

# Local Market Characteristics and Online-to-Offline Commerce: An Empirical Analysis of Groupon

Hui Li, Qiaowei Shen, Yakov Bart\*

August 2016

## Abstract

We investigate the factors that affect the growth of Groupon, the leading online daily deals platform. We concentrate on the online-to-offline (O2O) aspect of the business that differentiates it from other e-commerce platforms - its strong connection to local markets. We focus on travel cost and store density, the key local characteristics that affect consumer deal demand and merchant deal offering. Using a comprehensive longitudinal data set on deal offerings and sales across local markets and combining it with local market characteristics, we estimate a simultaneous equation model of the weekly number of deal offerings and deal sales characterizing the two-sided nature of the platform. We find that the word-of-mouth effect on the consumer side and the observational learning effect on the merchant side contribute to and reinforce the expansion of a two-sided platform. However, a larger number of deals intensifies the competition, which then lowers per deal sales and limits the number of deal offerings. We find that local characteristics have significant impact on both the deal demand and supply side. We further use model simulation to show how differences in growth patterns across markets may be driven by local characteristics, and decompose their relative impact on the demand and supply side. The paper provides managerial implications for firms specializing in O2O commerce.

Keywords: Groupon; daily deals industry; O2O; two-sided platform

---

\*Hui Li is assistant professor of marketing at the Tepper School of Business, Carnegie Mellon University (hui1@andrew.cmu.edu). Qiaowei Shen is professor of marketing at Guanghua School of Management, Peking University. Yakov Bart is assistant professor of marketing at the D'Amore-McKim School of Business, Northeastern University. This research was supported by the Wharton-INSEAD Center for Global Research and Education.

# 1 Introduction

Since the launch of Groupon in 2008, the U.S. daily deals industry has experienced 332% annual average growth over five years, reaching \$3 billion in sales by 2013 and expecting to climb to \$5.5 billion in 2016.<sup>1</sup> Yet how to further grow and sustain the business is the central question for many platforms. By analyzing the deal offerings and sales of Groupon across all the major markets in the U.S. from March 2010 to January 2012, we find that while the platform has grown substantially over time, the size of the platform in terms of the weekly number of deals offered and total sales does not necessarily grow with the size of the local population. In this paper, we study the factors that determine the growth and the size of daily deals platforms. In particular, we investigate how the “localness” of this business model affects its expansion.

Daily deals platforms represent a typical example of the emerging “online-to-offline” (O2O) business model that directs the customers acquired online to offline stores. In contrast to such e-commerce platforms as Amazon and eBay that focus on shifting consumer transactions from offline to online environment, daily deal platforms are closely tied to local offline markets. They serve as digital intermediaries that bring together local merchants and local consumers through coupons. While consumers purchase deals from the platform online, the coupons are redeemed in offline transactions with the local merchants who offer the deals.

Each daily deals platform is effectively a two-sided market with consumers or deal-hunters on one side and merchants that offer deals on the other side.<sup>2</sup> The size of the platform, measured by transaction volume, depends both on the number of deals available and sales of each deal. It is the result of the mutual choices made by consumers and merchants. On the merchant side, the decision to list deals on a deal platform depends on the expected return, which includes both the revenue from the deal sales and the

---

<sup>1</sup>Sources: <http://www.ibisworld.com/industry/daily-deals-sites.html> and [http://www.biakelsey.com/Company/Press-Releases/120917-U.S.-Consumer-Spending-on-Online-Deals-to-Reach-\\$3.6B-in-2012.asp](http://www.biakelsey.com/Company/Press-Releases/120917-U.S.-Consumer-Spending-on-Online-Deals-to-Reach-$3.6B-in-2012.asp).

<sup>2</sup>For two-sided market theories, see Rochet and Tirole (2003), Caillaud and Jullien (2003), Armstrong and Wright (2005), Armstrong (2006), and Hagiu (2009). Empirical studies of two-sided market include Rysman (2004), Kaiser and Wright (2006), Dubé, Hitsch and Chintagunta (2010), Song (2011), Zhu and Iansiti (2012), Jin and Rysman (2012).

advertising exposure to the potential customers using the online platform. In other words, deal offerings would be affected by the size of the other side of the platform. Merchants' decision to participate may also be affected by the perceived platform attractiveness and past experience with the platform.

On the consumer side, the decision to purchase a Groupon deal depends on several factors. One is the good-will toward the platform, which accumulates over time with the past transaction volume. More consumers buying deals in the past is likely to correspond to more chatter (word-of-mouth effect). In addition to deal characteristics such as price and discount level, deal sales also depend on the number of active deals on the platform. While more variety may often attract more consumers (e.g., Stephen and Toubia 2010), it can also create competition between deals in the same category, which then may decrease sales of a deal.

In addition to these factors typically considered in a two-sided platform, local characteristics are also likely to affect the supply and demand of deals given the nature of "online to offline" business. We focus on two key local market characteristics in addition to the typical demographic variables considered: travel cost and store density. Both characteristics are market-specific and neither controlled by the platform nor controlled by the merchants. On the supply side, merchants may have a stronger incentive to offer online deals to motivate consumer patronage when travel cost is high. The effect of store density on deal offering can be mixed. High store density implies intense local competition, which can encourage the use of online deals, but low margins can also restrict merchants' ability to offer deep discounts. It is therefore an empirical question as to which effect dominates.

Travel cost and store density also affect deal demand as they are related to the cost of deal redemption. Consumers obtain the value of a Groupon voucher only upon redeeming it at the local merchant offering the deal. We argue that these local characteristics may not only affect the overall demand but also interact with deal choices. For example, ease of travel (low travel cost) would make deals more accessible. This may lead to higher demand for deals in general and intensify the competition among deal offers. In contrast, high travel cost may discourage consumers from purchasing deals that require redemption

in remote locations, which makes deals offered in different locations less substitutable. Similarly, high store density may have a positive effect on deal purchase because it is related to ease of redemption. However, consumers are also more likely to have access to local offline deal options in a market with high store density, and therefore would be less prone to buy online deals.

We use a comprehensive data set on Groupon deals and local market characteristics to empirically estimate these effects. The data covers the period from March 2010 to January 2012 and includes all the major U.S. markets. We estimate a simultaneous equation model of weekly number of deal offerings and deal sales. We also incorporate a platform attractiveness that evolves over time and affects merchants and consumers choices. We use Kalman filter method together with the control function approach (Petrin and Train, 2010) to account for unobserved platform attractiveness and address the endogeneity of deal offerings and deal sales.

The estimation results suggest that local characteristics affect how consumers choose among deals as well as the supply of online deals. We find a significant deal substitution effect which counters the expansion of deal platform. This effect is moderated by the travel cost: higher travel cost is associated with lower deal substitution effect. In addition, we find that higher store density decreases consumers' sensitivity to online discounts and therefore may negatively affect deal demand. We also find that merchants are less likely to offer deals in markets with high store density. Our empirical findings demonstrate why population or market size in itself is not sufficient to explain the growth pattern and the size of an O2O platform in a local market. Local characteristics such as travel cost and store density also contribute to the growth of the platform.

We further examine the individual effect of different local characteristics including travel cost, store density and population size on platform growth (measured by transaction volume over time) through simulation. Overall, comparative statics show that population and travel cost contribute positively to platform growth while store density contributes negatively in this context. These results can provide some explanation to the different growth pattern observed across markets. We further decompose the local effect on the

supply side and demand side of daily deals. The simulation results indicate that travel cost plays a more important role on the consumer side, while store density have a stronger impact on the deal supply side.

Our research provides important insights for daily deals platforms. Early survey-based studies examined the profitability of running deals for merchants and deal users' redemption behavior and perceptions (Dholakia 2010, 2011a, 2011b). A stream of analytical literature studied group buying as a selling mechanism (Jing and Xie 2011; Hu, Shi and Wu 2013), as well as advertising (Edelman, Jaffe and Kominers 2016). On the empirical side, there have been studies on the industry-specific properties of daily deals markets such as the relationships between Groupon offers and Yelp reviews (Byers, Mitzenmacher and Zervas 2012), platforms' ability to poach merchants (Kim, Lee and Park 2013), and opportunities for social influence and observational learning among consumers (Luo et al. 2014). Wu, Shi, and Hu (2014) used panel data of deal sales to study the thresholds as a mechanism to simulate deal sales. Li and Wu (2013) provided evidence that observational learning and social media word-of-mouth are complements in driving deal sales. We explore both the consumer-side and merchant-side factors that affect deal sales and deal offerings. In particular, we emphasize the role of the offline local markets in the Groupon business and examine how the local characteristics moderate these effects in shaping the platform growth and size.<sup>3</sup>

Our research also sheds light on platforms operating in O2O commerce which integrates online consumer acquisition and offline businesses.<sup>4</sup> Technological advancement and increasing mobile usage have enabled a rapid growth of such platforms in various areas, leading some industry commentators to call it a "trillion dollar opportunity".<sup>5</sup> Previous research has focused on the differences and the competition between online and offline retail formats (e.g. Brynjolfsson and Smith 2000; Brynjolfsson, Hu and Rahman 2009; Forman, Ghose and Goldfarb 2009; Anderson et al. 2010). Yet the research is

---

<sup>3</sup>Our paper also relates to the studies of two-sided platform growth (Zhang et al. 2012, Ahn, Duan and Mela 2014).

<sup>4</sup>Examples include Airbnb for accommodations, OpenTable for restaurants, Shopkick for mobile marketing, Uber for transportation, Zocdoc for doctors.

<sup>5</sup><http://techcrunch.com/2010/08/07/why-online2offline-commerce-is-a-trillion-dollar-opportunity>.

scarce on O2O commerce where online and offline markets are not fighting with each other for survival, but work together to deliver consumer value. We consider both online two-sided platform characteristics and offline local conditions in our model and provide empirical evidence that local market characteristics are important in determining the expansion of such O2O platforms. We argue that simply adding traditional demographic variables (e.g. population size) as controls to the model is not sufficient, as some local market factors (e.g., travel cost and store density) have impact on the key underlying effects related to consumers' and merchants' choice behavior and online platform growth.

The rest of the paper is organized as follows. We first describe industry background and present the data. We then estimate a simultaneous model of deal sales and deal offerings and discuss how local characteristics affect the demand and supply of deals. We further examine the individual effect of local characteristics on platform growth through simulation. Finally, we conclude with a discussion of the implications of our results for practice, as well as directions for future research.

## 2 Industry Background and Data

### 2.1 Deal Data Description

Daily deals platforms such as Groupon and LivingSocial list deals offered by merchants by geographical areas. Platforms develop and maintain connections with local merchants. Those merchants who decide to list a deal on a platform sign a contract that specifies the product or service offered, price, duration, and discount rate. Consumers can choose to buy any available deal voucher and redeem it later from the merchant offering it. Platforms usually get about half of the total deal revenue, and the merchants get the other half (Dholakia 2011b). It is not always profitable for the merchants to put up deals in the short term, because the discount rate is high in general and the revenue is shared with the platform.<sup>6</sup> However, the merchants can use the deal as an advertisement to

---

<sup>6</sup>With a typical deal discounted at about 50%, and the platform sharing about 50% of the remaining revenue, a merchant is often left with 25% or less of the face value. Dholakia (2011b) conducted a survey

reach new consumers and gain returning consumers (Edelman et al. 2016).

We obtained data from a company that specializes in aggregating deal offerings from most of the daily deals websites in the major U.S. markets. For each local market, we observe the deal offerings and sales on the major daily deals websites between March 2010 and January 2012. Our empirical analysis focuses on the leading platform in this industry: Groupon. In our dataset, Groupon deals amount to 38% of all deal offerings and 69% of all sales revenue. The second major player is LivingSocial, whose deal offerings account for 10% of the overall deals and the revenue accounts for 21% of the industry total in our dataset.<sup>7</sup>

For each deal, we know the market and the category in which the deal was offered, and the basic characteristics of the deal, including the merchant information, starting date, ending date, price, value, and discount. For example, Groupon offered a deal “\$25 for \$50 worth of dry-cleaning service at Pure cleaners” on May 26, 2010 that ended on the second day in the New York area. In this case, the price is \$25, the value is \$50, and the discount is 50%. Table 1 provides the summary statistics of the deal offerings by Groupon over time and across markets. On average, we observe 50 deals per week across markets with a transaction volume (total deal sales) about 16457. On average 65% of the deals each week are offered by merchants new to Groupon. The repeated merchants on average use Groupon 2.5 times in our data period.<sup>8</sup>

Table 2 summarizes the deal characteristics by category.<sup>9</sup> The average duration of a deal is 2-3 days and the average price of a deal listed on Groupon is \$67 with median price at \$30. While the price and the value of the deals vary substantially, the discount rate of deals is relatively stable. About 41% of deals offer a discount rate of 50%, and another 30% fall in the range of 50%-60% discount rate.

Groupon business is highly localized. There are 99 markets or Metropolitan Statistical

---

of businesses who ran daily deals between August 2009 and March 2011. The results revealed that more than half (55.5%) profited from their daily deal promotion, whereas over a quarter (26.6%) lost money.

<sup>7</sup>During the observation period, hundreds of websites listed deals across markets. Most of these were small and local sites that offered deals sparsely, and many have disappeared during the observation period. No other deal site had more than a 3% share of total deal listings.

<sup>8</sup>For LivingSocial, 95% of deals each week are offered by first time users of the platform.

<sup>9</sup>We dropped the deals with durations of more than 3 weeks in our analysis, which accounts for 0.17% of Groupon deals.

Area (MSA) in the data. The largest daily deals markets in terms of the number of deal offerings are Chicago, New York, Los Angeles, Boston, Atlanta, San Francisco, San Diego, and Denver (in descending order). Note that the local launch date of the daily deals sites varies across the MSAs. To examine the industry dynamics across markets, we find that a comparison of the number of weeks since the first presence of daily deals sites in the market ("history week") is more relevant than the calendar week. In other words, as Groupon entered different markets at different time, inspecting the markets according to their growth stage yields a more accurate picture of the evolution dynamics than taking the simple average across markets at successive points of calendar time. Given the local nature of the business, it takes time for deal sites to build relationships with local merchants. Thus, the clock for a local market does not start to run until the first entry of daily deals sites in that market. We base our analysis on the history-week time measure in the following.

Figure 1 shows the average weekly deal offerings and transaction volume of the top 8 largest markets since the existence of Groupon in these cities. The upward growing trend is obvious. We also plot the number of weekly deal offerings by the size of local population for each of the four main categories in Figure 2. Interestingly, we find that the number of weekly deals does not grow with the size of the local population after a certain point, and this pattern is similar for all main categories. Why does it happen? We argue that roles of other local market factors need to be properly accounted for and introduce them next.

## **2.2 The Role of Local Characteristics**

A unique characteristic of the Groupon business model, in comparison with other e-commerce companies such as Amazon and eBay, is its "localness". The platform lists deals from local merchants and the coupons purchased online are redeemed in the local stores. Between traditional brick-and-mortar retail representing one extreme (strong local boundaries) and e-commerce platforms (e.g., Amazon) representing the other extreme



(with no geographical restrictions and “unlimited” trading areas) (Bell, Choi and Lodish, 2012), the daily deal platforms sit in the middle. While Groupon has the nature of online business, its expansion is subject to local conditions.

In the offline world, consumers do not travel far to grab a cup of coffee, to work-out in the gym or to dry clean their clothes. Retailers in these sectors mainly serve the customers nearby. The online deals break these natural boundaries. Not only do consumers become aware of more options, but they are also willing to travel a bit farther to try new products and services because of coupon incentive. However, the cost of redeeming the deals offline would still affect consumer choices. Studies have found that “the location is too far from home or office” is an important reason for consumers to forgo deal opportunities (Ardizzone and Mortara, 2014). The location issue, or the redemption effort more generally, is not only salient for daily deals but also found to be important for other types of businesses bringing online consumers to offline merchants, such as platforms offering mobile coupons (e.g., Dickinger and Kleijnen, 2008). Consider a circle with focal consumers’ location as its center. Consumers would only buy deals from merchants within the circle. The radius of the circle would depend on the travel cost: lower travel cost would suggest that consumers are willing to travel further distance to redeem a deal coupon. Conditional on this circle radius, the number of merchants that consumers would be willing to consider depends on the store density. We thus use travel cost and store density as the key local characteristics and discuss their implications in the following.

We posit that travel cost and store density may affect consumers’ intentions to purchase deals in multiple ways. The first is the direct effect on baseline demand. Higher store density and lower travel cost may positively affect deal sales because of ease of redemption. The second is the indirect effect through affecting the substitutability among online deals. Given multiple similar deals available for purchase, consumers are more likely to purchase the deals from nearby merchants when travel cost is high. High travel cost therefore reduces the substitution among deals as the merchants are effectively more differentiated by location. In contrast, ease of travel (low travel cost) would make deals

more accessible and intensify the competition. This idea is consistent with prior work on spatial differentiation originating from Salop (1979) and Hotelling (1929). We expect similar effects from store density: high store density is likely to strengthen competition among similar deals.

In addition, store density and travel cost may affect consumers' responsiveness to some deal characteristics. Studies have shown that consumers' online purchases become less sensitive to online price discounts as distance to offline stores decreases (e.g., Forman et al. 2009). Similarly in our context of daily deals, when there are more stores in the local area, consumers might be more likely to find good promotions offline and less likely to be attracted by online deals. To test for this hypothesis (higher store density in the local market decreases consumers' sensitivity to the discount rate of online deals) we consider the interaction effect of deal discount and store density. Likewise, travel cost might also interact with deal discount - ease of travel may increase consumers' sensitivity to discounts.

For our empirical analysis, we collected multiple measures of travel cost by area.<sup>10</sup> The first is the travel time index, defined as the ratio of travel time in the peak period to travel time at free-flow conditions. A travel time index of 1.30 indicates that a 20-minute free-flow trip takes 26 minutes in the peak period. The second is the area's yearly traffic delay per commuter, which is calculated as the extra time spent during the year traveling at congested speeds rather than free-flow speeds by private vehicles. The third measure is the congestion cost per commuter calculated as the yearly value of delay time (estimated at \$17.67 per hour of travel) and wasted fuel (estimated using state average cost per gallon for gasoline and diesel). The summary statistics of these variables across markets are reported in Table 3. Figure 3 (a) plots the relationship between travel time index and population. We find that cities with relatively small populations can also have high travel cost. To construct the variable 'store density', we first collected data about the number of local stores or merchants in each major category from the Esri Demographics and Business Database,<sup>11</sup> and then divided the number of stores in a category in an

---

<sup>10</sup>Source: 2015 Urban Mobility Scorecard, Texas A&M Transportation Institute.

<sup>11</sup>See [http://www.esri.com/data/esri\\_data/business-overview/business](http://www.esri.com/data/esri_data/business-overview/business) for the database details. The

area by the geographic size of the area. Interestingly, store density is not statistically significantly correlated with population size or travel index. Many cities with relatively small populations have relatively high store density as shown in Figure 3 (b).

Apart from the cost of deal redemption offline, we consider local characteristics that would directly affect the level of deal demand. Earlier exploratory studies have identified two main motivations for consumers to purchase online deals: saving money on things that they would have purchased anyway or trying new things (e.g., Dholakia and Kimes, 2011; Ardizzone and Mortara, 2014). The former is determined by the depth of discount while the latter can be proxied by local characteristics related to consumers' openness to new things. The measure we use is the percentage of conservative voters in a city.<sup>12</sup> Research in political science and psychology has found that political orientation is correlated with personal traits. Carney et. al (2008) found that openness to new experience is an important trait underlying the differences between liberals and conservatives, and that the liberals are more open to try new things in their daily life. Therefore, we conjecture that deals sales are higher in a more liberal (less conservative) city. The correlation between the percentage of conservative voters and population is -0.38, which indicates that cities with larger population sizes are more likely to be liberal and open to new things. We also collected other demographic variables at MSA level, including population size, gender ratio, average age, education level etc, summarized in the lower panel of Table 3.

We have discussed how local characteristics may affect the demand side of Groupon deals. In addition to the effect through deal sales, local characteristics can also have a direct impact on deal supply. We hypothesize that merchants are more likely to adopt online deals in markets where travel cost is high. When it is more inconvenient or more costly for consumers to physically reach a destination, merchants may have a higher motivation to utilize the online deal platform to advertise and to attract customers with

---

categories reported in the Esri Demographics and Business Database do not match perfectly with the categories in our data. We therefore use the number of "health and personal care stores" to correspond to Beauty and Fitness category, use the number of "sports/hobby/book/music stores" for Entertainment and "food and beverage stores" for Restaurant.

<sup>12</sup>Data source: "The Most Conservative and liberal Cities in the United States," The Bay Area Center for Voting Research (BACVR), available at <https://alt.coxnewsweb.com/statesman/metro/081205libs.pdf>.

coupon. We also expect deal supply to be affected by store density. On the one hand, high store density could imply stronger local competition and increase the need to attract new customers using online deals. On the other hand, lower profit margins associated with intense competition may restrict merchants' ability to offer deeper discounts. It is therefore an empirical question as which effect is likely to dominate. Merchants' incentive to adopt a daily deals online platform can also be affected by local demographic characteristics. For example, merchants in more conservative markets could be less open to trying new tools. Finally, we also expect other demographic variables such as population size to affect deal supply.

In the following, we empirically investigate the role of local characteristics and other factors in determining the expansion of Groupon from both customer and merchant sides.

## 3 Empirical Analysis

### 3.1 Model

We model the weekly number of deals and deal sales as a result of the simultaneous decisions of merchants and consumers. Technically, contracted merchants need to wait several months before the deals become active. Although the merchants may not know the exact dates of deal activation, we assume that they have correct expectations of future demand. For instance, merchants selling outdoor adventure items might expect higher demand during summer. Therefore, the market outcome in terms of deal offerings and sales can be approximated by the simultaneous model.

## Merchant side

On the deal offering side, we assume that the number of deals in category  $j$  of market  $m$  at period (week)  $t$  can be expressed by the following equation:<sup>13</sup>

$$(1) \ln N_{jmt} = v_{jmt} + \gamma_1 \ln S_{jmt} + \gamma_2 \ln \sum_j b_{jmt-1} + \alpha^{tul} Tulcost_m + \alpha^{den} Storeden_{jm} \\ + \alpha_Z Z_m + \alpha_W W_{jmt-1} + \alpha_T T_t + \alpha_D D_j + \epsilon_{jmt}^N$$

where  $v_{jmt}$  captures the category-specific unobserved state of platform attractiveness which evolves over time.

$$(2) \quad v_{jmt} = \delta^v v_{jmt-1} + \varphi_1 \ln N_{jmt-1} + \varphi_2 \ln N_{jmt-1}^c + \eta_{jmt}^v$$

The scalar  $\delta^v$  is the depreciation factor which captures the carryover effect of platform attractiveness or perceived quality. The perceived quality in the current period is affected by the number of deals in the past. More merchants using the platform in the last period may encourage more firms to try the platform this period, which captures the observational learning effect on the merchant side. We also allow the rival platform LivingSocial's popularity, i.e., the number of deals  $\ln N_{jmt-1}^c$ , to affect the attractiveness of the focal platform.  $\eta_{jmt}^v$  is the random shock that affects the platform attractiveness in the current period. One source could be the platform's periodic investment on the merchant side which is unobservable to researchers.

In addition to the perceived platform quality, the number of merchants choosing to list deals may depend on the (expected) deal sales,  $\ln S_{jmt}$ . Merchants may also use the platform for advertising purposes. Past studies found that the exposure to potential customers can be more valuable for merchants than the deal revenue in attracting merchants to run deals (e.g., Dholakia 2011a).  $b_{jmt-1}$  is category  $j$ 's transaction volume in the last period ( $b_{jmt-1} = N_{jmt-1} \cdot S_{jmt-1}$ ) and the term  $\sum_j b_{jmt-1}$  represents the total transaction volume across categories in the last period on the platform, which approximates the size

---

<sup>13</sup>We use log transformation of the number of deals and deal sales as these variables are highly skewed.

of the audience that the deal may be potentially exposed to.<sup>14</sup>

Local characteristics may also affect merchants' incentive to use online deal platform. In particular, we consider the effect of travel cost and store density on deal supply. As discussed in the previous section, we expect that merchants are more likely to offer online deals when the travel cost is high. Store density can be considered as a proxy for the degree of local competition. The overall effect of store density on deal offering would depend on which of the following two effects dominates. First, stronger competition may drive merchants to utilize the deal platform to attract new consumers. Second, stronger competition may cut margins and constraint merchants' ability to offer deals. The parameter  $\alpha^{den}$  captures the net effect. We also control for the set of usual demographic variables such as population size and income level, as well as the percentage of conservative voters in the market, summarized by vector  $Z_m$ .

Merchants' intention to use the platform in the current period can also be affected by the past experience with the platform. We use two measures on repeated usage of the platform. The first is the percentage of deals offered by returning merchants who have put up deals in this category before period  $t$ . The second is the average number of times that the merchants in this category have used Groupon before period  $t$ . The two are summarized by vector  $W_{jmt-1}$ . We also control for the time trend and seasonality effect. Specifically, we include the number of months since the platform first entered the market, its square term, year dummy and month dummy in  $T_t$ . Earlier market entrance by Groupon would naturally correspond to more merchants being familiar with the platform. Also, the number of deals in a category is likely to be affected by the time of the year. For example, there might be more (or fewer) entertainment deals in holiday season. These variables partially control for the planning role of the platform, which takes seasonality into consideration. Finally,  $D_j$  is the category dummy capturing category fixed effect. The error term  $\epsilon_{jmt}^N$  is the random shock that follows normal distribution.

---

<sup>14</sup>We use the lagged transaction volume to capture the advertising effect for model tractability reasons. Using current transaction volume would introduce additional simultaneity/endogeneity problem and quadruple the number of simultaneous equations (i.e., multiply by four categories). We assume that merchants may not have the correct expectation of all the categories and they use lagged transaction volume as a proxy for the size of the audience on the platform that can be reached in the current period.

## Consumer side

On the deal sales side, we model the per-deal sales in category  $j$  of market  $m$  at period (week)  $t$  as follows:

$$\begin{aligned}
 \ln S_{jmt} &= u_{jmt} + \gamma_3 \ln N_{jmt} + \gamma_4 \ln N_{jmt}^c + \beta_X X_{jmt} + \beta^{den} \text{Storeden}_{jm} + \beta^{ttl} \text{Tvlcost}_m \\
 &+ \gamma^{ttl} [\ln N_{jmt} \cdot \text{Tvlcost}_m] + \gamma^{den} [\ln N_{jmt} \cdot \text{Storeden}_m] \\
 (3) \quad &+ \beta^{DS} (\text{Disc}_{jmt} \cdot \text{Storeden}_{jm}) + \beta^{DT} (\text{Disc}_{jmt} \cdot \text{Tvlcost}_m) \\
 &+ \beta_Z Z_m + \beta_D D_j + \beta_T T_t + \epsilon_{jmt}^S
 \end{aligned}$$

where  $u_{jmt}$  is the unobserved platform attractiveness of category  $j$  from consumers' perspective which evolves as follows:

$$(4) \quad u_{jmt} = \delta^u u_{jmt-1} + \phi_1 \ln b_{jmt-1} + \phi_2 \ln b_{jmt-1}^c + \eta_{jmt}^u$$

It is a function of the transaction volume on the platform in the past. More consumers purchasing deals correspond to higher perceived platform attractiveness. We interpret it as the cumulative effect of word-of-mouth on the consumer side. We also allow the traffic of the same category and same market on the rival platform,  $\ln b_{jmt-1}^c$ , to affect the perception of the focal platform.  $\delta^u$  is the depreciation parameter. The larger the value (or smaller depreciation), the more persistent is the impact of the past performance on the current perception.  $\eta_{jmt}^u$  is the random shock that follows i.i.d normal distribution.

The second term in equation (3) is the effect of the number of deal offers in the category on per deal sales, which can be positive or negative. The parameter  $\gamma_3$  estimates the net effect of demand expansion (more variety) and the competition among similar deals. If the parameter is negative, it suggests that the competition effect dominates and more deals in the same category would cannibalize the sales of each other.<sup>15</sup>  $\ln N_{jmt}^c$  is the number of deals in the same category at the same week on LivingSocial in log form, and

<sup>15</sup>We also estimated a model with both  $\ln N$  and  $(\ln N)^2$  controlled. The coefficient on  $\ln N$  is similar to the current estimate while the coefficient in front of the square term is insignificant and of much smaller magnitude. We thus use the current specification.

$\gamma_4$  captures the potential competition effect from the rival platform.  $X_{jmt}$  is the set of average deal characteristics including duration, price and discount of category  $j$  at period  $t$  in market  $m$ . We expect higher price to discourage sales while deeper discount to attract more consumers.

Deal sales are also likely to be affected by local characteristics. First, we consider the direct effect on deal demand. In particular, we control for the store density of category  $j$  and the travel cost in the market. We also control for other local sociographic and demographic variables, including the percentage of conservative voters in the local market, summarized by the vector  $Z_m$ . Second, we consider the interaction effect of local characteristics with other effects. As discussed in the previous section, we conjecture that travel cost will affect consumers' choice among similar deals. High travel cost will further differentiate similar deals offered by merchants in different locations while low travel cost will decrease such differentiation and make the deals in the same category more substitutable. We therefore include the interaction term of travel cost with the (log) number of deals in a category to test the hypothesis. The parameter  $\gamma^{tol}$  captures the potential moderating role of travel cost. Similarly, we conjecture the interaction effect of store density with the (log) number of deals to be negative, increasing the substitution effect between deals. We also consider the interaction effect of store density with the deal discount level. With higher store density in the local market, consumers may be able to find promotions or deals more easily in a local store and become less sensitive to the discount rate of online deals. We also include the interaction between discount level and travel cost. Consumers could be more responsive to discounts under low travel cost.

In addition, we control for time trend and seasonality in the vector of  $T_t$ .  $D_j$  is category specific effect. The error term  $\epsilon_{jmt}^S$  captures the random demand shocks which follow normal distribution.



## 3.2 Estimation

There are two main issues in estimating the coefficients in this system described above. First, the number of deal offerings in a period and the deal sales are endogenously determined as depicted by equation (1) and (3). Second, the unobserved platform attractiveness to merchants and consumers,  $v_{jmt}$  and  $u_{jmt}$ , are unobservable and dynamically evolving over time. We use Kalman filter method together with control function approach (Petrin and Train, 2010) to address these issues.

We first discuss the identification of the simultaneous equations. The exclusion restrictions help to identify the structural parameters in equation (1) and (3). The variables of merchant experience  $W_{jmt}$  only affect the number of deal offerings but do not directly affect consumer choice. The deal characteristics  $X_{jmt}$  of the current period affect the current deal sales but do not directly affect the number of deals in the current period.<sup>16</sup> We also assume the lagged observations such as lagged number of deals and lagged deal sales are predetermined and therefore are not endogenous. This assumption can be problematic if the error terms  $\epsilon_{jmt}^N$  and  $\epsilon_{jmt}^S$  are serially correlated. In our case, we find that serial correlation is not a major concern after accounting for the time-evolving platform attractiveness.<sup>17</sup>

The estimation is complicated by the unobserved evolving platform attractiveness. We apply the Kalman filter method in estimating the parameters. Kalman filter has been used to estimate unobserved time-varying state such as product attractiveness (Wang et al., 2015) and advertising awareness and quality (Naik et al. 1998). The transition of the

---

<sup>16</sup>Deal characteristics such as discount rates are largely determined by Groupon and follow previous deals with similar merchants/products, according to typical contractual details described in numerous media publications that documented merchants' experiences. Groupon approaches every business with the standard terms of 50% discount rate and 50/50 revenue split (See <http://www.extramoneyblog.com/negotiating-with-groupon/>). Merchants might be able to negotiate the split of revenue if they are "well known within the community" or they have "a unique product offering" (See <http://blog.agrawals.org/2012/01/21/daily-deals-merchant-faq/>), but rarely did we find that they were able to negotiate the discount rates. Some merchants described the Groupon merchant agreement as "incredibly lopsided in favor of Groupon, as are most agreements where one party doesn't really have the ability to negotiate" (See <http://blog.agrawals.org/2011/06/07/an-analysis-of-the-groupon-merchant-agreement/>).

<sup>17</sup>To check for serial correlation, we use the method proposed by Godfrey (1994) for dynamic model with endogenous variables. The method entails two steps. In the first step, we recover the residuals from each equation,  $\hat{\epsilon}_{jmt}^N$  and  $\hat{\epsilon}_{jmt}^S$ . In the second step, we include the lagged residual in each equation as new regressor and re-estimate the model. The parameters in front of the lagged residuals are insignificant, which suggests that autocorrelation of the error terms is not a major concern.

states is described by equation (2) and (4), which can be expressed in the matrix form:

$$(5) \quad \begin{bmatrix} v_{jmt} \\ u_{jmt} \end{bmatrix} = \begin{bmatrix} \delta^v & 0 \\ 0 & \delta^u \end{bmatrix} \begin{bmatrix} v_{jmt-1} \\ u_{jmt-1} \end{bmatrix} + \begin{bmatrix} \beta_1 & 0 \\ \phi_1 & \phi_1 \end{bmatrix} \begin{bmatrix} \ln N_{jmt-1} \\ \ln S_{jmt-1} \end{bmatrix} + \Pi \cdot \Gamma_{t-1} + \begin{bmatrix} \eta_{jmt}^N \\ \eta_{jmt}^S \end{bmatrix}$$

where  $\Gamma_{t-1}$  is the vector of exogenous variables that affect the evolution of the states and  $\Pi$  is the associated parameter matrix. The error terms follow normal distribution,

$$(6) \quad \begin{bmatrix} \eta_{jmt}^N \\ \eta_{jmt}^S \end{bmatrix} \sim N[0, \varpi]$$

The observation equations are equation (1) and (3). The Kalman filter uses the current observations on  $\ln N_{jmt}$  and  $\ln S_{jmt}$  to predict the values of the unobserved states next period and uses the realizations to update the forecast using Bayesian theorem.

The estimation proceeds as follows.<sup>18</sup> In the first step, we use control function approach to address the endogeneity problem.<sup>19</sup> We regress the two endogenous variables  $\ln N_{jmt}$  and  $\ln S_{jmt}$  on all the exogenous or predetermined variables in the system respectively.<sup>20</sup> The residuals are then included as additional variables in the observation equations in the second step. Based on the recursive nature of Kalman filter, given a vector of parameters, we can calculate the mean and variance of the state variables of the current period based on the information up to the last period. Given the state variables, we have the prediction of outcome, i.e. the number of deals and deal sales. Therefore, we can construct the conditional log-likelihood function using the probabilities of observing the number of deals and deal sales in each category in each market at each period given the information set:

$$(7) \quad LL(\theta) = \sum_j \sum_m \sum_t \ln[Prob(Y_{jmt}|I_{t-1}, \theta)]$$

<sup>18</sup>We provide more details in the Appendix.

<sup>19</sup>Blundell, Kristensen and Matzkin (2013) discuss the use of control function approach in estimating simultaneous equation models.

<sup>20</sup>The F-statistics of the regression on  $\ln N_{jmt}$  is 740 and and F-statistics of the regression on  $\ln S_{jmt}$  is 360.

where  $\theta$  is the vector of parameters to be estimated,  $I_{t-1}$  is the information set available at period  $t$  and

$$(8) \quad Y_{jmt} \equiv \begin{bmatrix} \ln N_{jmt} \\ \ln S_{jmt} \end{bmatrix}$$

### 3.3 Results

We first estimate a baseline model without controlling for the local characteristics. The estimation results are reported in the first column of Table 4. First, we look at the results of state transition which determines the unobserved platform attractiveness on the merchant side and consumer side. On the merchant side, we find that the number of deals in the past, which can approximate the number of merchants using the platform in the past, has a positive effect on the perceived quality of the platform in the current period. This can be interpreted as the merchant side observational learning. More merchants adopting Groupon in the past increases the attractiveness of the platform. Interestingly, we find that the number of deals on the rival site LivingSocial also has a positive effect on the unobserved platform attractiveness. We interpret this as an industry spillover effect in the early stage. The past users of rival platforms are likely to switch to Groupon to list deals this period. On the consumer side, deal sales in the past have a positive effect on the perceived quality of the platform, which we interpret as the word-of-mouth effect on the consumer side. The sales on the rival platform has a negative but insignificant impact. The depreciation parameters  $\delta^v$  and  $\delta^u$  are both close to 0.59, which captures the accumulation of perceived qualities on either side.

Next, we examine the results of the two main equations which determine deal offerings and deal sales. For deal offerings, the expected sales has a positive effect on the number of deals. Total transaction volume on the platform, which we interpret as the potential audience exposed to such advertising, also has a positive and significant effect on deal offerings. This result offers empirical evidence that advertising to consumers on the

platform is an important motivation for merchants to run online deals. In addition, a higher percentage of returning merchants and more past experience with Groupon make merchants more willing to list deals in the current period.

For deal sales, we find that the number of deals in a category has a negative and significant effect on per deal sales. In other words, the substitution effect dominates any positive variety effect as the number of similar deals increases in a category. The competition from the same category on the rival platform exists but of a much smaller magnitude. In terms of deal characteristics, we find that deeper discount can increase deal sales and that higher price would decrease deal sales, which is not surprising. The effect of duration on deal sales is not significant.

Column (2) reports the results from the full model. We now focus on the effect of local characteristics (other main results are similar to the results of the baseline model). On the merchant side, we find that travel cost, measured by travel time index, has a positive and significant effect on deal offering. In other words, merchants are more likely to use online deals to attract consumers when the local travel cost is high. Store density, on the other hand, has a negative effect on deal supply. Although a priori strong competition can either encourage or discourage deal offerings, we find that the net effect of store density on deal supply is negative.

On the consumer side, we find that deal sales are significantly higher in markets with higher store density. Higher store density may suggest lower redemption cost, which has a positive effect on demand. The main effect of travel cost on deal sales is not significant. However, we find that the interaction of travel cost and the number of deals is positive and significant. Recall that the coefficient for the number of deals captures the net effect of deal substitution, which is negative. Then the positive interaction suggests that high travel cost effectively lowers the substitution effect among deals. The effect is highly robust when we use alternative measures of travel cost, as shown in Table 5. The results indicate that distance and redemption cost matter in consumer choice of deals. Consistent with this argument, we find that high store density or ease of redemption increases deal substitution. The interaction effect of store density and the number of

deals is negative although not significant. The local conditions also affect consumers' responsiveness to deal discounts. The interaction of store density and discount rate is negative and significant. It suggests that consumers are less sensitive to the magnitude of online deal discounts when having many stores in the local area. The interaction effect of travel cost and discount is not significant.

We now examine the impact of city demographic variables on deal supply and sales. We find that the number of deal offerings tends to be larger in cities with larger population, higher female percentage, higher income level and lower education. The level of conservativeness does not affect merchant side deal offering. On the demand side, deal sales tend to be higher in markets with larger population. Consistent with survey results (Dholakia and Kimes, 2011), younger people like Groupon more and markets with younger populations tend to have higher sales. Deal sales are also positively associated with household incomes. Interestingly, we find that deal sales are lower in more conservative cities.

In Figure 4, we present model fit and plot the observed and the predicted weekly number of deals and per deal sales for selected cities over time. Due to the nature of Kalman filter, the first few observations are used to initializing the updating process. The predictions quickly converge to the observed pattern. Overall, the model fits the data well.

## 4 Further Discussion

The estimation results above indicate that local characteristics can affect the expansion of Groupon from both the demand and supply side. To summarize, travel cost has a positive effect on deal sales by increasing differentiation among similar deals and lowering the substitution effect. Travel cost also has a positive effect on the number of deals offered in a market. Store density, on the other hand, has a negative impact on deal supply. On the demand side, store density has a positive main effect and a negative interaction effect

(through deal discount) on deal sales. Population size has a positive effect on both the merchant and consumer side. Other demographic variables affect deal offering and sales either in the same direction or only on one side. In this section, we use simulation to show how local characteristics such as travel cost, store density and population size contribute to the differentiated growth of Groupon across markets.

#### 4.1 Local Characteristics on Platform Growth

First, we conduct a comparative statistic study to examine the individual effect of different local characteristics on platform growth. In each exercise, we keep all other variables at the average values in the data while varying one of the three local characteristics: store density, travel cost and population size. For example, to disentangle the effect of travel cost on platform growth, we simulate deal offerings and deal sales for 20 periods for hypothetical markets with different travel costs while fixing the value of other variables at the mean value in the data. We then plot the transaction volume (number of deals multiplied by per deal sales) in log form in the last period against travel cost. The results are presented in Figure 5.

The results show that travel cost and population size contribute positively to the platform growth. Compare two otherwise similar cities where travel cost index in one city is 1.2 and the other is 1.4. The simulation result shows that by the end of the 20 periods, the transaction volume in the city with high travel cost would be 23% higher than in the other city. For example, consider Atlanta vs. Seattle. The population of Atlanta is 50% greater than that of Seattle (5.29 vs. 3.45 million). The store densities in the two cities are similar while the travel cost is higher in Seattle (travel index 1.38), compared to Atlanta (travel index 1.24). Figure 6 illustrates that Groupon was expanding faster in Seattle than in Atlanta. It suggests that in cities with moderate population size, travel cost can be a more important factor in determining the growth of deal platform.

The net impact of store density on transaction volume depends on the level of deal discount. When the deal discount is high (at the observed value of 55%), store density

has a net negative impact on both the demand and supply side. High store density lowers consumers' sensitivity to deal discounts. This negative interaction effect outweighs the positive main effect of store density on sales when deal discount is large. Together with the negative effect on the supply side, store density contributes negatively to the platform growth. For example, consider New York City vs. Washington DC. The population of New York City is 3.4 times that of Washington DC (18.9 vs. 5.61 million). The travel costs in the two cities are the same while the store density is much higher in New York City (15.0), compared with Washington DC (3.53). Figure 6 illustrates that the expansions of Groupon in the two cities are comparable despite the much larger population in New York. It suggests that high store density does not help or may even hurt the expansion of such online deal platform. However, when the average discount rate is low (e.g., at 25% level), the positive main effect of store density on sales dominates, and store density contributes positively to the platform growth. The results are presented at the lower panel of Figure 5.

Overall, comparative statics suggest that population and travel cost contribute positively to platform growth while store density contributes negatively to platform growth given the actual average deal characteristics on the platform. The observed growth patterns of the markets are consistent with model predictions. To illustrate, we focus on the set of cities with more than 50 weeks of platform presence and compare their growth rates. In Table 6, we list these cities in descending order in terms of growth rate and bracket them into three tiers.<sup>21</sup> The first tier cities have experienced the highest platform growth. The second tier cities have also grown fast but are behind the first tier. Closer examination reveals that one of the key local characteristics is not as favorable in these cities.<sup>22</sup> The cities in the third tier were growing slower and have at least two unfavorable local conditions. Consistent with our model prediction, population has a positive effect on platform expansion and cities with larger population size are generally ranked higher in the table. Yet store density and travel index also play significant roles. On

---

<sup>21</sup>The growth rate is measured by the average transaction volume in the last 10 periods relative to that in the first 10 periods for each market.

<sup>22</sup>The highlighted (in bold) numbers in the table are suggestive of unfavorable local characteristics.

the one hand, some cities with large population size are ranked lower due to unfavorable store density (e.g., New York City) or unfavorable travel cost (e.g., Atlanta) or both (e.g., Boston). On the other hand, some cities with small population size are ranked above larger cities because of favorable store density or travel cost (e.g., Seattle, Denver). We also plot the platform growth for selected cities in Figure 7. Although there might be various factors that contribute to the different expansion rate of the platform across markets, our empirical model provides at least some explanation of the observed growth pattern.

## 4.2 Local Characteristics on Consumer and Merchant Sides

After demonstrating how local characteristics affect platform growth overall, we proceed to decompose that overall impact of local market characteristics into relative effects on consumer and merchant sides by conducting two sets of simulation. In the first simulation, we suppress the effect of the local characteristics on the consumer side (i.e., the effect coefficients are set to zero) and keep the effect on the merchant side as estimated. In the second simulation, we suppress the effect of the local characteristics on the merchant side and keep the effect on the consumer side only. For each scenario, we simulate the growth pattern by varying the value of one of the local characteristics on one side of the platform. We then compare the log transaction volume at the end of the 20th period under the two scenarios with the full model where local effects on both consumer and merchant sides are allowed.

The results are presented in Figure 8. In each graph, the solid line represents the simulation results from the full model; the dashed line represents the scenario where the effect of local characteristics on merchant side is removed; the line with dots represents the case where the effect on the consumer side is removed. The first graph indicates that the effects of population on the two sides are comparable. Dropping the consumer-side effect or dropping the merchant-side effect results in a similar magnitude of drop in transaction volume. The second graph indicates that travel cost has stronger positive impact on the



consumer side than on the merchant side. Dropping the consumer side effect results in a much larger reduction in the platform growth than dropping the merchant-side effect. The third graph indicates that store density has stronger negative impact on the merchant side than on the consumer side. The effect of store density on transaction volume is negative (at the observed discount rate). Removing the negative effect on the merchant side leads to a higher increase in transaction volume than dropping the consumer-side effect.

To summarize, the simulation results suggest that local characteristics affect the demand and supply of Groupon deals differently: 1) travel cost has stronger (positive) effect on the demand side; 2) store density has stronger (negative) effect on the deal supply side; 3) population size has comparable effects on both sides.

## 5 Conclusion

In this paper, we examine the factors that affect the growth and the scale of Groupon, the leading daily deals platform. We estimate a simultaneous model of deal offerings and deal sales which captures the two-sided nature of the platform. The results indicate that there are significant word-of-mouth effect on the consumer side and observational learning effect on the merchant side which contribute to the expansion of the platform. However, there is also a significant deal substitution effect among the deals in the same category which prevents the platform size to grow further. We find that such competition within the platform is a more important factor than the competition from the rival platform in restricting the deal sales and the number of deals, at least during the period under investigation.

The uniqueness of the Groupon business model lies in its strong connection with the local market. We investigate how the local characteristics shape Groupon's expansion, leveraging the rich data across markets and the heterogeneity of markets. In addition to the conventional demographic variables such as population size, we find that local market characteristics that affect the cost of deal redemption offline have a significant impact on

consumers' deal choice behavior. In particular, we find that higher travel cost decreases the substitution effect and that the effect is highly robust. We also find that high store density decreases consumers' sensitivity to the deal discount rate. Furthermore, we find that travel cost and store density also affect the supply side. Specifically, high store density is associated with lower deal supply. Overall, our empirical results show that properly accounting for local market conditions is important for platforms with both online and offline components.

Current research also contributes to the literature more broadly, by examining the growth of two-sided platforms in O2O business. Addressing this problem is particularly relevant given the rapid growth of technology-enabled O2O platforms in various business contexts. We illustrate how these platforms differ from traditional two-sided online platforms by emphasizing the importance of local characteristics in determining the growth and scale of these platforms. The results offer applicable insights for other O2O platforms when evaluating market potentials and making entry decisions. In particular, the scale of the platform is not necessarily determined by the size of the market. Smaller cities with certain characteristics might provide better growth opportunities than larger cities.

There are several avenues for future research. First, it would be important to examine the interaction of online and offline characteristics in contexts other than daily deals platforms, where local market factors can play different roles. For example, while in our settings local market factors are not affected by the platform, one may argue that such O2O platforms as Uber may have direct impact on relevant local characteristics, such as travel cost. Second, while our data period covers time when mobile commerce was dominated by online, future research may benefit from examining how the impact of local market characteristics on two-sided platforms may evolve as consumers gradually shift from making purchases online to mobile transactions (sometimes termed "mobile-to-offline" or M2O commerce). Finally, we hope future work will be able to examine the industry for a longer horizon and incorporate potential long-term dynamic effects. For example, slow but steady decrease in profitability of running online deals may eventually change merchants' incentive and limit the platform growth. Also, while not salient during

our data period, the competition between different deal platforms is likely to become more important as the industry evolve, and future research may benefit from analyzing how such competition may interact with various local market conditions.

## References

- [1] Ahn, D. Y., Duan, J. A., and Mela, C. F. (2014) “Managing User Generated Content: A Dynamic Rational Expectations Equilibrium Approach,” working paper, Duke University.
- [2] Anderson, Eric T., Fong, Nathan M., Simester, Duncan I., and Tucker, Catherine E. (2010). “How Sales Taxes Affect Customer and Firm Behavior: The Role of Search on the Internet,” *Journal of Marketing Research*, 47 (2): 229–39.
- [3] Ardizzone, A. and Mortara, A. (2014) “Consumers’ Motivations and Daily Deal Promotions,” *The Qualitative Report* 2014, Vol. 19, Article 61, 1-15. Available at <http://www.nova.edu/ssss/QR/QR19/ardizzone61.pdf>.
- [4] Armstrong, M. (2006) “Competition in Two-sided Markets,” *RAND Journal of Economics*, 37(3), 668-691.
- [5] Armstrong, M. and Wright, J. (2007) “Two-sided markets, Competitive Bottlenecks and Exclusive Contracts,” *Economic Theory*, 32(2), 353-380.
- [6] Bell, D., J. Choi and Lodish, L. (2012) “What Matters Most in Internet Retailing, ” *MIT Sloan Management Review*, 54, 27 - 33.
- [7] Blundell, R., Kristensen, D., and Matzkin, R. L. (2013) “Control Functions and Simultaneous Equations Methods,” *American Economic Review*, 103(3), 563-569.
- [8] Brynjolfsson, E., Hu, Y., and Rahman, M.S. (2009) “Battle of the Retail Channels: How Product Selection and Geography Drive Cross-Channel Competition,” *Management Science* 55 (11), 1755-1765.
- [9] Brynjolfsson, E., Smith, M.D. (2000) “ Frictionless Commerce? A Comparison of Internet and Conventional Retailers ,” *Management Science*, 46 (4), 563-585.
- [10] Byers, J. W., M. Mitzenmacher, and Zervas, G. (2012) “Daily Deals: Prediction, Social diffusion, and Reputational Ramifications,” In *Proceedings of the fifth ACM international conference on Web search and data mining*, 543-552.
- [11] Caillaud, B. and Jullien, B. (2003) “Chicken and Egg: Competition among Intermediation Service Providers,” *RAND Journal of Economics*, 34(2), 309-328.

- [12] Carney, D. R., Jost, J. T., Gosling, S. D., and Potter, J. (2008) "The Secret Lives of Liberals and Conservatives: Personality Profiles, Interaction Styles, and the Things They Leave Behind," *Political Psychology*, 29(6), 807-840.
- [13] Dickinger, A., Kleijnen, M. (2008) "Coupons going wireless: determinants of consumer intentions to redeem mobile coupons." *Journal of Interactive Marketing*, 22(3), 23-39.
- [14] Dubé, J-P., G.J. Hitsch and P. K. Chintagunta (2010) "Tipping and Concentration in Markets with Indirect Network Effects," *Marketing Science*, 29(2), 216-249.
- [15] Dholakia, U. M. (2010) "How Effective are Groupon Promotions for Businesses?" working paper, available at SSRN: <http://ssrn.com/abstract=1696327>.
- [16] Dholakia, U. M. (2011a) "What Makes Groupon Promotions Profitable for Businesses?" working paper, available at SSRN: <http://ssrn.com/abstract=1790414>.
- [17] Dholakia, U. M. (2011b) "How Businesses Fare with Daily Deals: A Multi-Site Analysis of Groupon, LivingSocial, Opentable, Travelzoo, and BuyWithMe Promotions," working paper, available at SSRN: <http://ssrn.com/abstract=1863466>.
- [18] Dholakia, U.M. and Kimes, S.E., (2011) "Daily deal fatigue or unabated enthusiasm? A Study of Consumer Perceptions of Daily Deal Promotions," working paper, available at SSRN: [http://papers.ssrn.com/sol3/Papers.cfm?abstract\\_id=1925865](http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=1925865).
- [19] Edelman, B.,S. Jaffe and S. Kominers, (2016) "To Groupon or not to Groupon: the Profitability of Deep Discounts," forthcoming, *Marketing Letters*.
- [20] Forman, C., Ghose, A., Goldfarb, A. (2009) "Competition Between Local and Electronic Markets: How the Benefit of Buying Online Depends on Where You Live," *Management Science*, 55(1), 47-57.
- [21] Hagiu, A. (2009) "Two-Sided Platforms: Product Variety and Pricing Structures," *Journal of Economics and Management Strategy*, 18(4), 1011-43.
- [22] Godfrey, L.G. (1994) "Testing for Serial Correlation by Variable Addition in Dynamic Models Estimated by Instrumental Variables," *The Review of Economics and Statistics*, 76(3), 550-559.

- [23] Harvey, A.C., Ruiz, E. and N.G. Shephard, (1994) "Multivariate Stochastic Variance Models," *Review of Economic Studies*, 61, 247-264.
- [24] Hotelling, H. (1929) "Stability in Competition," *Economic Journal*, 39, 41-57.
- [25] Hu, M., Shi, M., and Wu, J. (2013) "Simultaneous vs. Sequential Group-Buying Mechanisms," *Management Science*, 59(12), 2805-2822.
- [26] Jin, G. Z. and Rysman, M. (2012) "Platform Pricing at Sports Card Conventions," *NBER Working Paper*, (No. w17959).
- [27] Jing, J. and Xie, J. (2011) "Group Buying: A New Mechanism for Selling Through Social Interactions," *Management Science*, 57(8), 1354-1372.
- [28] Kaiser, U. and Wright, J. (2006) "Price Structure in Two-sided Markets: Evidence from the Magazine Industry," *International Journal of Industrial Organization*, 24(1), 1-28.
- [29] Kim, B-C, J. Lee and H. Park (2013) "Platform Entry Strategy in Two-Sided Markets: Evidence from the Online Daily Deals Industry," working paper, Georgia Institute of Technology.
- [30] Li, X. and L. Wu (2013) "Observational Learning and Social-Network Word-of-Mouth: Evidence from Groupon," working paper, available at SSRN: <http://ssrn.com/abstract=2264411>.
- [31] Luo, X., Andrews, M., Song, Y. and Aspara, J. (2014) "Group-Buying Deal Popularity," *Journal of Marketing*, 78(2), 20-33.
- [32] Naik, P. A., Mantrala, M. K., and Sawyer, A. G. (1998) "Planning Media Schedules in the Presence of Dynamic Advertising Quality," *Marketing Science*, 17(3), 214-235.
- [33] Petrin, A. and Train, K., (2010) "A Control Function Approach to Endogeneity in Consumer Choice Models," *Journal of Marketing Research*, 47(1), pp.3-13.
- [34] Rochet, J. C. and Tirole, J. (2003) "Platform Competition in Two-sided Markets," *Journal of the European Economic Association*, 1(4), 990-1029.

- [35] Rysman, M. (2004) "Competition between Networks: A Study of the Market for Yellow Pages," *The Review of Economic Studies*, 71(2), 483-512.
- [36] Salop, S. (1979) "Monopolistic Competition with Outside Goods," *The Bell Journal of Economics*, 10, 141-156.
- [37] Song, M. (2011) "Estimating Platform Market Power in Two-sided Markets with an Application to Magazine Advertising," Simon School Working Paper, No. FR 11-22.
- [38] Stephen, A. T., and Toubia, O. (2010) "Deriving Value from Social Commerce Networks," *Journal of Marketing Research*, 47(2), 215-228.
- [39] Wang, Y., Lewis, M., Singh, V. (2015) "The Unintended Consequences of Counter-marketing Strategies: How Particular Antismoking Measures May Shift Consumers to More Dangerous Cigarettes," *Marketing Science*, 35(1), 55-72.
- [40] Wu, J., Shi, M., and Hu, M. (2014) "Threshold Effects in Online Group Buying," *Management Science*, 61(9), 2025-2040.
- [41] Zhang, K., Evgeniou, T., Padmanabhan, V., and Richard, E. (2012) "Content Contributor Management and Network Effects in a UGC Environment," *Marketing Science*, 31(3), 433-447.
- [42] Zhu, F. and Iansiti, M. (2012) "Entry into Platform-Based Markets," *Strategic Management Journal*, 33(1), 88-106.

## 6 Appendix

The estimation proceeds as follows. First, we regress the two endogenous variables  $\ln N_{jmt}$  and  $\ln S_{jmt}$  on all the exogenous or predetermined variables in the system respectively and obtain the residuals  $\hat{\epsilon}_{jmt}^N$  and  $\hat{\epsilon}_{jmt}^S$ .

The observation equations can be written in the vector form:

$$(9) \quad \begin{bmatrix} \ln N_{jmt} \\ \ln S_{jmt} \end{bmatrix} = \begin{bmatrix} v_{jmt} \\ u_{jmt} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{jmt}^N & 0 \\ 0 & \mathbf{Z}_{jmt}^S \end{bmatrix} \begin{bmatrix} \Upsilon_N \\ \Upsilon_S \end{bmatrix} + \begin{bmatrix} \tau_1 \hat{\epsilon}_{jmt}^N \\ \tau_2 \hat{\epsilon}_{jmt}^S \end{bmatrix} + \begin{bmatrix} \epsilon_{jmt}^N \\ \epsilon_{jmt}^S \end{bmatrix}$$

where  $\mathbf{Z}_{jmt}^N$  and  $\mathbf{Z}_{jmt}^S$  are all the control variables in the original equations respectively. The above equation can be further expressed compactly as:

$$(10) \quad Y_{jmt} = U_{jmt} + \mathbf{Z}_{jmt} \Upsilon + \tau \hat{\epsilon}_{jmt} + \epsilon_{jmt}$$

where  $\epsilon_{jmt} \sim N(0, V)$  and each element corresponds to a vector or matrix above with

$$(11) \quad Y_{jmt} \equiv \begin{bmatrix} \ln N_{jmt} \\ \ln S_{jmt} \end{bmatrix}$$

and

$$(12) \quad U_{jmt} \equiv \begin{bmatrix} v_{jmt} \\ u_{jmt} \end{bmatrix}.$$

Similarly, we can express the state transition equation (5) in vector form:

$$(13) \quad U_{jmt} = \delta U_{jmt-1} + \phi Y_{jmt-1} + \Pi \cdot \Gamma_{t-1} + \eta_{jmt}$$

where  $\eta_{jmt} \sim N(0, \varpi)$ .

We then follow the Kalman filter process to estimate the parameters (see Harvey 1994 and Naik et al. 1998).

1. Use the intercepts of the linear regression (equation (10)) as the initial value of  $U_{jm0}$ . Initial variance  $\Sigma_0$  of the unobserved states is assumed to be some large numbers.
2. Given the parameter guess and the information up to period  $t-1$ , the estimates of



the mean and variance of the state variables are:

$$U_{jmt|t-1} = \delta U_{jmt-1|t-1} + \phi Y_{jmt-1} + \Pi \cdot \Gamma_{t-1}$$

$$\Sigma_{jmt|t-1} = \delta \Sigma_{jmt-1|t-1} \delta' + \varpi$$

3. Given the state variables, the prediction error of  $Y_{jmt}$  and the associated variance are:

$$e_{jmt|t-1} = Y_{jmt} - U_{jmt|t-1} - \mathbf{Z}_{jmt} \Upsilon - \tau \hat{e}_{jmt}$$

$$S_{jmt|t-1} = \Sigma_{jmt|t-1} + V$$

4. Update the posterior of the unobserved state variables:

$$U_{jmt|t} = U_{jmt|t-1} + \Sigma_{jmt|t-1} [S_{jmt|t-1}]^{-1} e_{jmt|t-1}$$

$$\Sigma_{jmt|t} = \Sigma_{jmt|t-1} - \Sigma_{jmt|t-1} [S_{jmt|t-1}]^{-1} \Sigma_{jmt|t-1}$$

We can obtain the series for  $t = 1, 2, \dots, T$  by iterating the steps 2-4.

The conditional log-likelihood function can be written as:

$$(14) \quad LL(\theta) = \sum_j \sum_m \sum_t -\ln(2\pi) - \frac{1}{2} \ln |S_{jmt|t-1}| - \frac{1}{2} e'_{jmt|t-1} [S_{jmt|t-1}]^{-1} e_{jmt|t-1}$$

The parameters are estimated using maximum likelihood method.

Table 1: Summary Statistics

	Mean	Std.Dev.
Number of deals (weekly)	49.43	38.03
Number of transactions (weekly)	16457.28	19818.59
Ratio of first time merchant	0.65	0.23
Listings from the same merchant	2.49	11.53

Table 2: Summary Statistics: by Category

	Mean	Std.Dev.	Min	Mdn	Max
Total (110,237)					
Duration	2.79	1.40	0	3	20
Price	66.97	209.20	1	30	25000
Value	185.75	559.25	3	75	33800
Discount	56.52	10.10	0	52	99
Sold	328.21	2139.73	0	100	640000
Beauty (17.60%)					
Duration	2.93	1.33	0	3	13
Price	105.54	176.00	4	55	2999
Value	366.32	803.39	8	140	33800
Discount	59.89	11.62	5	55	97
Sold	194.72	442.25	0	85	17000
Fitness (7.22%)					
Duration	2.66	1.25	0	3	20
Price	45.88	63.58	4	35	2475
Value	150.35	175.28	8	110	4950
Discount	63.99	12.56	39	61	96
Sold	287.15	687.31	0	110	23731
Entertainment (16.68%)					
Duration	2.93	1.43	0	3	17
Price	56.03	126.02	2	30	8906
Value	122.46	242.25	3	64	12000
Discount	52.65	6.38	0	50	97
Sold	350.72	1498.56	0	94	95555
Restaurants (14.10%)					
Duration	2.38	1.12	0	2	10
Price	23.09	60.38	1	15	3600
Value	49.94	131.51	4	30	7600
Discount	51.68	3.57	27	50	98
Sold	701.14	1005.30	0	410	19753

Note : number or percentage of observations in the parentheses.

Table 3: Summary Statistics: Local Characteristics and Demographics

	Mean	Std.Dev.	Min	Mdn	Max
Population (in million)	1.98	2.68	0.12	1.13	18.92
Store density	8.19	10.80	1.44	5.73	92.25
Travel index	1.20	0.08	1.05	1.18	1.43
Yearly traffic delay per commuter (hours)	38.8	15.0	7	39	82
Yearly congestion cost per commuter (dollars)	901	341	153	927	1,834
Pct. of conservative voters	37.9	14.4	0.0	40.8	74.8
Female (%)	51.0	0.65	49.6	51.1	52.2
Age	30.0	0.81	27.8	30.1	33.0
Income (%)	47.3	5.1	33.1	47.2	60.5
Education (%)	29.9	5.9	15.5	29.2	46.8

*N*: 99 (MSAs)

*Note* : “Income” represents percentage of households with annual household income above \$50,000. “Education” represents percentage of households with bachelor degrees or above.

Table 4: Estimation Results

	Baseline		Full	
State transition equation				
<b>Merchant side <math>v</math></b>				
Platform attractiveness accumulation ( $\delta^v$ )	0.5862***	(0.0146)	0.6329***	(0.0087)
Number of deals ( $\varphi_1$ )	0.0346***	(0.0085)	0.0322***	(0.0054)
Rival's number of deals ( $\varphi_2$ )	0.0716***	(0.0067)	0.0233***	(0.0063)
<b>Consumer side <math>u</math></b>				
Platform attractiveness accumulation ( $\delta^u$ )	0.5909***	(0.0040)	0.6364***	(0.0057)
Transaction volume ( $\phi_1$ )	0.0169***	(0.0027)	0.0169***	(0.0028)
Rival's transaction volume ( $\phi_2$ )	-0.0066	(0.0044)	-0.0052	(0.0043)
Measurement equation				
<b>Merchant side <math>N</math></b>				
Deal sales ( $\gamma_1$ )	0.0689***	(0.0032)	0.0355***	(0.0034)
Transaction volume ( $\gamma_2$ )	0.0554***	(0.0007)	0.0873***	(0.0013)
Return	0.1361***	(0.0015)	0.2182***	(0.0021)
Past experience	0.0195***	(0.0002)	-0.0088***	(0.0011)
Conservative	-	-	5.531e-5	(0.0006)
Store density ( $\alpha^{den}$ )	-	-	-0.0096***	(0.0004)
Travel Index ( $\alpha^{trl}$ )	-	-	0.0100***	(0.0009)
Population	-	-	0.0096***	(0.0014)
Female	-	-	0.0021***	(0.0004)
Age	-	-	-0.0005	(0.0020)
Income	-	-	0.0011	(0.0012)
Education	-	-	-0.0020*	(0.0011)
Category dummy	Yes		Yes	
<b>Consumer side <math>S</math></b>				
Number of deals ( $\gamma_3$ )	-0.0943***	(0.0008)	-0.1157***	(0.0008)
Rivals' number of deals ( $\gamma_4$ )	-0.0228***	(0.0019)	-0.0100*	(0.0061)
Discount	1.5960***	(0.0110)	1.6315***	(0.0137)
Price	-0.0453***	(0.0011)	-0.0663***	(0.0007)
Duration	-0.0184	(0.0109)	0.0831***	(0.0011)
Number of deals $\times$ Travel Index ( $\gamma^{trl}$ )	-	-	0.0765***	(0.0020)
Number of deals $\times$ Store density ( $\gamma^{den}$ )	-	-	-0.0008	(0.0007)
Discount $\times$ Travel Index ( $\beta^{DT}$ )	-	-	0.0009	(0.0015)
Discount $\times$ Store density ( $\beta^{DS}$ )	-	-	-0.0311***	(0.0008)
Conservative	-	-	-0.0079***	(0.0010)
Store density ( $\beta^{den}$ )	-	-	0.1165***	(0.0011)
Travel Index ( $\beta^{trl}$ )	-	-	0.0010	(0.0018)
Population	-	-	0.0108***	(0.0003)
Female	-	-	-0.0004	(0.0008)
Age	-	-	-0.0018*	(0.0011)
Income	-	-	0.0019**	(0.0003)
Education	-	-	0.0003	(0.0008)
Category dummy	Yes		Yes	
$N$	9663		9663	
MLE obj.	25955.6		26287.4	

Note: Standard errors are in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. The estimated coefficients of the time fixed effects are omitted from the table to save space.

Table 5: Per Deal Sales Regression: Robustness Checks

	Baseline		(i)		(ii)	
State transition equation						
<b>Merchant side <math>v</math></b>						
Platform attractiveness accumulation ( $\delta^v$ )	0.6329***	(0.0087)	0.6329***	(0.0085)	0.6329***	(0.0082)
Number of deals ( $\varphi_1$ )	0.0322***	(0.0054)	0.0321***	(0.0052)	0.0322***	(0.0050)
Rival's number of deals ( $\varphi_2$ )	0.0233***	(0.0063)	0.0233***	(0.0062)	0.0233***	(0.0061)
<b>Consumer side <math>u</math></b>						
Platform attractiveness accumulation ( $\delta^u$ )	0.6364***	(0.0057)	0.6363***	(0.0054)	0.6364***	(0.0051)
Transaction volume ( $\phi_1$ )	0.0169***	(0.0028)	0.0169***	(0.0030)	0.0169***	(0.0030)
Rival's transaction volume ( $\phi_2$ )	-0.0052	(0.0043)	-0.0054	(0.0042)	-0.0052	(0.0041)
Measurement equation						
<b>Merchant side <math>N</math></b>						
Deal sales ( $\gamma_1$ )	0.0355***	(0.0034)	0.0352***	(0.0032)	0.0322***	(0.0035)
Transaction volume ( $\gamma_2$ )	0.0873***	(0.0013)	0.0868***	(0.0013)	0.0872***	(0.0015)
Return	0.2182***	(0.0021)	0.2182***	(0.0022)	0.2182***	(0.0023)
Past experience	-0.0088***	(0.0011)	-0.0084***	(0.0010)	-0.0088***	(0.0011)
Conservative	5.531e-5	(0.0006)	-0.0003	(0.0005)	2.768e-6	(0.0005)
Store density ( $\alpha^{den}$ )	-0.0096***	(0.0004)	-0.0090***	(0.0004)	-0.0096***	(0.0003)
Travel measure ( $\alpha^{trl}$ )	0.0100***	(0.0009)	0.0100***	(0.0009)	0.0100***	(0.0010)
Category dummy	Yes		Yes	Yes	Yes	Yes
Local demographics	Yes		Yes	Yes	Yes	Yes
<b>Consumer side <math>S</math></b>						
Number of deals ( $\gamma_3$ )	-0.1157***	(0.0008)	-0.1157***	(0.0012)	-0.1157***	(0.0010)
Rivals' number of deals ( $\gamma_4$ )	-0.0100*	(0.0061)	-0.0100*	(0.0062)	-0.0100*	(0.0060)
Discount	1.6315***	(0.0137)	1.6314***	(0.0142)	1.6314***	(0.0140)
Price	-0.0663***	(0.0007)	-0.0662***	(0.0006)	-0.0663***	(0.0005)
Duration	0.0831***	(0.0011)	0.0830***	(0.0013)	0.0831***	(0.0014)
Number of deals $\times$ Travel measure ( $\gamma^{trl}$ )	0.0765***	(0.0020)	0.0765***	(0.0017)	0.0765***	(0.0019)
Number of deals $\times$ Store density ( $\gamma^{den}$ )	-0.0008	(0.0007)	-0.0005	(0.0008)	-0.0008	(0.0008)
Discount $\times$ Travel measure ( $\beta^{DT}$ )	0.0009	(0.0015)	0.0009	(0.0013)	0.0009	(0.0016)
Discount $\times$ Store density ( $\beta^{DS}$ )	-0.0311***	(0.0008)	-0.0314***	(0.0008)	-0.0310***	(0.0007)
Conservative	-0.0079***	(0.0010)	-0.0075***	(0.0011)	-0.0078***	(0.0011)
Store density ( $\beta^{den}$ )	0.1165***	(0.0011)	0.1164***	(0.0010)	0.1165***	(0.0009)
Travel measure ( $\beta^{trl}$ )	0.0010	(0.0018)	0.0010	(0.0019)	0.0010	(0.0017)
Category dummy	Yes		Yes		Yes	
Local demographics	Yes		Yes		Yes	
$N$	9663		9663		9663	
MLE obj.	26287.4		26260.8		26276.9	

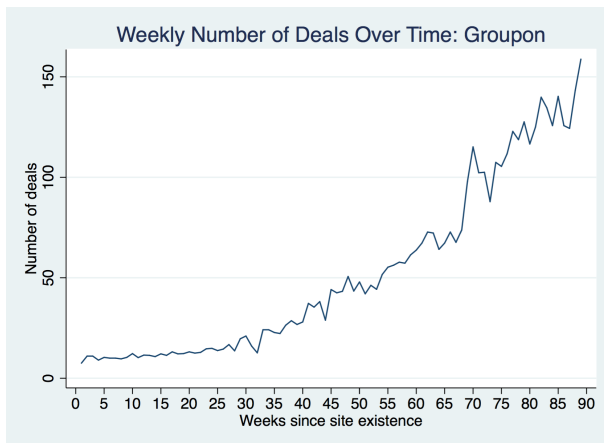
Note : Standard errors are in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. The baseline column uses travel time index as the measure of travel cost as in the main empirical model. Specification (i) and (ii) use hours delay and congestion cost as the measure of travel cost. The estimated coefficients of the time fixed effects and local demographics are omitted from the table to save space.

Table 6: City Growth and Local Characteristics

City	Population (million)	Store Density	Travel Index
Chicago	9.5	3.25	1.31
Dallas	6.4	2.88	1.27
Washington D.C.	5.61	3.53	1.34
Seattle	3.45	3.63	1.38
Pheonix	4.21	2.0	1.27
New York City	18.9	<b>15.0</b>	1.34
Los Angeles	12.8	<b>8.7</b>	1.43
Philadelphia	5.97	<b>5.67</b>	1.24
Houston	5.98	2.96	1.33
Atlanta	5.29	3.45	<b>1.24</b>
Denver	<b>2.56</b>	2.45	1.30
Las Vegas	1.95	<b>4.91</b>	<b>1.26</b>
Boston	4.56	<b>4.07</b>	<b>1.24</b>
Baltimore	2.71	<b>6.62</b>	<b>1.26</b>
Miami	5.58	<b>8.25</b>	<b>1.29</b>

Figure 1: The Growth Trend of Groupon in Top Cities

(a) Number of deals



(b) Transaction volume

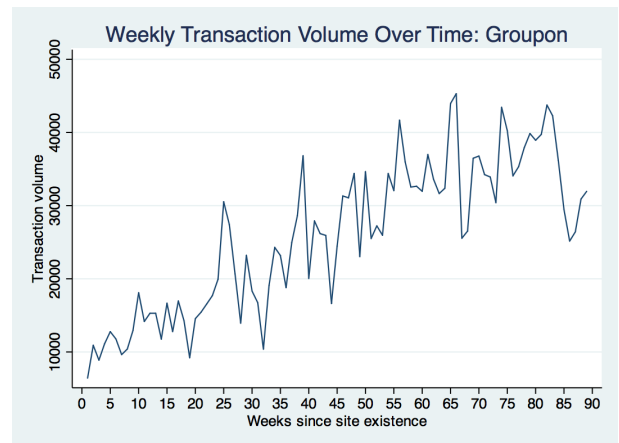


Figure 2: Number of Deals by Population

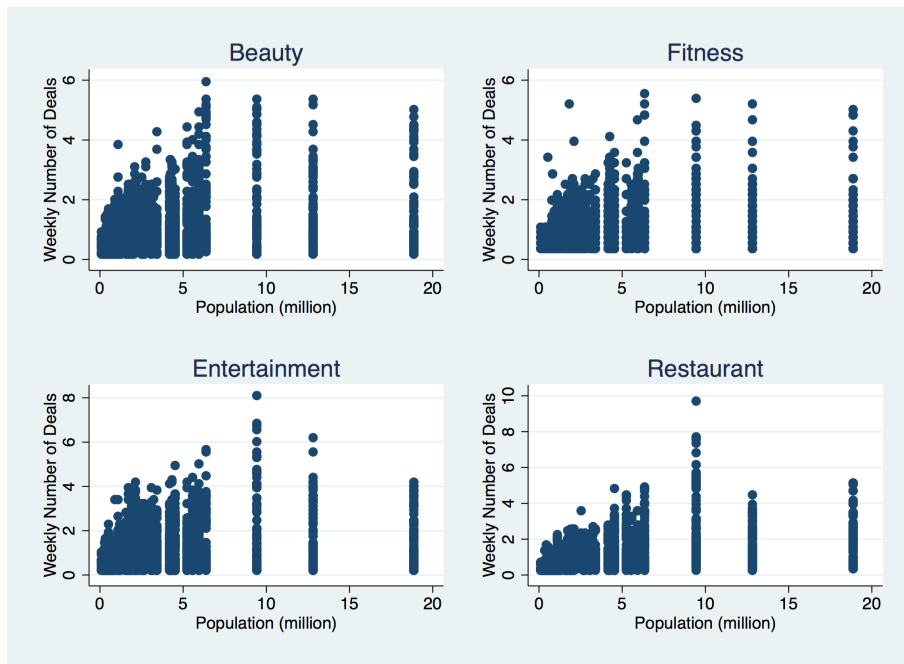
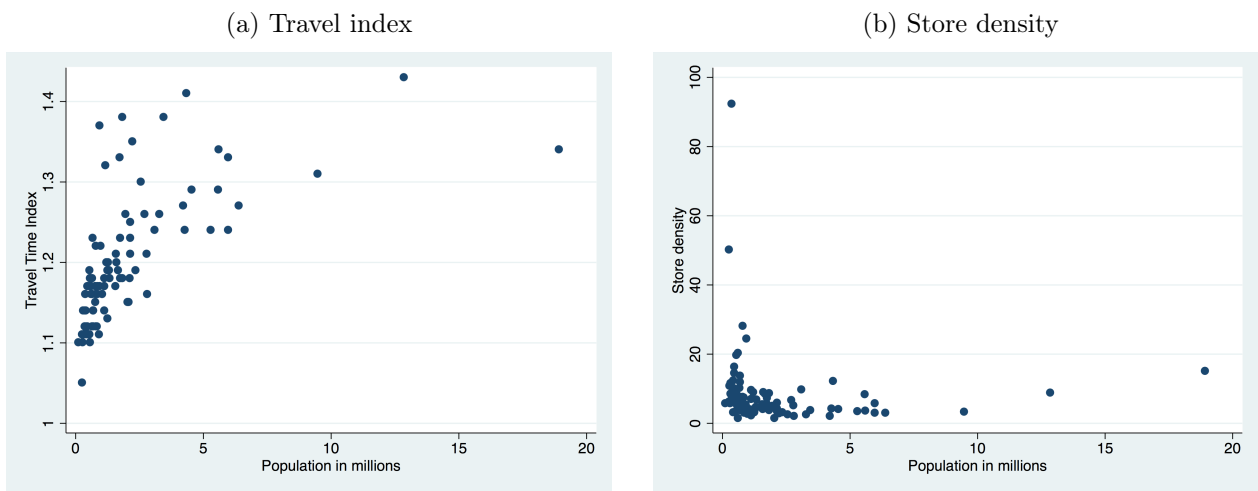
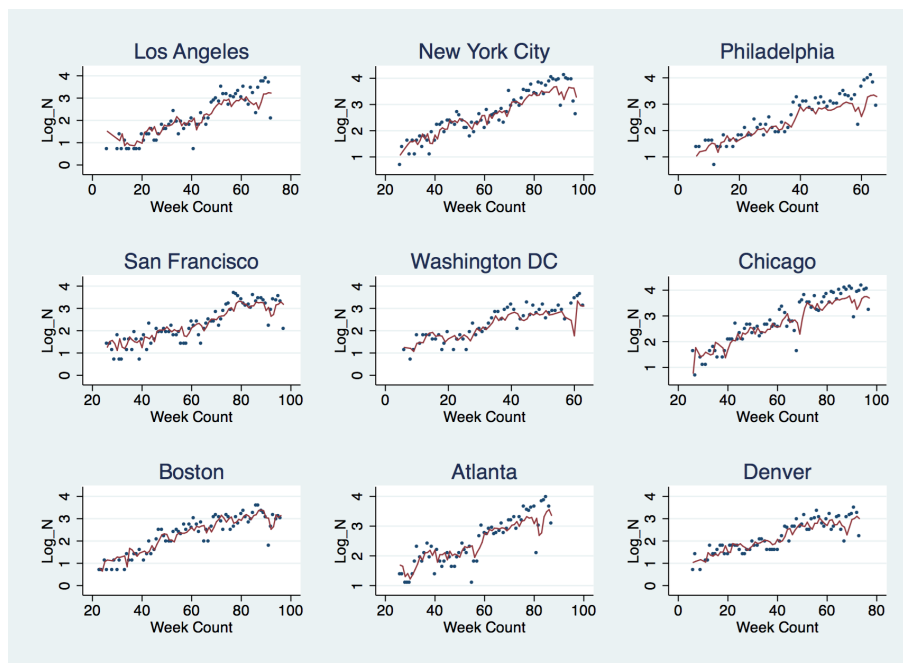


Figure 3: Local Characteristics

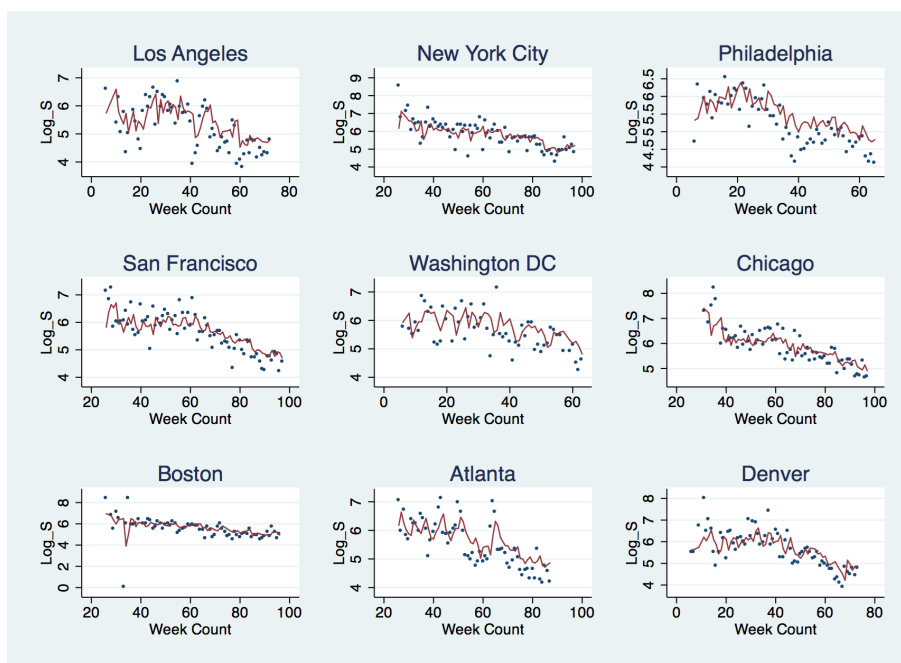


Note : Each point represents an MSA.

Figure 4: Model Fit



(a) Number of Deals



(b) Per Deal Sales



Figure 5: Comparative Statics

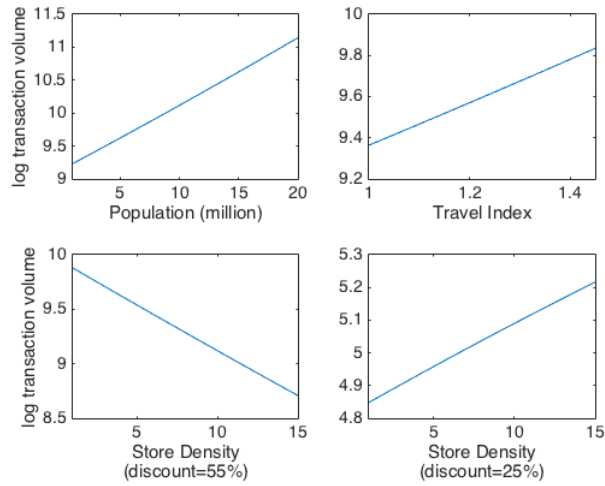
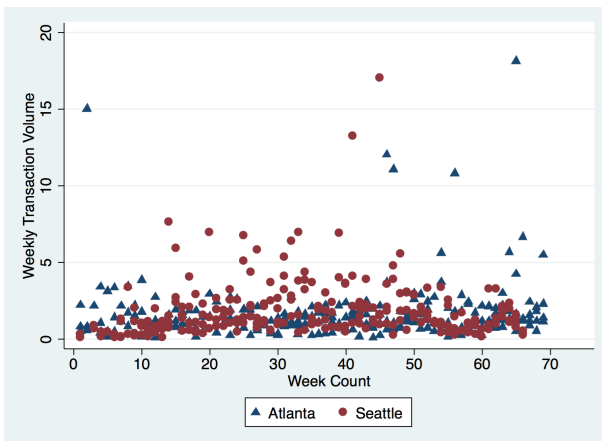
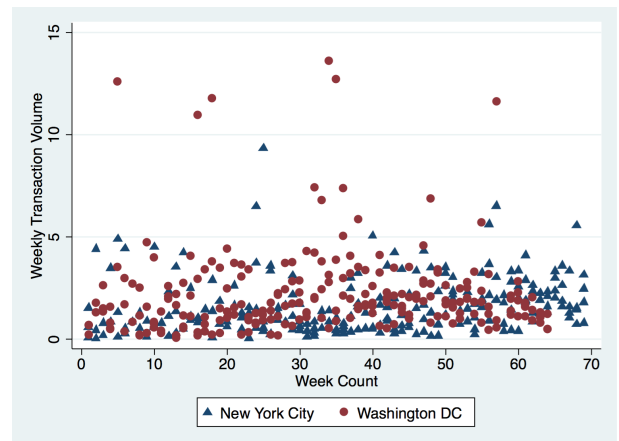


Figure 6: Weekly Transaction Volume in Selected Cities Over Time

(a) Atlanta and Seattle



(b) New York City and Washington DC



*Note:* Each point represents a category-level weekly transaction volume observation. As the average transaction volume differs by category, we normalize the values using the average transaction volume in each category so that we can focus on the overall growth pattern.

Figure 7: Selected City Growth

Y-axis: log transaction volume.

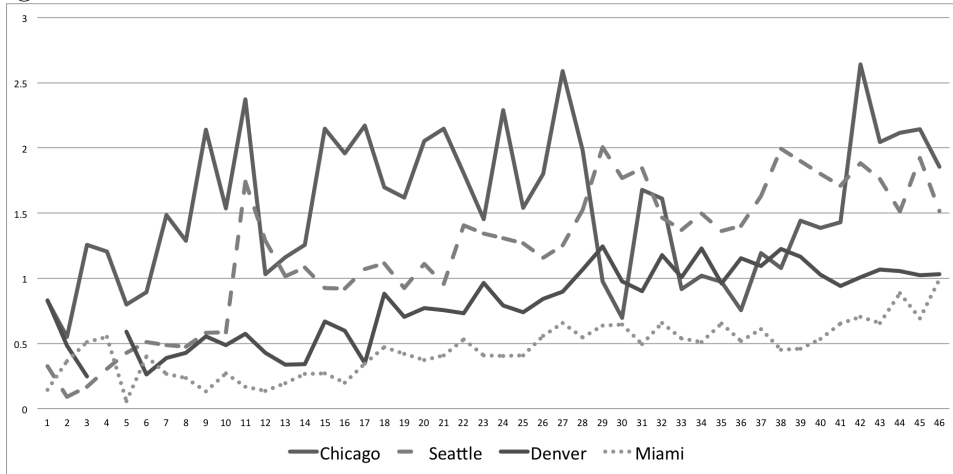


Figure 8: Comparative Statics: Consumer Side and Merchant Side

