NET Institute*

www.NETinst.org

Working Paper #12-24

September 2012

Do Online Marketplaces Become More Efficient Over Time?

Andrey Fradkin
Department of Economics
Stanford University

* The Networks, Electronic Commerce, and Telecommunications (“NET”) Institute, <http://www.NETinst.org>, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, “virtual networks” comprised of computers that share the same technical standard or operating system, and on network issues in general.

Do Online Marketplaces Become More Efficient Over Time? (Preliminary Do Not Cite)

Andrey Fradkin*

October 15, 2012

Abstract

An increasing proportion of transactions in two-sided markets are being mediated by online platforms. Presumably, agents choose to use online platforms because they have a lower transaction cost technology compared to alternatives. I use data from a growing online platform that matches travelers and hosts to study matching and transaction costs on online platforms. I show that the matching probability for guests has increased by 18% over a span of two years on the platform. I then show that the increase in efficiency holds even when controlling for the search intensity of guests, the change in the composition of transactions and aggregate market conditions at the time of search I demonstrate that guests are elastic with respect to the time it takes to make a transaction. I then investigate one potential reason for why guests are more successful over time: a decrease in the transaction costs required to book. I show that the time to book and the amount of communication required to book on the platform are falling over time. I then consider three possible explanation hypotheses for the reduction in required transaction costs: learning, reputation building

*Department of Economics, Stanford University. The author would like to thank Jon Levin, Liran Einav and Caroline Hoxby for invaluable advice. Financial support was provided by the NET Institute (www.NETinst.org) 2012 Summer Research Grant as well as the Shultz Graduate Fellowship. Data access was generously provided by Airbnb. The author was employed as a data scientist by Airbnb during part of the time that this research was conducted.

and platform policy. I show that all three are likely important for explain the increased efficiency over time.

1 Introduction

There has been an explosion in the number of transactions mediated by online platforms. Transactions in labor, housing, used goods, insurance and travel markets have all in some part shifted to online marketplaces. The above markets are typically characterized by frictions. Those frictions cause agents to incur transaction costs, suboptimal matching of agents and equilibrium vacancies. Presumably, agents choose to use online platforms to transact because online platforms have lower transaction costs than other options for completing the same transaction. Successful platforms, much like any other institutions, should have lower transaction costs compared to their alternatives[16]. The growth of online platforms raises the possibility that search markets may be made significantly more efficient with the proper institutions and technology. The goal of the paper is to document and explain the trajectory of transaction costs over the lifetime of a platform as a function platform innovations and endogenous agent behavior.

Economists have typically modeled platforms as passive actors that merely set prices and provide a centralized place for different types of agents to meet[19, 20, 21].

In the standard conception of a platform, the main transaction cost is a coordination friction. Agents choose where to transact conditional on the types of users that are present on a particular platform and the fees that the platform charges. Alternatively, instead of fees, platforms can manipulate other instruments in order to select their user base [22, 23, 10].

The view of platforms as manipulators of pre-defined instruments rather than as innovators is at odds with the fact that online platforms often invest heavily in technology and enact frequent policy changes [5]. Platforms develop matching technologies, devise rules for agent behavior, and provide auxiliary services to agents in their transactions. Platform investment in technology and policy might provide huge returns. The existence of an on-line listing service may simply be enough to generate an active marketplace. Since platforms tend to change many aspects of their site over time, I focus on measuring and describing the trajectory of matching probabilities and transaction costs rather than estimating the causal effect of a specific

change.

I use data from Airbnb, a website that connects people who have available space with travelers who are looking for a place to stay. Airbnb is the leading platform for renting out spare rooms and apartments in cities. The type of transaction that Airbnb is involved in is not new. For example, Craigslist has a listing service for short-term stays and Couchsurfing has facilitated stays between strangers since 2003. Further, the market for vacation home rentals and timeshares has existed for a long time. Nonetheless, Airbnb has experienced exponential growth since its founding in 2008 (Figure 1).

Figure 1: Number of Transactions over Time

The rapid growth of Airbnb suggests that it is a superior platform to transact on compared to alternatives. Transaction costs exist for both sides of the market: guests and hosts. I focus on the transaction costs faced by potential guests on the market. Guests face the cost of the time involved in

finding suitable rooms and the effort required in communication with hosts¹. Those costs can be wasted because there is a non-trivial chance that a transaction does not happen.

2 Platform Background

Airbnb describes itself as “a trusted community marketplace for people to list, discover, and book unique accommodations around the world – online or from a mobile phone².” Airbnb makes revenue by taking a percentage fee of every transaction that takes place on the platform. Hosts typically have a 3% fee while guests have a variable fee that ranges between 6% and 12% depending on the details of the transaction. Almost since the beginning of the website, it had a search engine, a messaging system, a review system for hosts, and a payments mechanism. The payments mechanism was an especially large deviation from previous platforms such as Craigslist, where payments were handled offsite. The fact that Airbnb holds payments in escrow until the stay of the trip made agents on the site a lot less weary of being scammed. If a dispute arose, Airbnb was in a unique position to resolve it. Furthermore, security deposits and cancelation policies could be force.

Airbnb has undergone tremendous changes in technology, policy and website design over its existence (Screenshots of the booking process are in section 9). The information displayed on each hosting page has increased. Specifically, the site has become a lot more comprehensive with regards to displaying amenities, house rules and neighborhood details. Airbnb has verified locations by providing hosts the option of having their space photographed by a professional photographer. The platform has also made the process of managing pricing and availability calendars easier for hosts. Hosts can set date specific prices, offer special deals, stipulate security deposits and offer cancelation policies. Each host’s response rate, last calendar update and average response time for messages is now displayed on the site. Encouraging hosts to accurately display their availability is especially important for Airbnb because it only wants to show searchers the properties that are available for the times of the stay. It is difficult to incentivize hosts to update calendar

¹There are physical transaction costs such as key exchange which are not the focus of the paper.

²<https://www.airbnb.com/home/press>

availability because hosts want to see all of the potential deals they can get, even if the room is likely unavailable. Further, hosts often do not want to expend the effort of keeping an up to date calendar.

In addition to each room having a dedicated page, each host also has a page. Host pages display other properties that the host manages, and reviews that the host has received while being a guest. Furthermore, hosts have the option of having their phone number verified by Airbnb. Both guests and hosts can connect their accounts to Facebook and other services. Airbnb then displays any common social or college connections between the two users on the search page. Common social connections or group affiliations can help reduce problems of informational asymmetry [9, 11].

The search processes has been undergone a series of changes. Firstly, Airbnb has frequently refined the search algorithm over the lifespan of the site. Modification of search display order has the the potential to greatly affect welfare [13]. For example Jeziorski and Segal 2009 show that simply altering the order of display ads can improve user welfare by 16%. Secondly, additional filters and better neighborhood definitions have been added to the search page. Lastly, users who never respond or behave badly are removed from search altogether. For searches that occur very close to the trip date, alternative search interfaces called “standby” and “Air Match” were developed that make it easier for guests to get quick responses.

The messaging process has been enhanced through added text messaging ability and through a mobile phone app. One of the main advantages of a mobile app is that hosts can easily see the characteristics of their guests and check the availability of their calendar. Further, because hosts and guests might speak different languages, the platform has worked hard to develop and deploy effective, automatic translation solutions. The platform has added default responses that decrease the need to write personalized responses. For example, a host can “pre-approve” in response to a guest’s message. A pre-approval allows a guest to book a room without any additional delay. Airbnb also allows some select hosts can elect to allow guests to “instant book” rooms without sending a message first.

Lastly, the platform is involved in a lot of offline activity complementary to its online efforts. The platform has a large, global customer support and concierge staff that assists users whenever they face problems. Furthermore, Airbnb contacts hosts to instruct them how to make more effective hosting pages and set better prices. The platform also offers a one million dollar insurance guarantee against major damages done to their property by guests.

3 The Search Process

3.1 What Are the Frictions in the Market?

The key friction that a guest faces is the lack of knowledge about the characteristics and availability of the properties in each market. Each match between guest and host on Airbnb is idiosyncratic. A room is a unique and discrete good that can only be occupied by one guest at a time. Furthermore, each host is a unique person that may be easy or difficult to interact with. Each room/host combination is heterogeneous on many potentially important vertical and horizontal characteristics (i.e. price, availability, location, room size, room type, wireless internet, reviews, gender, age, and interests). Hosts face a similar problem in their lack of knowledge about the characteristics of potential guests. Hosts not only have to interact with guests but also have to worry about potential damage or theft to their property. Like hosts, guests are heterogeneous on both easily observable characteristics such as gender and difficult to observe characteristics such as trustworthiness. The preferences of hosts add an additional level of uncertainty for guests because guests do not know which hosts will like them. The process of acquiring the information required to make a choice and booking a room is costly. The major non-monetary costs that guests face are the time spent waiting for a booking to happen and the effort required to search and communicate.

3.2 The Conversion Flow

There is a series of discrete steps in the process of booking a room.

1. Structured Search - In this step the potential guests enters the travel dates, number of guests and location into a search engine and receives a list of results. The search can then be refined.
2. Investigation - The potential guest then clicks on individual properties and host pages to learn about property characteristics such as amenities, reviews and house rules.
3. Message - If the potential guest likes a room, the guest sends a message to the host inquiring about room details and availability. The process of messaging typically takes time and several rounds of back and forth. Guests often send messages to multiple hosts at the same time.

4. Booking - If a transaction is agreed upon in messaging, the guest then clicks the “Book It” button. If the host accepts, the money is charged and taken in escrow by the platform. Some transactions occur without a messaging step if the host allows the “Instant Book” or “Book It” features.
5. Stay - After a trip is booked, there is further coordination needed to exchange keys and coordinate the details of the trip. Further, either side can cancel with a variable penalty rate.

At each of the above steps, the potential transaction may fail. A guest considering booking on Airbnb weighs the possibility of failure at each of the above steps and the realized cost of time and effort involved in making a transaction. The probability of failure and the cost of search can vary with the type of stay that a guest is looking for. For example, longer stays require a lot more commitment by both sides of the transaction. Therefore, each side requires more knowledge about the other. Such knowledge is typically obtained through more extensive communication. Transactions between parties who speak different languages may require more effort and may be more uncertain in ex-ante utility. Another example of a factor that affects transaction costs is the season of the stay. Hosts may be more likely to be available in popular seasons because they themselves are going on vacation. Similarly, market tightness varies over time due to seasonality of demand and platform advertisement. As demand increases relative to supply, the chance of messaging an unavailable host increases. Further, because multiple communications are sent at the same time, there is the potential of congestion externalities[2, 3, 8]. First, the probability of a simultaneous attempts to book the same room increases. Second, as “good” properties are booked, the remaining properties left may be adversely selected on transaction costs. For example, a property may be unavailable but may not have an updated calendar. As properties are booked approaching a trip date, a guest has an increased probability of messaging a host whose property looks good but is actually unavailable.

4 Data

The optimal dataset for studying transaction costs would contain information about every single action hosts and guests take with regards to AirBnB.

However, data on user clicks is not available historically. Instead, I focus on the subset of actions that have been measured since close to the start of the platforms existence: messages, threads and reservations. For each communication I observe the time a message was sent, who sent the message, the number of words in the message and any pre-defined templates that were associated with the message³. Each set of messages between a guest and a host composes a thread. A thread contains information about trip dates, the number of guests, the room of inquiry and a link to a reservation if a reservation resulted from a thread. There are two shortcomings with regards to the thread data. First, before late 2009 guests were not required to enter trip dates when sending an inquiry. Second, the link between a reservation and a thread is sometimes overwritten when a guest makes a reservation with the same host again. A reservation observation contains data on the transaction details such as the the time the reservation was submitted, the time it was approved (if approved), the time it was canceled (if canceled), the price paid, the trip dates and the number of guests.

Each host, room and guest also have data associated with them. For hosts and guests, the data includes demographics, language spoken and region of residence for a part of the sample. For rooms, the data includes all of the characteristics that a guest would see on the site. For example, it includes room type, number of bedrooms, the latitude and longitude, number of reviews, number of previous stays, the maximum room capacity and other details. Some of the data for the rooms is available historically. In order to compose my dataset, I aggregate individual message data such as word count and link it to threads. I then link each thread to a reservation. Lastly, I link each thread-reservation observation to the host, guest and room characteristics.

There are two types of direct proxies for transaction costs that I observe in the data: time variables such as response time and communication variables such as the number of words in a message. I measure transaction costs within two different contexts: a guest-host interaction and a search spell. A guest host interaction consists of a message or booking request and all subsequent interactions between that guest and host. Each search spell is composed of a set of actions by a potential guest in an attempt to find a stay in a particular market for a specific set of dates. A search spell begins with the

³Pre-defined templates were introduced in late 2011 and specify whether either a canned response(i.e. “Not Available”) or a mechanism of response (i.e. iPhone).

first contact between a guest and a host in a specific market for a specific week during which the trip would begin⁴. Each spell is associated with trip specific characteristics X and transaction cost related outcomes Y . The spell ends with a booked trip or an exit from the market. Although the platform has existed since 2008, I use data only beginning in 2010. Before 2010, users who sent messages were not required to enter check in and check out dates when sending a message to a host. Therefore, I cannot accurately define a “trip” that a user is likely to take from the structured message data. Details on how each search spell is constructed are discussed in Appendix A. Due to host experience, language barriers and market characteristics, transaction costs vary greatly by city. All results below are for City A unless otherwise noted. City A is an US city that is one of the largest and longest lived markets on the platform.

5 The Probability of Booking

The first step that a guest takes in the search for a room is using the search engine and browsing properties. I do not observe the first step of the search. However, I do observe all subsequent steps. figure 2 displays the overall probability of booking a room for any given search spell over time. The probability of success increases more than 10% over the period of the sample. The large increase in the probability of booking is surprising because all of the potential guests in the sample already have an intent to book. There are several hypotheses about why success rates improve over time on the platform other than an increase in the value of the match. Firstly, the composition of trip characteristics may have changed over time towards those that have a higher probability of success. Second, guests may have changed their search behavior in a way that makes successful matches more likely. Lastly, market conditions may have changed to make matches more likely.

The three most relevant aspects of a trip search are the number of nights,

⁴There are many cases in which a guest is searching for dates that are flexible on either end of the trip. Further, guests have the ability to stay with multiple hosts during the same trip. In those cases, the above definition of a trip spell will overestimate the number of potential trips. Alternatively, guests may be flexible with regards to other aspects of the trip such as the number of guests. In that case a search by one individual may actually also be a search for several individuals who could also search on their own. In that case, I may underestimate success percent because a combined trip with multiple guests may subsume several trips.

the number of guests and the time in advance of the stay in which the contact is made. I compute the Herfindahl index (HHI) for nights, weeks ahead and number of guests amongst search spells each month. figure 4 shows that the concentration of trip characteristics in the market decreases drastically over time. The decrease in contract type concentration suggests that more complex types of stays are being handled on the site more frequently.

In order to understand the impact of trip type on the probability of success, I estimate a linear regression controlling for trip type characteristics.

$$S_j = \beta X_j + \gamma t(j) + M_j + \epsilon_{jt} \quad (1)$$

S_{jt} is an indicator of a successful booking for trip j at time t . X_t is a vector of trip characteristics and M_t are dummies for the month of the year. Table 2 displays the results of the regression. Consistent with the hypothesis that more difficult to arrange transactions are being handled on the site, the overall trend in success rate increases from .59% per month to .62% per month. Further, search spells requiring more nights, more guests and those that are further ahead are all associated with lower success probabilities.

Another possibility for the increase in search trend is that search intensity increased. Search intensity can increase for various reasons, including the utility of potential matches, a decrease in search costs and learning about the returns to search⁵. I observe the number of contacts initiated by a searcher in the first 6 hours of the search spell. Approximately 80% of search spells are initiated with only 1 contact but there is a long tail of contacts. The more contacts a guest makes, the higher probability that at least one of those contacts is available and good enough to book. Second, I observe the number of words in the first message sent to a host. The number of words could indicate either more detail expressed to a host about the trip or more questions asked to the host about the property. One of the difficulties in the data is that I don't observe what the host's message says. However, I do observe when a host allowed a guest to book immediately without further communication and when a host responded with a "not available" template⁶. In

⁵The platform actively encourages guests to describe trip characteristics and to send multiple messages.

⁶A response of not available does not necessarily mean that a host is not available. The host may simply tag a message as not available because he does not like the guest. Further, a lot of rejection messages are not labeled as "not available". Lastly, the tagging system of responses was only enabled in 2012. A host cannot respond with a template if they are using a mobile phone.

order to measure the effect of word length on pre-approval and not available messages I run the regression (2).

$$A_{jh} = \beta X_j + \alpha(L)_j + \gamma(h) + M_j + \epsilon_{jh} \quad (2)$$

A_{jt} is an indicator of a pre-approval for trip j at time t or NA_{jt} is an indicator for a response with a not available tag, X_j is a vector of trip characteristics, $\alpha(L)$ are dummies for bins of message lengths, $\gamma(h)$ are host fixed effects and M_t are dummies for the month of the year. Table ?? displays the results of equation (2). The number of words in the first message is positively associated with the pre-approval of a trip and negatively associated with a response of non-available. The overall rate of pre-approval in the sample is 13%. The estimated negative change of 2% associated with having less than 15 words represents a 15% reduction in the probability of pre-approval. The regression results cannot be interpreted as causal because individuals who send short messages may be systematically different than those individuals who send longer messages. Nonetheless, the evidence is suggestive that word length matters for acceptance and rejection.

Table 4 column 1 demonstrates the effect of mean first word length and the initial number of contacts on the success probability of a search spell. The effects of both variables are economically large and significant although not necessarily causal. The search intensity may be correlated with intent to book. That is, those that derive the most value from booking on Airbnb, search the most. Regardless, the overall trend in success probability remains similar in magnitude and statistically significant.

Market factors may also be important for the success of a search spell. One way way in which market factors could affect success probability is that there are many searchers seeking to book the same rooms. I proxy for the probability of overlap by taking the ratio of the searchers searching in a particular week to the number of unique properties contacted by those searchers. The ratio of searchers to properties corresponds to the traditional “tightness” measure used in search models. Column 2 presents the results for the probability of success controlling for a measure tightness based on searches from one week for a specific trip week. Column 3 uses an alternate measure of tightness that uses the total searchers (across all search weeks) for a particular trip week. The trend in success probability remains of a similar magnitude in both cases. Another reason why market conditions may matter is that the more demand there is, the less good properties are available. In Column 4,

I interact the overall tightness measure with a polynomial of the number of weeks in advance of the trip that the contact is made. The reasoning is that when the market is especially tight, very few available or suitable properties will be available very close to the trip date. The coefficients on the tightness interactions confirm the theory. The probability of success is especially low if the contact is close to the trip week when markets are tight. Nonetheless, the overall trend in the success rate remains similar.

The above section has demonstrated that there is a significant trend in success probability. The lowest estimate of the trend in the estimates is .55% per month. That trend corresponds to a 15.4% increase in the probability of a search spell resulting in booking. The trend remains even after controlling for trip characteristics, search intensity and market conditions. The large increase in matching probability suggests that the platform has become a lot more efficient over time. The following sections explore possible reasons for the increase in efficiency.

6 Transaction Costs

There are two types of transaction costs involved in the observed search process: time and communication. If a search requires less communication and less time, then guests should be more likely to end up booking. Many of the changes that Airbnb has enacted are specifically targeted to reduce the above costs. For example, Airbnb has standardized information on the property, host and guest pages. The more information on the pages, the less information needs to be communicated in a message. Further, the more accurate the information is on the hosting page, the less likely the guest is to contact an unsuitable host. Similarly, Airbnb has also made it easier to communicate quickly by offering mobile apps and pre-defined responses. Below, I provide evidence that guests are elastic with respect to transaction costs. I then show that the transaction costs required to book have, on average, decreased over time. Lastly, I show that guests are now more likely to continue their search after an initial set of unsuitable matches.

The optimal experiment to identify a guest's elasticity for demand with respect to the time required to book is to randomly assign a response time of an acceptance to each guest sending a message. There is no such experiment in the data. However, there is a set of messages by guests that received a pre-approval from a host. A pre-approval allows a guest to book with one

click of a button. Some pre-approvals come quickly whereas others take a while to come in. figure 1 displays a histogram of the response times for pre-approvals in the sample. Most pre-approvals come within a span of 24 hours but over 25% come in on the second day and over 8% come in afterward. I estimate the effect of response time by estimating:

$$B_{jh} = \beta X_j + \mu(RT_{jh}) + \gamma(h) + M_j + \epsilon_{jh} \quad (3)$$

B_{jt} is the guest's decision to book a pre-approval, X_j are the trip characteristics, $\mu(RT)_{jh}$ are indicators for response time bins, $\gamma(h)$ are host-specific fixed effects, and M_j are month fixed effects. The results of equation (3) are displayed in Table ???. The baseline specification 1 uses the sample of all pre-approvals and does not include hosting level fixed effects. The probability of a guest accepting the booking decreases by over 10% for response that take between 24 to 48 hours as opposed to less than 6 hours. An alternative reason that longer responses yield less bookings is that guests are booking another room on the site instead. Specification 2 limits the sample to only those guests that send 1 initial message (about 50% of the sample of pre-approvals). The selected sample is likely to understate the true elasticity of booking with regards to time to book because those guests that really want to book quickly would send multiple messages. Nonetheless, the results are quantitatively similar. There is still a very large change in booking probability, the longer it took for the host to respond. Another concern about using response times is that properties that take longer to respond are worse. If longer response properties are worse, then guests would be less likely to book them even if they responded with a pre-approval. Alternatively, less motivated guests may message properties that take longer to respond. Specifications 3 and 4 include hosting level fixed effects. The effects of response time remain similar in magnitude and statistically significant. The large effect of response times on booking probability suggests a possible explanation for the increase in search success over time: that the time to book has decreased on the site.

There are four observed measures of time costs of search: no response, response time, time to book and trip time to book. In approximately 20% of cases, the host does not respond to the guest. When the host does respond, the median response time is 6.7 hours while the mean is 53 hours. The median response time conditional on a booking is 2.7 hours with a mean of 12 hours. Although most responses are quick, there is a long tail of long

responses.

I estimate the equation below in order to isolate the time trend from other changes on the platform.

$$TC_{jh} = \beta X_j + \alpha S_j + M_j + \epsilon_{jh} \quad (4)$$

Where TC_{jh} is a measure of transaction costs and S_j is a measure of the search intensity associated with trip j . ⁷ estimates the trends in time costs of search controlling for trip characteristics. The probability of no response is decreasing over time while the response time is increasing. The most likely theory for the two facts is that hosts who were not responding are now taking a long time to respond. The platform has begun to punish properties that do not respond by demoting them in search and sometimes removing them from search altogether. Therefore, hosts now make sure to respond even if it takes a long time and even if the answer to the guest is no.

Column 3 shows the results for the time it takes to book a room by a guest. The log transformation is more natural in this setting because response times seem to be exponentially distributed ⁷. The table shows that the time to book a room decreases at an average of 3.3% percent per month. Column 4 isolates the trend for the overall time to book for a guest. The decrease in time to book for a specific trip is 2.3% per month. The fact that guests have a large elasticity with respect to response times combined with the large decrease in response times over time suggests that a reduction in transaction costs could explain part of the increase in success probability.

It is worth noting that in many cases the guest does not book even if there a message and a quick pre-approval. The booking rate for guests who send one initial message and get a pre-approval in less than 6 hours is only 51%. Furthermore, only 16% of those guests who do not book after a quick pre-approval eventually book another room. Therefore, guests face significant uncertainty in their decision to book a room even if they've sent a message. One way to resolve the uncertainty about the utility of the room is to communicate with the host [6].

I observe four measures of the amount of communication required to book: the number of words sent to a host, the average word length, the number of

⁷In some cases, there are multiple trips being planned in the same message thread. In that case, it is difficult to classify whether the booked trip was actually the one that was being planned at the time of the first message. Therefore, there may be some classification errors.

distinct messages sent to a host before a booking and the number of words in the host’s response message. As [18] shows, it is generally not in the interest of the seller to disclose all information about the room. The platform has the ability to incentivize hosts to reveal information by manipulating the search algorithm. As more information is standardized on the website, the less information about the host should need to be communicated. The summary statistics for the amount communication is shown in Table 8. There are three overall trends to notice. First, the number of first words is increasing over time both for booked and not booked attempts. Second, the total number of words required to book is decreasing over time. Lastly, for communications that do not result in a book, more messages and words are sent over time. Table 9 shows that for those communications that end up in a booking, the total number of words and the number of messages decrease over time when controlling for trip characteristics. Furthermore, the host’s response contains fewer words on average over time. Presumably, as more information is displayed on the page, the host needs to respond with fewer details. Specification 5 controls for the length of the guest’s message. If a guest asks for more information, then the guest’s message should be longer. As expected, host’s responses are longer, the longer the guest’s initial message is. Nonetheless, there is a trend of -6.5% responded words per month for those threads that ended up being bookings. Table 10 shows that for the communications that do not end up in a booking, more messages are sent by the guest over time. The increased amount of messages for failed attempts suggests that guests are quitting later in the booking process on average. Lastly, host’s responses for failed attempts are decreasing in word length. It is worth investigating whether the trend is due to a decrease in information communicated, due to host learning or due to the standardization of language on the site.

The above section has demonstrated evidence that potential guests care about transaction costs and that the required transaction costs to book are decreasing over time. Another reason that trip success probability could go up is that after an initially unfavorable match, guests are more likely to continue their search now as opposed to in the past. Consider the following scenario. A guest sends n messages at time t . After waiting some time, the guest either does not hear a response or learns that the messages were sent to unsuitable hosts. The guest can then choose to continue the search or leave the site. I create a measure of continued search by counting the contacts that were sent by guests more than six hours after the first distinct contact. I limit the sample to those that do not book with a host they

initially sent a message to. The average number of contacts initiated during continued search almost doubles from an average of .78 to 1.47 between 2010 and 2012. Furthermore, the share of search spells that have at least one continued message increases from .27% to .4% between 2010 and 2012. The increase in continued search suggests that guests are finding it worthwhile to continue their search after an initial set of unsuitable matches. If the utility from sending a message increases then we should expect the probability of matching to increase as well. Indeed, the overall probability of matching for guests who do not succeed in the initial set of messages increases from 11% to 20% over the span of the sample.

In Table 11, I investigate whether the additional success is due to sending more contacts by a specific guest or due to a higher chance of any given continued contact succeeding by estimating equation (2) for the subsample of continuing searchers. Specification 2 controls for the number of continued contacts while specification 1 does not. Controlling for contacts, reduces the monthly increase in matching probability by a little under 50% from .46% to .28%. The matching probability per contact increases even after controlling for continued search intensity.

The above section has demonstrated that potential guests care greatly about transaction costs and that the platform has somehow managed to reduce the transaction costs required to book significantly from 2010-2012. Furthermore, guests have responded to the reduction in transaction costs by searching more intensively, even if their initial did not succeed. The next section will give descriptive evidence for the possible causes of the reduction in transaction costs.

7 Reputation and Learning to Reduce Transaction Costs

The time to book and the amount of communication required to book can fall over time for several reasons. One is that hosts are learning to screen guests, to include necessary information in their profiles and to respond quickly[4]. Another possible reason for the decrease in time to book is that guests face less uncertainty about hosts because hosts have accumulated reviews. The reputation channel has already been documented as important for platform growth on Odesk and Ebay [17, 7]. Lastly, the platform may have reduced

search costs through strategic choice of how to display information, technological improvements, certification and regulation [5, 1, 14].

Below, I show that all three of the above stories are probably important in explaining the decrease in transaction costs required to book. The baseline specification estimates the trend of the time to book over time controlling for trip characteristics and market conditions.

$$TC_{jh} = \beta X_j + M_j + \gamma(h) + \tau E_h + \epsilon_{jh} \quad (5)$$

The identification strategy for separating the effects of experience from reputation uses the fact that some hosts manage multiple properties. Each property is associated with a specific number of reviews at the time of the message. I proxy for the number of reviews with the number of previous bookings. Similarly, each host is associated with a prior number of bookings. In equation (5) I denote $\gamma(h)$ as either host fixed effects or hosting fixed effects and E_h as either the experience of the host or the property. The results of the estimation are displayed in Table 12. Specification 2 includes host level fixed effects. It shows that the overall decrease in book time within a host is approximately the same as the site average. The decrease in book time is not driven by selection onto hosts. Specification 3 includes the previous number of bookings for the host as a measure of the experience and reputation of the host. As expected, more experience results in quicker booking times. Nonetheless, there remains a significant trend of -2% in book time every month. Specifications 4 and 5 show analogous results for property level fixed effects. Accounting for property effects diminishes the time trend to a 1% decrease in time to book each month. Further, host experience is no longer significant when hosting level experience is used. The above results suggest that while selection onto lower transaction cost properties is definitely happening over the time period, transaction costs are decreasing in general. The more bookings properties have, the lower their transaction costs are. The fall in transaction costs could either be because of hosts learning about how to manage the specific property or because reviews demonstrated on the site lead guest to book quicker.

Table 13 restricts the sample to just those bookings that occurred within the first month of the existence of a property. New properties take less time to book over the sample period, even controlling for host experience. Specification 3 restricts the sample to just those properties whose host has no prior experience. The time to book for those properties decreases at a rate

of 2.4%. Since new hosts haven't had time to learn or build reputation, they must be influenced by some platform policy or technology that is changing over time.

8 Discussion

Two-sided search markets (such as the labor and housing market) play a prominent role in economics both as stand alone subjects of study and as inputs into macroeconomic models. Typically, when these markets are studied, the so-called matching technology “A” remains constant. Therefore, changes in equilibrium outcomes are attributed to either productivity or normal shocks. However, the increasing prominence of online platforms in those markets may also be contributing to changes in equilibrium outcomes. In fact, platforms have the ability and incentive to use non-price data in order to make matching more efficient. I show that, at least for the travel market, the assumption of a constant matching technology is unreasonable.

Specifically, I document a large positive trend in the probability that guests who contact a host find a room. Overall, the matching probability increases by 18% from early 2010 to mid 2012. I then investigate one possible reason for the increase in matching probability: a decrease in transaction costs required to book. First, I show that guests are highly elastic in their booking decisions with regards to host response times. A response time between 24 and 48 hours is associated with a 10% lower chance of booking than a response time of 6 hours or less. Then, I show that the time required to book and the amount of communication required to book falls over time. The decrease in transaction costs causes guests to search more intensively and for longer durations in 2012 as opposed to 2010. Lastly, I try to discover the reason for the trend in required transaction costs over time. I find evidence supporting the importance of both platform technology/policy as well as repetitional and experience effects.

References

- [1] Gian Luigi Albano and Alessandro Lizzeri. Strategic certification and provision of quality. *International Economic Review*, 42(1):267–283, February 2001.

- [2] J. W. Albrecht, P. A. Gautier, and S. B. Vroman. Matching with multiple applications. *Economics Letters*, 78(1):67–70, 2003.
- [3] James Albrecht, Serene Tan, Pieter Gautier, and Susan Vroman. Matching with multiple applications revisited. *Economics Letters*, 84(3):311–314, September 2004.
- [4] Lanier C. Benkard. Learning and forgetting: The dynamics of aircraft production. *American Economic Review*, 90(4):1034–1054, 2000.
- [5] K. Boudreau and A. Hagiu. Platform rules: Multi-sided platforms as regulators. *Available at SSRN 1269966*, 2008.
- [6] J. Brandts, G. Charness, and M. Ellman. Let’s talk: How communication affects contract design. 2012.
- [7] Luís Cabral and Ali Hortaçsu. THE DYNAMICS OF SELLER REPUTATION: EVIDENCE FROM EBAY*. *The Journal of Industrial Economics*, 58(1):54–78, 2010.
- [8] H. Chade and L. Smith. Simultaneous search. *Econometrica*, 74(5):1293–1307, 2006.
- [9] R. Fisman, D. Paravisini, and V. Vig. Cultural proximity and loan outcomes. Technical report, National Bureau of Economic Research, 2012.
- [10] M. Gentzkow and J. Shapiro. Media bias and reputation. Technical report, National Bureau of Economic Research, 2005.
- [11] Matthew O Jackson, Tomas Rodriguez-Barraquer, and Xu Tan. Social capital and social quilts: Network patterns of favor exchange. *American Economic Review*, 102(5):1857–1897, August 2012.
- [12] N. L. Jacquet and S. Tan. On the segmentation of markets. *Journal of Political Economy*, 115(4):639–664, 2007.
- [13] P. Jeziorski, I. Segal, et al. What makes them click: Empirical analysis of consumer demand for search advertising. *SSRN eLibrary*, 33, 2009.
- [14] Alessandro Lizzeri. Information revelation and certification intermediaries. *RAND Journal of Economics*, 30(2):214–231, 1999.

- [15] P. Morgan and R. Manning. Optimal search. *Econometrica: Journal of the Econometric Society*, pages 923–944, 1985.
- [16] D. C. North. *Institutions, institutional change and economic performance*. Cambridge university press, 1990.
- [17] Amanda Pallais. Inefficient hiring in entry-level labor markets. February 2012.
- [18] L. Rayo and I. Segal. Optimal information disclosure. *Journal of political Economy*, 118(5):949–987, 2010.
- [19] J. C. Rochet and J. Tirole. Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029, 2003.
- [20] J. C. Rochet and J. Tirole. Two-sided markets: a progress report. *The RAND Journal of Economics*, 37(3):645–667, 2006.
- [21] Marc Rysman. Competition between networks: A study of the market for yellow pages. *The Review of Economic Studies*, 71(2):483–512, April 2004.
- [22] E. Glen Weyl. A price theory of multi-sided platforms. *American Economic Review*, 100(4):1642–1672, September 2010.
- [23] Alexander White and Glen E. Wyl. Insulated platform competition. May 2012.

9 Figures

Figure 2: Probability of Success: Site Wide

Figure 3: Probability of Success in City A

Figure 4: Herfindahl–Hirschman Index for Transaction Type

Figure 5: Response Time Distribution of Pre-Approvals

10 Site Pictures

Figure 6: Splash Page

Figure 7: Search Results

Search results for Berlin, 09/12/2012 to 09/13/2012, 1 Guest. 1347 results (with 5 popular on Wish Lists).

Filters:

- Room type: Entire home/apt (669), Private room (635), Shared room (43)
- Price: \$10 to \$300+
- Connections: Social Connections (3)
- Neighborhood: Prenzlauer Berg (318)

Results:

Rank	Property Name	Type	Address	Price (Per night)	Reviews	Friends
1	Central, simple, charming	Entire home/apt	Mulackstraße, Berlin (Mitte)	\$84	44	2
2	Comfy Bed, Awesome Location	Private room	Revaler Straße, Berlin (Friedrichshain)	\$32	6	4
3	Design Apartment incl. 2 bikes	Entire home/apt	Bleiblicher Straße, Berlin (Neukölln)	\$82	30	29
4	Artist-jewel incl. 2 bikes	Entire home/apt	Schillerpromenade, Berlin (Neukölln)	\$91	27	29
5	Open room in central and hip area	Private room	Danziger Straße, Berlin (Prenzlauer Berg)	\$26	17	6

Figure 8: Hosting Page

View Nearby Properties

Quiet little flat in artsy Neukölln
Apartment - Private room • Karl-Marx-Platz, Berlin, Berlin 12043, Germany

Photos | Maps | Street View | Calendar

Room Details:

- From **\$49** Per Night
- Check in: mm/dd/yyyy
- Check out: mm/dd/yyyy
- Guests: 1
- Subtotal
- Book it!**
- Save to Wish List**

Description: This is a cozy little flat, fully furnished with everything you could

Amenities:

House Rules:

Room type: Private room

Host: Celine

Figure 9: Hosting Page: Bottom

I offer a great apartment in the very center of Kraków. My place is located in the old Jewish quarter called "Kazimierz", 5 minutes walk from the main square. It is in the old, 19th century house. The place is fully renovated and fashionably designed. While I am staying there I offer a separate private room with a double bed for 25\$ per night. That price is per room, not per person, so you don't have to pay an extra money if there is two of you :)

Of course you can use furnished and equiped kitchen, bathroom and huge (30 sq. m.) living room.

There is a wireless internet in my place, so you can easily connect with the internet from your laptop or mobile phone.

I travel a lot, so when I'm out, the whole apartment (2 bedrooms, living room, separate kitchen) is for rent for 60\$ per night.

The neighbourhood is very nice and safe, even late at night. There is a lot of great restaurants, pubs and cafes. At the same time, the quarter still keeps it's bohemian, artistic and unique atmosphere.

Room type:	Private room
Bed type:	Pull-out Sofa
Accommodates:	2
Bedrooms:	1
Bathrooms:	1
Extra people:	No Charge
Weekly Price:	\$165 /week
Monthly Price:	\$400 /month
Country:	Poland
City:	Krakow
Neighborhood:	Kazimierz
Size:	1076ft ² / 100m ²
Cancellation:	Flexible

Contact Me

[More about the host](#)

89% Response Rate

within a day Response Time

today Calendar Updated

How does Airbnb promote safety?

- Educate yourself about safety
- Protected by the \$1,000,000 Airbnb Host Guarantee
- 24/7 phone support
- Rich user profiles and reviews

Are you or your friends connected with this host?

Over 168 million social connections have been made. Join your friends and see where they've stayed.

Log in with Facebook

My Other Listings

Reviews (56) Friends (6)

OVERALL GUEST SATISFACTION

★★★★★

ACCURACY	★★★★★	COMMUNICATION	★★★★★
CLEANLINESS	★★★★★	LOCATION	★★★★★
CHECK IN	★★★★★	VALUE	★★★★★

Mariola is a nice host. Her apartment is very close to the main center and it's easy to walk around most of the city from her place. She was very helpful booking us a taxi to the airport and helping us with the bus and train system. Mariola and her place were highly recommended!

Figure 10: Message

✕

Send Message to Celine

Check in

mm/dd/yyyy

Check out

mm/dd/yyyy

Guests

1

Tell Celine what you like about their place, what matters most about your accommodations, or ask them a question.

☒ Reuse this message next time I contact a host
Contacting several places considerably improves your odds of a booking.

Can this host call you about your inquiry?

☒ Yes☐ No

Your number won't be revealed. They can only call from 9am to 9pm (EDT).
[Not your time zone?](#)

Send Message

11 Tables

Table 1: Site Wide Thread Level Transaction Costs

	2009	2010	2011	2012
% No Response	0.42	0.23	0.20	0.17
Mean Log Days to Response	0.97	0.77	0.78	0.64
Mean Log Hours to Response	1.59	1.70	1.76	1.53
% Attempted Reservation	0.17	0.11	0.14	0.18
% Booked	0.12	0.09	0.11	0.14
% Of Attempts Booked	0.73	0.76	0.81	0.79
Num Msgs If Not Booked	3.17	3.49	3.73	4.15

Table 2: Probability of Success

	(1) Booking	(2) Booking	(3) Booking
Date of First Contact (Months)	0.00583*** (0.000232)	0.00632*** (0.000235)	0.00619*** (0.000234)
Nights		-0.00931*** (0.000329)	
Weeks Ahead of Stay		0.00338*** (0.000281)	
Number of Guests		-0.0228*** (0.00132)	
Guest Dummies	No	No	Yes
Nights Dummies	No	No	Yes
Weeks In Advance Dummies	No	No	Yes
Month Dummies	No	Yes	Yes
R-Squared	0.00816	0.0218	0.0419
Observations	74440	74440	74440

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Probability of Acceptance or Rejection labelpreapproveregs

	(1)	(2)	(3)	(4)
	Pre-Approval	Pre-Approval	Not Available	Not Available
First Msg Words: <=15	-0.0340*** (-3.44)	-0.0212* (-2.38)	0.0222*** (3.37)	0.0187** (3.01)
First Msg Words: 16-50	-0.00947 (-1.07)	-0.000363 (-0.05)	0.0145* (2.45)	0.0125* (2.25)
First Msg Words: 51-100	0.000926 (0.10)	0.00393 (0.49)	0.00965 (1.63)	0.00901 (1.62)
First Msg Words: 101-200	0.00583 (0.64)	0.00486 (0.59)	0.00542 (0.89)	0.00473 (0.83)
Host-FE	No	Yes	No	Yes
Guest Dummies	Yes	Yes	Yes	Yes
Nights Dummies	Yes	Yes	Yes	Yes
Weeks In Advance Dummies	Yes	Yes	Yes	Yes
Month Dummies	Yes	Yes	Yes	Yes
R-Squared	0.0113	0.258	0.00559	0.187
Observations	98682	98682	98682	98682

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Probability of Success - Market Factors

	(1) Booking	(2) Booking	(3) Booking	(4) Booking
Date of First Contact (Months)	0.00612*** (26.62)	0.00550*** (23.54)	0.00644*** (27.58)	0.00659*** (28.08)
Log Num. Initial Contacts	0.0722*** (20.78)	0.0664*** (18.95)	0.0724*** (20.81)	0.0722*** (20.80)
First Word: <=15	-0.512*** (-36.14)	-0.510*** (-35.97)	-0.510*** (-36.04)	-0.509*** (-35.94)
First Word: 16-50	-0.387*** (-30.21)	-0.386*** (-30.09)	-0.386*** (-30.15)	-0.385*** (-30.11)
First Word: 51-100	-0.316*** (-24.62)	-0.315*** (-24.53)	-0.315*** (-24.59)	-0.315*** (-24.56)
First Word: 101-200	-0.274*** (-20.37)	-0.273*** (-20.26)	-0.274*** (-20.36)	-0.273*** (-20.34)
First Word: 201-300	-0.216*** (-10.13)	-0.216*** (-10.12)	-0.217*** (-10.17)	-0.218*** (-10.21)
Log Tightness (Book Week - Trip Week)		-0.0990*** (-15.05)		
Log Tightness (Trip Week)			-0.231*** (-8.13)	-0.574*** (-16.45)
Tightness * Weeks Ahead				0.0577*** (11.51)
Tightness * Weeks Ahead Sq.				-0.000845*** (-6.05)
Guest Dummies	Yes	Yes	Yes	Yes
Nights Dummies	Yes	Yes	Yes	Yes
Weeks In Advance Dummies	Yes	Yes	Yes	Yes
Month Dummies	Yes	Yes	Yes	Yes
R-Squared	0.0773	0.0801	0.0781	0.0814
Observations	74439	74439	74439	74439

t statistics in parentheses* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Probability Guest Books Given Pre-Approval labelguestbooks

	(1) Booked	(2) Booked	(3) Booked	(4) Booked
< 6 Hours	0.267*** (13.62)	0.313*** (11.15)	0.230*** (10.37)	0.259*** (7.71)
6-24 Hours	0.224*** (11.05)	0.263*** (9.05)	0.195*** (8.68)	0.218*** (6.43)
24-48 Hours	0.151*** (6.43)	0.186*** (5.50)	0.130*** (5.15)	0.148*** (3.88)
48-72 Hours	0.112*** (3.75)	0.107* (2.48)	0.109*** (3.42)	0.111* (2.26)
Only 1 Initial Contact	No	Yes	No	Yes
Host-FE	No	No	Yes	Yes
Guest Dummies	No	Yes	Yes	Yes
Nights Dummies	Yes	Yes	Yes	Yes
Weeks In Advance Dummies	Yes	Yes	Yes	Yes
Month Dummies	Yes	Yes	Yes	Yes
R-Squared	0.0300	0.0381	0.195	0.275
Observations	19323	9878	19323	9878

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Time Costs

	2010	2011	2012
Log Response Time	1.516 (2.081)	1.760 (2.124)	1.802 (2.141)
No Response	0.205 (0.404)	0.206 (0.405)	0.197 (0.398)
Log Thread Time to Book	3.173 (1.829)	3.002 (1.810)	2.564 (1.981)
Log Trip Time to Book	3.698 (1.818)	3.590 (1.833)	3.315 (2.100)

Table 7: Time Costs of Search

	No Response	Log Response Time	Log Thread Time to Book	Log Trip Time to Book
Time of Inquiry (Months)	-0.000300* (0.000152)	0.00968*** (0.000793)	-0.0337*** (0.00160)	-0.0234*** (0.00155)
Log Tightness (Book Week - Trip Week)	0.0130*** (0.00316)	-0.104*** (0.0175)	-0.144*** (0.0433)	-0.583*** (0.0465)
Log Nights Inquiry	-0.00385* (0.00155)	0.0410*** (0.00828)	0.0949*** (0.0176)	
Log Weeks Ahead Inquiry	-0.0283*** (0.00136)	0.177*** (0.00738)	0.533*** (0.0158)	
Log Guest Inquiry	-0.0157*** (0.00217)	-0.0495*** (0.0117)	0.0467 (0.0239)	
Trip Characteristic Dummies	No	No	No	Yes
Month Dummies	Yes	Yes	Yes	Yes
R-Squared	0.211	0.409	0.704	0.781
Observations	141988	142299	29106	28488

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Communication Costs by Year

	2010		2011		2012	
	Not Booked	Booked	Not Booked	Booked	Not Booked	Booked
Num. Words First Msg	56.63 (47.31)	68.87 (44.73)	59.84 (43.97)	70.90 (47.15)	63.91 (42.72)	73.00 (46.99)
Num. Total Words	78.72 (81.97)	146.8 (129.6)	81.93 (77.23)	139.5 (125.0)	87.34 (76.59)	130.8 (111.0)
Num. Msg	1.650 (1.212)	2.717 (2.035)	1.645 (1.180)	2.588 (1.977)	1.715 (1.192)	2.386 (1.887)
Observations	17506	2674	54780	9320	82331	15298

Table 9: Communication Over Time - Booked

	Log Word Length	Log Total Words	Num. Messages	Log Host Res. Wrd.	Log Host Res. Wrd.
Time (Months)	0.00000123 (0.0000828)	-0.00378*** (0.000342)	-0.0145*** (0.00177)	-0.0655*** (0.00146)	-0.0669*** (0.00144)
Log First Word Msg	-0.0320*** (0.00176)	0.780*** (0.00540)	0.183*** (0.0174)		0.389*** (0.0177)
Num. Messages		0.233*** (0.00366)			
Month Dummies	Yes	Yes	Yes	Yes	Yes
R-Squared	0.997	0.994	0.634	0.677	0.684
Observations	27599	27600	27600	23034	23025

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Communication Over Time - Not Booked

	Log Word Length	Log Total Words	Num. Messages	Log Host Res. Wrd.	Log Host Res. Wrd.
Time (Months)	0.000731*** (0.0000404)	0.000163 (0.0000881)	0.00370*** (0.000420)	-0.0337*** (0.000524)	-0.0355*** (0.000516)
Log First Word Msg	-0.0568*** (0.000709)	0.900*** (0.00149)	0.163*** (0.00423)		0.271*** (0.00596)
Num. Messages		0.289*** (0.00203)			
Month Dummies	Yes	Yes	Yes	Yes	Yes
R-Squared	0.996	0.997	0.673	0.786	0.790
Observations	155304	155312	155312	118400	118373

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Probability of Success - Continued Search

	(1) Booking	(2) Booking
Date of First Contact (Months)	0.00466*** (22.98)	0.00279*** (15.18)
Log Tightness (Trip Week)	-0.0624* (-2.49)	-0.127*** (-5.47)
First Word: <=15	-0.424*** (-22.18)	-0.373*** (-20.02)
First Word: 16-50	-0.332*** (-17.61)	-0.315*** (-17.18)
First Word: 51-100	-0.295*** (-15.63)	-0.283*** (-15.44)
First Word: 101-200	-0.295*** (-15.32)	-0.275*** (-14.72)
First Word: 201-300	-0.279*** (-10.77)	-0.251*** (-10.55)
Num. Contacts After Initial Search		0.0599*** (32.36)
Guest Dummies	Yes	Yes
Nights Dummies	Yes	Yes
Weeks In Advance Dummies	Yes	Yes
Month Dummies	Yes	Yes
R-Squared	0.0623	0.238
Observations	54199	54198

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Within Host TC - Log Thread Book Time

	(1)	(2)	(3)	(4)	(5)
	Log Book Time	Log Book Time	Log Book Time	Log Book Time	Log Book Time
Time of Inquiry (Months)	-0.0310*** (0.00159)	-0.0315*** (0.00225)	-0.0219*** (0.00335)	-0.0151*** (0.00405)	-0.0139*** (0.00415)
Previous Host Bookings			-0.00285*** (0.000838)		-0.00102 (0.00115)
Previous Prop. Bookings				-0.00857*** (0.00171)	-0.00755*** (0.00210)
Only New Properties					
Host FE	No	Yes	Yes	No	No
Property FE	No	No	No	Yes	Yes
Trip Characteristics	Yes	Yes	Yes	Yes	Yes
Host FE					
R-Squared	0.106	0.257	0.258	0.283	0.283
Observations	29185	29185	29185	29185	29185

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Time Costs of Search - New Properties

	Log Book Time	Log Book Time	Log Book Time
Time of Inquiry (Months)	-0.0236*** (0.00405)	-0.0212*** (0.00405)	-0.0237*** (0.00589)
Previous Host Bookings		-0.0798*** (0.0130)	
No Prev. Books	No	No	Yes
Trip Characteristics	Yes	Yes	Yes
R-Squared	0.145	0.155	0.194
Observations	3863	3863	1553

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix A: Defining a Search Spell

A search spell is composed of a set of contacts by a guest g for a trip at time k having characteristics X . I first drop inquiries that begin after 2013 or before 2010 because those trips should not exist in the dataset. I then drop inquiries for which there was a bug that caused the checkout date to be before the checkin date. I also drop inquiries that are for more than 28 nights as those constitute a sublet. I then drop threads that are associated with multiple checkin dates that are more than 7 days apart. Or with inquiries that were made in the same thread more than 120 days apart. The reason I drop those threads is that I cannot measure the transaction costs associated with each specific inquiry related to that thread. Furthermore, the later inquiries are likely to be repeat transactions which are interesting but would complicate the above study. I also drop inquiries for dates that are more than a year into the future.

I create groups of trips by first grouping the data by the week of the start of the trip and user. Each search spell consists of a set of messages and reservation attempts made by a guest. Some search attempts vary greatly in the number of nights. For example, a guest may have been looking for a 2 night stay and a 20 night stay beginning in the same week. Because those searches are very different, I grouped trip attempts further by number of nights using the following algorithm:

1. If the difference of the range of nights in inquiries for a trip-week is less than 3, I group those inquiries and calculate modal values of nights.
2. Otherwise, I subdivide attempts into group of 0-5 nights, 5 - 10 nights, 10 - 18 nights and 18 - 28 nights. I then use those groups in addition to the trip-week to define search spells.

Lastly, I keep only trip attempts that began after February 2010 or before August 2012. Before 2010, I could not observe the checkin dates of a lot of inquiries. Since some search spells started in 2009 but continued into 2010, I removed the first 2 months of 2010 from the sample. Similarly, I did not include August, September and October of 2012 in the sample because some search spells that began in those months have not ended yet.