Platform Entry Strategy in Two-Sided Markets: Evidence from the Online Daily Deals Industry*

Byung–Cheol Kim† Jeongsik Lee‡ Hyunwoo Park§

August 9, 2013

Abstract

We empirically study a new platform's entry strategy into two-sided markets in the online daily deals promotion industry where promotion sites such as Groupon and LivingSocial broker local merchants and local consumers. We find the incumbent Groupon shows a significant better deal performance due to its greater network size advantage in local consumers. Nonetheless, the late-mover LivingSocial's entry led to an intense rivalry with Groupon. How can a new platform successfully enter two-sided market with network effects working against the entrant? We find no evidence for a meaningful difference in deal terms. Instead, the entrant has developed a substantial number of deals from the incumbent-developed merchants. Several main findings include (i) the merchants with better track records are more likely to switch in the entrant's early industry tenure, (ii) the merchant switching decreases in the entrant's revenue and number of self-developed deals, and (iii) the switching deals shows a mean-preserving spread type of performance distribution of its self-developed deals'. We discuss our findings from the perspective of informational advantage to a second-mover.

JEL Codes: D40, L10, M20

Keywords: platform entry, two-sided markets, online daily deals, second-mover's advantage

*We gratefully acknowledge the financial support from the NET Institute (www.NETinst.org), 2012 summer grant. We thank Luis Cabral, Judith Chevalier, Jay Pil Choi, Pedro M Gardete, Yinghua He, Erik P. Johnson, Yaron Yehezkel, Yi Qian and all participants in the 11th Annual International Industrial Organization Conference in Boston, the 10th Atlanta Competitive Advantage Conference in Atlanta, TIGER-forum 2013, Conference on "The Economics of Intellectual Property, Software and the Internet" at Toulouse, France, 2013 NET Institute Conference on Network Economics at UC Berkeley, and the Fourth Annual Searle Center Conference on Internet Search and Innovation at Northwestern University for many useful discussion and comments. Early version of this paper was circulated under the title of "Dynamic Platform Competition in a Two-sided Market: Evidence from the Online Daily Deals Promotion Industry." All errors remain ours.

†Georgia Institute of Technology, School of Economics, 221 Bobby Dodd Way, Atlanta, GA 30332–0615, byung-cheol.kim@econ.gatech.edu.

‡Georgia Institute of Technology, Scheller College of Business, 800 West Peachtree St. NW, Atlanta, GA 30308-1149, jeongsik.lee@scheller.gatech.edu

§Georgia Institute of Technology, School of Industrial & Systems Engineering and Tennenbaum Institute, 755 Ferst Drive NW, Atlanta, GA 30332-0205, hwpark@gatech.edu

1 Introduction

Platforms are, in their classical meaning, where trains and passengers meet each other to fulfill travels. In many of today's businesses, especially those mediated by information technologies, platforms are where multiple, interested parties meet each other to fulfill transactions. Examples of platform-mediated businesses abound: Internet search engines, auctioneers, credit cards, and dating services, to name only a few. This multi-sided nature of platform-based businesses gives rise to quite distinct features from those in brick and mortar businesses. In particular, competition primarily takes place between platforms, each of which is supported by many individual players in each side of the platform. Consequently, "getting the two (or multiple) sides on board" is critical for them to function properly (Rochet and Tirole, 2006). Because of this cross-side linkage through the platform, the economic benefit of (and hence the willingness-to-pay for) the one side necessarily depends on the activity on the other side.[1] Particularly, when one platform has a size advantage in one side, in the presence of cross-side network effects, it is not uncommon to see the network effects oblige incumbent platform.

In this paper, we empirically study the question, "How can a new platform compete with an incumbent in two-sided markets where the network effect works against the entrant?" Our empirical setting is the online daily deals promotion market that broker local merchants and local consumers for highly discounted product/service through prepaid coupon sales. We compile comprehensive deal-level data including product and service descriptions, terms of the offer, merchant information, and deal outcomes from the market leader *Groupon* and its main competitor *LivingSocial*.[2] The availability of deal outcomes such as sold quantity is a significant advantage of our study because a key limitation of public data gathered from the Internet is lack of systemically organized information on sales quantities (Edelman, 2012).[3] In fact, most Internet-based businesses have little incentive to make their sales information publicly available. However, in the daily deals market, promotion sites publish the number of coupons sold, among others, for every deal. Doing so helps these sites

[1] Reflecting the growing importance of platforms in market competition, there have been significant theoretical developments in the study of multi-sided markets (Armstrong, 2006; Caillaud and Jullien, 2003; Hagiu, 2006; Hagiu and Jullien, 2011; Rochet and Tirole, 2003, 2006; Weyl, 2010). The main focus of these studies has been on the optimal pricing and payment structures that induce participation from both sides of the platform (Rochet and Tirole, 2006; Rysman, 2009). Despite subtantial theoretical progress, however, there has been a general paucity of empirical investigations.

[2] Though the market is crowded with many players offering discounted coupons in the U.S., LivingSocial is considered as the only serious rival of Groupon (Gupta, Weaver, and Rood, 2011) in terms of size and coverage of deal offers.

[3] To obtain sales quantities, researchers often do some work-arounds such as tracking sales rank or inferring sales from inventory changes.

to establish their credibility to potential consumers, thereby creating positive feedback from the consumer side, as well as to demonstrate the effectiveness of their promotion to potential merchants when developing deals. Furthermore, a majority of platforms (e.g., on-line auction sites, video games, search engines) compete in one aggregate market so that platform's entry in multi-markets at different timing cannot be observed. But, the daily deals industry involves *local* merchants to which *local* consumers pay physical visits to redeem the purchased coupons, which implies that the eventual match between a consumer and a merchant is necessarily constrained by the local-level competitive environment. Since both platforms, particularly LivingSocial as a second mover in this industry, have exhibited staged entries into regional markets, the variation in the entry timing across different geographic markets allows us to identify the entry strategy of the entrant and analyze its impact on the relative performance over time and across regions.

Taking advantage of these features in favor of our empirical study, we started to compare the two platforms' deal performance measured by the number of coupons sold in each deal after accounting for other likely determinants of sales. As theories on two-sided platforms generally predict, we find that the incumbent Groupon enjoys a considerable advantage in deal performance: on average, deals offered through Groupon sold 21.4% more coupons than the comparable deals offered through LivingSocial. We also find that Groupon's lead in the merchant-side performance appears largely attributable to its greater network size in consumer side.

Despite this size advantage that accrues to the incumbent platform, this industry has witnessed a fairly significant catch-up by LivingSocial, a late entrant. In a two-sided market, an entrant could use a "divide and conquer" strategy (Jullien, 2011) by attacking either the consumer side or the merchant side to ignite positive feedback from the other side. We thus first examine if LivingSocial offers to potential buyers more favorable terms than those of Groupon for a similar deal. We find little difference in all deal terms in discount rate, value and price in general. Given the lack of noticeable differences in deal terms, we turn to examine what the platform's strategic entry action might take place on the merchant side.

The entrant LivingSocial has developed a substantial number of deals from the merchants who already ran at least one promotion from the incumbent Groupon. More precisely, on average about 20% of new LivingSocial deals come from merchants that have previously promoted through Groupon. The proportion increases over time, reaching 25% toward the end of our data period. Our investigation reveals that the merchant switchings are related not only to the rank of their previoius deals, but to the entrant's industry tenure and local market performance and experience.

Specifically, we find that the merchants with better track records are more likely to switch in large cities at early entries. It is also interesting to discover that the merchant switching occurs more frequently when the entrant's local performance is worse from its self-developed deals. The entrant's local market experience measured by the number of self-developed deals at the local market is negatively related to the merchant switching.

How do the deals from switching merchants perform for the entrant? Our examination on the relative deal performance shows that the deals from switching merchants shows no difference in coupon sales on average compared to the deals developed internally. In fact, our quantile regressions indicate that the distribution of performance from switching merchants shows the pattern of mean preserving spread (MPS) of that of the entrant's own-developed deals. That is, the relative benefit of developing deals from the incumbent's previous merchants is to reduce a risk associated with its own deal development.

A distinct feature of this industry is that, for each deal, the identity of the deal-offering merchant and the deal outcomes are all available on the web. This information can thus provide a convenient, and extremely accurate, basis for building effective market intelligence including the potential market size, the geographic distribution of market demand, not to mention of the performance assessment of rival promotion sites. In particular, the promotion sites may use this information to solicit the merchants with prior deal experience; doing so would, all else equal, help the sites to lower the cost of developing the merchant side because these merchants have already revealed their preference for daily deals and are familiar with the business model. Our results are consistent with this insight. The merchant-related information is expected to take the higher value when an entrant does not have a long history of business experience and when a local market under new entry is loaded with more people of higher income and with numerous merchants, other things being equal. Also, the incentives to exploit the local merchant information for a new deal promotion decreases when its own deal development performs better and its local experience builds up. From the perspective of information-faciliated entry strategy, our study suggests that the information availability is likely to work in favor of the entrant who has to overcome the initial disadvantage in a relatively short period of time.

Our study complements and extends recent empirical studies on two-sided platforms. Rysman (2004) demonstrates the importance of network effects in the market for Yellow Pages directories. He finds that advertisers appreciate consumer usage and that consumers also value advertising, which suggests the existence of network effects. Zhu and Iansiti (2012) develop a formal model of entry in platform-based markets in which the relative strength of an entrant's indirect network effects is important to its success. They confirm this insight with data on Xbox's entry into the videogame industry. While our study also confirms the role of network size on platform performance, we focus on how *transaction-level* information affects the platform competition. Our paper is closely related to Seamans and Zhu (2012) who investigate the entry effect of Craigslist, a popular website for classified advertisement services, on the U.S. local newspapers. They find that, upon Craigslist's entry, local newspapers with their own ads manager cut ads rates more sharply than those without such manager. Our study complements theirs in that we, too, examine the effects of a new platform's entry on competitive market outcomes. A critical difference is that, while in Seamans and Zhu the competition occurs between a new type of platform (i.e., an online ads site) and a traditional type of platform (i.e., local newspapers), in our set-up two online platforms with an almost identical business model compete in a head-to-head fashion. In this regard, our study is also differentiated from Jin and Rysman (2012) who primarily focus on the relationship between platform pricing and competition in the U.S. sports card conventions.

More narrowly, our study joins a series of recent works on the online daily deals market. Despite the fast diffusion of the business model, research on this experimental market remains scanty and has mostly relied on case studies or survey methods. For instance, Edelman, Jaffe, and Kominers (2011) offer a theoretical model in which price discrimination and advertising can improve merchant profitability. Gupta, Weaver, and Rood (2011) conduct a comprehensive case study of Groupon and the daily deals market, and Norton et al. (2012) provide a similar case study of LivingSocial. Dholakia (2011) presents a survey of over 300 merchants across five major daily deal sites including Groupon and LivingSocial. We improve upon these studies by exploiting the full data of daily deals from two leading promotion sites to provide systematic analyses of the effects of multi-level market competition on the terms and performances of these deals. While Song et al. (2012) study consumer's shopping and redemption behaviors using individual-level panel data in one major daily deals site in South Korea, we focus on a new promotion site's entry into different local markets and the information-based deal development strategy. The rest of our paper is organized as follows. Section 2 offers a brief overview on the online daily deals promotion industry and market practices. Section 3 describes data and sample; we relegate detailed data merging process of the two different sites to Appendix. After we provide our empirical strategy in Section 4, we offer our results in Section 5-7. Section 8 closes this paper with discussion.

2 The Daily Deals Promotion Market

2.1 Two Market Leaders: Groupon and LivingSocial

Groupon is the current market leader with more than 11,000 employees as of June 2012. Only after three years of its first-ever deal, the 50%-discounted pizza coupon in Chicago, Groupon went public in November 2011, instantly raising $700 million that valued the company at close to $13 billion. The first mover's phenomenal success and the almost non-existent entry barriers prompted numerous followers who entered the market with similar services. By early 2011, there were more than 500 online sites that offered discounted daily deals in the U.S. (Norton et al., 2012). Among these, LivingSocial is considered as the only serious competitor of Groupon (Gupta et al., 2011) in terms of size and coverage of deal offers.[4]

LivingSocial sold its first discounted coupons in the U.S. in July 2009 for a restaurant in D.C.'s Chinatown. Since then, LivingSocial has rapidly expanded to reach 330 cities in North America (mostly in the U.S.) by August 2011. With a sequence of investments including Amazon's $175 million, LivingSocial was valued at more than $3 billion as of early April 2011, less than two years after its startup.

Figure 1 provides the geographic distributions of the two sites' business operation in the U.S. over time. The comparison between Panel (a) and Panel (c) illustrates Groupon's initial dominance, stemming from its first-mover status. The contrast between Panel (a) and Panel (b), and that between Panel (c) and Panel (d) show the rapid growth of both platforms that have penetrated numerous cities and states in less than two years. Lastly, Panel (b) and Panel (d) together illustrate the two platforms' head-to-head competition.

2.2 The Business Model

Despite some differences in origins, features and target consumers, promotion sites have several common cores in their business model. There are three parties involved in this market: consumer, promotion site, and merchant. A promotion site contacts a merchant (or, much less frequently, a merchant contacts a site) and presents the promotion proposal at specific terms including value, price, discount rate, expiration date, etc. Once both agree to the terms, the site informs their registered users ("subscribers") of the negotiated deal via email. Subscribers receive these offers

[4]There are some sizable players that specialize in certain types of deals, e.g., *Restaurant.com* (restaurants) and *TravelZoo.com* (travel). However, none of these rivals compete with Groupon in a way that LivingSocial does, particularly in terms of the geographic coverage and the variety of deal types.

on a daily basis, which they can also forward to their contacts through emails or social networking services. No consumer pays any usage or membership fees to the promotion site to receive deal information. There are even websites (called "aggregators") such as *Yipit.com* that specialize in aggregating and providing various daily deal promotions, free of charge. On receiving the deal offer, consumers make the purchase decision before the offer expires, typically within two days from the notification. Consumers who purchase the coupon are charged for the listed price in advance, unless the deal is nullified due to, for instance, insufficient demand. Some promotion sites such as Groupon set a minimum quantity called "tipping point," below which the deal is automatically revoked. Consumers then take the coupon to the offering merchant to redeem it for the product or service as specified in the promotion. Some coupons are reportedly left unclaimed even after the expiry, which is normally six months to one year from the issuance date (Gupta et al., 2011).

The promotion site earns a portion of the total revenue (= deal price × sold quantity), according to the pre-determined sharing rule. Dholakia (2011) reports that the merchant share ranges between 30% and 50%. The specific revenue sharing rule is unknown to the public and can vary across deals. However, a 50-50 rule seems to be the norm in this industry. Our own estimation of the split rule based on Groupon's IPO prospectus (SEC, 2011)[5] gives around 46% as the site's average share of the revenue. Figure 2 illustrates the flow of this business model.

2.3 Industry Characteristics

The daily deals promotion market has several distinctive features that provide an attractive context for our empirical investigation. First of all, this is a novel example of two-sided markets. Merchants value the number of subscribers, assuming a constant ratio of buyers to informed subscribers. Considering that consumers can feed the deal information to other potential buyers through various Internet channels, the value of adding one more registered user may increase more than linearly. Similarly, all else equal, consumers may prefer a site that offers a greater variety of deals.[6] Thus, promotion sites must develop many appealing deals in the merchant side to stay attractive to the consumer side. Groupon and LivingSocial as market leaders have stayed far ahead of other promotion sites in attracting merchants due to their size advantage in the consumer base, which in turn enabled them to appeal to a greater number of consumers.

Another notable feature is that market participants are exposed to uncertainty about their

[5] http://www.sec.gov/Archives/edgar/data/1490281/000104746911005613/a2203913zs-1.htm.

[6] Initially, Groupon maintained a single-deal-per-day rule in a given local market. Now it frequently offers multiple "daily" deals.

expected surplus from the transaction. Clearly, each party has an incentive for participation: the consumer gets a steep discount that is otherwise rarely obtained, the merchant advertises the business to a large, targeted audience without bearing any upfront cost, and the promotion site earns a decent cut of the revenue generated from the deal. However, these returns are in no way guaranteed. A merchant's eventual profit of running a promotion may well depend on various factors such as contracted sharing rule, costs of providing the promoted product or service, rate of consumer redemption, the number of coupons sold at the specified terms, composition among buyers between new and existing consumers, and "upside spending" at redemption.[7] Consequently, merchants running a promotion face uncertainty about ultimate profits. Dholakia (2011) reports that 26.6% of the surveyed merchants lost money in daily deal promotions and 17.9% just broke even.[8] Promotion sites are not immune to this uncertainty because the revenue of the sites are completely tied to the merchants' sales performance. Consumers also take risks by paying in advance for an advertised deal to grab the high discount rate. For instance, the consumer may find the restaurant too crowded with many other coupon users, which could ultimately ruin her dining experience. The consumer may also feel that the advertised value was set too high for the food and thus the "perceived" discount rate was in fact low. Since there are many of these factors that render the daily deal quite experimental, we suspect that the performance history such as popularity of the promotion site and the deal-offering merchant can play an important role in facilitating market transactions.

Fortunately, the sites in our dataset make public their deal performance such as the number of coupons sold, from which we can also compute the exact sales revenue from each deal. This fine-grained information is not available for most data that are scraped from the Internet (Edelman, 2012). In fact, most Internet-based businesses have little incentive to make their sales information publicly available. However, in the daily deals market, promotion sites publish the number of coupons sold, among others, for every deal. Doing so helps these sites to establish their credibility to potential consumers, thereby creating positive feedback from the consumer side, as well as to demonstrate the effectiveness of their promotion to potential merchants when developing deals.

We further notice that there is relatively little room for differentiation between promotion sites.

[7]For example, a customer redeeming a restaurant coupon may order drinks or desserts that are not covered by the coupon. Likewise, a dental patient redeeming a coupon for teeth whitening may opt for additional services if the dentist finds out some teeth problems and recommends a treatment.

[8]One may argue that the promotion aims at maximizing long-term profits rather than seeking profits in the current deal, as the advertising slogan goes: "sell at a discounted price and expect consumers to return for a full price." However, given the low predictability of various factors, the uncertainty about the deal performance appears to be a legitimate concern for the merchants.

The merchant exclusivity is rarely achieved as merchants can choose any site to offer their deals through. Establishing the consumer loyalty is also hard as consumers can multi-home and switch almost instantly at no cost. According to Dholakia (2011), 48% of the merchants indicated a willingness to repeat promotion but 73% of them considered a different daily deal site for their next trial. If this were true, the promotion sites might increasingly claim lower shares of revenue as the competition on the merchant side to attract deals gets intensified. The survey revealed another challenge for the promotion sites: the need to cultivate own loyal customers. In the survey, less than 20% of the daily deal buyers returned to the merchant later for a full-price purchase, indicating a fairly low customer loyalty formed through the daily deal promotion. Low differentiations between sites and generally weak loyalties from both sides of the market seem to be causing some industry observers to cast doubts on the market's overall prospect (e.g., Dholakia, 2011). In our data, we also observe frequent changes in the merchant-site matching. Moreover, because the two sites started business in different regions (i.e., Groupon in Chicago vs. LivingSocial in D.C.), the incumbent-entrant relationship also varies across geographic regions, which provide a variation that we can exploit to cross-check the estimated effect of an incumbent status.

3 Data

3.1 Data Sources

Our data came directly from the websites of Groupon and LivingSocial. We downloaded and parsed the entire population of deals that have been offered by these two promotion sites through the end of November 2011, beginning from the Groupon's first pizza deal in Chicago in November 2008. For each deal, we collect all available information such as product and service descriptions, terms of the offer, merchant information, and deal outcomes.

At the time of our data collection, both platforms provided a web-based interface in which a searcher could query with integer identifiers that these platforms had uniquely assigned to deals.[9] Using an automated web crawler written in Python, a computer programming language, we progressively scanned deals beginning from the integer one until no further query returned a result. Owing to this search method, we believe that our dataset is close to being exhaustive.

Each of the promoted deals contains detailed information about the deal such as the content (e.g., title, subtitle, tags (Groupon only), description of deal), terms of the offer (e.g., value,

[9]The last access date to these sites was February 24, 2012, after which LivingSocial seems to have blocked the web-based lookup of past deals using the numeric identifiers.

price, discount rate, start date (Groupon only), end date, expiration date, tipping point (Groupon only)), offering merchant (e.g., physical address, business name (Groupon only)) and others (e.g., sold quantity). Many of these information fields are common between the platforms, which enables the comparison of deal terms and performance across deals and between platforms. Figure A1 in online Appendix illustrates the sources of this information, annotated on typical deal promotions from these websites.

We augmented our dataset with several more variables that may reflect local socio-economic factors. From U.S. Census 2010, we used the county-level population, median household income and the proportion of population whose age ranges from 20 to 29. The inclusion of this particular age group and income was based on a survey by Accenture reporting that the primary consumers using daily deals are "young and affluent people."[10] In addition, to account for the possible impact of Internet accessibility on the demand for daily deals, we used the county-level high-speed Internet access data that the Federal Communications Commission publishes biannually since 2008. Each of these county-level variables was matched to individual deals according to the deal-offering merchant's physical address. Since the data were collected from two different sources, we went through a number of steps to ensure the comparability between the deals from each platform. Appendix A provides the details of the data integration process.

3.2 Sample

For the analysis, we constructed a sample from the raw data by the following procedure. We first excluded the deals that were offered outside the U.S. Since both platforms are operating globally, we chose only domestic deals to enhance comparability across deals. We also eliminated some obvious dummy deals that seemed to have been created for initial test purposes, such as those with negative prices or sold quantities.[11] Along with these dummies were dropped the deals that appeared to have obvious data errors such as zero quantity, zero or missing price and value, and discount rates outside the range of 0-100%. Together, these deals accounted for less than 2% of the observations.

We only included deals that were actually transacted. The Groupon data have a field that shows if a particular deal has been "tipped," which is to indicate if the deal coupons were sold more than the pre-set threshold (called the "tipping point"). Any deal that sells below this minimum purchase volume becomes void. About 7% of the Groupon deals were not tipped and hence were excluded.

[10] Wong, W. (2011, October 24) "Study finds daily deal websites limited to younger, more affluent users." *Chicago Tribune*.

[11] All of these dummy deals were clustered around the initial startup period of these platforms.

We further restricted the time window to the period from July 2009 to October 2011. July 2009 was the month that LivingSocial began to offer daily deals (Norton et al. 2012).[12] We ended the sample period at October 2011 to minimize potential truncation due to the timing of our data collection. Finally, we removed the deals that lacked information on merchant location. About 20% of the deals did not show the physical address of the merchant. Groupon had relatively more deals of this type (about 22% of its total deals) than LivingSocial (about 16%). These deals typically allow for redemption at any store and hence may confound the effect of local competition that we are interested in. Hence, we dropped these deals from the sample.[13]

The final sample used for our analysis had a total of 143,525 observations. Among these, 73% (104,764) were Groupon deals. Table 1 shows the breakdown of deals by platform and category. Overall, the *Arts & Entertainment* represented the largest category (22.4%). Other popular categories included *Beauty & Spas* (20.5%), *Restaurants* (20.0%), and *Health & Fitness* (15.4%). However, the two platforms seemed to differ somewhat in their focus of the deal portfolio: Groupon put a relatively greater emphasis on *Restaurants* while LivingSocial's portfolio was more heavily weighted towards *Arts & Entertainment*, *Beauty & Spas*, and *Health & Fitness.*

4 Empirical Strategy

4.1 Overview

As is seen in Figure 5, the entrant LivingSocial in more than 80% of US local markets has begun as a small marginal player, but quickly grew in size and, by the end of our data period, claimed over a third of the combined sales of the two platforms. Our main research question focuses on LivingSocial's entry strategy to overcome the initial disadvantage. The literature of two-sided markets combined with the literature of network effects predict that when one platform has a size advantage in one side, in the presence of cross-side network effect, the platform may gain a superior performance on the other side.[14] From this perspective, we start to examine if Groupon as incumbent in most local markets shows a performance advantage over the entrant LivingSocial and, if so, what drives such a difference; in particular, we focus on the role of the incumbent's relative size advantage in the group of consumers. The deal performance analysis is reported in Section 5.

[12] Despite this truncation, we included the entire period prior to July 2009 when constructing the variables that, by definition, traced back to the initial period such as entry timing and cumulative performance of platforms.

[13] Including these deals did not materially change the results reported in this paper.

[14] See Rysman (2009) for a well-written survey for these literature.

Given the two-sidedness of the industry, there are two ways that an entrant can quickly build up its business and thereby overcome the initial disadvantage. The entrant could either attack the consumer side by offering more favorable terms for a similar deal, thereby triggering positive feedback on the side of merchants who are attracted by a large consumer network. Alternatively, the entrant could try developing many attractive deals from the merchant side in order to attract consumers who actively search for such deals.

We first focus on the consumer side and analyze how LivingSocial compares with Groupon in deal terms such as discount rate, price and value, because these deal terms are most important for consumers' purchase decisions; all else equal, providing more attractive terms will help the entrant to win consumers. A meaningful deal term comparison is empirically feasible because we control for deal categories, local markets (i.e., divisions) and deal-offering time (i.e., year-month) as well as local socio-econ factors. According to a survey on daily deals (Edison Research, 2012)[15], 69% of LivingSocial subscribers also use Groupon and 37% of Groupon users also subscribe to LivingSocial. Given the prevalent multihoming (i.e., subscribing to both sites) on the consumer side as well as a highly competitive market situation for most merchants, we do not expect that both platforms offer considerably different deal terms for similar deal characteristics: we will examine this prediction in Section 6.

Then, we turn to the merchant side to examine any strategic actions by LivingSocial's entry into Groupon's territory since the search for attractive new merchants looks a critical part of the platforms' competitive strategies. Our data show that about 20.2% of merchants that have run a promotion with Groupon run another promotion with the entrant LivingSocial. We investigate the changes in the platform-merchant combination to examine under which situation the entrant develops more deals from the merchants of the incumbent. Specifically, we study the relationship of the likelihood of merchant switching to the entrant's overall industry tenure, to the entered divisions' characteristics, to the relative performance of their previous deals, and to the entrant's local market performance, etc. Finally, we also examine the relative deal performance of the deals from switching merchants compared to the deals developed internally. We report our results in Section 7.

[15] http://www.edisonresearch.com/wp-content/uploads/2012/04/Edison-Research-Daily-Deals-Report-2012.

4.2 Variables

For our research design, now we describe our variables for empirical specifications on deal performance, deal terms comparison and merchant switchings.

Deal Terms For each deal i, there are three deal terms of value, discount rate, and price of which the relationship is governed by $Price = Value \cdot (1 - Discount\ rate)$.

Deal Performance We used the sold quantity, q_i, to capture the sales performance of a deal i. One could alternatively use the gross revenue but, since revenue is equal to deal price times sold quantity, regressions of log-transformed revenue with a control of logged price produce almost identical coefficient estimates on the regressors but that of price. Hence, we keep the sold quantity for our primary performance measure.[16]

Measures for Incumbent-Entrant difference Note that deal i contains a number of different characteristics such as geographic division (d_i), deal category (c_i), deal date (t_i), mediating platform (w_i), and merchant location (l_i). We use *Groupon Deal* dummy as the indicator of deals offered by Groupon. Given that our sample covers the entire U.S. regions in which both platforms operate, this variable is expected to capture the incumbent advantage over the entrant in the performance of individual merchant deals. This is obviously a crude measure as it only dichotomizes between platforms and hence fails to capture possible effects due to, for instance, changes in the relative size of consumer network over time. Hence, we defined *Relative Platform Reputation* (RPR_i), more sophisticated measures of incumbent advantage in each division, as the difference between Groupon and LivingSocial in the total number of coupon sales cumulated up to $t_i - 1$ for the given divisional market. That is, and for each deal's division where t_i indicates the year-month of deal i. That is,

$$\text{Relative Plaform Reputation}_i(d) = \log(1 + \sum 1_{G(i)} \cdot S_j(d)) - \log(1 + \sum 1_{L(i)} \cdot S_j(d))$$

where $G(i) = \{j \mid t_j < t_i, d_j = d_i, w_j = G\}$, $L(i) = \{j \mid t_j < t_i, d_j = d_i, w_j = L\}$, and S_j is the sold quantity for deal $j \neq i$. This variable is thus a continuous and time-varying proxy of the size-based advantage of incumbent at each division market. In fact, to precisely measure the relative size of consumers, we would ideally like to count the number of "subscribers" who regularly receive deal offers from each platform at each point in time during the study period. However, such data are not available. Instead, we used the "realized" demand for the deals to proxy for the size of the consumers who are affiliated with the platform. To the extent that the consumer's propensity of

[16] We also tried total dollar amount instead of sold quantity. Results were very similar and available upon request.

purchasing a deal given the deal offer is comparable between two platforms, the actual number of coupon sales should be proportional to the size of the consumers affiliated with each platform.

Other Variables Our definition of the local market is a time-varying space that covers m months prior to the focal deal i, $0 < t_i - t_j \leq m$, within r miles of radius surrounding the focal merchant, $\|l_j - l_i\| \leq r$. Thus, the local market reflects the dynamic nature of competition among merchants and between platforms without being tied to a particular geographic boundary. We then defined *Local Density* (LD_i) as the total counts of daily deals offered by either of the platforms in the same deal category as that of deal i. This variable thus measures the "crowdedness" of the local market in the vicinity of deal i during the most recent period. That is, for each deal i, we generated

$$\text{Local Density}_i = \sum 1_{A(i)} \text{ s.t. } A(i) = \{j \mid 0 < t_i - t_j \leq m,\ \|l_j - l_i\| \leq r, \text{ and } \ c_j = c_i\},$$

where $1_{A(i)}$ is an indicator function that assigns one for any $j \in A(i)$ and zero otherwise. Figure 3 illustrates how our definition of local market matches with the contour of the map of deals in the case of New York City. We used $m = 3$ and $r = 3$ for our main specification.[17]

We also included a number of potential covariates of deal performance. *Prior Performance*, measured by the gross revenue from the immediate prior deal of the focal merchant, controls for cross-merchant differences in quality (e.g., popularity, convenience, attractiveness of deal terms). Note that, by construction, this variable is defined only for the merchants that offer more than one daily deals during the sample period. For the merchants that offer a deal for the first time, we assigned zero to this measure. To isolate the effect of this treatment, we separately included *First Deal*, a dummy indicating the deals offered for the first time by the same merchant. *Distance to Division Centroid* measures the geographic distance (in miles) between the deal-offering merchant and the division centroid which is computed as the geo-center of all the previous deal locations. This variable controls for differences in other demand-side characteristics associated with merchant locations. *Price* controls for the effect of this important deal term on the demand for deal i. *20's Ratio* represents the proportion of people whose age is between 20 and 29 in the corresponding county. *Median Household Income* measures the average purchasing power of people living in a county. Lastly, *High-Speed Internet Access*, coded as one of zero to six, captures Internet

[17] A more narrowly defined local market reduces an overlap with that of the geographic division, but contains fewer deals and tends to produce many zero values for *Local Density*. We experimented with alternative definitions of the local market by varying the time window (e.g., 1 month, 3 months, and all preceding months) and the radius (e.g., 3 miles, 5 miles and 10 miles). The results were robust to these variations.

accessibility in the region.[18] Table 2 presents the descriptive statistics of the variables in our data.

5 Determinants of Deal Performance

5.1 Estimation

We assume that the deal performance is a linear function of explanatory variables such as RPR, deal terms and local demographics. We interacted RPR with *Groupon Deal* dummy to identify potential incumbent advantages at each division market. We additionally controlled for temporal effects by including dummies for calendar year and month as well as for location-deal category-fixed effects by including division-category dummies. We used the ordinary least squares (OLS) for estimation. To reduce the problem of heteroskedasticity, we logged all skewed variables and used robust standard errors, clustered by division-category. The regression equation is given by

$$\log\ q_i(d,c,y,m) = \boldsymbol{X}_i\boldsymbol{\beta} + \delta_{d\times c} + \tau_y + \tau_m + u_i(d,c,y,m) \tag{1}$$

where $q_i(d,c,y,m)$ is the sold quantity of deal i in division d, category c, calendar year y and month m; $\mathbf{X}$ is the matrix of competition measures such as RPR and LD including demographic variables, first deal dummy, deal terms (price, discount rate), prior performance, and distance to division centroid; $\delta_{d\times c}$, τ_y, and τ_m are respective fixed-effect terms for division-category, year and month; and $u_i(d,c,y,m)$ is an error term. Note that, for this estimation, we separately control for calendar year and month because RPR_i is defined at the national level and hence is perfectly correlated with year-month fixed effects.

5.2 Results

Table 2 presents the results on the determinants of deal performance. Column 1 shows that, for a comparable deal, promotions through Groupon on average leads to 21.4% more coupon sales than those through LivingSocial, as indicated by the significantly positive coefficient on the *Groupon Deal* dummy. Thus it appears that the incumbent platform enjoys a considerable advantage over the entrant in terms of the national market-level demand. This performance premium is less likely to be due to selection (i.e., high quality merchants choose to place their deals with the incumbent) because we explicitly controlled for the prior merchant performance, which was a positive predictor

[18] The measure counts the number of residential fixed connections faster than at least 768 kilobits per second (kbps) downward and 200 kbps upward. The original data are coded, based on the percentage, for each county into one of six categories including zero: less than 20% is coded as 1, 20% to 40$ as 2, 40% to 60% as 3, 60% to 80% as 4, and above 80% as 5.

of the current deal performance. Note also that we included local demographic variables to control for socio-economic factors that might influence entry decision and deal performance.

Column 2 additionally included *Relative Platform Reputation*, our proxy for the relative size between the platforms in the divisonal market, along with its interaction with the *Groupon Deal* dummy. With the inclusion of this variable, the coefficient on *Groupon Deal* lost significance. We have just interpreted the positive coefficient on the *Groupon Deal* dummy as evidence of an overall incumbent advantage. In two-sided markets such as this one, the incumbent platform typically enjoys a greater positive feedback effect due to the larger network size in the consumer side. Reflecting this correlation, including a more fine-grained measure of the incumbent advantage in size completely absorbed the effect previously born by the incumbent dummy variable. Estimates on other variables remained nearly unchanged.

This incumbent advantage, however, was reduced by the entrant's penetration into regional markets: a higher competitive intensity at the division level boosted the performance of LivingSocial deals relative to that of Groupon deals. A greater local density also undermined the industry-wide incumbent advantage: all else equal, Groupon deals performed worse than LivingSocial deals in local areas where deals of the same category had been more densely offered.[19]

Coefficients on demographic variables were generally consistent with intuition: population, income, and the proportion of 20's in population were all positively related to sold quantity, though high-speed Internet accessibility was insignificant. This is understandable as Internet accessibility has little additional variation over the local household median income. First-time deals performed significantly better than subsequent deals on sold quantity. Consistent with the *law of demand*, a higher price led to lower sales. Merchants that were farther from the division centroid seemed to exhibit location disadvantage.

6 Competition on the Consumer Side

The result in Section 5 demonstrates the size advantage that accrues to the incumbent platform and hence a disadvantage for late entrants. Then, how did LivingSocial, a late entrant, attempt to overcome such disadvantage? One possible channel is for the entrant to attack the consumer side by offering more favorable terms for a similar deal, thereby triggering positive feedback on the side of merchants who are attracted by a large consumer network. This naturally motivates us to examine the actions on the consumer side by comparing the terms of deals offered by the two

[19]Since Groupon deals are, on average, more frequently offered in any given local market, this could reflect intra-platform competition in Groupon deals.

platforms.

6.1 Specification for Comparison of Deal Terms

Each deal is unique in its own way. Also, platforms exhibit different deal mixes in terms of geographic regions, product category and timing. Thus, to ensure a fair comparison, we controlled for division-category-year-month-fixed effects. As a result, we estimated the following equation:

$$deal\ term_i(d, c, y, m) = \alpha + \beta \cdot Groupon_i + \delta_{d\times c\times y\times m} + \varepsilon_i(d, c, y, m) \tag{2}$$

where $deal\ term_i(d, c, y, m)$ is one of three deal terms, offered for deal i in division d, category c, calendar year y, and month m; Groupon dummy takes on value one for a deal offered by Groupon and zero otherwise; $\delta_{d\times c\times y\times m}$ denotes the full set of division-category-year-month dummies; ε_i is an error term. Hence, the baseline group consists of the deals offered through LivingSocial in the same division, category, year and month as those through Groupon. Note that here we did not take logarithm of the deal terms in order to facilitate the interpretation of coefficients. To avoid possible bias due to extreme values, we excluded outliers (i.e., top 1% in each deal term) from estimation. The criteria for the outliers were either over \$1,000 in value, or over 90% in discount rate, or over \$250 in price.

In order to provide a more complete picture on the deal term comparison, we make the temporal distinction before LivingSocial's entry from the post-entry, and for the post-entry period we divide the divisions that LivingSocial entered early (within 11 months of their first-ever entry into this industry) and the other later-entered division. During the period when Groupon was a monopoly, Groupon offered 8,375 deal promotions in total. For post-entry period, there are similar number of deals, 55,708 and 58,435, respectively during the first 11 months and after that mark. Note that these early entry divisions are mostly large cities/metro areas with relatively higher income levels.

6.2 Results

Table 3 shows the comparison of deal terms. In the pre-entry era when Groupon was a monopoly, Groupon offered 55.77% discount on deals at a price of \$28.65 for a value of \$75.89 on average for its 8,375 promotions. For the post-entry, the coefficient on the constant term in the second column of Table 2-(a) indicate that, on average, LivingSocial offered the same 55.77% discount on deals with a value of \$95.28, selling them for a price of about \$35.05 to consumers.[20] When one compare those deal terms across different time windows (pre-entry, post-entry all, early-entered divisions and later-entered divisions), the entry of LivingSocial apprear to result in greater value and price, though the discount rate was not affected much.

However, more importantly, when we compare Groupon's deals and LivingSocial's in the same region, category and time, we find that both platforms generally offered quite similar deal terms. Specifically, for both value and price, there is not even difference of statistical difference between Groupon deals and LivingSocial's. The only difference in discount rate (0.16%p), albeit statistically significant at 5% level, did not appear economically meaningful. The result of no difference in deal terms strongly holds through the entire sample period, regardless of the timing of the entry.

To check the robustness of this result, we further controlled for other variables might also affect deal terms. These variables are the ones used in the previous section's performance analysis.[21]

$$deal\ term_i(d,c,y,m) = \alpha + \beta \cdot Groupon_i + \boldsymbol{X}_i \boldsymbol{\gamma} + \delta_{d \times c} + \tau_y + \tau_m + \varepsilon_i(d,c,y,m) \qquad (3)$$

The results, summarized in Table 4, generally confirm the basic findings. On average, Groupon offered marginally lower discounts by 0.37%p than LivingSocial, while deals of both platforms had similar value and price controlling for other factors. Looking at the effects of other controls on deal terms, platforms offered greater value and price where local deals were more densely offered. Deals offered for the first time appeared more lucrative as they commanded a significantly higher value and price than subsequent deals by the same merchant did. Merchants with a good track record generally commanded higher values and prices on their deals. A longer distance from the division centroid was associated with a slightly higher value and price. Among the demographics, both population and household income were positively associated with discount rates. Interestingly, prices were significantly lower in areas with greater population, suggesting that platforms may compete more fiercely in local markets with greater demand.

For another robustness check, we run seemingly unrelated regression (SUR) because the deal term comparision based on equation-by-equation standard ordinary least squares may have correlated errors across equations; if so, our estimates will not be efficient (Zellner, 1962). To lend some variations across equations, we included the variables that turn out significant at least at 10% level

[20] Deals of higher value often come with greater discounts but the relationship is not generally linear. Thus, the formula of $price = value \cdot (1 - discount\ rate)$ may not hold on the averages.

[21] To avoid collinearity, for this estimation we separately controlled for year and month effects, instead of year-month effects.

from the OLS regression reported in Columns 1-3 for the SUR estimation. As the SUR results in Columns 4-6 suggest, again we do not find differences of economic significance in deal terms to support any differentiation strategy in deal terms.

7 Entry into the Merchant Side

The previous section's result indicates that LivingSocial's increased penetration in this market is unlikely to have come through more aggressive deal terms on the consumer side. Thus, we turn to the merchant side to examine possible strategic actions that the entrant used to ignite positive feedback thereby overcoming the initial disadvantage in network size. In fact, the prevalent multi-homing by consumers, coupled with similar deal terms, is likely to make it critical for the promotion sites to continually search for new merchants and offer novel deals in order to remain competitive.

On the merchant side, the entrant appears to have one entrant-specific advantage over incumbent in that it can easily learn which merchant already had a business relationship with the incumbent. This is not uncommon in the Internet-based intermediating businesses where buyers are informed of sellers in open Internet websites. Moreover, for the online daily deals industry, each merchant's performance in the number of coupon sales is also available. In at least three ways, utilizing such information will help the entrant to build its business relatively quickly. First, it reduces the cost of searching for merchants who might be interested in offering daily deals because these merchants have already revealed a preference for this novel online marketing channel. Second, the learning that took place in prior deals also helps minimize the cost of executing the deals. Third, past performance in deals by these merchants provides a reasonable reference point and hence reduces uncertainty about the expected outcome of a deal.

7.1 Merchant Ranks and Switching over Time

To understand better the entrant's deal development strategy in the merchant side, we examine the changes in the platform-merchant combination. There are several new variables to be constructed for the investigation along this line. We define a dummy variable *Switching to LivingSocial* that takes value of one if a given merchant multi-homes (run at least one promotion from each of both sites) and there is at least one promotion with Groupon before another subsequent promotion with LivingSocial, and zero otherwise. We construct a new variable *Previous revenue percentile rank*. As its name speaks, for each switching deal i we track down a particular merchant's precedent deal j. Then, the percentile rank of the revenue of deal j is computed from the distribution of the revenues of all deals that were offered before t_i at the same division. *LS tenure* measures the time in months since LivingSocial started its business in the US market.

One main specification for merchant switching analysis is linear probability model (LPM) in OLS where the binary variable *Switching to LS* is regressed on *LS tenure*, *Previous revenu percentile rank*, and their interaction. Note that all predicted values from this specification are well inside between 0 and 1, ranging from 0.138 to 0.418 with the mean of 0.219.[22]

$$\begin{aligned} \textit{Switching to } LS_i(d,c) = \beta_0 + \beta_1 \cdot \textit{LS tenure} + \beta_2 \cdot \textit{Percentile rank} \\ + \beta_3 \cdot \textit{LS tenure} \times \textit{Percentile rank} + \delta_{d\times c} + \varepsilon_i(d,c) \end{aligned} \tag{4}$$

Before the above specification, let us check each variable in a progressive fashion. Model 1 shows that merchant switching declines over time; Model 2 shows that merchants with better track records are less likely to switch controlling for the time trend. More interesting, Model 3 specified above shows that the temporal decline in merchant switching is stronger for merchants with better track records, indicating an early focus on these merchants.

For further examination on the relationship of merchant switching to the entrant and the merchant's previous deal rank, we split the samples into the early-enterred division from the late-entered divisions as in section 6. As Models 4 and 5 show the contrast, the estimated coefficients are statistically significant and negative only for the early-entered divisions, while none of them is for the later-entered divisions. This patten carries over when we split the samples based on the division population. We define any division as a metro area when its population is over 0.5 million. Model 6 shows that the high-ranked merchants switch more in the entrant's earlier tenure in metro areas, while such feature disappears in non-metro areas as seen in Model 7. This result suggests that the entries to metro areas can be more important to the new platform to quickly establish a duopolistic competition with the incumbent in such a market of large demand.

7.2 Merchant Switching and Revenue/Number of Self-developed Deals

We just examined the time dimension associated with merchants switching and their rank in previous promotion. The other dimension of interest is under which situation the entrant may rely more on the merchants who already participated in this new type of online marketing channel. We

[22] We tried fixed-effect logit and probit as alternative specifications and find the results are robust and the marginal effect is almost the same as LPM. For limited space, we do not report all results in the paper as a table format, but they are all available upon request.

construct two new variables to address this inquiry.

First, we compute *Median revenue of self-developed deals* of the entrant LivingSocial at each division level. One presumption behind contructing this variable is that it measures the current performance level of the entrant at a given division. The lower the median revenue of self-developed deals, the more likely the entrant tries to make promotions from the incumbent's merchants and especially for the high-ranked merchants. Panel (a) in Table 7 reports the result of this examination. Models 1-2 show commonly that indeed the merchant switching is negatively related with this variable. Model 2 also confirms that the merchants doing better are less likely to switch to the new platform. However, we cannot find that the previous rank is associated with merchant switching through median revenue: the interaction term in Models 3-4 is not significant. This result can imply that when the entrant's division performance gets worse, the merchant switching occured without much selective scrunity.

Second, we also look into the *Number of self-developed deals*. While the Median revenue reflects the performance of self-developed deals, the Number may captures the deal development experience itself. Intuitively, as the entrant has developed more deals and know better the local market through this experience, it is expected that the importance of approaching to the incumbent merchants diminishes. Panel (b) in Table 7 confirms this intuition. Models 1-2 shows that the number of self-developed deals is negatively related to the merchant switching. Interestingly, Models 3-4 suggest that the lower number of deals is associated with more switching of higher-ranked merchants.

One important caveat for the above analysis is that we do not find the same qualitative results in the divisions where LivingSocial entered earlier than Groupon. That is, where LivingSocial played as an incumbent facing Groupon's entry, merchant switching does not depend on the proxies for local performance and experience.

7.3 Performance Comparison: Switching vs. Own-developed

How, then, do these deals contribute to the performance of the entrant? Recall that the switching merchants can help improve the entrant performance at least by lowering the cost of merchant development or by reducing the variance in deal outcomes. Although we do not observe the cost of merchant development, we have data that allow us to examine the outcome variation across deals. In particular, we can directly compare the outcomes of switching deals with those of deals developed internally. For a meaningful comparison between heterogeneous deals, we used the residuals recovered from the performance equation in Section 5. That is, we constructed the residuals in the following manner:

$$\widehat{u^k}_i = \log(q_i^k) - \widehat{\log(q_i^k)} \quad \text{for a given quadruple } \{d, c, y, m\} \tag{5}$$

where the superscript $k \in \{\text{switch}, \text{own}\}$. By this construction, $\widehat{u^k}$ for a deal can only differ in the way it was developed, i.e., switching from Groupon ("switch") versus developed internally ("own"). With this transformed measure of performance, we first examined the performance distribution between deal types.

Panel (a) in Figure 4 illustrates the kernel densities of $\widehat{u^k}$ for LivingSocial deals. The solid line represents 'switched' deals and the dotted line 'own' deals. Two observations stand out. First, the deals from switching merchants shows no difference in coupon sales on average compared to the deals developed internally. The OLS regression in Table 8 reports that the dummy for the merchants switched to LivingSocial is not statistically different from zero.[23] Second, as the quantile regressions in Table 8 indicates and Figure 4-(a) illustrates, the distribution of performance from switching merchants shows the pattern of mean preserving spread (MPS) of that of the entrant's own-developed deals. The switching deals have both thinner left-tail and right-tail than own deals and inter-quantile regressions also suggest that switching deals were also more predictable in performance, which unconditionally reduced the overall variance in outcomes of LivingSocial deals.

8 Discussion

8.1 Information-Based Entry Strategy

We have examined the entrant platform's entry strategy both in the consumer side through deal terms and in the merchant side through merchant switchings. Though we do not claim that the merchant switching is entirely driven by the entrant's favorable treatment to induce the incumbent's merchants to switch, our results are quite consistent with the presumption that the entrant make use of the information of the identity of the deal-offering merchant and the deal outcomes for new deal developments.

Suppose that merchants switched by their own preference, not by the entrant's solicitation. Assume that the preference is independent over time or in the ranking of their previous deals.

[23] A two-sample t-test of $\widehat{u^k}$ between switching deals and own deals cannot reject the null hypothesis of equal means between the two groups (t-stat $= -0.4684$).

Then, the temporal variation in merchant switching must be independent with merchants' track records. However, as we reported in section 7.1, the better merchants switch more in the entrant's early tenure. Similarly, the presence of such effects only in early timing and metro divisions is well supported by the entrant's information based entry strategy: the value of information of knowing who ran the daily deals promotions can be particularly high when the entrant needs to overcome the initial disadvantage in a relatively short period of time and/or in metro areas where there are many local merchants even in a given category so that a relatively small portion of merchants participated in the daily deals.

The information-based entry strategy is also well consistent with the findings for self-developed deals in Table 7. The incentives to exploit the local merchant information for a new deal promotion increases when its own deal development performs worse and its local experience is not that deep. Our finding also implies that the own deal developments have learning-by-doing to some extent and the less needs to reach out the rival's merchants. It could be the case that a platform could improve its efficiency in developing its own deals as it stays longer in the local market. Or, it is also possible that more consumers subscribe to the new platform's deal annoucement as the platform local activity continues.

8.2 Incumbent Response and Intensified Platform Competition

If merchant poaching was so frequently used by the entrant and doing so significantly helped it increase penetration and improve overall performance, such strategy would likely trigger an incumbent response. In our data, the proportion of switching merchants from the entrant LivingSocial to Groupon stayed around 7-8% during the first year of the entry but continued to increase up to 15% by two years after the entry. The mean proportion (8.4%), however, remained lower than that of LivingSocial, largely because Groupon offered many more deals during the period. Interestingly, merchant poaching turned out to be relatively more profitable for the incumbent. Panel (b) in Figure 4 shows the comparison of the distribution of $\widehat{u^k}$ for Groupon deals for the two groups, switch and own. The deals switching from LivingSocial were much less likely to generate a below-median sales performance among Groupon's new deals. Moreover, the distribution had a fat right-tail, suggesting that some of the poached deals delivered outstanding performance. On average, switching merchants' deals from LivingSocial to Groupon generated about 9% greater coupon sales for Groupon than the internally developed deals did, and this difference was statistically significant. As in the case of LivingSocial, merchant poaching helped Groupon as well to significantly reduce the variance in performance. Again, the availability of merchant and deal information allowed the incumbent to curtail the uncertainty of deal outcome by selectively soliciting merchants from the entrant platform.

The open information structure in this Internet-based industry allows the promotion sites to identify the merchants with prior deal experience and obtain detailed information on individual deals. This is particularly helpful for new entrants as it reduces the costs associated with searching, negotiating and contracting with merchants and the uncertainty surrounding deal performance. Clearly, this open information structure facilitates entries of new platforms into this industry. But, at the same time, the open information structure coupled with the already low switching cost for the merchants seems to make it harder for the platforms to build merchant loyalty. Thus, the information availability turns into a competition-intensifying channel over time, particularly as the incumbent fights back with an increased poaching of the new merchants that the entrant manages to develop at their own cost.

8.3 General Implication for Internet-based Intermediaries

The phenomenal advances in the information technology have dramatically reduced the cost of matching between different groups or agents. As a result, we observe a burgeoning of various web-based intermediary platforms that link between online and offline economic activities. We believe that our study will play a guiding role for future studies on these platform-mediated markets. Consider, for instance, the online vacation rental market. A number of websites (e.g., *Airbnb.com*, *VRBO.com*, and *Vacationrentals.com*) connect between homeowners who seek to rent out their places and travelers who look for short-term stays in lieu of hotels. This market exhibits most of the characteristics found in the daily deals market such as the heavy reliance on information technology, the critical involvement of local agents, low entry barriers, and considerable uncertainty associated with transactions. Intermediating platforms have strong incentives to publicize their transaction records to attract agents from both sides as the market is still deemed experimental and the transactional uncertainty that agents face is relatively high. Given the similarity in market conditions, we surmise that most of our findings on platform competition may easily apply to this emerging two-sided market. More generally, we hope that our study will encourage further fruitful research on the dynamic platform competition.

References

[1] Armstrong, Mark. "Competition in Two-Sided Markets." *RAND Journal of Economics* 37, no. 3 (2006): 668-91.

[2] Armstrong, Mark, and Julian Wright. "Two-Sided Markets, Competitive Bottlenecks and Exclusive Contracts.". *Economic Theory* 32, no. 2 (2007): 353-80.

[3] Bird, S., E. Klein, and E. Loper. *Natural Language Processing with Python*. O'Reilly Media, 2009.

[4] Caillaud, Bernard, and Bruno Jullien. "Chicken & Egg: Competition among Intermediation Service Providers." *RAND Journal of Economics* 34, no. 2 (2003): 309-28.

[5] Caminal, Ramon, and Carmen Matutes. "Endogenous Switching Costs in a Duopoly Model." *International Journal of Industrial Organization* 8, no. 3 (1990): 353-73.

[6] Chen, Yongmin. "Paying Customers to Switch." *Journal of Economics & Management Strategy* 6, no. 4 (1997): 877-97.

[7] Dholakia, Utpal M. "How Businesses Fare with Daily Deals: A Multi-Site Analysis of Groupon, Livingsocial, Opentable, Travelzoo, and Buywithme Promotions." *Working Paper* (2011).

[8] Edelman, Benjamin. "Using Internet Data for Economic Research." *Journal of Economic Perspectives* 26, no. 2 (2012): 189-206.

[9] Edelman, Benjamin G., Sonia Jaffe, and Scott Duke Kominers. "To Groupon or Not to Groupon: The Profitability of Deep Discounts." *Harvard Business School NOM Unit Working Paper* No. 11-063 (2011).

[10] Fudenberg, Drew, and Jean Tirole. "Customer Poaching and Brand Switching." *RAND Journal of Economics* 31, no. 4 (2000): 634-57.

[11] Gupta, Sunil, Ray Weaver, and Dharmishta Rood. "Groupon." *Harvard Business School Marketing Unit Case* No. 511-094 (2011).

[12] Hagiu, Andrei. "Pricing and Commitment by Two-Sided Platforms." *RAND Journal of Economics* 37, no. 3 (2006): 720-37.

[13] Hagiu, Andrei, and Bruno Jullien. "Why Do Intermediaries Divert Search?". *RAND Journal of Economics* 42, no. 2 (2011): 337-62.

[14] Jullien, Bruno. "Competition in Multi-Sided Markets: Divide and Conquer." *American Economic Journal: Microeconomics* 3, no. 4 (2011): 186-219.

[15] Kim, Byung-Cheol, and Jay Pil Choi. "Customer Information Sharing: Strategic Incentives and New Implications." *Journal of Economics & Management Strategy* 19, no. 2 (2010): 403-33.

[16] Norton, Michael I., Luc Wathieu, Betsy Page Sigman, and Marco Bertini. "What's the Deal with Livingsocial?". *Harvard Business School Marketing Unit Case* No. 512-065 (2012).

[17] Rochet, Jean-Charles, and Jean Tirole. "Platform Competition in Two-Sided Markets." *Journal of the European Economic Association* 1, no. 4 (2003): 990-1029.

[18] Rochet, Jean-Charles, and Jean Tirole. "Two-Sided Markets: A Progress Report." *RAND Journal of Economics* 37, no. 3 (2006): 645-67.

[19] Roson, Roberto. "Two-Sided Markets: A Tentative Survey." *Review of Network Economics* 4, no. 2 (2005): 3.

[20] Rysman, Marc. "Competition between Networks: A Study of the Market for Yellow Pages." *Review of Economic Studies* 71, no. 2 (2004): 483-512.

[21] Rysman, Marc. "The Economics of Two-Sided Markets." *Journal of Economic Perspectives* 23, no. 3 (2009): 125-43.

[22] Seamans, Robert, and Feng Zhu. "Technology Shocks in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers." *NET Institute Working Paper*, 2010.

[23] Song, M., E. Park, B. Yoo, and S. Jeon "Is The Daily Deal Social Shopping?: An Empirical Analysis of Purchase and Redemption Time of Daily-Deal Coupons", unpublished manuscript (2012), University of Rochester.

[24] Taylor, Curtis R. "Supplier Surfing: Competition and Consumer Behavior in Subscription Markets." *RAND Journal of Economics* 34, no. 2 (2003): 223-46.

[25] Taylor, Curtis R. "Consumer Privacy and the Market for Customer Information." *RAND Journal of Economics* 35, no. 4 (2004): 631-50.

[26] Weyl, E. Glen. "A Price Theory of Multi-Sided Platforms." *American Economic Review* 100, no. 4 (2010): 1642-72.

[27] Wilbur, Kenneth C. "A Two-Sided, Empirical Model of Television Advertising and Viewing Markets." *Marketing Science* 27, no. 3 (2008): 356-78.

[28] Zellner A. "An Ecient Method of Estimating Seemingly Unrelated Regression Equations and Tests of Aggregation Bias." *Journal of the American Statistical Association*, 57, (1962): 500-509.

[29] Zhang, Xiaoquan, and Feng Zhu. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia." *American Economic Review* 101, no. 4 (2011): 1601-15.

[30] Zhu, Feng, and Marco Iansiti. "Entry into Platform-Based Markets." *Strategic Management Journal* 33, no. 1 (2012): 88-106.

Tables

Table 1: Deals by platform and category

Category	Groupon	Share	LivingSocial	Share	Total (GP+LS)	Share
Arts and Entertainment	17,301	19.7%	9,864	25.6%	27,165	21.5%
Automotive	2,246	2.6%	748	1.9%	2,994	2.4%
Beauty & Spas	18,149	20.7%	8,748	22.7%	26,897	21.3%
Education	3,584	4.1%	1,709	4.4%	5,293	4.2%
Food & Drink	3,298	3.8%	1,966	5.1%	5,264	4.2%
Health & Fitness	12,836	14.6%	7,417	19.2%	20,253	16.0%
Home Services	187	0.2%	146	0.4%	333	0.3%
Nightlife	319	0.4%	601	1.6%	920	0.7%
Pets	258	0.3%	110	0.3%	368	0.3%
Professional Services	1,430	1.6%	702	1.8%	2,132	1.7%
Restaurants	19,650	22.4%	4,658	12.1%	24,308	19.3%
Shopping	7,000	8.0%	1,810	4.7%	8,810	7.0%
Travel	1,426	1.6%	111	0.3%	1,537	1.2%
Total	87,684	100.0%	38,590	100.0%	126,274	100.0%

Table 2: Descriptive statistics

	N	Mean	Std. Dev.	Min	Max
Discount rate	126,274	56.16	10.45	0.00	98.76
Value	126,274	122.65	331.42	1.00	30,000.00
Price	126,274	44.81	142.55	0.99	14,999.00
Sold quantity	126,274	357.70	549.89	1.00	25,000.00
(Dummy) Groupon deal	126,274	0.69	0.46	0	1
Divisional platform reputation	126,274	1.72	1.89	-8.70	12.35
Local density	126,274	1.26	1.06	0.00	4.89
(Dummy) First deal	126,274	0.70	0.46	0	1
Prior performance in revenue	31,540	11,860.82	19,017.37	2.00	716,782.00
Distance to division centroid	126,274	11.76	13.36	0.00	274.11
County population	126,274	1,358,169	1,867,241	1,175	9,818,605
Median household income	126,274	54,858	12,796	24,133	119,075
Ratio of 20's population	126,274	14.78	3.10	5.72	38.72
(Dummy) Switched to LS vs. repeating on GP	9,131	0.72	0.45	0	1
(Dummy) Switched to GP vs. repeating on LS	23,476	0.20	0.40	0	1
(Dummy) Switched to LS vs. self developed by LS	33,691	0.20	0.40	0	1
(Dummy) Switched to GP vs. self developed by GP	61,546	0.08	0.26	0	1

Table 3: Determinants of Deal Performance

	1 (Log) Sold quantity	2 (Log) Sold quantity
(Dummy) Groupon deal	0.214**	0.009
	(0.019)	(0.022)
Relative platform reputation		-0.109**
		(0.008)
Groupon dummy × Relative platform reputation		0.133**
		(0.008)
Local density	0.122**	0.116**
	(0.009)	(0.009)
Groupon dummy × Local density	-0.097**	-0.088**
	(0.010)	(0.010)
(Log) County population	0.047**	0.047**
	(0.010)	(0.010)
(Log) Median household income	0.423**	0.420**
	(0.055)	(0.055)
Ratio of household with highspeed Internet	-0.013	-0.013
	(0.013)	(0.013)
Ratio of 20's population	0.017**	0.017**
	(0.003)	(0.004)
(Dummy) First deal	2.655**	2.650**
	(0.086)	(0.087)
(Log) Price	-0.494**	-0.494**
	(0.009)	(0.009)
Discount rate	0.012**	0.012**
	(0.001)	(0.001)
(Log) Prior performance	0.309**	0.309**
	(0.010)	(0.010)
(Log) Distance to division centroid	-0.146**	-0.144**
	(0.009)	(0.009)
Constant	-2.369**	-2.092**
	(0.640)	(0.644)
N	125,990	125,990
F-stat	290.58	287.26
Adj. R^2	0.291	0.296

Note: Division-category, year and month fixed effects are included in all models. Robust standard errors, clustered by division-category, are in parentheses. ** denotes statistical significance at 1%.

Table 4: Comparison of deal terms between platforms

(a) Discount rate

	Pre-entry	Post-entry All	Post-entry Early div.	Post-entry Late div.
(Dummy) Groupon deal		0.158*	-0.064	0.380**
		(0.073)	(0.107)	(0.098)
Constant	55.772**	55.769**	56.012**	55.533**
	(0.000)	(0.051)	(0.075)	(0.067)
N	8,375	114,143	55,708	58,435
F-stat		4.69	0.36	14.99
Adj. R^2	0.323	0.264	0.264	0.266

(b) Value

	Pre-entry	Post-entry All	Post-entry Early div.	Post-entry Late div.
(Dummy) Groupon deal		0.377	-1.039	1.791
		(0.824)	(1.198)	(1.127)
Constant	75.888**	95.278**	102.041**	88.808**
	(0.000)	(0.572)	(0.842)	(0.773)
N	8,375	114,143	55,708	58,435
F-stat		0.21	0.75	2.52
Adj. R^2	0.234	0.196	0.209	0.181

(c) Price

	Pre-entry	Post-entry All	Post-entry Early div.	Post-entry Late div.
(Dummy) Groupon deal		0.029	-0.146	0.204
		(0.245)	(0.357)	(0.334)
Constant	28.654**	35.050**	37.422**	32.786**
	(0.000)	(0.170)	(0.251)	(0.229)
N	8,375	114,143	55,708	58,435
F-stat		0.01	0.17	0.37
Adj. R^2	0.293	0.241	0.244	0.237

Note: Division-category-year-month fixed effects are included in all models. Robust standard errors clustered by division-category are in parentheses. *, ** denotes statistical significance at 5%, and 1%, respectively.

Table 5: Comparison of deal terms between platforms

	OLS			SUR		
	1 Discount	2 Value	3 Price	4 Discount	5 Value	6 Price
(Dummy) Groupon deal	-0.366**	0.105	0.176	-0.152*	-1.451*	-0.533**
	(0.138)	(1.439)	(0.451)	(0.071)	(0.708)	(0.201)
Relative platform reputation	-0.100**	-0.180	0.124	-0.046†		
	(0.036)	(0.424)	(0.132)	(0.026)		
Groupon dummy $\times$ Relative platform reputation	0.167**	0.031	-0.198	0.107**		
	(0.041)	(0.457)	(0.147)	(0.027)		
Local density	-0.005	2.994**	1.003**		1.679**	0.715**
	(0.065)	(0.763)	(0.241)		(0.373)	(0.123)
Groupon dummy $\times$ Local density	0.096	-0.893	-0.210			
	(0.063)	(0.652)	(0.213)			
(Log) County population	0.376**	-0.494	-1.204**	0.427**		-1.137**
	(0.078)	(1.058)	(0.322)	(0.036)		(0.091)
(Log) Median household income	1.085**	0.191	-1.322	0.556**		
	(0.359)	(4.606)	(1.581)	(0.146)		
Ratio of household with highspeed Internet	-0.024	0.496	-0.063			
	(0.092)	(1.181)	(0.366)			
Ratio of 20's population	-0.043*	-0.578*	-0.148	-0.047**	-0.282**	
	(0.021)	(0.290)	(0.098)	(0.012)	(0.081)	
(Dummy) First deal	0.925	120.489**	38.322**		10.020**	3.107**
	(0.632)	(7.499)	(1.830)		(1.000)	(0.331)
(Log) Prior performance	0.174*	15.085**	4.786**	0.072**	2.582**	0.799**
	(0.074)	(0.918)	(0.224)	(0.007)	(0.129)	(0.040)
(Log) Distance to division centroid	-0.049	2.669*	1.149**		2.998**	1.171**
	(0.078)	(1.043)	(0.275)		(0.484)	(0.163)
Constant	38.840**	-25.159	26.907	49.540**	36.074	30.035
	(4.083)	(56.470)	(19.527)	(8.728)	(105.224)	(29.918)
N	113,943	113,943	113,943	113,943	113,943	113,943
F-stat	9.31	15.18	26.37			
χ^2				40,616.80	34,805.00	40,060.08
Adj. R^2	0.253	0.229	0.257	0.263	0.234	0.260

Note: We compute the SUR estimates using the same sample from OLS. Division-category, year and month fixed effects are included in all models. Robust standard errors, clustered by division-category, are in parentheses for OLS models. Standard errors are in parentheses for SUR models. †, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.

Table 6: Probability of switching to LS over entry timing and division population in reference to repeating with GP

	All sample			Entry timing		Division population	
	1	2	3	4 Early div.	5 Late div.	6 Metro	7 Non-metro
Time since LS started	-0.004**	-0.004**	0.005*	0.011**	-0.004	0.006*	0.003
	(0.001)	(0.001)	(0.002)	(0.003)	(0.004)	(0.002)	(0.006)
Previous revenue percentile rank		-0.067**	0.278**	0.441**	0.012	0.297**	0.151
		(0.014)	(0.073)	(0.084)	(0.126)	(0.077)	(0.203)
Interaction			-0.016**	-0.022**	-0.005	-0.016**	-0.011
			(0.003)	(0.004)	(0.006)	(0.004)	(0.009)
Constant	0.300**	0.340**	0.140**	-0.011	0.377**	0.126*	0.224
	(0.025)	(0.026)	(0.049)	(0.056)	(0.080)	(0.051)	(0.143)
N	14,798	14,798	14,798	6,996	7,802	11,332	3,466
F-stat	9.35	18.30	20.29	12.04	15.33	16.15	4.42
Adj. R^2	0.035	0.037	0.039	0.030	0.047	0.026	0.078

Note: Division-category fixed effects are included in all models. Robust standard errors, clustered by division-category, are in parentheses. *, ** denotes statistical significance at 5%, and 1%, respectively.

Table 7: Probability of switching to LS by experience(?) of LS

(a) Median revenue of self-developed deals

	1	2	3	4
(Log) Median revenue of self-developed deals	-0.072**	-0.074**	-0.093**	-0.089**
	(0.013)	(0.013)	(0.017)	(0.017)
Previous revenue percentile rank		-0.071**	-0.347†	-0.355†
		(0.014)	(0.183)	(0.183)
Median revenue × previous rank			0.033	0.034
			(0.021)	(0.021)
Time since LS started				-0.001
				(0.001)
Constant	0.825**	0.884**	1.044**	1.035**
	(0.107)	(0.107)	(0.148)	(0.147)
N	14,583	14,583	14,583	14,583
F-stat	31.93	30.60	20.92	15.62
Adj. R^2	0.037	0.039	0.039	0.039

(b) Number of self-developed deals

	1	2	3	4
(Log) Number of self-developed deals	-0.015*	-0.016**	0.004	0.007
	(0.006)	(0.006)	(0.008)	(0.014)
Previous revenue percentile rank		-0.068**	0.097*	0.096*
		(0.014)	(0.046)	(0.046)
Self-developed deal count × previous rank			-0.034**	-0.034**
			(0.009)	(0.009)
Time since LS started				-0.001
				(0.003)
Constant	0.293**	0.334**	0.240**	0.241**
	(0.028)	(0.029)	(0.038)	(0.038)
N	14,561	14,561	14,561	14,561
F-stat	6.43	16.33	14.81	11.34
Adj. R^2	0.034	0.036	0.037	0.037

Note: Division-category fixed effects are included in all models. Robust standard errors clustered by division-category are in parentheses. †, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.

Table 8: Performance comparison between switched and self-developed merchants

(a) LivingSocial ($N = 32,014$)

	OLS	Quantile regression					Inter-quantile regression	
		0.05	0.25	0.50	0.75	0.95	0.95-0.05	0.75-0.25
(Dummy) Switched to LS	0.009	0.341**	0.102**	-0.016	-0.125**	-0.175**	-0.516**	-0.227**
	(0.019)	(0.062)	(0.027)	(0.021)	(0.024)	(0.035)	(0.059)	(0.027)
Constant	-0.093**	-2.161**	-0.809**	-0.007	0.713**	1.706**	3.867**	1.522**
	(0.010)	(0.022)	(0.010)	(0.008)	(0.009)	(0.013)	(0.014)	(0.007)

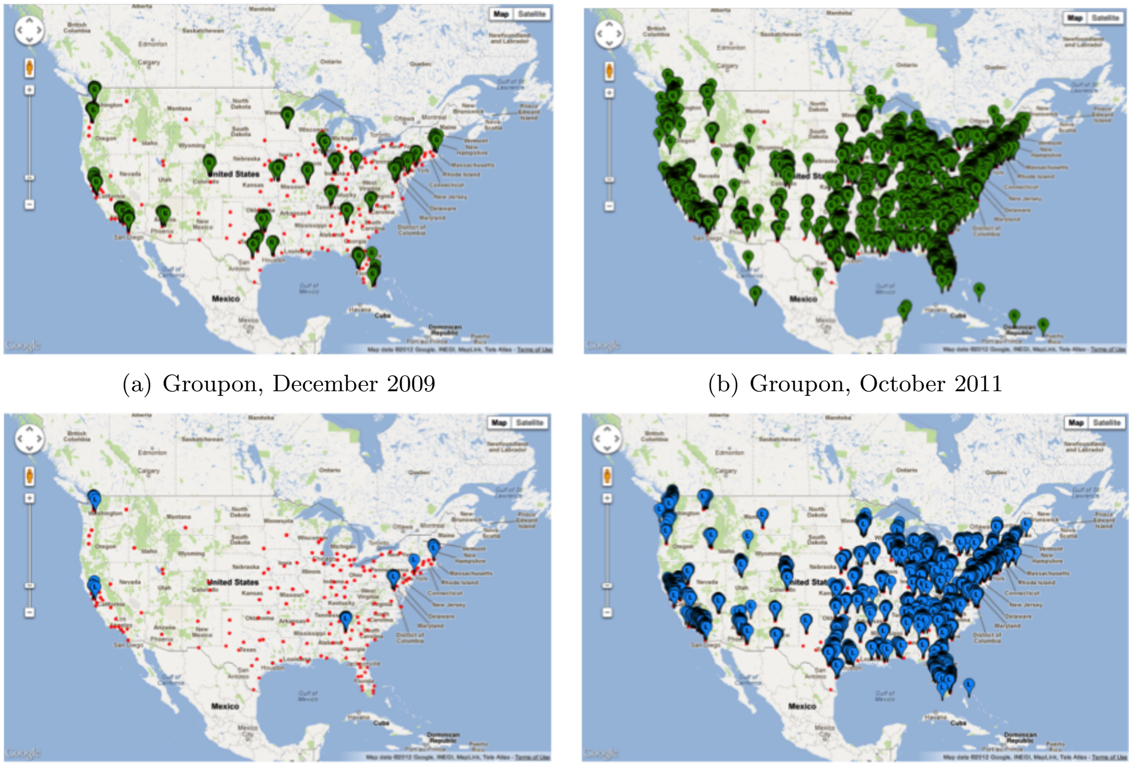
(b) Groupon ($N = 61,544$)

	OLS	Quantile regression					Inter-quantile regression	
		0.05	0.25	0.50	0.75	0.95	0.95-0.05	0.75-0.25
(Dummy) Switched to GP	0.089**	0.598**	0.225**	0.070*	-0.083**	-0.183**	-0.781**	-0.308**
	(0.020)	(0.074)	(0.035)	(0.028)	(0.024)	(0.027)	(0.051)	(0.033)
Constant	0.100**	-2.123**	-0.594**	0.266**	0.998**	1.686**	3.809**	1.592**
	(0.008)	(0.016)	(0.008)	(0.006)	(0.005)	(0.006)	(0.015)	(0.009)

Note: Baseline for all models is the deals developed by LS on its own. Division-category fixed effects are included in all models. Robust standard errors clustered by division-category are in parentheses. †, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.

Figures

Figure 1: Geographic distribution of Groupon and LivingSocial in the U.S.



(a) Groupon, December 2009

(b) Groupon, October 2011

(c) LivingSocial, December 2009

(d) LivingSocial, October 2011

Note: Each panel of (a)-(d) is a snapshot of deals distributed across the U.S. territory. Each green bubble represents a Groupon deal and each blue a LivingSocial deal. Each red dot indicates the center of a Groupon-designated geographic division (see Section 3 for the details).

Figure 2: Schematic illustration of the daily deals promotion market

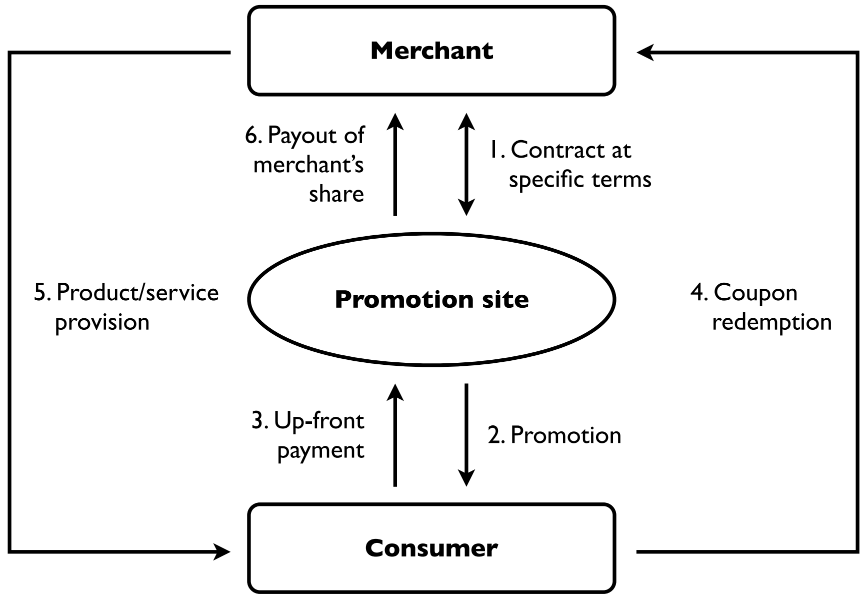
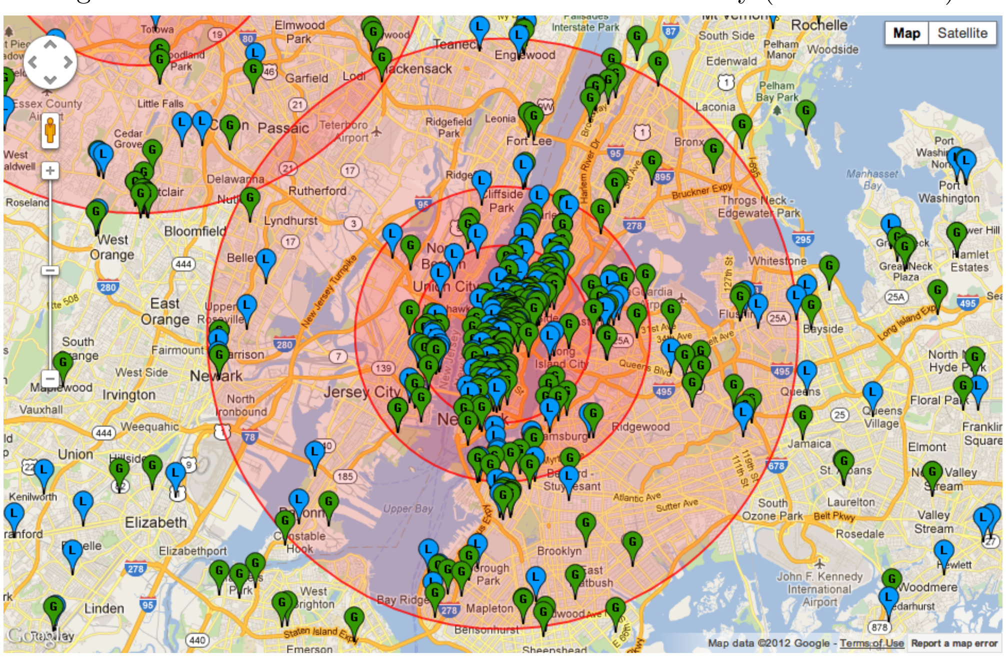
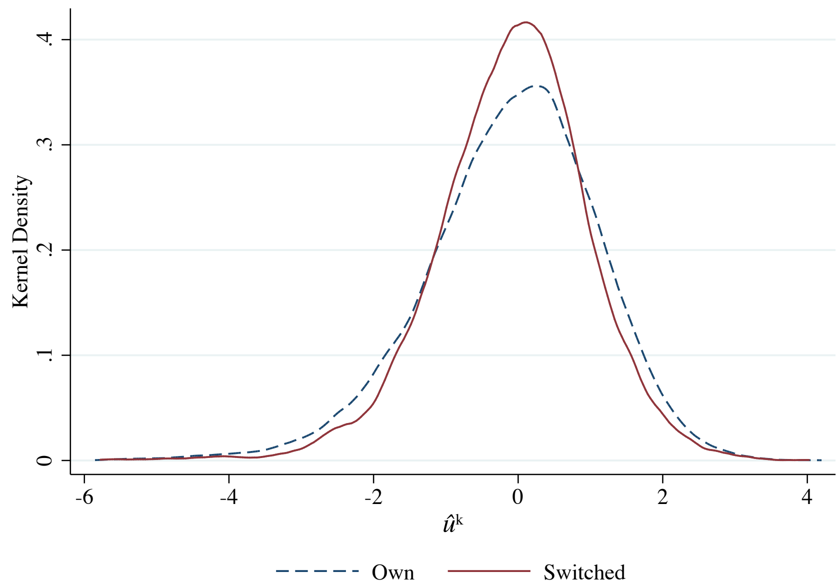


Figure 3: Boundaries of local market in New York City (October 2011)

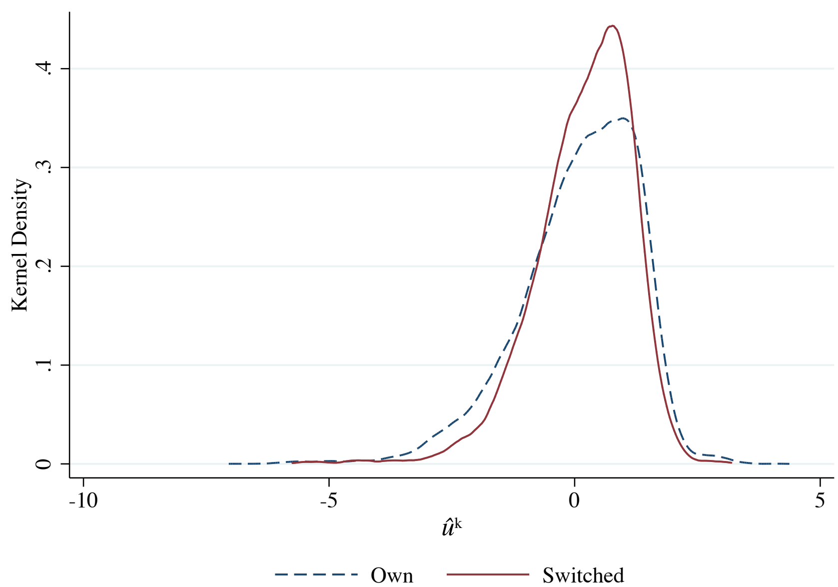


Note: The radii of the three concentric circles, from the smallest to the largest, are 3, 5, and 10 miles.

Figure 4: Performance comparison: own developed vs. switched



(a) LivingSocial



(b) Groupon

Appendix A: Data Integration

First of all, we unified the geographic unit of the deals within the U.S. territory. In the data, each platform utilizes their own way of classifying deals to specific geographic units, which are different from standard administrative units such as city or county. For instance, Groupon uses *divisions* and LivingSocial employs *areas*. These geographic units cluster the deals by broader location. More importantly, these units demarcate the boundary of business activities such as deal development and promotion.[24] Each division or area represents essentially a city-level metropolitan area (e.g., East Bay Area near San Francisco or Long Island in New York State). Though LivingSocial covers a smaller part of the U.S. than Groupon, the number of *areas* (220) is larger than that of *divisions* (155). Given our interest in the competition between platforms and the role of locality in the competition, we need to maintain consistency in matching the geographic unit of decision-making and define the measures at the corresponding level. Since Groupon is the dominant incumbent in this market and its divisions are closer to the boundaries of cities, we adopted Groupon's division system. Thus, we matched each of the LivingSocial deals with the closest division based on the merchant's street address.

We also developed a common deal categorization system. Each deal is, in principle, unique in the deal characteristics but is relatively homogeneous within the category it belongs to (e.g., restaurants, health and fitness, shopping, etc.). At a minimum, we need to control for the variations across deal categories. Groupon utilizes *tags* to characterize deals (e.g., steakhouses, personal trainers, pet stores, etc.). Each Groupon deal thus carries at least one and up to four tags that represent the nature of the deal. There are about 580 unique tags associated with Groupon deals. These tags can be aggregated into 18 broad tags ("categories," hereafter) defined by Groupon. We used these categories to classify the Groupon deals. However, there is no corresponding classification system for LivingSocial deals and hence we had to assign each of the LivingSocial deals into one of the Groupon categories based on the content of the deal. For this assignment, we utilized a machine learning algorithm that applied a Bayesian inference logic to match each deal with a category that best represented the deal content. We used as inputs for training the machine the entire list of vocabularies that appeared in *titles* and *subtitles* of the deals. For the implementation, we utilized the Natural Language ToolKit (Bird, Loper, and Klein, 2009). To ensure consistency in classification, we not only classified all LivingSocial deals but also re-classified each Groupon deal

[24] For instance, subscribers are only notified of deals offered in the geographic units they live in.

to a unique category using the same algorithm. Appendix C provides the details of this Bayesian category classification procedure. This process reduced the effective number of categories to 13, to which all deals in the dataset were uniquely assigned.

Further, we integrated the merchant information from both platforms to uniquely identify the merchants. Though both platforms provide the street addresses of the deal-offering merchants, they use different formats for the address field. More critically, LivingSocial does not state merchant names. These inconsistencies make it challenging to precisely match merchants between the two data sources. Thus, we used as merchant identifiers the geographic coordinates of latitude and longitude, measured to the sixth decimal point. Both platforms provided these coordinates for each deal in numerical forms. The difference of 10^{-6} degrees in latitude or longitude around the mid-point roughly translates into 3 feet. Thus, coordinates at this level of precision reasonably serve as valid identifiers.

To determine the timing of deal offer, we used the *end date*. One could alternatively use the *start date*, but such information was only available for Groupon deals. Moreover, the time gap between the start date and the end date was very short (the median was two days). In addition, we used the month as our time window for analysis. Thus, using the end date for determining the offer timing is unlikely to cause any material bias to our results.

A final note on the data concerns the sold quantity. We used this information, along with the sales revenue, to measure the outcome of a deal. The sold quantity in a Groupon deal was truncated at 1,000, 5,000, 25,000, 100,000 and 1,000,000 such that any deals falling between two marks were set for the lower bound (e.g., 1,001 units was recorded as 1,000). For about 8% of the deals, the sold quantity was reported truncated (about 80% of them were clustered at 1,000). In contrast, LivingSocial reported the actual number of units that the consumers purchased. We computed the sales revenue per deal by multiplying the price with the sold quantity. Hence, the actual sold quantity and the sales revenue for the quantity-truncated Groupon deals should always be greater than or equal to the corresponding figures in our data. This implies that any advantage of Groupon over LivingSocial we may find from our analysis on these outcome measures will be a conservative estimate. Our results hold even if we drop all deals that appear truncated.

Appendix B: Bayesian Category Classification Procedure

The procedure for the category classification, summarized in Figure 9, consisted of two steps. In the first step, we used the machine learning algorithm to train Bayesian classifiers for each of the 18 Groupon-designated broad tags (i.e., categories). Training the machine requires as inputs a set of vocabularies that are associated with each category so that, once trained, the classifier can compute using the Bayes' rule the probability that a deal containing a certain set of vocabularies belongs to a given category. Since Groupon assigns each deal to at least one category, we can use the entire deal description-category pairs to build a frequency-weighted vocabulary list for each of the 18 categories. To build these lists, we extracted the words from the titles and subtitles of the deals and tokenized them. For instance, a restaurant deal titled "Mexican food with soda" produces four tokenized title words (i.e., "Mexican," "food," "with," and "soda"). Then, these four words enter the list of vocabulary for the *Restaurants* category. We repeated this process for each and every Groupon deal, obtaining a full list of frequency-weighted vocabularies for each of the 18 categories. These profiles were then used to train the classifiers.

In the second step, we let the trained classifiers compute for each deal the probability that, given the tokenized title and subtitle, the deal belongs to a specific category. This probability is based on aggregated binary inferences (i.e., yes or no) on the tokenized title and subtitle words of the deal. It is thus possible that a deal gets classified into more than one category. In such case, we chose a category with the highest probability. For instance, the classifier recognizes a deal titled "Large pizza with bowling" by four tokenized title words (i.e., "large," "pizza," "with," and "bowling") and determines the probabilities that, given these words, this deal belongs to the *Restaurants* category (e.g., 0.991) or to the *Arts and Entertainment* category (e.g., 0.872). We then uniquely assigned this deal to the *Restaurants* category. We repeated this process for all Groupon and LivingSocial deals to ensure consistency in category classification between platforms. As five of the 18 categories received zero deal assignments, the effective number of categories became 13. For the implementation of this procedure, we used the Python Natural Language ToolKit (Bird et al., 2009).

Figure 5: Conceptual Diagram on How a Bayesian Classifier Assigns a Deal to a Category

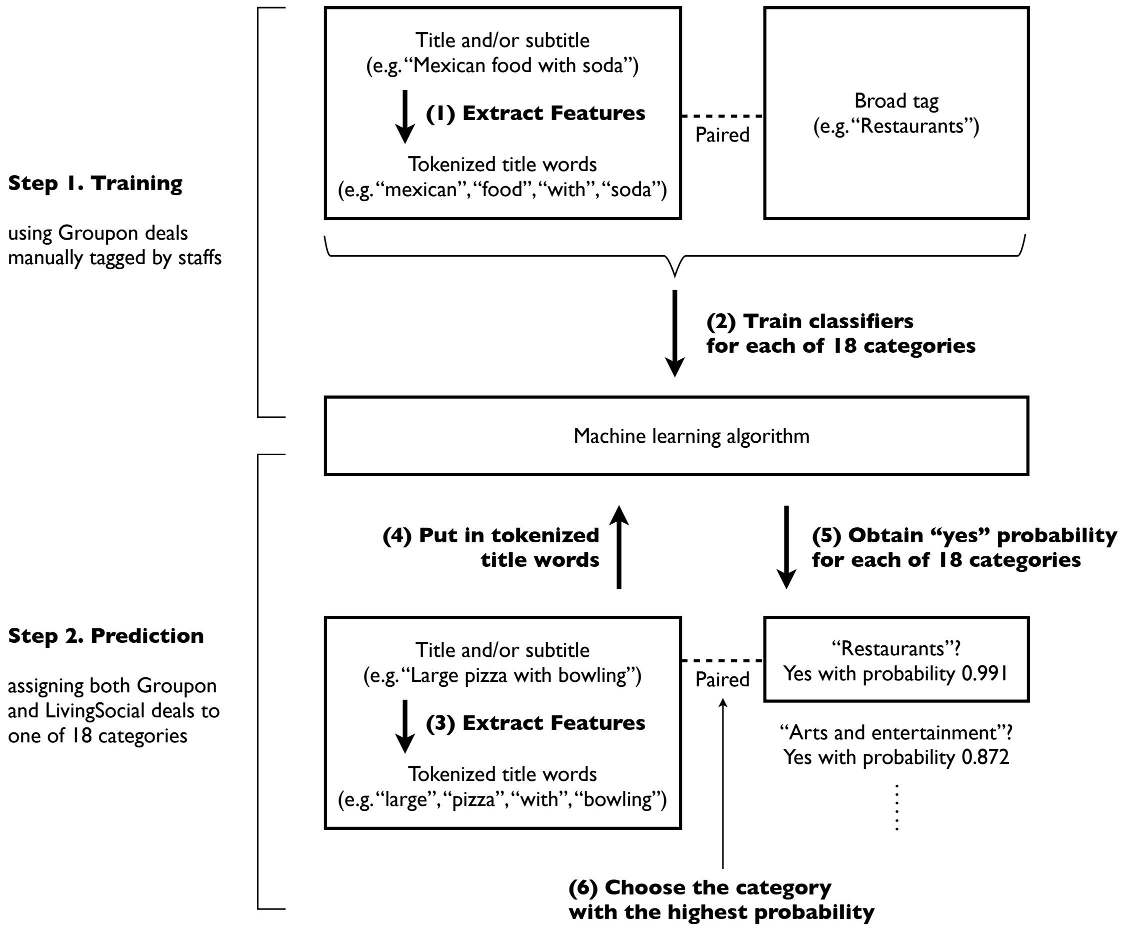
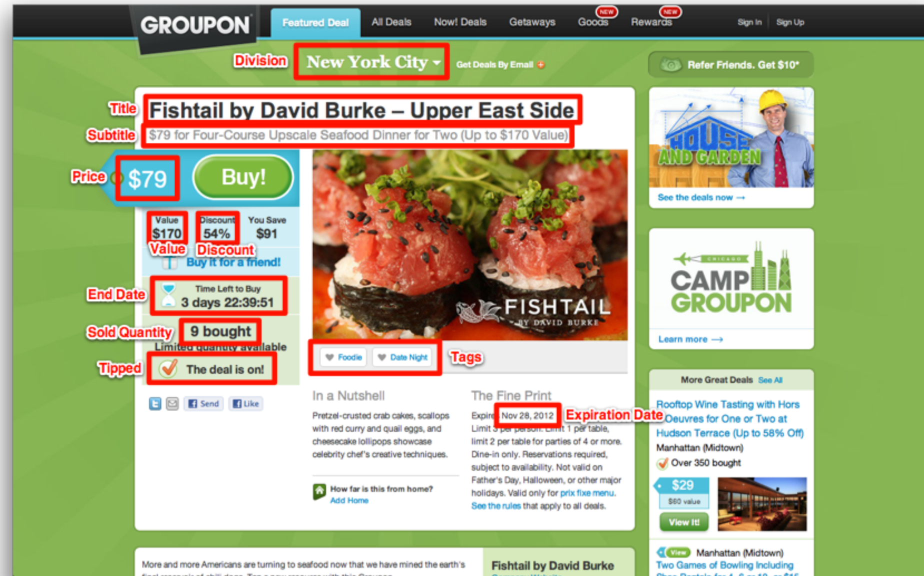
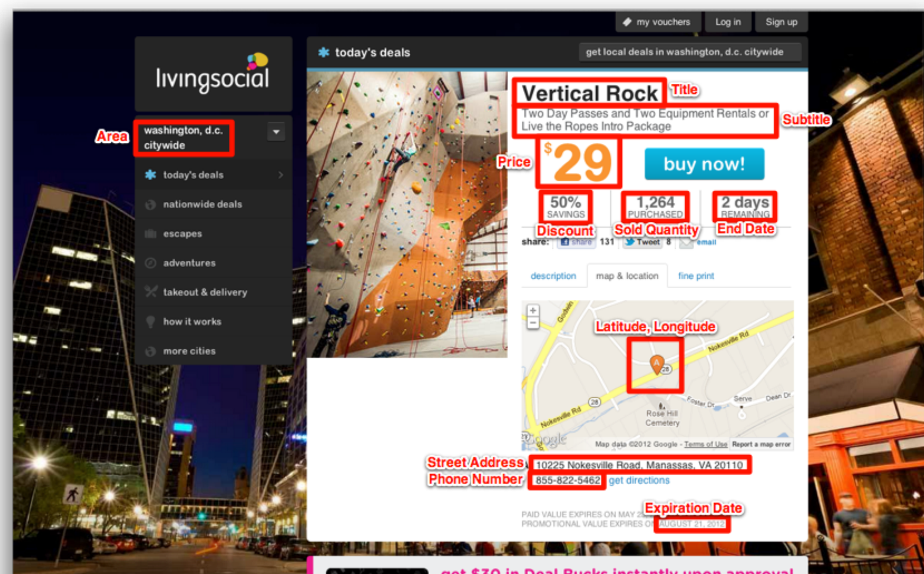


Figure 6: Examples of Deal Promotions at Groupon and LivingSocial



(a) Groupon



(b) LivingSocial