

Mobile App Introduction and Online and Offline Purchases and Product Returns

Unnati Narang,^a Venkatesh Shankar^b

^a Department of Marketing, Mays Business School, Texas A&M University, College Station, Texas 77843; ^b Center for Retailing Studies, Mays Business School, Texas A&M University, College Station, Texas 77843

Contact: unarang@mays.tamu.edu (UN); vshankar@mays.tamu.edu, <https://orcid.org/0000-0001-5476-3052> (VS)

Received: September 30, 2017

Revised: July 21, 2018; January 31, 2019; March 24, 2019

Accepted: March 25, 2019

Published Online in Articles in Advance: ■■ ■■, 2019

<https://doi.org/10.1287/mksc.2019.1169>

Copyright: © 2019 INFORMS

Abstract. How do a retailer’s mobile app adopters differ from nonadopters in their shopping outcomes, such as online and offline purchases and product returns? In this paper, we model the relationship between a retailer’s mobile app launch and the *frequency*, *quantity*, and *monetary value* of purchases in its online and offline channels, as well as product returns. We leverage data on a large retailer’s launch of a mobile app and use a difference-in-differences approach. Our results show that app adopters buy 33% more frequently, buy 34% more items, and spend 37% more than non-adopters in the period after app introduction. At the same time, they return 35% more frequently, 35% more items, and 41% more in dollar value. Combined, app adopters spend 36% more in net monetary value. Furthermore, app adopters’ purchases in both the online and offline channels increase after app launch. The time, location, and features of app use provide descriptive evidence of how the app aids shopping in other channels. *App-linked shoppers* (those who make a purchase within 48 hours of app use) use the app when they are close to the store of purchase and access the app for loyalty rewards, product details, and notifications. These insights offer important substantive implications.

History: K. Sudhir served as the editor-in-chief and Catherine Tucker served as associate editor for this article. This paper has been accepted for the *Marketing Science* Special Issue on Mobile Technologies.

Supplemental Material: Data and the web appendix are available at <https://doi.org/10.1287/mksc.2019.1169>.

Keywords: mobile marketing • mobile apps • channels • returns • difference-in-differences

1. Introduction

In recent years, the penetration of mobile devices has reached unprecedented levels. In 2018, 77% of the U.S. population owned a smartphone (Pew Research Center 2018). Mobile apps are increasingly dominating mobile device use. Mobile apps account for 87% of mobile usage at 2.5 daily hours for an average adult in the United States (*eMarketer* 2017), which constitutes the bulk of digital media time (Comscore 2016). Retail apps are among the fastest-growing app categories (Williams 2018), and the average U.S. adult has at least four retail apps on his or her phone (CNBC 2018). For example, nearly 70% of Walgreens’ customers engage through mobile devices, and more than 50% of app adopters use the app while shopping in stores (Hyken 2017). Furthermore, about 20% of all Starbucks transactions in stores originate from its “order and pay” app (*Forbes* 2015).

The purpose of this paper is to quantify the relationship between a retailer’s branded mobile app launch and its online and offline purchases and product returns using data from a large-scale U.S. retailer of video games, electronics, and wireless services. In particular, we evaluate the changes in online and

offline shopping behavior for app adopters relative to nonadopters, and we examine the app use patterns associated with purchases and returns in all the channels. We define app adopters as those who download *and* use the app. A mobile app may prompt shoppers to purchase more often, purchase more items, and spend more through the app. However, is app introduction and adoption linked to purchases in stores and online through the retailer’s website? Furthermore, although a mobile app can prompt shoppers to purchase, does it change their product returns and the net monetary value of purchases?

The nascent literature on mobile apps is silent on how app adopters differ from nonadopters in their online and offline purchases and product returns. Two related studies focus on aggregate loyalty points or total purchases (Einav et al. 2014, Kim et al. 2015) but not on purchases by channel. Wang et al. (2015) examine the effect of a campaign promoting the use of an app on online purchases. In a cross-platform/channel study, Xu et al. (2016) examine a tablet adoption’s relationship with online purchases via smartphones and desktops but not through the store. Prior research has also not examined the differences between

adopters and nonadopters with regard to product returns. Specifically, we address the following research questions:

- Do mobile app adopters have higher or lower frequency, quantity, and monetary value of purchases and product returns across all the channels relative to nonadopters?
- How do app adopters and nonadopters differ in their purchases in existing online and offline channels?
- What potential app use patterns are associated with purchases in all the channels?

We address our research questions using a unique data set relating to app introduction by a large retailer of video games, consumer electronics, and wireless services. The data set spans 18 months before and after introduction of the app. Teasing out the differences in app adopters' and nonadopters' shopping outcomes in the absence of randomization is a complex task. Major issues include endogeneity and self-selection of shoppers adopting the app. We mitigate the challenges posed by nonexperimental variation in the data in three ways: (1) by adopting propensity score matching to identify app nonadopters similar to adopters along their observed characteristics and comparing their shopping outcomes before and after app introduction, (2) by using the number of cell towers in the shoppers' zip code as a source of exogenous variation in app adoption to mitigate endogeneity as a result of unobservables, and (3) by conducting a series of robustness tests to rule out alternative explanations.

Our results show that app adopters buy 33% more frequently, buy 34% more items, and spend 37% more than nonadopters in the period after app introduction. At the same time, they return 35% more frequently, 35% more items, and 41% more in dollar value. Combined, app adopters offer 36% more in net monetary value (NMV) than do nonadopters after app introduction. Our analyses suggest that the time, location, and features of app use provide descriptive evidence of how the app aids shopping in other channels. App-linked shoppers (those who make a purchase within 48 hours of app use) use the app when they are close to the store of purchase and can access the app for loyalty rewards, product details, and notifications. App adopters also purchase a more diverse set of items (including less popular products) than do nonadopters. App adopters return products originally purchased in stores and through rewards more than those purchased online and without rewards.

Our research contributes to the mobile marketing and omnichannel marketing literatures in at least three important ways. First, we expand the scope of the mobile marketing literature by examining mobile app adopters' and nonadopters' purchases *and* returns. Second, unlike most prior mobile marketing studies

that focus on only online purchase outcomes, we consider both online *and* in-store purchases. Finally, we examine app use patterns in depth.

Our findings offer key managerial implications. We provide useful benchmarks for purchases and returns for managers to evaluate the app introduction decision. Our results also suggest that managers should plan for a higher number of store and website visits by app adopters. Managers can predict more purchases from adopters based on their time, location, and features used in the app. Finally, managers should plan on receiving greater returns from app adopters.

2. Related Literature

Mobile devices influence shoppers both in and out of a store (Shankar et al. 2016). Mobile apps may affect purchases in three major ways. First, mobile apps can provide anytime, anywhere information benefits to shoppers. Such benefits include product- and store-related information (Danaher et al. 2015, Fong et al. 2015, Dubé et al. 2017). Second, mobile apps can offer immediate access to shopping (Shankar and Balasubramanian 2009), potentially driving impulse buying through deals. Finally, apps may serve as convenient tools and reminders for shopping.

Two streams of research, one on mobile apps and the other on mobile device adoption, are closest to our research questions. First, the sparse literature on mobile apps considers how app usage relates to a limited set of purchase measures. Kim et al. (2015) study how the use of two app features influences shoppers' loyalty point accruals for an air miles reward program. They treat an app update with two new app features (not the app itself) as the intervention and find that point accruals across retailers increase. In a descriptive analysis of eBay's app usage, Einav et al. (2014) show that the app results in an immediate and sustained increase in the platform's aggregate revenues. Wang et al. (2015) show that a grocery retailer's promotional campaign for a mobile app results in a greater number of orders and spending. Gill et al. (2017) report similar findings for a business-to-business (B2B) app. Other papers examine related variables, such as purchase intent, website visits, and firm value (Bellman et al. 2011, Xu et al. 2014, Boyd et al. 2019).

Second, studies on mobile device adoption focus on the effects of a mobile device on purchase frequency and monetary value. Lee et al. (2016) find that purchase frequency is higher but the monetary value of each purchase is lower for shoppers who transact more using a mobile device than a desktop. Xu et al. (2016) study the effect of tablet adoption on commerce through different devices in online retailing and conclude that tablet commerce substitutes desktop commerce but complements smartphone commerce.

Table 1. Selected Related Literature and Our Contribution

Paper	Focus	DV ^a = FP	DV = QP	DV = VP	DV = FR	DV = QR	DV = VR	Other DV	Cross-channel effect in stores and online	Context
Prior research on app adoption										
Bellman et al. (2011)	Effect of app use on brand attitude and purchase intention							Purchase intent		Laboratory study
Einav et al. (2014)	Analysis of eBay’s mobile app adoption and platform revenues			✓						Online retail
Xu et al. (2014)	Effect of mobile app on demand at the mobile site							Site visit		Online news
Kim et al. (2015)	Effect of use of app check-ins and information lookups on loyalty point accruals							Loyalty point accruals		Air miles reward app
Gill et al. (2017)	Effect of manufacturer’s mobile app on B2B revenues	✓	✓	✓						B2B app of a tools manufacturer
Prior research on mobile device adoption										
Lee et al. (2016)	Effect of mobile shopping ratio (mobile vs. web) on purchases	✓		✓						Online retail
Xu et al. (2016)	Effect of tablet adoption on digital commerce via smartphones and PC devices			✓						Online retail
Our paper	Linkages between app introduction and purchase and returns across all channels	✓	✓	✓	✓	✓	✓	—	✓	Large retailer with a chain of stores and ecommerce site

^aDV refers to the key dependent variables used in the study, including frequency (FP), quantity (QP), and value (VP) of purchases and frequency (FR), quantity (QR), and value (VR) of returns.

Overall, they find that tablet adoption enhances e-commerce revenues.

Combined, these research streams suggest that mobile apps and devices positively affect aggregate purchases but do not inform us about how mobile apps relate to outcomes in different channels, such as store and web, at the shopper level. Nor do the research streams examine product returns.

Our study complements and extends these research streams as shown in Table 1. First, we study the relationship between a retailers’ mobile app launch (and subsequent adoption) and a variety of shopping outcomes, such as the frequency, quantity, value of purchases, and importantly, product returns. Second, we examine outcomes in existing online and offline channels of a retailer. Finally, we explore app use patterns that can potentially explain the relationship between app adoption and shopping in other channels.

We also draw from and extend the literature on cross-channel purchases and product returns. With regard to purchases, retail stores have a complementary relationship with the internet channel in the long run (Avery et al. 2012). Online sales increase after offline stores open because of gains in brand awareness in markets with low brand presence and gains in channel awareness (Wang and Goldfarb 2017, Bell et al. 2018). Furthermore, online returns decline because of enhanced product information offline (Bell et al. 2018). We apply these findings to our examination of purchases in both online and offline channels and to product returns after the app introduction.

3. Data and Research Setting

We collect data from a large U.S.-based retailer of video games, consumer electronics, and wireless services with 32 million customers. The gaming industry is a large \$99.6

billion industry and offers a ripe setting to examine shopping behaviors. Overall, 48% of the video game purchases are physical disc format games (Minotti 2016).

The retailer is similar to Walmart and PetSmart or any other brick-and-mortar chain with a relatively larger offline presence. The retailer's primary channel is its store network comprising more than 4,175 stores across the United States. In addition, it also has an e-commerce site, but only about 5% of its sales transactions are online, although the proportion is growing fast. The retailer typically allows product returns for a full refund within 30 days of purchase. Returns take place in stores.

Our focal independent variable/intervention is app introduction. The retailer introduced its app on July 1, 2014, without any targeted campaign. Approximately 6% of shoppers adopted the app in the 18 months after launch. The app allowed shoppers to browse the retailer's product catalog, get deals and offers, order online through the mobile browsers, and locate nearby stores to buy offline (including open hours, phone numbers, and driving directions). The app allowed shoppers to add products to an in-app cart linked to a checkout button. When shoppers clicked "Checkout," the app redirected them to the retailer's mobile site in the mobile browser to make an online purchase. In this way, this app is similar to several retailer apps (e.g., PetSmart). Web Appendix Figure A1 provides screenshots from the app.

We next describe the data used in our analysis. Mainly for security and privacy reasons, the retailer allowed us to access only a random sample of about 55,580 app adopters (from a population of about 2 million app adopters who started using the app for the first time in our data period) and 63,164 nonadopters (from the remaining 30 million shoppers).¹ The data on these shoppers include demographic information (age, gender, zip code) and loyalty program membership status. The firm provided us with two kinds of data for these shoppers, transactional data across channels and mobile app use data. From the transactional data, we have access to purchases in the firm's online and offline channels and to product returns. The online channel represents purchases made via the retailer's website, including those that come through the app checkout in the mobile browsers. In addition, we have data on product returns by the shoppers.

The mobile app data include the app features that shoppers access with time stamps. App features relate to app banner, product, stores, shopping, offers, rewards program, and notifications. We next describe these different features. The *app banner* refers to the home screen of the app that displays offers or any other updates. *Product-related features* refer to product search, details, catalog, and videos that provide information about products. Similarly, *store-related features* refer to store-related information, such as store check-in, locations, and directions. *Offers* refer to a list of offers in the app

or clicking on offer details to read more about a specific offer. The *rewards program* features include the ability to check loyalty reward points and redeem reward coupons. Finally, the *notification feature* refers to registering for and acting on push notifications. This feature categorization appears in Web Appendix Table A3.

For our research design, we collected data for a period of 18 months before (pre) and after (post) the app launch. This resulted in a panel from January 2013 to December 2015. The average interpurchase time of shoppers for the retailer's products is about 37 days. To avoid left censoring, we also collected data on the recency of purchases—that is, the day on which the shoppers made their last purchase if the last purchase was before the start of our data period.

We supplement these data with publicly available data on the number of cell towers by zip code from the U.S. Federal Communications Commission (FCC). This variable provides exogenous variation in who adopts the app, mitigating concerns about unobservable factors that might influence adopters differently from nonadopters. Higher cell tower areas are likely to get better connectivity on mobile devices and higher probability of downloading the app. Because most cell towers were constructed prior to 1990s, their location is not likely to be correlated with demand-led factors in recent years, such as population growth or demand for video games differentially in high and low cell tower zip codes. We explain the rationale for this instrument and rule out concerns for several potential omitted variables in Section 4.3.2. We show that the number of cell towers in a shopper's zip code is less likely to be correlated with omitted factors that may also drive demand (e.g., store presence, store quality, income levels) during our data period. The key variables, their operationalization, and descriptive statistics appear in Table 2.

4. Analyses

4.1. Relationship Between App Introduction and Shopping Outcomes: Descriptive Analysis

We define *app adopters* as shoppers who started using the app for the first time during our data period. *Nonadopters* are those who did not access the app even once during the study period.² A simple comparison of shopping outcomes shows that the average monetary value of purchases increased 19.25% (\$63.60–\$75.84 per month) for app adopters, whereas it decreased marginally by 2.92% (\$21.56–\$20.93 per month) for app nonadopters after app introduction ($p < 0.001$). However, the average monetary value of returns for app adopters also increased by 19.37% (\$5.37–\$6.41 per month) compared with app nonadopters, who experienced a marginal decrease of 8.23% (\$1.58–\$1.45 per month) in the same period ($p < 0.001$). Overall, the net monetary value increased by more than 19.23% for

Table 2. Variable Definitions and Descriptive Statistics

Variable	Operationalization	Mean	Std. dev.	Min	Max
<i>Frequency of purchases</i>	Number of purchase transactions in the time period	0.81	1.58	0	101
<i>Quantity of purchases</i>	Number of items bought in the time period	1.50	3.47	0	542
<i>Value of purchases</i>	Monetary value of purchases in the time period (\$)	43.94	117.61	0	8,993.81
<i>Frequency of returns</i>	Number of return transactions in the time period	0.11	0.48	0	51
<i>Quantity of returns</i>	Number of items returned in the time period	0.15	0.73	0	120
<i>Value of returns</i>	Monetary value of returns in the time period (\$)	3.56	29.73	0	5,150.63
<i>Net monetary value (NMV)</i>	(Monetary value of purchases – Monetary value of returns) (\$)	40.37	107.22	–1,937	8,993.81
<i>App Adopters (TREAT)</i>	Dummy indicating whether the shopper adopted the app (= 1) or not (= 0)	0.47	0.50	0	1
<i>Time Period (POST)</i>	Dummy indicating whether the period is before (= 0) or after (= 1) app launch	0.50	0.50	0	1
<i>Recency</i>	Number of days since the shopper’s last purchase	193.60	293.71	0.08	1,493.23
<i>Age</i>	Age of the shopper in years at the start of the data period	31.98	10.78	11.00	115
<i>Gender</i>	Gender of the shopper (female = 2, male = 1, unknown = 0)	0.67	0.63	0.00	2.00
<i>Distance to nearest store</i>	Distance in miles between the geographical centers of the shopper’s and the nearest store’s zip codes	3.90	7.21	0.00	574.48
<i>Number of stores</i>	Number of the focal retailer’s stores in the shopper’s zip code	0.56	0.72	0.00	4
<i>Loyalty program level</i>	Dummy indicating whether the shopper is enrolled in the basic (= 0) or professional (= 1) membership category on app introduction date	0.16	0.37	0.00	1
<i>Area population</i>	Population of the shopper’s zip code based on 2010 U.S. census	31,437	19,090	6.00	113,916
<i>Competitor stores</i>	Number of competing stores in the shopper’s zip code	0.29	0.52	0.00	4
<i>No. of unique products</i>	Number of unique stock-keeping units that the shopper buys	1.21	2.71	0.00	313
<i>No. of unique categories</i>	Number of unique categories that the shopper buys	0.84	1.53	0.00	29
<i>Pct. of top 100 products</i>	Spending on the top 100 products/Total spending	0.21	0.34	0.00	1
<i>Pct. of top 500 products</i>	Spending on the top 500 products/Total spending	0.47	0.42	0.00	1
<i>Cell towers</i>	Number of cell towers in the shopper’s zip code	6.99	6.93	0.00	52
<i>Precipitation</i>	Average precipitation level in the shopper’s zip code as reported by the NOAA in millimeters in June 2014	102.82	74.81	0.00	498.5
<i>Temperature</i>	Average air temperature of the shopper’s zip code in Celsius as measured by the NOAA in June 2014	23.53	3.57	0.96	36.38
<i>Download speed</i>	Percentage of the population in the shopper’s county with download speeds less than 6,000 KBps as reported by the FCC in June 2014	0.01	0.02	0.00	0.81
<i>Wireless access</i>	Percentage of the population in the shopper’s county with access to three or more wireless providers as reported by the FCC in June 2014	0.12	0.02	0.00	1

Notes. The statistics for the outcome variables (e.g., frequency, quantity, and value of purchases and returns) are averaged over the 36-month data period. NOAA, National Oceanic and Atmospheric Administration.

app adopters compared with a 2.53% decline for nonadopters. Notably, both the online and offline purchases are higher for app adopters than for nonadopters by 55.77% and 21.34%, respectively ($p < 0.001$). For the matched sample of nonadopters in Table 3, the comparison shows a similar direction but a larger magnitude of differences in treated and *matched* controls versus treated and *unmatched* controls over the same time period. Between the pre and post periods, the value of purchases decreased by 17.45% (from \$63.10 to \$52.09) for the matched control shoppers, and the value of returns decreased by 21.10% (from \$5.26 to \$4.15).

4.2. Econometric Model

To estimate the effect of mobile app introduction, the ideal approach would be to compare the shopping outcomes for shoppers who adopt the mobile app to the counterfactual—that is, to outcomes when the same shoppers do not adopt the mobile app. However, because we do not observe the counterfactual (a shopper cannot be both an app adopter and a non-adopter), and because the treatment is not randomly assigned (a shopper self-selects into adopting the app), we examine nonexperimental variation. Specifically, we compare the pre- and post-app introduction outcomes for app adopters and similar nonadopters.

Table 3. Model-Free Evidence: Mean Statistics

Variable	Treated pre period	Treated post period	Control pre period	Control post period	Matched controls pre period	Matched controls post period
Frequency of purchases	1.211	1.368	0.407	0.363	1.197	0.963
Quantity of purchases	2.256	2.524	0.570	0.665	2.214	1.737
Value of purchases	63.60	75.84	21.56	20.93	63.10	52.09
Frequency of returns	0.177	0.196	0.047	0.041	0.169	0.129
Quantity of returns	0.238	0.258	0.062	0.053	0.226	0.167
Value of returns	5.369	6.406	1.584	1.452	5.257	4.146
Net monetary value of purchases	58.236	69.438	19.975	19.481	57.85	47.94
Frequency of purchases—online	0.019	0.033	0.006	0.007	0.020	0.017
Quantity purchased—online	0.028	0.050	0.009	0.010	0.028	0.025
Value of purchases—online	1.41	1.997	0.495	0.425	1.571	1.085
Frequency of purchases—stores	1.192	1.335	0.400	0.356	1.177	0.946
Quantity purchased—stores	2.228	2.474	0.741	0.655	2.186	1.713
Value of purchases—stores	62.20	73.85	21.06	20.51	61.53	51.01

Notes. The pre period (post period) statistics are monthly averages over the 18-month period before (after) app launch across shoppers; online purchases in the post period for the treated include purchases on the mobile site after clicking checkout in app. Matched controls are nonadopters similar to adopters based on nearest neighbor matching using preperiod covariates.

After specifying the baseline regression model, in the next section, we outline our strategy to address the endogeneity of the individual app adoption decision using propensity score matching (Rosenbaum and Rubin 1983) and the Heckman selection model. We rule out several competing explanations for our results in the robustness checks. The complete list of analyses appears in Table 4.

We compare the change in outcomes for the app adopters 18 months before and 18 months after the retailer introduced the app to the change in outcomes for the nonadopters over the same time period. Formally, we specify our difference-in-differences linear regression model as

$$Y_{it} = \alpha_0 + \alpha_1 TREAT_i + \alpha_2 TREAT_i \times POST_t + \tau_t + \vartheta_{it}, \quad (1)$$

where i is the individual, t is the month, Y is the outcome variable (frequency, quantity, and monetary value of purchases and returns), $TREAT$ is a dummy variable denoting app adoption (1 if shopper i is an app adopter and 0 otherwise) to account for time-invariant unobserved differences in adopters and nonadopters, $POST$ is a dummy variable denoting the period (1 for the period after the app has been introduced and 0 otherwise), α is a coefficient vector, τ represents month fixed effects, and ϑ is an error term. The coefficient of $TREAT \times POST$ represents the change in app adopters' outcomes relative to nonadopters' outcomes between the pre and post periods.

The challenge in identifying the model specified in Equation (1) is that app adoption is endogenous. In other words, shoppers choose to adopt an app in anticipation of future purchase behaviors and because

Table 4. Overview of Analyses

Section	Analysis	Objective	Key insight/conclusion
4.3	Propensity score matched and Heckman two-stage linear regression	Quantifying the relationship between app introduction and shopping outcomes	App introduction is related with higher frequency; quantity; and monetary value of online and offline purchases, returns, and net monetary value of purchases.
5.2 and web appendix	Robustness checks (a) Alternative models: difference-in-differences without matching, and Poisson difference-in-differences (b) App adoption date as cutoff (c) App novelty (d) Alternative matching (e) Outliers (f) Shopper heterogeneity in deal proneness (g) Alternative sample	Ruling out alternative explanations for the results	App introduction effects are robust to alternative specifications and explanations, such as other adoption measures, time periods, app novelty, customer deal-proneness, and outliers.
6 and web appendix	App use patterns Time, geography, and features used	Identifying app use patterns associated with shopper purchases	Increased purchases seem to be associated with app use time, geography, and features.

of their inherent characteristics, both observed and unobserved (e.g., preference for the retailer). The selection challenges are posed by nonexperimental variation in our data similar to quasi-experiments (Tirunillai and Tellis 2017). Next, we describe empirical approaches for mitigating concerns for selection on observables and unobservables.

4.3. Endogeneity and Self-Selection

The first availability of the retailer’s mobile app represents a shock in our data period. The timing of the introduction of the app is exogenous to the shoppers, but the date on which an individual shopper adopts and starts using an app may be endogenous. Therefore, we use the app introduction date as the cutoff for defining the pre and post periods. In this way, our empirical strategy conservatively assumes that the benefits of the app start accruing right from the day of app introduction for all app adopters, irrespective of their self-selected time of app adoption (Manchanda et al. 2015). An alternative empirical strategy for estimating the effect of the app would be to treat the app adoption date for each adopter as the cutoff date for the intervention. This approach does not rule out endogenous app adoption timing because app adopters may decide to adopt the app in anticipation of increased future purchases. As a result, our main analysis reports the results from the first approach. However, as a robustness check, we estimate a model using the second approach that leverages individual adoption dates (see Section 5.2.2).

Arguably, one concern with using the app introduction date is the question of what exactly we are measuring. Is it the effect of app introduction or adoption by shoppers? Our context is similar to that of Manchanda et al. (2015). They study the launch of an online community that users join endogenously at different points in time. They use the launch of the community as the intervention and leverage the period before and after the community launch as their pre and post periods. However, they attribute the coefficients of *treatment effect* of launching the community to users’ act of joining the community.

In our setting, although we use the launch date to quantify the shift pre and post app introduction, the differences in shopping outcomes come from those who adopt the app (and not from their purchases even before they have adopted). We triangulate this through several analyses and falsification checks (see Section 4.3.2). In this sense, we measure the relationship between app introduction and shopping outcomes for the shoppers adopting the app. Therefore, our estimates are more representative of a local average treatment effect (LATE) than an average treatment effect (ATE) for the population.

In the absence of randomization on who adopts the app, we use several careful econometric techniques to examine the effects of app introduction. First, we use propensity score matching based on observed individual covariates. We conduct several tests to inspect the quality of the matches. Second, we estimate a two-stage Heckman selection model using the exogenous variation in the number of cell towers in the shopper’s zip code to mitigate selection on unobservables. Third, we carry out a set of additional falsification analyses. We next discuss our approaches for selection on observables and unobservables.

4.3.1. Selection on Observables: Propensity Score Matching.

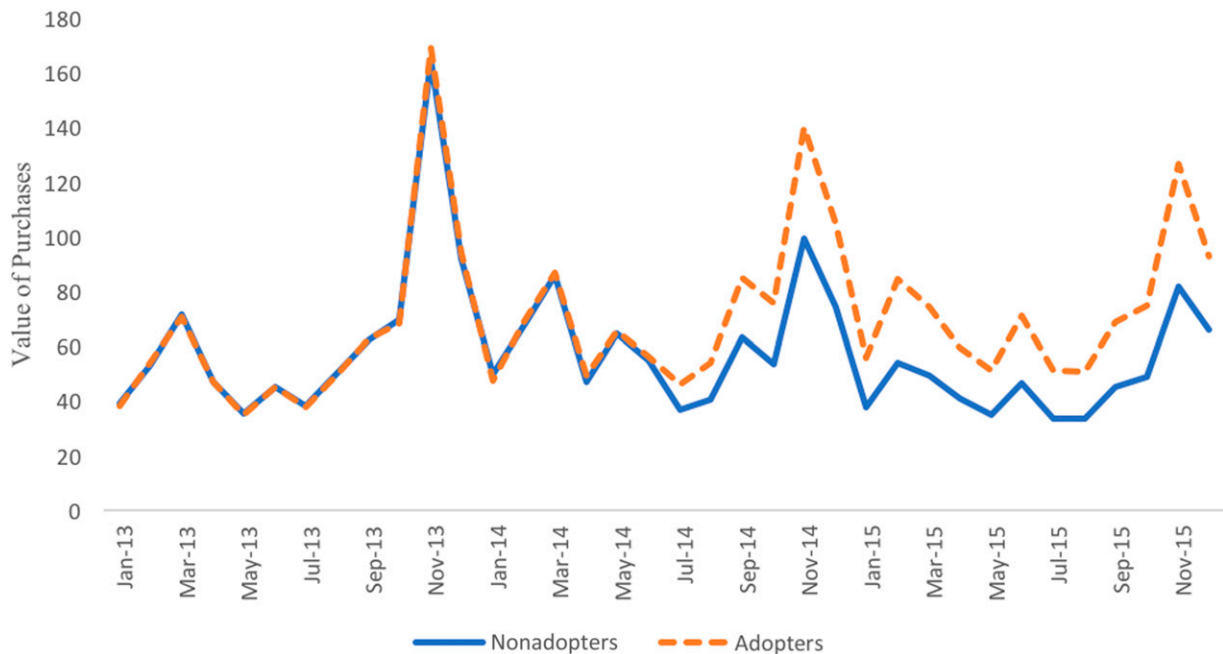
Propensity score matching allows us to match app adopters and nonadopters on observed demographic and behavioral covariates while tackling the curse of dimensionality (Rubin 2008, Guo and Fraser 2014). We begin by calculating each shopper’s propensity score, defined as the shopper’s probability of adopting the app. We do this using a binomial logit model (see Web Appendix Table A4). Next, we identify nonadopters similar to adopters based on the estimated propensity scores to create a control group. This approach is consistent with that of Rosenbaum and Rubin (1983) in creating a control group similar to the treated group in the distribution of observed covariates. We match each app adopter to a nonadopter based on the 1:1 nearest neighbor algorithm with replacement. Formally, if $P(X_i)$ is shopper i ’s propensity score, the treated shopper i is matched to the control individual j , where j is $\min\|P(X_i) - P(X_j)\|$, to create the matched pairs closest to each other (Wangenheim and Bayón 2007, Huang et al. 2012).

What factors explain the decision to adopt the app? Consistent with the existing research (Hung et al. 2003, Kim et al. 2015), we model app adoption as dependent on individual shopper demographics (e.g., age, gender), behavioral measures (e.g., pre-period spending and returns by month, recency, and frequency of purchase, online purchases), and other related measures (e.g., distance to the nearest store, number of stores in the shopper’s zip code, presence of competitor stores, loyalty program membership level on app introduction day) that are likely to influence shoppers:

$$\begin{aligned} \text{Propensity/Probability of App Adoption } A_i & \\ &= \frac{\exp(U_i)}{1 + \exp(U_i)} \end{aligned} \quad (2)$$

$$U_i = \gamma + \delta D_i + \varepsilon_i, \quad (3)$$

where i is the shopper, U is the utility from app adoption, D is a vector of covariates, (γ, δ) is a coefficient vector, and ε is an error term, distributed

Figure 1. (Color online) Monetary Value of Purchases Before and After App Introduction

Notes. The app was launched on July 1, 2014. In the pre period, the values of purchases are nearly identical for the adopters and nonadopters, hence appearing as a single line. The sharp peak in November 2013 reflects the launch of the Xbox One console, which sold more than 2 million units worldwide within the first month.

as double exponential (Wooldridge 2002). To test the goodness of our propensity score matches, we conduct a set of statistical analyses, including the standardized bias reduction test (Rosenbaum 2005).

Matching improves the percentage balance of propensity scores by 100%, making the matched treated and control groups comparable (see Web Appendix Table A5). This is important evidence that matching results in a valid control group. Another piece of evidence is visual verification that the pre period purchase trends for the treated and control groups are parallel (Figure 1).³ Web Appendix Table A6 presents the model estimates from the same comparison shown in Figure 1, based on the model in Equation (1), with treatment interacted with each monthly dummy. This check, together with the insensitivity of results to control variables (discussed in Section 5.2.1), assures us that we can reasonably mitigate selection concerns along observables.

Histograms showing the distribution of propensity scores for the treated and control groups before and after matching appear in Figure 2. The distributions visually look more similar after matching relative to before matching.

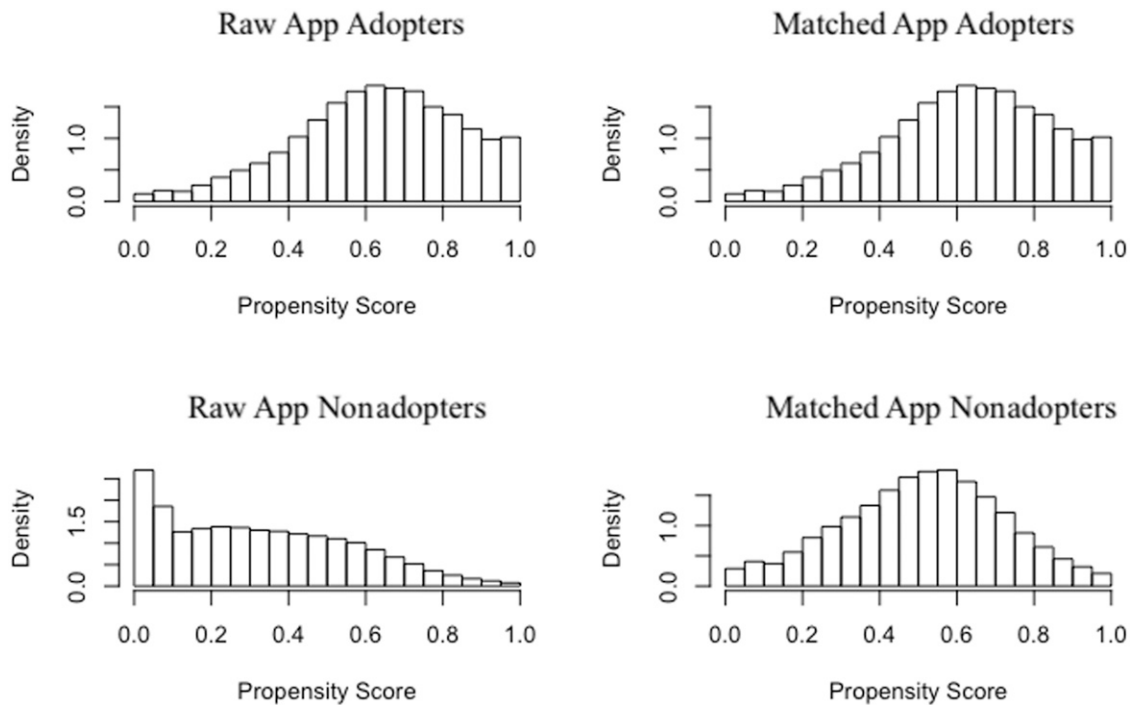
Matching estimates and inspecting common trends in our data ameliorate the self-selection concern to some extent. Nevertheless, we also carry out several other analyses to further reduce concerns for unobserved omitted variables (e.g., unobserved preference for the firm).

4.3.2. Selection on Unobservables: Heckman Correction and Other Analyses.

To more formally account for the nonrandomness of app adoption as a result of unobserved factors, in this section, we carry out two types of analyses. In the first set of analyses, we use (1) the two-stage Heckman correction procedure (Heckman 1979) and (2) an alternative control group based on future treated shoppers because they are more similar to app adopters than are nonadopters. In the second set of analyses, we exploit falsification tests to check that the estimates are not spurious by comparing app adopters' outcomes relative to nonadopters' outcomes in the periods after app introduction but prior to the adopters' specific adoption dates, consistent with Goldfarb and Tucker (2014).

We first describe the Heckman selection model. In the first stage, we model the choice to adopt the app using a probit model. App adoption is an endogenous variable in our outcome model because shoppers self-select into adopting the retailer's app. To identify the model, we require exogenous variation in app adoption. A potential source of such variation is the number of cell towers in the shopper's zip code. Presumably, a higher number of cell towers may be associated with better wireless connectivity on the shoppers' mobile devices. This is likely to influence app adoption. To meet the exclusion restriction, the number of cell towers should not be correlated with the error term ϑ_{it} . In other words, it should not drive demand through variables other than app

Figure 2. Distribution of Propensity Scores Before and After Matching



Note. Raw adopters' and nonadopters' plots represent pre-matching propensity scores.

adoption. Consistent with Bronnenberg et al. (2012) and Lundborg et al. (2017), we carefully consider and argue against several potential omitted variables that may be correlated with the number of cell towers.

First, in our context, high population growth areas or areas that experienced recent population surges could possibly endogenously drive the number of cell towers. If so, in such locations, the number of cell towers may be correlated with the demand-related factors. We analyze relevant data on population from recent years and the number of cell towers for the zip codes in our data. We examined the correlations between the number of cell towers constructed and the population in a zip code during 2010–2015. The correlation is low and positive at 0.072. Similarly, the correlation between the number of towers and the change in population between 2010 and the start year of our data—namely, 2013—is low and negative at -0.004 . Furthermore, we find that 90% of the cell towers in our data set were constructed prior to 2010, with 50% developed prior to 1990, before the birth of commercial internet. Thus, the changes in population and the resulting demand effects during our period of data would have less likely contributed to the variation we observe in cell towers across zip codes because of population growth. Furthermore, 42% of the cities in the data show stagnant or negative growth in population, ameliorating a concern that the emergence of cell towers in these already built-up cities could have been in response to population growth in recent years. In areas

where the locations of recently constructed cell towers map closely with population density, our exclusion restriction could fail.⁴ However, in our nationwide data, this is the case for only 0.6% of locations.

Second, a possible alternative variable through which the number of cell towers can directly affect shopping outcomes is store availability. However, the correlation between the number of towers and the number of stores in a zip code is low, at 0.18. During 2010–2015, the correlation between the number of new stores opened and the number of cell towers is even lower at 0.012. Nonetheless, we also control for population, the number of stores, and the distance of the shopper to the nearest store in our matching algorithm. Thus, both the app adopters and the matched nonadopters are similar in these variables besides the application of Heckman selection correction.

Third, another potential alternative variable through which the number of cell towers can directly affect shopping outcomes is store performance. Perhaps the retailer's stores in high cell tower regions perform better than stores in low cell tower regions because they might have superior staff and product assortment. To examine this possibility, we collect data on two commonly used measures of store performance, sales and sales per square foot (Gomez et al. 2004). We estimate the differences in these store performance measures across high and low cell tower zip codes based on the median split and find that there is no significant difference in the average store sales

($p > 0.10$) and average store sales per square foot ($p > 0.10$) during the data period.

Fourth, the number of cell towers could also affect shopping outcomes if high cell tower regions are those with stores offering better product assortment. We also confirm from the managers of the retail chain that they strategically vary neither the product assortments across stores nor between online and offline channels. This fact mitigates the concern that product assortment might vary between stores in high cell tower areas and stores in low cell tower areas.

Fifth, the number of cell towers may be endogenous because they might have been constructed in areas needing greater wireless connectivity. We find that most of the cell towers were built prior to 1990. However, even if the number of cell towers is highly correlated with the need for wireless connectivity, this demand is less likely to come from console buyers who primarily play games offline in the console and do not depend on connectivity to the internet. This focal retailer primarily sells console games. Furthermore, online game playing is a recent phenomenon, is still a small fraction of console game playing time, and occurred after most cell towers were built. About 70% console buyers play games that are offline, and their experience is not tied to better connectivity (Statista 2015). Anyone can purchase a game and play on his or her home console without being connected to the internet. Thus, in terms of scope and generalizability, our results are best interpreted in the context of retailing of an offline product (e.g., electronics, consoles, and offline games) rather than online and mobile games that require internet connection.

Sixth, the location of cell towers may be correlated with economic variables, such as income and employment. To examine this possibility, we analyzed the correlations by zip code. The correlation between the number of towers and the mean (median) household income in 2013 inflation-adjusted dollars is low at -0.13 (-0.11). Similarly, the correlation between the number of towers and employment rate (measured as the percent population 16 years and older in the labor force) is low at 0.03 .⁵

Next, we estimate a probit model to examine the strength of the first-stage effects. The exclusion restriction results in the following first-stage selection equation for modeling A_i , the probability of shopper i adopting the app:

$$\begin{aligned} \Pr(A_i = 1 | \text{CELLTOWER}_i, Q_i, \epsilon_i) \\ = \Phi(\partial_0 \text{CELLTOWER}_i + \partial_1 Q_i + \epsilon_i), \end{aligned} \quad (4)$$

where CELLTOWER is the number of cell towers in the shopper's zip code, Q is a vector of other covariates similar to those used in propensity score

matching, ∂ is a coefficient vector, and ϵ is an error term. We compute the inverse Mills ratio from this probit regression. We then augment the regression model in Equation (1) by including the inverse Mills ratio as an additional covariate in the second stage. Because the number of cell towers varies at the zip code level, we cluster standard errors by zip code to allow serial correlations and within-zip code correlations among shoppers (Bertrand et al. 2004).

Another approach we adopt for selection on unobservables is to use an alternative control group. In this approach, we treat future app adopters as a control group. Specifically, we consider a shopper who adopted the app in the 9 months between April 2015 and December 2015 (both months included) to be in the control group, whereas we view those who adopted the app in the first 9 months of app launch in July 2014 to be in the treatment group. We compare the outcomes of these groups of shoppers 9 months before and after app launch. This analysis is in the same spirit as the falsification analysis of Manchanda et al. (2015) and Goldfarb and Tucker (2011).

Finally, we report a falsification test to formally check whether app introduction has a strong relationship with shopping outcomes right from the time of launch or from when it is adopted by the shoppers. This test will help mitigate any concern about an unobserved event around app introduction that may be leading to spurious effects. Introduction by itself, in the absence of adoption, should not be linked to significant upticks in purchases despite a possible positive spillover effect of the news of app release. The way we conduct this falsification check is to first create different cohorts of app adopters based on adoption date. We then define the post-period observations as those from the months after app introduction but before the cohort's app adoption date. Thus, adopters of August 19, 2015, will have their post period defined as July 2014 through July 2015. Their pre period will be the same number of months before app launch (i.e., June 2013 through June 2014). We then estimate a regression model for these periods for the treated cohort with each adopting shopper's matched nonadopter. Although this test is not an exact replication of the full model because of the reduced number of post-period observations, it lends some confidence in the estimates by showing the lack of significant effects in the months before adoption.

5. Results and Robustness Checks

5.1. Results

We first report the results for the selection model. Table 5 shows the estimates from the first-stage self-selection probit model. A higher number of cell towers is associated with a higher adoption probability in the first stage ($p < 0.001$).

Table 5. First-Stage Probit Model Results

Variable	Coefficient (standard error)
<i>Cell towers</i>	0.0022*** (0.0006)
<i>Precipitation</i>	0.0003*** (0.0001)
<i>Temperature</i>	0.0036** (0.0012)
<i>Download speeds (less than 6,000 KBps)</i>	−0.0111 (0.0168)
<i>Wireless access</i>	0.0075 (0.0184)
<i>Age</i>	−0.0142*** (0.0004)
<i>Gender</i>	0.1566*** (0.0066)
<i>Recency</i>	−0.0018*** (0.0000)
<i>Loyalty program level</i>	−0.1902*** (0.0106)
<i>No. of stores in zip code</i>	−0.0168* (0.0067)
<i>Area population</i>	0.0000*** (0.0000)
<i>Distance to nearest store</i>	−0.0004 (0.0007)
<i>No. of competitor stores</i>	0.0201* (0.0083)
<i>Past frequency of purchases</i>	0.0003*** (0.0000)
<i>Past quantity of purchases</i>	−0.0003*** (0.0001)
<i>Past monetary value of purchases</i>	0.0001*** (0.0000)
<i>Past frequency of returns</i>	−0.0024*** (0.0010)
<i>Past quantity of returns</i>	−0.0019 (0.0062)
<i>Past monetary value of returns</i>	0.0225*** (0.0016)
<i>Past frequency of purchases—online</i>	−0.0001 (0.0039)
<i>Past quantity of purchases—online</i>	0.0406*** (0.0151)
<i>Past monetary value of purchases—online</i>	0.0145* (0.0074)
<i>Past quantity of purchases per order</i>	−0.0003*** (0.0001)
<i>Intercept</i>	−0.0084 (0.0363)

Notes. The number of cell towers in the shopper’s zip code provides the exogenous variation in app adoption; $N = 118,744$, Akaike information criterion = 127,565.6, and Bayesian information criterion = 127,798. Past period refers the 18-month period totals prior to app introduction.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

We include the inverse Mills ratio (IMR) obtained from the first stage in the second-stage outcome model for the propensity score-matched sample. The results for this model appear in Table 6. IMR is significant ($p < 0.001$). The coefficient of IMR is the fraction of the covariance between the decision to adopt the app and shopping outcome relative to the variation in decision to adopt the app. The coefficient of IMR is indeed negative (similar to Gill et al. 2017). This sign indicates that the selection correction term, resulting from omitted factors affecting shoppers, adjusts shopping outcomes downward.

The results in Table 6 show a positive and significant relationship between app introduction on the frequency, quantity, and monetary value of purchases and returns ($p < 0.001$). These estimates are more representative of a LATE than an ATE for the population. They are particularly relevant for app adopters whose decision to adopt the app is affected by the number of cell towers in the zip code, consistent with Sudhir and Talukdar (2015). App adopters spend \$23.26 more than nonadopters after app introduction and engage in a higher frequency ($\alpha_2 = 0.39, p < 0.001$) and quantity of purchases ($\alpha_2 = 0.75, p < 0.001$). Interestingly, relative to nonadopters, app adopters also return \$2.15 more of products and engage in a greater frequency ($\alpha_2 = 0.06, p < 0.001$) and quantity of

returns ($\alpha_2 = 0.08, p < 0.001$). Overall, app adopters’ net monetary value of purchases is 36.49% higher than those of matched pre-period nonadopters.⁶

Furthermore, both the online and offline purchases tend to be higher for app adopters relative to nonadopters (Table 7). The absolute effects are much larger for in-store purchases (\$22.18 additional amount) than for online purchases (\$1.07 additional amount). However, in percentage terms, the increase in online purchases for app adopters relative to matched nonadopters is 68% compared with about 36% for corresponding in-store purchases ($p < 0.001$).

Our findings show a significant change in shopper behavior in other channels. We find that app adopters buy 33% more frequently, buy 34% more items, and spend 37% more than nonadopters after app introduction. At the same time, they return 35% more frequently, 35% more items, and 41% more in dollar value compared with nonadopters after app introduction. Overall, the app is associated with a 36% increase in net monetary value of purchases across the channels. Both the online and offline purchases are higher for app adopters.

Similarly, the main model with an alternative control group based on future adopters produces robust estimates. These results appear in Table 8.⁷ In this method, we use matched future app adopters from April to

Table 6. App Introduction and Aggregate Shopping Outcomes

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
<i>(App Adopters × Post App Introduction)</i>	0.391*** (0.014)	0.745*** (0.027)	23.256*** (0.749)	0.059*** (0.004)	0.080*** (0.008)	2.149*** (0.217)	21.107*** (0.668)
<i>App Adopters</i>	0.018 (0.021)	0.051 (0.04)	0.746 (1.147)	0.009 (0.008)	0.013 (0.013)	0.134 (0.373)	0.612 (0.970)
<i>IMR</i>	-2.446*** (0.045)	-4.504*** (0.083)	-122.22*** (2.096)	-0.380*** (0.016)	-0.508*** (0.025)	-11.232*** (0.624)	-110.99*** (1.812)
<i>Intercept</i>	7.148*** (0.126)	13.123*** (0.233)	350.28*** (5.986)	1.094*** (0.045)	1.456*** (0.07)	32.731*** (1.81)	317.55*** (5.168)
Number of observations	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760
R-squared	0.156	0.117	0.098	0.047	0.038	0.012	0.099

Notes. Robust standard errors clustered by zip code are in parentheses; month fixed effects are included. IMR, inverse Mills ratio from the selection model.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

December 2015 as comparable nonadopters for app adopters from July 2014 to March 2015 and compare their outcomes before and after 9 months of app introduction. The results from this model are substantively similar.

Finally, we consider the pre adoption period for the app adopters as their post period relative to the pre launch period by adoption date-based cohorts. The estimates for nine such cohorts (randomly drawn from adoption dates between April and December 2015) appear in Table 9. We report the coefficient of the $TREAT \times POST$ variable for the adopters on each of the dates relative to their matched nonadopters. The post-period observations are those from the months after app introduction but before app adoption. From Table 9, most coefficients are insignificant. Even so, for the cohorts with significant coefficients in Table 9, we do not spot any systematic pattern. Thus, the availability of the app does not appear to be related to shopper behavior in the absence of adoption. The

results in Table 9 are for the shorter-term period relative to the main model period of 18 months. Thus, these estimates are best viewed as depicting short-term relationships.

5.2. Checking Robustness of Results and Ruling Out Alternative Explanations

We obtain consistent results when using individual and month fixed effects in our model specification. In addition, we perform several robustness checks and tests to rule out alternative explanations for the linkages we find between app introduction and purchases and returns.

5.2.1. Alternative Model Specifications. In addition to our preferred model with propensity score-matched linear regression and Heckman selection correction, we also estimate models without matching and using Poisson count data models for frequency and quantity variables. The results from these models replicate the

Table 7. App Introduction and Purchases by Channel

Variable	(1) Offline			(2) Online		
	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of purchases	Quantity of purchases	Value of purchases
<i>(App Adopters × Post App Introduction)</i>	0.375*** (0.014)	0.719*** (0.027)	22.183*** (0.737)	0.017*** (0.001)	0.026*** (0.002)	1.074*** (0.125)
<i>App Adopters</i>	0.019 (0.021)	0.051 (0.039)	0.902 (1.126)	0.000 (0.001)	0.000 (0.002)	-0.157 (0.131)
<i>IMR</i>	-2.411*** (0.044)	-4.454*** (0.082)	-119.91*** (2.063)	-0.035*** (0.002)	-0.05*** (0.003)	-2.306*** (0.122)
<i>Intercept</i>	7.042*** (0.125)	12.974*** (0.231)	343.72*** (5.894)	0.106*** (0.004)	0.149*** (0.008)	6.56*** (0.342)
Number of observations	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760	4,001,760
R-squared	0.050	0.038	0.051	0.003	0.002	0.003

Notes. Robust standard errors clustered by zip code are in parentheses; month fixed effects are included. IMR, inverse Mills ratio from the selection model.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table 8. Alternative Model with Future App Adopters as Control

Variable	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
<i>(App Adopters × Post App Introduction)</i>	0.262*** (0.014)	0.482*** (0.031)	15.534*** (1.196)	0.045*** (0.005)	0.063*** (0.007)	1.789*** (0.273)	13.745*** (1.084)
<i>App Adopters</i>	0.007 (0.023)	0.015 (0.048)	0.345 (1.484)	0.000 (0.007)	−0.003 (0.011)	0.054 (0.293)	0.299 (1.338)
<i>IMR</i>	−3.248*** (0.078)	−5.96*** (0.149)	−173.679*** (4.172)	−0.518*** (0.017)	−0.696*** (0.025)	−16.03*** (0.635)	−157.65*** (3.750)
<i>Intercept</i>	9.745*** (0.206)	17.911*** (0.398)	508.894*** (11.07)	1.55*** (0.048)	2.096*** (0.07)	44.981*** (1.721)	463.913*** (9.967)
Number of observations	1,388,772	1,388,772	1,388,772	1,388,772	1,388,772	1,388,772	1,388,772
R-squared	0.058	0.041	0.055	0.014	0.011	0.004	0.059

Notes. Robust standard errors clustered by zip code are in parentheses; month fixed effects are included. In this method, matched future app adopters from April to December 2015 are used as controls for app adopters from July 2014 to March 2015; their outcomes before and after nine months are compared. IMR, inverse Mills ratio from the selection model.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

findings from Tables 6 and 7 and appear in Web Appendix Tables A8–A11. Coefficients of the treatment effect from Web Appendix Tables A8 and A9 represent unconditional changes in outcomes as a result of the app, without conditioning on covariates via propensity scores. Web Appendix Tables A10 and A11 report the results from Poisson count data models. These results are substantively similar to those in Tables 6 and 7 for the matched samples. The insensitivity of the results to control variables suggests that the effect of unobservables relative to these observed covariates would have to be very large to change our results significantly (Altonji et al. 2005).

5.2.2. Alternative Intervention/Cutoff Date. We estimated an alternative model with a more nuanced pre- and post-period intervention/cutoff date based on individual app adoption. In our main model, we compared app adopters and nonadopters along their outcomes before and after app introduction. In the alternative model, we create a new setup where we treat the app adoption date for each treated shopper as the cutoff date for the intervention. In this model,

$POST_t$ in the regression from our estimation equation (1) becomes $POST_{it}$, a binary variable that indicates the post-app adoption period for each treated shopper and the matched counterpart. For example, if a shopper first adopted the app in September 2014, the indicator value is 1 for subsequent months and otherwise 0. The matched shopper for this adopter will share the same value. Thus, the comparison time window is the same for each matched pair (Xu et al. 2016). The results appear in Web Appendix Tables A12 and A13. The effects are robust and consistent with previous findings.

5.2.3. App Novelty Effect. An alternative explanation for the increased net monetary value of purchases after app adoption could be the novelty of the app. It is possible that the app triggers a heightened shopping response only because of a temporary novelty effect around its launch that fades after a few months. To test this explanation, we reestimate our models using different windows of time, including 3, 6, 9, and 12 months before and after introduction for the matched sample of adopters and nonadopters.

Table 9. Falsification Test with the Post Period Defined as After App Introduction but Before App Adoption

Cohort	Frequency of purchases	Quantity of purchases	Value of purchases	Frequency of returns	Quantity of returns	Value of returns	NMV of purchases
4/16/15	0.139 (0.14)	0.219 (0.287)	−2.511 (12.827)	0.064 (0.052)	0.068 (0.07)	6.218 (3.443)	−8.729 (11.557)
5/10/15	0.287 (0.147)	0.561 (0.316)	22.386* (10.948)	0.068 (0.042)	0.088 (0.059)	3.993 (2.407)	18.394 (9.957)
6/28/15	0.038 (0.11)	−0.179 (0.229)	−0.249 (10.03)	−0.03 (0.036)	−0.051 (0.064)	−1.28 (2.629)	1.031 (9.462)
7/18/15	0.159 (0.123)	0.058 (0.248)	5.235 (9.485)	0.023 (0.039)	0.025 (0.055)	0.398 (2.547)	4.837 (8.434)
8/19/15	0.168 (0.119)	0.164 (0.252)	−10.622 (9.674)	0.085* (0.04)	0.089 (0.052)	1.296 (1.576)	−11.918 (9.464)
9/22/15	0.337* (0.145)	0.718* (0.351)	19.026 (12.749)	−0.013 (0.04)	−0.002 (0.059)	−2.928 (2.83)	21.953* (11.109)
10/21/15	−0.127 (0.156)	−0.062 (0.291)	5.929 (10.365)	−0.06 (0.039)	−0.086 (0.061)	−2.602 (1.596)	8.531 (9.964)
11/8/15	0.033 (0.139)	0.039 (0.286)	−1.064 (11.122)	0.008 (0.03)	0.011 (0.04)	−1.076 (1.62)	0.013 (10.746)
12/15/15	0.056 (0.112)	−0.024 (0.24)	−12.333 (11.605)	0.058 (0.052)	0.084 (0.062)	−0.7 (3.237)	−11.633 (10.541)

Notes. Coefficients of $TREAT \times POST$ are reported in this table for cohorts based on their date of app adoption for randomly selected dates between April and December 2015. Robust standard errors clustered by zip code are in parentheses; month fixed effects are included.

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

The effects of the app introduction persist in these varying windows of time (see Web Appendix Tables A14 and A15).

5.2.4. Alternative Matching Methods. Our main analysis relies on the commonly used 1:1 nearest neighbor matching algorithm with replacement. In addition, we use the Mahalanobis metric, the nearest neighbor without replacement, and a refined caliper matching approach by defining the bandwidth within which to identify matched control units (Silverman 1986). We test using a bandwidth of 0.27 and a tighter 0.05 times the standard deviation of the propensity scores. Web Appendix Tables A16 and A17 report the results for these alternative matched samples. We find that these estimates are consistent with those from our proposed method.

5.2.5. Outliers. Another possible explanation for the effects could be outliers, such as the top spenders (and not the average shopper). We test this possible explanation by first removing the top spenders from our sample (the mean plus 3 standard deviations of spending by shoppers in the pre period) and then carrying out difference-in-differences methods as done earlier. Our estimates are robust to outliers as shown in Web Appendix Tables A18 and A19.

5.2.6. Shopper Heterogeneity. Whereas we match shoppers on a broad set of covariates, an untested alternative explanation is that deal-prone shoppers rather than app usage largely drive the effects. In general, the retailer did not send any unique offers to mobile app adopters that the nonadopters did not receive, or vice versa. Yet, to rule out the possibility that the actual redemption or use of offers in the pre period could influence the two groups differently, we repeat the analyses after removing deal-prone shoppers. Web Appendix Tables A20 and A21 report the estimates for the sample that did not use deals in the pre period. The results are substantively similar.

5.2.7. Alternative Sample. Whereas we analyze a sample of 55,580 app adopters (about 3% of all adopters) in our main analysis, we collect additional data on another random sample of 50,000 app adopters. We repeat the analyses for this sample using a propensity score-matched linear regression with Heckman selection correction from our main model. We find consistent results for this alternative sample. Web Appendix Tables A1 and A2 report these estimates.

6. App Use Patterns: An Exploratory Examination

Our results show that app adoption and not its availability is primarily associated with shopping

outcomes (see Table 9). In this section, we examine app adopters' app use and offer a descriptive analysis of the linkages between shoppers' app use patterns and purchase or return transactions.

Specifically, we examine three elements of app use—namely, time, geography, and features—in the 6 months following the launch of the app.⁸ We identify and report differences in these usage aspects in these 6 months between app adopters who subsequently make a purchase (*app-linked shoppers*; $n = 21,092$) within 48 hours after app use and those who do not (*non-app-linked shoppers*; $n = 8,934$). We chose 48 hours based on the distribution of the time gap between app use and purchase. We examined the distribution of time gap between app use and purchase time (within one day [24 hours], two days [48 hours], etc.). We found that the largest chunk (41%) of adopters' purchases after app introduction occur within 48 hours of app use.

6.1. Temporal App Use–Purchase Linkages

The differences between app-linked and non-app-linked shoppers reveal that app-linked shoppers use the app for a greater number of days (Web Appendix Figure A3), longer periods of time (Web Appendix Figure A4), a higher number of app sessions (Web Appendix Figure A5), and a longer dwell time (Web Appendix Figure A6) during each session ($p < 0.001$) than do non-app-linked shoppers in the six months after app launch.

Furthermore, app-linked shoppers use a greater number of app features across sessions. However, within a session, the number of features app-linked shoppers use does not significantly differ from those used by non-app-linked shoppers (Web Appendix Figure A7).⁹ Extended use of features over a longer period of time might possibly explain subsequent purchases by app-linked shoppers better than the diversity of feature use within a session does. An alternative explanation can be that the number of sessions is correlated with purchases—those who shop more also use the app more times.

6.2. Geographic App Use–Purchase Linkages

Of the 30,026 app adopters who adopted the app in the first six months, 9,940 enabled location tracking in the app. Their median (mean) distance from the store of purchase when they access the app is 3.78 (22.39) miles. Moreover, 20.05% of their transactions take place after app use within a mile of the store of purchase, with 65.36% of the transactions occurring after app use within five miles. App-linked shoppers (average of 21.62 miles) tend to use the app much closer to stores ($p < 0.001$) than do non-app-linked shoppers (average of 30.96 miles).

6.3. Feature Use–Purchase Linkages

A comparison of features used by app-linked and non-app-linked shoppers shows that (1) within a session, app-linked shoppers exhibit a higher frequency of use of features relating to the offer list, product availability check, and call store ($p < 0.001$), and (2) across sessions, app-linked shoppers display a higher frequency of use of features relating to the product search, product details, and rewards landing page ($p < 0.001$).

Additional analyses (see Web Appendix B) show that app adopters also purchase a more diverse set of items (including less popular products) than do nonadopters and tend to browse product and rewards information in the app prior to making purchases. App adopters return products originally purchased in stores and through loyalty rewards more than those purchased online and without rewards.

7. Managerial Implications, Conclusion, and Limitations

Our results offer key managerial implications. First, on the basis of the difference in net monetary value of purchases for app adopters relative to nonadopters from the propensity score–matched regression model with selection correction (\$21.11) and the most conservative estimate from the robustness checks (\$10.21) across the retailer’s 55,580 adopters, we estimate that the retailer’s net annual revenue increase after app introduction is approximately \$6.8 million–\$14.1 million from these adopters.

Second, the finding that app adopters’ purchases expand in both the online and offline channels suggests that managers should plan for app adopters visiting both the physical and online stores more often. Managers should expect more purchases from shoppers based on the time, location, and features used in the app of the shoppers. As a result, integrating individual-level shopper insights from the app into in-store and online experiences may help create additional value in these channels. Because shoppers are likely to access product information through the app, store associates can go beyond offering the standard product information.

In conclusion, our analyses offer several key insights that demonstrate novel results and extend extant research. First, mobile app adopters have a higher frequency, quantity, and monetary value of both purchases and returns relative to nonadopters. App adopters buy 33% more frequently, buy 34% more items, and spend 37% more compared with nonadopters in the period after app introduction. At the same time, they return 35% more frequently, 35% more items, and 41% more in dollar value. Overall, app adoption is associated with a 36% increase in net monetary value of purchases across all channels. Second, both online and offline purchases are higher for

app adopters than for nonadopters. Third, the time, location, and features of app use provide descriptive evidence of how the app aids shopping in different channels. Specifically, app-linked shoppers (those who make a purchase within 48 hours of app use) are those who use the app when they are close to the store of purchase and access the app for rewards, product details, and notifications.

Although our research is the first to quantify and explain the differences between app adopters and nonadopters of a retailer’s mobile app for a broad range of shopping outcomes, including returns, it has some limitations. First, our research relies on non-experimental variation in the data. Randomized field experiments, if feasible, could provide future researchers with unique opportunities for testing specific app-related manipulations. Second, our results capture the effects on app adopters who use the app at least once—not those who downloaded but never used the app. Third, the number of cell towers serves as a good source of exogenous variation in app adoption because it has little correlation with demand factors in our data. This exogeneity assumption or exclusion restriction will fall down in locations where the number of cell towers and population are highly correlated. Despite our careful analysis, if there are any other unobservable confounds (especially if correlated with the number of cell towers), we need to be cautious about a causal interpretation of our estimates. Furthermore, our estimates are more representative of a LATE than an ATE for the population. They are particularly relevant for app adopters whose decision to adopt the app is affected by the number of cell towers in the zip code (Sudhir and Talukdar 2015). Fourth, although we have examined the net monetary value of purchases, our data do not contain cost information. If cost data are available, it would be interesting to study the effect of app introduction on customer lifetime value. Fifth, we have data from only one retailer across channels. Future studies on mobile apps can examine data on multiple retailers, including pure play retailers with a growing brick-and-mortar presence (e.g., Warby Parker, Bonobos). Finally, future research can also examine the marketing mix strategies for improving adoption of and engagement through apps.

Acknowledgments

The authors thank participants at the Theory and Practice in Marketing (TPM) Conference and the Professors’ Institute meeting of Marketing EDGE, as well as Sridhar Narayanan, Harikesh Nair, Hema Yoganarasimhan, S. Sriram, and Raj Venkatesan for valuable comments.

Endnotes

¹ In subsequent robustness checks (Web Appendix Tables A1 and A2), we use an alternative sample of 50,000 app adopters and obtain similar results.

²The findings of subsequent empirical analysis apply to app adopters defined as those who download *and* use the app, not those who download but never use the app. This is reasonable because the latter are effectively nonadopters.

³We report the pre-period trends for the unmatched samples (raw data) in Web Appendix Figure A2. Together with Figure 1, it demonstrates that matching improves the pre-period trends, mitigating concerns for unobservables.

⁴The limits and scope of our exclusion restriction are similar in spirit to those of Jacobi and Sovinsky's (2016). They estimate marijuana demand using living in a city and temperature as exclusionary variables.

⁵Despite the thorough evaluation of the exclusion restriction and other falsification checks, we cannot completely rule out the possibility that there could be other unobserved factors not addressed by the number of cell towers variable (Goldfarb and Tucker 2011). If there is any such unobserved factor, we need to be cautious in interpreting our estimates as causal (Jacobi and Sovinsky 2016).

⁶We calculate the percentage change as the treatment effect relative to matched pre period control group's level of the variable reported in Table 3. For example, for NMV, it is (21.11/57.85), or 36.49%.

⁷The results from this model are similar for online and offline purchases, as reported in Web Appendix Table A7.

⁸In this analysis, we examine the 6-month period after app launch and consider adopters in this period (54% of all adopters in the main sample) to analyze the changes close to the app launch date for explaining the short-term differences between app-linked and non-app-linked shoppers. We subsequently considered different time periods and found similar results.

⁹We compute the within-session number of features used by calculating the average number of features a shopper clicks in the app in a given session (shopper-session level), and we calculate the across-session number of features through the cumulative sum of the number of features a shopper clicks across sessions in the six months after introduction. Therefore, this number can exceed the total number of features available in the app.

References

- Altonji JG, Elder TE, Taber CR (2005) Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *J. Political Econom.* 113(1):151–184.
- Avery J, Steenburgh T, Deighton J, Caravella M (2012) Adding bricks to clicks: Predicting the patterns of cross-channel elasticities over time. *J. Marketing* 76(3):96–111.
- Bell D, Gallino S, Moreno A (2018) Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Sci.* 64(4):1629–1651.
- Bellman S, Potter RF, Treleaven-Hassard S, Robinson JA, Varan D (2011) The effectiveness of branded mobile phone apps. *J. Interactive Marketing* 25(4):191–200.
- Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust difference-in-difference estimates? *Quart. J. Econom.* 119(1):249–275.
- Boyd ED, Kannan PK, Slotegraaf RJ (2019) Branded apps and their impact on firm value: A design perspective. *J. Marketing Res.* 56(1):76–88.
- Bronnenberg BJ, Dubé JPH, Gentzkow M (2012) The evolution of brand preferences: Evidence from consumer migration. *Amer. Econom. Rev.* 102(6):2472–2508.
- CNBC (2018) The cutting edge of customer experience is... apps? *CNBC.com* (July 6), <https://tinyurl.com/y8o4ukuv>.
- Comscore (2016) The 2016 U.S. mobile app report. White paper, Comscore, Reston, VA. <http://tinyurl.com/j43tjyw>.
- Danaher PJ, Smith MS, Ranasinghe K, Danaher TS (2015) Where, when, and how long: Factors that influence the redemption of mobile phone coupons. *J. Marketing Res.* 52(5):710–725.
- Dubé J-P, Fang Z, Fong N, Luo X (2017) Competitive price targeting with smartphone coupons. *Marketing Sci.* 36(6):944–975.
- Einav L, Levin J, Popov I, Sundaresan N (2014) Growth, adoption, and use of mobile e-commerce. *Amer. Econom. Rev.* 104(5):489–494.
- eMarketer (2017) eMarketer unveils new estimates for mobile app usage. *eMarketer* (April 11), <https://tinyurl.com/ydyz4hsm>.
- Fong NM, Fang Z, Luo X (2015) Geo-conquesting: Competitive locational targeting of mobile promotions. *J. Marketing Res.* 52(5):726–735.
- Forbes (2015) How mobile ordering can impact Starbucks' valuation. (October 30), <http://tinyurl.com/zrekrwp>.
- Gill M, Sridhar S, Grewal R (2017) Return on engagement initiatives: A study of a business-to-business mobile app. *J. Marketing* 81(4):45–66.
- Goldfarb A, Tucker C (2011) Privacy regulation and online advertising. *Management Sci.* 57(1):57–71.
- Goldfarb A, Tucker C (2014) Conducting research with quasi-experiments: A guide for marketers. Working paper, University of Toronto, Toronto.
- Gomez MI, McLaughlin EW, Wittink DR (2004) Customer satisfaction and retail sales performance: An empirical investigation. *J. Retailing* 80(4):265–278.
- Guo S, Fraser MW (2014) *Propensity Score Analysis: Statistical Methods and Applications* (Sage Publications, Thousand Oaks, CA).
- Heckman J (1979) Sample selection bias as a specification error. *Econometrica* 47(1):153–161.
- Huang Q, Nijss VR, Hansen K, Anderson ET (2012) Wal-Mart's impact on supplier profits. *J. Marketing Res.* 49(2):131–143.
- Hung SY, Ku CY, Chang CM (2003) Critical factors of WAP services adoption: An empirical study. *Electronic Commerce Res. Appl.* 2(1):42–60.
- Hyken S (2017) Walgreens: At the corner of technology and a better customer experience (CX). *Forbes* (July 8), <http://tinyurl.com/y96uddk4>.
- Jacobi L, Sovinsky M (2016) Marijuana on Main Street? Estimating demand in markets with limited access. *Amer. Econom. Rev.* 106(8):2009–2045.
- Kim SJ, Wang RJ-H, Malthouse EC (2015) The effects of adopting and using a brand's mobile application on customers' subsequent purchase behavior. *J. Interactive Marketing* 31(August):28–41.
- Lee J-Y, Zhuang M, Kozlenkova IV, Fang E (2016) The dark side of mobile channel expansion strategies. MSI Working Paper 16-119, Marketing Science Institute, Cambridge, MA.
- Lundborg P, Plug E, Rasmussen AW (2017) Can women have children and a career? IV evidence from IVF treatments. *Amer. Econom. Rev.* 107(6):1611–1637.
- Manchanda P, Packard G, Pattabhiramaiah A (2015) Social dollars: The economic impact of customer participation in a firm-sponsored online customer community. *Marketing Sci.* 34(3):367–387.
- Minotti M (2016) Video games will become a \$99.6B industry this year as mobile overtakes consoles and PCs. *Venture Beat* (April 21), <http://tinyurl.com/y7favjyc>.
- Pew Research Center (2018) Mobile fact sheet. Accessed February 5, 2018, <http://www.pewinternet.org/fact-sheet/mobile/>.
- Rosenbaum PR (2005) Sensitivity analysis in observational studies. Everitt BS, Howell DC, eds. *Encyclopedia of Statistics in Behavioral Science*, vol. 4 (John Wiley & Sons, New York), 1809–1814.
- Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41–55.
- Rubin DB (2008) For objective causal inference, design trumps analysis. *Ann. Appl. Statist.* 2(3):808–840.
- Shankar V, Balasubramanian S (2009) Mobile marketing: A synthesis and prognosis. *J. Interactive Marketing* 23(2):118–129.

- Shankar V, Kleijnen M, Ramanathan S, Rizley R, Holland S, Morrissey S (2016) Mobile shopper marketing: Key issues, current insights, and future research avenues. *J. Interactive Marketing* 34(May):37–48.
- Silverman BW (1986) *Density Estimation for Statistics and Data Analysis* (Chapman & Hall, London).
- Statista (2015) Share of console gamers in the United States who play online and offline games as of April 2015, by platform. Accessed November 27, 2018, <https://www.statista.com/statistics/420596/us-consumer-share-playing-console-games-by-type/>.
- Sudhir K, Talukdar D (2015) The “Peter Pan syndrome” in emerging markets: The productivity-transparency trade-off in IT adoption. *Marketing Sci.* 34(4):500–521.
- Tirunillai S, Tellis G (2017) Does offline TV advertising affect online chatter? Quasi-experimental analysis using synthetic control. *Marketing Sci.* 36(6):862–878.
- Wang K, Goldfarb A (2017) Can offline stores drive online sales? *J. Marketing Res.* 54(5):706–719.
- Wang RJ-H, Malthouse EC, Krishnamurthi L (2015) On the go: How mobile shopping affects customer purchase behavior. *J. Retailing* 91(2):217–234.
- Wangenheim FV, Bayón T (2007) Behavioral consequences of over-booking service capacity. *J. Marketing* 71(4):36–47.
- Williams R (2018) App Annie: Shopping app use jumped 54% in 2017. *Retail Dive* (January 12), <https://tinyurl.com/ycgnyj8o>.
- Wooldridge JM (2002) *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge, MA).
- Xu K, Chan J, Ghose A, Han SP (2016) Battle of the channels: The impact of tablets on digital commerce. *Management Sci.* 63(5):1469–1492.
- Xu J, Forman C, Kim JB, Ittersum KV (2014) News media channels: Complements or substitutes? Evidence from mobile phone usage. *J. Marketing* 78(4):97–112.