Higher Market Thickness Reduces Matching Rate in Online Platforms: Evidence from a Quasi-Experiment

Jun Li
Ross School of Business, University of Michigan, junwli@umich.edu

Serguei Netessine
The Wharton School, University of Pennsylvania, netessin@wharton.upenn.edu

Market thickness is a key parameter that can make or break a platform's business model. Thicker markets can offer more opportunities for participants to meet and higher chances that a potential match exists. However, they can also be vulnerable to potential search frictions. In this paper, using data from an online peer-to-peer holiday property-rental platform, we aim to identify and measure the causal impact of market thickness on matching rates. In particular, we exploit an exogenous shock to market size caused by a one-time migration of listings from other platforms, which gives rise to a quasi-experimental design. We find that increased market thickness actually leads to lower matching rates. Keeping search technology and other factors constant, doubling market size leads to a 15.4% reduction in traveler confirmation rate and 15.9% reduction in host occupancy rate. As a result, the platform lost 5.6% of potential matches because of the increased market size. We attribute the effect to increased search friction: travelers' search intensity increases by 18.3% when market size doubles. This effect is especially prominent when the matching needs to take place within a limited time. Our results offer insights for future empirical and theoretical research on matching markets. They also highlight the importance for platform owners to watch out for increased search frictions as markets grow and to invest in search technologies to facilitate more efficient search.

Key words: Matching markets, market thickness, matching supply and demand, quasi-experiment, difference-in-differences

1. Introduction

Do thicker markets have higher matching rates? Intuitively, they should. Size is critical to the survival and success of many matching markets. A larger number of participants on each side offers more opportunities for participants to meet and better chances that a match exists. In the recent rise of peer-to-peer marketplaces, platform owners have expended significant effort fueling growth. For instance, ride-sharing platforms, such as Uber, Lyft and DiDi, spend billions of dollars a year on marketing campaigns and incentives for drivers and passengers,[1] while other marketplaces, such as the holiday rental platform HomeAway, expand through aggressive acquisitions.[2]

A key difference between these peer-to-peer marketplaces and many traditional two-sided markets (such as the markets for credit cards, newspapers, and mobile phone service) is the existence of

[1] "Inside Uber's Battle with Its Chinese Rival: The Only Option Is to Choose a Side." *Business Insider.* August 6, 2011. "Lyft Is Spending Millions on an Ambitious Expansion Plan in the U.S." *Skift.* April 14, 2016.

[2] "The Good, the Bad and the Tragic: Stories of Acquisition for Growth." *Forbes.* March 17, 2013.

a high degree of heterogeneity—among market participants, and among the goods or services they offer or desire. In such markets, search is essential to finding the right match. However, a market with heterogeneity is vulnerable to various forms of search frictions (Arnosti et al. 2016, Kanoria and Saban 2017, Cullen and Farronato 2016, Fradkin 2017, Horton 2017, Bimpikis et al. 2017). Heterogeneity increases as the market grows, as does search friction. The presence of abundant options makes it more likely a potential match exists, but it can be more difficult to find.

More often than not, a need for matching comes with a time constraint. Some matching markets such as most labor markets, marriage markets, and markets for organ transplants, have implicit deadlines. Others have more explicit deadlines. For example, a traveler needs to find an accommodation before the intended day of travel. A homeowner needs to find a cleaning service before holding a party. A student needs to find a seat before the school year starts. With limited time to complete a match, the impact of search frictions becomes even more prominent. It is therefore unclear whether thicker markets will necessarily lead to higher matching probabilities, especially in matching markets with heterogeneity and time constraints.

Existing theoretical literature typically assumes matching markets exhibit increasing returns to scale or constant returns to scale (e.g., Diamond 1982, Pissarides 1990). Empirically, the majority of studies, almost all on labor markets, find evidence consistent with constant returns to scale. However, Petrongolo and Pissarides (2001, p.409) refers to these findings as "tentative," as most studies rely on aggregate time-series or cross-sectional observations, and "no study conducts a careful test of increasing returns to scale by testing, for example, whether the matching rate improves when the total number of participants increases."

This paper does exactly that. It studies the effect of market thickness on matching rates, exploiting an exogenous shock to market size. The fact that in our data some markets are impacted while others are not gives rise to a quasi-experimental design that allows us to make causal inferences. In particular, we use data from a leading peer-to-peer online holiday rental marketplace operating in Australia, with over 40,000 listings located in 3,200 areas and 77 regions. These data offer several advantages for studying the research question at hand. They cover a large number of geographically separated markets, which allows us to associate market characteristics, especially market thickness, with matching outcomes. They include information on inquiries and bookings for individual travelers, which allows us to study the property of the matching function at a micro level. Last and most important, the company's business strategy offers us a unique opportunity to exploit an exogenous shock to market size: migrations of listings from two previously acquired platforms expanded some of the markets significantly over a very short period.

The context also offers several appealing features that ensure the validity of our identification strategy: (1) all new listings migrated without selections either by the platform manager or by listing owners; (2) the timing of the event was idiosyncratic in nature: it was neither publicly announced nor anticipated by market participants; and (3) careful examinations of both control and treatment groups demonstrate parallel trends before the event took place.

Unlike previous literature, we do not find evidence of increasing or constant returns to scale. Rather, we find evidence of diminishing returns to scale. In the marketplace under study, further increases in the number of market participants lead to lower matching rates. Specifically, keeping search technology and other factors constant, when market size doubles, traveler confirmation rate goes down by 15.4% and host occupancy rate goes down by 15.9%. In total, the platform lost 5.6% of potential matches on each day due to the increased market size.

What drives the difference between our findings and those in prior research? Is it caused by differences in context, data disaggregation, or identification strategy? To investigate this question, we conduct additional analyses at the market level exploiting cross-sectional variations alone, as in previous studies, but not the exogenous shock. Without the use of the exogenous shock, we find constant returns to scale as previous studies do. But when using the exogenous shock as an instrument for market size, we again find decreasing returns to scale at the market level. In other words, the difference in our results is most likely driven by a better identification strategy, rather than a unique setting or simply that we use using disaggregated data.

The decrease in matching rates can be attributed to increasing search frictions in markets with more participants. We find that when market size doubles, the number of inquiries sent by travelers increases by 18.3% and the number of inquiries received by hosts increases by 19.6% keeping search technology and other factors constant. The travelers also end up paying higher transaction prices, and hosts end up receiving lower revenues. The impact of search friction becomes more severe when the deadline for matching is more imminent. When there is plenty of time left, an increase in search intensity does not affect matching probability and transaction prices. However, when the deadline is close in time, increasing search intensity leads to lower matching rates and higher transaction prices.

Our findings offer several implications for future research in matching. They emphasize the need for more careful empirical research to study the causal effect of market thickness on matching rates as well as to carefully evaluate of the properties of the matching functions used in modeling. They also highlight the need for theorizing the sources of market frictions, especially how deadlines affect matching rates.

Our findings also offer practical implications for rising peer-to-peer marketplaces. While many marketplace operators attempt to expand their platforms through aggressive marketing campaigns or acquisitions, it may not be an effective strategy if not accompanied by better search functionality. As the market grows, search frictions also become more prominent. Those once-thick markets could become thin because of such frictions. It is therefore crucial for platform operators to invest in technologies such as interactive decision aids, search engines and recommendation systems while expanding in size. Sometimes less obvious choices such as limiting the number of options may achieve better outcomes. Moreover, it is especially important to facilitate efficient search processes when the need for matching is time constrained.

2. Related Literature

Our paper contributes to the growing theoretical and empirical literature that studies peer-to-peer marketplaces, as well as the matching-market literature in general. In particular, our work is related to four streams of literature: (1) marketplace innovations; (2) effects of thickness in matching markets; (3) search frictions in online matching markets; (4) matching supply and demand in perishable-goods markets.

Marketplace Innovations. There is a burgeoning literature on innovative marketplaces fueled by increasing use of information technology and growing controls over the design, implementation, and operations of markets to better match supply and demand. Allon et al. (2012) shows how different levels of involvement by the moderating firm can affect market outcomes. Several papers study how information affects participants' behaviors and market outcomes (Moreno and Terwiesch 2014, Parker et al. 2016), while others study how price is determined or its effect on market outcomes (Cachon et al. 2017, Li et al. 2016, Kabra et al. 2017).

Effects of Thickness in Matching Markets. In labor markets, there is a long-standing discussion on whether matching markets are governed by increasing, constant, or decreasing returns to scale. Early theoretical work typically assumed increasing or constant returns to scale (e.g., Diamond 1982, Pissarides 1990). On the empirical side, with only a few exceptions, most evidence converges on finding constant returns to scale (see Petrongolo and Pissarides 2001 and references therein). Note that this line of empirical research predominantly focuses on labor markets because of data availability. Most studies are conducted using aggregate data on vacancies and unemployment. More importantly, the empirical evidence is largely drawn from observational variations, either times-series or cross-sectional, lacking clear identification strategies.

In contrast, we focus on a different context: online peer-to-peer marketplaces, which are often characterized by higher degrees of heterogeneity and need for more imminent matches. Moreover, we use individual-level search and matching data to control for many potential confounding factors that may also affect matching rates. Finally, we attempt to estimate the causal effect of market size, which to our knowledge has not been done before in this literature.

Relevant to our work, Cullen and Farronato (2016) uses individual-level data from an online peer-to-peer freelance marketplace to study the growth of online matching markets. It documents evidence consistent with constant returns to scale, but does not study this effect causally. Kabra et al. (2017) shows evidence of increasing returns to scale using data from a ride-hailing marketplace, which exhibits less heterogeneity than the holiday rental marketplace we study. Lastly, using data from an online business-to-business auction platform, Bimpikis et al. (2017) finds that consolidating the ending times of auctions increases the platform revenue by 6.5%, primarily because of bidder-participation friction.

Search Frictions in Online Matching Markets. This paper is closely related to theoretical and empirical work that studies search frictions in online matching markets. Compared to traditional marketplaces, online marketplaces such as eBay, Uber, and Airbnb substantially reduce costs associated with acquiring information. As a result, they present previously unavailable opportunities for buyers and sellers to meet, negotiate and trade. Bakos (1997) shows that reduction in buyers' search costs can avoid substantial allocational inefficiencies in differentiated markets. Brynjolfsson and Smith (2000) demonstrates empirically that online markets have lower frictions on many dimensions compared to conventional stores. Parker et al. (2016) demonstrates text-message services lead to higher efficiency in India's agricultural markets. Kroft and Pope (2014) finds that Craigslist, a platform that allows individual users to advertise jobs, house rentals, and items for sale, significantly lowered the apartment- and house-rental vacancy rate and the number of classified job advertisements in newspapers. Similarly, by allowing consumers to buy and sell used products, Craigslist has led to a 2 to 6% annual reduction in municipal solid waste in markets it entered (Dhanorkar 2017).

Even so, online marketplaces are not frictionless. Various forms of frictions still exist. First, there is still some amount of information asymmetry. Arnosti et al. (2016) finds that lack of job applicant availability information can lead to inefficient equilibria due to wasteful search in equilibrium. Using data from Airbnb, Fradkin (2017) finds that 15% of the initial contacts sent by guests were rejected by hosts due to stale vacancy. Using data from an online freelancing platform, Horton (2017) documents evidence that employers often cannot discern the availability of prospective workers, which causes their invitations to be rejected and reduces the job fill rate. Dhanorkar et al. (2015) finds that buyers' uncertainty about the products and sellers' commitments to transactions reduce the success rate of online exchange of material and waste.

Second, when information is abundant, individual users sometimes tend to engage in many searches and evaluations. This is not only costly, it can also result in users avoiding making decisions altogether because of information overload. As Hagiu and Jullien (2011) shows, platforms have incentives to induce consumers to search more than they would like. However, such behavior may have unintended consequences. Ghose et al. (2014) shows empirically that providing more information during the decision-making process may lead to fewer consumer purchases because of information overload. Using data from a hotel search platform, Koulayev (2014) estimates that it costs ten to thirty dollars to search a page of results, and attributes it to high cognitive costs of comparing various characteristics of newly found hotels.

Third, individuals may exert negative externalities on others and on the entire platform, which in turn leads to inferior matching outcomes. Romanyuk (2017) describes a cream-skimming externality: when sellers have limited capacity, a seller may reject a low-value buyer and strategically wait for a high-value buyer. As a result, he remains available and attracts a fraction of subsequent buyers who otherwise would contact other sellers, resulting in lower total surplus. Using data from eBay, Nosko and Tadelis (2015) empirically demonstrates a form of reputational externality, where one disappointing transaction may cause a buyer to update his beliefs about the quality of all sellers on the platform.

Platforms therefore have a significant role to play in designing marketplaces with fewer frictions and higher efficiencies. First, platforms could offer interactive decision aids such as filters, sorting mechanisms, and comparison tools to create customized search experience that better caters to individual preferences. Haubl and Trifts (2000) shows with lab experiments that such decision aids reduce consumers' search effort and increase decision quality.

Second, platforms could design better and more targeted recommendation systems to help users discover a potential satisfactory match more quickly. Fleder and Hosanagar (2009) shows that recommendation systems that explicitly promote diversity will enable better consumer–product matches. Ghose et al. (2012) develops a ranking system that ranks products based on their average surplus. De los Santos and Koulayev (2017) makes use of information available at the time of the consumer request and develops a customized ranking system based on its assessment of the consumer's preferences and the expected consumer type. In contexts in which consumers themselves are uncertain about their preferences and learn as they search, Dzyabura and Hauser (2017) shows that recommendations are more effective if they encourage consumers to search undervalued products and products with diverse attributes. Nosko and Tadelis (2015) demonstrates using a large-scale field experiment that prioritizing higher-quality sellers can reduce reputational externality and lead to higher buyer retention. Using eBay data, Dinerstein et al. (2018) find that a platform redesign promoting more relevant offers to consumers and prioritizing low price offers can help strengthen seller incentives to lower prices and hence reduce price-related search frictions.

Last, platforms could develop feedback and reputation systems to reduce frictions caused by information asymmetry (Tadelis 2016). Using data from an online intermediary for software development services, Moreno and Terwiesch (2014) shows participants are very responsive to the numerical reputation score and also to the unstructured reputational information. Using multiple field experiments conducted on Airbnb, Cui et al. (2016) demonstrates that reputation systems can effectively reduce matching inefficiencies caused by racial discrimination.

It is worth noting that sometimes restricting the amount of information disclosed or limiting the number of choices offered or actions available can actually lead to higher search efficiency and better matching outcomes. Ghose et al. (2017) shows that a carefully curated digest of social-media content during the early stages of consumer search can lead to a substantial increase in the conversion rate. Romanyuk (2017) find that partial information disclosure can reduce the cream-skimming externality and lead to higher total surplus than full information disclosure. Kuksov and Villas-Boas (2010) and Arnosti et al. (2016) both show that restricting the number of choices may actually lead to better outcomes. Lastly, Kanoria and Saban (2017) shows that to mitigate search frictions, a platform should force the more selective side of the market to reach out first, while disallowing the less selective side from doing so.

Matching Supply and Demand in Perishable-Goods Markets. Perishability of the good or service and limited capacity play a vital role in moderating the success of a market. A large body of literature in revenue management and dynamic pricing is devoted to better matching supply with demand in perishable markets (see Talluri and van Ryzin 2005 for an overview). As the deadline approaches and inventory depletes, prices evolve dynamically, and which consumers get served also changes. Recent research by Li et al. (2014), Lederman et al. (2014), and Tereyagoglu et al. (2017) studies empirically how such perishability affects consumer demand and how firms can better manage their depleting inventories in the airline, hotel, and theater industries, respectively. A key difference, however, between a peer-to-peer marketplace and a traditional market is that decisions are decentralized in the former. Overall, our study highlights the need for theory to better account for the effects of deadlines in matching markets.

3. Background and Data

This section describes the business model of the online peer-to-peer marketplace under study, and in particular, how matching works in this market. We then discuss the details of the data and features instrumental to our study.

3.1. Background

We receive data directly from an online peer-to-peer platform for holiday rentals in Australia. The company granted us data access in 2015. Most of the regions it operates in are tourist destinations. For example, the Gold Coast, Melbourne, Byron Bay, Sydney, Brisbane, Great Ocean Road, and the South Coast are among its top destinations. A property listed on the network is usually an entire apartment or house, rather than a room within it (by contrast, many Airbnb listings are of single rooms or partial units). In other words, a traveler will typically have the property all to himself during the contracted rental period, rather than sharing it with the host. A typical traveler (or a traveling group) stays for four to five nights. Over a third travellers travel with children. Because of the nature of these trips, the property is not just a place to stay. It is usually considered by travelers to be an important component of the holiday itself. For such reasons, travelers tend to plan early and spend a significant amount of time communicating with prospective hosts before making decisions.

The matching process works as follows. Property owners or managers join the platform by paying an annual fee, which enables them to list their properties as rentals. Similar to other online peer-to-peer rental networks such as Airbnb, listings typically include photos of the property, descriptions, and information on availability and prices. An interested traveler will send an inquiry to the owner of a listing after browsing its webpage. The inquiry is a built-in function of the platform, which automatically sends a message to the owner/manager's account and alerts her. Such an inquiry usually includes information about the trip, the traveler and questions regarding the listing. Through these inquiries, travelers attempt to confirm availability, ensure that pricing and property features are as advertised, discuss preferred payment method, and acquire any detailed information unavailable on the webpage. The platform does not offer instant booking as Airbnb does.

The owner/manager receiving the inquiry can then decide whether to reply and what to reply. She may answer the traveler's questions, and meanwhile ask for more information about the trip and the traveler himself. Due to the nature of the trips, owners are typically concerned about whether prospective guests will keep good care of the property during their stay and whether they will be able to pay on time. Therefore, in the communication, they attempt to acquire information regarding the purpose of the trip (e.g., bachelor party, family getaway, or business travel) and the background of the traveler or the traveling group (e.g., high school teenagers, college students, newly wedded couples, retired seniors). Sometimes, the owner will also include her email address or phone number to continue conversation in a more personal and timely manner. Eventually, based on the correspondence between the traveler and the owner, the two parties will determine whether to confirm the transaction. Once confirmed, we consider the traveler and the owner to be matched.[3] The two matched parties can arrange payment offline or online (using the platform's payment system). In sum, the matching process facilitated by the platform is more formal and curated than Craigslist but not as instantaneous as Airbnb.

[3] As in most travel reservations, there is a chance that the parties involved may cancel the reservation. Unfortunately, we do not have access to the cancellation data. The focus of the analysis is on the initial matching rate, and most often, cancellations are caused by unforeseen reasons after the initial matching has taken place.

3.2. Data

We acquire data on listing features, traveler inquiries, traveler bookings and owner bookings for the calendar year 2014. The listing data include all 150,000 listings that ever registered on the platform since its start in 2001, among which about 48,000 listings were still live by the end of 2014. The dataset provides information regarding the initial listed date, the location of the property (region, area, and geocode), number of bedrooms, maximum number of adults, and about 40 features associated with the property. These features are selected by the host. They include property types (e.g., house, apartment, cottage), what trip purposes the property is best for (e.g., corporate conference, family holiday, romantic getaway), nearby attractions (e.g., beach, national parks), amenities (e.g., spa, wheelchair access), guest accommodation (e.g., child-friendly, pet friendly), and others features such as special offers.

The traveler-inquiry and traveler-booking data include 3.91 million inquiries made throughout 2014. For each inquiry, we observe the date when it is sent, the listing it corresponds to, the destination of the trip, the check-in and check-out dates, the number of people (adults and children) on the trip, the daily price quoted, and whether it is confirmed. Among the 3.91 million inquiries, 350,000 (about 9%) were eventually confirmed.

Owners may also block out certain dates because of personal needs or for other reasons, and they do this dynamically. Once a date is blocked out by the owner, it will appear as unavailable to travelers on the calendar. Following the convention of the company, we refer to these block-outs as "owner bookings." The owner-booking data include 88,000 owner bookings in 2014. For each owner booking, we observe when and for which travel dates the booking is made.

Data from this platform offer several advantages in studying the impact of market thickness on search intensity and market rates. First, the data cover a large number of markets relatively separate from each other geographically, which allows us to associate market characteristics, and market size in particular, with matching outcomes. In some other matching markets, such as eBay, such geographical separation is not possible. Second, in addition to the final matching outcome, we also obtain inquiry histories for all travelers in all markets, which allows us to measure travelers' search efforts and examine how market thickness affects search intensity and in turn affects matching outcome. Fradkin (2017) is another study that uses guest-inquiry data, but only for one city. Last, and perhaps the most important, the acquisition and later the migration of listings provide us a unique opportunity to exploit an exogenous shock to market thickness, as we will elaborate in the next section.

The data are not without shortcomings. Unfortunately, because of business constraints, the company was unable to provide the web search history of each traveler before he communicated with owners. We therefore only observe a traveler if he sends an inquiry to a prospective host. We will address this data-censoring issue in the empirical analysis in the next section.

4. Empirical Framework

In this section, we introduce the exogenous shock we exploit to identify the causal impact of market thickness on matching rates. We then describe the empirical specification for travelers and hosts. Lastly, we discuss the validity of this empirical framework.

4.1. Shock to Market Thickness

To identify the effect of market thickness on matching rates, we note that how thick a market will be is often determined endogenously. In two-sided markets, the number of participants on each side is affected by levels of demand and supply, demand-side and supply-side network externalities, and cross-group externality (e.g., Rochet and Tirole 2003, Parker and Alstyne 2005, Armstrong 2006). Therefore, a direct cross-sectional comparison of thin and thick markets can be subject to many unobserved confounding factors that affect both market thickness and matching rates simultaneously. For example, higher demand levels or degrees of cross-group externality will lead to thicker markets (i.e., more participants on both sides). They also lead to higher matching probabilities since participants have a stronger intention to find a match. Such factors, if present but not fully observed, can lead to a positive correlation and an upward bias in estimating the causal effect of market thickness on matching rates.

In this paper, we exploit an exogenous shock to market thickness to estimate its causal effect. In March 2011, the company acquired two competing peer-to-peer vacation rental platforms. After the acquisition, both acquired platforms operated under the original domain and brand names for more than three years. In August 2014, the company migrated listings from the two acquired platforms and consolidated them with existing listings on its original platform. For duplicate listings, the original listing on the platform was kept. Nonduplicate listings were added as new listings to the platform. The consolidation led to a sharp increase in the number of listings available on the platform, as shown in Figure 1. Although about 30 new listings are added to the platform on an average day, a total of about 6,000 new listings were added during the four days when the consolidation happened, August 19–22, 2014. After the consolidation, traffic also started to migrate to the consolidated platform. Within a month, all traffic to the two acquired platforms was automatically re-directed to the consolidated platform.

How did the consolidation affect the platform's market thickness? From the supply side, the estimated number of live listings on the platform just before the consolidation was around 42,000 (see Appendix A for details). After the consolidation, an additional 6,000 unique listings were added, representing a 14% increase. From the demand side, daily site traffic (i.e., number of unique users per day) rose about 13% to 15% based on traffic data acquired from Alexa, as shown in Figure 2. That is, the magnitude of the increase in demand is on par with the increase in supply.

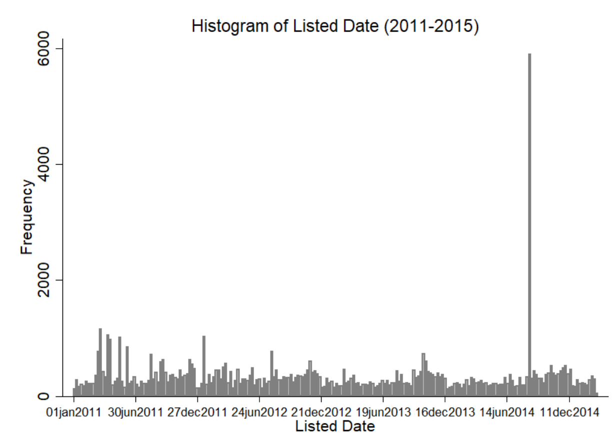

Figure 1 **Histogram of Listed Dates**

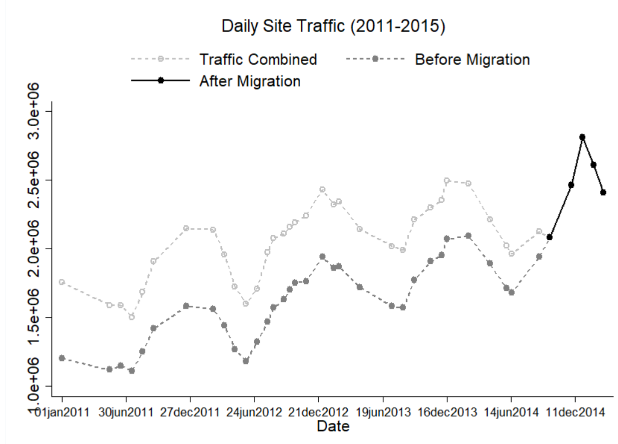

Figure 2 **Daily Site Traffic**

The shock therefore increased the thickness of the market, while keeping the supply–demand ratio largely unchanged. It therefore provides us an opportunity to estimate the causal impact of market thickness on matching rates.

4.2. Difference-in-Differences Approach

Although the migration of listings provides us with an exogenous shock to market thickness, direct comparisons of market performances between the before and after periods can result in biased estimates because they can be confounded with time trends that are present even in the absence of the shock. For instance, Figure 2 shows that Australia's holiday rental market exhibits a clear seasonality. Demand is the highest during the summer months (December to February), which also include the Christmas and the New Year holiday seasons, and the lowest during the winter months (June to August). Because of this seasonality, the demand is expected to increase starting from September, as more travelers book summer and holiday travels, even without the consolidation of listings. Besides seasonality, we also note that peer-to-peer holiday rental markets have in general enjoyed a steady growth in recent years, also shown in Figure 2, which can be another confounding factor to the effect in which we are interested.

To isolate the confounding effects, our empirical strategy relies on identifying markets that did not experience the shock as the control group. These control markets experienced the same time trend, though they did not experience an increase in thickness of the market. We then compare the changes before and after for both the treatment and control groups.

In particular, we notice that not all markets experienced an expansion in size after the consolidation. For those markets that experienced an expansion in size, the magnitude was not uniform across markets; it varied substantially. We now describe how we define whether a market experienced a discontinuous increase in thickness because of the consolidation. We do not necessarily want to characterize the addition of any new listing as an indicator of an increase in thickness.

Even without the consolidation, new hosts and new listings become available on a daily basis, as shown in Figure 1, and the speed at which new listings join varies across markets. Therefore, we define a discontinuous change in market thickness as relative to the normal speed at which new listings become available in that market. Specifically, we compare the average number of new listings added per day during the four-day consolidation period with the number of new listings added daily during the non-consolidation period. If the former is greater than the 90th percentile of the latter, then we consider the consolidation has caused a discontinuous increase in market thickness.[4] We then measure the magnitude of the increase of market thickness as the ratio of the marginal increase of new listings during the consolidation period relative to the number of live listings just before the period (see details in Appendix A). The marginal increase of new listings is defined as the total number of new listings in the consolidation period minus the average number of listings in a typical four-day period.[5]

In sum, the possibility to construct control and treatment groups before and after the exogenous shock to market thickness leads to a quasi-experimental research design. For travelers, we estimate difference-in-differences models of the form

$$y_{irm}^{Traveler} = f(\alpha + \beta_1 Tr_m + \beta_2 A_{ir} + \beta_3 Tr_m \cdot A_{ir} + \gamma X_{ir} + \delta W_m + \epsilon_{jdm}), \tag{1}$$

where the subscript i denotes traveler, r denotes trip, and m denotes market, that is, an area in a region. Variable $y_{irm}^{Traveler}$ denotes the outcome variable, that is, whether a potential traveler is matched or not, number of inquiries sent and price paid. The link function, f, concerns Logit, Poisson and OLS, respectively. Binary variable Tr_m indicates whether a market m experiences a discontinuous increase in market thickness because of the migration of listings. One can also specify Tr_m as a categorical variable with multiple levels stratified by the magnitude of the increase or a continuous variable measured by the percentage of increase, instead of a binary variable. A_{ir} denotes whether traveler i's search for trip r is conducted after the shock (August 23, 2014).[6] Matrix X_{ir} denotes traveler characteristics (e.g., country of origin) and trip characteristics (e.g., length of stay, weekend stay, number of adults and children, advance booking). Matrix W_m denotes market

[4] The results are consistent if we use other cutoffs such as the 75th percentile or 95th percentile. To avoid treating markets with trivial numbers of new listings in the four-day period as experiencing a discontinuous increase in thickness, we require that the market has to have a minimum of three new listings amounting to at least 1% increase relative to the number of existing live listings at the time in the market. The results are consistent if we do not impose this additional requirement though.

[5] Alternatively, one could measure the increase in thickness in terms of listing-days. It has a very high correlation (0.9952) with that calculated using number of listings. As a result, using the number of listing-days leads to almost identical results.

[6] To ensure clean definition of "before" and "after" periods, searches that started during the "before" period but ended during the "after" period are excluded from the sample.

characteristics m (e.g., urban vs. rural, business- or family-oriented). Error term ϵ_{irm} denotes the idiosyncratic shock to traveler i interested in trip r in destination m.

There are cases in which a traveler is flexible in his travel destination. As a result, we observe inquiries sent to listings in multiple destinations for the same trip. In such cases, we define treatment Tr_m and market characteristics W_m to be those of the primary market the traveler is interested in, where the primary market is defined as the first inquired market for a trip. We choose the first market a traveler inquires as his primary market because this market is likely to be given exogenously. Whether to search other markets can be influenced by the traveler's experience in the first market he inquires.[7] To ensure robustness of the results, we vary the definition of the primary market to be the most frequently inquired market for a trip. Alternatively, instead of associating Tr_m and W_m with a single primary market of our definition, we define them as the average across all markets inquired. As we will show in the results section, the results are robust to these alternative specifications.

As we mentioned previously, our data are censored since we do not observe when and where travelers intend to travel if they do not send any inquiry. Appendix B provides the details of how we correct for data censoring when estimating the effect of market thickness.

For listings, we estimate difference-in-differences models of a similar form:

$$y_{jdm}^{Listing} = f(\alpha + \beta_1 Tr_m + \beta_2 A_{jd} + \beta_3 Tr_m \cdot A_{jd} + \gamma D_d + \xi_j + \epsilon_{irm}), \tag{2}$$

where the subscript j denotes listing, d denotes day (or week), and m denotes market. Variable $y_{jdm}^{Listing}$ denotes the outcome variable, that is, whether a listing is occupied on specific stay date, and a listing's average daily revenue over a week (log-scaled). We take the average to smooth the observed value because many daily revenues are zero. The link function, f, is Logit and OLS, respectively. We define A_{jd} to equal one if the calendar date d (or first day of the calendar week) is after August 23, 2014, and zero otherwise. For a listing date, the definition of "after" can be tricky because even if a stay date is after August 23, 2014, it can be booked before August 23, 2014. Technically, these listing dates are not affected by the shock. Therefore, in our analysis, we exclude those listing-date combinations traveled after August 23, 2014, but booked before August 23, 2014, in both control and treatment markets. Note that excluding these listing-date combinations will lead to unbalanced samples for the treatment and non-treatment periods—while the non-treatment period include all listing-date combinations, the treatment period will be over-represented by listing-dates that failed to attract early bookings. Therefore, to allow for fair comparisons, we also exclude listing-dates during the non-treatment period that were booked

[7] For example, with more options available in the primary market of interest, a traveler may be less likely to search other markets. We indeed find evidence consistent with this conjecture.

before the start of the non-treatment period, i.e., June 18, 2014. Matrix D_d denotes stay-date characteristics (i.e., day of week, holiday dummies). Error term ξ_j denotes listing fixed effect, and ϵ_{jdm} denotes the idiosyncratic shock to listing j on day (or in week) d in market m.

4.3. Validity of Difference-in-Differences Approach

In applying the difference-in-differences framework to the data, it is important to carefully consider the validity of the quasi experiment created by the shock: migration of listings. Ideally, the treatment would be randomly assigned to markets. However, which market will experience an increase in thickness after the migration may not be completely random. In what follows, we investigate whether it can be endogenously determined or correlated with other factors that may affect matching rates.

First, all listings were migrated without filtering or selection by the company's management team. The decision to acquire these websites, which happened more than three years earlier, was driven by the corporate strategy to grow and expand in general, rather than to target a specific niche segment. Moreover, listing owners themselves did not select whether to migrate. Therefore, the typical selection bias that could threaten the validity of a quasi experiment is not likely to be present in our setting. However, it is still possible that even though listing owners do not actively select whether to be migrated, they may passively choose to not engage with the new platform—because of unfamiliarity, for instance. We consider this possibility carefully in our analyses and confirm that our results are consistent after accounting for it.

Second, even though there is no strategic selection from either the management or user side, new listings are, after all, not randomly assigned to markets. Rather, markets with more presence from the two previously acquired platforms will experience greater increase in thickness, whereas others will experience less increase or none at all. If these are different types of markets, then the comparison between the control and treatment groups may be threatened. Luckily, the validity of the difference-in-differences approach hinges upon the existence of the parallel trend. It does not require the control and treatment to be a priori the same in all aspects. It only requires that the two groups follow the same trend in the absence of the treatment, while the levels may differ because of differences in characteristics. This assumption can be investigated empirically using the data from the before period. In the subsequent section, we demonstrate that the parallel-trend assumption is indeed satisfied, which supports the validity of our empirical method.

Besides these considerations, the shock we exploit offers two other virtues. First, the change is sharp. Therefore, we expect a sharp and discontinuous change in the outcome variables, which can be visually examined to ensure that the effect is noticeable. Second, the timing of the consolidation was idiosyncratic, as it mostly depended on the capability of the IT team and was neither part of the original acquisition plan nor pre-announced publicly. Therefore, it is not likely that it was anticipated by listing owners or by travelers.

5. Descriptive Results

In this section, we present the treatment and control groups, discuss how trip, traveler, and market characteristics compare between treatment and control groups, and examine the validity of the difference-in-differences framework.

5.1. Treatment and Control Groups

We define a market as an area within a region (recall there are 3,200 areas and 77 regions in total). However, some of these areas are sparsely populated and rarely visited. We focus on areas with at least 35 listings just before the consolidation took place.[8] We are thus left with 281 markets, which represent 83.7% of inquiries and 82.8% of bookings on the platform. Using the definition of treatment and control groups described in the previous section, we identify 173 markets that experienced discontinuous increase in market thickness post-consolidation, with an average percentage increase of 29.6% as shown in Table 1. For the purpose of illustrating comparability and parallel trends between treatment and control groups, we further divide the treatment sample into two subgroups: with low and high levels of increase in thickness. However, we will account for the level of percentage increase in the subsequent empirical analysis. Using the 75th percentile of the percentage increase (17.1%) as the cutoff,[9] we identify 41 markets with relatively high increase of thickness, with an average of 34.6% increase, and 132 markets with relatively low increase of thickness, with an average of 8.0% increase. As Figure 3 illustrates, the level of increase in market thickness is fairly evenly distributed across markets in different geographical locations and of different sizes. Appendix C presents examples of markets that experience different changes in thickness.

Table 1 Treatment and Control Groups Based on Levels of Increase in Market Thickness

	No. of Markets	% Increase in Market Thickness	
		Mean	[5th percentile, 95 percentile]
No Increase	108	0.0%	[0.0%, 0.0%]
Increase	173	29.6%	[2.2%, 91.8%]
Low Increase	132	8.0%	[1.8%, 16.4%]
High Increase	41	34.6%	[17.7%, 62.7%]
Total	281	13.6%	[0.0%, 62.3%]

5.2. Comparability of Control and Treatment Groups

In this subsection, we examine whether the above construction of treatment and control groups leads to balanced samples of markets with similar trip, traveler, and market characteristics.

[8] The estimation results are robust to alternative cutoffs, 10 and 25.

[9] The results are robust to alternative thresholds at the 50th and 90th percentiles.

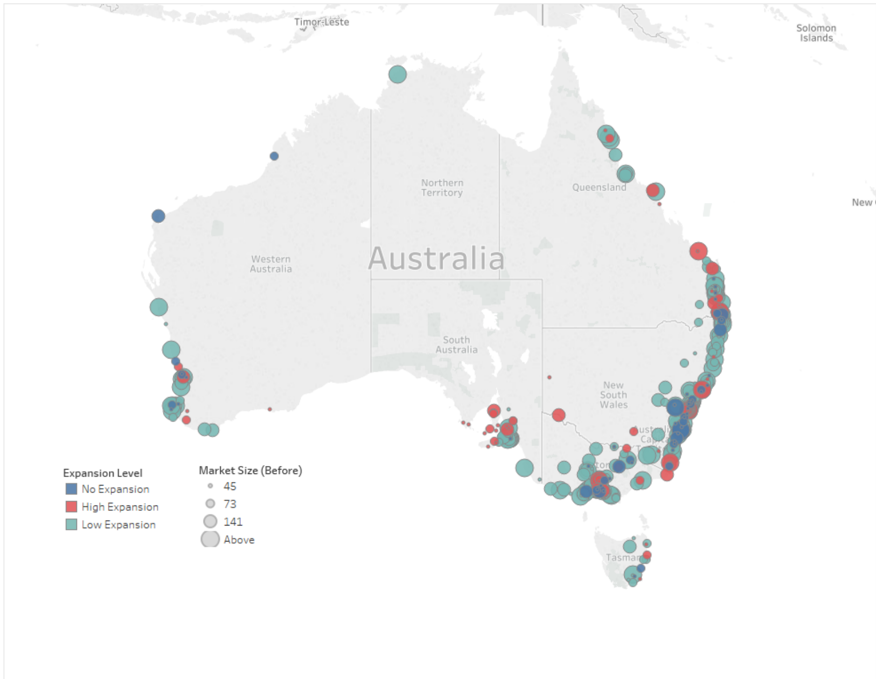

Figure 3 **Listing Expansions by Markets**

Our main analysis focuses on the period from 60 days before (June 18 to August 18, 2014) to 60 days after (August 23, 2014 to October 23, 2014) the migration of listings that took place in late August 2014.[10] We use the rest of the data to inform us about the live status of each listing and dates that are blocked during the study window, as well as for robustness analyses. During the four months of the study period, 940,000 inquiries were sent by travelers, with 83,000 of these eventually becoming confirmed matches. Note that a traveler may send out multiple inquiries for the same trip, sometimes with slight variations in destinations or travel dates. To accommodate such variations, we use similarities in destinations, travel dates and search dates to group inquires made by the same traveler for the same intended trip (see Appendix D for details). As a result, we associate 920,000 inquiries identified with traveler IDs for 456,000 trips. About 84%—that is, 771,000 inquiries and 372,000 trips—were in the 281 markets under study. Out of these 372,000 trips, 56,000 (about 15%) were eventually confirmed.

To assess the comparability of treatment and control markets, we first compare trip characteristics and traveler characteristics between these two groups during the before period. We construct trip-level characteristics along the following dimensions: length of stay, weekend stay, number of weekends, number of adults, number of children, search days in advance, and holiday travel. To identify relevant holidays for the period we study, we obtained 2014 and 2015 Australian national holidays from Australia.gov.au, and complemented this with the aggregated total inquiries for each

[10] The results are consistent if we instead focus on 30 or 90 days before and after the migration of listings.

travel date from our data. We therefore identify three holiday periods that are relevant for our analysis: Labor Day (October 3 to October 5, 2014), Christmas and New Year holiday season (December 20, 2014 to January 25, 2015), and Easter (April 3 to April 5, 2015). We create a dummy variable for each holiday that equals one if the travel dates contain any day during the holiday period and zero otherwise. The results are consistent if we define it as a continuous measure representing the percentage of travel days that fall into a holiday period. During the before period (i.e., two months prior to August 18, 2014), no inquires were made for the Easter holiday in the coming year. Therefore, it is only relevant for the after period.[11]

The upper panel of Table 2 shows that travelers and their trips are similar along all dimensions between treatment and control groups, and none of the differences in averages is statistically significant. Recall that there are cases in which a traveler sends inquiries to multiple markets. We assign the trip to the first inquired market by the traveler in these cases. Similarly, there are cases where travel dates may vary across inquiries within a trip. In such cases, we also define the travel dates of a trip as those of the first inquiry.

Next, we compare market characteristics between treatment and control markets. Although markets in the treated and control groups are similar along many dimensions, they are different along a few. We select the key characteristics among the thirty-five market characteristics based on t-test results comparing means between treated and control groups and based on collinearity of these variables (see Appendix E for details). The lower panel in Table 2 shows the summary statistics for these key market characteristics by treatment and control groups. Markets in the treated group tend to be larger in size, that is, they have more listings. They also seem more urban (fewer rooms per property, more apartments or serviced apartments, farther away from adventures and national parks), and more business oriented (more executive listings, more listings with wheelchair access and more listings for short-term rentals). This suggests that it may be important to control for these variables when estimating the effect of market thickness on matching rate. Even though these factors are likely already correlated with trip-level characteristics such as length of day or weekend stay, they may also be correlated with other unobserved characteristics of the traveler or the trip such that it will affect how willing a traveler is to reach a deal.

5.3. Parallel Trend and Difference-in-Differences at Macro Level

In this subsection, we first evaluate whether the parallel-trend assumption holds for the control and treatment groups before the migration of listings. We then examine whether it has led to a noticeable change in the number of inquiries sent by travelers, the number of listings that received inquiries, and matching rates in treatment markets relative to control markets.

[11] We verify that the results are not affected by the nonpresence of an Easter trip during the before period. Easter contains only 3 days. If we exclude trips that contain the Easter holiday, the results are consistent.

Table 2 Traveler, Trip and Market Characteristics by Treatment and Control Groups

			Treatment		
Variables	Total	Control	All	Low Increase	High Increase
Traveler-Trip Characteristics					
Length of Stay	4.308	3.840	4.513	4.286	5.269
	(4.321)	(3.786)	(4.521)	(4.095)	(5.651)
Weekend Stay	0.833	0.820	0.839	0.834	0.856
	(0.373)	(0.384)	(0.367)	(0.372)	(0.351)
No. of Weekends	1.145	1.102	1.163	1.149	1.210
	(0.513)	(0.489)	(0.522)	(0.515)	(0.542)
No. of Adults	4.155	4.385	4.055	4.159	3.708
	(2.892)	(3.025)	(2.826)	(2.919)	(2.460)
No. of Children	0.912	0.971	0.886	0.903	0.828
	(1.436)	(1.498)	(1.407)	(1.432)	(1.316)
Days in Advance	63.53	60.22	64.99	65.03	64.85
	(54.89)	(54.51)	(54.99)	(55.23)	(54.21)
Labor Day	0.0411	0.0402	0.0415	0.0426	0.0375
	(0.198)	(0.196)	(0.199)	(0.202)	(0.190)
Christmas and New Year	0.160	0.151	0.165	0.167	0.157
	(0.367)	(0.358)	(0.371)	(0.373)	(0.364)
Domestic	0.954	0.967	0.949	0.958	0.919
	(0.209)	(0.179)	(0.221)	(0.200)	(0.273)
Market Characteristics					
No. of Listings Per Market	114.9	87.02	132.0	138.7	107.8
	(85.87)	(46.77)	(99.20)	(99.23)	(97.58)
No. of Rooms Per Listing	2.902	3.056	2.808	2.930	2.372
	(0.544)	(0.441)	(0.581)	(0.520)	(0.592)
% of Apartment Listings	0.146	0.0947	0.178	0.142	0.306
	(0.139)	(0.0938)	(0.153)	(0.119)	(0.190)
% of Executive Listings	0.232	0.189	0.259	0.246	0.304
	(0.163)	(0.143)	(0.170)	(0.164)	(0.188)
% of Listings Close To Advendure	0.0195	0.0236	0.0169	0.0177	0.0140
	(0.0240)	(0.0313)	(0.0178)	(0.0180)	(0.0173)
% of Listings Close To National Parks	0.263	0.287	0.249	0.274	0.159
	(0.139)	(0.0938)	(0.153)	(0.119)	(0.190)

Note. Summary statistics during the before period (June 18 to August 18, 2014). Standard deviations are shown in parentheses.

Figure 4 displays these statistics for both treatment and control markets during the before and after periods. All data series are de-seasonalized. Specifically, we use Hodrick–Prescott decomposition to remove day-of-week effects from the time-series data. All subfigures demonstrate that treatment and control markets follow similar trends during the before period, a critical assumption to ensure validity of the difference-in-differences approach. Moreover, all subfigures demonstrate significant differences between treatment and control markets during the after period. The number of inquiries sent by travelers and the number of listings that received inquiries exhibit significant jumps right after the migration took place. We test for structural breaks with unknown break dates, and the results are displayed in Table 3. The test correctly identifies structural breaks on August 23 or 24, 2014, for the treatment markets right at the end of the migration period. The test did not identify any breaks with statistical significance for the control markets. After confirming that the event is real and impactful, we examine how the sharp increase in market thickness affected traveler- and listing-level matching rates. The bottom two figures demonstrate that traveler confirmation rate and listing occupancy rate have both gone down for the treatment group during the after period relative to the control group. In all subfigures, the changes in markets with high levels of thickness increase are larger than for those markets with low levels of increase.

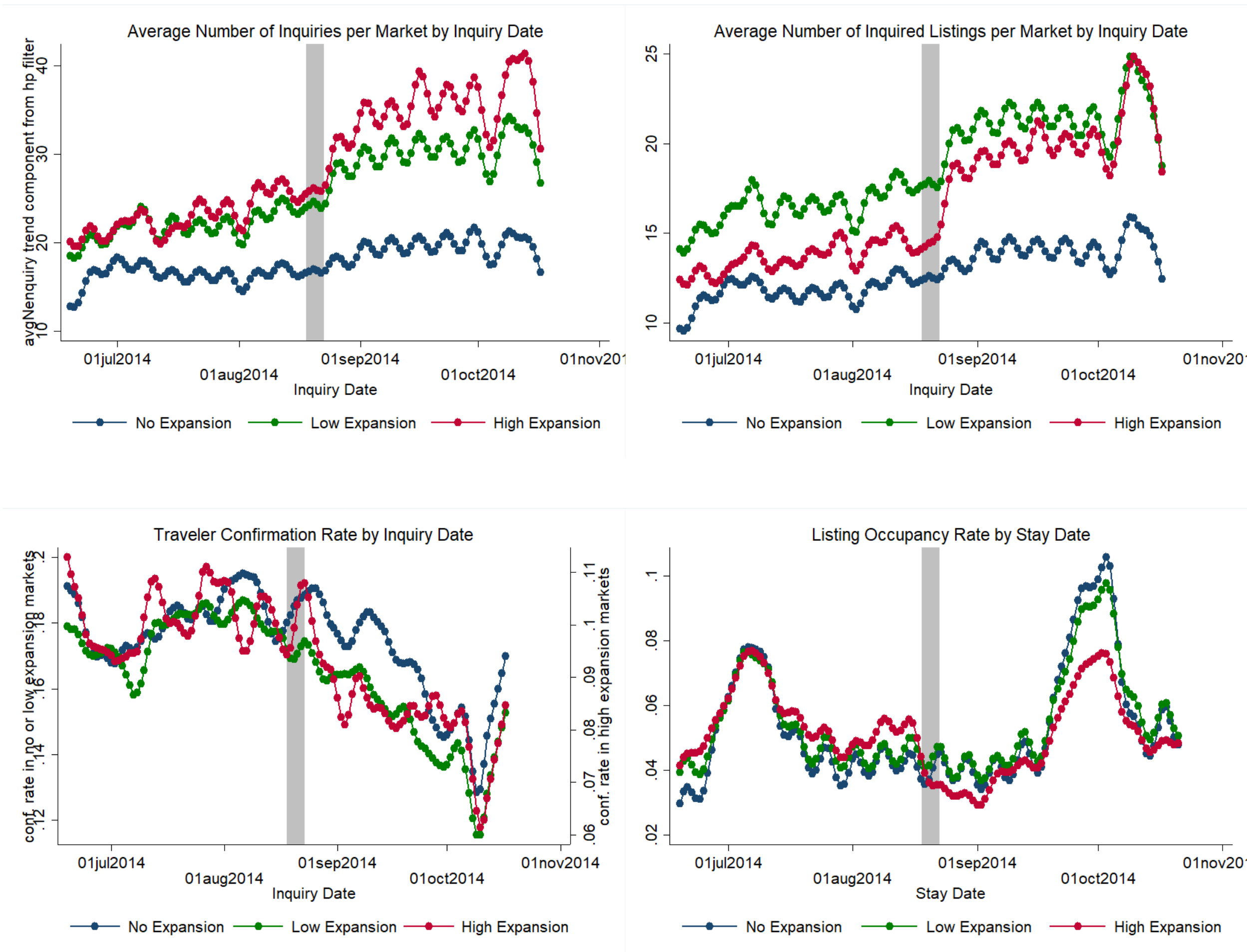

Figure 4 Difference-in-Differences Comparisons

Table 3 Test for a Structural Break with an Unknown Break Date

	No. of Inquiries		No. of Inquired Listings	
	Identified Break Date	p value	Identified Break Date	p value
No Increase in Thickness	31-Aug-14	0.3833	3-Aug-14	0.4323
Low Increase in Thickness	24-Aug-14	0.0347	23-Aug-14	0.1844
High Increase in Thickness	24-Aug-14	0.0545	23-Aug-14	0.0341

Note. p values are from supremum Wald tests.

While this analysis gives us the first evidence that increase in market thickness has led to lower matching rates, it does not properly adjust for trip, traveler, and market heterogeneity, or data censoring. In the next section, we will use the empirical specifications laid out previously to formally estimate the effect of market thickness on matching outcomes.

6. Estimation Results

In this section, we present the estimated effects of market thickness on matching probability and ensure our results are robust to alternative specifications and explanations. We then compare our results with those in the literature using a similar market-level analysis to examine what drives the difference in the results.

6.1. Effect of Market Thickness on Matching Rate

Table 4 shows the estimated effect of increase in market thickness on matching probabilities. We first review the effects on traveler confirmation rate. Columns (1) and (2) correct for data censoring and report average semi-elasticities for easy interpretation. Column (1) shows that a traveler's matching probability will decrease by 19.2% when market size doubles (i.e., increases by 100%). We calculate that for an average traveler, his chance of being matched is 5.6% lower if his destination is among those markets that expanded after the consolidation. On an average calendar date, about 330 trips were confirmed in these markets. The 5.6% reduction in traveler confirmation rate therefore translates into a loss of 20 potential matches on each calendar date. Given an average daily rate of 290.95 Australia dollars (AUD) and an average length of stay of 4.59 days, the platform thus lost 26,709 AUD worth of transactions on an average calendar date.

Table 4 Effect of Market Thickness on Matching Rate

	Traveler Confirmation Rate			Listing Occupancy Rate		
	All TVLRs (1)	Existing TVLRs (2)	Existing TVLRs (3)	All LSTGs (4)	Existing LSTGs (5)	Existing LSTGs (6)
Percent Increase in Thickness	-0.192***	-0.129*	-0.154**	-0.149*	-0.145*	-0.159**
	(0.060)	(0.069)	(0.070)	(0.080)	(0.081)	(0.094)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Add'l Controls			Yes			Yes
N	367672	288824	288824	959855	946798	946798
LogL	-149406.1	-120595.9	-97915.7	-210166.9	-203789.0	-207168.5

Note. Columns (1) to (3) report average semi-elasticities on treated markets during the after period. The estimates correct for data censoring. Standard errors are obtained through bootstrapping. Controls in these columns include main effects and trip, traveler, and market characteristics. Column (3) also includes average listing confirmation rate, average listing age (in days and log-scaled), and percentage of listings joined during the four-day migration period as additional controls. Columns (4) to (6) report coefficient estimates from fixed-effect logit models. Controls in these columns include main effects, listing fixed effects, and travel-date characteristics. Column (6) also includes average number of prior trips travelers have searched for on the platform as an additional control. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Results in Column (1) demonstrate that a traveler who visits the combined platform has a lower probability of being matched than a traveler who visits the original platform. One possible explanation for this lowered matching rate is that travelers on the merged and the original platforms are different. For example, if travelers on the two merged platforms are more difficult to match to begin with, either because they are more selective or less desirable, the average traveler matching rate will be lower for the combined platform. Moreover, travelers from the two merged platforms could have a lower matching rate because they are unfamiliar with the new system. To account for these possibilities, we conduct the same analyses on existing travelers from the main platform, that is, travelers who have sent at least one inquiry to some listings during the before period, which account for 74.3% of all traveler trips. The results are shown in Column (2). Existing travelers also suffered from a lower matching probability after the consolidation. When market size doubles, the matching rate for existing travelers decreases by 12.9%. For an average existing traveler, his chance of being matched is lowered by 3.9% if his destination is among those markets that expanded after the merger. In Appendix F, we verify that our results are robust to variations in the definitions of primary market and market characteristics.

The results for all travelers and existing travelers both confirm that the increase in market thickness led to a lower matching rate for travelers. Next, we examine how listing occupancy rate changed as a result of the change in market thickness. It is possible that the lower matching rate for travelers is a result of intensified competition from the demand side. If so, the supply-side matching rate could actually go up. Columns (4) and (5) display the effect of market thickness on listing occupancy rate. Similar to travelers, listing owners also experienced a decrease in matching rate. As shown in Column (4), the average occupancy rate decreases by 14.9% when market size doubles. A possible explanation for the lower occupancy rate is that migrated listings are more difficult to match. Another possible explanation is that the newly migrated owners are unfamiliar with the new system, and may appear less responsive to travelers' inquiries. To test whether these possibilities could explain our results, we conduct the same analyses on existing listings from the original platform. As shown in Column (5), when market size doubles, existing listings also experience a 14.5% reduction in occupancy rate. The results from the subsample analyses reassure that the lower matching rate is not simply driven by potentially lower matchability of the new participants from merged platforms. Since we do not have data on the two merged platforms during the before period, these subsample analyses are critical to ensure fair before-and-after comparisons.

We further note that in a two-sided market where each participant on one side can freely interact with any participant on the other, it is almost inevitable that some existing travelers would have interacted with a new listing owner migrated from the other two platforms, or some existing owners would have interacted with a new traveler too. Moreover, even if an existing traveler has not directly interacted with a listing owner from the merged platforms, but if the listing owners he contacted has interacted with other travelers from those platforms, the traveler himself is indirectly connected to other participants from those merged platforms. If travelers and owners from the merged platforms are indeed more difficult to match, it may affect the matching rate of those participants they directly interact with, and maybe others through indirect connections too (likely to a less degree). It thus seems that the "cleanest" approach would have been finding a subset of participants (both sides) from the main platform that is completely isolated from the rest (no direct or indirect connections). This "cleanest" subset exists in our data, but it represents only 3% of all participants. This is not surprising, because if one had to restrict a significant portion of participants to interact with whoever they would interact with without the merger of the platforms, it defeats the purpose of acquiring and merging platforms. However, we do want to ensure that the observed result is not a result of some sort of contagious process, but rather because of the treatment. We conducted further analyses in which we control for characteristics of participants that an existing traveler or listing owner directly interact with and see if the results would be significantly altered. The rationale is that, if the effects disappear or if their magnitudes are significantly smaller, then we might suspect that the lower matching rate can be a result of lower matchability of new participants that went "viral". The results are presented in Columns (3) and (6) of Table 4. In Column (3), in addition to previous controls, we also include average listing confirmation rate, average listing age (in days and log-scaled), and percentage of listings joined during the four-day migration period as controls. In Column (6), in addition to previous controls, we include average number of prior trips travelers have searched for on the platform.[12] As with Column (2) and (5), the analyses are conducted on the subsamples of existing travelers or existing listing owners only. We do not see any evidence of the effects going away, but they actually become more significant. This new evidence together with the previous ones reassure us that the resulting lower matching rate is most likely due to the increased market thickness, rather than differences of participants from different platforms, if such differences exist.

In addition, we study heterogeneous treatment effects across markets of different sizes. We find that the effect of an increase in market thickness on matching rate is more salient in larger markets than in smaller markets. This is likely because larger markets offer more outside options such as hotels and motels. Please refer to Appendix G for details.

6.2. Market Level Analysis

Although the literature largely demonstrates constant returns to scale in two-sided matching markets, we do not find evidence of increasing or constant returns to scale. Rather, we find evidence of diminishing returns to scale. We note our setting and empirical method differ from those used in the literature in several significant ways. First, our context, peer-to-peer online marketplaces, is different from previous studies, which are primarily concerned with labor markets. It may be that different types of matching markets exhibit different returns to scale. Second, our study is conducted at the micro level (i.e., the matching outcome of each trip or each listing day), whereas

[12] The number of prior trips is measured as starting from January 2014, since we only have access to traveler inquiry data in 2014.

previous studies were conducted at the macro level (e.g., aggregate unemployment rate of a city or a country). Third, a more profound difference between our study and previous studies is that we exploit an exogenous shock to market size to make causal inferences, whereas previous studies rely on either time-series or cross-sectional variations and lack clear identification strategies.

To understand the drivers of the differences in the results (context, aggregation, or identification strategy), and to test the consistency of our results relative to previous ones under otherwise similar empirical specifications, we conduct analysis at the market level. We first examine whether a thicker market has a higher or lower matching rate, exploiting the cross-sectional variations at the market level as in previous studies. Next, to estimate the causal effect on matching rate at the market level, we again exploit the listing migration as an exogenous shock to market thickness. This time, we use it as an instrument to the number of listings in the after period, and see if we will reach the same conclusions as previously obtained using the trip-level analysis or listing-day analysis.

Table 5 Effect of Market Thickness on Matching Rate: Market-Level Analysis

	Traveler Confirmation Rate			Listing Occupancy Rate		
	Before (1)	After (2)	After w/ IV (3)	Before (4)	After (5)	After w/ IV (6)
ln (No. of Listings)	0.004	0.001	-0.031**	0.000	-0.001	-0.006*
	(0.008)	(0.007)	(0.014)	(0.001)	(0.002)	(0.003)
Market Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
N	281	281	281	281	281	281
R-sq. (adj.)	0.205	0.212	0.158	0.200	0.218	0.193

Note. Columns (1), (2), (4), and (5) are estimated using OLS regressions. Columns (3) and (6) are estimated using 2SLS regressions. In Columns (3) and (6), we use the number of new listings joined during the migration period (log-scaled) as the instrument to the number of listings (log-scaled) in the after period. Only existing traveler and existing listings are included in the analysis of traveler confirmation rate and listing occupancy rate, respectively. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table 5 displays the results. Note that when calculating market-level matching rates during the after period, we include only existing travelers or existing listing owners from the main platform. Based on the results shown in Columns (1), (2), (4), and (5), one would reach a conclusion of constant returns to scale, consistent with the literature, albeit in a different context. Very interestingly, the results from the instrumental variables (IV) regression in Columns (3) and (6) once again point to decreasing returns to scale. That is, increases in market thickness lead to lower matching rates. This result not only demonstrates the robustness of our findings, it also indicates that the difference between our findings and those from the previous studies is not merely a result of differences in contexts or data aggregation, but is likely a result of a causal relationship. The magnitudes of the effects are comparable with those obtained from trip-level or listing-day analysis. Based on the estimates, when market size doubles, traveler confirmation rate decreases by

$3.1 \times \ln(2) = 2.15$ percentage points, equitant to a 14.0% reduction; listing occupancy rate decreases by $0.6 \times \ln(2) = 0.42$ percentage points, equivalent to an 14.2% reduction.

7. Search Friction and Market Congestion

Intuitively, one would probably expect that a higher market density leads to a higher matching rate. Our empirical results reject this hypothesis. We find evidence that increased market thickness may actually lead to worse matching outcomes, and we show this causally. In this section, we examine what drives this result.

First, we test whether the lower matching rate is a result of travelers exerting less search effort and quitting earlier after the consolidation. Table 6 displays the effect of market thickness on search intensity. Columns (1) and (2) show that travelers searched more (i.e., sent more inquiries per trip) in those markets that experienced an increase in thickness. Specifically, travelers' search intensity increases by 14.7% when market size doubles (i.e., increases by 100%). If we focus on those travelers who have had prior experiences on the main platform, their search intensity increases even more: by 18.3%. As a result, a listing owner receives 20% more searches per stay date, as shown in Columns (3) and (4). To formally test whether a traveler is more (or less) likely to quit search after the merger, we conduct survival analyses on travelers' search process. Columns (5) and (6) show that travelers are less likely to quit search and also less likely to be matched after the merger. In sum, this evidence demonstrates that the reduced matching rate is not a result of insufficient search effort. Rather, it could be a result of too much search, a possibility we will elaborate shortly.

Table 6 Effect of Market Thickness on Search Intensity

	Inquiries Sent Per Trip		Inquiries Received Per Date		Quit	Success
	All TVLRs (1)	Existing TVLRs (2)	All LSTGs (3)	Existing LSTGs (4)	Existing TVLRs (5)	Existing TVLRs (6)
Percent Increase in Thickness	0.147***	0.183***	0.208***	0.196***	-0.117***	-0.308***
	(0.021)	(0.021)	(0.008)	(0.008)	(0.029)	(0.090)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	367672	288824	3283400	3156276	683706	683706
LogL	-637834.0	-545934.1	-2904629.8	-2824627.3	-2874054.7	-520026.6

Note. Columns (1) and (2) estimate effect of market thickness on the number of inquiries sent per trip by travelers using zero truncated Poisson regression. Columns (3) and (4) estimate the effect of market thickness on the number of inquiries received per listing date using Poisson regression with listing fixed effects. Columns (5) and (6) are estimated using Cox survival model. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Second, even though the matching rate is lower, perhaps the increased search effort leads to better matched outcomes. That is, perhaps travelers find lower-priced properties that match their needs. Or perhaps listing owners receive higher revenues despite lower occupancy rates. Table 7 shows how the increase of market thickness affects prices paid by travelers and revenues received by listing owners. We do not find evidence of better matching outcomes. In fact, we find that travelers actually paid higher prices (9.8% higher) and listing owners suffered from lower revenues (7.9% lower).

Table 7 Effect of Market Thickness on Price Paid by Travelers and Weekly Revenue of Listing Owners

	Paid Price		Weekly Revenue	
	All TVLRs (1)	Existing TVLRs (2)	All LSTGs (3)	Existing LSTGs (4)
Percent Increase in Thickness	0.071*	0.098**	-0.094***	-0.079**
	(0.039)	(0.044)	(0.030)	(0.032)
Controls	Yes	Yes	Yes	Yes
N	50529	41192	577248	542529
R-sq. (adj.)	0.318	0.317	0.00339	0.00348

Note. Both price and revenue are log scaled. Columns (1) and (2) are estimated using OLS regressions. Columns (3) and (4) are estimated using fixed-effect regressions with listing fixed effects. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

How could travelers search more, but match less and pay higher prices? Similarly, how could listing owners receive more inquiries from interested travelers but end up with lower occupancy rates and lower revenues? The lower matching rate is likely caused by search frictions in congested markets. When markets become denser, both parties face a greater number of potential options. For travelers, the number of options available is evident once they enter the trip parameters and land on the search-results page. For listing owners, although demand is not directly observable to them, they will start to see an increase in their popularity after receiving multiple inquiries from travelers. Many studies have shown that individuals have difficulty making decisions when facing an abundance of choices (Jacoby et al. 1974, Iyengar and Lepper 2000, Bertrand et al. 2010). To understand whether this is likely the case in our context, it is helpful to take a closer look at the matching process. As we discussed in Section 3.1, the matching process is not instantaneous. Because of information asymmetry, travelers and listing owners need to spend time and effort to find out whether there is a potential match between what they need and what the other party can offer. During his communication with a listing owner, a traveler is trying to figure out whether the property offers features desirable for the intended trip; a listing owner is trying to ensure that the traveler is capable of making the payment on time and will properly maintain the property during the stay. Because of uncertainties, a traveler (a listing owner) could be communicating with multiple listing owners (travelers) simultaneously. While both parties obtain information from each other and make comparisons, valuable time could have passed. When travelers are ready to decide, their favorite choice may no longer be available (due to information asymmetry), or the other party may not be ready to commit yet (due to information overload), so they are forced to seek either pricier options or quit altogether and seek outside options. As a result, travelers either end up paying more for their trips or simply end up unmatched. An alternative explanation for higher transaction prices is that listing owners raised their prices after the consolidation. We find it is unlikely to be the case given that the quoted price by listing owners (conditional on receiving an inquiry) remains the same post consolidation. Since we do not observe prices when there is no inquiry received for a stay date, we cannot fully examine this possibility, which also leaves open opportunities for future research. Yet a third explanation for higher transactions prices is that higher search frictions prevented customers from fully exploring the market and finding lower prices. However, this would seem inconsistent with the finding that travelers sent more inquiries after the shock. While we believe this friction is likely what causes the matching rate to decrease when the market size increases, without access to data on the timing and content of the communication between travelers and listing owners, we have to leave it for future research to further examine the form and extent of this friction.

We expect this congestion effect to be more severe when there is less time to match. This is because if the need for match never expires, travelers could search indefinitely, and they eventually can find a match because unmatched listing owners are still available. However, when a matching deadline is close, the time pressure will force travelers to give up more frequently and seek outside options. This is what we observe in Table 8. When the intended travel dates are less than one month ahead, doubling market size leads to a 28.6% increase in the number of inquiries sent, 40.0% decrease in confirmation rate and 26.3% increase in transaction prices, compared to only 12.8% increase in the number of inquiries and no change in confirmation rate and transaction prices when the intended travel date is more than one month away. This result confirms that while increase in market thickness leads to higher search intensity for trips that are nearer or farther in the future, it only leads to worse matching outcomes when the matching deadline is close.

Table 8 Effects of Market Thickness by Closeness to Matching Deadlines (Existing Travelers Only)

	No. of Inquiries Sent		Confirmation Rate		Price Paid	
Days In Advance	≤ 30 (1)	> 30 (2)	≤ 30 (3)	> 30 (4)	≤ 30 (5)	> 30 (6)
Percent Increase in Thickness	0.286*** (0.036)	0.128*** (0.025)	-0.400*** (0.078)	0.076 (0.115)	0.263*** (0.071)	0.011 (0.056)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	111163	177661	111131	177661	16353	24839
LogL	-192929.0	-352361.1	-46802.2	-73573.9		
R-sq. (adj.)					0.231	0.332

Note. Results are for existing travelers. Price is log scaled. Columns (3) and (4) show semi-elasticities corrected for censored data. Standard errors in Columns (3) and (4) are obtained through bootstrapping. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Finally, we would like to point out some caveats in interpreting these results. We should note that the results hold true when the matching process and the search technology remain unchanged as a market becomes thicker. If, for example, the platform makes the matching process more instantaneous (such as the instant-booking function offered by Airbnb) or limits the number of options presented to travelers through more refined search, it may be able to reduce search friction and market congestion, which presents an interesting area for future empirical research. Second, the observed effects are best interpreted as applicable for markets that reach a certain scale. When markets are sufficiently small, growth in market sizes may generate different effects. Our analysis focuses on markets that have at least thirty-five listings. However, for those markets in their early growth stages that have only a handful of listings, it is possible that search frictions do not play a dominant role, and increase in market thickness could lead to better matching outcomes. Last, our results show that higher market thickness can lead to lower matching rates for participants on both sides of the market, but it does not necessarily indicate lower overall matching efficiency. Lower matching rates mean that participants are less likely to find successful matches, but those who do may have found better matches. The higher average transaction price may be an indication that listings are matched to travelers who value them more. While the average listing occupancy rate was lower due to the increased market thickness, occupancy rate for those popular listings did not experience statistically significant change (see Appendix G for details). Without data on matching quality, one cannot fully examine how consumer surplus would change as a result of thickness change, which we leave for future research.

8. Conclusion

Matching-market equilibrium often depends critically on the shape of the matching function—whether it exhibits increasing, constant or decreasing returns to scale—which is perhaps best tested empirically. However, answering this question is challenging because market size is typically determined endogenously, making causality difficult to establish. We address this challenge by exploiting an exogenous shock to market thickness caused by a one-time migration of listings from other platforms, which gave rise to a quasi-experimental design.

We find several interesting results. Thicker markets are not necessarily associated with higher matching rates. On the contrary, we find that an increase in market thickness may lead to a lower matching probability for both travelers and hosts, keeping search technology and other factors unchanged. In our context, the platform lost 5.6% of potential matches each day because of the increased market size. This effect can be attributed to increased search friction. We find that facing a larger number of options, travelers search more. Since the matching process is not instantaneous, both travelers and listing owners end up conversing with more potential partners before they can reach a decision, at which point, the other party may not be ready to commit yet or may be no longer available. As a result, travelers are forced to quit their search process and seek outside options. The effect of increases in market size on matching rate is even more prominent when the matching deadline is close.

Our findings have important implications for both theory and practice. They highlight the need for carefully designed empirical research to measure the causal impact of market thickness on matching rates. They also call for careful evaluations of the properties of matching functions and their implications in modeling work. Practically, peer-to-peer marketplace operators should pay close attention to search frictions as markets grow, and consider investments or interventions to alleviate congestions or externalities caused by search frictions. This is particularly important when the market exhibits high degrees of heterogeneity and when needs for matching are imminent.

Of course, our study is not free of limitations, which provides several opportunities for future research. First, we do not have access to granular search data. Such data would allow a more in-depth analysis of travelers' search processes, which might offer insights on how platform operators can help users avoid extensive and inefficient searches and find a match quickly. Second, we analyze matching rate, but we do not have direct measures of matching quality (e.g., satisfaction after matching). It would be interesting to investigate whether the quality of matches is better in larger markets, despite a decrease in matching probability. Third, we observe only transacted prices but not listing prices, so we are not able to fully examine effects of market thickness on prices or the potential effects of supply-side price changes on matching rates and search intensity. Instead, we measure the overall effects of the market thickness on matching rates, search intensity and (transacted) prices. Indeed, it would be interesting for future research to measure the effect on listing price and to decompose the overall effects of the market thickness to direct and indirect effects by estimating the interplay of prices, search and matching rate. To disentangle the direct and indirect effects, one would potentially need multiple instruments because prices, search and matching rates are determined through a system of simultaneous equations. Last, this paper measures the short-term effect following the increased market thickness. In the long run, hosts may decide to charge lower prices after realizing that the conversion rate is low, become less active, or even exit the platform. The platform operator may decide to invest in advanced search technology or provide incentives to keep participants active. There will be a different equilibrium.

References

Allon, Gad, Achal Bassamboo, Eren B. Cil. 2012. Large-scale service marketplaces: The role of the moderating firm. *Management Science* **58**(10) 1854–1872.

Armstrong, Mark. 2006. Competition in two-sided markets. *The RAND Journal of Economics* **37**(3) 668–691.

Arnosti, Nick, Ramesh Johari, Yash Kanoria. 2016. Managing congestion in matching markets. Working Paper, Stanford University, Stanford, CA.

Bakos, J. Yannis. 1997. Reducing buyer search costs: Implications for electronic marketplaces. *Management Science* **43**(12) 1676–1692.

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, Jonathan Zinman. 2010. What's advertising content worth? evidence from a consumer credit marketing field experiment*. *The Quarterly Journal of Economics* **125**(1) 263–306.

Bimpikis, Kostas, Wedad J. Elmaghraby, Ken Moon, Wenchang Zhang. 2017. Managing market thickness in online b2b markets. Working Paper, Stanford University, Stanford, CA.

Brynjolfsson, Erik, Michael D. Smith. 2000. Frictionless commerce? a comparison of internet and conventional retailers. *Management Science* **46**(4) 563–585.

Cachon, Gerard P., Kaitlin M. Daniels, Ruben Lobel. 2017. The role of surge pricing on a service platform with self-scheduling capacity. *Manufacturing & Service Operations Management* Published Online.

Cui, Ruomeng, Jun Li, Dennis Zhang. 2016. Discrimination with Incomplete Information in the Sharing Economy: Evidence from Field Experiments on Airbnb. Working Paper, Emory University, Atlanta, GA.

Cullen, Zoe, Chiara Farronato. 2016. Outsourcing tasks online: Matching supply and demand on peer-to-peer internet platforms. Working Paper, Harvard University, Cambridge, MA.

De los Santos, Babur, Sergei Koulayev. 2017. Optimizing click-through in online rankings with endogenous search refinement. *Marketing Science* **36**(4) 542–564.

Dhanorkar, Suvrat. 2017. Environmental benefits of internet-enabled c2c closed loop supply chains: A quasi-experimental study of craigslist. *Management Science* Forthcoming.

Dhanorkar, Suvrat, Karen Donohue, Kevin Linderman. 2015. Repurposing materials and waste through online exchanges: Overcoming the last hurdle. *Production and Operations Management* **24**(9) 1473–1493.

Diamond, Peter A. 1982. Aggregate demand management in search equilibrium. *Journal of Political Economy* **90**(5) 881–894.

Dinerstein, Michael, Liran Einav, Jonathan Levin, Neel Sundaresan. 2018. Consumer price search and platform design in internet commerce. *American Economic Review* **108**(7) 1820–59.

Dzyabura, Daria, John Hauser. 2017. Recommending products when consumers learn their preferences. Working Paper, New York University, New York City, NY.

Fleder, Daniel, Kartik Hosanagar. 2009. Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management Science* **55**(5) 697–712.

Fradkin, Andrey. 2017. Search, matching, and the role of digital marketplace design in enabling trade: Evidence from airbnb. Working Paper, MIT Sloan, Cambridge, MA.

Ghose, Anindya, Panagiotis G. Ipeirotis, Beibei Li. 2012. Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content. *Marketing Science* **31**(3) 493–520.

Ghose, Anindya, Panagiotis G. Ipeirotis, Beibei Li. 2014. Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science* **60**(7) 1632–1654.

Ghose, Anindya, Panagiotis G. Ipeirotis, Beibei Li. 2017. Modeling consumer footprints on search engines: An interplay with social media. Working Paper, New York University, New York City, NY.

Hagiu, Andrei, Bruno Jullien. 2011. Why do intermediaries divert search? *The RAND Journal of Economics* **42**(2) 337–362.

Haubl, Gerald, Valerie Trifts. 2000. Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Science* **19**(1) 4–21.

Horton, John. 2017. Buyer uncertainty about seller capacity: Causes, consequences, and a partial solution. Working Paper, NYU, New York City, NY.

Iyengar, Sheena S., Mark R. Lepper. 2000. When choice is demotivating: Can one desire too much of a good thing?. *Journal of Personality and Social Psychology* **79**(6) 995 – 1006.

Jacoby, Jacob, Donald E. Speller, Carol A. Kohn. 1974. Brand choice behavior as a function of information load. *Journal of Marketing Research* **11**(1) 63–69.

Kabra, Ashish, Elena Belavina, Karan Girotra. 2017. Designing incentives to scale marketplaces. Working Paper, INSEAD, Fontainebleau, France.

Kanoria, Yash, Daniela Saban. 2017. Facilitating the search for partners on matching platforms: Restricting agent actions. Working Paper, Columbia Business School, New York City, NY.

Koulayev, Sergei. 2014. Search for differentiated products: identification and estimation. *The RAND Journal of Economics* **45**(3) 553–575.

Kroft, Kory, Devin G. Pope. 2014. Does online search crowd out traditional search and improve matching efficiency? evidence from craigslist. *Journal of Labor Economics* **32**(2) 259–303.

Kuksov, Dmitri, J. Miguel Villas-Boas. 2010. When more alternatives lead to less choice. *Marketing Science* **29**(3) 507–524.

Lederman, R., M. Olivares, G. van Ryzin. 2014. Identifying competitors in markets with fixed product offerings. Working Paper, Columbia Business School, New York, NY.

Li, J., N. F. Granados, S. Netessine. 2014. Are consumers strategic? Structural estimation from the air-travel industry. *Management Science* **60**(9) 2114–2137.

Li, Jun, Antonio Moreno, Dennis J Zhang. 2016. Pros vs joes: Agent pricing behavior in the sharing economy. Working Paper, University of Michigan, Ann Arbor, MI.

Moreno, Antonio, Christian Terwiesch. 2014. Doing business with strangers: Reputation in online service marketplaces. *Information Systems Research* **25**(4) 865–886.

Nosko, Chris, Steven Tadelis. 2015. The limits of reputation in platform markets: An empirical analysis and field experiment. Working Paper, University of Chicago, Chicago, IL.

Parker, Chris, Kamalini Ramdas, Nicos Savva. 2016. Is it enough? evidence from a natural experiment in india's agriculture markets. *Management Science* **62**(9) 2481–2503.

Parker, Geoffrey G., Marshall W. Van Alstyne. 2005. Two-sided network effects: A theory of information product design. *Management Science* **51**(10) 1494–1504.

Petrongolo, Barbara, Christopher A. Pissarides. 2001. Looking into the black box: A survey of the matching function. *Journal of Economic Literature* **39**(2) 390–431.

Pissarides, Christopher A. 1990. *Equilibrium Unemployment Theory*. Blackwell, Oxford, UK.

Puhani, Patrick A. 2012. The treatment effect, the cross difference, and the interaction term in nonlinear difference-in-differences models. *Economics Letters* **115**(1) 85 − 87.

Rochet, Jean-Charles, Jean Tirole. 2003. Platform competition in two-sided markets. *Journal of the European Economic Association* **1**(4) 990–1029.

Romanyuk, Gleb. 2017. Ignorance is strength: Improving the performance of matching markets by limiting information. Working Paper, Harvard Univeristy, Cambridge, MA.

Tadelis, Steven. 2016. Reputation and feedback systems in online platform markets. *Annual Review of Economics* **8**(1) 321–340.

Talluri, K.T., G.J. van Ryzin. 2005. *The Theory and Practice of Revenue Management*. Springer, New York.

Tereyagoglu, Necati, Peter Fader, Senthil K. Veeraraghavan. 2017. Multi-attribute loss aversion and reference dependence: Evidence from the performing arts industry. *Management Science* Forthcoming.

Acknowledgments

We thank the corporate sponsor for making the data available.

Online Supplement

Appendix A: Definition of Live Listings

Given that the status of a listing changes dynamically over time, characterizing the live status of listings is a constant challenge for many peer-to-peer short-term rental platforms. A listing can become temporarily unavailable if the property is occupied for a certain time period during the year, regardless of whether the owner makes this information transparent on the listing. Also, a listing can become stale because of lack of activity—for example, if the owner does not maintain the listing or does not respond to inquiries from travelers. We define a listing as live at a historical date if one of the following four criteria is met: (1) at least one guest has inquired about the listing on that date or within three months prior to the date; (2) the owner has accepted at least one inquiry or taken at least one action in changing her calendar on that date or within three months prior to the date; (3) the date has been reserved by a traveler (not the owner); (4) the initial listed date is within three months of the date. We also vary these cutoffs from three months to six months, and the conclusions are consistent. Based on these procedures, we identify 41,651 live listings on August 18, 2014, just before the migration of listings took place.

Appendix B: Bias Correction For Censored Data

For a traveler i interested in trip r in destination m, let N_{irm} denote the number of inquiries he sent, and $Y_{irm} = 1$ if he is matched eventually, and 0 otherwise. For illustration, we let TrA_{irm} denote the interaction term $Tr_m \cdot A_{ir}$, and Z_{irm} denote the set of covariates, that is, $\{Tr_m, A_{ir}, X_{ir}, W_m\}$. Then Equation (1) can be simplified as

$$y_{irm} = f(\alpha + \beta TrA_{irm} + \theta Z_{irm} + \epsilon_{irm}).$$

The effect of the treatment equals $Pr[Y_{irm} = 1|TrA_{irm} = 1, Z_{irm}] - Pr[Y_{irm} = 1|TrA_{irm} = 0, Z_{irm}]$ with binary treatment indicator (Puhani 2012), or

$$\frac{\partial Pr[Y_{irm} = 1|TrA_{irm}, Z_{irm}]}{\partial TrA_{irm}}$$

in the case of continuous treatment levels.

Note that we do not observe when and where travelers intend to travel if they do not send out any inquiry to any listing owner. If we ignored these travelers, we would instead obtain only treatment effects conditional on $N_{irm} >= 1$, which would yield biased estimates for the unconditional effects that we intend to measure.

To obtain an unbiased estimate of the unconditional effect on matching probability, we note the following:

$$\begin{aligned} Pr[Y_{irm} = 1|TrA_{irm}, Z_{irm}] &= Pr[Y_{irm} = 1|N_{irm} >= 1, TrA_{irm}, Z_{irm}]Pr(N_{irm} >= 1|TrA_{irm}, Z_{irm}) \\ &+ Pr[Y_{irm} = 1|N_{irm} = 0, TrA_{irm}, Z_{irm}]Pr(N_{irm} = 0|TrA_{irm}, Z_{irm}) \\ &= Pr[Y_{irm} = 1|N_{irm} >= 1, TrA_{irm}, Z_{irm}]Pr(N_{irm} >= 1|TrA_{irm}, Z_{irm}). \end{aligned}$$

In the last row in the above equation, $Pr[Y_{irm} = 1|N_{irm} >= 1, TrA_{irm}, Z_{irm}]$ can be estimated directly from the data using a logit model. Meanwhile, $Pr(N_{irm} >= 1|TrA_{irm}, Z_{irm})$ can be calculated based on the estimates from zero-truncated Poisson regression.[13] Assume N_{irm} follows a Poisson distribution with rate λ_{irm}, where $\lambda_{irm} = \exp(\alpha^N + \beta^N TrA_{irm} + \theta^N Z_{irm})$. Then we have

$$Pr(N_{irm} >= 1|TrA_{irm}, Z_{irm}) = 1 - e^{-\lambda_{irm}} = 1 - e^{-\exp(\alpha^N + \beta^N TrA_{irm} + \theta^N Z_{irm})}.$$

With binary treatment indicator, the treatment effect can be estimated as follows:

$$\begin{aligned}
& Pr[Y_{irm} = 1|TrA_{irm} = 1, Z_{irm}] - Pr[Y_{irm} = 1|TrA_{irm} = 0, Z_{irm}] \\
= \; & Pr[Y_{irm} = 1|N_{irm} >= 1, TrA_{irm} = 1, Z_{irm}]Pr(N_{irm} >= 1|TrA_{irm} = 1, Z_{irm}) \\
& - Pr[Y_{irm} = 1|N_{irm} >= 1, TrA_{irm} = 0, Z_{irm}]Pr(N_{irm} >= 1|TrA_{irm} = 0, Z_{irm}).
\end{aligned}$$

For the purpose of interpretation, we calculate semi-elasticity, that is, percentage change in the outcome variable:

$$\frac{Pr[Y_{irm} = 1|TrA_{irm} = 1, Z_{irm}] - Pr[Y_{irm} = 1|TrA_{irm} = 0, Z_{irm}]}{Pr[Y_{irm} = 1|TrA_{irm} = 0, Z_{irm}]}.$$

With continuous treatment levels, the treatment effect can be estimated similarly. From the logit model, let us specify,

$$Pr[Y_{irm} = 1|N_{irm} >= 1, TrA_{irm}, Z_{irm}] = \frac{\exp(\alpha^Y + \beta^Y TrA_{irm} + \theta^Y Z_{irm})}{1 + \exp(\alpha^Y + \beta^Y TrA_{irm} + \theta^Y Z_{irm})} \equiv \rho_{irm}.$$

Then, the un-conditional effect of market density on matching outcome can be calculated as follows:

$$\begin{aligned}
& \frac{\partial Pr[Y_{irm} = 1|TrA_{irm}, Z_{irm}]}{\partial TrA_{irm}} \\
= \; & \frac{\partial Pr[Y_{irm} = 1|N_{irm} >= 1, TrA_{irm}, Z_{irm}]}{\partial TrA_{irm}} Pr(N_{irm} >= 1|TrA_{irm}, Z_{irm}) \\
& + Pr[Y_{irm} = 1|N_{irm} >= 1, TrA_{irm}, Z_{irm}] \frac{\partial Pr(N_{irm} >= 1|TrA_{irm}, Z_{irm})}{\partial TrA_{irm}} \\
= \; & \beta^Y \rho_{irm}(1 - \rho_{irm})\big(1 - e^{-\lambda_{irm}}\big) + \beta^N \rho_{irm} \lambda_{irm}^2 e^{-\lambda_{irm}} \\
= \; & \beta^Y \rho_{irm}(1 - \rho_{irm}) + \Big(\beta^N \lambda_{irm}^2 - \beta^Y (1 - \rho_{irm})\Big)\rho_{irm} e^{-\lambda_{irm}}.
\end{aligned}$$

In the last row in the above equation, the first term is the conditional marginal treatment effect on matching probability, and the second term is the correction term for censored data.

For the purpose of interpretation, we again calculate semi-elasticity, that is, percentage change in the outcome variable when the explanatory variable changes by one unit:

$$\begin{aligned}
& \frac{\partial Pr[Y_{irm} = 1|TrA_{irm}, Z_{irm}]/Pr[Y_{irm} = 1|TrA_{irm}, Z_{irm}]}{\partial TrA_{irm}} \\
= \; & \beta^Y (1 - \rho_{irm}) + \Big(\beta^N \lambda_{irm}^2 - \beta^Y (1 - \rho_{irm})\Big) e^{-\lambda_{irm}}.
\end{aligned}$$

Appendix C: Sample Markets and Changes in Thickness

Table A1 provides examples of markets that experience different changes in thickness. In particular, it shows the top fifty markets by size, the existing number of listings, and the number of new listings added during the consolidation period in each market. For example, Torquay Surfcoast (in Great Ocean Road), Fremantle (in Perch), and Inverloch (in Phillip Island and Gippsland) are three areas of similar size before the consolidation, with 268, 251, and 258 live listings, respectively. During the consolidation period, 69, 23, and 3 new listings were added to each area, corresponding to 25.7%, 9.2%, and 1.2% increase in the number of listings, respectively.

[13] To ensure that the results are not driven by the functional form of Poisson, we also conducted the analysis assuming negative binomial distribution, and the results are consistent.

Table A1 Top Fifty Markets by Size and Changes in Market Thickness

State	Region	Area	No. of Existing Listings	No. of New Listings	Increase in Thickness
QLD	Gold Coast	Surfers Paradise	396	184	High
VIC	Great Ocean Road	Torquay Surfcoast	268	69	High
VIC	Melbourne	Melbourne City	292	58	High
NSW	North Coast	Hawks Nest	223	56	High
NSW	North Coast	Port Macquarie	178	31	High
NSW	Northern Rivers and Byron Bay	Yamba	275	50	High
NSW	South Coast	Merimbula	182	60	High
NSW	Sydney	Bondi Beach	212	84	High
VIC	Bays and Peninsulas	Sorrento	385	8	Low
VIC	Bays and Peninsulas	Blairgowrie	274	13	Low
VIC	Bays and Peninsulas	Rye Bayside Beaches	241	14	Low
QLD	Cairns and Barrier Reef	Port Douglas	394	59	Low
VIC	Great Ocean Road	Apollo Bay	344	7	Low
VIC	Great Ocean Road	Lorne	340	8	Low
VIC	Macedon and SpaCountry	Daylesford	259	7	Low
NSW	North Coast	Forster	323	4	Low
NSW	Northern Rivers and Byron Bay	Byron Bay	669	75	Low
WA	Perth	Fremantle	251	23	Low
VIC	Phillip Island and Gippsland	Cowes	549	67	Low
NSW	Port Stephens	Nelson Bay	460	8	Low
QLD	Queensland Islands	Hamilton Island	360	36	Low
QLD	Queensland Islands	North Stradbroke Island	290	11	Low
NSW	South Coast	Mollymook	224	18	Low
QLD	Sunshine Coast	Noosa Heads	280	6	Low
VIC	Wineand High Country	Bright	231	16	Low
QLD	Cairns and Barrier Reef	Palm Cove	200	31	Low
QLD	Cairns and Barrier Reef	Cairns	180	26	Low
NSW	Central Coast	Avoca Beach	205	11	Low
SA	Fleurieu Peninsula	Victor Harbor	202	16	Low
QLD	Gold Coast	Broadbeach	193	29	Low
VIC	Great Ocean Road	Port Fairy	219	6	Low
VIC	Great Ocean Road	Ocean Grove	208	36	Low
SA	Limestone Coast	Robe	202	14	Low
NSW	North Coast	Coffs Harbour	203	34	Low
NT	Northern Territory	Darwin	208	33	Low
NSW	South Coast	Kiama	202	21	Low
WA	South West	Margaret River	222	5	Low
NSW	Sydney	Sydney City	207	13	Low
VIC	Bays and Peninsulas	Rye Ocean Bayside Beaches	236	2	None
NSW	Central Coast	Terrigal	210	2	None
VIC	Great Ocean Road	Anglesea	285	4	None
NSW	Hunter	Hunter Valley	196	2	None
NSW	Northern River sand Byron Bay	Kingscliff	207	3	None
VIC	Phillip Island and Gippsland	Inverloch	258	3	None
NSW	Snowy Mountains	Jindabyne	327	3	None
NSW	South Coast	Culburra Beach	217	2	None
NSW	South Coast	Vincentia	212	2	None
WA	South West	Dunsborough	211	4	None
QLD	Sunshine Coast	Sunshine Beach	235	1	None
NSW	Sydney	Manly Beach	247	2	None

Note. No. of existing listings refers to the number of listings that were live just before the migration took place. No. of new listings refers to the total number of listings joined during the four-day migration period.

Appendix D: Trip Definition

To characterize a trip, we first need to define a trip. The platform associates an inquiry with a traveler ID whenever possible, where a traveler is usually identified with an IP address or account log-in. About 99% of inquiries in our data were identified with a traveler ID. Traveler IDs alone do not allow us to link multiple inquiries for the same trip. With the traveler IDs and trip parameters, we are able to link multiple inquiries sent by the same traveler for the same trip, even if they were performed on different dates. It is important to accurately associate search sessions observed on different dates with the correct trip to measure the total search efforts exerted and corresponding matching outcome at the traveler–trip level.

We use destination, travel dates and search date to evaluate whether inquiries made by the same traveler were likely intended for the same trip. Inquiries associated with the same trip likely have the same destination, have the same check-in and check-out dates and are sent in adjacent time periods. However, some travelers may be flexible in the trip destination or travel dates. For example, a traveler may be interested in a weekend family getaway at one of the several beaches on the Gold Coast, and they may be flexible about whether to arrive on Friday night or Saturday morning. We follow the procedure described below to define trips based on geographical and temporal separation, while still allowing for some flexibility along both dimensions. We first sort inquiries sent by a traveler by region and inquiry date. We then define a new trip if at least one of the following criteria is met for an inquiry: (1) it is the first inquiry observed in all of the data for the traveler within the region; (2) the check-in date associated with the inquiry is more than one week apart (either before or after) from that of the previous inquiry sent by the same traveler to a listing in the same region; (3) the inquiry date is at least one month (thirty days) apart from the previous inquiry date by the same traveler to a listing in the same region. Once we define these criteria, length of stay does not vary as much, so we do not define an additional criterion based on check-out dates.

By focusing on inquiries sent to the same region rather than restricting them to the same area, we allow a certain level of flexibility in terms of the trip destination. By allowing a seven-day window between check-in dates, we allow flexibility in terms of the exact date of travel. By allowing a thirty-day window between inquiries, we allow travelers time to plan a trip and carefully consider options before making a commitment.[14] Table A2 presents a few examples illustrating these considerations. We thus associated 3.87 million inquiries with identified traveler IDs to 1.89 million trips. While each trip is associated with two inquiries on average, its distribution has a long tail.

Appendix E: Selection of Market Characteristics

Table A3 shows the results of t-tests of thirty-five different market characteristics (appearing in at least 0.1% of listings) in the two groups. Although markets in the treated and control groups are similar along many dimensions, they are different along a few. Based on the correlation Table A4, we further remove three variables (serviced apartment, wheelchair access, and short-term rental) because of high collinearity. We are then left with six key variables: number of listings in a market, number of rooms, percentage of executive listings, percentage of apartment listings, percentage of listings close to adventure, and percentage of listings close to national parks.

[14] These parameters are carefully chosen based on manual examination of inquiry histories and their associated parameters such as dates and destinations. To ensure robustness, we also tried restricting a trip to the same area instead of a region, a fourteen-day window between check-in dates, and a twenty-one- or a forty-day window between inquiries; all results are consistent.

Table A2 Examples of Trip Definitions

Traveler ID	Inquiry ID	Inquiry Date	Region	Area	Check-In	Check-Out	Trip ID
4	558518	28-Jun-14	SouthWest	MargaretRiver	21-Dec-14	24-Dec-14	1
4	424788	13-Sep-14	Perth	PerthCentral	24-Dec-14	29-Dec-14	2
4	172124	13-Sep-14	Perth	PerthCentral	24-Dec-14	29-Dec-14	2
4	648256	13-Sep-14	Perth	PerthCentral	24-Dec-14	29-Dec-14	2
8	783407	3-Jul-14	GoldCoast	SurfersParadise	21-Dec-14	28-Dec-14	3
8	232760	3-Jul-14	GoldCoast	SurfersParadise	21-Dec-14	28-Dec-14	3
8	860254	4-Jul-14	GoldCoast	GoldCoastCentral	24-Dec-14	28-Dec-14	3
8	140013	7-Jul-14	GoldCoast	SurfersParadise	21-Dec-14	28-Dec-14	3
23	29506	16-Jan-14	GreatSouthern	Albany	16-Mar-14	17-Mar-14	4
23	924595	16-Jan-14	SouthWest	MargaretRiver	14-Mar-14	16-Mar-14	5
23	704883	16-Jan-14	SouthWest	MargaretRiver	14-Mar-14	16-Mar-14	5
23	370824	21-Nov-14	SouthWest	Bunbury	12-Mar-15	16-Mar-15	6
27	788378	21-Mar-14	Sydney	Parramatta	25-Jun-14	4-Jul-14	7
27	114118	21-Mar-14	Sydney	Parramatta	25-Jun-14	4-Jul-14	7
27	193811	21-Mar-14	Sydney	Parramatta	25-Jun-14	4-Jul-14	7
27	791935	19-Jun-14	Sydney	Parramatta	27-Oct-14	6-Nov-14	8
27	567829	19-Jun-14	Sydney	Parramatta	27-Oct-14	6-Nov-14	8
32	638136	5-Jul-14	GreatOceanRoad	Lorne	6-Aug-14	7-Aug-14	9
32	564835	5-Jul-14	GreatOceanRoad	Lorne	6-Aug-14	7-Aug-14	9
32	862884	5-Jul-14	GreatOceanRoad	TorquaySurfcoast	6-Aug-14	7-Aug-14	9
32	815769	5-Jul-14	GreatOceanRoad	Lorne	6-Aug-14	7-Aug-14	9
32	830287	5-Jul-14	GreatOceanRoad	ApolloBay	6-Aug-14	7-Aug-14	9
32	299069	5-Jul-14	GreatOceanRoad	Warrnambool	6-Aug-14	7-Aug-14	9
32	726029	6-Jul-14	GreatOceanRoad	TorquaySurfcoast	5-Aug-14	7-Aug-14	9

Note. Traveler ID, Inquiry ID, and Trip ID are all de-identified.

Appendix F: Robustness to Definition of the Primary Market and Market Characteristics

Recall there are cases in which a traveler is interested in multiple destinations (markets) for a trip. In such cases, we define a primary market of interest and the associated market characteristics in our trip level analysis. Specifically, we define the primary market as the market first inquired by the traveler. Alternatively, one can also define the primary market as the most frequently inquired market by the traveler, and in case of a tie, the first inquired market is used. Lastly, instead of defining one primary market, one can also use the average market characteristics of all inquired markets in the analysis. The results of all three options are presented in Table A5. The results are consistent under different definitions. In the main analysis, we adopt the first approach because the first inquired market is more likely to be determined exogenously, while whether to search alternative markets can be an endogenous decision based on information received in the first market inquired. For example, with more options available in the primary market of interest, a traveler may be less likely to search other markets. We indeed find evidence consistent with this conjecture as shown in the table below.

Appendix G: Heterogeneous Effects of Market Thickness

This subsection first presents heterogeneous treatment effects across markets of different sizes. As shown in Table A6, the effect of an increase in market thickness on matching rate is more salient in larger markets than in smaller markets. This is likely because there are more outside options, such as hotels and motels, in larger markets than in smaller ones.

Table A3 Comparison of Market Characteristics between Treatment and Control Markets

Category	Variables	Control	Treatment	Diff.
Basic	Number of Listings	91.713	126.503	34.790***
	Number of Rooms	3.017	2.838	-0.179***
	Max. Number of Adults	3.72	3.634	-0.086
Property Type	Executive	0.178	0.228	0.050***
	Apartment	0.095	0.162	0.066***
	Bed and Breakfast	0.040	0.041	0.001
	Cottage/Lodge/Villa	0.147	0.156	0.009
	Farm	0.025	0.021	-0.003
	Hotel	0.011	0.015	0.004
	House	0.270	0.253	-0.017
	Resort	0.039	0.055	0.016
	Serviced Apartment	0.020	0.055	0.034***
Trip Purpose	Conference Corporate	0.041	0.054	0.013
	Family Holiday	0.382	0.384	0.002
	Romantic	0.299	0.317	0.017
	Weddings	0.090	0.094	0.004
Nearby Locations	Trail	0.117	0.113	-0.004
	Adventure	0.021	0.015	-0.006**
	Beach	0.365	0.332	-0.032
	Fishing	0.314	0.299	-0.015
	Golf Courses	0.300	0.320	0.02
	National Parks	0.291	0.249	-0.041*
	Nature	0.378	0.344	-0.034
	Ski	0.030	0.022	-0.009
	Surf Spots	0.242	0.205	-0.037
	Waterfront	0.157	0.159	0.002
	Wineries	0.185	0.191	0.006
Amenities	Spa	0.082	0.098	0.016
	Wheelchair Access	0.045	0.058	0.013**
Guest Accommodation	Child-friendly	0.300	0.306	0.006
	Gay-friendly	0.287	0.304	0.018
	Large Groups	0.112	0.109	-0.004
	Pet-friendly	0.159	0.145	-0.014
Other	Weekend Getaway	0.347	0.317	-0.03
	Short-term Rental	0.244	0.285	0.041**

Note. Diff. = Treatment - Control. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table A4 Correlation Table of Key Market Characteristics

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	Number of Listings	1.000								
(2)	Number of Rooms	0.009	1.000							
(3)	Executive	-0.030	-0.026	1.000						
(4)	Apartment	0.038	-0.477	0.535	1.000					
(5)	Serviced Apartment	-0.030	-0.320	0.456	0.677	1.000				
(6)	Adventure	0.013	0.169	0.281	0.019	-0.035	1.000			
(7)	National Parks	-0.105	0.408	0.343	-0.117	-0.152	0.306	1.000		
(8)	Wheelchair Access	-0.001	0.043	0.669	0.359	0.314	0.254	0.270	1.000	
(9)	Short-term Rental	-0.131	0.027	0.811	0.479	0.397	0.151	0.460	0.544	1.000

We also examine how the occupancy rate changes for popular listings as a result of the market thickness increase. We measure popularity using total number of inquiries, total number of bookings, total number of inquired days and total number of booked days during the before period, and define popular listings using the upper quartiles of these measures. The results are presented in Table A7. While the average listing occupancy rate was lower due to the increased market thickness, occupancy rate for those popular listings did not experience statistically significant change.

Table A5 Effect of Market Thickness on Traveler Matching Rate

	First Inquired		Most Inquired		Average Mkt Characteristics	
	All TVLRs (1)	Existing TVLRs (2)	All TVLRs (3)	Existing TVLRs (4)	All TVLRs (5)	Existing TVLRs (6)
Percent Increase in Thickness	-0.192***	-0.129*	-0.121*	-0.094	-0.143**	-0.123**
	(0.060)	(0.069)	(0.063)	(0.077)	(0.073)	(0.062)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	367672	288824	369395	290547	367672	288824
LogL	-149406.1	-120595.9	-149997.0	-121195.8	-149215.2	-120408.9

Note. This table shows the robustness of the results to different definitions of primary market and market characteristics. Columns (1) and (2) define the primary market as the market first inquired by a traveler, while Columns (3) and (4) define the primary market the as most inquired market by a traveler. In case of a tie, the primary market is the first inquired market among the most inquired ones. Travelers are included in the analysis if their primary destinations are among the markets under study. Therefore, the number of observations can vary slightly under different definitions of primary market. Columns (5) and (6) do not use the characteristics of the average market as controls rather than those of the primary market. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table A6 Heterogeneous Effects of Market Thickness by Market Size

	Traveler Confirmation Rate	Listing Occupancy Rate
	Existing TVLRs (1)	Existing LSTGs (2)
Percent Increase in Thickness	-0.026	-0.131
	(0.082)	(0.082)
Percent Increase in Thickness × Market Size	-0.322***	-0.144*
	(0.012)	(0.077)
Controls	Yes	Yes
N	350996	946798
LogL	-146134.8	-207170.7

Note. Column (1) reports average semi-elasticity on treated markets during the after period. The estimates correct for data censoring. Standard errors are obtained through bootstrapping. Controls in Column (1) include main effects and trip, traveler, and market characteristics. Column (2) reports coefficient estimates form fixed-effect logit model. Controls in Column (2) include main effects, listing fixed effects, and travel-date characteristics. Market size is log-scaled and centered. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Table A7 Effect of Market Thickness on Occupancy Rate for Popular Listings

	(1)	(2)	(3)	(4)
Percent Increase in Thickness	0.031	-0.029	0.007	-0.022
	(0.098)	(0.093)	(0.123)	(0.116)
Controls	Yes	Yes	Yes	Yes
N	412638	572437	384554	407391
LogL	-118002.8	-149596.7	-101052.0	-106795.9

Note. Analyses are conducted on existing listings only. Columns (1) to (4) define popular listings as upper quartiles based on total number of inquires, total number of bookings, total number of inquired days and total number of booked days during the before period, respectively. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.