects (i.e. consumers can only buy some products exclusively from one channel). Therefore, we exclude these four categories from our analysis. The fraction of sales of these four categories only accounts for 0.63% of the total revenue during the sample period, and as such unlikely to have any meaningful impact on our findings.

B Additional Summary Statistics

Table A2: Customer-Month Summary Statistics - Pre-Period (2019-01 to 2019-12)

Variables	mean	std	min	25%	50%	75%	max	count
Transactions across both Channels								
Spend	329.24	1387.46	0	0	31.40	454.86	2340617.16	9,345,799
Quantity	3.38	25.05	0	0	1.00	4.00	44158.00	9,345,799
Order	0.77	1.05	0	0	1.00	1.00	155.00	9,345,799
UniqueItem	1.99	4.04	0	0	1.00	3.00	4191.00	9,345,799
UniqueBrand	1.60	2.40	0	0	1.00	2.00	332.00	9,345,799
UniqueSubcategory	1.42	1.98	0	0	1.00	2.00	64.00	9,345,799
UniqueCategory	1.20	1.54	0	0	1.00	2.00	18.00	9,345,799
Transactions in Offline Channel								
Spend	297.04	1375.49	0	0	0	393.33	2340617.16	9,345,799
Quantity	3.14	24.88	0	0	0	3.00	44158.00	9,345,799
Orders	0.72	1.03	0	0	0	1.00	155.00	9,345,799
UniqueItems	1.89	4.02	0	0	0	3.00	4191.00	9,345,799
UniqueBrands	1.52	2.38	0	0	0	2.00	332.00	9,345,799
UniqueSubcategories	1.34	1.97	0	0	0	2.00	64.00	9,345,799
UniqueCategories	1.13	1.54	0	0	0	2.00	18.00	9,345,799
Transactions in Online Channel								
Spend	32.20	204.41	0	0	0	0	47187.28	9,345,799
Quantity	0.23	2.87	0	0	0	0	799.00	9,345,799
Orders	0.05	0.27	0	0	0	0	18.00	9,345,799
UniqueItems	0.10	0.65	0	0	0	0	68.00	9,345,799
UniqueBrands	0.09	0.50	0	0	0	0	30.00	9,345,799
UniqueSubcategories	0.08	0.45	0	0	0	0	16.00	9,345,799
UniqueCategories	0.07	0.38	0	0	0	0	11.00	9,345,799

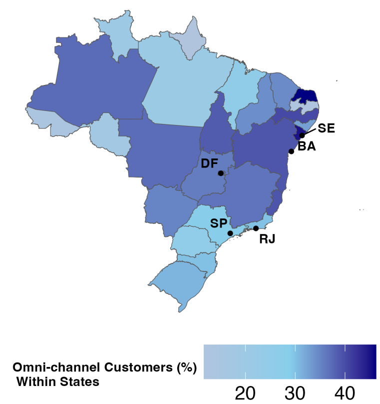
Table A3: Customer-Month Summary Statistics - Post-Period (2020-05 to 2023-08)

Variables	mean	std	min	25%	50%	75%	max	count
				Transactions across both Channels				
Spend	336.55	712.65	0	0	0	447.96	301134.12	80,171,932
Quantity	2.69	8.46	0	0	0	2.00	3440.00	80,171,932
Order	0.67	1.06	0	0	0	1.00	430.00	80,171,932
UniqueItem	1.56	2.96	0	0	0	2.00	649.00	80,171,932
UniqueBrand	1.26	2.16	0	0	0	2.00	194.00	80,171,932
UniqueSubcategory	1.09	1.75	0	0	0	2.00	48.00	80,171,932
UniqueCategory	0.93	1.37	0	0	0	2.00	17.00	80,171,932
UniqueVisitedStores	0.33	0.51	0	0	0	1.00	21.00	80,171,932
				Transactions in Offline Channel				
Spend	212.23	555.43	0	0	0	167.97	163070.60	80,171,932
Quantity	1.91	6.58	0	0	0	1.00	3239.00	80,171,932
Orders	0.46	0.88	0	0	0	1.00	430.00	80,171,932
UniqueItems	1.20	2.73	0	0	0	1.00	649.00	80,171,932
UniqueBrands	0.96	2.00	0	0	0	1.00	194.00	80,171,932
UniqueSubcategories	0.83	1.63	0	0	0	1.00	48.00	80,171,932
UniqueCategories	0.70	1.29	0	0	0	1.00	17.00	80,171,932
				Transactions in Online Channel				
Spend	124.32	450.03	0	0	0	0	276399.66	80,171,932
Quantity	0.78	5.10	0	0	0	0	3440.00	80,171,932
Orders	0.21	0.59	0	0	0	0	130.00	80,171,932
UniqueItems	0.37	1.23	0	0	0	0	141.00	80,171,932
UniqueBrands	0.31	0.95	0	0	0	0	44.00	80,171,932
UniqueSubcategories	0.28	0.81	0	0	0	0	22.00	80,171,932
UniqueCategories	0.25	0.68	0	0	0	0	13.00	80,171,932

C Appendix to Descriptive Analysis

In addition to income, age, and gender, a consumer's geographic location is also correlated with their channel choice. Figure A1 shows the fraction of omnichannel consumers relative to total customers for each state as of 2023-08-31 (the last day of the observation period). We see that the fraction of omnichannel consumers of a state ranges from 11.7% to 46.2%, with the mean and median being 31% and 34.3%, respectively.

Figure A1: Fraction of omnichannel customers to total customers within states



Tables A4 and A5 show the cross-sectional and within-consumer analysis comparing consumers' monthly purchase behaviors in terms of volume and variety.

Table A4: Cross-sectional Comparisons of omnichannel and Offline-only relative to Online-only customers

DVs:	Quantities	Orders	Unique Items	Unique Brands	Unique Subcategories	Unique Categories
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
omnichannel	1.558***	0.253***	0.901***	0.674***	0.546***	0.413***
	(0.011)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
Offline-Only	0.220***	-0.024***	0.390***	0.282***	0.225***	0.152***
	(0.011)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
Control Variables						
Age	Yes	Yes	Yes	Yes	Yes	Yes
Gender	Yes	Yes	Yes	Yes	Yes	Yes
Zipcode	Yes	Yes	Yes	Yes	Yes	Yes
HouseholdIncome	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-effects						
YearMonth	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	94,415,101	94,415,101	94,415,101	94,415,101	94,415,101	94,415,101
R^2	0.031	0.072	0.052	0.056	0.058	0.061
Within R^2	0.003	0.015	0.008	0.010	0.010	0.010

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

Table A5: Customers' Purchase Behaviors Before/After Switching to omnichannel

DVs:	Quantities	Orders	Unique Items	Unique Brands	Unique Subcategories	Unique Categories
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Post_Switch	0.941***	0.283***	0.237***	0.184***	0.156***	0.154***
	(0.012)	(0.001)	(0.004)	(0.003)	(0.002)	(0.002)
Fixed-effects						
Customer	Yes	Yes	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	27,239,397	27,239,397	27,239,397	27,239,397	27,239,397	27,239,397
R^2	0.47335	0.38419	0.36149	0.33476	0.31787	0.31193
Within R^2	0.00119	0.00704	0.00057	0.00064	0.00070	0.00113

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

Table A6: Customer-Month Summary Statistics By Customer Types - Pre-Period (2019-01 to 2019-12)

Variables	Mean	Std	Min	25%	50%	75%	Max	Count
				Omnichannel Customers				
Spend	403.91	718.01	0	0	89.14	586.28	112782.01	3,656,069
Quantity	4.08	10.90	0	0	1.00	4.00	2049.00	3,656,069
Order	0.86	1.13	0	0	1.00	1.00	155.00	3,656,069
UniqueItem	2.30	3.65	0	0	1.00	3.00	355.00	3,656,069
UniqueBrand	1.83	2.64	0	0	1.00	3.00	127.00	3,656,069
UniqueSubcategory	1.61	2.16	0	0	1.00	3.00	40.00	3,656,069
UniqueCategory	1.34	1.66	0	0	1.00	2.00	16.00	3,656,069
				Offline-only Customers				
Spend	281.09	1712.82	0	0	0	376.53	2340617.16	5,457,929
Quantity	2.97	31.50	0	0	0	3.00	44158.00	5,457,929
Order	0.73	1.01	0	0	0	1.00	152.00	5,457,929
UniqueItem	1.83	4.34	0	0	0	3.00	4191.00	5,457,929
UniqueBrand	1.47	2.25	0	0	0	2.00	332.00	5,457,929
UniqueSubcategory	1.32	1.87	0	0	0	2.00	64.00	5,457,929
UniqueCategory	1.12	1.47	0	0	0	2.00	18.00	5,457,929
				Online-only Customers				
Spend	285.16	510.89	0	0	0	473.76	44832.55	231,801
Quantity	1.87	6.98	0	0	0	2.00	495.00	231,801
Order	0.48	0.59	0	0	0	1.00	10.00	231,801
UniqueItem	0.89	1.57	0	0	0	1.00	42.00	231,801
UniqueBrand	0.75	1.19	0	0	0	1.00	21.00	231,801
UniqueSubcategory	0.70	1.05	0	0	0	1.00	14.00	231,801
UniqueCategory	0.64	0.89	0	0	0	1.00	10.00	231,801

Table A7: Customer-Month Summary Statistics By Customer Types - Post-Period (2020-05 to 2023-08)

Variables	Mean	Std	Min	25%	50%	75%	Max	Count
				Omnichannel Customers				
Spend	431.31	833.50	0	0	0	602.63	301134.12	38,500,906
Quantity	3.40	10.10	0	0	0	3.00	3440.00	38,500,906
Order	0.82	1.20	0	0	0	1.00	236.00	38,500,906
UniqueItem	1.87	3.28	0	0	0	3.00	322.00	38,500,906
UniqueBrand	1.49	2.37	0	0	0	2.00	114.00	38,500,906
UniqueSubcategory	1.29	1.90	0	0	0	2.00	39.00	38,500,906
UniqueCategory	1.08	1.48	0	0	0	2.00	15.00	38,500,906
				Offline-only Customers				
Spend	242.73	563.99	0	0	0	272.93	109274.30	34,797,896
Quantity	2.12	6.49	0	0	0	2.00	1174.00	34,797,896
Order	0.54	0.92	0	0	0	1.00	430.00	34,797,896
UniqueItem	1.36	2.74	0	0	0	2.00	649.00	34,797,896
UniqueBrand	1.10	2.01	0	0	0	2.00	194.00	34,797,896
UniqueSubcategory	0.96	1.66	0	0	0	1.00	48.00	34,797,896
UniqueCategory	0.82	1.32	0	0	0	1.00	17.00	34,797,896
				Online-only Customers				
Spend	280.78	569.29	0	0	0	437.39	110682.95	6,873,130
Quantity	1.60	6.72	0	0	0	1.00	3370.00	6,873,130
Order	0.51	0.69	0	0	0	1.00	130.00	6,873,130
UniqueItem	0.86	1.55	0	0	0	1.00	141.00	6,873,130
UniqueBrand	0.74	1.19	0	0	0	1.00	32.00	6,873,130
UniqueSubcategory	0.67	1.01	0	0	0	1.00	20.00	6,873,130
UniqueCategory	0.61	0.87	0	0	0	1.00	11.00	6,873,130

D Appendix to DID Analysis

Table A8: Fraction of Offline Spend over Month - Organic vs Covid Switchers

Dependent Variable:	Fraction of Offline Spend
Model:	(1)
Variables	
Constant	0.4854^{***}
	(0.0079)
Covid_Switcher	0.0830^{***}
	(0.0112)
t	-0.0021^{***}
	(0.0003)
t_*_Covid_Switcher	0.0010^{**}
	(0.0005)
Fit statistics	
Observations	80
R^2	0.83067
Adjusted R^2	0.82398

*Signif. Codes: ***: 0.01, **: 0.05*

Table A9: Heterogeneous Effect by Age: Fraction of Offline Spend over Month - Organic vs Covid Switchers

Dependent Variable:	Fraction of Offline Spend	
Model:	(1) Senior Switchers	(2) Young Switchers
Variables		
Constant	0.4981^{***}	0.4739^{***}
	(0.0087)	(0.0076)
Covid_Switcher	0.0729^{***}	0.0904^{***}
	(0.0123)	(0.0108)
t	-0.0022^{***}	-0.0023^{***}
	(0.0004)	(0.0003)
t_*_Covid_Switcher	0.0018^{***}	0.0002
	(0.0005)	(0.0005)
Fit statistics		
Observations	80	80
R^2	0.81312	0.83427
Adjusted R^2	0.80574	0.82773

*Signif. Codes: ***: 0.01, **: 0.05*

Table A10: Covid vs Organic Switcher Main Results - Offline Metrics

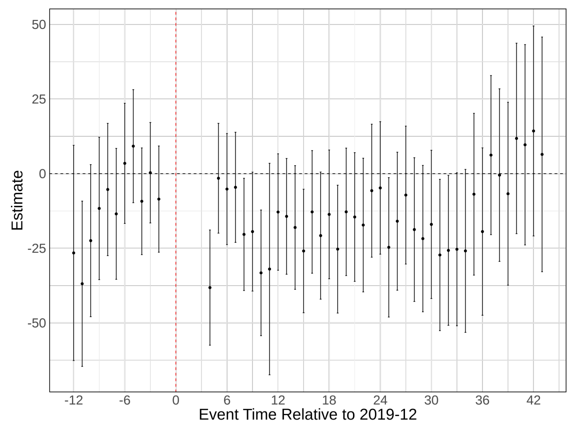
DVs:	Quantities	Orders	Unique Items	Unique Brands	Unique Subcategories	Unique Categories
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Covid_Switcher * Post_covid	-0.015	0.033^{***}	0.096^{***}	0.091^{***}	0.078^{***}	0.062^{***}
	(0.083)	(0.008)	(0.027)	(0.019)	(0.015)	(0.012)
Fixed-effects						
Customer	Yes	Yes	Yes	Yes	Yes	Yes
YearMonth	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	1,874,756	1,874,756	1,874,756	1,874,756	1,874,756	1,874,756
R^2	0.44630	0.36109	0.34530	0.32087	0.30775	0.30325
Within R^2	1.07×10^{-7}	4.65×10^{-5}	3.2×10^{-5}	5.49×10^{-5}	6.08×10^{-5}	6.43×10^{-5}

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

E Appendix for Validity and Robustness Checks

E.1 Statistical Test for Parallel Trend

Figure A2: Parallel Pre-Trend Assumption - Total Spend



We now provide a formal statistical test for the parallel trend assumption for the total spend outcome using an event study regression (i.e., leads-lags relative time regression). We incorporate a series of period dummy variables into the model to decompose the pre-treatment periods. Specifically, we estimate the relative-time model specified below:

$$\begin{aligned} y_{it} = Customer_i + YearMonth_t + \sum_j \theta_j (PRE_{it}(j) \cdot \mathbb{I}\{Covid\_switcher_i\}) \\ + \sum_k \theta_k (POST_{it}(k) \cdot \mathbb{I}\{Covid\_switcher_i\}) + \epsilon_{it} \end{aligned} \tag{A1}$$

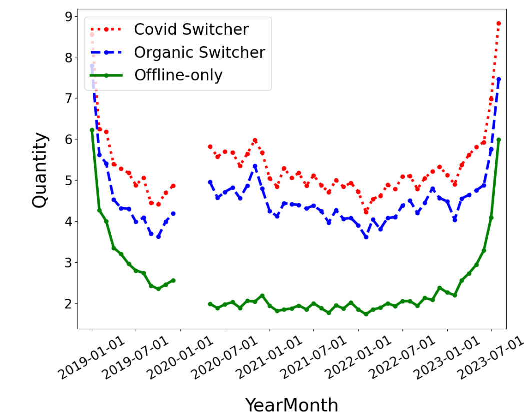
In this model, the interaction term $\sum_j \theta_j (PRE_{it}(j) \cdot \mathbb{I}\{Covid\_switcher_i\})$ allows us to examine if there are any differential trends across the two groups before the treatment (i.e. switching to online channel). We label the year-month period from January 2019 to August 2023 with ascending integers from 1 to 56, excluding the periods used for defining organic and covid switchers between January 2020 and April 2020 (periods 13 to 16) from the estimation. Therefore, the pre-treatment period ranges from 1 and 12, and the post-treatment period ranges from 17 and 56. $PRE_{it}(j)$ is an indicator that equals 1 if period t is j months prior to December 2019 (period 12). Hence, the coefficient θ_j for $j = -J, -(J-1), \ldots, -1$ captures the pre-treatment trend in shopping behaviors between covid switchers and organic switchers. Similarly, $POST_{it}(k)$ is an indicator that equals 1 if period t is k months after May 2020 (period 17), and θ_k captures the dynamic treatment effect. The pre-treatment parallel trend assumption is valid only if θ_js are not significantly different from 0. We set period 12 (i.e., December 2019) as the reference period (i.e., normalize the coefficient for December 2019 to zero) and consider a model with full lags and leads. The estimated coefficients and their 95% confidence intervals are presented in Figure A2. As we can see, the estimated coefficients are not significantly different from zero, i.e., the parallel pre-trend assumption seems valid. [20] This suggests that the DiD estimation of treatment effects was not falsely inflated by any trends prior to the treatment.

[20] All pre-treatment coefficients are insignificantly from zero except for lag 11. For the parallel trend assumption to hold, it is permissible for a small number of pre-treatment coefficients to be statistically significant (provided the majority of these coefficients are statistically insignificant), see Gopalan et al. (2020) and Chapter 5 of Angrist and Pischke (2008).

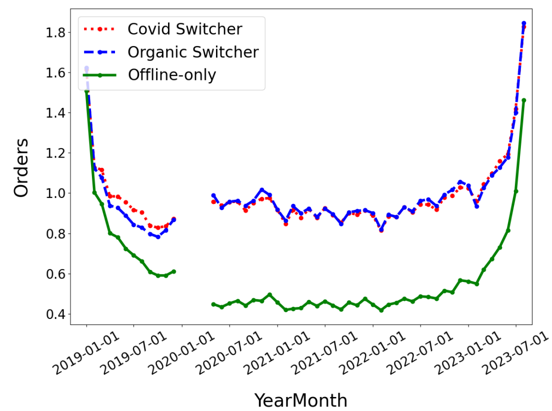
E.2 Additional Parallel Pre-Trend Plots

Figure A3: Parallel Pre-Trend Assumption

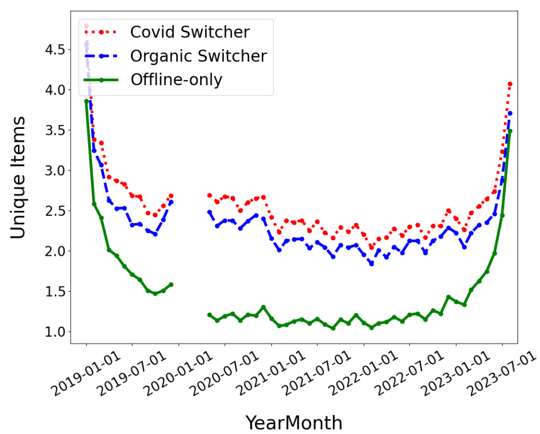
(a) Quantity



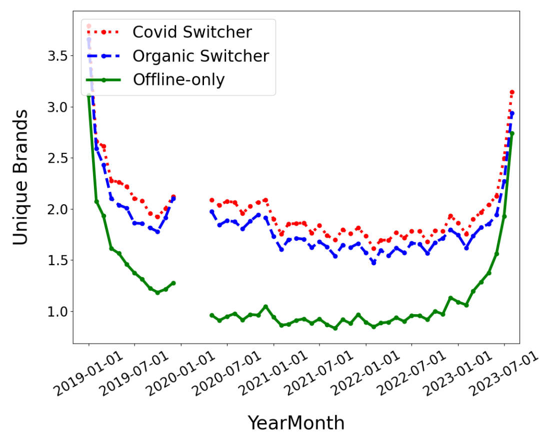
(b) Orders



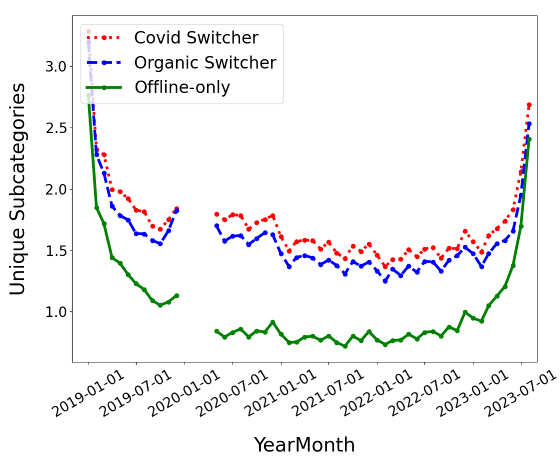
(c) Unique Items



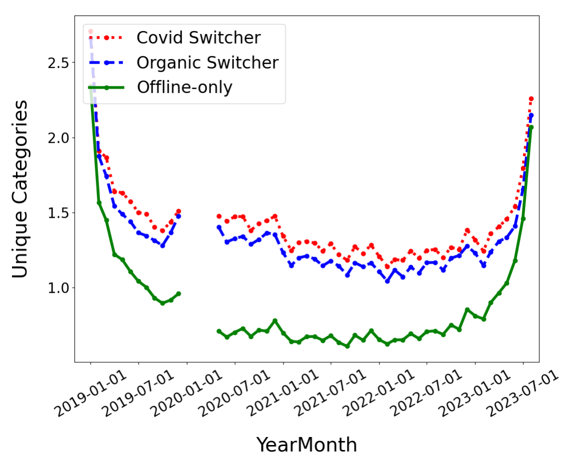
(d) Unique Brands



(e) Unique Subcategories



(f) Unique Categories



E.3 Re-defining the Organic Switcher Group

Table A11: Covid vs Organic Switcher (Redefining Organic Switchers + DID) - Spend

Dependent Variables:	Total Spend		Offline Spend		Fraction of Offline Spend	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	450.4^{***}		450.4^{***}		1.00^{***}	
	(3.92)		(3.92)		(2.87×10^{-16})	
isCovidConvert	27.5^{***}		27.5^{***}		0.00	
	(5.29)		(5.29)		(0.00)	
Post_Covid	86.8^{***}		-208.4^{***}		-0.501^{***}	
	(4.44)		(3.45)		(0.002)	
Covid_Switcher * Post_covid	-3.05	10.7	37.7^{***}	46.9^{***}	0.070^{***}	0.060^{***}
	(5.69)	(6.17)	(4.63)	(4.80)	(0.003)	(0.003)
Fixed-effects						
Customer		Yes		Yes		Yes
YearMonth		Yes		Yes		Yes
Fit statistics						
Observations	2,294,367	2,294,367	2,294,367	2,294,367	1,237,640	1,237,640
R^2	0.00114	0.38872	0.01141	0.36269	0.15420	0.50300
Within R^2		5.73×10^{-6}		0.00021		0.00107

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

E.4 DiD Models with Propensity Score Matching and Inverse Propensity Score Weighting

Figure A4: Propensity Score Model Predictions of Covid and Organic Switchers

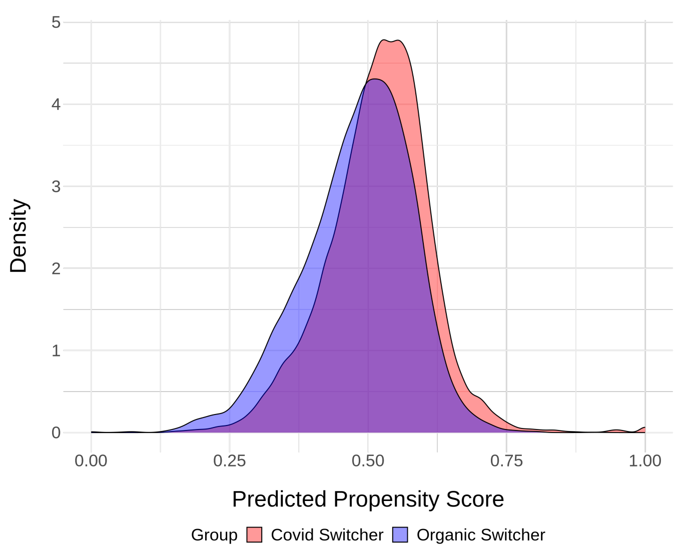


Table A12: Propensity Score Logit Model - Organic vs. Covid Switcher

	Dependent variable:
	Covid_Switcher
Gender Missing	-0.032
	(0.027)
Gender Female	0.043^{**}
	(0.018)
Household Income ($<$ 1000)	-0.150^{***}
	(0.023)
Household Income ($\geq$ 2000)	0.144^{***}
	(0.020)
Household Income Missing	0.107^{***}
	(0.035)
Tenure (Days)	0.001^{***}
	(0.0001)
CumulativeSpend	-0.00001
	(0.00001)
CumulativeOrders	0.012^{***}
	(0.003)
AvgSpendPerMonth	-0.0002
	(0.0001)
AvgQuantityPerMonth	0.007^{***}
	(0.002)
AvgOrdersPerMonth	-0.201^{***}
	(0.029)
AvgUniqueItemsPerMonth	0.037^{***}
	(0.013)
AvgUniqueBrandsPerMonth	-0.058^{***}
	(0.022)
AvgUniqueSubcategoriesPerMonth	-0.017
	(0.025)
AvgUniqueCategoriesPerMonth	0.092^{***}
	(0.026)
AvgUniqueVisitedStoresPerMonth	0.107^{***}
	(0.041)
Age	Yes
State	Yes
Observations	46,045
Log Likelihood	$-48{,}258.240$
Akaike Inf. Crit.	96,762.480
Note:	***p<0.01; **p<0.05

AvgUniqueVisitedStoresPerMonth measures the average number of distinct stores a customer visits per month before 2019-12-31. The coefficients of *Age* and *State* dummy variables are not disclosed due to a large number of categories.

Table A13: Customer-level Summary Statistics of Offline Purchase Behaviours Variables: Re-weighted Sample Based on Inverse Propensity Score. Columns 2 and 3 show the weighted mean of variables of covid and organic switchers with weights being the inverse of propensity.

Variable	Covid Switchers mean	Organic Switchers mean	Diff	T-stats	P-value
Tenure (Days)	230.63	230.47	0.16	0.12	0.90
CumulativeSpend	3611.46	3613.91	-2.46	-0.04	0.96
CumulativeOrders	7.42	7.40	0.02	0.20	0.84
AvgSpendPerMonth	459.13	459.01	0.13	0.02	0.98
AvgQuantityPerMonth	4.72	4.78	-0.06	-0.58	0.56
AvgOrdersPerMonth	0.95	0.95	0.00	0.24	0.81
AvgUniqueItemsPerMonth	2.90	2.89	0.01	0.20	0.84
AvgUniqueBrandsPerMonth	2.33	2.32	0.00	0.14	0.89
AvgUniqueSubcategoriesPerMonth	2.03	2.03	0.00	0.14	0.89
AvgUniqueCategoriesPerMonth	1.65	1.65	0.00	0.12	0.90

Table A14: Covid vs Organic Switcher (Propensity Score Matching + DID) - Spend

Dependent Variables:	Total Spend		Offline Spend		Fraction of Offline Spend	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	440.0***		440.0***		1.00***	
	(6.04)		(6.04)		(3.04×10^{-15})	
Covid_Switcher	25.6***		25.6***		0.00	
	(6.90)		(6.90)		(0.00)	
Post_Covid	108.6***		-196.3***		-0.514***	
	(6.18)		(5.07)		(0.004)	
Covid_Switcher * Post_covid	-12.5*	-7.41	38.0***	40.6***	0.082***	0.073***
	(7.04)	(7.13)	(5.82)	(5.91)	(0.004)	(0.005)
Fixed-effects						
Customer		Yes		Yes		Yes
YearMonth		Yes		Yes		Yes
Fit statistics						
Observations	1,772,310	1,772,310	1,772,310	1,772,310	962,275	962,275
R^2	0.00180	0.39934	0.00970	0.36450	0.17164	0.50065
Within R^2		2.57×10^{-6}		0.00012		0.00128

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

Table A15: Covid vs. Organic Switcher (IPTW + DID) - Spend

Dependent Variables:	Total Spend		Offline Spend		Fraction of Offline Spend	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	451.2^{***}		451.2^{***}		1.00^{***}	
	(6.57)		(6.57)		(1.08×10^{-14})	
Covid_Switcher	-0.487		-0.487		0.00	
	(7.28)		(7.28)		(0.00)	
Post_Covid	104.5^{***}		-196.8^{***}		-0.505^{***}	
	(5.85)		(5.01)		(0.004)	
Covid_Switcher * Post_covid	-8.15	-8.07	38.9^{***}	38.5^{***}	0.064^{***}	0.058^{***}
	(6.77)	(7.02)	(5.72)	(5.98)	(0.004)	(0.004)
Fixed-effects						
Customer		Yes		Yes		Yes
YearMonth		Yes		Yes		Yes
Fit statistics						
Observations	1,874,756	1,874,756	1,874,756	1,874,756	1,016,120	1,016,120
R^2	0.00177	0.39540	0.00932	0.36357	0.17095	0.50145
Within R^2		3.56×10^{-6}		0.00013		0.00094

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

E.5 Full Panel Data - Analysis

Table A16: Covid vs Organic Switcher (Full Panel Analysis) - Spend

Dependent Variables:	Total Spend		Offline Spend		Fraction of Offline Spend	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Constant	432.8***		432.8***		1.00***	
	(5.32)		(5.32)		(4.3×10^{-15})	
Covid_Switcher	32.9***		32.9***		0.00	
	(6.28)		(6.28)		(0.00)	
Post_Covid	-1.83		-241.3***		-0.514***	
	(5.04)		(4.42)		(0.003)	
Covid_Switcher * Post_covid	-5.09	0.729	26.6***	32.3***	0.083***	0.071***
	(5.92)	(6.19)	(5.21)	(5.07)	(0.004)	(0.004)
Fixed-effects						
Customer		Yes		Yes		Yes
YearMonth		Yes		Yes		Yes
Fit statistics						
Observations	2,220,198	2,220,198	2,220,198	2,220,198	1,016,120	1,016,120
R^2	0.00021	0.39347	0.01691	0.35617	0.17300	0.50164
Within R^2		2.87×10^{-8}		9.3×10^{-5}		0.00136

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

F Profitability Analysis

F.1 Missing Value Imputation

As discussed in the main text, we have data on prices, margins, and costs for each item sold from January 2020 to August 2023, but we lack cost data for 2019 because of issues with the data provider's upgrade to a new data management platform in 2020. Thus, we need to impute the missing costs and margins for the 2019 data. We do this in two steps.

Step 1: We impute the cost of each item in 2019 via an inflation-adjusting approach by using the unit cost and the inflation data. Based on conversations with our retail collaborator, we learned that procurement contracts are typically updated on an annual basis and that unit cost increases for the majority of items are largely due to inflation (or in sync with inflation rates). Hence, we calculate the average cost for each item j in each channel c (i.e., online and offline) and each year (i.e., 2020, 2021, and 2022), denoted as $cost_{jc,2020}$, $cost_{jc,2021}$, and $cost_{jc,2022}$. [21] We choose Brazil's Extended National Consumer Price Index (IPCA) in June (the midpoint of the year) as the average level of IPCA for each year, and obtain a set of unit cost estimates for each item j in

[21]We do not consider the cost value in 2023 as the basis for the cost imputation, since we do not observe the data for the full year and cannot compute the yearly average cost for 2023.

channel c, denoted as s_{jc}, by adjusting the ratio of IPCA index in Equation A2.[22]

$$s_{jc} = \{cost_{jc,2020} \times \frac{ICPA_{2019}}{ICPA_{2020}}, cost_{jc,2021} \times \frac{ICPA_{2019}}{ICPA_{2021}}, cost_{jc,2022} \times \frac{ICPA_{2019}}{ICPA_{2022}}\} \tag{A2}$$

We can then take the average of the elements of set s_{jc} for each item-channel combination and calculate the imputed cost of item j in channel c in 2019, as $cost_{jc,2019}$.

Step 2: Since we observe the price, we can then simply calculate the margin for each item-channel combination in 2019 as

$$margin_{jc} = price_{jc} - cost_{jc,2019}. \tag{A3}$$

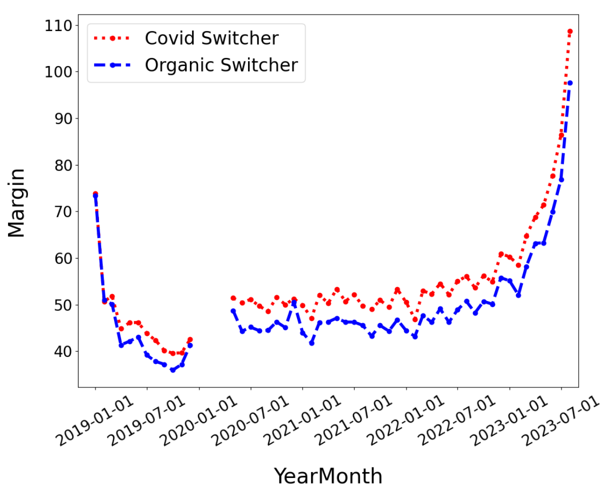
We then aggregate these margins at the customer-month level to get the customer-month panel data for the profitability measures in 2019.

F.2 Validity and Robustness Checks

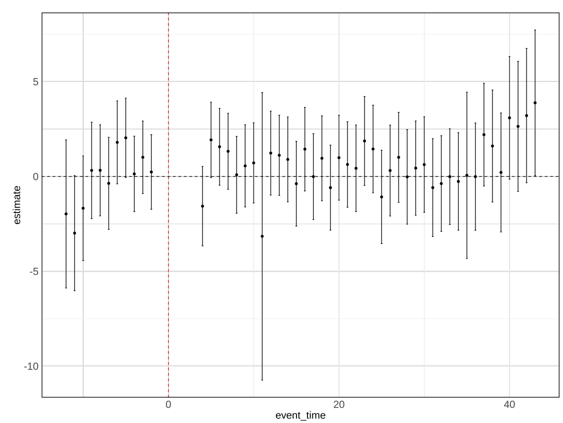
Figure A5 presents the parallel trend plot for both organic and covid switchers, and Figure A5b confirms that the parallel trend assumption for total margin is satisfied, since all pre-treatment coefficients are insignificant from 0.

Figure A5: Parallel Pre-Trend Assumption - Total Margin

(a) Time Series Visualization



(b) Pre-Trend Test via Event Study Model



The profitability analyses in § 4.5 show that, on average, covid switchers are more profitable than organic switchers after switching to omnichannel. However, one could argue these may simply reflect differences in the retailers' assortment differences across the two channels. To examine if this is the case, we now focus on the 25,258 SKUs that are common to both online and offline channels. In other words, we calculate the margin of each customer-month by excluding those items that are exclusive to a specific channel.[23] Table

[22]We obtain Brazil's Broad National Consumer Index (IPCA) from the Brazil Institute of Geography and Statistics (IBGE). The IPCA data is available under the Historic Serie section of the following URL:https://www.ibge.gov.br/en/statistics/economic/prices-and-costs.

[23]In total, there are 37,980 unique (SKUs) purchased at least once during our sample period. 8,538 SKUs are exclusively bought offline, and 4,184 SKUs are exclusively bought online.

A17 presents the DID results based on the profitability measures using these 25,258 common items. We see that the incremental profit of covid switchers is 9.65 units greater than that of organic switchers, after switching to omnichannel. These findings are consistent with the main finding in Table 15.

Table A17: Difference-in-Difference Margin Result Using only Common Items (Covid vs. Organic Switchers)

Dependent Variable:	Total Margin	
Model:	(1)	(2)
Variables		
Constant	141.7***	
	(1.65)	
Covid_Switcher	10.4***	
	(1.95)	
Post_Covid	28.6***	
	(1.79)	
Covid_Switcher * Post_covid	8.36***	9.65***
	(2.12)	(2.26)
Fixed-effects		
Customer		Yes
YearMonth		Yes
Fit statistics		
Observations	1,870,685	1,870,685
R^2	0.00205	0.30584
Within R^2		3.05×10^{-5}

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05*

References

Joshua Angrist and Joern-Steffen Pischke. Mostly harmless econometrics (princeton uni, 2008.

Maithreyi Gopalan, Kelly Rosinger, and Jee Bin Ahn. Use of quasi-experimental research designs in education research: Growth, promise, and challenges. *Review of Research in Education*, 44(1):218–243, 2020.