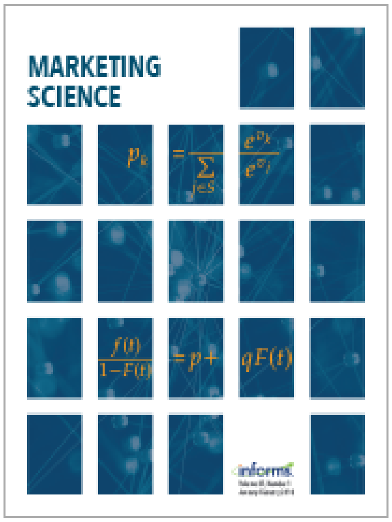


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Search and Learning at a Daily Deals Website

Mantian (Mandy) Hu, Chu (Ivy) Dang, Pradeep K. Chintagunta

MARKETING SCIENCE
Vol. 38, No. 4, July–August 2019, pp. 609–642
ISSN 0732-2399 (print), ISSN 1526-548X (online)
http://pubsonline.informs.org/journal/mksc/

Search and Learning at a Daily Deals Website

Mantian (Mandy) Hu,[a] Chu (Ivy) Dang,[a] Pradeep K. Chintagunta[b]

[a] CUHK Business School, The Chinese University of Hong Kong, Hong Kong; [b] Booth School of Business, University of Chicago, Chicago, Illinois 60637

Contact: mandyhu@baf.cuhk.edu.hk, http://orcid.org/0000-0001-6168-4600 (MH); dangchu@baf.cuhk.edu.hk (CD); pradeep.chintagunta@chicagobooth.edu (PKC)

Received: September 9, 2017
Revised: September 13, 2018; December 29, 2018
Accepted: January 21, 2019
Published Online in Articles in Advance: July 8, 2019

https://doi.org/10.1287/mksc.2019.1156

Abstract. We study consumers' purchase behavior on daily deal websites (e.g., Groupon promotions) using individual clickstream data on the browsing history of new subscribers to Groupon between January and March 2011. We observe two patterns in the data. First, the probability that a given consumer clicks on a merchant in the emailed newsletter declines over time, which seems to be consistent with the notion of consumer "fatigue"—a phenomenon highlighted by the popular press. Second, the probability that the consumer makes a purchase conditional on clicking increases over time, which seems contrary to the notion of "fatigue." To reconcile these two observations, we propose a model that rationalizes these patterns and then use it to provide insights for companies in the daily deal industry. When consumers first subscribe to a daily deal website, they are unlikely to be fully informed about the quality of the deals offered on that site. The daily newsletter provides only the price and some limited information about that day's featured deal. To learn more about quality, consumers need to click on the emailed newsletter; this takes them to the deal's website, where they invest time and effort to learn about the deal's quality. Such a search for information is costly. Furthermore, consumers do not know about the quality of deals they may receive in the future. Given the cost of searching and the uncertainty about the quality of future deals, consumers are more likely to search early on (i.e., click on the newsletter) in their tenure. As they learn about the distribution of the quality of deals on Groupon, they require less searching, resulting in a decline in clicks over time. As learning accumulates, consumers are better at recognizing the position of a deal in the quality distribution of Groupon deals and are therefore more likely to purchase the clicked deals. This results in an increase in the conditional probability of purchasing. We formulate a dynamic model of search and Dirichlet learning based on the above characterization of consumer behavior. We show that the model is able to replicate patterns in the data. Next, we estimate the parameters of the model and provide insights for managers of daily deal websites based on our findings and policy simulations.

History: Avi Goldfarb served as the senior editor and Anja Lambrecht served as associate editor for this article.
Funding: Financial support was received from the Research Grants Council of the Hong Kong SAR, University Grants Committee [10.13039/501100002920, ref. no. CUHK24500214]; and the Kilts Center for Marketing at the Booth School of Business (PKC).
Supplemental Material: Data and the online appendix are available at https://doi.org/10.1287/mksc.2019.1156.

Keywords: learning model • Dirichlet updating • dynamic search model • deep learning

1. Introduction

Daily deal websites offer discounts on goods and services for a given period (typically 24 to 36 hours) and have attracted considerable attention in the global marketplace in recent years. In 2015 the revenues from daily deal sites in the United States reached $4 billion, up from $2 billion in 2011, with 329 businesses in the industry employing 20,757 people (IBISWorlds Daily Deals Sites market research report).[1] Roughly 4 in 10 American adults use online deal websites, according to a survey by YouGov in 2014.[2] Groupon is currently the dominant player in the daily deal industry. By the third quarter of 2016, it had approximately 50.8 million active users worldwide. Groupon works with more than 1 million merchants[3] and was generating approximately $3.1 billion in revenue by the end of 2015.[4] However, despite the company's fast growth in its early years, the industry started to suffer from what the media calls "Daily Deal Fatigue" or "Groupon Fatigue" (Dholakia and Kimes 2011).[5] "Fatigue," as used by these media sources, describes a phenomenon in which consumers become tired of the deals on these sites, which then experience a decline in the number of clicks on their daily deals. Such a concern, if valid, challenges the sustainability of the daily deal business model. The concern is so widespread that some have even predicted the death of daily deals and daily deal websites.[6] The issue of fatigue, however, goes beyond the daily deal phenomenon and has also been observed in other contexts. Indeed, from the perspective of the popular press, businesses built on similar models, such as online dating websites (e.g., Tinder and OKCupid)[7] and flash sales websites (e.g., Gilt), seem to be suffering from the same issue.[8]

Before describing the study, it is useful to outline how daily deal websites such as Groupon operate. Upon registration, new subscribers receive an emailed newsletter every day from the website [as shown in Figure 1(a)]. That email contains one featured deal of the day, with a captioned subject line and more information in the email body. The deals provided usually have a short expiry window (typically ranging from one to three days). Subscribers thus face a trade-off between purchasing the current deal and waiting for another ("better") one. A consumer receiving the newsletter has to decide whether (s)he wants to know more about the deal by clicking on the newsletter and getting redirected to the deal's web page [as shown in Figure 1(b)], which requires the consumer to expend time and/or effort gathering information. Once on the website, consumers receive more information about the deal itself, such as the number of people who have already purchased it that day, the time left before it expires, and whether the deal has yet "tipped."[9] This information helps the consumer assess the "quality" of the deal. The consumer then decides whether to make a purchase. Although new subscribers may not know initially what to expect in terms of the quality of the deals offered by the merchants through the site, over time visits to the deal webpages help consumers better understand the (distribution of) quality of those merchants and deals. An important feature of this process is that to obtain information about the deal, the consumer needs to expend some effort to visit the deal's web page.

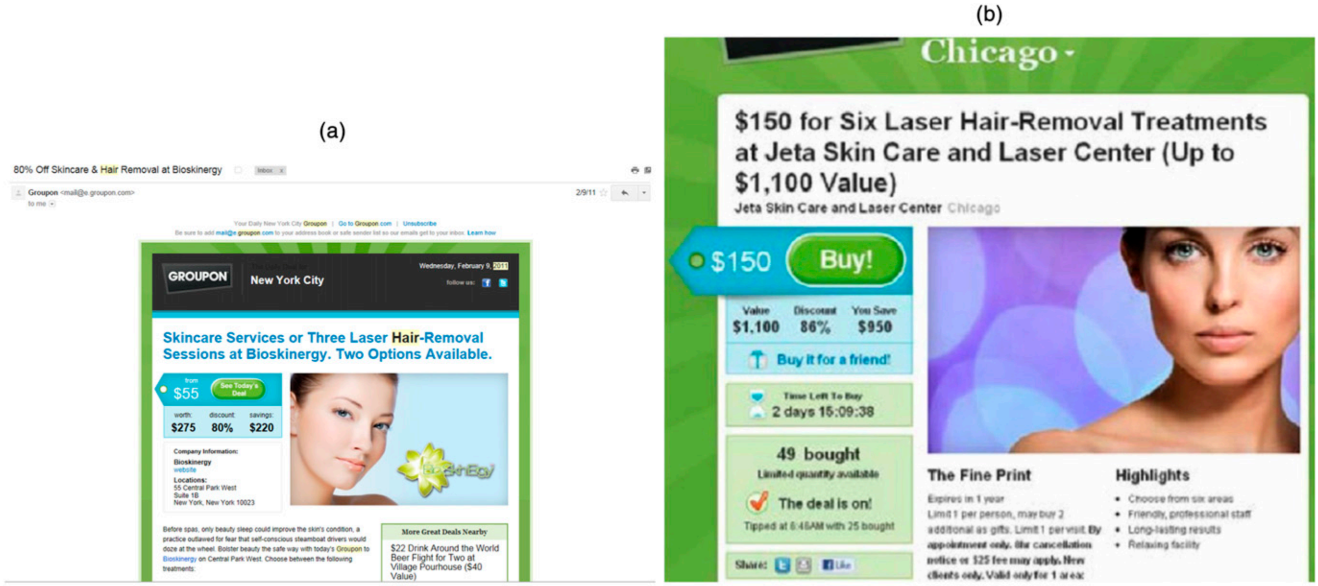
As previously noted, the term "Groupon fatigue" or "daily deal fatigue" has been used by the mass media in conjunction with waning interest in daily deals. Although the press does not formally define these terms, there are two possibilities that could lead to the phenomenon of declining clicks that the terms describe: (1) "within"-consumer behavioral changes over time—that is, a given consumer becomes less likely to click on deals over time); and (2) "across"-consumer changes over time—that is, customers leave the site or fewer subscribe, such that the number of active consumers declines over time, or it is also possible that new subscribers are less interested in the deals than are previous cohorts. Our focus in this paper is on the changes in "within"-customer behavior over time, because our data window is not large enough to draw conclusions about the evolving customer base. If there is indeed evidence of waning interest in Groupon, as evidenced in our case by a given consumer's clicking less over time, then the industry needs to formulate ways to overcome this issue. Despite its importance for the very survival of the industry, consumer behaviors on such sites have not been investigated much in the literature. Dholakia and Kimes (2011), for example, surveyed consumer perceptions of daily deals and found no evidence of the fatigue phenomenon.[10]

Our main objective in this paper is to provide empirical insight into how newly acquired customers behave on daily deal websites and to examine the empirical validity of concerns about within-consumer deal fatigue among new subscribers. On the basis of what we observe in the data, we propose an empirical model to characterize such behavior. This model helps us evaluate alternative designs for the consumer experience, such as those that increase the effort required for consumers to gather information about the featured deal (i.e., sending out newsletters with featured deals from more varied categories). This model also provides insight into the approaches that daily deal websites may take to further encourage purchases.

We obtained our data from a third-party online research firm in the United States. This proprietary data set consists of the complete clickstream history within the browsing sessions of people who newly subscribed to the Groupon service between January and March 2011. Thus, we observe consumers from the beginning of their associations with Groupon. The data reveal two patterns. First, as noted in the popular press, we find that the probability that a consumer clicks on a merchant in the emailed newsletter indeed declines over time. This is consistent with the notion of within-consumer "fatigue." By itself, this finding does not bode well for Groupon. After all, if consumers' interest is declining, then the site's expected revenue from its subscriber base is sure to be affected.[11] The second pattern revealed by the data, however, is a potential source of optimism for Groupon. In particular, we find that the probability that a consumer makes a purchase, conditional on clicking, actually increases over time. Together, these patterns suggest that even though consumers become more selective in terms of exploring the offers received, they are more likely to yield revenue for the site as time passes.

Our proposed structural model of search and learning tries to explain the empirical within-consumer patterns seen over time in both a decline in click-throughs and an increased probability of purchasing conditional on clicking. We explain the feature of daily deal websites through which consumers are forced to expend effort to obtain information on a deal and to purchase it as a "search cost." Unlike traditional search models (Weitzman 1979), whereby consumers actively seek out information on a number of alternatives to resolve their uncertainties before making a choice among

Figure 1. (Color online) Example of Groupon Newsletter and Web Page in Early 2011



Groupon Newsletter in early 2011 Groupon Webpage in early 2011

various alternatives, in our case, consumers evaluate whether to purchase the specific deal offered each day via the daily newsletter. Furthermore, the next specific deal that the consumer encounters is not within the consumer's control—it is simply the deal offered in the newsletter the following day that may or may not be relevant to the recipient. Thus the search process in this context can be construed as being "passive" rather than an active search for the next deal. Although the presence of a search cost enables us to explain whether the consumer clicks on the newsletter to go to the deal's web page, it does not explain the decline in clicks over time. To rationalize this behavior, we postulate that new subscribers to a site are uncertain about the nature of the deals available to them. By clicking on the newsletter and visiting the web page of the deal, they are able to (partially) resolve this uncertainty over the "quality" of the deals on the site. Over time, they learn about the deals that the site offers.[12] This resolution of uncertainty, combined with the cost of obtaining deal information, results in a pattern of declining clicks (or "deal fatigue"). As consumers become better informed about the quality of deals, they also become better at choosing which deals are of higher quality. Thus, consumers are more likely to purchase the clicked deals as they resolve their uncertainty regarding the position of a deal in the quality distribution of Groupon deals.[13] This behavior contributes to the increased probability of a purchase conditional on clicking that we observe in the data. The "waiting to learn" motivation makes our model inherently forward-looking.

Because the quality of deals available on Groupon is central to our understanding of users' behavior on the site, we require a measure of quality that captures the information that consumers are able to observe. To construct this measure, we use deep learning techniques applied to the text and image data on deals gathered from the Groupon website (refer to Section 3 for details). In particular, we use the following: (a) deal attributes summarized by Groupon (e.g., the size of the discount, the original price, the length of the redemption period, and whether the deal is "tipped" or not); (b) text information on the web page (e.g., deal title, deal description, fine prints, highlights); (c) image information (e.g., the size of the picture, the color of the picture, and whether there is a picture of a human face, a picture of the product, or a brand logo). We first use subjects from Amazon Mechanical Turk (MTurk) to rate randomly assigned webpages of a subset of deals on a scale from 1 to 9. Then we associate the above features of the deals with the measure of quality obtained from the MTurk subjects by training a convolutional neural network (CNN). We then use this algorithm to help us project deal meta-information onto the quality metric for all of the remaining deals in our data. Our measure of quality therefore captures all of the observable information on the web page. We compare the results from our quality metric with those from an alternative measure adapted from Liu et al. (2017).

On the basis of the model estimates, we replicate the two patterns observed in the data. The consumers who are uncertain about the quality of future deals face a trade-off between clicking on the deal and purchasing it now and waiting for the next deal. These consumers tend to click on more deals at the beginning of their subscription tenure to learn about the quality distribution of deals. As their knowledge accumulates their incentive to learn declines, and they click on fewer deals. As the consumers resolve their uncertainty over the position of the deal in the deal quality distribution, they are more likely to purchase the deal they click on. This results in the pattern of increasing purchase probability. We find that learning is a key force driving the patterns in our setup; the model without learning fails to replicate these patterns. With the model in place, we then turn to counterfactuals that may be of interest to the firm. Because the two main features of our paper are search and learning, we provide counterfactual analyses corresponding to each process. In particular, we investigate the roles of search costs and learning in influencing the company's revenues. Eliminating search costs can save consumers the time needed to learn about the true quality distribution. However, given that subscribers in our data begin with high prior beliefs about deal quality on Groupon, a strategy that delays learning about the true quality could benefit the site. In such a situation we also find that the site can generate more revenue by changing consumers' prior quality beliefs or by increasing the variance in the quality distribution of the deals.

The rest of this paper is organized as follows: In Section 2, we review the literature. Section 3 describes the data collection and the construction of our measure of deal quality. Section 4 describes the two behavioral trends observed in the data and provides explanations. Section 5 provides the reduced-form evidence for learning. Section 6 presents the model. The empirical application is found in Section 7. Section 8 offers managerial implications. Finally, in Section 9, we conclude and discuss future extensions of our model.

2. Literature Review

Our research mainly relates to three streams of literature—on search, on learning, and on daily deals. We summarize these streams in Table 1. In the literature on search, one area focuses on understanding consumers' search behaviors and, specifically, whether they engage in sequential (Stigler 1961, McCall 1970, Weitzman 1979, Kim et al. 2010, Chen and Yao 2016) or simultaneous searching behaviors (Honka 2014). In addition, research such as De los Santos et al. (2012) and Honka and Chintagunta (2016) also proposes identification strategies to distinguish between the two types of search. Another area focuses on search costs—how to identify them and how to quantify their impact on consumer demand (Hortaçsu and Syverson 2004, Hong and Shum 2006, Koulayev 2010, Seiler 2013). Studies such as those by Rust (1987), Song and Chintagunta (2003), Erdem et al. (2003), and Hendel and Nevo (2006) utilize dynamic structural models to account for consumers' forward-looking behavior, but search is not considered; consumers make a purchase decision by trading off purchases in the present period with waiting to obtain a higher utility in the next period. Hartmann and Nair (2010) used a two-stage search model to study intertemporal demand for tied goods, in which the decision process consists of a search stage and a purchase stage. Seiler (2013) considered consumers' price search for detergents, also under a two-stage decision process. We take Seiler's model as a benchmark, because we endogenously model consumers' clicking (or "search") decisions upon receiving a newsletter in the first stage and then the purchase decision in the second stage. Whereas the previous literature has treated the choice of which store to visit as exogenous, Mojir and Sudhir (2016) modeled the store visit decision as endogenous. Our research question focuses on the intertemporal behavior of subscribers on a daily deal website (i.e., we treat the subscription decision as given). As subscribers of a daily deal website, consumers automatically receive a newsletter every day. This alleviates the concern of the endogenous "store visit" problem in our case.

Our research also relates to the Bayesian learning models in the literature on discrete-choice demand (cf. Erdem and Keane 1996, Crawford and Shum 2005). In these studies, consumers are uncertain about some aspect of their demand functions, which are resolved over time via signals that the consumers obtain. Such learning has been modeled in a Bayesian fashion using a variety of distributional assumptions. Perhaps the most widely used is the parametric normal Bayesian updating rule (cf. Erdem and Keane 1996, Crawford and Shum 2005, Zhang 2010). For example, Zhao et al. (2013), Sriram et al. (2015), and Ching and Lim (2016), among others, incorporated learning into consumer choice models, but consumers were not assumed to be forward-looking. Erdem and Keane (1996), Crawford and Shum (2005), and Zhang (2010) embedded consumer learning within dynamic discrete-choice models.

Rothschild (1974) theorized the use of a nonparametric Bayesian method to study optimal search under two types of uncertainty: about current prices and about the underlying process that generates prices. Specifically, he usd Dirichlet priors to model a searcher's belief about the unknown price distribution and updated these in a nonparametric Bayesian fashion as new price quotes arrive. Koulayev (2013) brought the search model with Dirichlet priors into the empirical domain. He developed a novel characterization of optimal search that leads to closed-form, easily computable, ex ante probabilities of purchasing products. De Los Santos et al. (2017) took a step further by modeling nonparametric learning using more general Dirichlet process priors. Taking advantage of an

Table 1. Literature

	Without dynamic forward looking	With dynamic forward looking	
	Search model/optimal stopping	One-stage	Two-stage
No learning	Stigler (1961) McCall (1970) Weitzman (1979) Hortaçsu and Syverson (2004) Hong and Shum(2006) Kim et al. (2010) De Los Santos et al. (2012) Honka (2014) Honka and Chintagunta (2016) Chen and Yao (2016)	Rust (1987) Erdem et al. (2003) Song and Chintagunta (2003) Hendel and Nevo (2006)	Hartmann and Nair (2010) Seiler (2013) Mojir and Sudhir (2016)
Learning			
Parametric learning	Zhao et al. (2013) Ching and Lim (2016)	Erdem and Keane (1996) Crawford and Shum (2005) Zhang (2010) Chintagunta et al. (2009) Chintagunta et al (2012) Sriram et al. (2015)	
Nonparametric learning	Rothschild (1974) Koulayev (2013) De Los Santos et al. (2017)		Our model

individual-level data set containing information on consumers' search sequences, De Los Santos et al. (2017) developed a method to estimate consumer search costs for differentiated products when consumers have only partial information about the distribution that they sample from.

In this study, we embed consumer learning within a two-stage dynamic search model involving the search decision and the purchase decision. Consumers are uncertain about the distribution of the quality of the deals on Groupon and learn about this distribution by searching and gathering related information through the various deals. We use a nonparametric Dirichlet updating rule by which consumers update their beliefs about quality distribution. The Dirichlet prior places fewer restrictions than does the typically assumed normal distribution; the updating rule is simple to implement and quick to compute; and in our context, consumers learn about the distribution of deal quality rather than the true quality of the product because, in this case, the product changes over time and the same deal does not repeat in the data.

Our research contributes not only to the literature on search and learning but also to the literature on daily deals. Dholakia and his colleagues wrote a series of reports on the daily deal industry (cf. Dholakia 2010, Dholakia 2011, Dholakia 2012, and Dholakia and Kimes 2011). Dholakia (2010) and Dholakia (2012) surveyed the business side of the market and found no evidence of deterioration for small and medium-sized business in the performance of daily deal promotions over the surveyed year. Dholakia and Kimes (2011) surveyed consumer perceptions on daily deals and found no evidence of daily deal fatigue. Dholakia and Kimes (2011) developed a conceptual framework specifying the determinants of a profitable Groupon promotion and empirically tested it using their surveyed data. Wu et al. (2015) quantified the economic value of daily deals for merchants by proposing a structural model, empirically confirming the long-run profitability of the business model for merchants. Luo et al. (2014) investigated the effect of deal popularity and social influence-related factors on purchase likelihood and redemption time using a data set of 30,272 customers of a group-buying website. Hu and Winer (2016) tested the effects of the "tipping point" on Groupon using augmented clickstream data and found that the tipping point can alter consumers' decisions and affect sales. Our research also uses augmented clickstream data in which we observe the detailed click history of individual consumers, but we focus only on newly registered customers and their behaviors. Our study is the first to structurally model the behaviors of new subscribers on daily deal websites with the objective of examining the empirical validity of concerns over deal fatigue.

3. Preliminary Data Analysis and Quality Measure Construction

The data come from a third-party, U.S.-based research firm that tracks the online browsing behavior of a sample of more than two million people in the United States by collecting their clickstream data. For new Groupon users who subscribed between January and

March 2011, we extracted the complete clickstream data within their browsing sessions whenever they logged on to the Groupon website. Thus, we are able to observe consumers from the beginning of their association with Groupon. Our data contain, but are not limited to, individuals' entire history on Groupon, such as their subscription date, their click and view history, and their purchase behaviors with the associated time stamps. Furthermore, on the basis of this clickstream information, we are able to trace back the original deal webpages; this allows us to crawl the site for detailed information on Groupon deals. Overall, the data set contains 26,523 records, 14,096 featured deals, and 10,951 new subscribers who visited Groupon at least once during the sample period. The deals cover 46 states and 171 cities in the United States and contain a variety of categories, such as restaurants, arts and entertainment, shopping, and beauty and spas. The detailed statistics of the data set are shown in Table 2.

3.1. Assembling the Data for Analysis

3.1.1. Step 1: Gathering Information on the Clicked Deals. Each Groupon subscriber receives a daily newsletter via email with a featured deal of the day and several less prominently placed deals. The unit of observation for our analysis is the individual day. The consumer can choose to click on a deal offered in the newsletter; hence, our data contain both instances—when there is a click and when there is no click on a newsletter deal. If a click does occur, we then observe whether the consumer purchases the deal. By observing the consumer's clickstream data, we can infer whether (s)he clicked on a newsletter deal and arrived at the deal's page on Groupon. An example of a uniform resource locator (URL) from Groupon's clickstream data when the customer clicks the web page of a Groupon deal is as follows:

$www.groupon.com/san-diego/deals/jc-golf-san-diego?utm\_campaign = jc-golf-san-diego\_\&emailutm\_medium = email\&utm\_source = newsletter\&c = title\&addx = {*}{*}{*}{*}19 \& utm\_content = sandiego\_feed \& user = {*}{*}{*}{*}@{*}{*}{*}{*}.com \& d = deal \& s = more\_deals\_for\_you \& p = {*}{*}{*}{*}1 \& divison = {*}{*}{*}{*}9 \& date = 20110315\&fbxdfragment.$

Table 2. Raw Data Summary

Data	Summary
Covered time period	January 1, 2011 to March 31, 2011
No. of records	26,523
Average record size	180 bytes
New subscribers within three months	10,951
Repeated visitors within three months	9,611
Distinct featured deals	14,096

From the embedded query strings in the URL, we can obtain information about the deal, the city to which it is targeted, and whether it comes from the emailed newsletter or other sources such as search engines. The query strings also reveal whether the deal is featured in the newsletter that day by the value of utm campaign (in the example above, the deal "jc-golf-san-diego" is the featured deal). Our clickstream data also reveal whether the consumer purchased the deal that was clicked on.

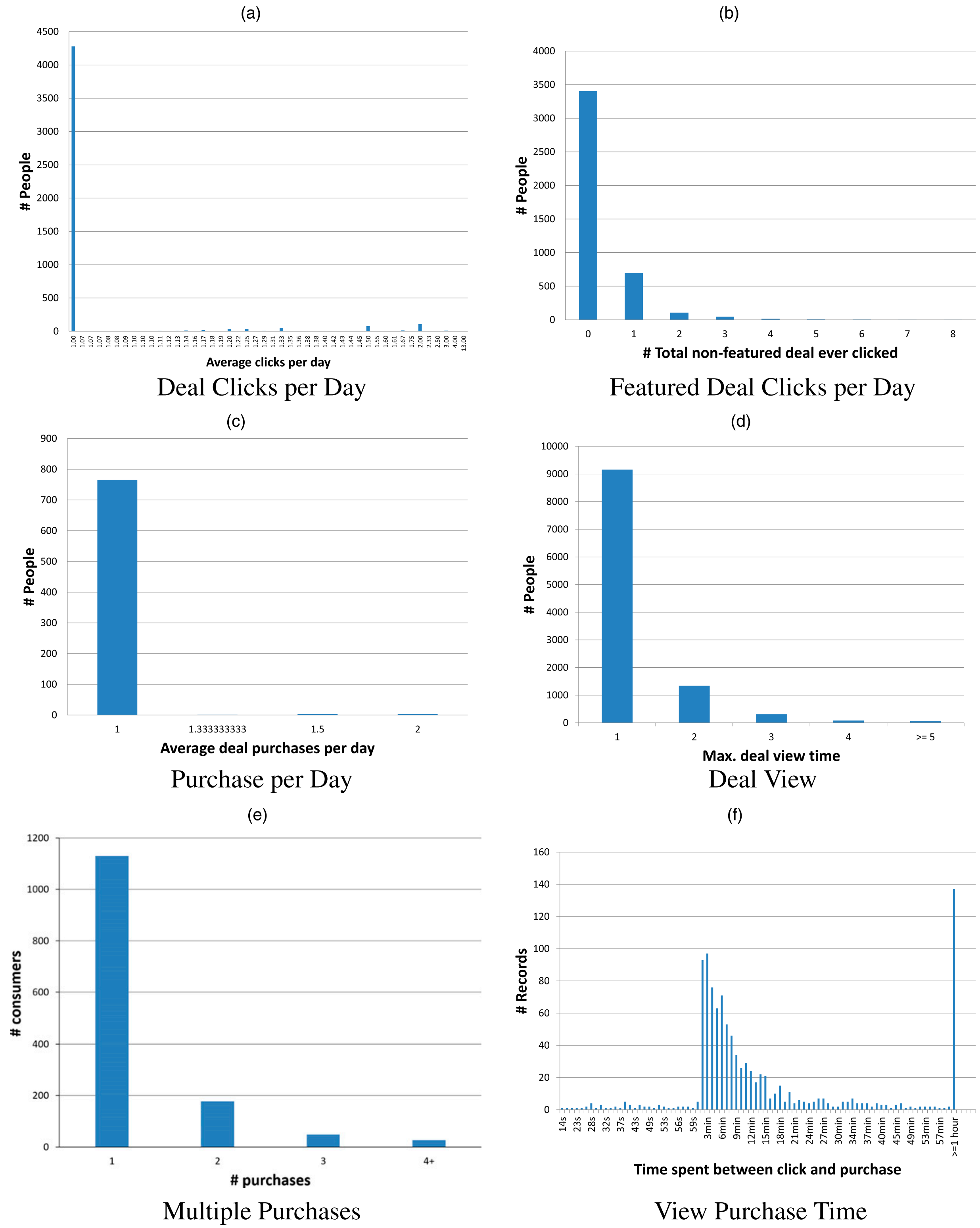
3.1.2. Step 2: Focusing on the Featured Deals. Next, we examine the data on the clicked deals gathered in the first step. We note the following. First, throughout the sampling period, 90% of subscribers clicked on only one newsletter deal on a given day [Figure 2(a)]. Second, among those who clicked on only one newsletter deal per day, 79.50% only clicked on the featured deal throughout the entire sampling period [Figure 2(b)]. In a typical Groupon newsletter in 2011 [Figure 1(a)], the main body is occupied by the featured deal, whereas the sidebar includes several additional deals. Of the subscribers who clicked on newsletter deals (including those who clicked on more than one newsletter deal per day), 82.66% only clicked on the featured deal. Combining this with the first observation, we narrow our focus to subscribers' daily decision on whether to click on the featured deal.

Third, 99.09% of the subscribers in the data purchased at most one deal per day [Figure 2(c)], whereas 93.66% of those who made a purchase only purchased a featured deal in the course of the sampling period. Given these observations, we focus only on subscribers' purchase decisions regarding the featured deal from the newsletter on a specific day.

Fourth, 83.61% of the subscribers only checked a deal once [Figure 2(d)], whereas 89.86% of deals were checked only once by a given subscriber. We do not observe purchases of previously clicked deals. This shows that for subscribers, the purchase of Groupon deals is a sequential decision and there is no recall.

3.1.3. Step 3: Augmenting the Data for Each Consumer with Nonclicked Deals. How then do we observe the deals that a consumer did not click through on a particular day? We infer this from the deals that other subscribers click on (both existing and new) in that city on the same day. One caveat to this in our data is that there are days in certain cities during which no subscriber clicked on the Groupon deal. However, such cases are rare in the data. Figure 3 shows the distribution of nonmissing days by cities ranked by number of new subscribers. We rank all of the cities according to the total number of new subscribers and place these on the x axis. The cities on the left-hand side have more new subscribers than the cities on the

Figure 2. (Color online) Empirical Data Analysis



right. The y axis represents the number of days that we observe any click of the featured deal (max = 90, mean = 87 days, sd = 6 days). We use this plot to illustrate that missing values do not pose a significant problem (note that our data span a three-month period). When we do encounter a missing value, we replace it with the unexpired featured deal of the previous day.

3.1.4. Step 4: Considering Other Data Issues. A potential concern is when an email inbox such as Gmail (Google's free webmail service) automatically classifies the Groupon newsletter as spam; in such cases, the subscriber does not see the newsletter at all. By implication, the subscriber is not skipping the deal intentionally. Unfortunately, our data do not discern this type of instance. However, owing to the way that clickstream data are collected, individuals who have logged onto Groupon at least once during the three-month collection period are included in the sample; as such, it is unlikely that the individuals whose Groupon newsletter is classified as spam are included in the sample (unless the classification happens partway through the data period). Our results hold for all individuals who normally receive the Groupon newsletter on a daily basis.

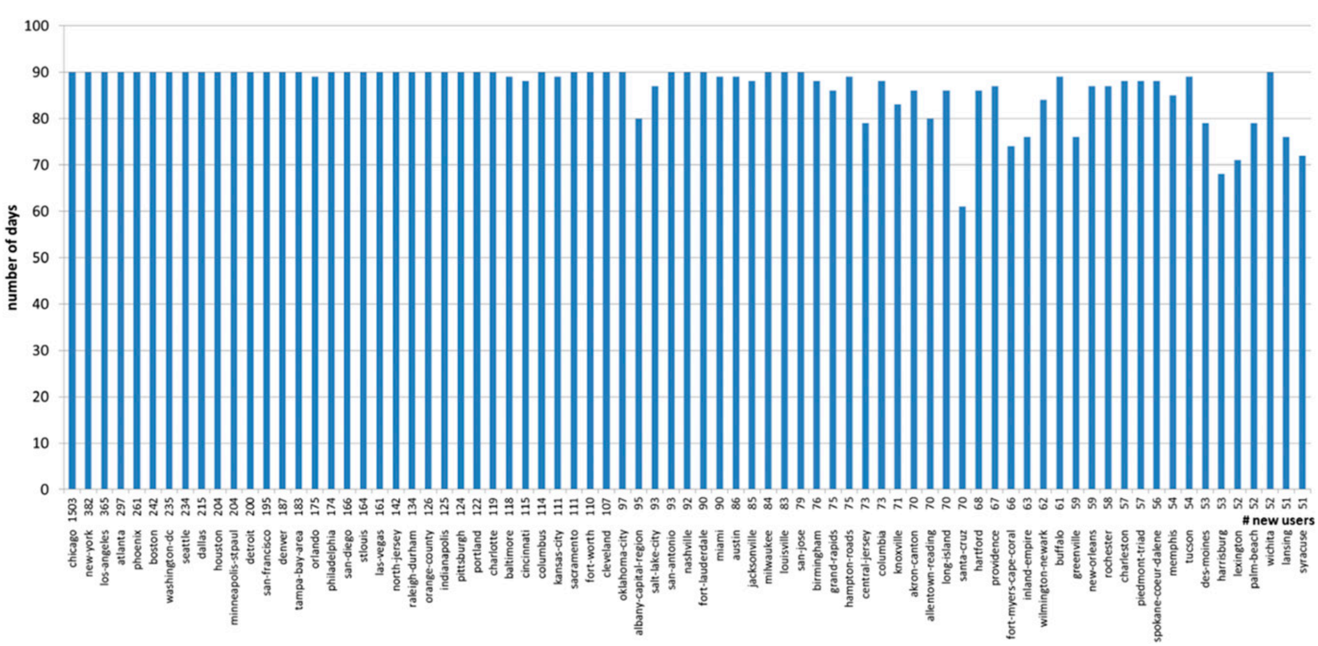
Another issue is that we have three months of data on subscribers who joined Groupon beginning in January 2011. We assume that these consumers do not drop out of our sample. If a consumer does not click, we model it as if they skipped the deal of that day.

3.1.5. Step 5: Data Used in the Analysis. When the data assembling process is complete, we have gathered information on a sample of consumers who (i) received a newsletter each day; (ii) decided whether to click on the featured deal; and (iii) conditional on clicking, chose whether to purchase the deal. We organize the data as follows: individual, day, featured deal, deal characteristics, consumer's clicking decision, consumer's purchase decision, and so forth.

3.2. Constructing a Measure of Deal Quality

Quality can be defined as superiority or excellence (Zeithaml 1988). Perceived quality is a perceptual, conditional, and somewhat subjective attribute and may be understood differently by different people. Therefore, it is a challenge to come up with a measure of the quality of the deals on the Groupon website. In marketing, researchers tend to use surveys to collect quality measures (e.g., Phillips et al. 1983, Parasuraman et al. 1985, Zeithaml 1988, Aaker and Jacobson 1994). However, it would not be efficient in terms of either cost or time to measure the quality of all 24,520 deals in the data using surveys. Another approach is to infer quality indirectly from observed information, such as brand alliances (Rao and Monroe 1989, Rao et al. 1999) waiting time (De Vany and Saving 1983), or total product sales (Orhun et al. 2015). In our context, however, the use of brand alliances or waiting time is not applicable. Price only explains a small proportion of the variation in the deal quality (see Section 5). Using total product sales may be problematic because popularity does not mean high quality. The number of coupons sold for a deal may be due largely to a lower price, not necessarily to higher quality. Using this proxy of quality may lead to the estimate of the price effect being biased downward. One possible approach is to use existing customer ratings of the deals. However, there are two potential problems: first, the rating may be about the retailer (e.g., the quality of a restaurant) rather than the deal (e.g., how satisfying the discount is); and second, the ratings are from the subset of customers that purchased the deal, so demand still exerts some influence.

Figure 3. (Color online) Number of Nonmissing Days vs. Rank of New Subscribers



To construct our measure of deal quality, we use surveys to elicit perceptions of deal quality for a subset of deals in our data. We then use deep learning techniques to relate these elicited quality measures to the text and image data on the evaluated deals. We then use this calibrated relationship to predict the quality levels of the other deals in our data. In Online Appendix A, we provide the details of this approach. Specifically, the process involves three stages: the labeling stage, the training stage, and the prediction stage. In the labeling stage, we obtain the original web page for each Groupon deal included in our data. We then recruit workers on MTurk[14] to label the quality of each deal on a scale from 1 to 9 (Liu et al. 2017). Each MTurk worker (MTurker) is given a random sample of the original deal web pages to rate. Also, each web page is rated by several MTurkers.[15] We use the round number of the average rating across these mTurkers as the quality value for that deal. Figure 4 shows examples of the surveys used.

Next, in the training stage, we gather the meta-information from each deal's web page. This includes (a) the deal attributes summarized by Groupon (e.g., the size of the discount, the original price, the length of the redemption period, and whether the deal is "tipped")[16]; (b) text information on the web page (e.g., deal title, deal description, fine print, highlights); and (c) image information (e.g., size of picture, color of picture, and whether there is a picture of a human face, a picture of the product, or a brand logo). With the MTurk quality ratings and the meta-information, we then train a CNN model on the data collected from MTurk workers. Figure 5 explains the model that we use, and Online Appendix A provides more details. We feed the model with raw text and image data to automatically discover feature representations (i.e., the deal quality in our case). Last, in the prediction stage, given the trained model, we then assign quality ratings (1–9) to the rest of the deals in our data.

This quality measure has the following advantages. First, it takes all kinds of observed information on the web page into consideration.[17] Second, the diverse content information guarantees variance in the quality measure. Third, the deep learning trained results can be used to predict the quality of new restaurant deals. If Groupon data have been collected for a new deal, the trained model is ready to be used to predict its quality. Fourth, as already demonstrated by Liu et al. (2017), deep learning–based natural language processing (NLP) methods outperform classical NLP methods, such as the "bag-of-words" and "parts-of-speech" models. We apply various machine learning–based models in addition to a deep learning–based model. Among these, CNN performs the best.[18] Last, researchers are exploring the possibility of combining machine learning techniques and structural modeling. One way is to use machine learning to generate new variables and inputs to a structural model. Motivating us to explore this method is the approach of Gechter and Tsivanidis (2017), which used a CNN and satellite images to identify slums, building a structural model to study the efficiency and equity of land policy.

Figure 6 shows the distribution of deal quality across various quality levels (in addition to the conditional purchase probabilities for deals at each level). The higher the number (from 1 to 9), the better the quality. The quality of the majority of deals is between 4 and 7. However, the distribution does not pass the normal distribution test (with Shapiro-Wilk Normality Test $W = 0.964$, $p < 0.000$), a feature that we accommodate in our model formulation. The conditional probability increases with the bin number, indicating that the higher the quality of a deal, the more likely the deal is to be purchased.

3.3. Behavior on Daily Deal Websites

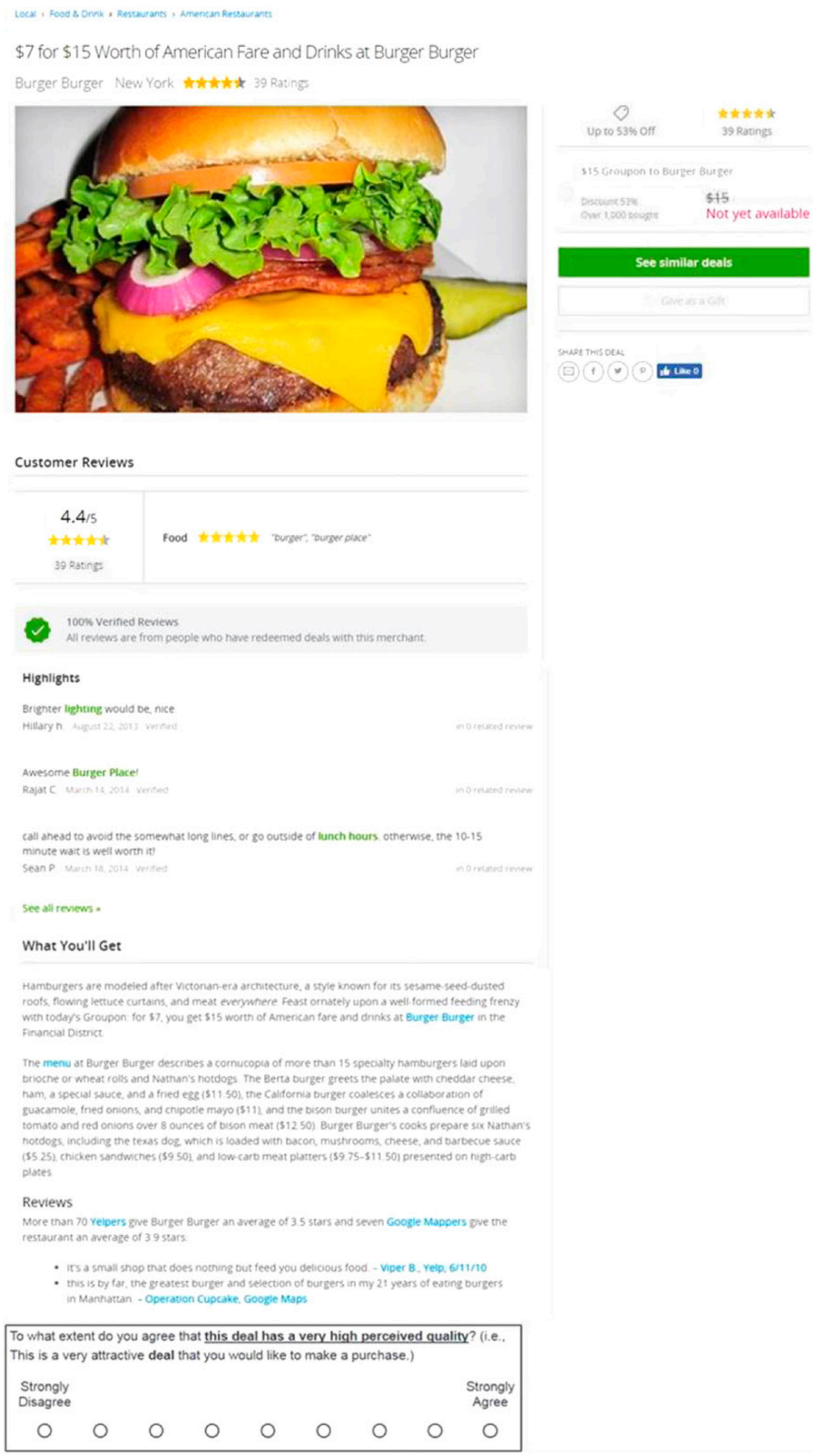
In addition to the features already discussed in the context of focusing on featured deals, the data reveal other behavioral patterns of consumers that are relevant when specifying our model.

On average, each subscriber clicked on and viewed approximately 2.3 deals over the three-month period, with an average purchase rate, conditional on clicking, of 8%. The low purchase rate indicates that subscribers rarely make purchases, even when they receive deals that could be better in price and quality than previously purchased deals. Our data also show that subscribers made multiple purchases during the observation window, although this number is low. Of the 10,951 new subscribers, 2.4% purchased more than one deal [Figure 2(e)]. Consumers clicked on multiple deals before they made a purchase [Figure 2(f)]. On average, they clicked on 2.08 (with standard deviation 2) deals before they made any purchase. This provides some support for the notion that search costs are associated with making purchases on the site.

3.4. Targeting of Deals?

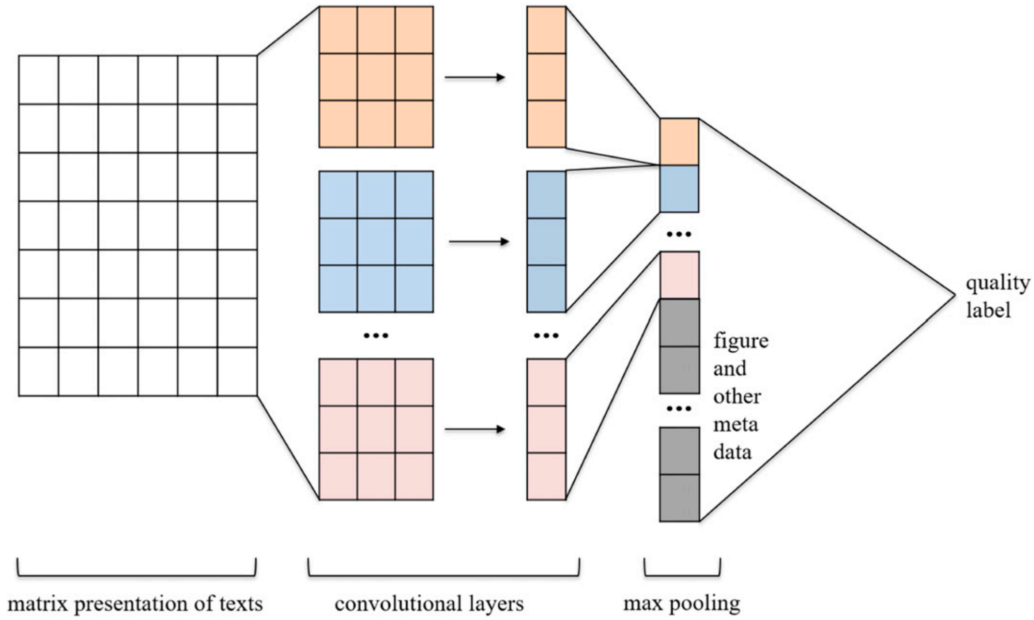
An obvious concern when studying the behavior of consumers on Groupon is that their behavior may be influenced by activities undertaken by Groupon to encourage clicking on and purchase behaviors; specifically, the targeting of deals by Groupon in early 2011. Most problematically, the consumers may have been targeted by Groupon on the basis of their past click and purchase behaviors. The presence of such unobservables, which are correlated with the quality of the deals that consumers received by email and their subsequent clicking and purchasing behaviors, leads to an endogeneity problem. As we show here, however, it does not seem that Groupon was

Figure 4. (Color online) Survey of the Deep Learning-Based Deal Quality Measure



targeting consumers in this way during our sample collection period, with the exception that the company sent deals in a given category of interest that customers specified when registering. By implication, everyone saw the same deals in a given category. First, in the presence of targeting, we expect a significant difference between the probability of receiving a deal in the same category before and after purchasing a deal from that category. This is a within-person comparison. We find that probability to be

Figure 5. (Color online) Training Model



25.59% before making a purchase and 20.67% after making a purchase; a *t*-test shows no significant difference between the two (p = 0.26). Second, we compare two groups of people: those who purchased deals from within a certain category ("buy") and those who never bought a deal in this category ("no buy"). We calculate the probability of receiving featured deals of the purchased category from the newsletter for both groups in different cities. Additionally, 92 (53.80%) of the 171 cities have a "buy" group with a higher probability of receiving deals in the same category, whereas there are 79 cities (46.20%) where the "no buy" groups have a higher probability. We cannot reject the null hypothesis (p = 0.16) that the two groups are the same. One possible reason why we cannot find any evidence of targeting is that at that point in the company's history, Groupon was focusing on local deals. In its early days there were seldom many deals from a specific category in a given local area on a specific day. As such, it would not have been possible to customize deals. Indeed, if targeting improved during the sample period, then we may reasonably expect to see more clicking over time instead of the opposite—that is, unless consumer learning about the deals was faster than the firm's learning about the consumer.

Figure 6. (Color online) Conditional Purchase Probability vs. % Deals in Bin

3.5. Implications of the Data for the Model

The findings of our preliminary examination of the data reveal the following.

1. Given the relative frequencies of clicking and purchasing, we focus on consumers' decisions to click on and purchase the featured deal in the newsletter sent to subscribers each day.
2. We need to account for the sequential nature of decision making and to allow for multiple purchases.
3. Search costs, in addition to deal characteristics, play an important role in consumers' deal purchase behavior.

4. Data Patterns and Potential Explanations

4.1. Data Patterns

In the data, we track individual subscribers over time from their initial registration on Groupon. Thus, we observe within-subscriber patterns of behavior from their inception. We see two patterns of interest in the data. First, we find that the probability that a consumer clicks on a merchant in the emailed newsletter declines over time. This pattern of decline is consistent with the "fatigue" thesis, according to which consumers' interest in daily deals decreases over time. Simultaneously, we find that the probability that a consumer makes a purchase conditional on clicking actually increases over time.

We graph the two patterns in Figure 7(a) and (b). Figure 7(a) plots the click probability (y axis) against tenure on the site (x axis), and Figure 7(b) shows the conditional and unconditional purchase probabilities (y axis) against tenure (x axis). The probabilities are averaged across all individuals with the same tenure.[19] We also include two solid lines, which represent the linear fitted trend lines for each data series.

4.2. Potential Explanations

There may be several alternative explanations for our data patterns. Here we discuss why we believe that most of these apply to one of our data patterns but not to both simultaneously. We list these in Table 3, with predictions regarding click and purchase behaviors based on each of these explanations.

A potential alternative explanation for the declining click rate in our data is the novelty effect. New subscribers are excited about Groupon and frequently click on deals that take them to the company's website. As time passes, however, customers gradually become bored with the deals and hence click on these email links less often. However, if boredom is what reduces clicks, it is unlikely that this also explains the increasing trend in conditional purchases. One possible explanation for the increase in conditional purchase probability over time is preference changes due to past purchases, which induce loyalty (or a high "switching cost"). In the reduced form analysis in the following section (Table 5), we find that past purchases have a positive effect on clicks, indicating an inertia or loyalty effect of past purchases on clicks but a *negative* effect on purchases. This seems to indicate the presence of a budget constraint that limits consumers' purchases. Hence, inertia alone is unlikely to be at play.

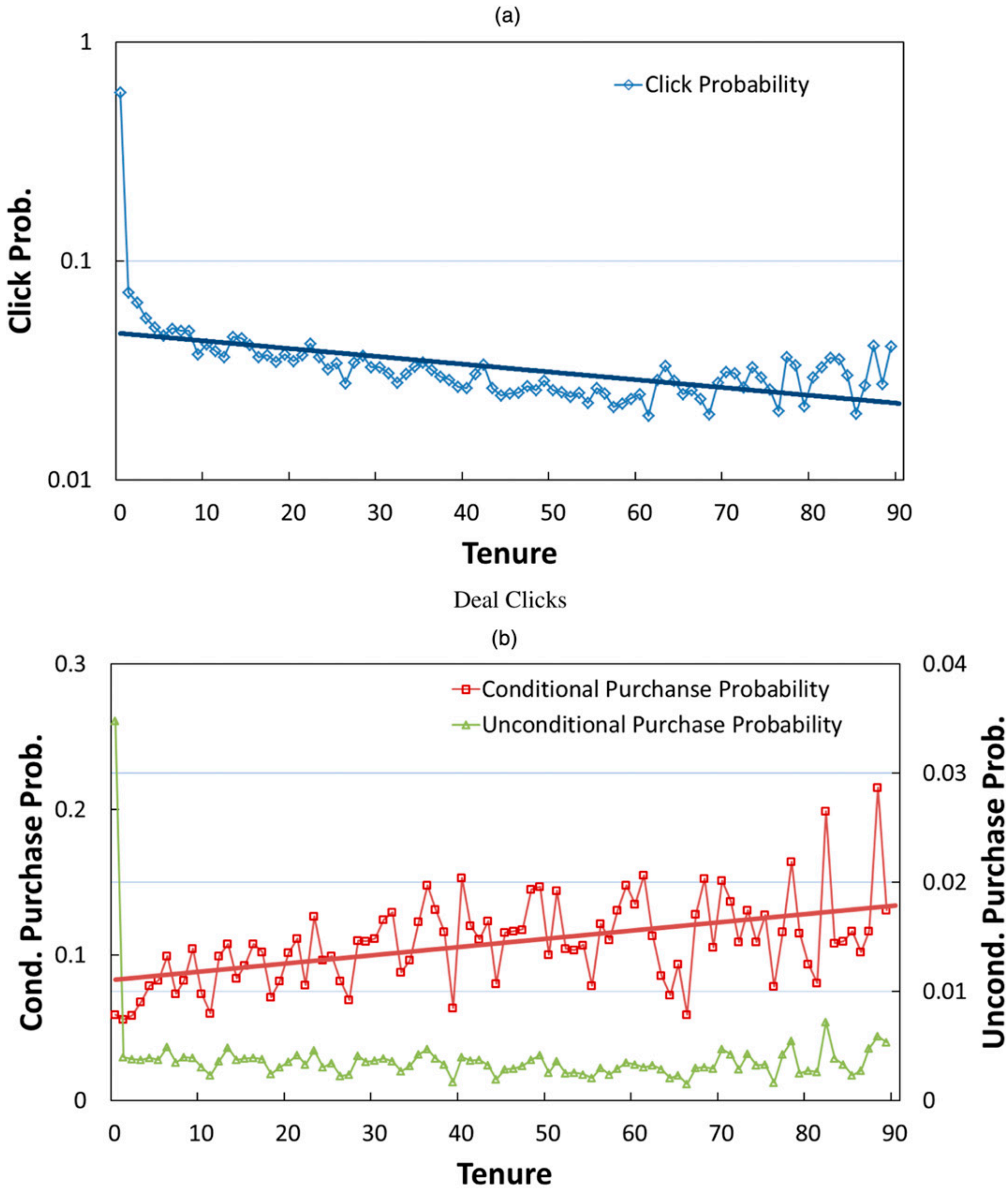
Because budget constraints may play a role,[20] it is important to control for previous purchases to address this concern. Budget constraints can serve as an alternative explanation for the "within"-customer decline in click probability but cannot explain increases in the within-customer conditional purchase probability (with a limited budget, customers are less likely to click over time, and they should also become less likely to make a purchase). However, if we treat the "budget constraint" as a cross-customer difference, this may explain the increase in conditional purchase probability that we observe at the aggregate level but not the declining click rate (people with higher budgets are more likely to purchase, but they are also more likely to click). However, our focus is on changes in within-consumer behavior. As such, "budget constraints" alone is an insufficient alternative explanation for reconciling both data patterns.

We note that the above set of alternative explanations is by no means exhaustive. Furthermore, the explanations for our data patterns may be more nuanced than our proposition and may involve combinations of the above explanations. Given these potential ambiguities, we acknowledge that what we propose is only one plausible explanation for the patterns observed in our data.

4.3. Search with Learning

To reconcile the two competing patterns observed in the data, we invoke the notion of consumer learning. Upon receiving the newsletter from daily deal websites, consumers can quickly obtain some limited information about the featured deal, such as its original price [$220 in the illustrated example in Figure 1(a)], the discount rate (80%), the price after discount ($55), and a visual depiction of the deal. If the consumer is interested in the deal, (s)he can click the newsletter, which redirects to the deal's web page [i.e., Figure 1(b)]. The web page contains complete information about the deal. In addition to information already revealed in the newsletter, the web page displays information about the time left to purchase the deal ("2 days, 15 hours, and 9 minutes"), the number of coupons sold ("49 bought"), the status of the tipping point ("The deal is on!"), and the deal's fine print (such as the length of the redemption period and the maximum number of coupons that can be purchased per person).

Figure 7. (Color online) Data Trends



Deal Clicks

(Un)Conditional Deal Purchases

When making a purchase decision, the consumer faces a trade-off between purchasing the current deal on the basis of the quality of information revealed on the web page and waiting for another deal in the future. However, as a new subscriber, the consumer is uncertain about the quality of future deals; expectations about the quality of deals offered by Groupon need to be formed. This, in turn, is based on prior knowledge of Groupon deals in addition to the deals that the consumer clicks on and views after subscribing. Consumers use deal information from the web page to update their knowledge regarding the quality of those deals that Groupon offers. This comprises the "learning" stage of our model.

Upon receiving the newsletter from the daily deal website, the consumer decides whether to click on the newsletter to obtain more information about the deal. A search cost is incurred if the consumer decides to visit the page to obtain further information. This may be associated with the loading time of the web page or the effort spent reading through the details. The search cost in our context is also related to the relevance of Groupon deals to the subscribing customer. Because Groupon determines what the consumer sees, the majority of the deals received may not be relevant to the recipient. Also, because we do not filter new subscribers according to the number of clicks or purchases observed in the data, there could be

Table 3. Hypotheses and Predictions

Hypothesis		Predictions: Declining click trend	Increasing conditional purchase trend	Making better choice over time
Novelty effect	New subscribers are excited about Groupon so they go to the website and click deals frequently. As time passes by, however, they gradually get bored and hence will click less.	Yes	No	No
Loyalty effect or switching cost	Preference changes due to past purchase, which induces loyalty	No	Yes	No
Budget constraint	Consumers cannot keep purchasing deals on Groupon	Yes	No	No
Search with learning	Consumers search for product information and learn about the quality distributions of deals	Yes	Yes	Yes

several for whom the deals are not relevant; hence they do not click or purchase. Such behavior can be reconciled by a large search cost. We refer to the click decision stage as the "search" stage. Consumers tend to click on more deals in the beginning, thereby learning about the quality of Groupon deals. As their knowledge accumulates the incentive to learn declines, and they click on fewer deals. As the uncertainty about the position of a specific deal's quality in the quality distribution of Groupon deals is resolved with learning, the consumer is more likely to purchase the clicked deal because there is a declining incentive to wait to purchase the deal. This produces the data pattern of increasing conditional purchase probabilities that we observe in the data.

A key question to be answered is, what does the consumer learn on the web page? Note that several characteristics are available directly in the newsletter. By visiting the web page, the consumer learns about some additional features but also about where this deal stacks up relative to other deals that (s)he may have encountered on the site. Because a different deal presents itself to the consumer in each period (each day), we are interested in the consumer's knowledge of the distribution of deals and how this is influenced by information from a specific deal, which, in turn, influences subsequent behavior on the site. As noted in the data section, we refer to this as the distribution of deal quality. In our context, learning is thus unlike the traditional concept of learning in the brand choice literature (see, e.g., Erdem and Keane 1996), in which the consumer considers the same set of products in each period and learns about each product over time on the basis of the signals that (s)he receives by buying and consuming the product. In our case, we are not interested in the purchase of a specific deal or some small set of deals (because these exact deals are unlikely to reappear in the data).

Table 4. Residual Quality Regression

	Dependent Variable (DV): Quality: Estimate	Standard error	
Discount	−0.0248	0.0002	***
Original price	−0.0471	0.0011	***
Picture color	0.0130	0.0020	***
If has face in the picture	0.0066	0.0005	***
Title text length	−0.0410	0.0300	
Category fixed effects		Yes	
Adjusted R^2		0.0581	
Observations		24,520	

$^{***}p < 0.01$.

5. Reduced-Form Analysis for Learning

In this section, we use our data to provide evidence for learning. First, we choose the first and the last months for a consumer with Groupon.[21] Then, in Figure 8(a), we plot the mean and standard deviation of the purchased deal quality for the first and last months averaged across all subscribers. Compared with the first months with Groupon, subscribers purchase deals with higher quality (4.41 vs. 5.84, $p < 0.000$) and with lower quality variance (1.73 vs. 1.39, $p < 0.05$) in the last month. Figure 8(b) uses the quality/price ratio to control for price, and the effects are even more pronounced. Of course, this pattern may be due to a gradual improvement in the deals available on Groupon. We plot the evolution of deal quality on Groupon during the sampling period in Figure 9. We graph the average quality of deals over time. There is no significant time trend (coef. = −0.00083, $p = 0.27$), suggesting that the increase in the purchased deal quality is not driven by an improvement in the quality of Groupon deals. Later, in the model specification, we assume that subscribers are learning from a *stationary* quality distribution.

5.1. Quality and Search

In the click stage, when consumers are deciding whether to click on the links in the newsletter and search for more quality information, they are only able to observe a small portion of information on deal

Table 5. Reduced-form Analysis: Linear Fixed Effects Model

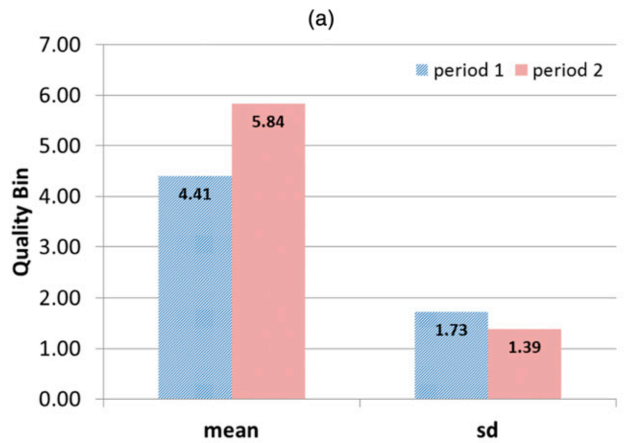
	(1)	(2)	(3)	(4)	(5)
	Baseline	Tenure	Past Click	Tenure&Past Click	(4)+Past Purchase
DV: if click	Estimate	Estimate	Estimate	Estimate	Estimate
Fitted quality	0.0092***	0.0110***	0.0106***	0.0108***	0.0108***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Residual quality	0.0016	0.0008	0.0010	0.0009	0.0009
	(0.0018)	(0.0018)	(0.0017)	(0.0017)	(0.0017)
Past purchase					0.0085***
					(0.0015)
Past click			−0.0296***	−0.0283***	−0.0290***
			(0.0003)	(0.0003)	(0.0003)
Tenure		−0.0009***		−0.0001***	−0.0002***
		(0.0000)		(0.0000)	(0.0000)
Current price	0.0004	0.0010***	0.0009***	0.0010***	0.0010***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	527,298	527,298	527,298	527,298	527,298
	Baseline	Tenure	Past Click	Tenure&Past Click	IV+Past Purchase
DV: If purchase (conditional)	Estimate	Estimate	Estimate	Estimate	Estimate
Quality	0.0048***	0.0048***	0.0048***	0.0048***	0.0035**
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0017)
Past purchase					−0.2538***
					(0.0071)
Past click			0.0005***	0.0011***	0.0081***
			(0.0002)	(0.0003)	(0.0012)
Tenure		0.0001***		0.0001	0.0017***
		(0.0000)		(0.0002)	(0.0002)
Current price	−0.0037*	−0.0037*	−0.0037*	−0.0037*	−0.0038**
	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0019)
Individual FE	Yes	Yes	Yes	Yes	Yes
Observations	25,459	25,459	25,459	25,459	25,459

Note. Standard error in parentheses.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

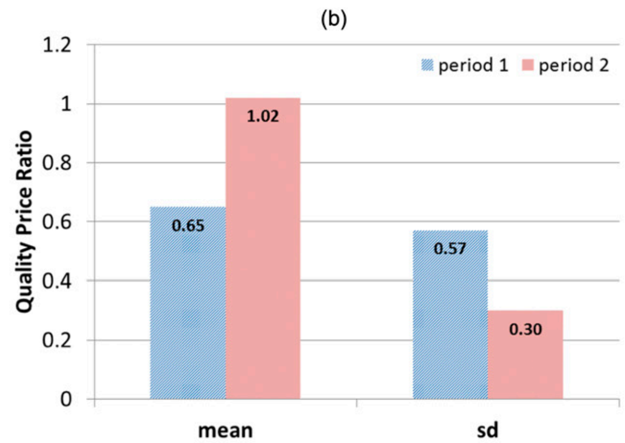
quality, for example, the deal price, the discount rate, and a picture of the deal. This quality information affects the consumer's decision to click and proceed to the web page. When making a purchase decision, the consumer observes the complete information regarding the deal's quality. That information includes the deal's expiration date, the location of the merchant, Groupon's rating of the deal, the merchant's rating, and a detailed description of the deal, in addition to the partial information observed before clicking. To account for difference in information about the deal's quality at the two stages, we divide deal quality into two parts: that observed before clicking, and the remaining quality information. We do this by regressing our measure of deal quality on the characteristics observed at the click stage, including discount rate (discount), the original price of the deal (original price), the characteristics of the picture shown in the deal (e.g., the picture color and whether there is a face), and the length of the deal title. Table 4 shows the results from this regression. The observed quality information before clicking is found to account for only a small portion of variation in deal quality (the adjusted R^2 is 0.06). This indicates that the majority of deal quality information is only revealed *after* clicking. We use this as a motivation for the consumer's need to click (and search).

On the basis of the regression described above, we compute the fitted quality ($X\beta$ part of the regression) and residual quality (ϵ part of the regression). Next, we estimate two baseline models. The first is a linear probability model at the individual-day level, with the click decision serving as the dependent variable, and the second is a linear probability model of the conditional purchase decision as the dependent variable. For the former, we include the fitted and residual quality (quality_fit and quality_res) separately as regressors; for the second, we use only the (total) quality measure because all of the deal quality information is observed after clicking. Column (1) of Table 5 shows

Figure 8. (Color online) Empirical Evidence of Consumer Learning: Purchased Deals



First and Last Month Analysis on Quality



First and Last Month Analysis on Quality Price Ratio

the results from these regressions. As expected, the fitted quality has a significant positive effect on clicking, but the residual quality, revealed after clicking, does not. Overall quality significantly affects purchase decisions. These results demonstrate the different effects of partial versus complete quality information on click and purchase behaviors.

5.2. Time Trend in Click, Conditional Purchase, and Unconditional Purchase

Column (2) of Table 5 reflects the patterns in the data that we show in Figure 7, that is, the declining time trend (represented by the tenure variable of the time since subscription) in clicking and the increasing time trend in conditional purchases, after controlling for price after discount (current_price), quality, and individual fixed effects. Table 6 shows the results for unconditional purchases. These indicate that it is important to study what happens after clicking (i.e., conditional purchase decisions). If we look only at the declining click trend or the unconditional purchase trend, we might conclude that there is indeed a "fatigue" or novelty (wear-out) effect. However, when we look at the trend in conditional purchases together with the trend in clicking behaviors, we can recognize the nuanced nature of behavior that might underlie our data.

One point to note is that the current price has a positive effect on clicks but a negative effect on purchases. We attribute the former to the possible role that price plays in signaling quality at the click stage (over and above the indirect effect of price on quality via our formulation of the quality measure). The latter simply reflects the economic effect of price.

5.3. Effects of Past Purchases, Past Clicks, and Tenure on Clicks and (Conditional) Purchases

We compare the results of the previous model with consumers' tenure as Groupon subscribers (controlling for price, quality, city, and individual fixed effects) with those of several alternative models, including a model with the number of past clicks on Groupon deals (pastClick) instead of tenure [column (3) of Table 5], a

Figure 9. (Color online) Quality of Featured Deals over Time

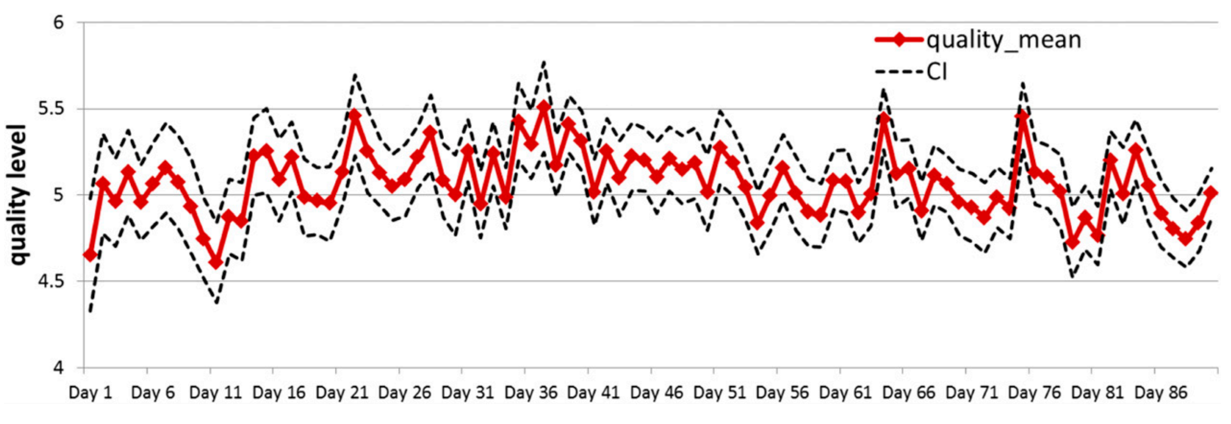


Table 6. Reduced-Form Analysis: Unconditional Purchase

	DV: If purchase (unconditional)		
	Estimate	Standard error	
Quality	0.0003	0.0001	***
Tenure	−0.0001	0.0000	***
Current price	−0.0003	0.0001	***
Individual FE		Yes	
Observations		527,298	

***$p < 0.01$.

model with past clicks and tenure [column (4) of Table 5], and a model with past clicks, past purchases (pastPur), and tenure [column (5) of Table 5].

The past purchase variable is the cumulative number of deals purchased by each consumer. This reflects any budget constraints or switching costs. Past clicks represent the cumulative number of deals that each consumer has clicked on. It reflects the amount of learning a consumer has accumulated. These are both alternative explanations for state dependence (Dubé et al. 2010). Any remaining temporal effect is reflected in the tenure variable.

First, we replace tenure in column (2) of Table 5 with past clicks. The results are shown in column (3) of Table 5. Consumers learn about Groupon deals' quality when they click on the newsletter; therefore, past clicks serve as a proxy for learning. The significant negative effect of past clicks on click behavior and the positive effect on purchases indicate that the more a consumer learns about the quality of Groupon deals, the less likely (s)he is to click new deals. However, the more knowledge the consumer accumulates, the more likely (s)he is to purchase the deal after clicking. This is consistent with our characterization of observed behavior as learning.

Next, we add tenure back into the previous model; the results are shown in column (4) of Table 5. For the click decision, the effect of tenure does not fully disappear but is attenuated after controlling for past clicks (from −0.0009 to −0.0001). This may demonstrate the existence of some other temporal phenomenon related to clicking in addition to learning (e.g., some residual fatigue). For the conditional purchase decision, the effect of tenure goes away after controlling for past clicks. So the time trend for clicks and conditional purchases can be captured by past clicks as a measure of learning, and the signs are consistent with our learning explanation.

Last, in column (5) of Table 5, we add the past purchases variable to the analysis. Past purchases capture the budget constraint or switching cost. This has a significant positive effect on clicking (coef. = 0.0085, s.e. = 0.0015), indicating that past purchases have an inertia effect on click behaviors. This variable also has a significant negative effect on conditional purchases (coef. = −0.2538, s.e. = 0.0071), indicating possible budget constraints. Moreover, the signs of the past purchase variable are opposite to the signs of past clicks (a proxy for learning) for both behaviors (with 0.0085 vs. −0.0290 for clicks, and −0.2538 vs. 0.0081 for purchases). This gives rise to the separation of learning and switching costs in the structural model.

6. Methodology

We build a two-stage dynamic search model (e.g., Hartmann and Nair 2010, Seiler 2013, Mojir and Sudhir 2016) with consumer learning. In the first stage of the model, the consumer forms expectations about the deal quality distribution and decides whether (s)he wants to click on the newsletter and visit the deal's web page. Once (s)he has clicked on and visited the web page, the consumer decides whether to purchase the deal or wait for another on the basis of the information that they gather from the web page. Visiting the page also allows consumers to learn from that information and use it to update their posterior perceptions of the quality distribution of the deals that the website offers.[22]

We use a nonparametric learning approach (i.e., Dirichlet learning) to characterize the learning process (Koulayev 2013, De Los Santos et al. 2017). One distinction that we draw with a majority of the learning literature (e.g., Erdem and Keane 1996), which assumes that consumers know the distribution of quality (e.g., normal) and learn about the mean (and sometimes the variance) of the distribution, is that we assume that new Groupon subscribers do not know the distribution of deal quality and need to learn about the distribution itself. As a result, we incorporate Dirichlet learning into the model (Rothschild 1974, Koulayev 2013, De Los Santos et al. 2017). A Dirichlet prior is less restrictive than, for example, a normal distribution in terms of the distribution shape and seems better suited to the distribution in Figure 6. We assume that the consumer has a prior on the quality distribution of Groupon deals before registration. In particular, the consumer knows that quality can take a finite number of values but does not know the probability of each quality level occurring. On the basis of the deals clicked on, the consumer learns about the probability of each quality level occurring, thereby learning about the *distribution* of deal quality rather than the "true" quality of the product (as in, e.g., Erdem and Keane 1996). However, the consumer can never predict the exact quality of the next featured deal with certainty, even after learning about the distribution of deals.

We note that our assumption that consumers learn about deal quality distributions is not easy to distinguish empirically from the assumption that consumers learn about the mean of the quality distribution with normal priors, as in traditional Bayeasian learning models.[23] On an intuitive level, given our setting, in which consumers are faced with a sequence of different restaurant deals over time, we believe that our assumption that consumers learn about the quality distribution of deals is more appropriate.

6.1. Two-Stage Forward-Looking Search Model

Our structural model incorporates nonparametric Bayesian learning in a two-stage dynamic search model (e.g., Hartmann and Nair 2010, Seiler 2013, Mojir and Sudhir 2016). In those studies, the researchers formulated a two-stage dynamic demand model that took into account search costs. In the first stage, the consumer decides whether to search for price; in the second stage, the consumer decides whether to make a purchase. However, such a model does not consider the consumer's learning behavior, which our model does.

In Section 1, we introduced the consumer's decision-making process. We now build a model to reflect such a process. On the basis of our preliminary analysis, we focus on consumers' click and purchase decisions for the featured deal in the emailed newsletter that they receive each day. As we show below, consumers' daily decisions can be formulated as a dynamic discrete choice problem. We note that the basic model borrows heavily from Seiler (2013), as we endogenously model the search decision; we then extend the model to account for subscribers' learning behavior.

6.1.1. Flow Utility. We use the subscript "cs" to denote the click stage, "ps" to denote the purchase stage, "i" to denote consumers, and "t" to denote the time period. Additionally, "d_{it}^C" is the decision indicator at the click stage, and $d_{it}^C = 1$ ($d_{it}^C = 0$) means the consumer decided to (not) click the featured deal in the newsletter. Similarly, "d_{it}^P" is the decision indicator at purchase stage. $d_{it}^P = 1$ ($d_{it}^P = 0$) means the consumer decided to (not) purchase the featured deal. At each stage, corresponding to each decision, the consumer has a flow utility.[24]

At the click stage, to obtain information on the featured deal, the consumer incurs a search cost associated with clicking and going to the deal's web page. The search cost is denoted by c_{it}. The flow utility that (s)he gets is

$$u_{cs,it}^1 = -c_{it} + \varepsilon_{cs,it}^1 \\ = \bar{u}_{cs,it}^1 + \varepsilon_{cs,it}^1. \tag{1}$$

Similarly, if (s)he decides not to click the deal, no search cost is incurred, and (s)he gets the following flow utility:

$$u_{cs,it}^0 = \eta \cdot pp_{it} + \varepsilon_{cs,it}^0 \\ = \bar{u}_{cs,it}^0 + \varepsilon_{cs,it}^0, \tag{2}$$

where pp_{it} is the number of deals that an individual i purchased before period t. Daily deals are coupons that can be redeemed within a certain period. As such, the deals that a consumer has already purchased but not redeemed could influence his or her opportunity costs, which in turn affects whether (s)he clicks on a subsequent deal. We use this term to capture the effects of switching cost or budget constraint. Adding this term is motivated by our earlier finding in the raw data that Groupon subscribers skipped better deals after making a purchase decision and that there is a negative coefficient on the number of previously purchased deals in our regressions [see column (5) of Table 5, where η captures the corresponding effect]. We reset $pp_{it} = 0$ after the redemption periods of the purchased deals expire. The deals' average redemption period is approximately 195 days (with standard deviation of 82 days) for our data. We also checked our results for sensitivity to this reset period. The pp_{it} variable is different from the "inventory" effect for storable goods in Seiler (2013) but similar to the "consumption stock" effect in Hartmann (2006), in that we allow it to increase over time but reset it periodically. With this term, we are able to explain why people do not keep purchasing deals even when better ones come up after a purchase. The effect does not last after resetting the variable, which ensures stationarity.

If the consumer decides to click, (s)he enters the second stage, in which more information about the featured deal is revealed. Upon opening the emailed newsletter, the price of the featured deal P_{it} is revealed to the subscriber; we assume that there is no cost to obtain price information. We use Q_{it} to denote a specific deal's quality level as revealed to the consumer after clicking through to the merchant's site. On the basis of the newly obtained information (i.e., the deal's quality level), (s)he updates his/her beliefs about the quality distribution (we introduce this learning mechanism in Section 6.2). We denote the prior quality distribution at period t for consumer i as $f_{it}^{pr}(Q)$ and the corresponding posterior as $f_{it}^{po}(Q)$. We also assume that the consumer is risk neutral. Therefore, if (s)he decides to purchase the deal, then the flow utility obtained is determined by the current deal (i.e., its price and quality) and the previous purchase status (i.e., the number of previously purchased deals):

$$u_{ps,it}^1 = \beta_0 + \beta_P \cdot P_{it} + \beta_Q \cdot Q_{it} + \eta \cdot pp_{it} + \varepsilon_{ps,it}^1 \\ = \bar{u}_{ps,it}^1 + \varepsilon_{ps,it}^1. \tag{3}$$

If, however, (s)he decides not to purchase, (s)he gets utility

$$u_{ps,it}^0 = \eta \cdot pp_{it} + \varepsilon_{ps,it}^0 \\ = \bar{u}_{ps,it}^0 + \varepsilon_{ps,it}^0. \tag{4}$$

When deciding to click on a deal in the newsletter, the consumer anticipates that (s)he must make another decision within the same time period and knows that (s)he is going to receive more *nondiscounted* utility, as defined by the purchase-stage flow utility. Therefore, if the consumer decides to search, the total utility in the current time period is simply the sum of the flow utilities from both decision-making stages:

$$U_{it}(d_{it}^C = 1) = \max\{u_{ps,it}^0, u_{ps,it}^1\} + u_{cs,it}^1 = \max\{u_{ps,it}^0, u_{ps,it}^1\} + \bar{u}_{cs,it}^1 + \varepsilon_{cs,it}^1. \quad (5)$$

When deciding not to click, the total utility in the current time period is equal to the flow utility at the click stage (because the consumer does not enter the second stage, no purchase utility is obtained).

$$U_{it}(d_{it}^C = 0) = u_{cs,it}^0 = \bar{u}_{cs,it}^0 + \varepsilon_{cs,it}^0 \quad (6)$$

Note that before clicking, the current period flow utility in the purchase stage is unknown. Therefore, the consumer has to form an expectation about the flow utility in the purchase stage, knowing that (s)he makes the optimal purchase decision conditional on the information that (s)he obtains in the purchase stage.[25] This information is not yet available to the consumer in the click stage, and therefore the consumer cannot perfectly predict his/her decision in the purchase stage.

Let $\tilde{\varepsilon}_{cs,it} = (\varepsilon_{cs,it}^0, \varepsilon_{cs,it}^1)$ be the vector of idiosyncratic shocks at click stage in time t, and $\tilde{\varepsilon}_{ps,it} = (\varepsilon_{ps,it}^0, \varepsilon_{ps,it}^1)$ is the corresponding idiosyncratic shock vector at purchase stage. Note that different sets of error terms enter the model before and after the clicking decision.

6.1.2. The Dynamic Optimization Problem. Formally, a consumer chooses an infinite sequence of decision rules $D = \{d_{it}\}_{t=0}^{\infty}$ to maximize his/her expected present-discounted sum of utilities[26]

$$\max_{\{d_{it}\}_{t=0}^{\infty}} E\left\{\sum_{t=0}^{\infty} \delta^t \cdot U_{it}(d_{it}) | x_{it}, \tilde{\varepsilon}_{it}\right\}, \quad (7)$$

where $d_{it} = \{d_{it}^C, d_{it}^P\}$, $\tilde{\varepsilon}_{it}$ includes all of the idiosyncratic shocks and δ is the discount factor. x_{it} is the vector of observed state variables for subscriber i in period t.

6.1.3. Value Functions. In the model, the consumer potentially has two consecutive decisions to make in one period. First, the consumer has to decide whether to click on the deal in the newsletter to obtain more information about the deal. If (s)he does not click, then (s)he does not to make any other decision in the current period. If (s)he decides to click the deal, then (s)he receives further information about the quality distribution and has to decide whether to make a purchase. Therefore, we define two different value functions, which depend on whether the consumer clicked on the deal or not: (1) the value function at click stage $V_{cs,it}$; (2) the value function at purchase stage $V_{ps,it}$. These depend on one another and must be solved simultaneously. As in Rust (1987), we define the expectation of value function at the click (purchase) stage $EV_{cs,it}$ ($EV_{ps,it}$), integrated over the realization of $\tilde{\varepsilon}_{cs,it}$ ($\tilde{\varepsilon}_{ps,it}$) as

$$EV_{cs,it} = EV_{cs}(x_{cs,it}) = \int_{\tilde{\varepsilon}_{cs,it}} V_{cs,it}(x_{cs,it}, \tilde{\varepsilon}_{cs,it}) \cdot dF_{\tilde{\varepsilon}_{cs,it}} \quad (8)$$

$$EV_{ps,it} = EV_{ps}(x_{ps,it}) = \int_{\tilde{\varepsilon}_{ps,it}} V_{ps,it}(x_{ps,it}, \tilde{\varepsilon}_{ps,it}) \cdot dF_{\tilde{\varepsilon}_{ps,it}}. \quad (9)$$

By assuming the error terms ($\tilde{\varepsilon}_{cs,it}$, $\tilde{\varepsilon}_{ps,it}$) follow independent and identically distributed extreme value distributions and also making the standard conditional independence assumptions[27] (cf. Rust 1987), we obtain closed-form expected value functions with respect to each decision stage[28]:

$$\begin{aligned} EV_{cs,it} &= EV_{cs}(x_{cs,it}) \\ &= \log\{\exp(\bar{v}_{cs,it}^0) + \exp(\bar{v}_{cs,it}^1)\} \\ &= \log\Big\{\exp\Big(\bar{u}_{cs,it}^0 + \delta \cdot E\left[EV_{cs}(x_{cs,it+1}) | x_{cs,it}, d_{it}^C = 0\right]\Big) \\ &\quad + \exp\Big(\bar{u}_{cs,it}^1 + E\left[EV_{ps}(x_{ps,it}) | x_{cs,it}, d_{it}^C = 1\right]\Big)\Big\} \end{aligned} \quad (10)$$

$$\begin{aligned} EV_{ps,it} &= EV_{ps}(x_{ps,it}) \\ &= \log\{\exp(\bar{v}_{ps,it}^0) + \exp(\bar{v}_{ps,it}^1)\} \\ &= \log\Big\{\exp\Big(\bar{u}_{ps,it}^0 + \delta \cdot E\left[EV_{cs}(x_{cs,it+1}) | x_{ps,it}, d_{it}^P = 0\right]\Big) \\ &\quad + \exp\Big(\bar{u}_{ps,it}^1 + \delta \cdot E\left[EV_{cs}(x_{cs,it+1}) | x_{ps,it}, d_{it}^P = 1\right]\Big)\Big\}. \end{aligned} \quad (11)$$

The expected value functions only depend on observed state variables.[29] The term $EV_{ps,it}$ can be interpreted as the inclusive value of clicking the featured deal in time period t, excluding the search cost. At the click stage, the consumer makes a trade-off between the expected utility of this inclusive value less the search cost, and the utility of not clicking.

6.1.4. State Variables and Transition Rules. There are two decision stages in one time period, and the consumer enters the second stage if and only if (s)he clicks; otherwise, (s)he just makes one decision in the current period and then enters the next. There are two sets of state variables: click-stage state variables $x_{cs,it}$ and purchase-stage state variables $x_{ps,it}$. Corresponding to the decision sequences, we have the following possible state transitions:

1. From the *click stage* in the current period to the *click stage* in the next period: $x_{cs,it} \rightarrow x_{cs,it+1}$
2. From the *click stage* in the current period to the *purchase stage* in the current period: $x_{cs,it} \rightarrow x_{ps,it}$
3. From the *purchase stage* in the current period to the *click stage* in the next period: $x_{ps,it} \rightarrow x_{cs,it+1}$

Before clicking in period t, the consumer has information on the price of the current featured deal from the newsletter and partial information about the quality levels of the deals.[30] We use the vector $\vec{\alpha}_{it}^{pr}$ to denote a consumer's knowledge of the deal quality distribution ($f_{it}^{pr}(Q)$) on the Groupon website at click stage in the current period. The expected value function therefore depends on $\vec{\alpha}_{it}^{pr}$. We provide the derivation in Section 6.2. Additionally, we assume that the customers know with certainty the number of deals that they previously purchased (i.e., their inventory). Therefore, the state variables at this stage are the price of the deal, the number of previously purchased deals, and prior knowledge of quality:

$$\begin{aligned} x_{cs,it} &= \{P_{it}, pp_{it}, \vec{\alpha}_{it}^{pr}\} \\ &= \{P_{it}, pp_{it}, \{\alpha_{it}^{pr,1},\ \alpha_{it}^{pr,2}, \dots, \alpha_{it}^{pr,N}\}\}. \end{aligned} \tag{12}$$

If the consumer enters the purchase stage, then more information is revealed and (s)he receives the quality signal Q_{it} for the given deal; the consumer then uses this signal to update his or her knowledge about the distribution of quality. After updating, the distribution becomes $f_{it}^{po}(Q)$, and the knowledge about this distribution is denoted as $\vec{\alpha}_{it}^{po}$. The purchase-stage state variables are

$$\begin{aligned} x_{ps,it} &= \{P_{it}, Q_{it}, pp_{it}, \vec{\alpha}_{it}^{po}\} \\ &= \{P_{it}, Q_{it}, pp_{it}, \{\alpha_{it}^{po,1},\ \alpha_{it}^{po,2}, \dots, \alpha_{it}^{po,N}\}\}. \end{aligned} \tag{13}$$

To complete the definition of state variables, we note that P_{it} remains unchanged within period t; there is thus no transition in P_{it} from the click stage to the purchase stage. Here, P_{it+1} is the price revealed to consumer i (and researcher) in the next period. We make the assumption that the consumer has rational price beliefs and follows the empirical distribution of prices, which we approximate with a normal distribution with the empirical mean and empirical standard deviation derived from the data. The evolution of pp_{it} is straightforward. pp_{it+1} is equal to $pp_{it} + \Delta pp_{it}$, where Δpp_{it} is the number of deals purchased in period t. In our case, Δpp_{it} is either 0 or 1. We lay out the detailed transition rules in the online appendix.

What are the factors that drive a consumer to make the first click? From the model's perspective, four factors can potentially influence the consumers' first click decision, as follows. (1) The high initial priors of consumers. For example, new subscribers on Groupon may have very high expectations of the deals that the site provides. (2) The value of learning. This means that the value of clicking outweighs the value of not clicking, with the additional information gained from learning. (3) Price. A deal with a very low price induces the consumer to click, given that at the click stage they are already aware of the price of the deal. (4) The idiosyncratic shock ϵ. This shock might be a sudden shopping need, or a recommendation from a friend.

6.2. Dirichlet Learning Process

The subscriber updates his or her belief about the quality distribution according to a Dirichlet learning process. The Dirichlet distribution is widely used in Bayesian statistics because it is a conjugate before the multinomial distribution.

Assume that there are N quality levels (i.e., grids): $\vec{q} = \{q_1, q_2, \dots, q_N\}$ and the sampling probability of each level at period t is denoted using the vector $\vec{\rho}_t = \{\rho_{1t}, \rho_{2t}, \dots, \rho_{Nt}\}$, which is distributed according to a Dirichlet distribution of order N with density

$$f(\rho_{1t}, \rho_{2t}, \dots, \rho_{Nt}) = \frac{\Gamma\left(\sum_{n=1}^{N} \alpha_{nt}\right)}{\prod_{n=1}^{N} \Gamma(\alpha_{nt})} \prod_{n=1}^{N} \rho_{nt}^{\alpha_{nt}-1}. \tag{14}$$

If the prior expected value of each ρ_{nt} is given by $E[\rho_{nt}] = \alpha_{nt}/W_t$ and $W_t = \sum_{n=1}^{N} \alpha_{nt}$, then the corresponding posterior is

$$E[\rho_{nt+1}] = \begin{cases} \dfrac{\alpha_{nt}}{W_t + 1} & \text{if } q_n \text{ is not sampled} \\ \dfrac{\alpha_{nt} + 1}{W_t + 1} & \text{if } q_n \text{ is sampled.} \end{cases} \tag{15}$$

This completes our definitions of state variables and their transition rules. For a more detailed description of the Dirichlet learning process, please refer to Online Appendix D. We now explicitly provide the expected value functions for the click and purchase stages (a detailed derivation can be found in Online Appendix B).

$$\begin{aligned} EV_{cs,it} &= EV_{cs}(P_{it}, \vec{\alpha}_{it}^{pr}, pp_{it}) \\ &= \log\Bigg\{\exp\Bigg(\eta \cdot pp_{it} + \delta \cdot \int_{P_{it+1}} EV_{cs}(P_{it+1}, \vec{\alpha}_{it}^{pr}, pp_{it+1}) \\ &\qquad \cdot dF(P_{it+1})\Bigg) \\ &\qquad + \exp\left(-c_{it} + \sum_{n=1}^{N} \left(EV_{ps}(P_{it}, \vec{\alpha}_{it}^{po}(q_n), pp_{it}) \cdot \rho_e^n\right)\right)\Bigg\} \end{aligned} \tag{16}$$

$$
\begin{aligned}
EV_{ps,it} &= EV_{ps}(P_{it}, Q_{it}, \vec{\alpha}_{it}^{po}, pp_{it}) \\
&= \log\Bigg\{\exp\Bigg(\eta \cdot pp_{it} + \delta \cdot \int_{P_{it+1}} EV_{cs}(P_{it+1}, \vec{\alpha}_{it}^{po}, pp_{it+1}) \\
&\qquad \cdot dF(P_{it+1})\Bigg) \\
&\quad + \exp\Bigg(\beta_0 + \beta_P \cdot P_{it} + \beta_Q \cdot Q_{it} + \eta \cdot pp_{it} + \delta \\
&\qquad \cdot \int_{P_{it+1}} EV_{cs}(P_{it+1}, \vec{\alpha}_{it}^{po}, pp_{it+1}) \cdot dF(P_{it+1})\Bigg)\Bigg\}.
\end{aligned} \tag{17}
$$

6.3. Terms in the Likelihood Function

On the basis of our assumptions for the idiosyncratic shocks in both stages, the probability that a customer clicks on a deal and the probability of making a conditional purchase can be derived as follows, using their corresponding choice-specific counterparts:

$$P_{it}^{c} = \frac{\exp(\bar{v}_{cs,it}^{1})}{\exp(\bar{v}_{cs,it}^{0}) + \exp(\bar{v}_{cs,it}^{1})}, \tag{18}$$

$$P_{it}^{pur|c} = \frac{\exp(\bar{v}_{ps,it}^{1})}{\exp(\bar{v}_{ps,it}^{0}) + \exp(\bar{v}_{ps,it}^{1})}. \tag{19}$$

- The probability of observing a consumer click and purchase the deal is ($d_{it}^{C} = 1$, $d_{it}^{P}=1$): $P_{it}^{c} \cdot P_{it}^{pur|c}$.
- The probability of observing a consumer click the deal but not make a purchase is ($d_{it}^{C} = 1$, $d_{it}^{P}=0$): $P_{it}^{c} \cdot (1 - P_{it}^{pur|c})$.
- The probability of observing a consumer not clicking is ($d_{it}^{C} = 0$, $d_{it}^{P}=0$): $1 - P_{it}^{c}$.

The probabilities derived above can now be used to form the likelihood function using the decision indicator at both stages.

$$
\begin{aligned}
L_{it} &= \left(P_{it}^{pur|c} \cdot P_{it}^{c}\right)^{d_{it}^{P}} \\
&\quad \times \left(\left[\left(1 - P_{it}^{pur|c}\right) \cdot P_{it}^{c}\right]^{d_{it}^{C}} \cdot \left[1 - P_{it}^{c}\right]^{1-d_{it}^{C}}\right)^{1-d_{it}^{P}}
\end{aligned} \tag{20}
$$

The total log-likelihood function is simply $LL = \sum_i \cdot \sum_t \log L_{it}$.

6.4. Estimation and Identification

We adapt the two-stage dynamic learning model from the Keane and Wolpin (1994) method, which was developed in Crawford and Shum (2005). It assumes an infinite horizon and is based on the nested fixed-point algorithm (Rust 1987). As in Rust (1987), given a set of parameter values, the inner loop computes the value function and evaluates the likelihood function; the outer loop searches for parameters that maximize the likelihood function. In particular, we use the Berndt-Hall-Hall-Hausman (Berndt et al. 1974) algorithm to conduct the outer loop optimization. When computing the value functions, we adopt the approximation method used in Crawford and Shum (2005).[31]

6.4.1. Empirical Identification. We now briefly consider the model's structural features and variations in the data that help to identify the model parameters. From the purchase decision, variations in deal prices and quality levels, in addition to the number of deals previously purchased, help us identify the price (β_P), quality (β_Q), and previous purchase number (η) coefficients. As in Koulayev (2010), search cost (c_{it}) is identified through variations in click decisions (i.e., click or not click), together with the information on previously clicked deals. Unlike Seiler (2013), in which the search stage is latent, we observe individual click behavior in the first stage, thereby allowing for identification of the search cost parameter. The rate of the declining click trend and the increasing conditional purchase trend, in addition to the stationary quality level distribution (shown previously), helps us identify the learning parameters, that is, the Dirichlet priors ($\vec{\alpha}_{it=0}^{pr} = \{\alpha_{it=0}^{pr,1}, \alpha_{it=0}^{pr,2}, \ldots, \alpha_{it=0}^{pr,N}\}$). Because the discount factor is difficult to estimate (cf. Rust 1994, Erdem and Keane 1996), we fix the value of δ at 0.995.

6.4.2. Dirichlet Assumption. To gain a better understanding of the Dirichlet learning process, we investigate how well it can recover the true distribution under different conditions. We vary (1) the true distribution, (2) the number of grids in the prior, (3) the updating times, and (4) the distributional assumption of the prior. The results are shown in Online Appendix E. We find that no matter what the prior or the true distribution are, the Dirichlet learning process can successfully recover the true distributions after just a few rounds of updating.

6.4.3. Monte Carlo Simulation. We conduct Monte Carlo simulations to verify our estimation algorithm and to examine our estimation strategy. The main objective is to confirm that our estimation procedure can recover the model's unknown parameters. The simulation is run as follows. We randomly generate a panel data of 500 agents with 200-day observations. For each day, each agent receives a newsletter with one featured deal. Then, the agent makes a decision about whether to click on the deal; if a click occurs, then the agent has to decide whether to make the purchase. An observation includes data on the deal price, quality level, click decision, and purchase decision. The price and quality are generated from a normal distribution. The stochastic term in the utility specification is randomly drawn from a standard type

I extreme value distribution, and each observation has a different draw for each stage–decision combination. The flow utility gained at each stage follows the definition in Section 6.1.1.

Table 7 presents the results of the Monte Carlo simulation. Column (1) provides the search cost and utility parameters used to generate the data. Column (2) presents the mean and standard deviation of our model estimates across 30 replications. The results show that we are able to recover the true parameters using our model specification and estimation strategy.

7. Empirical Application

7.1. Data Set Used in Structural Model Estimation

For the empirical estimation, we focus on featured *restaurant* deals only.[32] The featured deal is not always a restaurant deal; when there is no featured restaurant deal on some days or in some cities, consumers do not have an observation recorded. Hence, these instances do not enter the estimation. After all, when the choice is not available, we cannot model the corresponding consumer behavior. Our structural model analysis is therefore conditional on a consumer receiving a featured restaurant deal on a given day. Given that the newsletters were not targeted at the consumer level during the data collection period, conditioning on the arrival of a featured deal still captures consumers' reactions to the featured deal.

Among the 18 categories of deals predefined by Groupon, restaurant deals are the largest, accounting for approximately 20% of all deals. Deals are differentiated products that may vary in different aspects. Focusing on one major category simplifies our analysis. We leave the extension of our model to multiple categories to future research.

Summary statistics for the data used for estimation are presented in Table 8. In total, we have 108,698 observations from 10,485 subscribers. The average price of restaurant deals is approximately 13.23 U.S. dollars (USD), with an average discount of 51%. The empirical distribution of the price after discount, the original price, and the discount rate for restaurant deals are shown in Figure 10, (a)–(c). Owing to the lack of variation in the discount rate, we do not include it as a variable in the estimation. The mean quality of deals is around quality bin 6. Figure 10(d) presents the empirical distribution of restaurant deals' quality. It shows a bimodal pattern, which is captured well by our Dirichlet assumption. In terms of subscribers' characteristics, on average, consumers had been registered with Groupon for approximately one month, with some joining early in our data period and others joining later. The click rate in the data was approximately 4.6%; the conditional and unconditional purchase rates were 11% and 0.49%, respectively.[33] We also include the summary statistics of the variables from all of the categories in Table 8 as comparison.

Table 7. Monte Carlo Simulation

	(1) True coefficient	(2) Coefficient	Standard error
Utility			
Constant	−3	−2.830	0.360
Price/100	−3	−3.028	0.101
Quality	2	1.976	0.106
Past purchase η	−0.05	−0.048	0.014
Baseline cost: k	1	0.985	0.028
Dirichlet priors			
α_1	1	1.011	0.331
α_2	2	1.996	0.428
α_3	1	0.981	0.318
Observations	—	100,000	

Note. Column (2) presents the mean and standard error of the parameter estimates across 30 replications.

7.2. Findings

We estimated the two-stage dynamic model with Dirichlet learning using the data and present the results in Table 9. The results of the proposed model are intuitive. Lower prices (coef. = −0.351, s.e. = 0.041) and better quality (coef. = 0.190, s.e. = 0.027) are related to higher purchase likelihood. The cost of search is positive and significant (coef. = 3.141, s.e. = 0.012). This indicates that consumers incur a cost when they click on the deal to go to the web page; this cost counterbalances the benefit of learning, which results in a decreasing click rate. The effect of previous purchases is negative and significant (coef. = −0.038, s.e. = 0.014). This effect shows that "inventory" discourages purchases on Groupon (Seiler 2013). For the estimation of Dirichlet learning, we divide the priors into nine grids, from low to high values. Grids 1 through 8 are not different from zero. Grid 9 is significant and has a large value (coef. = 1.407, s.e. = 0.183). This suggests that according to consumers' prior beliefs, the merchants on Groupon are of high quality. They adjust this belief as they learn about the clicked deals. This makes sense because the data concern new subscribers. If consumers do not think Groupon deals are high quality, then they are unlikely to register to shop on Groupon at all. As a result, two aspects of the learning model help generate the data trends we observe. First, consumers learn to have a more precise idea of what the quality distribution looks like; second, they start off being overly optimistic about future deals.

As a comparison, we used the two-stage dynamic search model without learning as a benchmark (Seiler 2013). We assumed that consumers do not learn about the distribution of deals: consumers are forward-looking and maximize their expected present-discounted

Table 8. Summary Statistics

	Restaurant				All categories			
	Mean	Min	Max	SD	Mean	Min	Max	SD
Deal								
Price (USD)	13.23	2	1,500	17.47	37.28	1	12500	109.64
Discount (%)	51	40	83	2.62	55.79	19	100	9.51
Original price (USD)	25.99	3.33	3,000	34.74	740.60	3	999,999	25,308
Quality[a]	5.98	1	9	1.66	5.35	1	9	1.67
User								
Tenure (days)	32.27	0	89	21.58	31	0	89	21.96
If click	0.0457	0	1	0.2089	0.0483	0	1	0.2144
If purchase	0.0049	0	1	0.0696	0.0038	0	1	0.0615
If purchase (conditional)[b]	0.1065	0	1	0.3085	0.0785	0	1	0.2690
Cumulative click	0.4001	0	20	0.8585	1.77	0	84	0.43
Observations	108,698				527,298			

Note. SD, standard deviation.
[a]We used deep learning techniques to construct the measure of deal quality and divided the numbers into 9 categories, with 1 indicating lowest and 9 the highest quality.
[b]Conditional on clicking.

Figure 10. (Color online) Deal Characteristics: Restaurant Deals

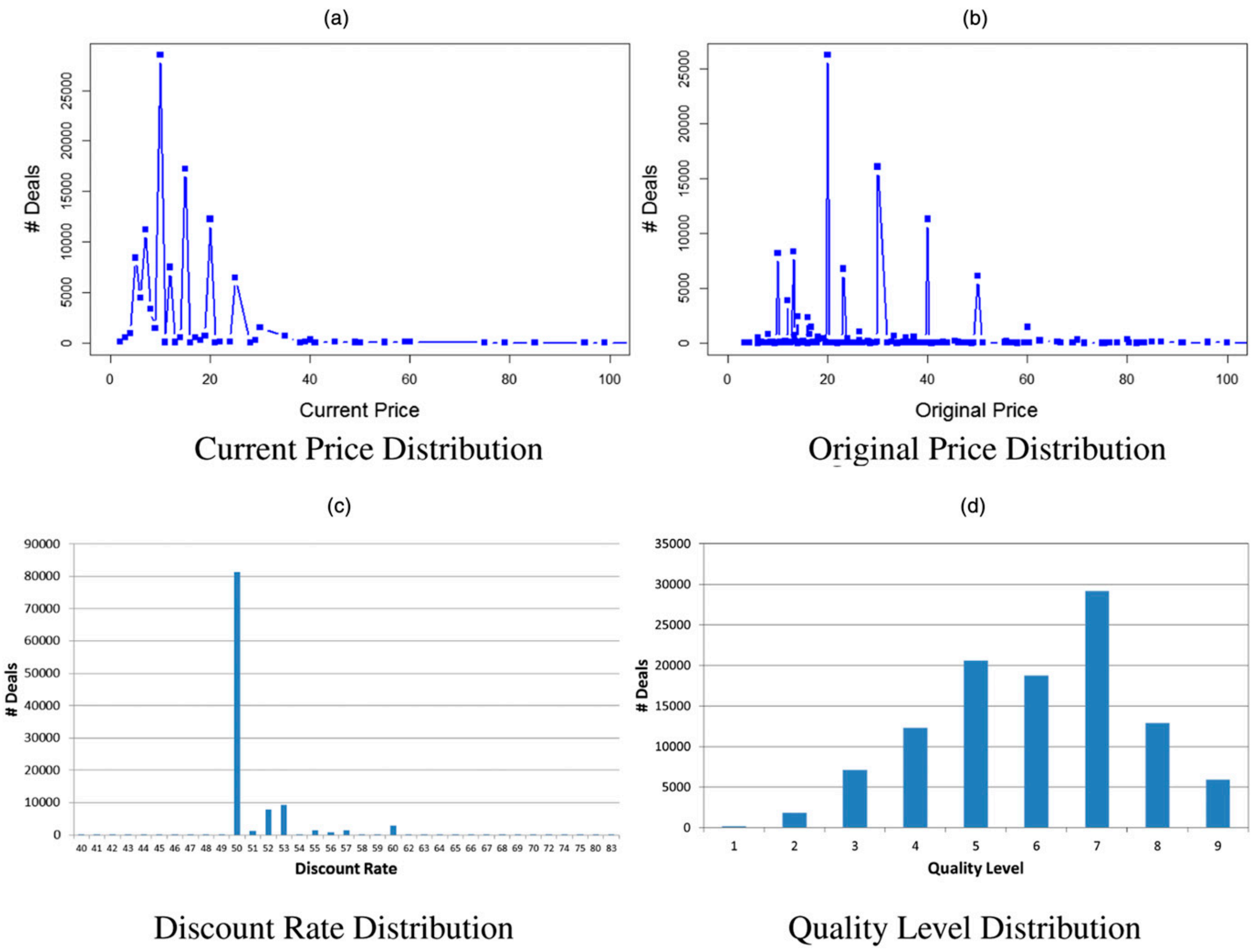


Table 9. Estimation Results

	Forward-looking without learning[a]		Forward-looking with Dirichlet learning	
	est.	s.e.	est.	s.e.
Utility				
Constant	−2.628***	0.935	−2.592***	0.906
Price/100	−0.356***	0.042	−0.351***	0.041
Quality	0.203***	0.029	0.190***	0.027
Past purchase: η	−0.041***	0.015	−0.038***	0.014
Baseline cost	3.082***	0.020	3.141***	0.012
Dirichlet priors				
α_1	—	—	1.003E-05	2.52E-4
α_2	—	—	1.001E-05	2.51E-4
α_3	—	—	1.002E-05	2.57E-4
α_4	—	—	1.002E-05	2.53E-4
α_5	—	—	1.000E-05	2.50E-4
α_6	—	—	1.010E-05	2.55E-4
α_7	—	—	1.012E-05	2.56E-4
α_8	—	—	1.009E-05	2.60E-4
α_9	—	—	1.407***	0.183
Observations	108,698		108,698	
-LL	22,304.365		22,172.101	
AIC	44,618.73		44,372.202	
BIC	44,666.71		44,506.551	

[a]We assume consumers are forward-looking without learning the quality distribution, and the prior belief is what we estimated using forward-looking with Dirichlet learning model (i.e., a high prior).
***$p < 0.01$.

sum of utilities; the click and purchase decisions are determined by the price and quality of the current deal and the number of previous purchased deals. Consumers are still assumed to face an intertemporal tradeoff (purchase now or wait for the next deal), and although consumers have knowledge of the price distribution and have beliefs about the quality distribution, they still face uncertainty about future deals and recognize that purchasing now lowers their utility from future purchases. Most state variables remain the same, only that now we do not have state variables related to learning (i.e., α_{it}^{pr} and α_{it}^{po}). To be specific, the click-stage state variables are $x_{cs,it} = \{P_{it}, pp_{it}\}$ and the purchase stage state variables are $x_{ps,it} = \{P_{it}, Q_{it}, pp_{it}\}$. In this model, we assume that consumers know the empirical distribution of deal prices and that they have a high prior belief (what we estimated using our proposed model as in Table 9) about the deal quality distribution. However, because consumers are not learning, this belief is not updated. As in the proposed model, both P_{it} and Q_{it} are observed by both consumers and researchers. Here, P_{it} comes from a normal distribution with the empirical mean and empirical standard deviation, and Q_{it} comes from the empirical categorical distribution. The estimates are comparable to those of the proposed model, with negative effects of price and number of previous purchased deals, and positive effects of quality and search cost.

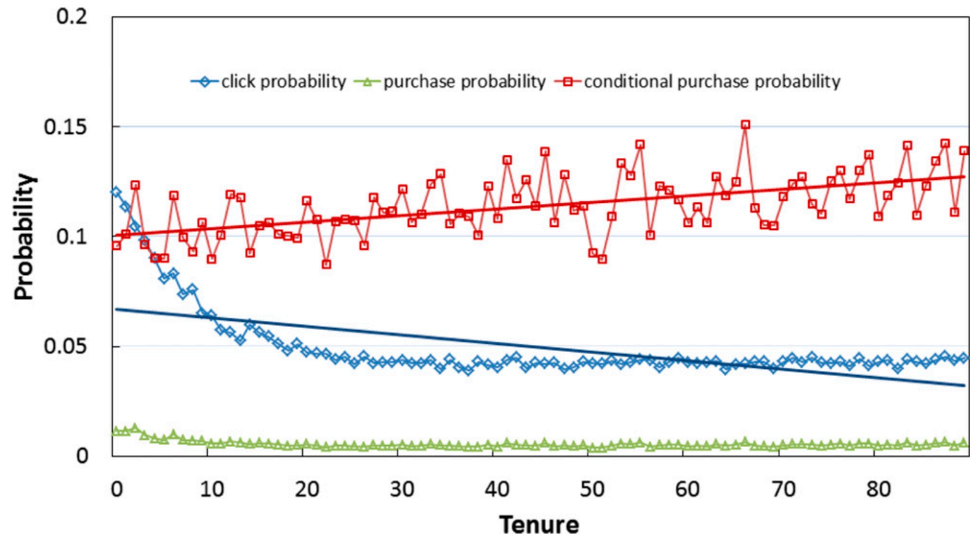
The model with learning has a higher likelihood and lower values of Akaike information criterion (AIC) and Bayesian information criterion (BIC), indicating a better fit to the data. Allowing consumers to learn the quality distribution over time seems to better reflect the variation observed in the data than a model without learning. On the basis of the estimation results, we then simulate each individual's click and purchase behaviors. As in Figure 7, we plot the click probability and conditional purchase probability against the tenure with Groupon in Figure 11. We can see that using the estimated model, we are able to successfully replicate the two observed data patterns.

Consumers click less because the incentive to learn about quality decreases. However, once the uncertainty of the quality in the quality distribution is resolved through learning, consumers purchase deals with high quality on Groupon. Indeed, this is what the simulation from our proposed model indicates; specifically, with beliefs about quality getting closer to the true distribution, the quality of deals purchased improves with tenure (Figure 12). This is a further validation of the model.

7.3. Model Without Learning

What is the benefit of incorporating learning in our context? As we have already shown, the model that does not incorporate subscriber learning does not fit the data as well as the proposed model. Further, simulating data from the model without learning reveals that the model is unable to recover the two observed data patterns—that is, the declining click

Figure 11. (Color online) Simulated Data Trends Using the Model with Learning



probability and the increasing conditional purchase probability. Figure 13 illustrates the simulation results. We simulate the behavior of 500 agents for a 60-day period and repeat this process 1,000 times. Within each iteration, we generate data using the structural model specification[34] with the estimates of the proposed model. We are unable to replicate the two patterns observed in the data.

Without learning, consumers do not exhibit a declining tendency to click on the newsletter's deals, because their expectation or knowledge of the quality of future deals is not updated; thus there is no benefit to searching as long as there is a search cost. Furthermore, without learning, the subscribers are unable to choose better deals; purchase probability conditional on clicking thus does not change over time. Although the estimates are similar, the proposed two-stage dynamic search model with learning is able to capture the underlying patterns in the data, whereas the model without learning is unable to generate either of the observed data patterns.

7.4. Robustness Checks

7.4.1. Alternative Measure of Quality. As a robustness check, we construct an alternative deal quality measure also by using deep learning techniques. This measure is based on reviewer comments on deals, as proposed in Liu et al. (2017). It follows the same generation procedure as the proposed quality measure. Both generate similar findings. However, our proposed measure is better in terms of capturing all of the observable information on the deal web page and generating a better model fit.[35] After all, the review comments are from the subset of customers that purchased the deal; hence demand still has some influence.

Figure 12. (Color online) Simulated Mean and Confidence Interval of Quality of Purchased Deals

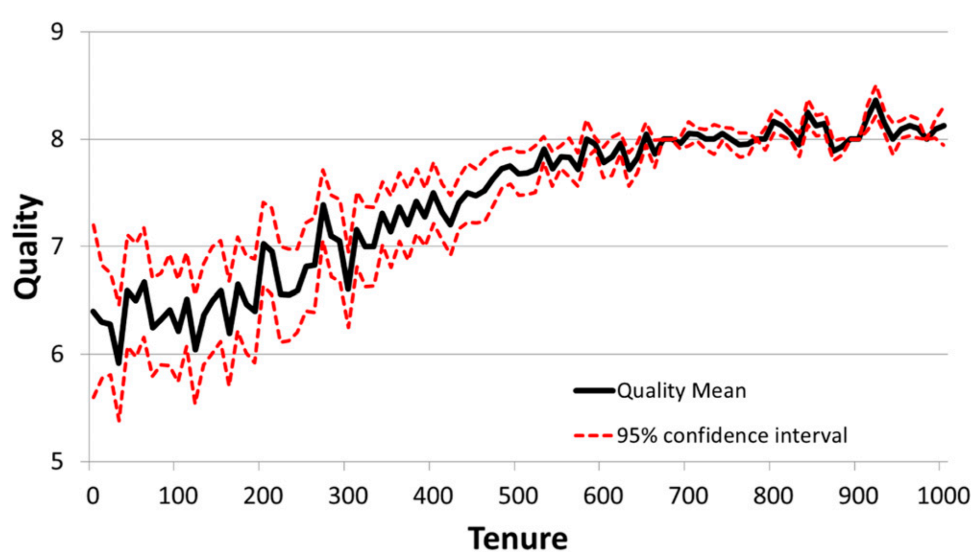
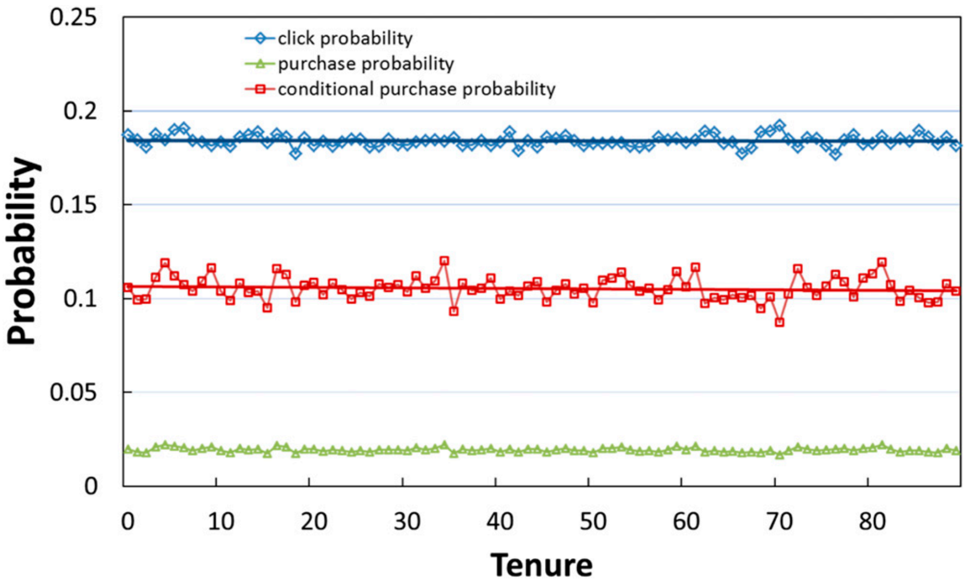


Figure 13. (Color online) Simulated Data Trends Using the Model Without Learning



7.4.2. Stationarity. ***7.4.2.1. Market Stationarity.*** To rule out the possibility that the trends observed in the data are driven by heterogeneity in the new Groupon subscribers across different periods, we focus on the subset of new subscribers who joined only in January (the beginning of the sample period) and run the model on this part of our data. The results, shown in Table 10, are almost identical to those generated using all new subscribers across our sample period. This indicates that the profile of customers joining did not change much during our data period. When looking across a longer duration (e.g., multiple years as opposed to our three months), it is certainly possible that cross-customer heterogeneity explains the two patterns observed in the data. Nevertheless, we find that in the context of our short-duration data, our explanation based on within-customer behavioral change is viable.

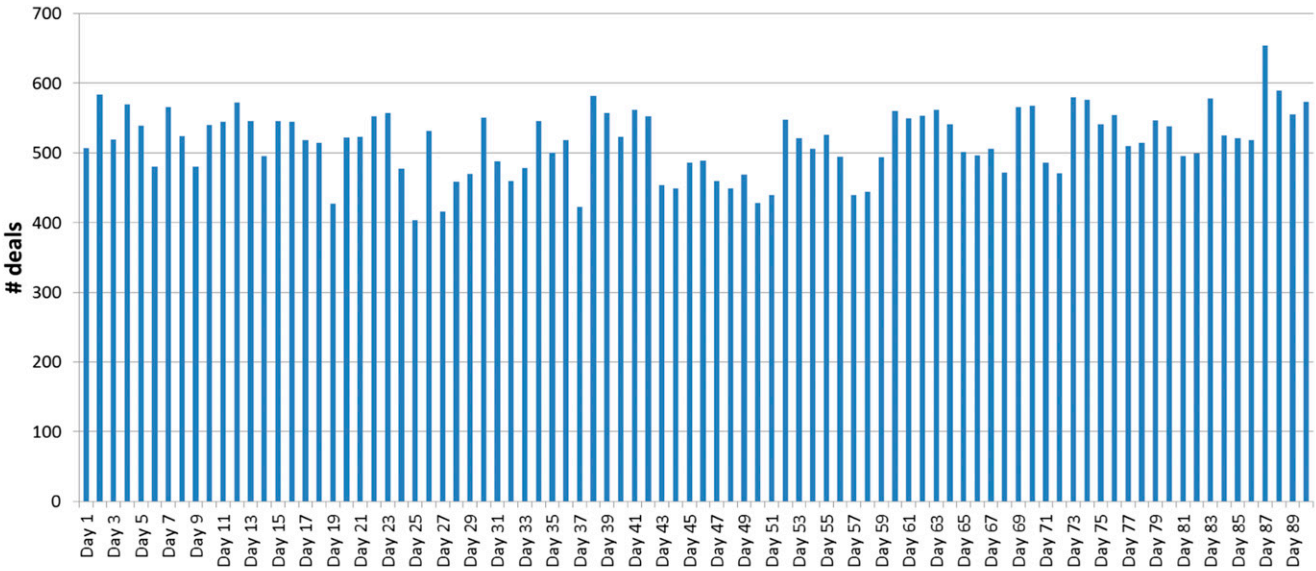
Table 10. Robustness Check: Customers Joined in January

	Forward-looking without learning[a]		Forward-looking with Dirichlet learning	
	est.	s.e.	est.	s.e.
Utility				
Constant	−2.621***	1.108	−2.580***	1.023
Price/100	−0.354***	0.048	−0.352***	0.052
Quality	0.210***	0.030	0.189***	0.025
Past purchase: η	−0.042***	0.013	−0.039***	0.012
Baseline cost	3.118***	0.036	3.147***	0.025
Dirichlet priors				
α_1	—	—	1.01E-05	2.50E-4
α_2	—	—	1.00E-05	2.48E-4
α_3	—	—	1.02E-05	2.51E-4
α_4	—	—	1.00E-05	2.50E-4
α_5	—	—	1.00E-05	2.50E-4
α_6	—	—	1.02E-05	2.51E-4
α_7	—	—	1.03E-05	2.49E-4
α_8	—	—	1.00E-05	2.58E-4
α_9	—	—	1.486***	0.364
Observations		65,972		65,972
-LL		11,911.14		11,841.54
AIC		23,931.08		23,711.09
BIC		24,058.44		23,838.45

[a]We assume consumers are forward-looking without learning the quality distribution, and the prior belief is what we estimated using forward-looking with Dirichlet learning model (i.e., a high prior).

***$p < 0.01$.

Figure 14. (Color online) Number of Deals over Time



7.4.2.2. Supply-Side Stationarity. We illustrate supply-side stationarity from three perspectives: (1) stability in the numbers of deals; (2) stability in the quality of deals; and (3) stationarity in the selection rule that determines the featured deal.

Throughout the sample period, each customer received one featured deal per day. The number of featured deals thus remained constant. Furthermore, the number of deals in each day was very stable during the three months of our observation period. We plot the total number of unique deals over time (Figure 14). There is no significant time trend. We also regress the number of total unique deals over time, and the result rejects a significant time trend in the number of unique deals (coef. = 0.27, $p = 0.16$).

Figure 9 demonstrates that the deals' quality was stable throughout the three months of our observation period; and the regression result reveals no significant time trend (coef. = −0.00083, $p = 0.27$) in the quality of deals.

That said, even if there is a trend on the supply side, it should not affect our analysis. For example, if deal quality improves over time, then this likely increases the purchase likelihood without decreasing click probability. If deal quality worsens, then this likely decreases the click probability without increasing purchase likelihood.

Finally, to illustrate the stationarity of the featured deal selection rule, we split the time period into two halves and regress the variable "if_featured" on deal characteristics such as original price, discount rate, Yelp rating, whether the deal is of limited quantity, the maximum number of purchases, deal category, and deal-day type (i.e., one-day deal, two-day deal, etc.) for each of the two time periods. The results are shown in Table 11. Although the featured deals tend to be restaurant deals with higher discount rates and larger maximum quantities, this rule is quite stable over time. The stationarity of the selection rule thus seems like a reasonable assumption in our case.

8. Counterfactual Analysis

Because the two main features of our paper are search and learning, we conduct counterfactual analyses corresponding to each of them using our proposed model. In particular, we investigate the roles of search costs and learning in influencing Groupon's revenues. To compute outcomes for Groupon, we assume that each successfully recommended deal purchased by the user creates 50% net revenue for the firm with the current deal price.[36] In this way, we can compute the revenue under each counterfactual.

8.1. Counterfactual Analysis on Search Cost

To evaluate the magnitude of search cost, we first set the estimated search cost parameter[37] to zero. Next, we compute how much the price of the restaurant has to change to compensate for the elimination of the search cost. This change in the deal price gives us a dollar metric for the magnitude of the search cost.

The above counterfactual analysis is conducted in the following steps. First, using the structural estimates including the estimated search cost $\hat{c}$, we simulate how consumers click and purchase using the data set from the paper. Second, we set the search cost parameter estimate to zero while keeping the other estimated parameters unchanged. Meanwhile, we change the deal price level by giving each price the

Hu, Dang, and Chintagunta: *Search and Learning*
Marketing Science, 2019, vol. 38, no. 4, pp. 609–642, © 2019 INFORMS

Table 11. Featured Deal Selection

	Dependent variable: if featured deal		
	All deals	First half	Second half
	(1)	(2)	(3)
Origin price	−0.000 (0.000)	0.000 (0.000)	−0.000 (0.000)
Discount	0.0005** (0.0002)	0.0002*** (0.0001)	0.001*** (0.0003)
Yelp rating	0.001 (0.002)	0.001 (0.002)	0.001 (0.002)
If limited quantity	−0.001 (0.004)	0.001 (0.005)	−0.003 (0.005)
Max purchase quantity	0.0003** (0.0001)	0.001** (0.0002)	0.0004** (0.0002)
Time left to deadline	0.001*** (0.000)	0.001*** (0.000)	0.002*** (0.000)
Category: Automotive	0.009 (0.011)	0.026 (0.017)	0.008 (0.016)
Category: Beauty and spas	−0.002 (0.006)	0.008 (0.008)	−0.006 (0.008)
Category: Education	−0.036*** (0.011)	−0.040*** (0.015)	−0.046*** (0.016)
Category: Financial Services	0.089 (0.068)	0.088 (0.101)	0.084 (0.111)
Category: Food and drink	0.009 (0.007)	0.010 (0.009)	0.011 (0.009)
Category: Health and fitness	−0.016** (0.007)	−0.012 (0.010)	−0.026*** (0.010)
Category: Home services	−0.018 (0.012)	−0.012 (0.017)	−0.022 (0.017)
Category: Legal services	0.064 (0.173)	0.058 (0.181)	0.068 (0.169)
Category: Nightlife	0.005 (0.013)	−0.013 (0.019)	−0.004 (0.018)
Category: Pets	−0.033 (0.027)	−0.016 (0.040)	−0.015 (0.041)
Category: Professional services	−0.004 (0.010)	−0.008 (0.014)	−0.004 (0.014)
Category: Public services	0.027 (0.110)	0.029 (0.125)	0.015 (0.177)
Category: Real estate	−0.257* (0.141)	−0.216 (0.143)	0.017 (0.247)
Category: Restaurants	0.025*** (0.006)	0.033*** (0.008)	0.024*** (0.008)
Category: Shopping	0.009 (0.006)	0.015* (0.008)	0.005 (0.008)
Category: Travel	0.009 (0.011)	0.004 (0.016)	−0.002 (0.016)
Day type: 2 day	−0.017*** (0.004)	−0.014** (0.006)	−0.021*** (0.006)
Day type: 3 day	−0.005 (0.004)	−0.005 (0.006)	−0.013** (0.006)
Day type: 4 day	0.012 (0.007)	0.014 (0.011)	0.011 (0.011)
Day type: 5 day	0.012 (0.014)	0.044** (0.021)	0.016 (0.021)
Day type: 6 day	0.049 (0.031)	0.013 (0.051)	0.066 (0.045)
Day type: 7 day	0.008 (0.018)	0.023 (0.027)	0.016 (0.027)
Day type: 8 day	0.061* (0.034)	0.029 (0.048)	0.069 (0.047)
Day type: 9 day	0.103 (0.244)	0.100 (0.202)	0.115 (0.232)
Day type: 10 day	0.106** (0.049)	0.108 (0.069)	0.126 (0.079)
Day type: 15 day	0.180 (0.244)	0.160 (0.203)	0.200 (0.212)
Day type: 17 day	0.093 (0.244)	0.095 (0.247)	0.082 (0.247)
Day type: 29 day	0.035 (0.174)	0.021 (0.161)	0.055 (0.247)
Observations	24,520	12,244	12,256

Note: Standard errors in parentheses.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

scale factor γ.[38] The simulated (unconditional) purchase rate is 0.41% in this case. The price scale factor changes the price level faced by each consumer in the data so it affects the final purchase rate. We let the program loop through various scale factor values and find "the" scale factor γ^* that produces the same purchase rate as with the estimated search cost. Finally, the dollar equivalent search cost is calculated by the final price scale factor γ^* multiplied by the average price, less the average price (i.e., $cost = (\gamma^* - 1) \cdot \overline{price}$). The search cost dollar value that we obtained by using the aforementioned method is approximately \$389 when we include all of the new subscribers in our data, regardless of whether they had ever clicked on or purchased a deal on Groupon (see the first column in Table 12).[39] Changing the search cost affects the frequency of purchase. So to monetize search cost, we want to see how much the price of the deal has to change to result in the same purchase rate when we eliminate search cost.

The search cost in our context is large. This has to do with the relevance of Groupon deals to the subscribing customer. Recall from our description of the data that we consider all new subscribers to the Groupon platform during the focal time period. We include all subscribers regardless of whether they click on or purchase deals in any category (i.e., not just restaurants) on the platform. Note that this is in contrast with most studies in the literature that have estimated search costs. In our case, however, because Groupon determines what the consumer sees, the majority of the deals received may not be relevant to the recipient. Additionally, because we do not filter the new subscribers, there could be several for whom

Table 12. Counterfactual Analysis: Search Cost Simulation

	(1) Original estimation sample	(2) People who clicked a restaurant deal at least once	(3) People who clicked a restaurant deal at least once and made at least one purchase on Groupon	(4) People who clicked restaurant deals at least three times and made at least one purchase on Groupon
Click rate	0.0457	0.1251	0.1521	0.2600
Purchase rate	0.0049	0.0133	0.0550	0.0644
Constant	−2.592	−1.3080	−0.2118	−0.8205
Price /100	−0.351	−1.5170	−2.2921	−3.5350
Quality	0.19	0.0178	0.0354	0.1264
Cost	3.141	3.2355	3.0459	2.5189
Past purchase	−0.038	−0.0152	−0.0082	−0.0137
Prior: 9th grid	1.407	2.6085	7.7301	11.6739
Cost in dollar value (divided by price coef.)	$895	$213	$133	$71
Cost in dollar value (structural simulation)	$389	$98	$56	$26

Note. The click rates and purchase rates reported here are the samples average values.

the deals are not relevant; hence they do not click or purchase. One way in which the model reconciles such behavior is by assigning a very high search cost that then dissuades the consumers from searching and clicking. To demonstrate the magnitude of search cost computed in our context changes with the relevance of the deals, we filtered the data in a variety of ways and then computed the counterfactual search cost as illustrated above. The results are reported in columns (2) to (4) of Table 12. One of the subsamples includes only consumers who have clicked on a restaurant deal at least once [column (2) of Table 12]. Not surprisingly, the click rate found in this subsample is much higher than that in the original sample (original sample, 4.57%; current subsample, 12.5%). We estimate the model according to the subsample and then calculate the search cost following the aforementioned strategy. The magnitude of the search cost estimated is $98. If we further exclude people who never made a purchase on Groupon (not just for restaurant deals, but for deals in any category), then the magnitude of the search cost is reduced to approximately $56 [see the last row in column (3) of Table 12]. If we only include people who have clicked restaurant deals at least three times and have made at least one purchase on Groupon (again, in any category), then the search cost is further reduced to $26. The simulation therefore indicates that the magnitude of the search cost (as computed in our context) changes with the relevance of the deals.

Another counterfactual analysis of search cost is to examine the impact of search on learning. The purpose of search in our model is to learn about the deal quality distribution, which allows consumers to identify a good deal and make a purchase decision. To study the impact of search costs on learning in our context (where consumers incur the cost to learn about the quality distribution), we used the amount of time that the consumer takes to achieve a certain level of learning under different levels of search cost. We simulate the average time that it takes for consumers to learn approximately 50%, 75%, and 90% of the true quality distribution of deals when search cost is fixed at the estimated value and when it is eliminated.[40] The results are in Table 13. We find that people are learning faster when we eliminate the search cost. We define the time it takes to learn the true mean quality of the distribution of deals at the estimated search cost parameter as the benchmark. On average, it takes approximately 45% of that time to learn 50% of the true quality. When we eliminate the search cost, it only takes 34% of the total time. In general, eliminating search costs saves 11%–16% of the time to learn. However, from the company's perspective, on the basis of the counterfactual analysis of investigating the role of learning on company revenue (the next section), expedited learning may not be a good idea for Groupon (because the mean of the prior mean distribution is higher than the mean of the distribution of true deal quality). So we do not recommend that the company should cut search cost. Instead, sending out newsletters with featured deals from more varied categories, rather than focusing on a certain category, could increase the search cost owing to the potentially low relevance of the deal. This is likely to benefit the firm.[41]

8.2. Counterfactual Analysis on Consumer Learning

For the counterfactuals about learning, there are two key aspects to consider: (1) the role of the prior belief about the quality distribution; and (2) the role of the true quality distribution of deals that generates the signals that the consumer receives each day. These signals, in conjunction with prior belief, influence the

Table 13. Counterfactual Analysis: Change Search Cost

	Time to learn 50% of true quality (%)	Time to learn 75% of true quality (%)	Time to learn 90% of true quality (%)
With cost[a]	45	65	85
Without cost	34	55	69

Notes. We simulate the behavior of 500 agents for a 1,000-day period, and we repeat this process 1,000 times. The times to learn about a certain level of true distribution are the average of 30 simulations. Note that we normalize the time by the time needed to learn 100% true quality when search cost = 3.

[a]Benchmark model using estimates in Table 9 (i.e., search cost = 3).

speed with which consumers learn about the true quality distribution of deals. Because our data pertain to consumers who actually sign up to receive Groupon deals, it is reasonable for these consumers to begin with a prior belief about the high quality of deals on the website. Our empirical results corroborate this intuition. The following question then arises. What happens if we change the nature of the prior distribution? Firms can influence consumers' prior beliefs through advertising. They can also follow Groupon's practice by creating a web page to selectively summarize real deals appearing on Groupon.[42] In the counterfactual, we consider the following scenarios in which consumers begin with (1) low prior beliefs; (2) uniform prior beliefs (the noninformative case); and (3) rational prior beliefs. Table 14 provides the results from this counterfactual; we express the results relative to the base case, which is the high prior distribution. The table tells us that in all cases, the firm is worse off than in the current situation. The click rate is 60% lower, and the firm makes $1,840 less in revenue for low prior case. For the uniform case, the click rate is 69% lower, and the revenue $2,100 less. Additionally, for the rational prior, the click rate is 67% lower, and the revenue loss is $2,043 relative to the high prior case. Initially, the high prior belief about quality encourages the consumer to click more often, with the clicking resulting in finding a potential match with a specific deal. In the absence of this incentive, both clicks and purchases decline, thereby lowering the firm's revenues. Therefore, by collectively listing on the top deals on their website, Groupon does seem to be setting up a high prior belief among consumers. Furthermore, advertising may help acquire new customers while simultaneously affecting consumers' prior beliefs about deal quality on the website.

Next, we look at the impact of different true quality distributions. This counterfactual is relevant because the daily deal website decides what deals to include and what to list as a featured deal. They also negotiate discount rates with the retailers. The consumers begin with a high prior belief, so there are two possibilities. The first is if the true quality of deals offered on Groupon match up with this high prior belief. In this case, the speed with which consumers learn is not relevant because when coupled with the high prior belief, high-quality deals generate more revenues and clicks for the daily deal website. However, if the mean of the true quality distribution is lower than the consumer's prior belief, then the firm is better off if the distribution of deals has a higher variance than a lower variance. A higher variance is preferable to a lower variance (with the same mean) because this slows down learning about the true quality distribution. In other words, getting a few signals (i.e., draws from this distribution) may not be enough for consumers to resolve their uncertainty regarding the deal quality distribution. This, combined with the higher prior belief, results in more clicks and revenues for Groupon. Thus, in this counterfactual we hold the average quality of deals fixed but change the variance of the distribution by examining a few scenarios: (1) the same quality distribution as in the data; (2) a distribution with the same mean quality as (1) but with higher variance; and (3) a distribution with the same mean quality as (1) but with lower variance. The results are shown in Table 15.

Table 14. Counterfactual Analysis: The Role of Prior Belief

Prior	Relative change in click rate[a] (%)	Relative change in purchase rate (%)	Difference in revenue (USD)
Low	−60.24	−61.48	−1,840
Empirical[b]	−67.31	−68.18	−2,043
Noninformative	−69.05	−70.10	−2,100

Notes. We simulate the behavior of the consumers in our estimated sample given the estimates in Table 9 of the paper, changing only the initial prior. The result is the average of 1,000 simulations.

[a]All the results are the relative change to the baseline: the high prior case (i.e., using the prior we estimated from the data).

[b]The deal quality distribution in the data set.

Table 15. Counterfactual Analysis: The Role of the True Quality Distribution

Quality distribution	Click rate	Purchase rate	Revenue (USD)
Empirical [a]	0.0488	0.0016	3,070
High variance [b]	0.0540	0.0021	4,000
Low variance [c]	0.0486	0.0015	2,885

Notes. We simulate the behavior of 5,000 agents for a 60-day period given the estimates in Table 9 of the paper. The result is the average of 1,000 simulations. Note that the revenue column stands for an average total revenue of 5,000 new subscribers for a 60-day period. The three quality distributions investigated here have the same mean (5.84) but different variance: empirical (2.86) vs. high variance (15.28) vs. low variance (1.13).

[a]The same quality distribution as in the data.

[b]With deal quality level either 1 or 9.

[c]With the deal quality level being 4, 5, 6, or 7.

We find that compared with the quality distribution in the data, the quality distribution with high variance has the highest click rate, purchase rate, and revenue (5.40%, 0.21%, and $4,000, respectively). Additionally, the quality distribution with low variance has the lowest click rate, purchase rate, and revenue (4.86%, 0.15%, and $2,885). This counterfactual indicates that if the mean of the prior distribution for a consumer is high quality and the true mean is of lower quality, Groupon is better off having a few very-high-quality and a few very-low-quality deals that average out to an intermediate quality rather than having all deals being of that intermediate quality level (where the speed of learning is high).

9. Discussion

Analysts have been gloomy about the prospects of daily deal websites and have attributed this, at least in part, to what they call "daily deal fatigue" (Dholakia and Kimes 2011). The empirical data presented in this paper support some of the behavior underlying the phenomenon of deal fatigue; however, we also found a reason for some optimism. In particular, we observed a declining probability that a consumer clicks on a merchant in the emailed newsletter over time but an increasing probability of the consumer making a purchase conditional on clicking. The former trend is consistent with the notion of "fatigue," whereas the latter is a potential source of optimism for daily deal websites, such as Groupon. Together, these patterns suggest that although the consumer becomes more selective in terms of exploring the offers received, (s)he is more likely to yield the site revenue as time passes.

We proposed a model of search and learning that tries to explain these observations. On the basis of the model estimates, we were able to replicate the two observed patterns observed in the data.[43] On daily deal websites, the consumers who are uncertain about the quality of future deals constantly face a trade-off between clicking on the deal to purchase it now and waiting for the next one. Consumers start off overly optimistic about deal quality, clicking on more deals in the beginning to learn about their quality distribution. As consumers' knowledge accumulates the incentive to learn declines, and they click on fewer deals. Because clicking is less about learning the quality of deals over time, the motivation for clicking is more likely to be to actually purchase the deal—hence the data pattern of increasing purchase probabilities. Through our simulations, we showed that consumers' learning and being forward-looking are two conditions that help us replicate the patterns observed in the data. Modeling these subscriber behaviors, we were able to evaluate the effect size of search costs. Companies can affect consumers' learning by increasing the uncertainty in deal quality and affecting prior beliefs about deal quality, to generate more revenue.

From a methodological perspective, we combined a dynamic two-stage search model with Dirichlet learning on the distribution of quality, about which consumers are uncertain. Compared with other parametric learning models, such nonparametric learning is flexible and appropriate in contexts where consumers are not learning about the true quality of a specific product or service. Given the importance of characterizing deal quality in our analysis, we leverage state-of-the-art deep learning techniques to construct our quality measure.

Because of the constraints of the data, we were only able to examine consumer behavior on the Groupon website in our empirical analysis, and with only a focus on restaurant deals. In the future, we can extend the above analysis to more deal categories while allowing for the possibility that subscribers may have unique learning trajectories for each deal category, in addition to analyzing the nature of deals on the site as a whole.

Acknowledgments

The authors thank the senior editor, the associate editor, and the reviewers for their feedback; Andrew Ching, Tulin Erdem, Ali Hortaçsu, Jun Kim, Xing Li, Tesary Lin, Olivia Natan, Stephan Seiler, S. Sriram, and Raluca Ursu; attendees of the 2016 and 2017 Marketing Science Conferences and attendees of the 2016 and 2017 Marketing Dynamics Conference; and seminar participants at Guanghua School of Management at Peking University, University of Florida, CUHK Business School and School of Management and Economics at CUHK (Shenzhen), for their helpful comments. This paper will be an essay in the second author's dissertation. The authors contributed equally to this research. The usual disclaimer applies.

Endnotes

[1] See http://www.ibisworld.com/industry/daily-deals-sites.html.

[2] See https://today.yougov.com/topics/lifestyle/articles-reports/2014/10/06/discount-massages-what-new-york-women-want-their-d.

[3] See https://www.statista.com/statistics/273245/cumulative-active-customers-of-groupon/.

[4] See https://biz.yahoo.com/e/160212/grpn10-k.html.

[5] The terms have been used in articles in major mass media such as the *New York Times*, CNN, and Forbes.

[6] See http://business.time.com/2012/11/16/is-the-daily-deal-model-dying-a-slow-death/.

[7] See https://www.theatlantic.com/health/archive/2016/10/the-unbearable-exhaustion-of-dating-apps/505184/.

[8] See https://www.cnbc.com/id/47507737.

[9] "Tipping" means that enough people have purchased the coupon that the merchant actually runs the promotion. This ensures that the merchant generates enough revenue from the entire promotion to justify running the deal.

[10] We recognize that the overall concept of fatigue therefore goes beyond what we consider in this paper and could be a critical issue for companies.

[11] As a robustness check (discussed later), we show using only the subset of new subscribers who joined in January that clicking behavior declines over time. This confirms a waning interest *within* consumers on the platform over time, rather than across consumers.

[12] As we describe later, we operationalize learning in terms of the distribution of deal quality on the website.

[13] Suppose that a newly subscribed consumer starts off with a very high prior belief about deal quality on Groupon (9 on a 9-point scale). Before the consumer learns about the true quality distribution of deals, when (s)he encounters a deal with quality rating 7, (s)he is not sure whether it is a high-quality deal. Therefore, the consumer may be reluctant to purchase the deal. However, after the uncertainty is resolved through learning, a deal of quality 7 is seen as high quality because the true average quality is 6. At this point the consumer is more likely to purchase the deal.

[14] MTurk is a crowdsourcing marketplace for simple tasks, such as surveys and text analysis. Various fields, such as marketing (Liu and Tang 2015, Liu et al. 2017), psychology (Buhrmester et al. 2011), and computer science (Feng et al. 2009) have leveraged this platform for online data collection. The seminal work by Buhrmester et al. (2011) demonstrated that MTurk can be used to obtain high-quality data inexpensively and rapidly. In the text-mining community, it is also standard practice to rely on MTurk workers to assign labels to text data (Feng et al. 2009).

[15] Refer to Online Appendix A for the details of this step.

[16] Note that we are including original price and discount as potential drivers of our quality measure. The extent to which they influence quality is captured via the process we describe below. In addition, we include price separately in the model to reflect the economic trade-offs facing consumers.

[17] Here, we assume that consumers and researchers have the same information that is observable based on the meta information of the deal web page.

[18] In Online Appendix A, we include a section of model comparisons to demonstrate the superior performance of the proposed deep learning model.

[19] The drop in the second period owes to registration on the Groupon website usually being accompanied by one's intention to view and purchase a deal. As such, the click and purchase rates in the first period are extremely high.

[20] We assume that the budget allocated to a particular category, for example restaurants, is stable during the three-month period.

[21] For example, if a consumer joined at the beginning of February, then her first month is February and her last month is March.

[22] The proposed model replicates the two main patterns in the data but not all the patterns reflected in our data.

[23] As a robustness check, we compared the estimation results with those obtained with parametric Bayesian learning. The proposed model provides a better fit to the data, owing to the nature of the empirical quality distribution, as shown in Figure 6. The results are available upon request

[24] The definition of flow utility follows Seiler (2013), the utility at either decision stage in a particular time period.

[25] As in Seiler (2013), the use of the max-operator in the above equation constitutes a slight abuse of notation. To be more accurate, the consumer makes a choice in the purchase stage that maximizes the present discounted value and not the flow utility. The maximization is therefore relative to the choice-specific value function, which we subsequently make clear.

[26] Specifically, the *total* flow utility gained in period t with respect to each decision is expressed as follows:

$$U_{it}(d_{it}^C = 0) = u_{cs,it}^0,$$
$$U_{it}(d_{it}^C = 1, d_{it}^P = 0) = u_{cs,it}^1 + u_{ps,it}^0,$$
$$U_{it}(d_{it}^C = 1, d_{it}^P = 1) = u_{cs,it}^1 + u_{ps,it}^1.$$

[27] Specifically, these conditional independence assumptions are as follows:

$$p(x_{cs,it+1}, \tilde{\varepsilon}_{cs,it+1} | x_{cs,it}, \tilde{\varepsilon}_{cs,it}, d_{it}^C = 0) = p(\tilde{\varepsilon}_{cs,it+1} | x_{cs,it+1}) \cdot p(x_{cs,it+1} | x_{cs,it}, d_{it}^C = 0),$$
$$p(x_{ps,it}, \tilde{\varepsilon}_{ps,it} | x_{cs,it}, \tilde{\varepsilon}_{cs,it}, d_{it}^C = 1) = p(\tilde{\varepsilon}_{ps,it} | x_{ps,it}) \cdot p(x_{ps,it} | x_{cs,it}, d_{it}^C = 1),$$
$$p(x_{cs,it+1}, \tilde{\varepsilon}_{cs,it+1} | x_{ps,it}, \tilde{\varepsilon}_{ps,it}, d_{it}^P = j) = p(\tilde{\varepsilon}_{cs,it+1} | x_{cs,it+1}) \cdot p(x_{cs,it+1} | x_{ps,it}, d_{it}^P = j), \quad j \in \{0,1\}.$$

[28] The detailed derivation of value functions can be found in Online Appendix B.

[29] We postpone the detailed introduction of state variables in Section 6.1.4.

[30] If the user is brand new, without any click history, this knowledge corresponds to the subscriber's prior beliefs, that is $f_{it}^{pr}(Q)$, where $t = 0$.

[31] A detailed explanation of the method can be found in the appendix provided by Crawford and Shum (2005).

[32] With the results of the reduced form analysis, we demonstrate that data on featured restaurants only and on the entire sample, with all featured deals, are comparable. The results are available upon request.

[33] A Pearson product-moment correlation coefficient is computed to assess the relationship between price and quality. There is a very weak negative correlation between the two variables for all categories of deals ($r = -0.079$, $df = 527300$, $p < .000$), which is even weaker for restaurant deals ($r = 0.034$, $df = 108700$, $p < .000$).

[34] In this case, the proposed model does not have the learning part.

[35] The results are available from the authors.

[36] See http://www.investopedia.com/articles/active-trading/080515/how-groupon-makes-money.asp.

[37] The numerical value of the estimated search cost parameter is 3.141 (Table 9).

[38] We preferred this approach to adding a fixed amount to each price given the variation in prices across deals.

[39] The magnitude of the search cost is related to the relevance of the Groupon deals to the consumers. If we include only consumers who have clicked on a restaurant deal at least once (indicating that the Groupon deals are more relevant to this subsample of consumers compared with all of the new subscribers we use for estimation), the

search cost is $98. If we further exclude people who have never made a purchase on Groupon from all categories, the search cost is reduced to approximately $56. If we only include people who have clicked on restaurant deals at least three times and have made at least one purchase of any category on Groupon, then the search cost falls to $26. These results are available upon request.

[40] We assume the learning journey starts with the mean of the prior distribution and ends with the mean of the empirical distribution of deals.

[41] This may sound counterintuitive. However, creating frictions to make something less simple may sometimes benefit the firm (see https://www.nytimes.com/2018/12/12/technology/tech-friction-frictionless.html).

[42] See https://www.groupon.com/occasion/top-deals.

[43] The model does not replicate all the data patterns beyond the central ones in the paper.

References

Aaker DA, Jacobson R (1994) The financial information content of perceived quality. *J. Marketing Res.* 31(2):191–201.

Berndt ER, Hall BH, Hall RE, Hausman JA (1974) Estimation and inference in nonlinear structural models. *Ann. Econom. Soc. Measurement* 3(4):653–665.

Buhrmester M, Kwang T, Gosling SD (2011) Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspect. Psych. Sci.* 6(1):3–5.

Chen Y, Yao S (2016) Sequential search with refinement: Model and application with click-stream data. *Management Sci.* 63(12): 4345–4365.

Ching AT, Lim H (2016) A structural model of correlated learning and late-mover advantages: The case of statins. Working Paper No. 2662286, Rotman School of Management, University of Toronto, Toronto, ON, Canada.

Chintagunta PK, Jiang R, Jin GZ (2009) Information, learning, and drug diffusion: The case of Cox-2 inhibitors. *QME* 7(4): 399–443.

Chintagunta PK, Goettler RL, Kim M (2012) New drug diffusion when forward-looking physicians learn from patient feedback and detailing. *J. Marketing Res.* 49(6):807–821.

Crawford GS, Shum M (2005) Uncertainty and learning in pharmaceutical demand. *Econometrica* 73(4):1137–1173.

De Los Santos B, Hortaçsu A, Wildenbeest MR (2012) Testing models of consumer search using data on web browsing and purchasing behavior. *Amer. Econom. Rev.* 102(6):2955–2980.

Dholakia UM (2010) How effective are Groupon promotions for businesses? Working paper, Rice University, Houston.

Dholakia UM (2011) What makes Groupon promotions profitable for businesses? Working paper, Rice University, Houston.

Dholakia UM (2012) How businesses fare with daily deals as they gain experience: A multi-time period study of daily deal performance. Working paper, Rice University, Houston.

Dholakia UM, Kimes SE (2011) Daily deal fatigue or unabated enthusiasm? A study of consumer perceptions of daily deal promotions. Working paper, Rice University, Houston.

Dubé JP, Hitsch GJ, Rossi PE (2010) State dependence and alternative explanations for consumer inertia. *RAND J. Econom.* 41(3): 417–445.

De Vany AS, Saving TR (1983) The economics of quality. *J. Political Econom.* 91(6):979–1000.

Erdem T, Imai S, Keane MP (2003) Brand and quantity choice dynamics under price uncertainty. *Quant. Marketing Econom.* 1(1):5–64.

Erdem T, Keane MP (1996) Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Sci.* 15(1):1–20.

Feng D, Besana S, Zajac R (2009) Acquiring high quality non-expert knowledge from on-demand workforce. *Proc. 2009 Workshop People's Web Meets NLP: Collaboratively Constructed Semantic Resources* (Association for Computational Linguistics, Stroudsburg, PA), 51–56.

Gechter M, Tsivanidis N (2017) Efficiency and equity of land policy in developing country cities: Evidence from the Mumbai mills redevelopment. Working paper, University of Pennsylvania, University Park.

Hartmann WR (2006) Intertemporal effects of consumption and their implications for demand elasticity estimates. *Quant. Marketing Econom.* 4(4):325–349.

Hartmann WR, Nair HS (2010) Retail competition and the dynamics of demand for tied goods. *Marketing Sci.* 29(2):366–386.

Hendel I, Nevo A (2006) Measuring the implications of sales and consumer inventory behavior. *Econometrica* 74(6):1637–1673.

Hong H, Shum M (2006) Using price distributions to estimate search costs. *RAND J. Econom.* 37(2):257–275.

Honka E (2014) Quantifying search and switching costs in the US auto insurance industry. *RAND J. Econom.* 45(4):847–884.

Honka E, Chintagunta P (2016) Simultaneous or sequential? Search strategies in the U.S. auto insurance industry. *Marketing Sci.* 36(1):21–42.

Hortaçsu A, Syverson C (2004) Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds. *Quart. J. Econom.* 119(2):403–456.

Hu MM, Winer RS (2016) The tipping point feature of social coupons: An empirical investigation. *Internat. J. Res. Marketing* 34(1): 120–136.

Keane MP, Wolpin KI (1994) The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. *Rev. Econom. Statist.* 76(4): 648–672.

Kim JB, Albuquerque P, Bronnenberg BJ (2010) Online demand under limited consumer search. *Marketing Sci.* 29(6):1001–1023.

Koulayev S (2010) Estimating demand in online search markets, with application to hotel bookings. Working Paper (09-16), Federal Reserve Bank of Boston, Boston.

Koulayev S (2013) Search with Dirichlet priors: Estimation and implications for consumer demand. *J. Bus. Econom. Statist.* 31(2): 226–239.

Liu Q, Tang Y (2015) Construction of heterogeneous conjoint choice designs: A new approach. *Marketing Sci.* 34(3):346–366.

Liu X, Lee D, Srinivasan K (2017) Large scale cross category analysis of consumer review content on sales conversion leveraging deep learning. NET Institute Working Paper No. 16-09, Stern School of Business, New York.

Luo X, Andrews M, Song Y, Aspara J (2014) Group-buying deal popularity. *J. Marketing* 78(2):20–33.

McCall JJ (1970) Economics of information and job search. *Quart. J. Econom.* 84(1):113–126.

Mojir N, Sudhir K (2016) Price search across stores and across time. Working paper, Yale University, New Haven, CT.

Orhun Y, Venkataraman S, Chintagunta PK (2015) Impact of competition on product decisions: Movie choices of exhibitors. *Marketing Sci.* 35(1):73–92.

Parasuraman A, Zeithaml VA, Berry LL (1985) A conceptual model of service quality and its implications for future research. *J. Marketing* 49(4):41–50.

Phillips LW, Chang DR, Buzzell RD (1983) Product quality, cost position and business performance: A test of some key hypotheses. *J. Marketing* 47(2):26–43.

Rao AR, Monroe KB (1989) The effect of price, brand name, and store name on buyers' perceptions of product quality: An integrative review. *J. Marketing Res.* 26(3):351–357.

Rao AR, Qu L, Ruekert RW (1999) Signaling unobservable product quality through a brand ally. *J. Marketing Res.* 36(2): 258–268.

Rothschild M (1974) Searching for the lowest price when the distribution of prices is unknown. *J. Political Econom.* 82(4):689–711.

Rust J (1987) Optimal replacement of GNC bus engines: An empirical model of Harold Zurcher. *Econometrica* 55(5):999–1033.

Rust J (1994) Structural estimation of Markov decision processes. *Handbook Econometrics* 4:3081–3143.

Santos BDL, Hortaçsu A, Wildenbeest MR (2017) Search with learning for differentiated products: Evidence from e-commerce. *J. Bus. Econom. Statist.* 35(4):626–641.

Seiler S (2013) The impact of search costs on consumer behavior: A dynamic approach. *Quant. Marketing Econom.* 11(2):155–203.

Song I, Chintagunta PK (2003) A micromodel of new product adoption with heterogeneous and forward-looking consumers: Application to the digital camera category. *Quant. Marketing Econom.* 1(4):371–407.

Sriram S, Chintagunta PK, Manchanda P (2015) Service quality variability and termination behavior. *Management Sci.* 61(11): 2739–2759.

Stigler GJ (1961) The economics of information. *J. Political Econom.* 69: 213–225.

Weitzman ML (1979) Optimal search for the best alternative. *Econometrica* 47(3):641–654.

Wu C, Liang YS, Chen X (2015) Is daily deal a good deal for merchants? An empirical analysis of economic value in the daily deal market. Working paper, University of British Columbia, Vancouver, BC, Canada.

Zeithaml VA (1988) Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *J. Marketing* 52(3): 2–22.

Zhang J (2010) The sound of silence: Observational learning in the us kidney market. *Marketing Sci.* 29(2):315–335.

Zhao Y, Yang S, Narayan V, Zhao Y (2013) Modeling consumer learning from online product reviews. *Marketing Sci.* 32(1):153–169.