

DOCTORAL DISSERTATION

**A STRUCTURAL ANALYSIS OF SHARING ECONOMY LEVERAGING LOCATION AND  
IMAGE ANALYTICS USING DEEP LEARNING**

by

Shunyuan Zhang

submitted to the  
David A. Tepper School of Business  
in partial fulfillment for the requirements for the degree of  
DOCTOR OF PHILOSOPHY  
in the field of Industrial Administration  
at Carnegie Mellon University

DISSERTATION COMMITTEE:

Kannan Srinivasan (co-chair)

Param Vir Singh (co-chair)

Nitin Mehta

Tridas Mukhopadhyay

Anindya Ghose

Carnegie Mellon University, May 2019

© Shunyuan Zhang, 2019

All Rights Reserved

## **ABSTRACT**

The global sharing economy, e.g., AirBnB and Uber, is projected to generate roughly \$335 billion by 2025. The rise of sharing economy has drawn enormous attention from academia and led to policy intervention debates. However, three questions that are essential to a better understanding of sharing economies remain unanswered: 1) can we identify, from unstructured data (product images), the key dimensions of interpretable attributes that affect consumers' choices, and provide guidelines for sharing economy platform for optimizing images to improve the product demand, 2) can a scalable economic model be developed to disentangle factors that influence AirBnB hosts' decisions on the type of property photos to post, and to explore photograph policies that platforms such as AirBnB can employ to improve the profitability for both the hosts and the platform, and 3) are there demand interactions/externalities that arise across sharing economies to provide policy implication. This dissertation contributes to the relevant literature by filling the gap. To achieve this objective, I apply economic theory to a large-scale demand data leveraging advanced machine learning techniques in computer vision and deep learning models.

In the first chapter, I investigate the economic impact of images and lower-level image factors that influence property demand in AirBnB. Employing Difference-in-Difference analyses on a sixteen-month AirBnB panel dataset spanning 7,423 properties, I find that units with verified photos (taken by AirBnB photographers) generate 8.9% more demand, or \$3,500 more revenue per year on average. Leveraging deep learning techniques to classify aesthetic quality of more than 510,000 property photos, I show that 41% of the coefficient of verified photos is explained by the high image quality in these photos. Next, I identify 12 human-interpretable image attributes from photography and marketing literature relevant for real estate photography that capture image quality as well as consumer taste. I quantify (using computer vision algorithms) and characterize unit images to evaluate the empirical marginal effects of these interpretable attributes on demand. The results reveal that verified images not only differ significantly from low-quality photos, but also from high-quality unverified photos on most of these features. The treatment effect of verified photos becomes statistically insignificant once controlling for these 12 attributes, suggesting that AirBnB's photographers not only improve the quality of the image but also align it with the taste of potential consumers. This implies there is significant value in optimizing images in e-commerce settings on these attributes. From an academic standpoint, this study provides one of the first large-scale empirical evidence that directly connects systematic lower-level and interpretable image attributes to product demand. This contributes to, and bridges, the photography and marketing (e.g., staging) literature, which has traditionally ignored the demand side (photography) or did not implement systematic characterization of images (marketing). Lastly, these results provide immediate insights for housing and lodging e-commerce managers (of AirBnB, hotels, realtors, etc.) to optimize product images for increased demand.

In the second chapter, I investigate how AirBnB hosts make decisions on the quality of property images to post. Prior literature has shown that the images play the role of advertisements. Particularly, compared to lower quality amateur images, high quality professional images can increase the present demand by approximately 9% (Zhang et al. 2018). However, there exist a large number of amateur images on AirBnB, even when AirBnB was providing professional photography service *for free* to all the hosts. I posit that the host's decision on what quality of images to post depends not only on the advertising impact of images on the present demand and on the cost of images, but also on the impact of images on the future demand. Thus, some hosts would be hesitant to post professional images because professional images can create unrealistically high expectations for the guests, especially if the actual property is not as good as what the images portray and if the hosts are unable to provide a high-level service to match those expectations. This would result in the satisfaction level of guests to decrease, who would then write a bad review or not write any review at all; and since the number/quality of reviews is one of the key drivers in generating new bookings, this will adversely affect the future demand. I build a structural model of demand and supply, where the demand side entails modeling of guests' decisions on which property to stay, and the supply side entails modeling of hosts' decisions on what quality of images to post and what level of service to provide in each period. I estimate the model on a unique one-year panel data consisting of a random sample of 958 AirBnB properties in Manhattan (New York City) where I observe hosts' monthly choices of the quality of images posted and the level of service provided. The key findings are: 1) guests who pay more attention to images tend to care more about reviews, 2) hosts incur considerable costs for posting above-average quality of image, and 3) hosts are heterogenous in their abilities in investing service effort. In counterfactual analyses, I compare the impact of the current photography policy (offering free high-level images to hosts) and of two proposed policies (offering a menu of free medium-level images to hosts) on the property demand. I show that the proposed policies, though dominated by the current policy in the short-run, outperform the currently policy in the long-run. Noticeably, hosts who might end up using amateur images to avoid the dissatisfactory gap under the current policy, now use free medium-level images to make more revenues under the proposed policy.

In the third chapter, I examine how ride sharing services such as Uber/Lyft affect the demand for home sharing services such as AirBnB. The existing research has largely focused on the impact of sharing economy on incumbent industries while ignoring the interactions among sharing economies. In this study, I examine how ride sharing services such as Uber and Lyft affect the demand for home sharing services such as AirBnB. The identification strategy hinges on a natural experiment where Uber and Lyft exited Austin in May 2016 in response to the introduction of new regulations in Austin that targeted ride sharing services. Applying the Difference-in-Difference approach on a 9-month balanced longitudinal data spanning 7,300 AirBnB properties across 7 US cities, I find that the exit of Uber/Lyft led to a decrease of

9.6% in the AirBnB property demand, which is equivalent to a decrease of \$6,482 in the annual revenue to the host of an average property. I further find that the exit of Uber/Lyft reduced the (geographic) demand dispersion of AirBnB. The demand became more concentrated in areas with access to better public transportation services. Moreover, the properties farther from downtown experienced greater decreases in their demand in the absence of Uber/Lyft. The results indicate that Uber and Lyft affect the demand for AirBnB properties primarily by reducing the transportation costs to and from AirBnB properties that otherwise have poor access to transportation services. The research effort is a first step toward understanding the positive externalities between sharing economies and provides policy implication.

## ACKNOWLEDGEMENTS

I wish to first thank my advisors: Professor Kannan Srinivasan and Professor Param Singh. I would like to express my sincere gratitude to Professor Kannan Srinivasan, who has been, and continuous to be, a tremendous mentor for me. I received a great deal of support and invaluable guidance from him and was constantly enlightened by the discussion with him throughout the development of this dissertation. I am deeply indebted to Professor Param Singh for his support, motivation, and patience. Since my first day in the Ph.D. program at Tepper, Professor Param Singh gave me endless support and utmost encouragement.

I would like to thank the rest of my dissertation committee members: Professor Nitin Mehta, Professor Tridas Mukhopadhyay, and Professor Anindya Ghose, for their great support and advice. I am immensely thankful to Professor Nitin Mehta, who provided me with unreserved support and invaluable advice for my research and the job market. Every discussion with him has benefited me a lot. I am grateful to Professor Tridas Mukhopadhyay. I have benefited a lot from his great support and invaluable advice and I also gained precious experiences from him. I also want to express my gratitude to Professor Anindya Ghose for his guidance and support. Professor Anindya Ghose gave tremendous advice and insightful comments in helping me develop my first research paper.

I thank Professor Dokyun Lee for his time and effort on the projects we coauthored, which leads to part of the dissertation. I am grateful to many other professors for taking the time to discuss research with me and for giving insightful comments on my dissertation papers: Yan Huang, Manmohan Aseri, Hui Li, Tim Derdenger, Christopher Olivola, Jeff Galak, Tong Lu, Maryam Saeedi, David Childers, Ariel Zetlin-Jones, Linda Argote, Oliver Hahl, Sunkee Lee, Joseph Xu, Sridhar Tayur, Ramayya Krishnan, Beibei Li, Rahul Telang, Xiao Liu, Vineet Kumar, Liye Ma

I thank my fellow Ph.D. students and my friends at CMU: Yingjie Zhang, Zijun Shi, Runshan Fu, Nikhil Malik, Siddhartha Sharma, Yijin Kim, Serim Hwang, Jinwoo Kim, Mi Zhou, Yangfan Liang, Xuege Zhang, for sharing research and other aspects of the Ph.D. life and for giving comments on my presentations. I also thank Lawrence Rapp and Laila Lee for all they have done for us Ph.D. students to make the past five years more pleasant.

Finally, I would like to express my deepest gratitude to my parents for supporting me spiritually throughout my life. I also thank my in-laws for their encouragement and support. A special thanks to my loving husband, Ran An, for his wholehearted support and for the full faith in me. This dissertation would not have been possible without his unreserved love and unending inspiration. Most importantly, I thank my wonderful daughter, Ann An, for the joy she brought to my life and the motivation she made me to come this far.

# Table of Contents

ABSTRACT.....	v
ACKNOWLEDGEMENTS.....	viii
Chapter 1 How Much is an Image Worth? Airbnb Property Demand Analytics Leveraging A Scalable Image Classification Algorithm.....	1
1.1 Introduction.....	1
1.2 Relevant Literature.....	4
1.3 Empirical Framework.....	5
1.3.1 Data Description.....	5
1.3.2 Definitions and Measures of Key Variables.....	6
1.3.3 Analysis on Property Images.....	10
1.4 Methods and Results.....	13
1.4.1 Difference-in-Difference (DiD) Analysis.....	13
1.4.2 Propensity Score Weighting (PSW) Method.....	13
1.4.3 Model Specification and DiD Estimator.....	14
1.4.4 Validating the DiD Model.....	16
1.4.5 The Impact of Verified Photos on Property Demand: Image Characteristics and Aesthetic Quality.....	19
1.4.6. The Impact of Verified Photos on Property Demand: Interpretable Image Attributes.....	21
1.5 Discussion and Conclusion.....	34
References.....	36
Appendix to Chapter 1.....	40
I. Classifying Image Quality using Deep Learning-based Classification Model.....	40
II. Algorithm and Concepts for Image Attributes Computation.....	45
III. Measurement of Interpretable Image Attributes.....	48
IV. Room Type Classification.....	50
Training a Room Type Classifier.....	51

V. Robustness Tests and Excluding Alternative Explanation.....	52
Chapter 2 Do Lower-quality Images Lead to Greater Demand on AirBnB?.....	65
2.1 Introduction.....	65
2.1.1 Research Objectives and Main Findings.....	67
2.1.2 Literature Review.....	69
2.2 Research Context and Descriptive Statistics.....	70
2.2.1 Research Context .....	70
2.2.2 Data Description and Measures of Key Variables .....	70
2.2.3 Reduced-form Evidence.....	75
2.3 Model.....	80
2.3.1 Model Overview and Timing of Events.....	80
2.3.2 Property Market-Share Model and Hosts' Revenues.....	82
2.3.3 Individual Host's Cost .....	87
2.3.4 Individual Host's Per-period Payoff .....	90
2.3.5 State Variables .....	91
2.3.6 Individual's Optimization Problem.....	94
2.3.7 A Dynamic Game and Equilibrium Concept .....	95
2.3.8 Unobserved Heterogeneity.....	96
2.4 Estimation Strategy and Identification.....	97
2.4.1 Identification .....	99
2.4.2 Estimating Demand-Side Model.....	100
2.4.3 Estimating Supply-Side Model .....	102
2.4.4 Computation Challenges .....	104
2.5 Estimation Results .....	104
2.5.1 Property Market Share Estimates (Airbnb Demand) .....	104
2.5.2 Host Choice of Images and Effort Estimates (AirBnB Supply).....	109
2.6 Policy Simulation.....	111

2.6.1 Should AirBnB Provide Medium-Quality instead of High-Quality Images for Free?.....	111
2.6.2 Should AirBnB Discriminate Hosts on Offering Photography Policy.....	114
2.7 Discussion and Conclusion .....	115
Reference .....	117
Appendix to Chapter 2 .....	120
I. Technical Notes on Deep Learning-based Image Quality Classification .....	120
II. Technical Notes on Estimation Strategy .....	122
Chapter 3 Demand Interactions in Sharing Economies: Evidence from a Natural Experiment Involving Airbnb and Uber/Lyft .....	129
3.1. Introduction.....	129
3.2. Literature Review.....	133
3.3. Research Context and Empirical Framework .....	134
3.3.1. Observational Data on Airbnb .....	135
3.3.2. Access to Public Transportation: Walkscore.com .....	136
3.3.3. Natural Experiment—Uber/Lyft’s Austin Exit.....	136
3.3.4. Definitions of Key Variables and Descriptive Statistics.....	137
3.4. Empirical Strategy and Results .....	139
3.4.1. Difference-in-Difference (DiD) Model and Weighting Method.....	139
3.4.2 Parallel Pre-Treatment Trends .....	141
3.5. Robustness Checks.....	144
3.5.1. Validating PSW Strategy .....	144
3.5.2. Matching Estimator of the Effect of Uber/Lyft’s Exit .....	145
3.5.3. Seasonality Examination (on Prior Years’ Data).....	146
3.5.4. Random (Shuffled) Treatment Test .....	148
3.5.5. Inclusion of Additional Controls.....	149
3.5.6. Excluding Alternative Explanations .....	150
3.6. Empirical Extensions .....	151

3.6.1. Effect by Access to Transportation..... 151

3.6.2. Effect by Property Luxuriousness..... 153

3.6.3. Effect by Geographic Locations ..... 155

3.7. Conclusion ..... 156

References..... 158

## Chapter 1

# How Much is an Image Worth? Airbnb Property Demand Analytics Leveraging A Scalable Image Classification Algorithm

### 1.1 Introduction

The global sharing economy market has been rapidly increasing in recent years and is projected to generate roughly \$335 billion by 2025 (PwC report 2015). Airbnb, the world’s largest home sharing platform, was recently valued at 20% higher than Marriott and hosted 25% more guests per night than Hilton Worldwide (Winkler and Macmilan 2015). Airbnb has thus become one of the most prominent sharing economy platforms for travelers to choose lodgings and for hosts to generate income by renting out their properties.

Despite its success, Airbnb faces a significant problem in solving the uncertainty that consumers face when evaluating property quality. The inefficiency of information transfer regarding the hosted units—especially from inexperienced hosts—has introduced significant transactional friction and loss of users. Reports show that the quality uncertainty facing potential consumers leads many of them to choose trusted hotel brands over Airbnb (PwC report 2015, Ufford 2015). Airbnb deploys several features to alleviate quality uncertainty, including customer reviews, host verification, detailed description of the property, and property images.

In particular, the property images provide visual information and reduce uncertainty about experiential aspects (e.g., cleanliness, mood) of units in ways that written reviews and descriptions cannot. However, in contrast to hotel images that are taken by professional photographers, most Airbnb property images are taken by hosts, who are amateurs. And therein lies the inefficiency of information transfer, causing uncertainty for potential guests. Furthermore, hosts often complain that the property photos they take are of poor quality and actually make the property appear smaller than it is. To address this concern, in 2011 Airbnb launched a “photography program,” which gives interested hosts (free) access to local professional photographers who are assigned by the company to visit and take photos of the host’s property. An image that is shot and uploaded by Airbnb’s professional photographer is shown with a “verified” mark that appears below the photo. Figure 1 shows how drastically improved and different an image of a room looks when it is shot by an Airbnb professional versus an amateur photographer.

**Figure 1 Compare Unverified to Verified Photo**

Unverified



Verified



It is unclear, however, whether the effect of the professional photography program will be positive, due to the potential improvements in the property images, or be negative, due to the concern of overselling the property. In fact, the photography program has raised much controversy among Airbnb hosts and consumers. On Airbnb's host forum, some hosts mentioned that verified photos may oversell/misrepresent the property and may incur a negative impact.<sup>1</sup> The Airbnb photography program raises a series of questions: 1) Do verified photos lead to an increase in demand? 2) If so, is such an increase due to higher quality of verified images? Or is it due to the additional trust arising from a professional photographer acting as a verified source of proper representation of the facilities? Or is it due to potential differences in high quality images taken by professional photographers affiliated with Airbnb versus others? 3) If good images drive demand, what are the key characteristics of a good image for an Airbnb property? 4) Finally, for these characteristics, can a scalable model be developed for rapid and real-time classification of the images?

To answer these questions, we collected a panel data of 13,000 Airbnb listings with over 510,000 property images in 7 U.S. cities, from January 2016 to April 2017. The dataset contains rich information about a property's monthly reservations (we obtain actual availability, bookings, and blocks by the host), photos, price, and other detailed information about property and hosts. One unique feature of this data is the variation in property images, both across units and over time periods. We can observe properties as they transit from having unverified photos to having verified photos.

Our research analysis is at the intersection of methods in econometrics and computer vision, and draws upon theories from marketing and professional photography literature to define underlying dimensions of a good image that improves economic outcome in an e-commerce setting. We derive multiple dimensions in image attributes that potentially play significant roles in determining economic outcome (e.g., product demand). We see this study as one of the first analyses that structures unstructured data (images) in a

<sup>1</sup> <http://airhostsforum.com/t/professional-photography/3675/35>.

systematic manner to connect directly to economic outcomes—a step towards content engineering paradigm in e-commerce.

As a first step, we use Amazon Mechanical Turk to classify a random (stratified) set of pictures into binary categories of high- and low-quality images using experts. This manual classification must be analyzed to develop a scalable model. To accomplish this, we rely on the developments in computer vision and deep learning. Taking pixel-level information of the images as the input, we use methods known to the field to build a Convolutional Neural Network (CNN) to classify the aesthetic quality for each image in the training sample. The CNN model is optimized to extract a hierarchical set of features from images and learn the “relationship” between the set of features and the image’s label (high- versus low-quality).

Using our trained CNN image quality classifier, we classify unlabeled images in an algorithmic and scalable way into the two categories. While achieving effective classification, the high-dimensional CNN-extracted features are not very helpful in providing managerially relevant information on the drivers of image quality. To accomplish this, we identify three major components, namely, composition, color, and figure-ground relationship. The components are identified based on research in photography literature and consumer behavior literature in psychology and marketing. Twelve dimensions of image attributes (see Table 5) form the basis for these components. Computer vision methods are available to score images on these twelve attributes. We find that not only do high-quality images differ from low-quality images on these attributes, but also verified images differ from other high-quality images on these attributes. These attributes together capture not only image quality but also taste. Given this quantification of the attributes, we can prescribe the actionable recommendations for improving images for Airbnb.

Employing Difference-in-Difference (DiD) analysis in conjunction with machine learning techniques to measure quality of images on a large scale, we report four main findings. First, we find that property will be nearly 8.985% more frequently booked by having verified photos. This effect is positive and significant even after controlling for other sources of information such as guest reviews. Second, we move a step further to explore potential sources of the effect of verified photos. We find that the estimated coefficient of verified photo is largely absorbed after we incorporate photo characteristics in the demand model. Particularly, the coefficient reduces by 41.0% when we control for image quality, suggesting that a significant portion of the effect of verified photos comes from high quality in these photos. Third, using automated computer vision algorithms to score the meaningful image attributes along 12 lower-level dimensions, we investigate what makes a good Airbnb property photo. We do so at the level of three image components, namely composition, color, and figure-ground relationship. Results show that verified photos not only differ from low-quality photos but also from high-quality photos taken by external photographers on these attributes. After controlling for the three components, the estimated coefficient photo verification becomes statistically insignificant. The results suggest that most of the effect of the verified photos comes

from the 12 image attributes which together capture not only quality but taste. In comparison to amateurs as well as external professionals, Airbnb professional photographers better capture the attributes that matter for Airbnb property demand. Lastly, a marginal effect analysis of the 12 attributes on property monthly revenue finds interesting empirical insights related to images. For example, the average increase in predicted revenues associated with one standard deviation improvement in the attributes suggest that potential impact differs<sup>2</sup>. The results further imply that crisp, balanced, and organized view of the unit could attract more customers. Interestingly, attributes that are harder to distinguish or grasp by non-photographers, such as Figure-Ground relationship, which measures how much an image's figures are distinguishable from its background, had lower impact at improvements.

Our research effort makes several key contributions. *First*, this is among the first papers to dissect image attributes according to photography literature and connect them to a direct economic outcome. While impact of images has been studied in marketing literature on advertisement and product images, most studies only relate a few isolated image features to consumer perception. In contrast, we theorize three key interpretable image components that are major sources in the effect of product images on product demand and relate them directly to product demand. Our empirical marginal effect analysis on the interpretable attributes highlights that color may be the most important component in affecting property demand. Beside color, composition and figure-ground relationship are also significant factors that make a good Airbnb property image. *Second*, insights from our paper can guide the image content engineering efforts in the context of the short-term lodging (Airbnb and hotels) and real estate markets. For example, our algorithms for image quality classification and image feature extraction can be adopted by photographers (of firms and hosts) to check the quality and to identify shortcomings of their photographs, and ultimately improve image-based information transfer for their products. Our analysis provides a real-time, scalable model for implementation. Our image analytics algorithm efficiently computes an extensive set of image attributes (~1.06 seconds per image), on a 2-Intel-Haswell (E5-2695 v3) CPU. The analytics step can be scaled up with powerful multiple-thread computing. *Lastly*, as unstructured data is gathering more importance (Netzer et al. 2012, Liu et al. 2017), we demonstrate that extracting information from images and embedding them in sound econometric models can address substantive business problems.

## 1.2 Relevant Literature

A few studies find that the existence of product images plays a positive role in providing product information and reducing quality uncertainty. Images can easily and accurately copy and represent product features that may not be easily conveyed through text. A good product image provides an accurate

---

<sup>2</sup> The extra revenue associated with improved image attributes are computed by averaging over all observations in the data.

visualization of the product, improving a potential customer's confidence level in judging the quality (Shedler and Manis 1986).

Besides providing product information and reducing quality uncertainty, images can also be used as visual messages to persuade consumers and product viewers. There is an evolving stream of marketing literature that studies the impact of images on consumers' perception of products (e.g. Larsen et al. 2004, Meyers-Levy and Peracchio 1992, Peracchio and Meyers-Levy (1994, 2005), Mitchell and Olsen 1981, Gorn et al 1997, Miller and Kahn 2005, Valdez and Mehrabian 1994, Scott 1994). Despite directly linking images to economic outcome, several of these studies largely focused on whether the images exist or not, while ignoring image attributes.

While relevant marketing literature has considered the effect of images or visual elements, our paper differs from existing papers in the following ways:

1) Extant studies focus on viewers' emotional arousal and are restricted to either certain isolated image features (e.g., Gorn et al. (1997, 2004), who look at color) or to high-level image content/style (e.g., Hagtvedt and Patrick (2008), who look at whether an image is "art"). In contrast, we identify the attributes (image features) along which any photograph can be evaluated from the art and photography literature's point of view. We then study the impact of these attributes in combination. Further, in most contexts, the images studied in the relevant literature are of high quality (Bertrand et al 2010). However, in e-commerce, including the sharing economy, a number of the product images are user-generated and are of low quality.

2) Extant studies infer the images' effect on consumers' perception by interpreting results from small-scale survey data collected in a laboratory setting. In contrast, we compute the potential impact of said attributes on economic outcomes measured from large-scale field demand data. Our approach is to directly relate images (and image features) to the economic outcome—namely, demand.

## **1.3 Empirical Framework**

### **1.3.1 Data Description**

We randomly selected 13,000 listings (properties) from 7 cities in the United States (Austin, Boston, Los Angeles, New York, San Diego, San Francisco, and Seattle) and then collected data from January 2016 to April 2017. For each property host in our dataset, we accessed the host's public user profile on Airbnb.com. From the user profile, we obtained the personal information provided to Airbnb by the host. Specifically, we know the date when the host became a member on Airbnb.com, and whether the host has a verified Airbnb account. For each property, we collect information on its characteristics that are static, including location (city, zip code), property type (e.g., house vs. apartment), property size (i.e., the number of beds), amenities (e.g., pool, AC, close to beach), and capacity (maximum number of guests to accommodate). This type of information is static since it is unlikely to change over time. We also collect information on property's characteristics that are dynamic, which may change over time: property bookings, nightly prices,

guests' reviews, property photos and whether the property has verified photos or not. Below we describe the measures of key variables that are used in our analysis and summarize their measurement statistics.

### **1.3.2 Definitions and Measures of Key Variables**

Since this study analyzes panel data with Difference-in-Difference model (i.e., DiD model, see Section 4), we begin with the key definitions in DiD analysis.

#### **Treatment and Untreated Group, and Treatment Status**

The panel data spans 16 periods from January 2016 to April 2017, with each period spanning a month. We define a “treatment” as “having verified photos.” A property is “treated” if it is observed to have verified property photos. The sample for our main analysis consists of 7,423 unique properties that did not have a verified photo by January 2016. Out of these properties, 212 had verified photos by the end of April 2017 (constituting the treatment group), and the remaining 7,211 properties did not have any verified photos throughout the observation window (constituting the untreated group). We define an indicator variable,  $TREAT_i$ , which equals 1 (0) if property  $i$  belongs to the treatment group (untreated group). We further define an indicator variable,  $AFTER_{it}$ , which equals 1(0) for a property  $i$  if period  $t$  is after (before) the period when property  $i$  was first observed to have verified photos. For example, if a property got verified property photos in March 2016, then the variable  $AFTER_{it}$  for this property equals 0 for periods January and February 2016 and equals 1 for periods March 2016 and afterwards. Hence, the treatment status indicator,  $TREATIND_{it}=TREAT_i * AFTER_{it}$  equals 1 if the property  $i$  is treated in period  $t$  and equals 0 if otherwise.

#### **Property Demand**

We purchased listing-level booking data from a company that specializes in collecting Airbnb property demand data. The booking data includes the number of days in a month that a property is open (i.e., the property was available to be booked), booked (i.e., the property was booked by a guest), or blocked (i.e., the property was marked as “unavailable” by the host, without a real booking). For each property  $i$  in each period  $t$ , property demand is measured as occupancy rate, i.e., the fraction of days that a property is booked out of the days that the property is open. We further scale the fraction by 100. For example, if a property in March was made open for 24 days and booked for 6 days, then its demand in that month  $DEMAND_{it}=(6/24)*100=25.00$ .

#### **Property Price and its Instrument**

Property price for a property  $i$  in period  $t$ ,  $NIGHTLY\_RATE_{it}$ , refers to the average over property  $i$ 's nightly price for days in period  $t$ . Property price is endogenous because it is correlated with random demand shock in the current period that also affects property demand. To address the endogeneity concern, we use a set of instrument variables (IVs) for price. Following the extent literature, we first include competitor products' characteristics to serve as IVs (Berry et al. 1995, Nevo 2001). The logic is that competitors' characteristics

are unlikely to be correlated with the unobserved shocks in the focal product’s demand. However, the proximity in product characteristics space between a product and its competitors influence the competition, and as a result, influence the product markup and the price<sup>3</sup>. In addition, we collect cost-related variables—the factors that enter a product’s cost/supply side but not demand side. For this study, we use local (zip code level) residential utility fee obtained from OpenEI and local rental information collected from Zillow<sup>4</sup>. The logic is that these factors serve as an indirect measure of cost and enter price through affecting the cost. However, they’re unlikely to be correlated with the unobserved factors in demand on the short-term lodging market.

### Property Photos

Property photos refer to the set of photos posted on the property webpage in a period. Three variables are measured from the data on property photos: the photo quantity, the photo quality, and the distribution of types of photographed places. The variable  $IMAGE\_COUNT_{it}$  is the number of photos of property  $i$  available on its webpage during period  $t$ . Due to the large number of property photos (over 510,000 images),<sup>5</sup> we leverage machine learning techniques to automatically assess the image quality of these photos. We build a supervised image quality classifier that classifies images into two categories— “high quality” and “low quality.” Then we calculate the average image quality  $IMAGE\_QUALITY_{it}$  over all photos associated with property  $i$  in period  $t$ . Since each photo’s quality is a binary response (0 or 1), variable  $IMAGE\_QUALITY_{it}$  is a real number between 0 and 1. For example, if property  $i$  had 10 images in period  $t$ , with 8 images classified as high quality, we have  $IMAGE\_COUNT_{it} = 10$  and  $IMAGE\_QUALITY_{it} = (8*1+2*0)/10=0.8$ . Lastly, we include the distribution of photographed room types because professional photographers may know which types of places appeal more to the guests and hence present more of these aspects of properties. Specifically, we compute the proportion that each of the five room-types—bathroom, bedroom, kitchen, living-room, and outdoor—is photographed. Then for each property  $i$  in period  $t$ , the distribution is represented by a vector  $\{BATHROOM\_PHOTO\_RATIO_{it}, BEDROOM\_PHOTO\_RATIO_{it}, KITCHEN\_PHOTO\_RATIO_{it}, LIVINGROOM\_PHOTO\_RATIO_{it}, OUTDOOR\_PHOTO\_RATIO_{it}\}$ .

---

<sup>3</sup> In this paper, we compute IVs based on property type, listing type, property capacity, number of reviews, which are not directly in hosts’ control. A property’s competitors are defined as properties in the same zip code.

<sup>4</sup> Open EI dataset provides average residential, commercial and industrial electricity rates by zip code by compiling data from ABB, the Velocity Suite and the U.S. Energy Information Administration dataset 861: <https://openei.org/doe-opendata/dataset/u-s-electric-utility-companies-and-rates-look-up-by-zipcode-feb-2011>; Zillow Research provides average home values by zip code, at a particular home size (# bedrooms): <https://www.zillow.com/research/data/>

<sup>5</sup> The images data contains all images, associated with all properties, collected during data collection periods. That is, it includes images for properties that were verified before the observation started (and hence are not included in the sample for the DiD analyses) and all images updated/added/deleted during observation periods.

Table 1 shows summary statistics for the key variables at the group level. Above each variable is its short description. To show the overall trend of the changes in key variables, we report statistics for the pre-treatment period (January 2016), when none of the properties in sample were treated, and for the post-treatment period (April 2017), when all the properties in treatment group were treated. As shown, for units in the control group, the image quality stayed approximately the same, from 0.29 in January 2016 to 0.30 in April 2017. However, for the units in the treatment group, the image quality drastically improved from 0.27 in January 2016 to 0.77 in April 2017. The improvement in the image quality is consistent with our expectation that photos shot by Airbnb professional photographers are of high quality.

**Table 1 Summary Statistics of Airbnb Properties**

	Control Group		Treatment Group		t-test	
	Mean	Std. Dev.	Mean	Std. Dev.	diff.	t-stats
<b>Pre-treatment (January 2016)</b>						
<i>DEMAND</i> (occupancy rate * 100)	31.07	35.93	32.57	32.30	-1.49	-0.66
<i>IMAGE_QUALITY</i>	0.29	0.27	0.27	0.25	0.02	1.03
<i>IMAGE_COUNT</i>	12.78	9.49	14.48	10.38	-1.71*	-2.58
<i>BATHROOM_PHOTO_RATIO</i>	0.21	0.17	0.22	0.16	-0.01	-0.98
<i>BEDROOM_PHOTO_RATIO</i>	0.30	0.20	0.29	0.19	0.01	1.21
<i>KITCHEN_PHOTO_RATIO</i>	0.11	0.11	0.10	0.10	0.01	0.77
<i>LIVINGROOM_PHOTO_RATIO</i>	0.19	0.17	0.18	0.16	0.00	0.46
<i>OUTDOOR_PHOTO_RATIO</i>	0.19	0.20	0.20	0.20	-0.01	-1.13
<i>SECURITY_DEPOSIT</i>	181.97	347.75	202.77	333.07	-20.80	-0.98
<i>CLEANING_FEE</i>	48.02	53.04	54.69	58.95	-6.66	-1.78
<i>MAX_GUESTS</i>	3.19	2.11	3.50	2.23	-0.30*	-2.14
<i>SUPER_HOST</i>	0.09	0.29	0.15	0.35	-0.05*	-2.36
<i>INSTANT_BOOK</i>	0.11	0.31	0.11	0.31	-0.00	-0.16
<i>MINIMUM_STAY</i>	2.62	2.71	2.57	2.92	0.05	0.26
<i>RESPONSE_RATE</i>	91.00	16.29	92.25	14.31	-1.25	-1.36
<i>RESPONSE_TIME</i> (minute)	261.07	365.66	225.12	338.39	35.95	1.66
<i>NIGHTLY_RATE</i>	179.74	249.31	170.15	240.80	9.58	0.62
<i># RESERVATION DAYS</i>	5.49	8.79	6.62	8.45	-1.13*	-2.08
<i>REVIEW_COUNT</i>	16.41	26.25	20.56	26.98	-4.14*	-2.41
<i>HAST_RATING</i>	0.68	0.47	0.78	0.41	-0.10***	-3.81

<i>RATING_COMMUNICATION</i>	9.71	0.64	9.74	0.49	-0.02	-0.66
<i>RATING_ACCURACY</i>	9.50	0.76	9.54	0.56	-0.05	-1.16
<i>RATING_CLEANLINESS</i>	9.21	0.97	9.26	0.82	-0.06	-0.93
<i>RATING_CHECKIN</i>	9.67	0.67	9.72	0.48	-0.05	-1.45
<i>RATING_LOCATION</i>	9.39	0.82	9.40	0.70	-0.01	-0.20
<i>RATING_VALUE</i>	9.24	0.79	9.34	0.60	-0.10*	-2.21
<i>ZILLOW_RENTAL</i>	2459.95	800.40	2474.75	845.05	-14.80	-0.27
<i>UTILITY</i>	0.18	0.03	0.17	0.04	0.00	1.93

---

**Post-treatment (April 2017)**

<i>DEMAND</i> (occupancy rate * 100)	36.79	39.46	46.58	38.37	-9.79**	-3.26
<i>IMAGE_QUALITY</i>	0.30	0.27	0.77	0.22	-0.47***	-29.37
<i>IMAGE_COUNT</i>	16.22	11.98	19.61	11.53	-3.39***	-3.80
<i>BATHROOM_PHOTO_RATIO</i>	0.21	0.16	0.21	0.15	-0.01	-0.68
<i>BEDROOM_PHOTO_RATIO</i>	0.30	0.19	0.30	0.18	0.01	0.39
<i>KITCHEN_PHOTO_RATIO</i>	0.11	0.10	0.09	0.10	0.01	1.55
<i>LIVINGROOM_PHOTO_RATIO</i>	0.19	0.15	0.18	0.17	0.00	0.08
<i>OUTDOOR_PHOTO_RATIO</i>	0.20	0.19	0.21	0.21	-0.01	-0.66
<i>SECURITY_DEPOSIT</i>	149.53	325.71	179.82	280.53	-30.29	-1.39
<i>CLEANING_FEE</i>	41.51	55.93	46.02	51.01	-4.51	-1.14
<i>MAX_GUESTS</i>	3.37	2.26	3.50	2.31	-0.14	-0.77
<i>SUPER_HOST</i>	0.19	0.39	0.29	0.45	-0.10**	-2.94
<i>INSTANT_BOOK</i>	0.11	0.32	0.16	0.37	-0.05	-1.81
<i>MINIMUM_STAY</i>	2.36	3.75	2.56	4.06	-0.21	-0.67
<i>RESPONSE_RATE</i>	96.63	12.17	95.82	12.50	0.80	0.83
<i>RESPONSE_TIME</i> (minute)	150.42	266.86	167.72	287.81	-17.31	-0.78
<i>NIGHTLY_RATE</i>	237.97	319.49	210.14	146.81	27.83	1.96
<i>#RESERVATION_DAYS</i>	8.59	10.35	10.97	10.13	-2.37**	-3.03
<i>REVIEW_COUNT</i>	32.89	49.33	41.80	46.37	-8.91*	-2.48
<i>HAST_RATING</i>	0.82	0.38	0.84	0.36	-0.02	-0.80
<i>RATING_COMMUNICATION</i>	9.80	0.45	9.82	0.40	-0.02	-0.64
<i>RATING_ACCURACY</i>	9.60	0.60	9.66	0.51	-0.05	-1.17
<i>RATING_CLEANLINESS</i>	9.40	0.75	9.40	0.73	-0.00	-0.04
<i>RATING_CHECKIN</i>	9.79	0.45	9.78	0.45	0.01	0.18
<i>RATING_LOCATION</i>	9.47	0.67	9.50	0.65	-0.03	-0.50

<i>RATING_VALUE</i>	9.36	0.64	9.39	0.57	-0.03	-0.53
<i>ZILLOW_RENTAL</i>	2459.80	822.06	2426.22	823.28	33.58	0.53
<i>UTILITY</i>	0.18	0.04	0.17	0.04	0.01	1.88
<b>Time-invariant</b>						
<i>APARMENT</i>	0.67	0.47	0.60	0.49	0.07*	2.09
<i>ENTIREHOME</i>	0.59	0.49	0.61	0.49	-0.02	-0.52
<i># BEDS</i>	1.70	1.22	1.81	1.21	-0.11	-1.42
<i>POOL</i>	0.08	0.28	0.10	0.30	-0.02	-0.98
<i>BEACH</i>	0.01	0.10	0.02	0.14	-0.01	-1.04
<i>AC</i>	0.98	0.13	0.99	0.06	-0.01	-1.51
<i>PARKING</i>	0.44	0.50	0.50	0.50	-0.06	-1.89
<i>INTERNET</i>	0.98	0.13	0.99	0.09	-0.01	-1.62
<i>TV</i>	0.74	0.44	0.79	0.41	-0.05*	-1.99
<i>WASHER</i>	0.59	0.49	0.60	0.49	-0.02	-0.53
<i>MICROWAVE</i>	0.09	0.29	0.15	0.36	-0.06**	-2.64
<i>ELEVATOR</i>	0.23	0.42	0.20	0.40	0.03	1.03
<i>GYM</i>	0.10	0.29	0.11	0.31	-0.01	-0.55
<i>FAMILY_FRIENDLY</i>	0.23	0.42	0.19	0.39	0.04	1.67
<i>SMOKE_DETECTOR</i>	0.46	0.50	0.55	0.50	-0.09**	-2.94
<i>SHAMPOO</i>	0.35	0.48	0.45	0.50	-0.10**	-3.22

\* p<0.05 \*\* p<0.01 \*\*\* p<0.001

### 1.3.3 Analysis on Property Images

#### 1.3.3.1 Classifying Images into High or Low Quality

We combine techniques from deep learning and computer vision to build a supervised learning classifier that classifies images into high or low quality. The classifier achieves a high accuracy in predicting an image’s label as high quality versus low quality. We build a deep learning-based classifier through the following three steps:

##### *Training set construction*

We chose a stratified random sample of images from our dataset to tag and evaluate with Amazon Mechanical Turk (AMT), a crowdsourcing platform for human intelligence tasks. A stratified (by crude metric of quality) random sample is necessary as it ensures that the sample is balanced as well as random. For each image, we asked Turks (i.e., workers) to rate the image based on its image quality on a 1–7 Likert scale, where 1 is “very bad” and 7 is “excellent”. We provided example photos with different levels of image quality from “very bad image” to “excellent image.” We also provided guidance to the Turks with

detailed instruction on how to evaluate images—for example, “visually pleasing” and “clearly shows room/house features.” Further, each image is evaluated by five qualified Turkers. In the Appendix we report details on image tagging using AMT. The image labels were further converted to binary (“high quality” vs. “low quality”), following practices in computational aesthetics literature (Datta et al. 2006). The training set from this exercise resulted in 1,155 images getting classified as high quality images and 1104 getting classified as low quality images.

#### *Training step*

*Convolutional Neural Networks (CNN) Approach:* We apply Convolutional Neural Networks (CNN), an emerging deep learning framework widely applied in the field of computer vision that has shown breakthrough performance tasks including object recognition and image classification (Krizhevsky et al. 2012, Simonyan and Zisserman 2015). Our CNN image quality classifier, as shown in Figure 2, represents the architecture of a classic CNN model. The CNN consists of a sequence of neural network layers, with each layer extracting features from the output from the previous layer (the first layer extracts features from the input image, which is simply a pixel-valued matrix) and summarizes the features to the next layer. The key component in CNN is a convolution kernel (or convolution filter), represented by an  $n$  by  $n$  weighting matrix that, given the intermediate output from the previous layer, extracts features through a matrix dot product operation between the weighting matrix and the intermediate output. The sequence of layers in CNN learns a representation of the input image by extracting a hierarchical set of image features. The model learns a relationship between the extracted features and the labels and is optimized to extract the features that have the most discriminative power on predicting the labels (e.g., image quality). To reduce overfitting problems in the training step, we employ method of data augmentation and implement a real-time (during training) image transformation over each image in the training sample, by randomly 1) flipping input image horizontally, 2) rescaling input image within a scale of 1.2, and 3) rotating the image within  $20^\circ$ . This method introduces random variation in the training sample, increasing the training set size and reducing the overfitting concern (Krizhevsky et al. 2012).

*Transfer Learning and Fine-tuning the Parameters of the CNN:* A CNN consists of a set of filters, each represented by a matrix. Because a deep learning model has many such filters, the huge number of the parameters, i.e., weights of the matrices, requires a large data-set of images to train the model. Due to the limited training data, we leverage transfer learning to train our CNN. Specifically, we take a widely applied CNN model, VGG-16, which was trained on over 1 million images (Simonyan and Zisserman 2015), as our framework. We fine-tune the parameters based on our training set of Airbnb property images, with the pre-trained parameters of VGG-16 serving as a starting point in the training. We randomly choose 80% of the training set images for training the CNN and use the remaining 20% of images as a hold-out sample to

test the performance of the trained CNN. On average, 90.4% of the hold-out samples were correctly classified; that is, the generalization error rate is less than 10%. The classifier’s high accuracy in predicting image quality ensures a valid interpretation of our results regarding the effect of image quality. In the Appendix, we provide detailed description on the architecture and technical details of the training step of the CNN image quality classifier.

### Prediction step

Once the relationship between the image features and image labels is learned in the training step, the trained classifier is used on unlabeled images in the sample to predict the image quality. The classifier, taking as input an unlabeled property image, extracts the hierarchical set of image features, with the parameters of the trained classifier fixed. The classifier then predicts the label on the output layer (Figure 2) and assigns “1” to high-quality images and “0” to low-quality images.

**Figure 2 Description of Architecture and Layer Description of the CNN Classifier**



**Filters:** The number of convolution windows (i.e., #number of feature maps) on each convolution layer.

**Zero Padding:** Pads the input with zeros on the edges to control the spatial size of the output. Zero padding has no impact on the predicted output.

**Max-pooling:** Subsampling method. A 2×2 window slides through (without overlap) each feature map at that layer, and then the maximum value in the window is picked as representation of the window. Reduces computation and provides translation invariance.

### 1.3.3.2 Categorizing Images into One of Five Room Types

We compute the distribution of the types of rooms photographed in the set of property images. First, we build a deep learning model to automatically classify the room type (bathroom, bedroom, kitchen, living-room, and outdoor) for any given property images. Leveraging transfer learning with a deep learning model that was pre-trained on a large scene classification dataset Places205 (Zhou et al. 2014), we optimize the classifier on a dataset we collected, which consists of 54,557 images of “bathrooms”, 59,082 images of “bedrooms”, 88,030 images of “kitchens”, 81,819 images of “living-rooms”, and 5,734 images of “outdoors”. This classifier achieves an average accuracy of 95.05% on hold-out sample. Next, for each property image, we use the trained room-type classifier to categorize the type of scene/place, given the

image. In the appendix, we provide detailed description for the training set and technical notes for the training step of our room-type classifier.

## **1.4 Methods and Results**

We implement a Difference-in-Differences (DiD) analysis, which is a popular strategy for treatment effect evaluation (Heckman et al. 1997). A propensity-score-based weighting strategy is combined with the DiD analysis for addressing the endogeneity concern.

### **1.4.1 Difference-in-Difference (DiD) Analysis**

The implementation of DiD analysis requires identifying a treatment group and a (comparable) control group. In this study, the treatment group consists of 221 properties, which did not have a verified photo by January 2016, but obtained at least one verified photo by April 2017. The control group consists of 7,211 properties, which did not get a verified photo throughout the observation window.

In an ideal setting, where the two groups are comparable, the impact of photo verification will be reflected by the difference in their demands in the post-treatment period. However, one potential concern in our research setting is that the two groups are not comparable and the treatment is endogenous. Indeed, properties are not randomly assigned to the photography program, but rather hosts self-select to join. As shown in Table 1, the treatment and the control groups differ in some pre-treatment covariates. If the differences that affect hosts' decisions on whether to join the photography program also affect property demand, then we cannot simply attribute any observed difference in the changes in demand to the treatment (Athey and Imbens 2006). To address the concern of endogenous treatment, we adopt an identification strategy where the Propensity Score Weighting (PSW) method is combined with DiD analysis.

### **1.4.2 Propensity Score Weighting (PSW) Method**

Propensity score is defined as the probability of an individual unit receiving a treatment, conditional on a set of observed covariates (Rosenbaum and Rubin 1983). Propensity scores are effective in “balancing” samples and are widely used as sampling weights to make the two groups comparable on covariates (Rosenbaum 2002).

In practice, true propensity score is often unknown. Hence the propensity scores are estimated from samples with modeled as a function of observed covariates. That is, the propensity score of unit  $i$ ,  $ps_i$ , is computed via a specified function  $ps_i = f(\mathbf{X}_i\boldsymbol{\beta})$ . Here  $\mathbf{X}_i$  is a  $1 \times M$  dimensional vector of pre-treatment observed covariates of unit  $i$  and  $\boldsymbol{\beta}$  is an  $M \times 1$  dimensional vector of parameter for  $\mathbf{X}$ .

The model finds a set of parameters that maximize the data likelihood of observing treatment assignments in the sample (Rosenbaum and Rubin 1983). In this study, the parameters are estimated via a logistic regression, with the treatment assignment modeled as a binary choice. The selection of  $\mathbf{X}$  is based on performing a covariates balance check (see appendix). That is, the differences in the means of covariates

between treatment and control groups should be minimized. With parameter vector  $\beta$  estimated, we approximate for each unit  $i$  (with observed covariates  $\mathbf{X}_i$ ) the propensity score  $\widehat{ps}_i(\mathbf{X}_i)$ , which is used in computing sampling weights.

### Computing Sample Weights Based on Propensity Scores

We use propensity score for a weighting strategy—Inverse Probability of Treatment Weighting (IPTW) in the DiD analysis (Austin and Stuart 2015). This weighting strategy is widely applied and is suggested to achieve more precise estimations (with minimal estimation bias and lower mean squared error), compared to some of the propensity score methods. IPTW method assigns a weight to each unit by inverting its propensity score. Specifically, the weight for unit  $i$  is defined as

$$\omega_i(T, \mathbf{X}_i) = \frac{T}{\widehat{ps}_i(\mathbf{X}_i)} + \frac{1 - T}{1 - \widehat{ps}_i(\mathbf{X}_i)}$$

where  $\widehat{ps}_i(\mathbf{X}_i)$  is the estimated propensity score of unit  $i$  computed with its observed covariates  $\mathbf{X}_i$ .  $T$  is a dummy variable that equals 1 if  $i$  is in treatment group and is 0 if otherwise.

A validation on the PSW strategy via a balance check on covariates is performed to ensure the propensity-scores-based sampling weights effectively remove imbalances existing in the two groups (see the appendix for technical details and results). Next, having obtained sampling weights from propensity scores, we implement DiD analysis on the re-weighted sample, i.e., run a weighted regression.

#### 1.4.3 Model Specification and DiD Estimator

Our DiD estimator is obtained with a Weighted Least Squares (WLS) regression, where sampling weights are computed using estimated propensity scores. Let  $DEMAND_{itcym}$  denote the demand for property  $i$  (in city  $c$ ) in year  $y$  and month  $m$  (further let  $t$  denote the period for year  $y$  and month  $m$ ), which can be modeled as

$$DEMAND_{itcym} = INTERCEPT + \alpha TREATIND_{it} + \lambda CONTROLS_{it} + PROPERTY_i + SEASONALITY_{cym} + \varepsilon_{it} \quad (1)$$

where  $TREATIND_{it}$  is the treatment status indicator, which equals 1 if property  $i$  has received treatment in period  $t$  and equals 0 if otherwise. The key coefficient  $\alpha$  estimates the percentage change in property demand caused by having verified photos, compared to property demand without verified photos.  $\varepsilon_{it}$  is a random shock in period  $t$  on property  $i$ 's demand, assumed to follow an i.i.d. normal distribution. The vector  $CONTROLS_{it}$  represents a set of control variables that may be correlated to property demand, e.g., the property rules and the consumer reviews<sup>6</sup>. We include the property fixed effect term,  $PROPERTY_i$ , to

---

<sup>6</sup> Specifically, the vector  $CONTROLS$  includes two metrics that measure hosts' responsiveness,  $RESPONSE\_RATE$  (percentage of responding to a guest's message or request) and  $RESPONSE\_TIME$  (number of minutes to respond to

capture time-invariant factors that may impact property demand, such as geographic information and property-specific characteristics. Also included is the time fixed effect term  $SEASONALITY_{cym}$ , which captures any city-specific trend in property demands (i.e., we allow each city to have its own seasonality pattern).

As shown in Table 2, the estimated coefficient of the key variable  $TREATIN$  suggests a positive significant treatment effect. Specifically, under this model, having verified photos would lead to an increase of 8.985% in the property occupancy. Since on average an untreated property is made open 18.1 days per month, if getting verified photos, on average this effect corresponds to  $18.1 \text{ days/month} * 8.985\% * 12 \text{ months/year} = 19.5$  additional booked days in a year (or 1.6 additional days in a month) for an average property. In terms of revenue, treatment brings in  $\$179.5/\text{day} * 19.5 \text{ days} = \$3500.3$  more per year in revenue for an average untreated property<sup>7</sup>.

**Table 2 Difference-in-Difference Model: The Impact of Verified Photos on Property Demand**

VARIABLES	Main DiD Model (Equation 1)	
	ESTIMATES	Robust S.E.
<i>TREATIND</i>	<b>8.985***</b>	<b>1.660</b>
<i>log REVIEW_COUNT<sub>t-1</sub></i>	9.375***	0.930
<i>NIGHTLY_RATE</i>	-0.146***	0.0320
<i>INSTANT_BOOK</i>	4.156**	1.361
<i>CLEANING_FEE</i>	0.0808***	0.0184
<i>MAX_GUESTS</i>	0.260	1.117
<i>RESPONSE_RATE</i>	0.0699	0.0430
<i>RESPONSE_TIME (minute)</i>	-0.000477	0.00161
<i>MINIMUM_STY</i>	0.133	0.131
<i>SECURITY_DEPOSIT</i>	0.00177	0.00201
<i>SUPER_HOST</i>	3.801*	1.494

a guest), the minimum number of stay nights for booking,  $MIN\_STAYS$ , the maximum number of guests to stay,  $MAX\_GUESTS$ ,  $SECURITY\_DEPOSIT$ , the money that a guest will be charged, upon investigation and approved by Airbnb, if the host reports damages after the hosting and makes a claim for damages,  $CANCELLATION\_STRICT$ , whether the rule on cancelling a booking is strict (1) or not strict (0),  $SUPER\_HOST$ , whether the host has (1) a badge of ‘super host’ or not (0), determined by consumers reviews, responsiveness etc.,  $BUSINESS\_READY$ , whether the property has (1) business-related amenities or not (0),  $HAS\_RATING$ , which indicates that the average guest ratings are computed and presented on the property page (1) or not (0), as well as the interaction terms of  $HAS\_RATING$  and the multi-dimensional ratings.

<sup>7</sup> The properties in our sample in the pre-treatment period (January 2016) are priced at \$ 179.5 per night on average.

<i>BUSINESS_READY</i>	1.806	0.985
<i>CANCELLATION_STRICT</i>	1.016	1.271
<i>HAS_RATING</i>	14.32	12.25
<i>HAS_RATING</i> × <i>COMMUNICATION</i>	-0.212	1.420
<i>HAS_RATING</i> × <i>ACCURACY</i>	0.878	1.211
<i>HAS_RATING</i> × <i>CLEANLINESS</i>	-1.344	1.133
<i>HAS_RATING</i> × <i>CHECKIN</i>	-2.060	1.526
<i>HAS_RATING</i> × <i>LOCATION</i>	-0.757	1.183
<i>HAS_RATING</i> × <i>VALUE</i>	2.141	1.176
<i>INTERCEPT</i>	30.06***	6.683
Fixed Effect	Property	
Seasonality	City-Year-Month	
Num. Observations	76901	
R-squared	0.6608	

\* p<0.05 \*\* p<0.01 \*\*\* p<0.001

#### 1.4.4 Validating the DiD Model

We implement a set of analyses to validate our DiD model combined with PSW strategy. We begin with a falsification check that examines the critical “common trends in pre-treatment periods” assumption followed by a random (shuffled) treatment test and Rosenbaum bounds analysis for selection on unobservables.

##### 1.4.4.1 Falsification Checks on the Pre-treatment Trends

The validity of the causality of the DiD approach (Equation (1)) relies on a critical assumption of pre-treatment common trends. That is, the (weighted or matched) two groups should have common trends in their demands in the periods prior to the treatment (Angrist and Pischke 2008).

A method of examining common trend assumption is relative-time model, with the inclusion of pre-treatment periods. Following the extant literature (e.g., Wang and Goldfarb 2017), we implement falsification checks by decomposing the pre-treatment periods into a series of dummies of the periods prior to the treatment— $PRE_{it}(j)$ :

$$\begin{aligned}
 DEMAND_{icym} = & INTERCEPT + \alpha TREATIND_{it} + \sum_j \beta_j (PRE_{it}(j) \cdot TREAT_i) \\
 & + \lambda CONTROLS_{it} + SEASONALITY_{cym} + PROPERTY_i + \varepsilon_{it}
 \end{aligned} \tag{2}$$

We set the period prior to the treatment month as the reference period (i.e., normalizing its coefficient to zero) and consider a three-period interval prior to the reference period for better interpretability (Autor

2003). Specifically, we let  $PRE(1)$ ,  $PRE(2)$ , and  $PRE(3)$  stand as a dummy for the periods 1 month, 2 months, and 3 months prior to the treatment period, respectively. Furthermore, we let  $PRE(4)$  represent all the periods spanning from the beginning (i.e., January 2016) towards the period 4 months prior to treatment month. For the properties that did not have enough pre-treatment periods (for example, for properties that became verified in February 2016, there was only one pre-treatment month), the period dummies are simply zeros.

The set of coefficients  $\beta_j$  allows us to validate the DiD model by examining the trend lines in the property demand prior to treatment. For the DiD analysis and corresponding causal inferences to be valid, there should be common pre-treatment trends in the property demand between the (weighted) treatment and (weighted) control group. That is, we expect that  $\beta_j$  cannot be positive and significant for validating the DiD model.

Table 3 reports the estimated results for the falsification test. As can be seen, the coefficients for the period dummies  $\beta_j$  are statistically not significantly different from zero. The set of  $\beta_j$  does not exhibit an increasing trend in the property demand for the treatment units, compared to the control units, towards the adoption of treatment. This suggests that the estimated treatment was not a falsely significant result that either began prior to the treatment or was caused by idiosyncratic shock that is potentially associated with both the treatment assignments and with property demand.

**Table 3 Falsification Checks on Pre-Treatment Trends: A Relative-Time Model**

VARIABLES	Relative-time Model (Equation 2)	
	ESTIMATES	Robust S.E.
$PRE(4) * TREAT$	2.230	2.503
$PRE(3) * TREAT$	-0.894	3.164
$PRE(2) * TREAT$	1.302	3.102
$PRE(1) * TREAT$ (reference month)	--	--
$TREATIND$	9.988***	2.311
<b>Property (Non-Photo) Characteristics</b>		
$\log REVIEW\_COUNT_{t-1}$	9.338***	0.933
$NIGHTLY\_RATE$	-0.147***	0.0320
$INSTANT\_BOOK$	4.166**	1.363
$CLEANING\_FEE$	0.0811***	0.0184
$MAX\_GUESTS$	0.199	1.108
$RESPONSE\_RATE$	0.0730	0.0431

<i>RESPONSE_TIME (minutes)</i>	-0.000403	-0.00161
<i>MINIMUM_STAY</i>	0.132	0.131
<i>SECURITY_DEPOSIT</i>	0.00175	0.002
<i>SUPER_HOST</i>	3.764*	1.495
<i>BUSINESS_READY</i>	1.821	0.985
<i>CANCELLATION_STRICT</i>	1.05	1.271
<i>HAS_RATING</i>	14.57	12.22
<i>HAS_RATING</i> × <i>COMMUNICATION</i>	-0.229	1.421
<i>HAS_RATING</i> × <i>ACCURACY</i>	0.978	1.21
<i>HAS_RATING</i> × <i>CLEANLINESS</i>	-1.403	1.134
<i>HAS_RATING</i> × <i>CHECKIN</i>	-2.117	1.521
<i>HAS_RATING</i> × <i>LOCATION</i>	-0.716	1.182
<i>HAS_RATING</i> × <i>VALUE</i>	2.118	1.177
<i>INTERCEPT</i>	29.48***	6.792
Fixed Effect	Property	
Seasonality	City-Year-Month	
Num. Observations	76901	
R-squared	0.6609	

\* p<0.05 \*\* p<0.01 \*\*\* p<0.001

#### 1.4.4.2 Selection on Unobservables

Since propensity scores are computed based on observed variables, a concern with the propensity score-based method is that there may be hidden bias if there are unobserved variables affecting the selection process (i.e., the treatment assignment) and the outcome variables simultaneously. To assess the sensitivity of our estimation to a potential hidden bias, we implement Rosenbaum bounds test (Rosenbaum 2002). Rosenbaum bounds evaluate how much the change in odds ratio of participation, due to unobservables, would be required to nullify the treatment effect identified by propensity score method. One should be more confident about the inference of the estimation results, if it would require a greater change in the odds ratio, caused by the unobservables, to overturn the estimated treatment effect.

Our results on Rosenbaum bounds test (table presented in the appendix), with examination on the Hodges-Lehmann's estimates (Rosenbaum 1993), suggest that, for a positive estimated treatment effect on property demand to be overturned, the potential unobserved factors affecting treatment assignment process would have to be large enough to increase odds ratio of participation by at least 55%. The results of our sensitivity analysis are on the same order of the results obtained in the extant literature (Sun and Zhou 2014,

Manchanda et al. 2015, Li et al. 2016, DiPrete et al. 2004), which reported Gamma ranging from 1.2 to 1.6. Hence, we are confident that our study is robust, to some extent, to the hidden bias caused by hypothetical unobserved factors that affect the selection process. In the appendix, we provide detailed results and discussions.

#### 1.4.4.3 Additional Robustness Tests

In addition to the analysis for examining the common-trend assumption in the DiD model, and the Rosenbaum test for assessing sensitivity of the PSW method to potential unobservables, we perform a complete list of tests that verify the robustness of our main findings. The list of tests includes 1) addressing concern of potential inflated effect of verified photos in the long-term if Airbnb’s ranking algorithm favors properties using professional images, 2) testing whether properties with particular amenities adopted verified photos when these amenities became more attractive ((e.g., pool, close to beach in summer seasons), and 3) investigating whether the main finding of a positive treatment effect was driven by unobserved enhancement to the property quality or to the hosting quality. We provide detailed descriptions about these tests and provide results in the appendix.

#### 1.4.5 The Impact of Verified Photos on Property Demand: Image Characteristics and Aesthetic Quality

Our main analyses in section 4.4 establish a positive causal effect of verified photos on property demand. Next, we move a step further to explore and understand the source(s) of the effect of verified photos. To do so, we first perform a large-scale image analytics to extract and measure image characteristics from property images. Next, we incorporate these image-related variables in our DiD model. That is, we estimate a demand equation (see Equation 3) where the images of a property are captured through three variables: *IMAGE\_COUNT*, the number of property images, and *IMAGE\_QUALITY*, the average aesthetic quality of property images, and *{ROOMTYPE\_PHOTO\_RATIO}*, the distribution of the types of photographed scenes. In section 3.3 and the appendix, we describe the definitions of the variables and the technical details regarding their measurements.

$$\begin{aligned}
 DEMAND_{itcym} = & INTERCEPT + \alpha TREATIND_{it} + \mu IMAGE\_COUNT_{it} \\
 & + \gamma IMAGE\_QUALITY_{it} + \rho_1 BATHROOM\_PHOTO\_RATIO_{it} \\
 & + \rho_2 BEDROOM\_PHOTO\_RATIO_{it} + \rho_3 KITCHEN\_PHOTO\_RATIO_{it} \\
 & + \rho_4 LIVING\_PHOTO\_RATIO_{it} + \lambda CONTROLS_{it} + SEASONALITY_{cym} \\
 & + PROPERTY_i + \varepsilon_{it}
 \end{aligned} \tag{3}$$

Table 4 reports the estimation results, where difference between column (1) and (2) is that the model specification in (1) does not include *IMAGE\_QUALITY*. As shown in the table, the estimated coefficient of the key variable, *TREATIND*, decreases to 7.453 (in column (1)), with part of the coefficient absorbed by

*IMAGE\_COUNT*, which has a positive and significant coefficient. The estimated coefficient of *TREATIND* reduces by 41% ( $(7.453-4.397)/7.453=41\%$ ) once *IMAGE\_QUALITY* is controlled for (the estimate reduces to 4.397, in column (2)). This significant reduction suggests that a large portion of the positive treatment effect comes through the improved aesthetic quality in the verified images. However, there is a significant residual impact (the coefficient of *TREATIND* is positive and significant) even when we control for image quality.

**Table 4 Difference-in-Difference Model: Controlling for Property Images**

VARIABLES	(1)		(2)	
	Without Image Quality		Including Image Quality	
	ESTIMATES	Robust S.E.	ESTIMATES	Robust S.E.
<i>TREATIND</i>	<b>7.453***</b>	<b>1.777</b>	<b>4.397*</b>	<b>2.190</b>
<b>Property (Non-Photo) Characteristics</b>				
<i>log REVIEW_COUNT<sub>t-1</sub></i>	9.754***	0.944	9.570***	0.942
<i>NIGHTLY_RATE</i>	-0.183***	0.0325	-0.187***	0.0325
<i>INSTANT_BOOK</i>	3.768**	1.357	3.664**	1.349
<i>CLEANING_FEE</i>	0.0931***	0.0188	0.0955***	0.0187
<i>MAX_GUESTS</i>	0.247	1.098	0.285	1.094
<i>RESPONSE_RATE</i>	0.0868*	0.0427	0.0886*	0.0427
<i>RESPONSE_TIME (minute)</i>	-0.000232	0.00159	-0.000151	0.00160
<i>MINIMUM_STY</i>	0.172	0.132	0.171	0.133
<i>SECURITY_DEPOSIT</i>	0.00211	0.00201	0.00210	0.00200
<i>SUPER_HOST</i>	3.999**	1.495	3.890**	1.494
<i>BUSINESS_READY</i>	1.805	0.977	1.813	0.974
<i>CANCELLATION_STRICT</i>	1.282	1.277	1.308	1.278
<i>HAS_RATING</i>	15.82	12.27	14.96	12.28
<i>HAS_RATING</i> × <i>COMMUNICATION</i>	0.0499	1.418	0.0702	1.423
<i>HAS_RATING</i> × <i>ACCURACY</i>	0.588	1.209	0.681	1.206
<i>HAS_RATING</i> × <i>CLEANLINESS</i>	-1.214	1.133	-1.144	1.141
<i>HAS_RATING</i> × <i>CHECKIN</i>	-2.260	1.525	-2.267	1.516
<i>HAS_RATING</i> × <i>LOCATION</i>	-0.785	1.186	-0.745	1.193
<i>HAS_RATING</i> × <i>VALUE</i>	2.124	1.176	2.023	1.182
<i>AFTER</i> × <i>POOL</i>	5.841	4.610	6.053	4.608
<i>AFTER</i> × <i>BEACH</i>	-11.29	10.91	-10.39	10.66

<i>AFTER</i> × <i>AC</i>	0.331	3.978	-0.198	3.963
<b>Property Image Characteristics <sup>(+)</sup></b>				
<i>log IMAGE_COUNT</i>	6.874***	1.724	5.518**	1.777
<i>BATHROOM_PHOTO_RATIO</i>	0.777	8.457	0.779	8.439
<i>BEDROOM_PHOTO_RATIO</i>	-2.575	7.539	-2.128	7.499
<i>KITCHEN_PHOTO_RATIO</i>	18.04	11.27	17.20	11.22
<i>LIVINGROOM_PHOTO_RATIO</i>	-12.07	8.430	-12.41	8.405
<i>IMAGE_QUALITY</i>			8.984**	3.371
<i>INTERCEPT</i>	18.42*	8.221	19.55*	8.209
Fixe Effect	Property		Property	
Seasonality	City-Year-Month		City-Year-Month	
Num. Observations	76901		76901	
R-squared	0.6623		0.6628	

+ Note the coefficient for *OUTDOOR\_PHOTO\_RATIO* is not estimated.

\* p<0.05 \*\* p<0.01 \*\*\* p<0.001

#### 1.4.6. The Impact of Verified Photos on Property Demand: Interpretable Image Attributes

Analysis in section 4.5 highlights that image quality is an important source of the effect of verified photos. However, the residual effect of *TREATIND* implies there are key factors of images that go beyond the image quality measure and explain more variation in the property demand. To explore this, we identify key attributes that are human-interpretable, and borrowed from photography and staging literature, to characterize their associations with property demand. Below, we start with a subsection where we theorize the key dimensions.

##### 1.4.6.1 What Makes a Good Property Image?

Multiple image attributes may affect consumers' perceptions and choices. To investigate what the key attributes are and how much they determine Airbnb property demand, it is crucial to understand what makes a good property image for Airbnb. This section defines the key dimensions along which a photograph can be compared and categorized. We first borrow from the art and photography literature to define the image attribute dimensions that would be relevant for property photos. Next, drawing from the literature that studies the role of images in viewer perception, we theorize how each attribute would affect property demand. The photography literature highlights 12 image attributes categorized in 3 components—composition, color, and figure-ground relationship—to evaluate an image or an art work (Freeman 2007,

Datta et al. 2006, Wang et al. 2013). Together these features capture not only quality but also taste. We discuss each of the attributes in detail.

*Component: Composition*

Composition is the arrangement of visual elements in the photograph that would lead the eyes of viewers to the center of focus (Freeman 2007). An expert photographer uses compositional technique to help viewers quickly identify an element that would act as the center of focus (Grill and Scanlon 1990). What compositional technique is appropriate to use depends on the context. Three compositional techniques are relevant for real estate photography.

**Attribute 1. Diagonal Dominance.** A photographer can guide the eyes of viewers through an effective use of leading lines. The two diagonals of a photograph serve as leading lines. In a diagonally dominant photograph, the most salient visual elements are placed close to the two diagonals (Grill and Scanlon 1990). Furthermore, in a rectangular frame, the longest straight lines are the diagonals. If a photographer leads a viewer's eye along a diagonal, it would give the viewer a perception of spaciousness. Hence, we posit that images that are diagonally dominant would lead to greater property demand. For example, in Figure 3, the image on the right is more diagonally dominant than the image on the left. It is likely viewers will perceive that the image on the right represents a larger room than the one on left.

**Figure 3 Compare Images on Diagonal Dominance**

Image Without Diagonal Dominance

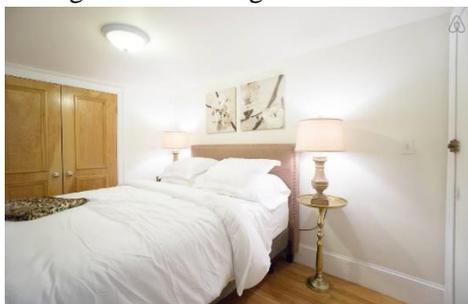
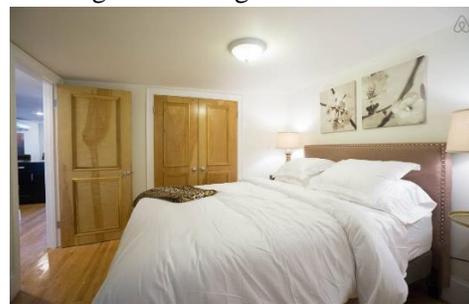


Image With Diagonal Dominance



**Attribute 2. Rule of Thirds.** An image can be divided into nine equal parts by its (imaginary) horizontal and vertical third lines. The rule of thirds (ROT) states that the main visual elements should be placed along the imaginary third lines or close to the four intersections of the lines (Krages 2005). These off-center focal points introduce movement in the photograph, making the image aesthetically pleasing and dynamic (Meech 2004). For example, in Figure 4, the image on the right follows the ROT better than the image on the left. Hence, when looking at the image, a viewer's attention first goes to the bed and then its counterpoint—the other vertical third line. In comparison, the image on the left appears static and it is not obvious to viewers what the focus and key objects are. Therefore, we suggest that images that follow the

ROT would lead to greater property demand, as they would effectively engage viewers by making images more aesthetically pleasing and dynamic.

**Figure 4 Compare Images on Rule of Thirds**

Image (Relatively) Doesn't Follow Rule of Thirds

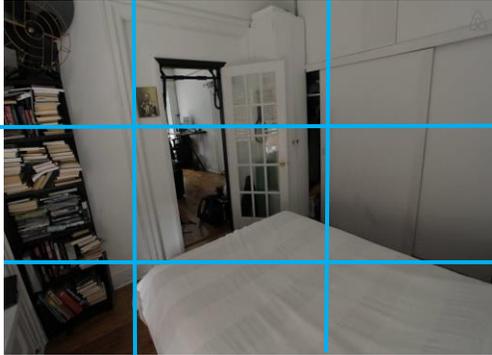


Image (Relatively) Follows Rule of Thirds



**Figure 5 Compare Images on Visual Balance**

Image without Visual Balance



Image with Visual Balance



Attributes 3 and 4. Visual Balance Intensity and Visual Balance Color. Visual Balance relates to the distribution and arrangement of visual elements (Kragas 2005). We evaluate the visual balance of an image from two aspects—intensity and color. If an object is split in half and both sides of the object are mirror images of each other, then the object is considered visually balanced (the extreme case is perfect symmetry). Humans subconsciously consider visual balance to be aesthetically pleasing and it raises visual interest (Arnheim 1974, Bornstein et al. 1981). Visually balanced real estate images give viewers the feeling of order and tidiness and minimizes cognitive demands (Machajdik and Hanbury 2010, Kreitler and Kreitler 1972). Hence, we argue that visually balanced images would lead to greater property demand. In Figure 5 the image on the right is more visually balanced than the image on the left. The image on the right can be processed very quickly and gives a sense of order and cleanliness.

*Component: Color*

Color is one of the most significant elements in photography. Color is widely believed to affect the level of emotional arousal in viewers (Gorn et al 1997,2004). Building on past research, Gorn et al (1997) explain

the two dimensions of arousals—the first that goes from boredom to excitement and the second that goes from tension to relaxation. Excitement is preferred to boredom and relaxation is preferred to tension. Three dimensions of color—hue, saturation (chroma), and brightness (value) can affect the level of arousal. Each of these dimensions has been widely studied in marketing literature, particularly in the context of web design, product package, and advertisement design (Gorn et al 1997, Gorn et al 2004, Miller and Kahn 2005). In addition to these three dimensions, we also discuss another attribute, image clarity, which is affected by the combination of the three.

Attribute 5. Hue. Hues (e.g. red, green, blue) are believed to be a major driver of emotion. Warm hues (such as red and yellow) elicit higher levels of excitement (Gorn et al. 2004, Valdez and Mehrabian 1994). In contrast, cool hues (such as blue and purple) elicit higher levels of relaxation. Hence, we argue that images with warm hues would lead to greater property demand. While the warmth in an image would be affected by the colors of the subject (such as walls, furniture, etc.), a photographer can also further warm up (or cool down) a picture by varying its hues' values during post processing. In Figure 6, we present a cool photo of a living room on the left and a warm photo of a living room on the right.

**Figure 6 Compare Images on Hue (Cool Color vs. Warm Color)**

Image with Cool Color

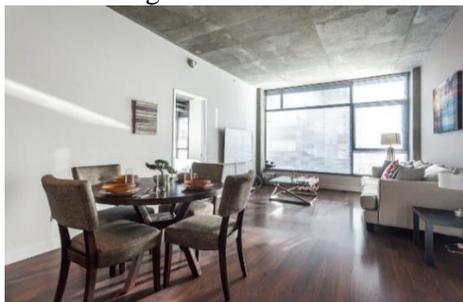


Image with Warm Color



Attribute 6. Saturation. Saturation refers to the richness of color. Highly saturated images reflect colorfulness while low-saturated images contain low levels of pigmentation. More saturated colors are perceived to be associated with happiness and purity, while less saturated colors are associated with sadness and distress (Valdez and Mehrabian 1994, Gorn et al. 2004). Hence, real estate images with saturated colors would induce positive emotions in the viewers and lead to a greater demand. To illustrate the difference in emotion arousal, we present two images of the same room in Figure 7. The only difference between the images is that the image on the right is more saturated than the one on the left.

**Figure 7 Compare Images on Saturation**

Image with Low Saturation

Image with High Saturation



Attributes 7 and 8. *Brightness and The Contrast of Brightness*. Photography literature identifies two image attributes regarding image illumination: brightness and its contrast. Brightness is the overall illumination level of an image. The contrast of brightness describes whether the illumination is evenly distributed over the image with smooth flow. Viewers prefer bright images as they induce a sense of relaxation but do not affect the level of excitement (Gorn et al 1997, Valdez and Mehrabian 1994). Furthermore, sufficient illumination makes the content of an image clear to viewers, because images convey information through pixel brightness. Unevenly distributed brightness may induce a feeling of harshness. For example, in Figure 8, the image on the right has a higher level of and more uniform illumination than the image on the left. We conjecture that property photos where brightness is sufficient and evenly distributed would lead to greater property demand.

**Figure 8 Compare Images on Illumination**

Image with Low and Uneven Illumination



Image with High and Uniform Illumination



Attribute 9. *Image Clarity*. Clear color reflects the intensity of hues in HSV (i.e., Hue, Saturation, Value) space (Levkowitz and Herman 1993). An image is “dull” if it has a color combination mainly consisting of desaturated colors or has near-zero hue intensities in some color channels (He et al. 2011). Amateur photographers often shoot dull photos, inducing a so called “haze effect” that leads to local regions of the image being unclear to viewers and makes the regions look ill-focused. For example, in Figure 9 we present two photos with the right one having higher clarity and the left one having poor clarity. Images with high clarity, we anticipate, would lead to greater property demand because image clarity reduces the friction in information transfer.

### Figure 9 Compare Images on Clarity

Image with Dull Color



Image with Clear Color



### Figure 10 Compare Images on Figure-Ground Relationships

9a. Clear Separation of Figure from Ground



9b. Unclear Separation of Figure from Ground



*Component: Figure-Ground Relationship*

**Attributes 10, 11, and 12. Area Difference, Color Difference, and Texture Difference.** The figure-ground relationship of an image is evaluated from three aspects—area, color, and texture. The principle of figure-ground relationship is one of the most basic laws of perception and is used extensively by expert photographers to plan their photographs. In visual art, the figure refers to the key region (i.e., foreground) and the ground refers to the background of the figure. Figure-Ground (F-G) relationship describes the separation between the figure and the ground. The figure-ground principles follow the gestalt theory, which states that things that share visual characteristics such as size, color, and texture are seen as belonging

together by viewers (Arnheim 1974). Hence, for a region to become more salient, size, color, and texture are the characteristics for which the region and its surroundings should differ. Research on consumers’ responses to advertisement designs suggests that images with clear figure-ground relationships get greater attention from viewers (Schloss and Palmer 2011, Larsen et al 2004). Hence, we argue that images with clear figure-ground separation would lead to greater property demand. In Figure 10, we present one set of images where the figure is clearly separable from the ground and another set where the separation is not obvious.

**1.4.6.2 Measurement of Image Attributes and the Statistics**

We begin with the measurement of these attributes and present the statistics of how professional versus amateur property images score along the key dimensions. In the task of image attributes measurement, computer vision algorithms are first used to process images, extract image features, and then finally measure image attributes. An example of an image processing task is to segment images into patches and to detect key/salient regions (regions considered to be important in an image). After regions of interest are detected, subsequent computation is done for measuring image attributes. The steps for computing image attribute measurements after image processing implementation are provided in an Appendix to this paper. Table 5 summarizes the list of 12 key attributes and the brief description for each.

**Table 5 List of 12 Image Attributes and the Descriptions**

<b>COMPONENT</b>	<b>ATTRIBUTE</b>	<b>DESCRIPTION</b>
<b>Composition</b>	1 Diagonal Dominance	Evaluates how close the key region in an image is placed to the diagonals.
	2 Visual Intensity Balance	Evaluates whether an image has key objects that are symmetric around its vertical central line.
	3 Visual Balance Color	Evaluates if an image has vertically balanced colors.
	4 Rule of Thirds	Image divided into nine equal parts by four horizontal /vertical lines. Evaluates how close the key object is placed to the four intersections of the four lines.
	5 Warm Hue	Portion of warm colors (yellow, orange, etc.) in an image.
	6 Saturation	Evaluates the richness/vividness of image colors.

	7	Brightness	Evaluates the overall image illumination level.
<b>Color</b>	8	Contrast of Brightness	Evaluates whether the illumination distribution is uniform across the whole image.
	9	Image Clarity	Evaluates whether image colors have sufficient intensity.
<b>Figure-Ground</b>	10	Size Difference	Difference in area between image's figure and ground.
<b>Relationship</b>	11	Color Difference	Difference in color between image's figure and ground.
	12	Texture Difference	Difference in texture between image's figure and ground.

### *Statistics of Image Attributes*

Based on the measures of the image attributes, we compute measurements for each property image in our dataset. We divide the property images in our dataset into three groups of photos and look at whether/how one group differs from another on the attributes. The three groups that we construct are as follows:

**Group LQ:** Consists of all low-quality images. This group contains 368,626 images.

**Group HQ\_UN:** Consists of all unverified images that are of high quality. This group contains 69,380 images.

**Group HQ\_V:** Consists of all verified images (these are all high quality). This group contains 72,608 images.

Table 6 summarizes the statistics for the image attributes by groups. We report the means of image attributes for images in each group, with the standard deviations presented in parentheses under the means. The last (rightmost) column compares high quality unverified (group HQ\_UN) and high quality verified (group HQ\_V) images along the dimension of each image attribute. We present the differences between the group means for each attribute measurement, along with the two-sample t-statistics reported in parentheses under the difference. The differences in means where images in group HQ\_UN and in group HQ\_V are statistically different (at 5% significance level) are in bold.

**Table 6 Summary Statistics — Mean (Standard Deviation) of Image Attributes and Compare Verified to Unverified High-Quality Images**

COMPONENT	IMAGE ATTRIBUTE	LQ Low Quality 368,626 Obs. Mean (Std. Dev.)	HQ_UN High Quality Unverified 69,380 Obs. Mean (Std. Dev.)	HQ_V High Quality Verified 72,608 Obs. Mean (Std. Dev.)	HQ_V V.S. HQ_UN Difference (t-statistic)
<b>Composition<sup>8</sup></b>	Diagonal Dominance	-0.342	-0.281	-0.236	<b>0.045***</b>
	Visual Balance	-0.865	-0.774	-0.757	<b>0.017***</b>
	Visual Balance Color	-59.281	-53.093	-50.096	<b>2.997***</b>
	Rule of Thirds	-0.147	-0.089	-0.089	0.0003
<b>Color</b>	Warm Hue	0.738	0.751	0.789	<b>0.038***</b>
	Saturation	59.023	73.942	73.683	-0.259
	Brightness	136.029	154.212	175.802	<b>21.590***</b>
	Contrast of	60.601	58.029	53.996	<b>-4.033***</b>
	Image Clarity	0.324	0.413	0.595	<b>0.182***</b>
<b>Figure-Ground Relationship</b>	Size Difference	-0.405	-0.181	-0.140	<b>0.041***</b>
	Color Difference	23.090	33.054	39.063	<b>6.009***</b>
	Texture Difference	0.043	0.057	0.059	<b>0.002***</b>

Standard deviations in parentheses (for the rightmost column: t statistics in parentheses)

\* p<0.05 \*\* p<0.01 \*\*\* p<0.001

Table 6 shows that low-quality images rate poorly on these image attributes in comparison to high-quality images. More interestingly, the unverified high-quality images also perform poorly on most of these image attributes in comparison to the high-quality verified images. This result indicates that there is systematic difference in high-quality images taken by Airbnb photographers versus others.

<sup>8</sup> Note composition measurements are negative because the composition attributes are evaluated by distances (see section 3.3.2). Hence, we subtract distances from zero, so that the absolute magnitudes stay the same while the directions are reversed. Thus, a greater value of the composition measurements suggests a better performance on that composition attribute. For example, a higher value of diagonal dominance suggests that the image is more diagonal dominant.

### 1.4.6.3 Relating Interpretable Image Attributes to Property Demand

We incorporate the 12 human-interpretable attributes in the demand equation and investigate each attribute’s association with the property demand. For each property image, we measure the 12 key image attributes. The image attribute measurement for a property in a period is then averaged across all images associated with the property in said period. We then perform a DiD analysis on Airbnb property estimating the demand model as specified in Equation (4).

This analysis would also help us examine potential sources of the residual impact. For example, if the significance of the estimated coefficient of *TREATIND* goes away once we added the key image attributes, then it would suggest that the key attributes are major parts of the effect size and that the effect of other factor (such as the verification seal) is statistically insignificant beyond our list of observables. The interpretation might be that the residual impact is explained through systematic difference along the 12 interpretable image attributes in images taken by Airbnb photographers versus high quality images taken by external photographers.

$$\begin{aligned}
 DEMAND_{itcym} = & INTERCEPT + \alpha TREATIND_{it} + \mu IMAGE\_COUNT_{it} \\
 & + \rho_1 BATHROOM\_PHOTO\_RATIO_{it} \\
 & + \rho_2 BEDROOM\_PHOTO\_RATIO_{it} + \rho_3 KITCHEN\_PHOTO\_RATIO_{it} \\
 & + \rho_4 LIVING\_PHOTO\_RATIO_{it} + \eta IMAGE\_ATTRIBUTES_{it} \\
 & + \lambda CONTROLS_{it} + SEASONALITY_{cym} + PROPERTY_i + \varepsilon_{it}
 \end{aligned} \tag{4}$$

Table 7 presents the results from estimating Equation (4) **Table 7**<sup>9</sup>. Noticeably, with the effects of image attributes being teased out, the coefficient of the key variable *TREATIND* reduces to 1.721 and is statistically insignificant. Recall that the combination of the 12 attributes summarizes the key dimensions in image that would each affect property demand. This result combined with results from Table 6 suggests that the treatment effect is primarily due to the fact that Airbnb professional photographers capture these 12 interpretable attributes much better than other photographers.

**Table 7 Difference-in-Difference Model: Controlling for Interpretable Image Attributes**

COMPONENT	VARIABLES	ESTIMAT	
		ES	Robust S.E.
	<i>TREATIND</i>	1.721	6.575
	<b>Property (Non-Photo) Characteristics</b>		
	<i>log REVIEW_COUNT</i>	9.279***	0.927

<sup>9</sup> For ease of understanding, we use standardized values for image attributes (variables are normalized to zero-mean and unit-variance).

<i>NIGHTLY_RATE</i>	-0.194***	0.0325
<i>INSTANT_BOOK</i>	3.245*	1.351
<i>CLEANING_FEE</i>	0.0955***	0.0185
<i>MAX_GUESTS</i>	0.159	1.093
<i>RESPONSE_RATE</i>	0.0946*	0.0424
<i>RESPONSE_TIME (minute)</i>	-0.000203	0.00158
<i>MINIMUM_STY</i>	0.159	0.130
<i>SECURITY_DEPOSIT</i>	0.00189	0.00208
<i>SUPER_HOST</i>	3.614*	1.518
<i>BUSINESS_READY</i>	2.182*	0.971
<i>CANCELLATION_STRICT</i>	1.670	1.256
<i>HAS_RATING</i>	14.02	11.76
<i>HAS_RATING</i> × <i>COMMUNICATION</i>	0.0507	1.393
<i>HAS_RATING</i> × <i>ACCURACY</i>	0.156	1.200
<i>HAS_RATING</i> × <i>CLEANLINESS</i>	-0.879	1.112
<i>HAS_RATING</i> × <i>CHECKIN</i>	-2.198	1.472
<i>HAS_RATING</i> × <i>LOCATION</i>	-0.473	1.156
<i>HAS_RATING</i> × <i>VALUE</i>	2.059	1.165
<i>AFTER</i> × <i>POOL</i>	6.789	4.455
<i>AFTER</i> × <i>BEACH</i>	-9.047	10.31
<i>AFTER</i> × <i>AC</i>	1.474	2.410

---

**Property Image Characteristics**

---

<i>log IMAGE_COUNT</i>	4.530*	1.824
<i>BATHROOM_PHOTO_RATIO</i>	3.431	8.526
<i>BEDROOM_PHOTO_RATIO</i>	0.282	7.642
<i>KITCHEN_PHOTO_RATIO</i>	15.57	11.28
<i>LIVINGROOM_PHOTO_RATIO</i>	-10.39	8.484

---

**12 Human-Interpretable Image Attributes**

---

Image Component	Image Attribute			
<b>Composition</b>	1	<i>DIAGONAL_DOMINANCE</i>	2.516**	0.945
	2	<i>VISUAL_BALANCE_INTENSITY</i>	4.618***	1.350
	3	<i>VISUAL_BALANCE_COLOR</i>	8.869***	2.143

	4	<i>RULE_OF_THIRDS</i>	3.537**	1.106
	5	<i>WARM_HUE</i>	4.715*	2.363
	6	<i>SATURATION</i>	3.920*	1.927
<b>Color</b>	7	<i>BRIGHTNESS</i>	3.434*	1.679
	8	<i>CONTRAST_OF_BRIGHTNESS</i>	-4.897*	2.411
	9	<i>IMAGE_CLARITY</i>	6.212**	2.175
<b>Figure-</b>	10	<i>SIZE_DIFFERENCE</i>	3.807*	1.541
<b>Ground</b>	11	<i>COLOR_DIFFERENCT</i>	2.728*	1.372
<b>Relationship</b>	12	<i>TEXTURE_DIFFERENCE</i>	2.313*	1.090
		<i>INTERCEPT</i>	23.13**	8.546
Fixed Effect			Property	
Seasonality			City-Year-Month	
Num. Observations			76901	
R-squared			0.6670	

\* p<0.05 \*\* p<0.01 \*\*\* p<0.001

As Table 7 shows, the coefficients for the 12 interpretable features are all statistically significant. Moreover, the sizes of the estimated coefficients highlight differential associations with property demand. For example, *IMAGE\_CLARITY* is suggested to be the most important attributes in color component. In addition, the sign of the estimated coefficient for each attribute is consistent with the photography and staging literature that we theorize in section 4.6.1. For example, the coefficients for the four attribute measurements in composition component are all positive and statistically significant. The interpretation is that a good property image should have the photographed objects arranged in a way that make the room look visually balanced (hence tidier and more organized) and more spacious. The contrast of brightness attribute is the only image attribute that has a negative coefficient. The interpretation is that the contrast of brightness attribute should be reduced to make a property image preferable to viewers. This is because an image with high contrast of brightness lacks smooth illumination flow and often comes across as harsh to viewers or may generate unclear local regions in the image.

#### *Empirical Marginal Effects of 12 Interpretable Image Attributes*

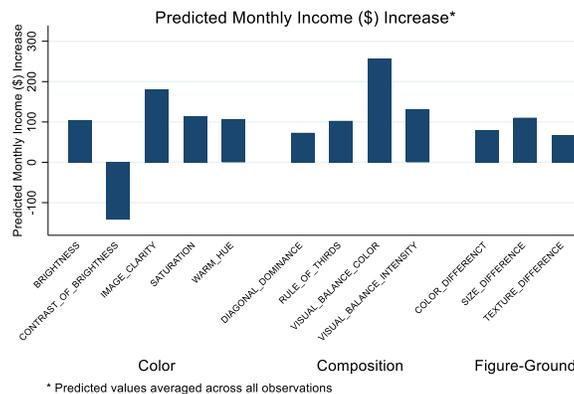
To better understand how each 12 interpretable image attributes may play a role, we present the empirical marginal effects by computing the increase in the predicted host income (monthly revenue) associated with one standard deviation increase in each attribute. We calculate this for each attribute by averaging the increase in unit-specific revenue per month across all units in the data. Please see the appendix for technical details.

First, we show what an increase of 1 SD may look visually in Figure 11. The figure shows the original photos and +1SD for two example attributes, brightness and saturation. Brightness (saturation) increase is associated with \$104 (\$114) increase in monthly income on average. Figure 12 presents the mean of the marginal impacts across all observations in bar chart.

We first note that the increased revenue per month ranges from \$67 to \$257. Given the average price per night in our dataset, the predicted increase per month is substantial. Among the 12 attributes, visual balance color is associated with the highest increased average dollar value at \$257. Image clarity follows second at \$180. Taken together, two attributes' impacts suggest that crisp, balanced, and organized view of the unit could attract more customers. In contrast, attributes that are harder to distinguish or grasp by non-photographers such as diagonal dominance and rule of the thirds had relatively lower impact at \$72 and \$102, respectively. Similarly, Figure-Ground (F-G) relationship, which measures how figures in images are more distinguishable from the background, had lower impact at improvements of \$79, \$67, and \$110, for color-, texture-, and size F-G relationships, respectively. Interestingly, the brightness, often the first attribute to be edited by amateur photographers to make the photos look better, relatively had lower impact at \$103.

Summarizing, the exploratory empirical marginal effect analyses provide support that 1) the benefits from image content engineering could be substantial, 2) potential impact differs, and 3) features that are readily distinguishable to consumers such as visual balance and clarity show relatively higher empirical effects.

**Figure 11 Average Marginal Impact of Image Attribute on Average Predicted Monthly Income**



**Figure 12 Examples of Images with One Std. Dev. Increase in Image Attributes**

Attribute	Original Photo	+1 SD in Attribute
-----------	----------------	--------------------

**Brightness**



**Saturation**



### **1.5 Discussion and Conclusion**

This study identifies the economic impact of images on product demand in the context of Airbnb. We employ DiD analysis combined with the propensity score weighting method to investigate the effect of Airbnb's photography program by looking at the impact of verified photos on property demand. We further explore potential sources of the impact by incorporating relevant image characteristics and find that image quality explains a significant portion of the effect of verified photos. We apply methods in the literature on computational aesthetics in computer vision to automatically assess the image quality of property photos on a large scale with high-generalized accuracy. Furthermore, we identify 12 key attributes in 3 components in image features that are relevant to property demand, with each dimension potentially affecting property demand through information representation or emotion arousal.

Estimation results suggest that having verified photos would lead to an increase of 8.98% in the property demand on average, or an additional \$3500.3 in the annual revenue. Exploring the potential sources underlying, we find that a large portion of the impact of verified photos is explained through high aesthetic quality in these images. Our results reveal that the photos taken by Airbnb professionals differ from external professionals on several key image attributes. As a result, the Airbnb verified photos provide a higher return than the external professional photos. Our findings suggest that most of the effect of the verified photos comes through the 12 key image attributes.

The image attribute analysis enables us to capture the subtle differences between property images and to compute the empirical marginal effects, directly from the data, of each attribute on property revenue. The estimation results suggest that color attributes (including attributes of image clarity, warm hues, saturation, brightness, and contrast of brightness) are important, which validates relevant marketing literature. An interesting finding is that image clarity, despite potentially being the most decisive color attribute for improving property demand, was largely overlooked in past studies. One explanation is that those studies primarily looked at professional images, which have good image clarity. However, ignoring image clarity may create a significant problem for Airbnb, since many of the properties may suffer from having low-quality images, which often have poor image clarity. We also find that image composition, which is largely unstudied in the marketing literature, plays an essential role in determining property demand. The results suggest that images that follow the rule of thirds, and images that are diagonally dominant and visually balanced would lead to greater property demand. It should be noted that specific effects of composition attributes might differ from one context to another. For example, one would not like to make portraits diagonally dominant. However, for property images, diagonal dominance makes the property look more spacious and hence more preferable. Finally, results from image attribute analysis suggest that better figure-ground separation leads to greater property demand. The findings support and extend marketing literature in advertising images, which finds that salient product images receive more attention from viewers and lead to better product perception. A good figure-ground separation is achieved by contrasting the figure to the background in subtle ways. Our results suggest that the separation works the best when it is based on size difference, followed by color difference and difference in textures.

Altogether, this paper relates property images to property demand in both high-level and lower-level dimensions in image features. An exploratory empirical marginal effect analyses supports that the potential impacts of the image attributes on product demand differ and that the benefits from image content engineering could be substantial. Certain industries could benefit from the documented differential effects. For example, home renting markets such as Airbnb and VRBO (Vacation Rental By Owners) could more efficiently resolve the issue of quality uncertainty, by incentivizing their hosts to present high-quality property images. This potentially leads to a greater aggregated property demand, i.e., greater market share. Hosts on the home renting markets also benefit from receiving higher property demand. Our findings also apply to related industries, such as real estate (Zillow.com, Redfin.com, RE/MAX, etc.) and hospitality. For example, a platform such as Zillow.com could use our results to more precisely predict listing sales by estimating how much the attributes of home images contribute to the listing demand, and even launch platforms to improve listing images. Furthermore, our study is among the first to directly link property photos to property demand to identify the economic impact of image attributes. Our demand-driven results could serve as a guideline for creating staging plans or photographs that improve the demand for a property.

This paper thus contributes to the (staging) photography literature, which primarily focused on the effect of image features on aesthetics and did not look at the direct economic impact of image features. Lastly, the identified differential effects of image attributes could have implications for content engineering of product images.

There are a few limitations to this research. First, the quality of property images is not perfectly predicted<sup>10</sup>. Though the high accuracy (90.4%) of the deep learning classifier minimizes the impact of the misclassifications, future studies may consider further improvement in the predication performance, if more labeled data and advanced deep learning model become available. Second, we ignore the user search process on Airbnb. Typically, a potential guest would surf through several properties on an Airbnb search page under certain user specified criterion (e.g., location and dates). In this case, the property image displayed on the search result page may influence the candidate properties that the guest chooses to further evaluate. We do not have access to consumer search processes and hence cannot explicitly incorporate relevant information in our analysis. As more data (on search process and transaction, etc.) become available to researchers, these limitations will open up exciting avenues for future research. Lastly, our analyses on potential impacts of image features on demand are not causal due to the nature of observation data (e.g., the 12 attributes were not randomly assigned in any given photo). Having said that, our work still adds value to relevant fields as this is among the first studies that theorize and dissect key image attributes and connect them to a direct economic outcome at large scale.

Finally, we note that the proposed method and framework can also be applied to other contexts. Visual data-images and videos (since videos can be viewed as sequences of images) have become some of the most effective marketing tools. It has also become the primary way people share information (for example, consumers post and view images on Instagram and Yelp). Our paper on studying the effects of images in key dimensions is a step toward better understanding and leveraging visual data in various markets. Though the magnitude of the effects may differ across markets, our findings have valuable implications for both researchers and market practitioners.

## References

Angrist, J. and Pischke, J. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press Dec 15, 2008

Arnheim, R. 1974. *Art and Visual Perception: A Psychology of the Creative Eye*. Berkeley: University of California Press.

---

<sup>10</sup> In our analyses, image quality enters the econometric model as the mean quality measure over a set of property images. Though a few images could be misclassified during the machine learning step, the computed mean quality is relative consistent with the true mean quality, because some of the misclassifications are averaged out. Hence the misclassification would not be a big concern in our case.

- Athey, S., G.W. Imbens. 2006. Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica* **74**(2) 431–497.
- Autor, D. 2003. Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing. *Journal of Labor Economics* **21**(1): 1-42.
- Austin, P.C., E.A. Stuart. 2015. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine* **34**(28): 3661–3679.
- Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica* **63**(4), 841- 890.
- Bertrand, M., Duflo, E., Mullainathan, S. 2004. “How Much Should We Trust Differences-in-Differences Estimates?”, *The Quarterly Journal of Economics* 119(1): 249–275.
- Bornstein, M.H., K. Ferdinandsen, G.G. Charles. 1981. Perception of Symmetry in Infancy. *Developmental Psychology* **17**(1) 82–86.
- Datta, R., D. Joshi, J. Li, J.Z. Wang. 2006. Studying Aesthetics in Photographic Images Using a Computational Approach. *ECCV* **3953** 288–301. Springer, Heidelberg (2006).
- Freeman, M. 2007. *The photographer's eye: Composition and Design for better digital photos*. Focal Press; 1st edition.
- Gorn, G.J., A. Chattopadhyay, J. Sengupta, S. Tripathi. 2004. Waiting for the Web: how screen color affect time perception. *Journal of Marketing Research* **41**(2) 215-225
- Gorn, G.J., A. Chattopadhyay, T. Yi, D.W. Dahl. 1997. Effects of Color as an Executional Cue in Advertising: They're in the Shade. *Management Science* **43** (10) 1387-400.
- Grill, T., M. Scanlon. 1990. *Photographic Composition*. Watson-Guption.
- Hagtvedt, H., V.M. Patrick. 2008. Art Infusion: The Influence of Visual Art on the Perception and Evaluation of Consumer Products. *Journal of Marketing Research* **45**(3) 379-389.
- He, K., J. Sun, X. Tang. 2011. Single image haze removal using dark channel prior. *Pattern Analysis and Machine Intelligence, IEEE Transactions* **33**(12) 2341-2353.
- Heckman, J., H. Ichimura, P. Todd. 1997. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies* **64**(4) 605–654.
- Krages, B. 2005. *Photography: The Art of Composition*. Allworth Press, U.S.

- Kreitler, H., S. Kreitler. 1972. *Psychology of the arts.* Duke University Press. ISBN 13: 9780822302698.
- Krizhevsk, A., I. Sutskever I., G.E. Hinton. 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* 1097–1105.
- Larsen, V., D. Luna, L. A. Peracchio. 2004. Points of view and pieces of time: a taxonomy of image attributes. *Journal of Consumer Research* **31**(1) 102–111.
- Levkowitz, H., G.T. Herman.1993. GLHS: A generalized lightness, hue and saturation color model. *CVGIP: Graphical Models and Image Processing.* **55** (4): 271–285. [doi:10.1006/cgip.1993.1019](https://doi.org/10.1006/cgip.1993.1019)
- Li, Jun and Moreno, Antonio and Zhang, Dennis, Pros vs Joes: Agent Pricing Behavior in the Sharing Economy (August 28, 2016). Available at SSRN: <https://ssrn.com/abstract=2708279>
- Liu, L.G., R.J. Chen, L. Wolf, D. Cohen-Or. 2010. Optimizing photo composition. *Computer Graphics Forum* **29**(2) 469–478.
- Liu, L., D. Dzyabura, and B. Mizik.2017. Visual Listening In: Extracting Brand Image Portrayed on Social Media. *Working Paper.* Available at SSRN: <https://ssrn.com/abstract=2978805>.
- Machajdik, J., A. Hanbury. 2010. Affective image classification using features inspired by psychology and art theory. in *Proceedings of the international conference on Multimedia ACM* 83–92.
- Manchanda, P., G. Packard, and A. Pattabhiramaiah. 2015. Social Dollars: The Economic Impact of Customer Participation in a Firm-Sponsored Online Customer Community. *Marketing Science* **34**(3):367-387.
- Marchesotti, L., F. Perronnin, D. Larlus, G. Csurka. 2011. Assessing the aesthetic quality of photographs using generic image descriptors. In *ICCV*, pp. 1784-1791.
- Meech, S. 2004. *Contemporary Quilts: Design, Surface and Stitch.* Batsford, London, UK.
- Meyers-Levy, J. and L. A. Peracchio.1992. Getting an Angie in Advertising: The Effect of Camera Angle on Product Evaluations. *Journal of Marketing Research* **29** 454–461.
- Miller, E.G., B.E. Kahn. 2005. Shades of Meaning: The Effect of Color and Flavor Names on Consumer Choice. *Journal of Consumer Research* **32**(1) 86–92.
- Mitchell, A., J.O. Olsen.1981. Are Product Attribute Beliefs the Only Mediator of Advertising Effects on Brand Attitude. *Journal of Marketing Research* **18** 318–32.
- Netzer, O., R. Feldman, J. Goldenberg, M. Fresko. 2012. Mine Your Own Business: Market-Structure Surveillance Through Text Mining. *Marketing Science* **31**(3) 521–543.

- Nevo, A. 2000. Mergers with differentiated products: The case of the ready-to-eat cereal industry. *RAND Journal of Economics* 31(3): 395-421.
- Peracchio, L.A. and J. Meyers-Levy.1994. How Ambiguous Cropped Objects in Ad Photos Can Affect Product Evaluations. *Journal of Consumer Research* **21** 190–204.
- PwC Report. 2015. Consumer Intelligence Series: The Sharing Economy. Retrieved April 2015, <http://www.pwc.com/us/en/industry/entertainment-media/publications/consumer-intelligence-series/sharing-economy.html>.
- Rosenbaum, P.R. 1993. Hodges-Lehmann Point Estimates of Treatment Effect in Observational Studies. *Journal of the American Statistical Association* 83: 1250–1253.
- Rosenbaum, P.R., and D.B. Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* **70**(1) 41–55.
- Rosenbaum, P. R. 2002. *Observational Studies*. 2nd ed. New York: Springer.
- Schloss, K. B., S.E. Palmer, S. E. 2011. Aesthetic response to color combinations: preference, harmony, and similarity. *Attention, Perception, & Psychophysics* **73**(2) 551–571. <http://doi.org/10.3758/s13414-010-0027-0>.
- Scott, Linda M. 1994. Images in Advertising: The Need for a Theory of Visual Rhetoric. *Journal of Consumer Research* **21** 252–273.
- Shedler, J., M. Manis. 1986. Can the Availability Heuristic Explain Vividness effects. *Journal of Personality and Social Psychology* **51** 26–36.
- Simonyan, K., A. Zisserman. 2015. Very deep convolutional networks for large-scale image recognition. *In International Conference on Learning Representations (ICLR)*
- Sun, M., Zhu, F. 2013. Ad Revenue and Content Commercialization: Evidence from Blogs. *Management Science* **59**(10):2314-2331.
- Ufford, S. 2015. The Future of the Sharing Economy Depends On Trust, Forbes. Retrieved February 10 2015, <http://www.forbes.com/sites/theyec/2015/02/10/the-future-of-the-sharing-economy-depends-on-trust/#21a3370b58ff>.
- Valdez, P., A. Mehrabian. 1994. Effects of color on emotions. *Journal of Experimental Psychology: General* **123**(4) 394–409.
- Villas-Boas, J., R. Winer. 1999. Endogeneity in brand choice models. *Management Science*. **45** (10) 1324–1338.

Wang, K., A. Goldfarb. 2017. Can Offline Stores Drive Online Sales?. *Journal of Marketing Research* Vol. LIV (October 2017), 706–719.

Winkler, R., D. Macmillan. 2015. The Secret Math of Airbnb’s \$24 Billion Valuation. Retrieved June 17, 2005, <https://www.wsj.com/articles/the-secret-math-of-airbnbs-24-billion-valuation-1434568517>.

Zhou, B. et al. Learning deep features for scene recognition using places database. *Proceeding NIPS'14 Proceedings of the 27th International Conference on Neural Information Processing Systems*. Vol. 1: 487-495.

## **Appendix to Chapter 1**

### **I. Classifying Image Quality using Deep Learning-based Classification Model**

#### **1. Training Set Construction**

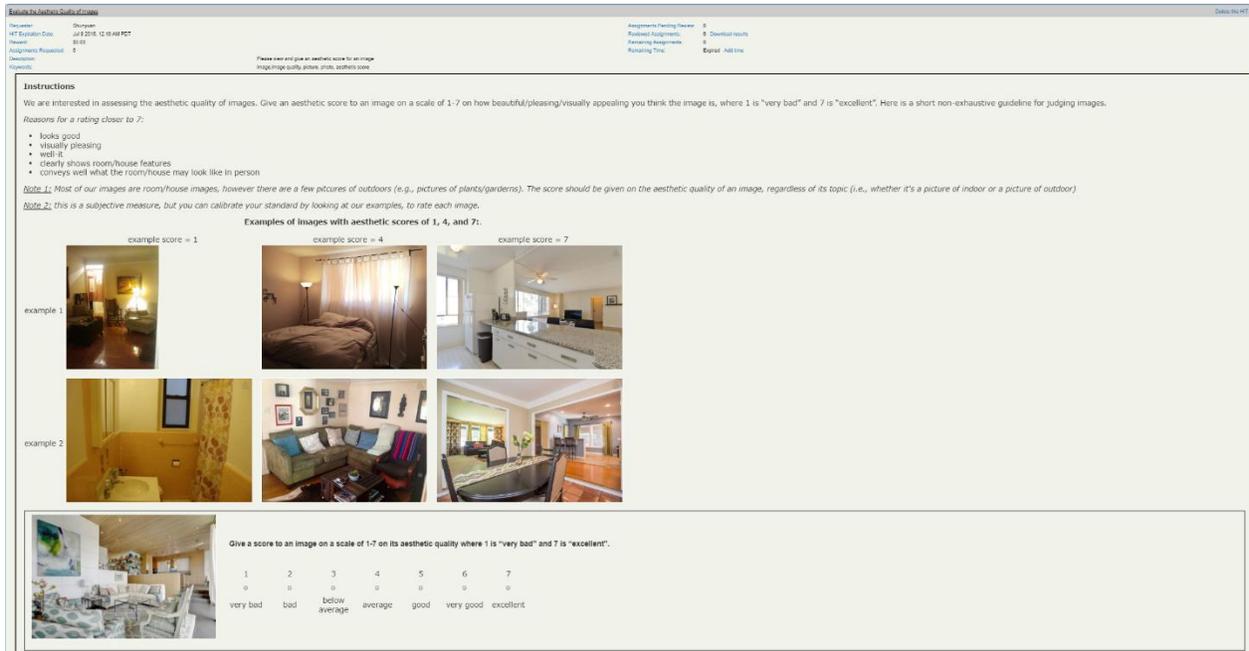
##### *Image Quality Assessment Survey on AMT*

We describe the steps of the dataset (for training our classifier) using Amazon Mechanical Turk (AMT). AMT is one of the platforms of Amazon Web Services (AWS) that enables users to outsource small tasks to a large group of workers at a relatively low cost. AMT has been widely used for human intelligence tasks such as data collection and data cleaning. As a crowdsourcing method, AMT has been found to be quite efficient and accurate (Laws et al. 2011, Casalboni 2015).

To construct a labeled training set for supervised learning, we select 3,000 Airbnb property images from our dataset and ask the Amazon Mechanical Turker (AMT) to tag each image based on the image quality. When selecting images for AMT tagging, full random sampling may not be optimal, since we didn’t have knowledge on the distribution of image quality beforehand. We used stratified random sampling to ensure that there are a sufficient number of images evaluated and labeled for different image qualities. A stratified (by a crude metric of quality) random sample is necessary, as it ensures that the sample is balanced as well as random. Specifically, we randomly selected 500 images from the pool of verified images, since these images are guaranteed to be taken by professional photographers and are most likely to be of high quality. Then, from all the unverified images, we chose 8,000 images that “look bad” (this step was done by a human judge), from which 500 images were randomly sampled. From the unverified images, we also chose 5,000 images that look in between “excellent” and “very bad,” from which 500 images were randomly sampled. Lastly, we randomly sampled 1,500 images from the whole sample. Constructing the AMT data in such a way ensures we will have a sufficient random sample of images from each stratum (subgroup of images in terms of image quality). Furthermore, we manually went through the selected images to make sure that no image was repeatedly sampled.

For the AMT tagging task, we create a survey instrument comprising a set of questions where a Turker is asked to assign a score to a displayed image on its aesthetic quality. To provide Turkers an accurate guideline for image evaluation, we borrow instructions from professional photography forums, as well as from Airbnb’s guidelines on how to shoot good property photos. The quality measurement is based on a 1–7 Likert scale, where 1 is “very bad” and 7 is “excellent.” To ensure high quality and consistent responses from the Turks, we require that Turkers have an approval rate of higher than 95% and have completed at least 50 approved tasks. Furthermore, we provide detailed instruction on where we define aesthetic quality of images and show example photos to guide the Turkers. Figure A1 shows an example of a question that a Turker will have in this survey.

**Figure A1 Example of AMT Aesthetic Quality Assessment Task**



### Constructing Training Set from Tagged Images

After the AMT survey, we obtain 3,000 tagged images, each evaluated by 5 Turkers, to be preprocessed for constructing our training set. Following previous studies on aesthetic quality (Datta et al. 2006, Datta et al. 2008, Marchesotti et al. 2011), we compute for each image  $i$ , the mean aesthetic score  $score_i$  averaged across five Turkers. We then set two thresholds  $\theta_1 = \overline{score} + gap/2$ , and  $\theta_2 = \overline{score} - gap/2$ , and annotate an image  $i$  “high quality” if its average score  $score_i \geq \theta_1$  and “low quality” if  $score_i \leq \theta_2$ .  $\overline{score} = 4.5$  is the mean score of all 3,000 images. Images with average score  $\theta_1$  and  $\theta_2$  are excluded from training set, leaving an artificial “gap” between “high quality” images and “low quality” images. As argued in Datta et al.’s (2006) paper, the reasons for creating a gap between high- and low-quality images is that

close aesthetic scores (e.g., 4.5 and 4.4) are unlikely to reflect that the images differ in their aesthetic quality; rather, they represent noise in the Turkers’ quality measurement process. Note that increasing the value of *gap* makes the task of distinguishing high- and low-quality images easier; however, it leads to a smaller training set since images lying in the “gap” are dropped. To choose an optimal value of *gap*, we vary the value of *gap* and select the value that gives the best performance of trained classifier on a hold-out set (Datta et al. 2006, Marchesotti et al. 2011). In our study, *gap*=0.8 was chosen, which left us a training set of 2,259 images. Note that since 50% of the training set was selected purely at random, in consideration of selection, we repeated the analysis with 50% and found similar results in subsequent image labeling tasks.

## 2. Image Quality Classifier Training

### *The Architecture of Convolutional Neural Networks (CNN) Framework*

After a training set is constructed, the next task is to build an image quality classifier using labeled data. We apply Convolutional Neural Networks (CNN), an emerging deep learning framework widely applied in the field of computer vision, which has been shown to perform very well for tasks including object recognition and image classification (Krizhevsky et al. 2012, Simonyan and Zisserman 2015).

As shown in Figure A2, a CNN model is constructed by a sequence of layers, with each consisting of multiple “neurons”. The number of neurons can vary from one to thousands. These layers of neuron implement matrix multiplication on an input, which is represented as a multi-dimensional matrix, generating an output (again represented by a multi-dimensional matrix) as the input for the next layer. The sequence of layers makes the neural network “deep”.

In the deep learning framework, the images are the first input. In our training task, we first resized all the images to 224×224 (in pixel). We then read pixel intensity of each image and represent the image with a 3-dimensional array (matrix) that contains pixel information in the 3 channels (RGB). We resize the images to alleviate the computational burden and to be aligned with the pretrained VGG16 model (we describe VGG16 model below).

The last output layer predicts binary label for its input, which, after passing through the whole network, is an  $N$ -dimensional vector extracted from the image. For any image in our training set (represented by  $IMG^k$ ), the output layer applies a sigmoid function and predicts the label as a high- versus low-quality:

$$\widehat{\text{Label}}(\text{Image Quality}|IMG^k) = \begin{cases} 1 \text{ (high quality)} & \text{if } \frac{1}{1 + \exp(-(X^T W_1 + W_0))} \geq 0.5 \\ 0 \text{ (low quality)} & \text{if } \frac{1}{1 + \exp(-(X^T W_1 + W_0))} < 0.5 \end{cases}$$

where  $X^T$  represents the output from the layer preceding the output layer (e.g., in our model, it's the FC2 layer, which produces a 4096\*1 vector),  $W_1$  represents the weight parameters and  $W_0$  represents the bias (i.e., a constant) connecting the preceding layer to the output layer.  $\frac{1}{1+\exp(-(X^T W_1 + W_0))}$  is the probability that this image is of high quality, given the extracted vector  $X^T$  and the weights  $W_1$  and  $W_0$  on the output layer.

Throughout the CNN model, there are a sequence of such weights on each layer, and the weights define the intermediate extracted vectors from each layer, including  $X^T$  as described above. These weights are adjusted during the training process, so as to optimize the model's performance on predicting the images in the training set.

### *Operation of a Few Key Layers in CNN*

In the CNN, there are a few of layers that improve the performance of CNN. Below we describe convolution layer, zero padding layer, and max-pooling layer.

#### **Convolution Layer**

The convolution layer is the most important and unique layer in the CNN. A convolution layer consists of a stack of so called convolution filter or convolution kernel. For example, the two convolution layers in Layer Block A (shown in Figure A2) consist of 64 and 128 convolution filters, respectively. A convolution filter is simply a matrix with each element representing a numeric value. For example, in Layer Block A the convolution layers have a size of 3×3 and hence consists of 9 such numeric values<sup>11</sup>. Such a matrix, treating an image or an intermediate input as a matrix, operates a dot production by “sliding” through the input. Therefore, for an input with relative large size (e.g., 224×224), a 3×3 convolution filter operates dot production for every 3×3 patch on that input matrix. The nice features of convolution operation are that: 1) it reduces the dimensionality of parameters, and 2) it well explores and reserves the (local) spatial relationships of the input. Particularly, an intuitive example of the second feature is that if a convolution kernel extracts a particular oriented edge of an object, then operating this kernel on every small square (e.g., 3×3) of an image would extract all edges in that direction from the image. Many of such kernels that extract edges would extract edges in all directions, potentially constructing the contour of an object. As can be seen in **Figure 2**, each of the blocks consists of varying numbers of convolutional filters (e.g., 64, 128, 256, and 512 filters). Hence, these kernels extract features from an input data, which represents the extracted features from the preceding layers. Towards the output layer in the CNN, the filters combined extract

---

<sup>11</sup> The size of a convolution layer is a choice of the model architecture. 3×3 is a widely-used choice. Another common choice is 5×5.

higher- and lower-level features. That is, the CNN is able to extract a hierarchical structure of features, that are related to predict the output labels.

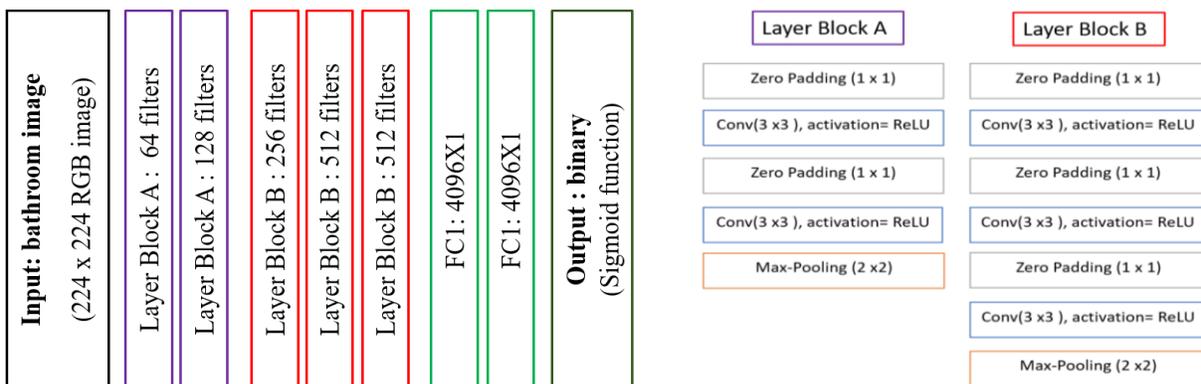
### Zero-padding Layer

The zero-padding layer adds numerical arrays consisting of all “0” values to the edge of a given intermediate output from a layer. The size of zero-padding layer is a hyperparameter in the CNN model. Typically, as done in our model, the intermediate output is padded with  $1 * M$  ‘0’ vectors on each side such that the width and the height both increase by 1 after the zero-padding value. Note that the zero-padding layer does not affect the extracted features through the layers since zeros will not contribute to the matrix multiplication. However, zero-padding layer has a nice feature of allowing us to control the spatial size of the intermediate outputs. Particularly, it can prevent the outputs from reducing too quickly after layers of convolution operations.

### Max-pooling Layer

It’s a common practice in CNN to insert a max-pooling layer in-between the successive convolution layers. A max-pooling layer is a small square filter. In our model, the max-pooling filter is a  $2 \times 2$  matrix. Similar to the operation of convolution filter, a max-pooling layer applies to every  $2 \times 2$  square patch on an input data. However, the function of a max-pooling layer is to pick and preserve only the maximum value in that  $2 \times 2$  square. Adding max-pooling layers can reduce the spatial size of the intermediate features and the dimension of the trained parameters in the model. Particularly, it helps to efficiently prevent the problem of over-fitting.

**Figure A2 Description of Architecture and Layer Description of the CNN Classifier**



**Filters:** Indicate the number of convolution windows (i.e., number of feature maps) on each convolution layer.

**Zero Padding:** Pads the input with zeros on the edges to control the spatial size of the output. Zero padding has no impact on the predicted output.

**Max-pooling:** Subsampling method. A  $2 \times 2$  window slides through (without overlap) each feature map at that layer, and then the maximum value in the window is picked as representation of the window. Reduces computation and provides translation invariance.

### *Training the CNN*

We randomly split the dataset into training set and validation set, with 80% of the examples forming the training set and the remaining being used as validation set. To reduce the overfitting problem in the training step, we employ method of data augmentation and implement a real-time (i.e., during training) image transformation over each image in the training sample, by randomly 1) flipping input image horizontally, 2) rescaling input image within a scale of 1.2, and 3) rotating the image within  $20^\circ$ . This method introduces random variation in the training sample, increasing the training set size and reducing the overfitting (Krizhevsky et al. 2012).

To effectively learn features from a relatively small sample size, we follow the idea of “transfer learning” by building our model on top of an existing well-trained CNN model and fine-tune it. Since the features extracted from images are generic to some extent (e.g., almost all CNN extract edge information at the first layer), transfer learning is quite common in deep learning and is suggested to be an effective approach for tackling problems with limited data (Zhang et al. 2015, Girshick et al. 2014, Lin et al. 2015). In this study, we adopted VGG16 (Simonyan and Zisserman 2015) since it’s a conceptually simple and popular pre-trained model. Specifically, we removed the last fully connected layers from the original VGG16 model since they contain more data-specific features. We then add the output layer as the last layer. Figure presents the architecture of our (VGG16-based) image classification model. Hence, the parameters are initiated with the pre-trained weights except for the output layers, where the parameters were initialized following LeCun’s uniform scaled initiation method (LeCun et al. 1998). Then we fix the parameters on the first 25 layers and fine-tune the model. The model is trained on the training set on a NVIDIA K80 GPU, and then the performance is tested on the hold-out set at the end of training. The optimization is implemented with adaptive method of gradient descent (Adadelata optimization, Zeiler 2012) on each mini-batch of 16 examples.

## **II. Algorithm and Concepts for Image Attributes Computation**

We define key concepts that are used in the process of image attribute computation (the measurement of the image attributes is discussed in the following section). The key concepts include image saliency, key/salient region, and figure-ground. For each concept, we first discuss its definition, then the image algorithm (steps) to detect, to extract, or to compute the concept.

## 1. Visual Saliency

The basic unit in image saliency is visual saliency at the level of image pixel. Then, the overall saliency score for a local patch of an image can be computed based on each pixel saliency within the local region.

### *Definition*

Saliency describes a concept originating from visual unpredictability. In images, it's often captured by variations (such as boundaries and changes in colors). Studies in cognitive psychology and computer vision investigate how humans process and pay attention to visual information, and find that we allocate our attention to parts of the information (e.g., the regions of an image) while (cognitively) “ignore” other parts. Visual uniqueness is “salient” in the sense that it easily gets attention from viewers.

### *Calculation*

In general, models proposed for calculating visual saliency are based on local contrast to the surroundings. The contrasts are determined using features including colors, intensity, edge density and orientation. A simple example is the “gradient” of pixel intensity. A pixel with great contrast is assigned a high saliency value.

## 2. Salient Region

### *Definition*

Following the definition of visual saliency, salient regions are defined as regions on an image that are “salient”—the regions with high overall saliency score.

### *Detection*

The detection of a salient region requires three steps. 1) Segment an image into local “patches”; 2) assign a saliency score to each patch; 3) merge similar patches into a “region”; and 4) find the most salient region.

- 1) *Image segmentation*: Image segmentation generally involves grouping pixels of an image into multiple parts, where pixels within each part are similar to each other. Segmentation can be edge-based (detected edges are assumed to define boundaries of objects), color-based (e.g., clustering pixels based on their colors), and others. The logic of a segmentation algorithm is to compare two intensity differences: the difference across the boundary of “patches” and the difference between two neighboring pixels within the same patch.
- 2) *Assign saliency score to a patch*: Each pixel is assigned a saliency value based on the calculation described in the calculation of pixel saliency value. The saliency score of a patch is calculated by averaging over all pixels within the patch.
- 3) *Merge patches into a region*: If neighboring patches have similar colors, then merge them into a larger region.

- 4) *Salient region*: The salient region is then found by picking the region with the highest average salient score.

### 3. Figure-Ground

#### *Definition*

Figure is the foreground, and ground is the background of the figure. For one image, only one figure is detected and only one ground is found. This is different from detecting salient regions, where multiple regions could be detected, and we just rank them by their salient scores and pick the region with the highest.

#### *Detection*

First detect and extract the foreground (figure) from an image, then the ground is “the rest of the image.” Detection of foreground is an “extension” of image segmentation, where a pixel is either assigned a value of 1 (foreground) or 0 (background).

#### *Detailed Algorithm*

We use the state-of-art model for foreground extraction—the GrabCut model (Rother et al. 2004). An image here is treated as a graph, with each pixel a node and pixel similarity an edge. GrabCut implements EM (Expectation-Maximization) algorithm and mincut algorithm to iteratively assign foreground/background label to pixel and to cut the graph into two sub-graphs—one representing foreground and the other representing background.

- 1) Initially, an arbitrary rectangle separates the image into two parts—pixels in the rectangle are labelled “1” (foreground) and pixels outside are labelled “0” (background). The initial position of the rectangle can be arbitrary. Alternatively, we can specify the rectangle location or hard label some pixels if we have good prior knowledge on where the foreground might be.
- 2) A Gaussian Mixture Model (GMM) is trained with EM algorithm, based on the distribution of pixel color statistics. From the GMM we get the probability of each pixel belonging to a particular mode (or cluster). That is, the GMM labels a pixel as “probable foreground” versus “probable background”.
- 3) A graph is built, with each node representing a pixel, and the edge weight between two pixels representing the pixel similarity. Similar pixels will get a low edge weight, and vice versa. Pixel similarity can be computed based on pixel intensity, pixel color, or pixel texture.
- 4) On the graph, two additional nodes are created—“source node” and “sink node”. All pixel nodes labeled as “probable foreground” (“probable background”) are connected to source node (sink node), with edge weight between each pixel node and the source node (sink node) representing the probability that the pixel belongs to foreground (background).

- 5) Then the graph is cut into TWO parts implementing mincut algorithm through minimizing a cost function. The cost function is the sum of edge weights over all edges that are cut. That is, mincut algorithm penalizes a cut if this cut will cause two similar pixels to be separated into two subgraphs. Intuitively, if two pixels both have high probability of being foreground (background), then we want them to be labelled “1” (“0”) at the end of this iteration.
- 6) Repeat steps 2)~5) till pixel labeling convergences.

### III. Measurement of Interpretable Image Attributes

With the computer vision algorithm described above, we implement image processing tasks to segment images into patches and to detect key/salient regions (regions considered to be important in an image). After regions of interest are detected, subsequent computation is done for measuring image attributes. This section discusses the steps for computing image attribute measurements after image processing implementation.

#### *a. Composition*

Four image attributes are categorized in the composition component. How well an image performs on a particular attribute is evaluated by “distance,” such as the distance between two pixels. A smaller distance indicates a better performance in the specific composition attribute. For all four composition attributes, we compute the distance metrics, then subtract the metrics from zero.

**Diagonal Dominance** (Attribute 1): Diagonal dominance captures how close an image’s key region is placed to the two diagonals of the image. For an image, we first identify the key region and then measure the weighted Manhattan distance from the key region to each diagonal (Liu et al. 2010, Wang et al. 2013).<sup>12</sup> The measurement of diagonal dominance is computed by subtracting the minimum weighted distance from zero. Hence, a greater value of diagonal dominance measurement suggests that the image is more diagonally dominant.

**Rule of Thirds** (Attribute 2): An image is divided into nine equal parts with two (imaginary) equally spaced vertical lines and two (imaginary) equally spaced horizontal lines. We calculate the Euclidean distance from the centroid of the key region to each of the intersections (Wang et al. 2013). If the minimum distance is small, then the image follows the rule of thirds with its key region close to at least one intersection. The measurement of rule of thirds is computed by subtracting the minimum distance from zero. Therefore, if an image follows rule of thirds more closely, it has a greater value for the rule of thirds measurement.

**Visual Balance Intensity** (Attribute 3): In this measurement, we “split” the image along its vertical central line. On each half of the image, we identify a key region and compute the distance from its centroid to the

---

<sup>12</sup> The Manhattan distance between two points on an image is measured as the number of pixels between them, with only horizontal and vertical path from one point to the other allowed.

vertical central line (Liu et al. 2010). A relative distance measure is calculated by subtracting the shorter distance from the longer and dividing the difference by the longer distance. The measurement of visual balance intensity is then computed by subtracting the relative distance from zero. A greater value for this measure suggests that the image is more (vertically) visually balanced on pixel intensity.

**Visual Balance Color** (Attribute 4): The color measurement of visual balance compares the left half to the right half of an image based on colors. We calculate the Euclidean distance in color intensity (i.e., RGB channel) between each pixel and its symmetrical pixel (symmetric around vertical line). The measurement of visual balance color is computed by subtracting the average difference from zero. Therefore, a greater value of visual balance color suggests that the image is more visually balanced on colors around its vertical central line.

#### *b. Color*

Five image attributes are computed in color component. The measurements are taken with pixel intensities or related values (e.g., hue and saturation).

**Warm Hue** (Attribute 5): The warm hue measurement captures the “warmth” of an image, which is defined by the relative portion of warm hues (e.g., yellow) to cool hues (e.g., green). The measurement is computed in the HSV (Hue, Saturation, and Volume) space. We calculate the pixel hues that fall outside the cool hue range (30–110) on the hue spectrum (Wang et al. 2013). If an image contains more warm hues, such as yellow and orange, it will have a greater value for warm hue measurement.

**Saturation** (Attribute 6) We compute the pixel saturation value in HSV space. The measurement of saturation is computed by averaging the saturation values across all pixels on the image. A greater value for the saturation measurement indicates a higher saturation (e.g., the image contains more saturated colors).

**Brightness** (Attribute 7): The brightness of an image evaluates the overall illumination level. We calculate the intensity of each pixel then average the intensity values across all pixels on the image. A “brighter” image has a greater value for the brightness measurement.

**Contrast of Brightness** (Attribute 8): The contrast is calculated as the standard deviation of pixel intensity over the whole image. Thus, a lower contrast of brightness measurement suggests that the brightness is more evenly distributed across the image.

**Image Clarity** (Attribute 9): The measurement of image clarity captures the portion of hues with sufficient intensity. We measure the pixel brightness on a [0–1] scale, then compute the portion of pixels with brightness falling into the region of [0.7–1] (Wang et al. 2013). A clear image has a great value for the image clarity measurement.

### *c. Figure-Ground Relationship*

The figure-ground relationship is described by the “difference” between the figure and its ground in three metrics—the size, the color, and the texture. An image with good figure-ground relationship has a clearly separable figure and ground, that is, has greater differences.

**Size Difference** (Attribute 10): The size difference attribute compares the size of the figure to the size of the ground. We detect the figure and the ground of an image and calculate each size ratio respective to the whole image (Cheng et al. 2011). Size difference is then computed by subtracting the ground’s size ratio from the figure’s size ratio. Thus, a greater value for size difference measurement indicates that the image has a figure that occupies a relatively larger area, making the figure stand out from the ground.

**Color Difference** (Attribute 11): The color difference attribute captures the difference in colors between the figure and the ground. We compute the Euclidean distance between the mean color of the figure and that of the ground. A high value of color difference measurement suggests that the figure and the ground contain distinct colors. The figure is thus easily distinguished from its ground due to the contrast of their colors.

**Texture Difference** (Attribute 12): Texture difference measures the difference between the figure and the ground in terms of “texture,” which is captured by edge density within a local region. For the figure and the ground, we operate the Canny edge detector to detect edge and then compute edge density, respectively. The measurement is the absolute difference between the two densities. A great value for texture difference measurement suggests that the figure and the ground have a clear separation based on textures.

## **IV. Room Type Classification**

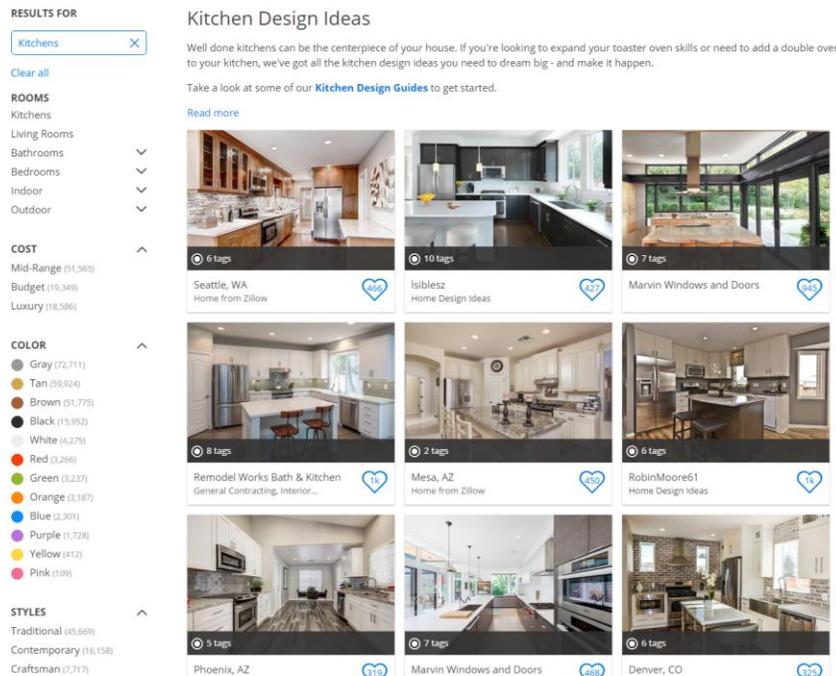
We build a deep learning model to automatically categorize the type of the photographed scene. The goal is to compute, given the type of each photo, the distribution of the types of rooms photographed in the property images. Controlling the distribution in our demand model helps us to address the concern that maybe professional photographers know which types of places appeal more to the consumers and present these aspects of properties.

We build a deep learning model to automatically classify room types (bathroom, bedroom, kitchen, living-room, and outdoor) for any given property images. Leveraging transfer learning with a deep learning model that was pre-trained on a large scene classification dataset Places205 (Zhou et al. 2014), we optimize the classifier on a dataset we collected, which consists of 54,557 images of “bathrooms”, 59,082 images of “bedrooms”, 88,030 images of “kitchens”, 81,819 images of “living-rooms”, and 5,734 images of “outdoors”. The average classification accuracy is 95.05% on a hold-out set across the five categories.

## Training Set of Indoor/Outdoor Photos

To train a room type classifier, we need a large amount of room images with room type labeled. We would like to use the original images from Airbnb.com, however the images on property webpage are not labeled, and it will incur a high labor cost to manually (e.g., Amazon Mechanical Turk) label images for us. Hence, we crawl data from real estate relevant website, where tons of indoor/outdoor images are classified into multiple categories. For example, Figure 13 shows the portion of a webpage displaying 23 images of “kitchens” of multiple properties.

Figure 13 Real Estate Web Page of Kitchens



From the web site, we collected images across the five categories of room types—bathroom, bedroom, living-room, kitchen, and outdoor. We then split the dataset, using 80% of the dataset as training set and the remaining 20% as hold-out test set.

## Training a Room Type Classifier on Collected Training Set

We use VGG16 ConvNet model pre-trained on Places205 dataset (Zhou et al. 2014) and fine-tune the parameters on our training set<sup>13</sup>. The model was used for a task to classify 205 categories of places. To transfer the pre-trained model to our study, we remove the output layer in the pretrained model and then add an output layer that is designed for our specific task. The added output layer is a 5\*1 vector, with each

<sup>13</sup> The pre-trained VGG16 model (architecture and the parameters) is available to be accessed here: <http://places.csail.mit.edu/downloadCNN.html>

element indicating the predicted probability of assigning the corresponding label. Note that to make sure all the probabilities sums up to 1 (i.e., we assume a room belongs to only one category), Softmax function for calculating predicted probability. For example, the probability that an image is assigned room type “k” is computed as:

$$\text{Prob}(\text{Room Type} = k | X^T) = \frac{X^T W_1^k + W_0^k}{\sum_{j=0}^4 X^T W_1^j + W_0^j}$$

Where  $j=0, \dots, 4$  represents the room type of {bathroom, bedroom, kitchen, living-room, outdoor}, respectively.  $X^T$  represents the output from FC2 (the 2<sup>nd</sup> fully connected layer),  $W_1^j$  represents the weight connecting FC2 to the  $j^{\text{th}}$  node on the output layer, and  $W_0^j$  represents the bias connecting FC2 to the  $j^{\text{th}}$  node on the output layer.

## V. Robustness Tests and Excluding Alternative Explanation

This section reports a series of analyses to test the robustness of our main results and/or excludes alternative explanation. We begin with the validation on propensity score method used in our empirical model: Propensity Score Weighting Strategy and Sensitively Assessment on unobservables.

### 1) Validating Propensity Score Method

To ensure that the Propensity Score approach effectively eliminates potential systematic imbalance between the treatment and control groups, one needs to show that the propensity scores have balanced the covariates on matched or weighted samples.

We implement a balance check, which compares, over the covariates, the weighted means of treatment group,  $\bar{X}_{\text{treatment}} = \frac{\sum_{i \in \text{treatment}} \omega_i X_i}{\sum_{i \in \text{treatment}} \omega_i}$ , and of control group,  $\bar{X}_{\text{control}} = \frac{\sum_{i \in \text{control}} \omega_i X_i}{\sum_{i \in \text{control}} \omega_i}$ . Here,  $X_i$  is a 1\*M dimensional vector of pre-treatment observed covariates of unit  $i$  and  $\omega_i$  is the sample weight for unit  $i$ , computed based on the estimated propensity scores.

**Table 1** presents, for each variable  $X^m$  ( $m=1, 2, \dots, M$ ), the weighed group means and a test for the difference in the means. As shown with the t-stats in the table, the weighted samples are not statistically different at the 95% significance level. That is, the systematic differences in the weighed samples are negligible after performing PSW method. Hence, we validated that our PSW method has effectively eliminated the imbalances in the sample and that our weighted treatment and control groups are comparable in the observed covariates that may affect the treatment selection process.

**Table 8 Propensity Score Weighting Validation: Covariates Balance Check**

---

**Weighted Means in**

---

VARIABLES	Treated	Untreated	T-test	
			t	p-value
<i>REVIEW_COUNT</i>	20.56	19.88	0.27	0.790
<i>IMAGE_QUALITY</i>	0.27	0.25	1.00	0.316
<i>IMAGE_COUNT</i>	14.48	15.1	-0.67	0.506
<i>NIGHTLY_RATE</i>	170.15	191.36	-1.14	0.257
<i>MINIMUM_STY</i>	2.57	2.57	-0.00	1.000
<i>MAX_GUESTS</i>	3.5	3.67	-0.82	0.410
<i>RESPONSE_RATE</i>	92.25	91.19	0.79	0.431
<i>RESPONSE_TIME (minute)</i>	225.12	260.98	-1.18	0.238
<i>SUPER_HOST</i>	0.15	0.11	1.05	0.292
<i>INSTANT_BOOK</i>	0.11	0.11	-0.14	0.888
<i># BLOCKED DAYS</i>	9.51	8.32	1.10	0.271
<i># RESERVATION DAYS</i>	6.62	6.74	-0.16	0.877
<i>PARKING</i>	0.5	0.49	0.09	0.929
<i>POOL</i>	0.1	0.08	0.76	0.445
<i>BEACH</i>	0.02	0.02	0.34	0.737
<i>INTERNET</i>	0.99	1	-0.58	0.563
<i>TV</i>	0.79	0.81	-0.55	0.579
<i>WASHER</i>	0.6	0.57	0.81	0.419
<i>MICROWAVE</i>	0.15	0.13	0.64	0.523
<i>ELEVATOR</i>	0.2	0.21	-0.22	0.826
<i>GYM</i>	0.11	0.13	-0.69	0.490
<i>FAMILY_FRIENDLY</i>	0.19	0.2	-0.56	0.576
<i>SMOKE_DETECTOR</i>	0.55	0.52	0.62	0.534
<i>SHAMPOO</i>	0.45	0.44	0.09	0.929
<i>BATHROOM_PHOTO_RATIO</i>	0.22	0.21	1.05	0.295
<i>BEDROOM_PHOTO_RATIO</i>	0.29	0.29	-0.26	0.795
<i>KITCHEN_PHOTO_RATIO</i>	0.1	0.1	-0.15	0.880
<i>LIVINGROOM_PHOTO_RATIO</i>	0.18	0.19	-0.52	0.602
<i>OUTDOOR_PHOTO_RATIO</i>	0.2	0.21	-0.13	0.898
<i>SECURITY_DEPOSIT</i>	202.77	225.19	-0.65	0.517
<i># OPEN DAYS</i>	21.49	22.68	-1.10	0.271
<i>CANCELLATION_STRICT</i>	0.26	0.29	-0.89	0.371

# BEDS	1.81	1.96	-1.17	0.242
APARTMENT	0.6	0.61	-0.27	0.786
ENTIRE_HOME	0.61	0.64	-0.64	0.522

## 2) Sensitivity Analysis on Propensity Score Method (Rosenbaum bound tests)

For the propensity score estimation step, we included a rich set of variables and their interactions for our propensity score estimation step. Inclusion of a complete set of covariates reduced the chance of our main results being affected due to variables that were not accounted for when computing the propensity scores.

Despite the long list of included covariates for propensity score estimation, there is still a chance that there are omitted variables that affect one’s decision on professional program participation decision. As is commonly acknowledged, propensity scores are computed based on observed variables; therefore, there may be hidden bias if there are unobserved variables affecting the selection process (i.e., the treatment assignment) and the outcome variables simultaneously. To assess the sensitivity of our estimation to a potential hidden bias, we implement a widely-adopted approach of sensitive analysis—Rosenbaum bounds test (Rosenbaum 2002).

The logic of Rosenbaum bounds analysis is as follows. Suppose the participation probability of unit  $i$  is  $P_i(\text{treated}_i = 1|x_i, u_i) = f(\beta x_i + \gamma u_i)$ , where  $x_i$  and  $u_i$  are vectors of observed and unobserved variables, respectively.  $\gamma$  is 0 if there are no unobserved variables affecting treatment selection process. Units  $i$  and  $j$  with  $x_i = x_j$ , have the same probability of receiving the treatment if and only if  $\gamma(u_i - u_j) = 0$ . Rosenbaum bounds evaluate how much of a change in the ratio of odds of participation, due to unobservables, would be required to nullify the treatment effect identified with the propensity score method. The inference of the estimation results could inspire more confidence, if it would require a greater change in the odds ratio, caused by the unobservables, to overturn the estimated treatment effect.

The Rosenbaum bound test results are provided in Table 9. Since our main DiD analysis identified a positive effect of verified photos on property demand, we are more concerned with potential upward (positive) than downward (negative) bias in the DiD estimator. Hence in Table 3 we are mostly interested in the column of sig+. The results on Gamma suggest that even when a hypothetical unobservable increases the odds ratio to 1.4 times greater, our causal inference on the positive treatment effect of professional images on property demand, identified with the propensity score method, will be robust at the 95% confidence level (and it remains positive significant at the 90% confidence level until Gamma is increased to 1.55). The results suggest that for a positive estimated treatment effect on property demand to be overturned, the potential unobserved factors affecting the treatment assignment process would have to be large enough to increase the odds ratio of participation by at least 60%. Moreover, if we look at Hodges-

Lehmann’s estimates (Rosenbaum, 1993), they suggest a more robust result, as the upper (t-hat+) and the lower (t-hat) bounds do not contain 0 at least up until 1.6.

The results of our sensitivity analysis are on the same order of the increase in Gamma obtained in the extant literature (Sun and Zhou, 2014; Manchanda et al., 2015; Li et al., 2016; DiPrete et al., 2004), which reported Gamma ranging from 1.2 to 1.6. Though the treatment effect will be insignificant if the unobservable is large enough to change propensities further, this doesn’t mean that the propensity score method is invalid. This is because Rosenbaum bounds analysis gives us a lower bound of confidence on making causal inference in the worst scenario with hypothetical hidden selection bias—note that the hidden bias due to the unobservable does not necessarily exist. For this reason, we are confident that our study is robust, to some extent, to the hidden bias caused by hypothetical unobserved factors affecting the selection process.

**Table 9 Sensitivity Analysis: Rosenbaum Bound Test**

<b>Gamma</b>	<b>Significance Level</b>		<b>Hodges-Lehmann Point Estimate</b>		<b>95% Confidence Interval</b>	
	<b>sig+</b>	<b>sig-</b>	<b>t-hat+</b>	<b>t-hat</b>	<b>CI+</b>	<b>CI-</b>
<b>1</b>	0.000205	0.000205	13.9456	13.9456	5.50612	24.1935
<b>1.05</b>	0.00052	0.000075	12.9032	15.8307	4.83871	25
<b>1.1</b>	0.00119	0.000027	12.5	17.5824	3.22581	26.4631
<b>1.15</b>	0.002482	9.50E-06	11.2903	18.3333	2.2043	27.7778
<b>1.2</b>	0.004777	3.30E-06	10.2716	19.3548	1.25353	29.6461
<b>1.25</b>	0.008569	1.10E-06	9.05707	20.2419	-4.40E-07	30
<b>1.3</b>	0.014442	3.80E-07	8.06452	21.4286	-4.40E-07	30.8405
<b>1.35</b>	0.02304	1.30E-07	6.45162	22.4194	-4.40E-07	32.0079
<b>1.4</b>	0.035005	4.20E-08	6.45161	23.5526	-4.40E-07	33.1349
<b>1.45</b>	0.050917	1.40E-08	5.50612	24.1935	-4.00E-06	33.8095
<b>1.5</b>	0.071236	4.50E-09	5	24.7878	-1.0326	35.0824

<b>1.55</b>	0.096249	1.50E-09	4.07337	25.8064	-2.01613	35.7692
<b>1.6</b>	0.126037	4.70E-10	3.22581	26.7374	-3.10676	36.7374

---

Gamma: log odds of differential assignment due to unobserved factors

sig + (-): upper (lower) bound of significance level

t-hat + (-): upper (lower) bound of Hodges-Lehmann point estimate

CI + (-): upper (lower) bound of confidence interval ( $\alpha=0.95$ )

### 3) Addressing Concern of Inflated Long-term Effect: Running Analyses on Shorter-Period Sample

This is an alternative approach to address the concern of possible long-term inflation in the treatment effect if Airbnb’s search algorithm favors the properties with professional photos. To address the concern, we estimate our main DiD specification on a set of subsamples in which we included a shorter period span (8 months) for each treated unit:

Specifically, for each treated unit, the observation spans exactly 4 periods per property before and after treatment adoption. For the untreated (control) units, we used the full periods for estimation. This is because for untreated units, there is no such ‘reference period’ to which we can define the pre- & post- span.

In Table 10 we present the estimation results from the robustness analysis on the selected subsets. As shown, the estimated coefficients of the key variable, *TREATIND*, stays consistently positive and significant. The consistent estimated treatment effect from estimating shorter periods adds confidence that our main finding was not driven by long-term inflation caused by Airbnb’s search ranking algorithm.

**Table 10 DiD Robustness Tests: Selecting Shorter Periods of Samples (limit sample to 8 period/months span)**

<b>On subset:</b>		
<b>4 pre- and 4 post- treatment periods for treated &amp; whole-period for untreated units</b>		
<b>VARIABLES</b>	<b>ESTIMATES</b>	
<i>TREATIND</i>	10.33***	1.624
<i>log REVIEW_COUNT<sub>t-1</sub></i>	10.03***	0.768
<i>NIGHTLY_RATE</i>	-0.175***	0.0262
<i>INSTANT_BOOK</i>	4.436***	1.169
<i>CLEANING_FEE</i>	0.0898***	0.0125
<i>MAX_GUESTS</i>	-2.142*	1.034
<i>RESPONSE_RATE</i>	0.0767*	0.0348
<i>RESPONSE_TIME (minute)</i>	-0.00152	0.00136
<i>MINIMUM_STAY</i>	0.0279	0.0873
<i>SECURITY_DEPOSIT</i>	0.00298*	0.00143
<i>SUPER_HOST</i>	1.989*	1.003
<i>BUSINESS_READY</i>	0.902	0.919
<i>CANCELLATION_STRICT</i>	0.521	1.045
<i>HAS_RATING</i>	6.380	8.310

<i>HAS_RATING</i> × <i>COMMUNICATION</i>	0.204	1.336
<i>HAS_RATING</i> × <i>ACCURACY</i>	0.00741	1.056
<i>HAS_RATING</i> × <i>CLEANLINESS</i>	-2.015	1.035
<i>HAS_RATING</i> × <i>CHECKIN</i>	-1.142	1.398
<i>HAS_RATING</i> × <i>LOCATION</i>	0.805	1.074
<i>HAS_RATING</i> × <i>VALUE</i>	1.552	0.998
<i>INTERCEPT</i>	44.81***	6.246

---

Fixed Effect	Property
Seasonality	City-Year-Month

---

Observations	75406
R-squared	0.6943

---

Robust standard errors in parentheses; \* p<0.05 \*\* p<0.01 \*\*\* p<0.001

#### 4) Adding Interaction Terms with Meaningful Amenities

One concern regarding the self-selection issue is that properties with particular amenities may be more likely to adopt the treatment. For example, if some amenities make the properties more attractive in particular seasons (e.g., pool or beach in summer season), and the hosts adopt the treatment at that time, then some of the increase in demand would be brought by those attractive amenities. The effects of these amenities on demand (which are property-fixed characteristic) cannot be fully taken care of by the property fixed effect terms, as amenities' effect may be time-varying (for example, a pool has a greater effect in summer than in winter).

To address this concern, when estimating the propensity scores (for Propensity Score Weighting Strategy), we obtained additional data on property amenities and incorporated the amenity information (e.g.,

AC, pool, whether the property is close to a beach) to account for possible factors that may be correlated with both the property demand and with the hosts' treatment adoption decision in particular seasons that cannot be captured by property fixed-effects. In addition, in the model specification, we added interaction terms of dummy *AFTER* and meaningful amenities (e.g., pool, beach, AC) to account for the higher effect of these property time-invariant variables after treatment or in particular seasons.

In Table 11, we present the estimation results. Column (1) presents results from estimation Equation (2a), which is our main specification, and column (2) presents results from estimation Equation (2a), where we incorporated interacting *AFTER* with meaningful amenities. The consistent estimated results confirm a positive and significant treatment effect of more than an 8% increase in the property occupancy rate after controlling for area-specific seasonality as well as time-varying effect of particular time-invariant property amenities. In the panel of "Interacting with Meaningful Property Amenities", we present the coefficients of the interaction terms. As can be seen below, all coefficients are insignificant.

$$\begin{aligned}
 DEMAND_{it} = & INTERCEPT + \alpha_3 TREATIND_{it} + \lambda CONTROLS_{it} + SEASONALITY_{cym} \\
 & + PROPERTY_i + \varepsilon_{it}
 \end{aligned}
 \tag{2a}$$

$$\begin{aligned}
 DEMAND_{it} = & INTERCEPT + \alpha_3 TREATIND_{it} + \lambda CONTROLS_{it} + \vartheta_1 AFTER \times POOL \\
 & + \vartheta_2 AFTER \times BEACH + \vartheta_3 AFTER \times AC + SEASONALITY_{cym} \\
 & + PROPERTY_i + \varepsilon_{it}
 \end{aligned}
 \tag{2b}$$

**Table 11 Difference-in-Difference Model: The Impact of Verified Photos on Property Demand**

VARIABLES	Main DiD Model (Equation 2a)		Interacting with Amenities (Equation 2b)	
	ESTIMATES	Robust S.E.	ESTIMATES	Robust S.E.
<i>TREATIND</i> ( <i>TREAT</i> × <i>AFTER</i> )	<b>8.985***</b>	<b>1.660</b>	<b>8.668***</b>	<b>1.747</b>
<b>Property (Non-Photo) Characteristics</b>				
<i>log REVIEW_COUNT</i> <sub><i>t-1</i></sub>	9.375***	0.930	9.403***	0.933
<i>NIGHTLY_RATE</i>	-0.146***	0.0320	-0.149***	0.0320
<i>INSTANT_BOOK</i>	4.156**	1.361	4.168**	1.362
<i>CLEANING_FEE</i>	0.0808***	0.0184	0.0814***	0.0185
<i>MAX_GUESTS</i>	0.260	1.117	0.305	1.111
<i>RESPONSE_RATE</i>	0.0699	0.0430	0.0693	0.0430

<i>RESPONSE_TIME (minutes)</i>	-0.000477	0.00161	-0.000591	0.00161
<i>MINIMUM_STAY</i>	0.133	0.131	0.136	0.131
<i>SECURITY_DEPOSIT</i>	0.00177	0.00201	0.00171	0.00199
<i>SUPER_HOST</i>	3.801*	1.494	3.781*	1.494
<i>BUSINESS_READY</i>	1.806	0.985	1.791	0.984
<i>CANCELLATION_STRICT</i>	1.016	1.271	1.040	1.271
<i>HAS_RATING</i>	14.32	12.25	14.56	12.24
<i>HAS_RATING</i> ×	-0.212	1.420	-0.145	1.421
<i>COMMUNICATION</i>				
<i>HAS_RATING</i> × <i>ACCURACY</i>	0.878	1.211	0.837	1.211
<i>HAS_RATING</i> × <i>CLEANLINESS</i>	-1.344	1.133	-1.323	1.134
<i>HAS_RATING</i> × <i>CHECKIN</i>	-2.060	1.526	-2.118	1.526
<i>HAS_RATING</i> × <i>LOCATION</i>	-0.757	1.183	-0.767	1.183
<i>HAS_RATING</i> × <i>VALUE</i>	2.141	1.176	2.142	1.173

#### Interacting with Meaningful Property Amenities

<i>AFTER</i> × <i>POOL</i>			6.157	4.692
<i>AFTER</i> × <i>BEACH</i>			-10.77	10.83
<i>AFTER</i> × <i>AC</i>			1.034	3.990
<i>INTERCEPT</i>	30.06***	6.683	30.52***	6.662
Fixed Effect		Property		Property
Seasonality		City-Year-Month		City-Year-Month
Num. Observations		76901		76901
R-squared		0.6608		0.6609

\* p<0.05 \*\* p<0.01 \*\*\* p<0.001

### 5) Testing Changes in Property's or Host's Unobserved Quality (via multi-dimensional ratings)

One possibility for a self-selection issue is that properties or hosts are self-selected to adopt Airbnb professional images when there are some substantial changes in the hosting quality delivered to the guests. Such a change could be, for example, an upgrade of the house/room, or delivery of warmer/better services to the guest. This would introduce an upward bias in the estimated coefficient, as they happened at the same time as the treatment adoption. Although we are not able to observe and control for everything, there are a few things that help to control for or to alleviate this issue:

- 1) In the demand model, we add measurement of host responsiveness (host response rate and host response time) to address the concern that hosts being more responsible in the post-treatment period.
- 2) In the demand model, we add a complete set of multi-dimensional guest ratings to capture and account for any potential changes in the stay experience or in the hosting quality.
- 3) As we show below, for the properties that have review ratings available, we compare the average ratings of review along multi-dimensional aspects a few periods before the treatment versus after. The goal is to examine whether there are substantial changes, along with the adoption of treatment, in the property characteristics that may be unobserved to us but was captured in guest ratings.

To implement the robustness test in 3), we estimate the following specification:

$$Rating_{it} = INTERCEPT + \alpha_3 TREATIND_{it} + \lambda CONTROLS_{it} + SEASONALITY_{cym} + PROPERTY_i + \varepsilon_{it}$$

where the specification is the same as our main demand model, with the dependent variable replaced with one of the following guest ratings that can capture, to some extent, how good a stay or a host was: communication, accuracy, cleanliness, and check-in. Metrics on ‘cleanliness’ capture potential changes in the property, while metrics on ‘communication’ capture the hosting quality--the quality of communication between a host and his/her guest. The coefficient  $\alpha_3$  hence captures the changes in ratings along a particular dimension after treatment adoption.

As shown in Table 12, the coefficients of *TREATIND* are insignificant across all specifications, suggesting that for the treated units, before and after adopting verified photos, there was no substantial changes in the stay or hosting quality delivered to guests.

**Table 12 Robustness Test: Changes in Multi-dimensional Guest Ratings after Verified Photo Adoption**

VARIABLES	D.V.: Multi-dimensional Guest Review Rating			
	Communication	Accuracy	Cleanliness	Check-in
<i>TREATIND</i>	-0.0155 (0.0118)	0.0175 (0.0167)	-0.0261 (0.0193)	-0.0245 (0.0177)
<i>log REVIEW_COUNT<sub>t-1</sub></i>	-0.0371 (0.0189)	0.0277 (0.0265)	-0.0582* (0.0286)	0.0261 (0.0285)
<i>NIGHTLY_RATE</i>	0.00189*** (0.000297)	-0.000621 (0.000344)	0.00123* (0.000546)	0.000400 (0.000394)

<i>INSTANT_BOOK</i>	-0.0134 (0.0153)	0.0154 (0.0187)	-0.0174 (0.0197)	0.0162 (0.0158)
<i>CLEANING_FEE</i>	-0.000325 (0.000208)	0.000604 (0.000378)	-0.000343 (0.000665)	-0.000104 (0.000318)
<i>MAX_GUESTS</i>	-0.0101 (0.00728)	0.00335 (0.00814)	0.0449*** (0.0125)	0.0106 (0.00988)
<i>RESPONSE_RATE</i>	-0.000951** (0.000332)	0.000425 (0.000367)	-0.000782 (0.000504)	0.000130 (0.000357)
<i>RESPONSE_TIME</i> (minutes)	-0.00000573 (0.0000131)	0.000000617 (0.0000126)	-0.00000225 (0.0000163)	-0.0000180 (0.0000133)
<i>MINIMUM_STAY</i>	-0.00129** (0.000395)	-0.000554 (0.000612)	-0.000550 (0.000739)	0.000187 (0.000565)
<i>SECURITY_DEPOSIT</i>	0.00000253 (0.0000107)	0.0000174 (0.0000112)	0.0000104 (0.0000168)	-0.00000397 (0.0000101)
<i>SUPER_HOST</i>	-0.0246* (0.0117)	0.0169* (0.00826)	0.0188 (0.0154)	0.00510 (0.00905)
<i>BUSINESS_READY</i>	-0.00618 (0.00943)	0.0166 (0.0109)	0.0270 (0.0140)	0.00700 (0.00982)
<i>CANCELLATION_STRICT</i>	-0.0281 (0.0147)	0.0284 (0.0164)	-0.0152 (0.0200)	0.0236 (0.0194)
Constant	9.608*** (0.0759)	9.546*** (0.0919)	9.157*** (0.121)	9.511*** (0.106)
Fixed Effect	Property City-Year-	Property City-Year-	Property City-Year-	Property City-Year-Month
Seasonality	Month	Month	Month	City-Year-Month
Observations	45386	45386	45386	45386
R-squared	0.8771	0.8931	0.9155	0.8562

## **Can Lower-quality Images Lead to Greater Demand on AirBnB?**

**Shunyuan Zhang**

David A. Tepper School of Business  
Carnegie Mellon University, Pittsburgh, PA  
[shunyuaz@andrew.cmu.edu](mailto:shunyuaz@andrew.cmu.edu)

**Nitin Mehta**

Rotman School of Management  
University of Toronto, Toronto, Canada  
[Nmehta@rotman.utoronto.ca](mailto:Nmehta@rotman.utoronto.ca)

**Param Vir Singh**

David A. Tepper School of Business  
Carnegie Mellon University, Pittsburgh, PA  
[psidhu@cmu.edu](mailto:psidhu@cmu.edu)

**Kannan Srinivasan**

David A. Tepper School of Business  
Carnegie Mellon University, Pittsburgh, PA  
[kannans@cmu.edu](mailto:kannans@cmu.edu)

## Chapter 2

### Do Lower-quality Images Lead to Greater Demand on AirBnB?

#### 2.1 Introduction

AirBnB, a peer-to-peer short-term lodging marketplace provider offering near 5 million listings across 81,000 cities, has hosted over 300 million host arrivals since its start in 2008<sup>14</sup>. Previous studies on AirBnB suggest that the determinates in property demand include property type (apartment, house etc.), property size (number of bedrooms), property price, property location, guest reviews, host service, and property image quality (Zhang et al. 2008, Li and Srinivasan 2018, Li et al. 2016). Only a few factors, i.e., image quality, service, and price—are in immediate control of the hosts. The focus of this paper is on investigating how the AirBnB hosts make decisions on the quality of photographs to post.

Property images, as a way of providing visual information about a listing to the consumers<sup>15</sup>, can effectively increase the demand for an AirBnB property. In the context of photos, Zhang et al. (2018) investigated the impact of photo quality on AirBnB demand. Analyzing property images with a deep learning model, they classified the image quality along two levels, high-quality (professional images) and low-quality (amateur images)<sup>16</sup>. They found that the photographs played the role of an advertisement and found the quality of photographs had a significant impact on the property demand, with high-quality images increasing a property's present booking by 14.3%. The strong advertisement impact of images on present demand arises because of the special context of AirBnB. Specifically, AirBnB properties, relative to hotels' rooms that are fairly standardized, have a large variation in terms of styles, characteristics, and hosts. Moreover, most consumers, with a large number of alternatives on the platform, do not repeatedly choose and stay in the same property. In addition, analyzing written details of the property and comparing those across properties can be very onerous and time consuming. Consumers, in order to ease decision-making, hence rely heavily on visual information, which can be easily accessed and processed, to skim through and quickly compare lots of the lodging alternatives. Based on the images, consumers form an expectation on the quality of their stay at the property and accordingly make their decision on which host's property to stay.

---

<sup>14</sup> <https://press.atAirBnB.com/fast-facts/>.

<sup>15</sup> In this paper we use listing and property interchangeably, guest and consumer interchangeably.

<sup>16</sup> The image quality is classified with a convolution neural network (CNN) that analyzes the basics very high-dimensional pixel information in the images in training set, extracting a hierarchical set of image features that have the most prediction power on the image quality label.

Despite the importance of the quality of property images in enhancing the demand, the reality is that there exist a large number (approximately two-thirds) of amateur (low-quality) images on AirBnB. One possible explanation for the low adoption of high quality images is that high quality images are costly for the hosts, as most of them would be amateur photographers. However, this does not completely explain the result—in 2011, AirBnB offered highest quality professional images to all the hosts by sending their professional photographers to the property and shoot, process and post the photos for the hosts. Not only was this program free for all hosts, it also required very little effort from the host's part. Moreover, all hosts were made aware of this program (explain how they were made aware). To AirBnB's surprise, only thirty percent of the hosts used the AirBnB professional photography program after its launch. This result is intriguing since we would expect the demand to explode when a high quality product/service is offered for free (Shampanier et al. 2007).

In this paper, we provide an explanation for hosts' behavior. We posit that the host's decision on the quality of images to post depends not only on the advertising impact of the photos on present demand and the cost of photos, but also on the impact of the photos on the satisfaction level of the guest post consumption, which would then in turn impact the future demand. The last point follows from the reference dependence literature which suggests that the images create a reference point for the guest in terms of what quality to expect, and their satisfaction level post consumption individual's utility from consuming a product depends not only on the realized outcome, but also on her reference point—the individual's pre-consumption expectation (Kahneman and Tversky 1979, Tversky and Kahneman 1991, Koszegi and Rabin 2006). Particularly, individuals tend to react severely to a 'dissatisfaction or disappointment gap'—i.e., when the actual outcome turn out to be worse than the expectation (Genesove and Mayer 2001). Thus, some hosts would be hesitant to post high quality photographs (even if they were free) because they can create unrealistically high expectations for the consumers, especially if the actual property is not as good as what the photos portray and if the hosts are unable to provide a high level of service to match the high expectations. As a result, the consumers' satisfaction level would decrease, and they would either leave a bad review or would not write any review at all<sup>17</sup>; and this lack of reviews would in turn adversely impact the future demand of future guests of that property.

---

<sup>17</sup> Dissatisfied guests instead of writing bad reviews, they tend to not to write a review. As previous studies on online reviews suggest, the 'silence' in online reviews actually reflects customer dissatisfaction (see Dellarocas and Wood (2008)'s work on 'the sound of silence'), because customers with bad experience tend to choose not to leave a review (Masterov et al. 2014, Nosko and Tadelis 2015). This is consistent with the observations that online reviews tend to be positive, possibly because giving a negative rating is costly for the consumers. Particularly, with a field experiment on AirBnB that involves encouraging consumers to leave reviews, Fradkin et al (2018) found that guests get more utility from leaving a positive review, and they also don't like to misrepresent their experiences. As a result, a guest is less likely to leave a review, if she is unsatisfied with her stay experience.

In summary, professional property images, although more expensive, can help generate bookings for the AirBnB hosts in the current period, since consumers with imperfect information rely heavily on images to make their lodging decisions. On the other hand, professional images can lead to a dissatisfaction gap if the actual property is not as good as what the professional images portray or if the hosts are unable to provide a high level of service to match the high expectations. Our goal is to disentangle the aforementioned factors that influence the host's decision on the type of photographs to post, and to explore policies that platforms such as AirBnB can employ to improve the hosts' adoption of professional photos and thereby improve the profitability of both the hosts and AirBnB. To achieve this goal, we have the following objectives, which we explain as follows.

### **2.1.1 Research Objectives and Main Findings**

The *first* objective of this paper is to model hosts' periodic (monthly) decisions on the quality of property images to post, and the quality of service to provide. The image decision entails choosing between three quality levels of images—low, medium, and high<sup>18</sup>. The service decision entails choosing between two levels of service: high and low. The image decision affects the host's profits in the short run through the costs associated with preparing, shooting and editing the particular quality-level images, and through the impact of images on the present demand. And it affects the host's future profits via the following mechanism: professional images come with a risk of increasing the guests' dissatisfaction gap. This decreases their likelihood of writing reviews, which then negatively impacts the future demand. The service decision impacts the host's profits in the short run through the costs (good services come with a cost), and in the long run by impacting the guests' satisfaction level and their subsequent likelihood of writing reviews.

To achieve the first objective, we need to model both the guests' decisions on which property to choose (the demand side) and the hosts' decision on the quality of images to post and the level of service to provide (the supply side). The property demand is a function of property characteristics including property images, number of reviews and prices. The property supply function models how hosts make images and service decisions, taking into account the impact of their actions on the current and future utility.

Regarding the demand side, we estimate a random-coefficient logit model (Berry et al. 1995) using AirBnB properties' aggregate monthly market-share data. The guests' utilities are modeled as functions of property characteristics which include property images, number of reviews and property prices. Heterogeneous consumers form expectations on the lodging alternatives, based on their preferences of the property attributes, which include property image quality, prices and the number of reviews. Regarding the

---

<sup>18</sup> This is the aggregate photo quality decision across all photos posted.

supply side, we model the hosts' image quality and service decisions as outcomes of their long-term profit maximization.

The *second* objective is to estimate the model using a panel data consisting of 900+ individual AirBnB hosts' choices of images and service over time. The data contains rich information on property characteristics, property reservation days and monthly revenues, and guest's reviews. There are two unique features in this data: 1) we observe the dynamics in property images and whether an option of AirBnB's free professional photos was available to the host, from which we know hosts' periodic image decisions and infer the associated costs<sup>19</sup>; 2) the periodic reviews a property received from its guests, from which we infer the dissatisfaction gap between image-induced expectation and realized property as well as the invested effort in service. Our key empirical findings are: a) guests have heterogeneous and correlated preferences on property attributes. Particularly, guests who value professional images more, also value the number of reviews more. Thus, for a property that faces such a pool of consumers, using professional images may have a high marginal effect on generating booking in current period. Yet, the 'penalty' in the future is also likely to be high, as these consumers value highly the number of reviews. b) hosts have a considerable degree of heterogeneity in their ability (cost) in investing in service and in values of their outside options. Such heterogeneity results in hosts self-selecting to choose different quality levels of images.

Our *third* objective is to explore image-related policies that a peer-to-peer platform such as AirBnB can employ to effectively improve the overall property performance and service quality. To do so, we run two counterfactuals. In the first counterfactual, we examine three image policies. The first policy is the same as AirBnB's professional photography program where it provides the highest-level professional images for free to all hosts (current policy). The second policy is an alternative policy in which AirBnB instead provides medium quality-level images for free to all hosts (proposed policy 1). In both policies A and B, we allow for hosts to self-select on whether they would adopt the free AirBnB program or not. The third policy is the baseline policy in which AirBnB were not to offer any photography service to the hosts (baseline). We find that, both policies A and B, compared to the baseline policy, substantially improve the average property demand across all properties. Interestingly, policy B, though dominated by the policy A in the short-run (for the first four periods), outperformed policy A in the long-run (12.4% vs 7.6%).

Our results indicate that, medium-level images, compared to high-level images, despite forming a smaller expected utility for the consumers, has a greater effect on property booking in the long-run as they,

---

<sup>19</sup> AirBnB offered professional photos during our observation window. However, AirBnB's photography program says it can be offered to the same listing for once. Hence, for roughly one-third of the properties in our sample, which were observed to already have AirBnB-offered photos by the start of our observation, if they were to change their photos with another batch of professional photos, they would incur a cost on their own.

with lower risks of creating a dissatisfactory gap, help hosts to obtain new reviews. Moreover, individual hosts who might end up using amateur (bad-quality) images to avoid the dissatisfactory gap under current policy, now use free medium-level images to make more revenues under proposed policy 1. In the second counterfactual, we explore an alternative policy in which AirBnB were to offer a menu of image quality choices for free. The menu includes both high- and medium- level of property images (images examples are provided) and allow the hosts to self-select which program they want. Comparing with the proposed policy in the first simulation, we find that this policy performance the best in the long-run by improving average property demand by 16.2%.

### **2.1.2 Literature Review**

We start with discussing the relatively new stream of literature on the sharing economy platform AirBnB. A few recent studies looked at the reputation system, i.e., consumer reviews, on AirBnB (Zervas et al. 2015, Proserpio et al. 2017, Fradkin et al 2018). These studies mainly focused on documenting a particular aspect existing in the reputation system, e.g., reciprocity, without investigating how the reputation system affects consumers' choices of demand and drives hosts' choices of supply. Another stream of studies investigated the impact of AirBnB's entry on the incumbent lodging industry, namely hotels (Li and Srinivasan 2018, Zervas et al. 2017). There are a couple of distinctions between these papers and our paper. 1) these papers focused on quantifying the impact of AirBnB's supply on hotels, while ours aims to understand how AirBnB's supply choices (including image and service choices) are endogenously determined by the hosts. 2) these papers largely treat AirBnB properties as a homogeneous party (or categorized properties into a few sub-groups based on the property type), without taking into account the heterogeneity across the properties and hosts. An exception is the work of Farronato and Fradkin (2018), which incorporates hosts' heterogeneous variable costs in modeling their supply decisions. However, they did not consider the heterogeneity in terms of a property's quality. We argue that, given different properties at different states, hosts' supply choices could be different even they have the same variable cost. (3) none of these studies investigate the role of property images. Our paper differs from the existing study as we have access to a panel data that consist rich information on hosts' periodic image choices. An exception is Zhang et al. (2018), where property images are analyzed and scored by the image quality. However, they employed a quasi-experimental method (difference-in-difference) to make a causal link between property images and property demand, treating image choices as if they were exogenously given. In contrast, we endogenize the image decisions and explain how host heterogeneity drives the observed choices of images and service. This heterogeneity nature of peers, though unobserved, plays a significant role in driving the market equilibrium outcome. This distinction allows us to estimate a more comprehensive model of hosts' decisions, to resolve the observed puzzle—why many hosts did not use AirBnB's professional photos

despite free, and to answer our research question posed in section 1.2—what else can AirBnB do with its image-related policy to improve the market equilibrium outcomes?

As discussed above, guests' post-consumption behavior, influenced by potential dissatisfaction gap, affects hosts' pre-consumption decisions. Here we briefly discuss how behavioral argument of consumer dissatisfaction gap is related to and used in our paper. The idea that utility depends on both the actual outcome and the alternative outcome that could have occurred has been posited in a large body of literature in psychology and behavior economics (Gul 1991; Kahneman and Tversky 1979). Specifically, studies on reference dependence suggest that individual's utility from consuming a product depends not only on the realized experience, but also on the reference point—the individual's pre-consumption expectation (Koszegi and Rabin 2006). Hence, with the same actual outcome, individuals may react differently, if they had different expectations and experienced different 'gaps' (Mas 2006). We conjecture that in the context of AirBnB, guests' post-consumption behavior is affected by both the realized stay experience after their arrivals and the expected experience they had when seeing the property images. Particularly, since people react more severely to a 'loss' than to a 'gain' (Genesove and Mayer 2001), if the reality did not meet the expectation, a dissatisfaction gap significantly reduces the guests' likelihood of writing a review.

## **2.2 Research Context and Descriptive Statistics**

### **2.2.1 Research Context**

Our research context is AirBnB—one of the largest sharing-economy platform for peers to list their spare rooms and to find short-term lodgings. AirBnB now offers near 5 million listings in over 81,000 cities. Since its foundation in 2008, AirBnB has hosted more than 300 million guest arrivals. AirBnB makes revenue from charging a service fee of 9~12% proportional to each transaction.

### **2.2.2 Data Description and Measures of Key Variables**

Our sample consists of 958 randomly selected AirBnB properties in Mahanttan, New York. For each property in the sample, we collected property time-invariant characteristics, including property's location, type, size, and capacity. We constructed a panel data of the 958 properties spanning 12 months (January 2016-December 2017). For each property in each month, we obtained dynamic information about the property's demand (i.e., the number of reservation days) as well as the property's supply (i.e., whether a listing was active in a particular month). Such dynamic information also includes property nightly rate, guests review, and property images. Below we describe the definitions of key variables used in our analyses.

#### **Property Characteristics**

We obtained property characteristics, defined by the following variables: 1) *EntireHome*, which equals 1 (0) if the property is listed as an entire (shared) place, 2) *Apartment*, which equals 1 (0) if the property is

an apartment (or not, e.g., condo or house), 3) *Bedrooms*, *Bathrooms*, and *Beds*, which indicates the number of bedrooms, bathrooms and beds, respectively, 4) *MaxGuests*, which indicates the maximum number of accommodated guests), 5) *DriveTime*, the driving commute time (in minutes) from each property's address to the downtown area). The driving time is further scaled by 1/10 in the analyses, and 6) *WalkScore*, a score 0-100 based the evaluation of the available nearby amenities such as restaurants, malls, parkings etc., and 7) the area code associated with each property (Manhattan is categorized into 10 subareas or neighborhoods)<sup>20</sup>.

### **Property Reservation**

We purchased listing-level reservation data from a third-party company that specializes in collecting AirBnB property booking data. One unique feature in the reservation data is that, they distinguish real booking (days when the property was booked by a guest) from blocking (days when hosts marked the property as 'unavailable'). Since blocking days do not reflect the actual property demand, we used only the number of reservation days to construct our demand measurement. For each property  $i$  in month  $t$ , variable  $ReserveDays_{it}$  indicates the number of days that  $i$  was booked in that period.

**Market Share and Market Size** A property's market share reflects its demand on the lodging market. It's defined as the number of property-nights sold in each period (month), divided by the market size. Market size measures the total number of nights (including AirBnB property, hotel, and other alternatives) that could be possibly sold to the travelers, approximated from combining New York City tourism trend report in 2016 and the distribution of hotels across the five boroughs in NYC<sup>21</sup>. In our study, we used a constant market size of 2,400,000 across the periods. The seasonality trend is captured through a series of period fixed effects we incorporated in the property demand equation (see section 4.1). Variable  $MarketShare_{it} = ReserveDays_{it}/MarketSize$  indicates property  $i$ 's market share in month  $t$ .

### **Property Nightly Rate**

Property  $i$ 's nightly rate,  $NightlyRate_{it}$  is computed as the average of daily prices over the days in period  $t$ . We further take the logarithm form of the average nightly rate. As in many other markets, property prices

---

<sup>20</sup> In Manhattan, the 10 areas are: Central Harlem, Chelsea and Clinton, East Harlem, Gramercy Park and Murray Hill, Greenwich Village and Soho, Lower Manhattan, Lower East Side, Upper East Side, Upper West Side, Inwood and Washington Heights. Refer to the following website for more details: <https://www.health.ny.gov/statistics/cancer/registry/appendix/neighborhoods.htm>

<sup>21</sup> 60.5 million travelers visited New York in 2016. About 48% of the market are day trips hence do not consume any lodging, and the remaining on average consumed 2.4 nights each trip. [https://www.nycgo.com/assets/files/pdf/new\\_york\\_city\\_travel\\_and\\_tourism\\_trend\\_report\\_2017.pdf](https://www.nycgo.com/assets/files/pdf/new_york_city_travel_and_tourism_trend_report_2017.pdf); <http://mycrains.crainsnewyork.com/stats-and-the-city/2017/tourism/hotel-occupancy-rate-by-year>.

Though we don't have tourism data on the borough level, we can approximate the market size for each borough from the hotel distribution in NYC. This is because hotels choose locations with high demand—how the hotels geographically distribute capture how the demand distribute across the areas.

may be correlated with random shocks in the demand which are unobserved to the researchers. To address the endogeneity issue, as we will describe in section 4, a set of instruments were used. These instruments are correlated with  $NightlyRate_{it}$ , yet should be uncorrelated with demand shocks in the current period  $t$ . We also include the number of reviews in the previous period in the instrument variable set, as the number of reviews may affect how a property's nightly rate is set, yet it is uncorrelated with the aggregate demand shocks in current period. Further, following Li and Srinivasan (2018), we collect local (i.e., in the same zip code) rental information from Zillow to serve as an instrument for the monthly average Airbnb property nightly rate<sup>22</sup>. Finally, we include local utility fee in the set of instruments for  $NightlyRate_{it}$ , as the utility cost can influence the property price however is unlikely to be correlated with unobservables that are correlated with property demand. Specifically, we obtain average residential electricity rates by zip code<sup>23</sup>.

### **Property Active**

A host can choose to make a property temporarily 'inactive'. For example, if a host feels constantly managing a listing (e.g., checking property page and updating property availability calendar) is a little demanding while the returns from managing is little, then he may choose to "snooze/un-list" the listing for a while. Such, the property will be temporarily not viewed. If afterwards, he/she decides to activate the listing, all the current records remain the same. Variable  $Active_{it}$  equals 1 (0) if property  $i$  was active (inactive) in period  $t$ .

### **Property Review Count**

We collected data property reviews posted on property page. Specifically, we count the accumulated number of posted reviews for property  $i$  till the beginning of period  $t$ ,  $NumReview_{it}$ . Since number of reviews may be correlated with unobserved property characteristics (as we will describe in section 3, consumers' likelihood of writing a review is influenced by the property quality), in the demand function, we take the approach of instrument variables to deal with the endogeneity issue. Specifically, two variables serve as instruments for  $NumReview_{it}$ . The first is BLP instruments—the sum of number of review of other properties and the second is the number of hotel rooms in each zip code. The latter serves as an instrument as the supply from hotels is unlikely to be correlated with a property's unobserved attributes, however can influence the evolution of  $NumReview_{it}$  through competition that affects the number of bookings a property can receive in a month.

---

<sup>22</sup> See Zillow Rental Index for more details: <http://www.zillow.com/research/data/#rental-data>.

<sup>23</sup> National Renewable Energy Laboratory provides average residential, commercial and industrial electricity rates by zip code for both investor owned utilities (IOU) and non-investor owned utilities by combing data from [ABB, the Velocity Suite](#) and the [U.S. Energy Information Administration dataset 861](#). For more details, see <https://openei.org/datasets/dataset/u-s-electric-utility-companies-and-rates-look-up-by-zipcode-2016>.

## Property Ratings

We also obtained monthly data on the lodging experiences, rated along multiple dimensions, from peers who have stayed at the property. Specifically, for each property  $i$  in month  $t$ , guests can rate the property, on a 0-10 scale, on its accuracy in description, cleanliness, host communication before/during the stay, check-in smooth/convivences, overall value/experience, and location.

## Host Response Time and Service Effort

Airbnb's algorithm automatically records and compute the average time (minutes) that a host responded to the consumers in the past 30 days, denoted by *ResponseTime*. The algorithm tracks the communicate between a host and his/her guest during the process of making a reservation, prior to a stay, and during the stay. Such communication could include requesting information about the property, asking details regarding check-in, or any question that a guest may ask during a stay. From the html source code of each property page, we obtain the response time (as in minutes) that the algorithm rated each host for every month. We further categorize *ResponseTime* and create dummy variable *HighEffort* equal 1 if and only if *ResponseTime* is less than 1 hour<sup>24</sup>. Hence, in our study, we look at two different levels of effort—high- and low- effort. Since effort is unobserved to researchers (also unobserved to a guest ex-ante), we use a metrics that measure host's responsiveness to proxy for the level of effort that she invested in a particularly period., Though responding to a guest within 1 hour does not guarantee that a problem/question is resolved within 1 hour, this at least reflect how serious (the attitude) the host is about the guests' communication.

## Property Photos

AirBnB hosts post photos of their place on the property page. To capture the dynamics in the set of photos when hosts updated their property images, we measure the aesthetic quality of images as a time-variant variable. Specifically, leveraging computer vision techniques, we built a scalable deep learning model that automatically classifies any property image into one of three categories—namely, “high-quality”, “medium-quality”, and ‘low-quality’. The set of photos posted for property  $i$  in period  $t$  was then represented by its average image quality. For example, if property  $i$  had 10 images in period  $t$ , with 8 images classified as high-quality, 1 image classified as medium-quality, and 1 image classified as low-quality, then the average image quality  $ImageQuality_{it} = (8 * 1.0 + 1 * 0.5 + 1 * 0)/10 = 0.85$ .

Next, in our structural model, we discretized the image quality and categorize images into low-, med, and high- 3 quality levels. Then dummy variable *HighImage<sub>jt</sub>* is 1 if and only if  $ImageQuality_{it} > 0.75$ . Dummy variable *MedImage<sub>jt</sub>* is 1 if and only if the  $0.75 \geq ImageQuality_{it} \geq 0.5$  Lastly, dummy

---

<sup>24</sup> Our results stay qualitatively consistent when we use different criteria (i.e., setting the threshold at ‘responding within a few hours’).

variable  $LowImage_{jt}$  is 1 if and only if  $ImageQuality_{it} < 0.5$ <sup>25</sup>. Hence, for any property  $i$  in any period  $t$ , its image can be expressed with a tuple of binary variables  $(MedImage_{jt}, HighImage_{jt})$ . Property  $i$  had low-level of quality property images in period  $t$  if and only if  $(MedImage_{jt}, HighImage_{jt}) = (0,0)$ .

#### *A Deep Learning Classifier to Measure Image Quality*

We leverage techniques from computer vision and deep learning to build a classifier that, for any given input property image, predicts its image quality as high- versus low- quality. Specifically, we first construct a training set consisting of 3,000 (stratified) randomly selected AirBnB property images, with each the image quality is manually evaluated and labeled by five Amazon Mechanical Turkers (AMT)<sup>26</sup> on a 1-7 score. The image quality for each image is then computed as the mean score averaged across the scores assigned the five raters who evaluated the images. We discretize the image quality to create a quality label for each image, with label of ‘low-quality’, ‘medium-quality’, and ‘high-quality’ corresponds to an average score that are in the range of 1-3, 3-5, and 5-7, respectively.

Next, we build a Convolutional Neural Networks (CNN), a deep learning framework with a series of breakthroughs in vision tasks such as image classification and object recognition (Krizhevsky et al. 2012, Simonyan and Zisserman 2015). We then apply supervised learning and train the CNN classifier on the collected training set. The classifier is optimized by extracting image features that have predictive power on its label (image quality) and by learning the relationship between extracted features and the label. The classifier achieved a high accuracy of 86.7% on a hold-out set. Lastly, we apply the optimized CNN classifier to all property images in our sample to automatically predict the image quality label for each. In appendix, we provide detailed description on machine learning steps and on the architecture of the CNN classifier as well as technical notes on the training process.

Finally, Table 13 summarizes the statistics for the key variables.

**Table 13 Summary Statistics of AirBnB Properties**

<b>Variables</b>	<b>Observations</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
<i>EntireHome</i>	11496	0.654489	0.475555	0	1

<sup>25</sup> Alternatively, we could use mode (i.e., majority) of the image quality to represent to quality level of images associated with a property in a period. We obtained consistent results. This is because we observe that hosts post images that have quality concentrating in the same level—there is very limited behavior of mixing different image quality levels. In the given example, using mode quality would give us a high-level quality for the specific set of images, where the majority (8 out of 10) of the images are high-quality. Using average quality would also give us a high-level quality, as the average image quality is above the cutoff for high-level (i.e., 0.75). In regard to the thresholds of discretizing the average image quality, we tried other partition points (e.g., 0.7 as cutoff for high-level and 0.4 as cutoff for medium-level), the results remain qualitatively unchanged.

<sup>26</sup> <https://www.mturk.com/>.

<i>Apartment</i>	11496	0.938413	0.240414	0	1
<i>Bedrooms</i>	11496	1.180585	0.733769	0	6
<i>Bathrooms</i>	11496	1.094468	0.365832	0	4
<i>Beds</i>	11496	1.655532	1.051104	0	9
<i>MaxGuests</i>	11496	3.187717	1.901601	1	16
<i>MinimumStay</i>	11496	2.766962	2.865382	1	35
<i>HostExperienceYear</i>	11496	3.144262	1.481851	1	8
<i>WalkScore</i>	11496	98.17537	4.010409	62	100
<i>DriveTime (minutes)</i>	11496	13.57516	11.28168	1	56
<i>LocalUtilityRate</i>	11496	0.2163	0.043404	0.158837	0.249082
<i>LocalRentalIndex (Zillow)</i>	11496	2517.88	433.3505	1395	4525
<i>ReserveDays</i>	8622	16.12097	10.92966	0	31
<i>NightlyRate</i>	8622	228.6123	261.8237	28	5000
<i>Active</i>	11496	0.7621	0.433032	0	1
<i>NumReviews</i>	11496	40.63718	44.58707	0	383
<i>OverallRating</i>	9245	92.05116	5.606863	50	100
<i>CommunicationRating</i>	9245	9.724608	.4868291	8	10
<i>AccuracyRating</i>	9245	9.435803	.630361	7	10
<i>CleanlinessRating</i>	9245	9.116387	.8229696	5	10
<i>CheckinRating</i>	9245	9.637642	.5566269	6	10
<i>LocationRating</i>	9245	9.44186	.6645177	6	10
<i>ValueRating</i>	9245	9.136182	.6105189	6	10
<i>HighEffort</i>	8622	0.4066342	0.491234	0	1
<i>HighImg</i>	11496	0.227561	0.425607	0	1
<i>MedImg</i>	11496	0.16365	0.429303	0	1
<i>LowImg</i>	11496	0.608789	0.499669	0	1

---

### 2.2.3 Reduced-form Evidence

We explore patterns in the data and presents a series of reduced-form analyses that suggest 1) both image quality and number of review have positive impact on AirBnB property’s present booking (demand) and 2) 4) high-quality of images come with a risk of creating a ‘negative/dissatisfaction gap’ for consumers and could adversely affect the future demand.

### 2.2.3.1 Regressing Demand: The Impact on Images and Reviews on Property Bookings

We run the following regression, as specified in Equation (3), to see how image quality and the number of review affect a property’s present demand.

$$ReserveDays_{jt} = Intercept + \beta_1 ImageQuality_{jt} + \beta_2 NumReviews_{jt-1} + \beta_3 Controls_{jt} + Property_j + Period_t \quad (3)$$

where dependent variable—property demand—*ReserveDays<sub>it</sub>* indicates the number of reserved days for property *j* in period *t*. *ImageQuality<sub>jt</sub>* refers to the property’s image quality in period *t*, level, with low-quality serving as the baseline (i.e., its coefficient is normalized to zero). *NumReviews<sub>it-1</sub>* indicates the number of reviews that property *i* has accumulated till the end of period *t-1* (i.e., till the beginning of period *t*). Lastly, *Controls<sub>it</sub>* is a vector of control variables. *Period<sub>t</sub>* are time fixed-effects are incorporated to capture seasonality patterns in property demand.

As can be seen in Table 14, improving the quality of posted property images will, else being equal, lead to a greater property demand for the current period. The positive coefficient of *NumReviews* suggests the number of reviews is a key driver in generating property bookings, with a greater significance than the coefficient of image quality. Hence, in the long-run, to consistently get booking, it is essential for a host to be able to grow the reviews. This is particularly crucial, when the peers of the host are accumulating more reviews.

**Table 14 Regress Property Bookings on Image Quality and Number of Reviews**

VARIABLES	D.V. # Reservation Days Equation (3) <sup># +</sup>
<i>NumReviews</i>	0.105*** (0.00524)
<i>MedImage</i>	0.942** (0.3106)
<i>HighImage</i>	1.584*** (0.425)
<i>NightlyRate</i>	-0.0095*** (0.00284)
Observations	8622
<i>R</i> <sup>2</sup>	0.6903
Fixed Effect	Property

Seasonality	Monthly
Robust Standard errors in parentheses	
* p<0.05 ** p<0.01 *** p<0.001	
# property time-varying characteristics (overall rating, service effort, MaxGuests, and MinStays) are controlled.	
+ model is regressed over samples when a property was ‘active’.	

### 2.2.3.2 Law of Motion: The Impact of Image Quality on Guests’ Post-Consumption Review-Writing Behavior

We present evidence that higher-quality of images can reduce the guests’ post-consumption satisfaction and hence reduce their probability of writing a review. As explained in section 1, on Airbnb, guests who are unsatisfied tend to walk away without writing a review. Moreover, a departure from the expected outcome influences one’s post-consumption satisfaction (Koszegi and Rabin 2006). Hence, to capture how an expectation-realization gap plays a role in affecting Airbnb guests’ post-consumption satisfaction, in Equation (4) we implement a logistic regression on the probability that a consumer, upon his/her stay, will write a review, as a function of the ‘gap’ between expectation and realization and other relevant factors.

$$WriteProb_{jt} = \alpha_0 + \alpha_1(PropertyQuality_j - ImageQuality_{jt}) + \alpha_2Effort_{jt} + \alpha_3NightlyRate_{jt} + \alpha_4Control_{jt} + PropertyQuality_j + Period_t \quad (4)$$

where the dependent variable  $WriteProb_{jt}$  is measured the proportion of bookings occurred in period  $t$  for property  $j$  that led to a review. Key coefficient  $\alpha_1$  captures the impact of the realization-expectation gap, the service effort, the property’s nightly rate on guests’ likelihood of writing reviews after their stays.  $Period_t$  are time fixed-effects are incorporated to adjust the property price based on seasonality and to capture patterns that possibly correlate with the overall guests’ writing review behavior. Dummy variable  $Effort_{jt}$  is 1 if and only if the guests stayed in property  $j$  during time  $t$  were provided with a high-level effort of service (i.e., we normalize the coefficient for low-level service effort to 0). The control variables include the number of reviews  $NumReview_{jt}$ , overall review rating  $OverallRating_{jt}$ , and time-variant property characteristics such as maximum number of accommodated guests  $MaxGuests_{jt}$ , and minimum number of stay nights  $MinStays_{jt}$ .

Two points are worth noting in the realization-expectation gap,  $PropertyQuality_j - ImageQuality_{jt}$ . First, it allows bi-directional departures (i.e., the realized outcome—property quality could exceed or not meet the expected outcome—image quality). This bi-directional gap is consistent with hosts and guests’

beliefs or experiences that a positive gap ( $PropertyQuality_j > ImageQuality_{jt}$ ) increases a guest's post-consumption satisfaction while a negative gap ( $PropertyQuality_j < ImageQuality_{jt}$ ) decreases the satisfaction<sup>27</sup>. Second, for the same property, we assume its quality (realization) is time-invariant across the one-year panel in our study.

Re-writing Equation (4), we obtain the following property fixed-effect logistic regression:

$$\begin{aligned}
 WriteProb_{jt} &= \alpha_0 + a_1PropertyQuality_j + PropertyQuality_j - \alpha_1ImageQuality_{jt} \quad (5) \\
 &+ \alpha_2Effort_{jt} + \alpha_3NightlyRate_{jt} + \alpha_4Control_{jt} + Period_t \\
 &= \alpha_j + \alpha_{high}HighImage_{jt} + \alpha_{med}MedImage_{jt} + \alpha_{effort}Effort_{jt} \\
 &+ \alpha_{price}NightlyRate_{jt} + \alpha_4Control_{jt} + Period_t
 \end{aligned}$$

where  $\alpha_j = \alpha_0 + a_1PropertyQuality_j + PropertyQuality_j$  captures property fixed effect. Dummy variables  $HighImage_{jt}$  and  $MedImage_{jt}$  equals 1 if and only if the aggregated image quality level for property  $j$  in period  $t$  is high-quality, and medium-quality, respectively. That is, the baseline image quality here is low-level, and  $\alpha_{high}$  ( $\alpha_{med}$ ) captures the impact, for the same property, of updating images from low-level to high-level (medium-level) on guests' likelihood of writing a review.

As can be seen in Table 15, key coefficients  $\alpha_{high}$  and  $\alpha_{med}$  are negative, suggesting that higher-quality property images reduce the likelihood that a guest, upon his/her stay, i.e., having observed the realized quality of property, will write a review for that property. Particularly, using high-quality of images has a greater negative impact than using medium-quality images ( $-0.682 < -0.441$ ) on generating new reviews from the guests. In addition, as we anticipated, the positive estimated coefficient for service effort— $\alpha_{effort}$ —suggests that investing a high-level effort in providing good service to the guests can effectively increase their likelihood of writing a review. Interestingly, the coefficient for property's average nightly rate— $\alpha_{price}$ —is insignificant at the 0.05 significance level, suggesting that once controlling for image quality and service effort, the seasonality property price does not play a role in affecting guests' post-consumption satisfaction and their likelihood of writing reviews. The explanation is that, though property price is correlated with property quality, the factors that seasonality-adjusted price captures, such as the property's size, type, amenities, and location, are listed on property page and known to the guests

---

<sup>27</sup> See hosts' discussions on their strategic thinking on how a positive or negative gap would influence the satisfaction: <https://airhostsforum.com/t/worth-paying-a-photographer/12724/15>. In additional, we observed, from the guests' textual comments, that exceeding the guests' expectation with a property quality better than the image quality, can improve guests' satisfaction.

beforehand. Consumers' perceptions about these factors do not change before and after they have arrived in the property. As a result, they do not have impact on the guests' post-consumption satisfaction.

The results in Table 14 combined with Table 15 suggest two trade-off problems for an Airbnb host. First, if a host uses higher-quality images to make the property look nice, he/she will attract more property bookings. However, this may adversely impact the property demand in the future if he/she is unable to get new reviews from the guests as he/she cannot deliver the stay experience (in terms of property quality and/or service quality) that meets the guests' higher-expectation. Second, though a host can purposefully decrease the expectation for the guests by using low-quality images to improve the guests' post-consumption satisfaction, he/she may be unable to effectively generate new reviews from guests as a result of few number of transactions (property bookings) occurred.

**Table 15 Law-of-Motion: Regressing Review-Writing Probability**

Equation (5)	
VARIABLES	D.V.: WriteProb <sup>(#)(+)</sup>
	0.1794**
<i>HighEffort</i>	(0.0618)
<i>MedImage</i>	-0.421**
	(0.1493)
<i>HighImage</i>	-0.682***
	(0.1838)
<i>NightlyRate</i>	-0.2201
	(0.2407)
	0.0378
<i>MaxGuests</i>	(0.0792)
<i>MinStays</i>	-0.0110
	(0.0134)
<i>NumReviews</i>	0.0273***
	(0.0042)
<i>OverallRating</i>	0.01052
	(0.0201)
Observations	7546
Log pseudolikelihood	-2910.09

Fixed Effect	Property
Seasonality	Monthly

---

# Robust standard errors are presented in parentheses

+ regressed over samples with property received positive number of bookings

p<0.05 \*\* p<0.01 \*\*\* p<0.001

## 2.3 Model

The model-free evidences presented in section 2.3 indicate that image quality, affects both the present property demand (prior-consumption), and that the guests' (post-consumption) likelihood of writing reviews. Particularly, if the higher-quality images do not match the realized lower-quality properties and/or the host does not provide a good service that meet guests' expectation, unsatisfied guests are likely to choose not a write a review. In the long-run, this host may end up losing property demand and revenue as he/she is unable to grow the number of reviews, which is a key driver in generating bookings as consumers rely on the number of reviews to make decisions (particularly when the review ratings are seriously inflated)<sup>28</sup>. Knowing this, using high-quality image may not be the best interest of an Airbnb host, even when the images are available for free, if the host's property has low quality and/or the host is unable to deliver good service.

### 2.3.1 Model Overview and Timing of Events

One of our main objectives is to predict the dynamic choice of property images and investment in service for each AirBnB host, given his/her own 'type'. To do so, we build a structural model that incorporates hosts' ability in investing in service and the true quality of their properties. This is a dynamic game model, in which individual hosts, other lodging alternative providers including their peers, make monthly decisions about the quality of posted property images and the amount of effort for providing guest service. We assume that hosts are rational and forward-looking, with their objectives to maximize the total discounted utility flow summed over all periods forward.

Figure 14 illustrates the timing/sequence of the events for our model. At the beginning of each period (month)  $t$ , an individual host  $j$  observes her current state of 1) the current aggregated images quality of her property, categorized into low-, med-, and high- quality states, 3) the total number of guests reviews she accumulated till now, and 4) the states of her peers  $-j$ . Next, she makes a decision on whether or not to keep the listing active for this period. If she decides to 'snooze/de-active' the listing, then the model for the current period ends and she receives a realized value from choosing an outside option. The model will begin

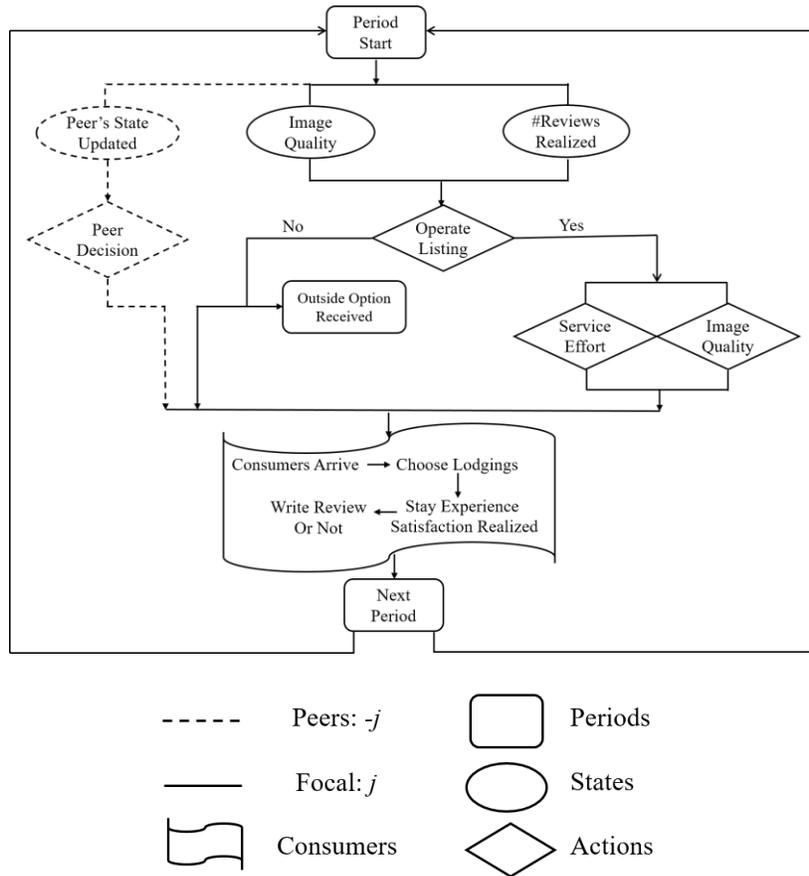
---

<sup>28</sup> Airbnb reported that United States homes were rated 4.8 out of 5 stars with 26,000,000+ reviews.

again for her at the beginning of period  $t+1$ , with her individual state remain the same and her peers' states updated according to peers' actions in period  $t$ . If she chooses to operate the listing in period  $t$ , then she pays a cost for the operation regardless of received property bookings in this period. Then she makes a decision on: 1) choosing the aggregate image quality from the three quality levels of low-, medium-, and high- for period  $t$ , and 2) choosing an effort between low- versus high- levels to invest in providing service in period  $t$ . If she updates the images to a different quality level, she pays a cost of preparing and posting images associated with that quality level. She incurs a cost associated with the level of invested service effort. After every individual host has made decisions, the consumers/guests will 1) observe the properties and their characteristics, including the image quality and number of reviews, 2) form an expectation on each property and choose one lodging alternative (guests are allowed to choose an outside option such as a hotel room), 3) receive, after having arrived in the property, a realized lodging experience by observing the property quality and being hosted with a certain quality of service, and 4) decide whether or not to leave a review, based on the post-consumption satisfaction. Next, each individual host receives a total amount of the revenue generated from renting out their property. Furthermore, for each reserved day, she pays a cost for hosting the guests associated to the amount of invested service effort. At the beginning of each period, a host chooses the action that maximizes her summed discounted profit ( $V$ ). Lastly, each individual property's number of reviews is updated, corresponding to the number of bookings received and their guests' post-consumption (review-writing) behavior. The model then moves to period  $t+1$  and the sequence of events is repeated monthly.

A host's per-period profit can be simply decomposing into revenue she makes from renting out the listing and costs of her actions. Hence, to construct individual host's objective function, we need to first estimate a property demand model from which a host can compute her property's market share and the corresponding revenue.

**Figure 14 Timing of Events for Each Month**



### 2.3.2 Property Market-Share Model and Hosts' Revenues

As a main component of hosts' objectives, the revenues of renting out their places comes through the population of consumers/travelers choosing lodging alternatives on the lodging supply market. Notably, there are three challenges arise for our study. First, the lodging market consists of a large number of differentiated products. Besides hotels and other lodging options, AirBnB properties themselves may be quite distinct from each other—in terms of property's type, size, location etc... Second, consumers are heterogeneous in their preferences on the lodging features. Third, our data on demand (i.e., property bookings) is at an aggregated market-level. That is, we only observe property bookings as aggregated responses from individual consumers' choices of lodgings, without knowing the trajectories of who booked particular properties.

To resolve above challenges, we an aggregate-demand model introduced in the seminal work of Berry, Levinsohn, and Pakes (Berry et al. 1995, hereafter BLP). The appealing BLP framework has been widely

applied in economics and marketing, as it uses readily-available aggregate level sales data, allowing for unobserved individual consumer heterogeneity and producing a more realistic product substitution pattern (Davis 2006, Houde 2006, Nevo 2001, and Sudhir 2001).

Suppose there are  $J$  AirBnB properties (i.e., products) on Manhattan (New York City) lodging market. In each period, in total  $I$  consumers choose at most 1 property from the  $J$  alternatives. Consumers are also allowed choose an outside option (denote as  $j=0$ , e.g., choosing a hotel or staying at friend's home). Each property is viewed as a set of property attributes  $X$ , on which consumers evaluate to make decisions on which property to choose. That is, the utility that consumer  $i$  choose alternative  $j$  in period  $t$  can be written as

$$u_{ijt} = \begin{cases} X_{jt}\theta^i + \eta_{jt} + \epsilon_{ijt} & \text{if } i \text{ chooses from the } J \text{ products} \\ \epsilon_{ijt} & \text{if an outside option is chosen} \end{cases}$$

where  $X_{jt}$  is a 1 by  $K$  product-attribute vector (as described in section 3.2) and  $\theta^i$  captures consumer  $i$ 's preferences over the  $K$  attributes. Specifically,  $X_{jt} = \{Period_t, Area_j, EntireHome_j, Apartment_j, Bedrooms_j, Bthrooms_j, Beds_j, MaxGuests_j, MinimumStay_j, NumReviews_j, MedImage_j, HighImage_j, DriveTime_j, WalkScore_j, NightlyRate_j\}$ , where  $Period_t$  is period (month) fixed effects included to capture seasonality and  $Area_j$  a area (neighborhood) fixed effects to capture the geographic-related popularity.  $\eta_{jt}$  is a common aggregate demand shock across consumers. The idiosyncratic shock  $\epsilon_{ijt}$  follows an *i.i.d.* distribution  $F_\epsilon(\epsilon)$ .

To incorporate possible heterogeneity in the consumers' preferences, we further model individual-specific preference  $\theta^i$  as an independent draw from the preference distribution  $F_\theta(\theta, \omega)$  characterized by parameter  $\omega$ . Each consumer chooses the alternative that gives him/her the highest utility, and the aggregated (i.e., integration over the population) response of  $I$  consumers' choices gives us the market share for alternative  $j$  in period

$$ms_{jt} = \int_{\theta^i, \epsilon_{it} | u_{ijt} \geq u_{ilt}, \forall l \neq j} dF_\theta(\theta, \omega) dF_\epsilon(\epsilon) \quad (6)$$

Following the convention in related literature (Berry et al. 1995, Dube et al. 2012), we model consumer preferences  $F_\theta(\theta, \omega)$  follow a normal distribution with  $\omega = (\bar{\theta}, \Sigma)$  that characterizes the means and the covariance matrix of the  $K$ -dimension parameter vector. Further, idiosyncratic shock  $\epsilon_{ijt}$  is assumed be drawn from type-I Extreme Value distribution. Then Equation (6) can be written as

$$ms_{jt} = \int ms_{ijt} \phi(\theta^i | \bar{\theta}, \Sigma) d\theta^i = \int \frac{\exp(X_{jt}\theta^i + \eta_{jt})}{1 + \sum_{l=1}^J \exp(X_{lt}\theta^i + \eta_{lt})} \phi(\theta^i | \bar{\theta}, \Sigma) d\theta^i \quad (7)$$

where  $ms_{ijt} = \frac{\exp(X_{jt}\theta^i + \eta_{jt})}{1 + \sum_{l=1}^J \exp(X_{lt}\theta^i + \eta_{lt})}$  indicates the probability that consumer  $i$  chooses property  $j$  from the choice set  $j=0, 1, \dots, J$  ( $j=0$  indicates the outside option, with its coefficient normalized to one for identification's purpose) in period  $t$ . Here individual-specific preference  $\theta^i = \bar{\theta} + v^i$ , with  $v^i$  an independent draw from  $N(\mathbf{0}, \Sigma)$ . Hence  $\bar{\theta}$  reflects the average preference in the population on the  $K$  property attributes and  $v^i$  quantifies individual's deviation from mean preference  $\bar{\theta}$ . The covariance matrix of the normal distribution,  $\Sigma$ , thus captures the extent of consumer heterogeneity and correlations, if any, between the preferences.

Finally, with our market share data on  $J+1$  alternatives spanning  $T$  periods, we obtain  $J*T$  demand equations. For each  $j=1, 2, \dots, J$  and  $t=1, 2, \dots, T$ , we write a market share as specified in Equation (8):

$$\begin{aligned} ms_{jt} &= \int \frac{\exp(X_{jt}\bar{\theta} + \eta_{jt} + X_{jt}v^i)}{1 + \sum_{l=1}^J \exp(X_{lt}\bar{\theta} + \eta_{lt} + X_{lt}v^i)} \phi(v^i | \mathbf{0}, \Sigma) dv \\ &= \int \frac{\exp(\mu_{jt} + X_{jt}v^i)}{1 + \sum_{l=1}^J \exp(\mu_{lt} + X_{lt}v^i)} \phi(v^i | \mathbf{0}, \Sigma) dv \end{aligned} \quad (8)$$

where  $\mu_{jt} = X_{jt}\bar{\theta} + \eta_{jt}$  is the “mean utility” for alternative  $j$  common across consumers in period  $t$ .

As can be seen from Equation (8), the set of market shares,  $ms_t = (ms_{1t}, ms_{2t}, \dots, ms_{Jt})'$ , given the observed covariates  $X_t = (X_{1t}, X_{2t}, \dots, X_{Jt})'$  and the preference distribution  $N(\bar{\theta}, \Sigma)$ , is a function of the demand shocks  $\eta_t = (\eta_{1t}, \eta_{2t}, \dots, \eta_{Jt})'$ . As we will discuss in section 4, following Jiang et al. (2009), we specify distribution for  $\eta_{jt}$  then evaluate the likelihood function for estimating the demand equations. Though the Bayes estimators, compared to GMM estimators, require an additional distributional assumption (i.e., the assumption on  $\eta_t$ ) to derive the likelihood, they have a couple of prominent advantages: 1) The MCMC (Monte Carlo Markov Chain) method implemented in deriving Bayes estimators provide a natural and unified framework for conducting inference (from the stationary posterior distribution) on the functions of model parameters such as price elasticity and markups. However, in the GMM framework, one would have to implement extra computations outside the model parameter estimation procedures, e.g., through bootstrap methods (Nevo 2001) to obtain asymptotic standard errors of these nonlinear functions of the model parameters. 2) Jiang et al. (2009) conduct simulation experiments and show that GMM estimators' asymptotic standard errors understate the true variance in the (simulated) samples and that Bayes estimators have lower MSE (Mean Squared Error) than GMM estimators. 3) The distributional assumption on the aggregate demand shocks give flexibility when conducting policy simulations or computing price elasticities. However, in a GMM framework, one would either impose a value of zero for demand shocks or use the realized demand shocks when

computing price elasticities. 4) Jiang et al. (2009) show that Bayes estimators are quite robust to possible misspecifications on the i.i.d. normal distribution assumption of  $\eta_t$  (a departure from normality, independence, homoscedasticity etc.).

### 2.3.2.1 Addressing Endogeneity

A couple of variables in the demand equation suffer from potential endogeneity issue. Below we first describe how the endogeneity concern is addressed in a Bayesian framework and then introduce the set of Instrument Variables (IV) that we use for each of endogenous variables.

In the presence of endogenous variables in the BLP model, a conventional approach is GMM method that exploit the orthogonality conditions (Berry et al. 1995, Nevo 2000). Specifically, suppose the observed covariates  $X_{jt}$  can be decomposed as  $X_{jt} = \{W_{jt}, P_{jt}\}$ , where  $P_{jt}$  indicates the endogenous variable that may be correlated with demand shocks ( $\eta_{jt}$ ) that are unobserved to the researchers and  $W_{jt}$  are all other exogenous variables. GMM method requires to find a set of instrumental variables for  $P_{jt}$ — $Z_{jt}$  that is orthogonal to  $\eta_{jt}$  to construct moment conditions.

In our Bayesian-BLP framework, following Jiang et al. (2009), we use a Bayesian approach to instrumental variable (see Rossi et al. (2005)) to address the endogeneity issue. Similar to a ‘first-stage’ in a classic 2SLS estimation, the following linear regression relates  $P$  to  $Z$ :

$$P_{jt} = Z_{jt}\delta + \xi_{jt}$$

where  $\xi_{jt}$  is a stochastic shock. The endogeneity of variable  $P_{jt}$  arises as  $\xi_{jt}$  is correlated with the common demand shock  $\eta_{jt}$ . We specify the distributional assumption on  $\xi$  and  $\eta$ :

$$\begin{pmatrix} \xi_{jt} \\ \eta_{jt} \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \right)$$

Recall that  $\mu_{jt} = X_{jt}\bar{\theta} + \eta_{jt} = [W_{jt}, P_{jt}]\bar{\theta} + \eta_{jt}$  and that market shares  $ms_t$  are functions of  $\mu_{jt}$  and  $\eta_{jt}$  (see Equation (8)), the endogeneity properly captured through writing the joint distribution of  $(P_t, ms_t)$  as a function of  $\begin{pmatrix} \xi_{jt} \\ \eta_{jt} \end{pmatrix}$ .

As can be seen, the key step here is to find a set of instrumental variables (IV) for each potential endogenous variable. In our demand model, there are three variables that may be endogenous—property price, image quality, and number of reviews. Below we discuss the endogeneity concern and describe the set of IV used for each variable.

### *Addressing Endogeneity in Property Price*

As in the previous literature, price in our study is endogenous, i.e., they may be correlated with demand shocks that are unobserved to the researchers. Following the extent literature using aggregated market share model, we first include so called “BLP instruments” in the set of IV for property price (Berry et al. 1995, Nevo 2001). These instruments use own product’s characteristics and the sum of competitor products’ characteristics. The logic for the former is that, the own characteristics are determined before the prices are set, hence they are unlikely to be correlated with time-variant unobserved demand shocks. The logic for the latter is that competitors’ characteristics are unlikely to be correlated with the unobserved shocks in a product’s demand. However, the proximity in product characteristics space between a product and its competitors influence the competition, and as a result, influence the product markup and the price. In addition, we use cost-based instruments—the factors that enter a product’s cost side but not demand side, i.e., product-specific cost shifters (BLP 1999, Dube 2012). For this study, we use local (zip code level) residential utility fee obtained and rental information (collected from Zillow). The logic is that these factors serve as an indirect measure of cost and enter price through affecting the cost on the supply side. However, they’re unlikely to be correlated with the unobserved factors on the demand side.

### *Addressing Endogeneity in Choice of Image Quality*

Hosts’ choices of image quality are endogenous in the sense that hosts’ incentive to use high-quality images to attract more bookings when the overall demand level varies. Moreover, a host’s decision on image quality is affected by the quality of her property. As introduced above, we use “BLP instruments”—the sum of quality of competitors’ property images. In addition, we include the following variables as an instrument for the choice of image quality. Property location—some locations or neighborhoods may have higher supply of local photographers, which make it easier for the host to hire a professional photographers on her own; Listing type—it may be easier to manage a professional photography shooting for a property that is listed as entire home, compared to a shared place (especially if the host has a roommate or sublease living in the same property); The number of years of experience as an Airbnb host—a host with more years of hosting experiences may be more likely to hire a professional photographers on her own, because she has more experience or knowledge, compared to a host with less experience, in knowing how to manage the whole process.

### *Addressing Endogeneity in Number of Reviews*

The number of reviews on the demand side is endogenous because they may be correlated with unobserved property characteristics (as shown in section 2.3, the guests’ likelihood of writing a review is dependent on the property’s quality). We use two sets of instruments to address the endogeneity concern.

The instruments are uncorrelated with the focal property's unobserved quality, however can influence, through competition, the evolution of number of reviews by affecting the number of bookings a property can receive. The first set of instruments are the 'BLP instruments'—the sum of number of reviews of other Airbnb properties in the same neighborhood, as the number of reviews of local properties are unlikely to be correlated with one's unobserved property quality but is correlate with one's property quality via competition. The second set of captures coemption through the supplied lodging alternatives—the number hotel rooms in the same neighborhood. Similarly, the number of supplied hotel rooms influences a property's number of review via affecting one's received property bookings, however is uncorrelated with one's unobserved quality.

### 2.3.2.2 Hosts' Revenues

With the market share model specified in section 3.2.1., a host  $j$  can approximate the expected market share and hence the expected revenue in any given period  $t$  assuming he knows  $\{X_{jt}\}_{j=1}^J$  or expected  $\{X_{jt}\}_{j=1}^J$ . Being able to make money is a major motive that people choose to host on AirBnB. In Equation (9), we specify the indirect- payoff from renting the property  $j$  in period  $t$  to capture the monetary motive for a host:

$$\begin{aligned}
 Revenue_{jt} &= ReservationDays_{jt} \cdot NightlyRate_{jt} & (9) \\
 &= (ms_{jt} \cdot MarketSize) \cdot NightlyRate_{jt} \\
 &= \left( \int \frac{\exp(\mu_{jt} + X_{jt}v^i)}{1 + \sum_{l=1}^J \exp(\mu_{lt} + X_{lt}v^i)} \phi(v^i | \mathbf{0}, \Sigma) dv \right) \cdot MarketSize \\
 &\quad \cdot NightlyRate_{jt}
 \end{aligned}$$

where  $ms_{jt}$  indicates the property  $j$ 's market share in period  $t$ , as defined in Equation (8), given the property characteristics for the set of alternatives,  $\{X_{lt}\}_{l=1}^J$ , the mean utility for the set of alternatives,  $\{\mu_{lt}\}_{l=1}^J$ , and the heterogeneity in the guests' (consumers') preferences over these characteristics,  $v^i \sim N(\mathbf{0}, \Sigma)$ .

### 2.3.3 Individual Host's Cost

#### Individual-Specific Cost: Investing Effort in Guest Service

Before a guest booked a property, she does not have much interaction with the host. After she has arrived, for each day of her stay, she receives guest service provided by the host. Such service may include communicating with the guest when then check in/leave to guarantee a smooth transition, keeping the place clean and air fresh, leaving message/travel guide in the room, and answering guests' questions (e.g., regarding dining options in that city or how to use an appliance in the room). A survey on AirBnB

suggests that guests care about the quality of received service<sup>29</sup>. In fact, unfriendly/irresponsive hosts and unpleasant conditions are major sources that could lead to serious unsatisfactory in the lodging experiences<sup>30</sup>.

However, providing high-quality service is costly to the hosts. Particularly, some hosts, because of their occupations and/or the location of their property, may find it difficult to frequently check their messages and promptly respond to the guests. Thus, we assume that investing more effort for hosting guests cost more to the same host and that investing the same amount of effort may have different costs for different hosts (i.e., hosts have different ability in investing service effort). For property  $j$  in period  $t$ , the cost of providing service,  $ServiceCost_{jt}$  is specified as below

$$ServiceCost_{jt} = \begin{cases} \lambda_j^{low} & \text{if investe low effort} \\ \lambda_j^{high} & \text{if investe high effort} \end{cases}$$

where  $\lambda_j^{low}$  and  $\lambda_j^{high}$  indicates the cost for  $j$  to invest low effort and high effort in service, respectively. For identification's purpose, we further normalize  $\lambda_j^{low}$  to 0 and instead estimate the relative marginal effort cost  $\lambda_j^{effort} = \lambda_j^{high} - \lambda_j^{low}$  for  $j$ . We allow for heterogeneity in individual's marginal effort cost assume  $\lambda_j$  to be i.i.d. draws from a normal distribution with mean  $\bar{\lambda}$  and variance  $\sigma_\lambda^2$ — $\lambda_j^{effort} \sim N(\bar{\lambda}, \sigma_\lambda^2)$   $\bar{\lambda}$  then captures the cost of investing high effort in service to an average host and  $\sigma_\lambda^2$  reflects the variation in the host population.

Finally, Equation (10) summarize the service cost that host  $j$  incurs in period  $t$ :

$$ServiceCost_{jt} = \begin{cases} 0 & \text{if HighEffort}_{jt} = 0 \\ \lambda_j^{effort} & \text{if HighEffort}_{jt} = 1 \end{cases} \quad (10)$$

### **Individual-Specific Cost: Operating Listing (i.e., Opportunity Cost)**

At the beginning of each period, a host can choose to make her listing 'active', i.e., to operate the listing in current period, or to 'snooze' her listing, i.e., to temporally exit from Airbnb (no operation on Airbnb) for that period. A host incurs a listing-operation cost for managing an 'active' listing on Airbnb. Listing-operation cost include the activities that a host take to remain the listing 'active', such as managing the property page, keeping the property availability calendar updated, and possible social cost as the neighbors may be unhappy with an AirBnB listing in the neighborhood/building<sup>31</sup>, and potential opportunity cost due to listing their property on AirBnB.

<sup>29</sup> <https://www.asherfergusson.com/airbnb/>.

<sup>30</sup> <https://www.asherfergusson.com/AirBnB/>.

<sup>31</sup> <https://www.nbclosangeles.com/news/local/I-Team-Investigation-Short-Term-Rentals-Property-AirBnB-415128373.html>.

We assume an individual-specific operation cost<sup>32</sup> and denote property  $j$ 's operation cost with  $\lambda_{operate}$ . Further, for identification's purpose we normalize the cost of keeping an inactive listing in a month to zero (i.e., normalizing the operating cost and outside option value to zero). A host hence incurs zero cost and receives zero revenue from Airbnb in that period if she snoozes the listing.

Lastly, Equation (11) specifies the structure of the operation costs. Note that, unlike individual-specific effort cost, which the host incurs for every booked day, the operation cost is incurred as a fixed cost at the beginning of each period, regardless of the realized booking in that period.

$$OperationCost_{jt} = \begin{cases} 0 & \text{if } Active_{jt} = 0 \\ \lambda_j^{operate} & \text{if } Active_{jt} = 1 \end{cases} \quad (11)$$

### **Common Cost: Photography Cost**

In addition to the heterogeneous costs of investing service effort and operating listings, there is a key component in the cost that a host may incur—cost of posting property images. Image-posting cost include things such as organize/clean the place, taking photos, do post-processing, and then upload photos. We assume that the photography costs are common across the hosts for two reasons—1) The cost of hiring a professional photographer in a specific market (e.g., Manhattan) are likely to be relatively the same across subareas in that market, and 2) our sample does not observe sufficient variation, at the individual property level, in updating their aggregate-quality level of the property images<sup>33</sup>.

Since we categorize property images' aggregate quality into a low-, med-, and high- 3 level, the cost of posting images is likely to differ across the levels. For example, it is easy for a host to take amateur images (with their smartphone phone camera). But to take a med-level image, someone may need to organize the place, ask her friend who can help or spend a whole day of taking lots of photos (trying different scene organization, camera angle, illumination etc.) to pick some good from, and then edit/post process the photos. Taking high-level photos is likely to be the costliest, as one may have to clean and prepare the place, then pay a professional photographer to take photos for her. Thus, we assume that the cost of posting property images for a host is:

---

<sup>32</sup> In reality, these costs are likely to heterogeneous across the hosts/properties. If a neighborhood has very strict policy on home-sharing platform or the neighbors are more again home-sharing, then hosting a property in this neighborhood is likely to be more costly than other hosts. Additionally, the locations of properties introduce variation in the opportunity cost, as local rental (lodging) popularity leads to different outside option values for the properties across geographic areas.

<sup>33</sup> In reality, these costs may vary across hosts. For example, if a host herself is a professional photographer, then we expect that the cost of taking a high-quality photo for her to be lower than for hosts who are amateurs. However, we think the variance in hosts' ability of shooting professional photos are likely to be small as most of them are amateurs. By 2011, very few of the photos on Airbnb were professional, which motivated the company to launch its professional photography program was launched in 2011 (see the report in 2012 from Joe Zadeh, Product Lead at Airbnb).

$$ImageCost_{jt} = \tilde{\lambda}_{high}I(HighImage) + \tilde{\lambda}_{med}I(MedImage) + \tilde{\lambda}_{low}I(LowImage)$$

where  $I(\cdot)$  is an indicator function and  $\tilde{\lambda}_{high}$ ,  $\tilde{\lambda}_{med}$ , and  $\tilde{\lambda}_{low}$  refers to the cost of posting high-, med-, and low- quality level images, respectively. We further normalize the  $\tilde{\lambda}_{low}$  to 0 and instead identify  $\lambda_{high} = \tilde{\lambda}_{high} - \tilde{\lambda}_{low}$  and  $\lambda_{med} = \tilde{\lambda}_{med} - \tilde{\lambda}_{low}$ . We do so for identification's purpose, as we do not observe that hosts post no image. Moreover, we assume that a cost is incurred only when the quality level of images is updated. For example, if a host had high-level image in  $t-1$  period and decides to remain those images for period  $t$ , then there is no cost of 'posting' high-quality images.

Lastly, Equation (12) specifies the structure of the photography costs. Note that the photography cost is incurred as a fixed cost at the beginning of each period, depending on a property's current image quality and the image quality decision, before the property bookings are realized in that period. That is, a host incurs a photography cost if and only if he updates the image quality to a different quality level that is not low-level.

$$ImageCost_{jt} = \begin{cases} \lambda^{MedImg} & \text{if } MedImage_{jt} = 1 \text{ and } MedImage_{jt-1} = 0 \\ \lambda^{HighImg} & \text{if } HighImage_{jt} = 1 \text{ and } HighImage_{jt-1} = 0 \end{cases} \quad (12)$$

### 2.3.4 Individual Host's Per-period Payoff

As discussed above, an individual host's per-period payoff can be decomposed into 1) revenue making from renting out the property, 2) effort cost of investing on service, 3) costs of updating property images, and 4) cost of operating the listing. An individual host  $j$  can choose from 7 possible combinations of actions (3 levels of images\*2 levels of effort + operating/snoozing action), denoted by a finite set  $A_j = \{1, 2, \dots, 7\}$ . In every period  $t$ , every host  $j$  makes a choice  $a_{jt} \in A_j$ . We further let  $a_t = (a_{1t}, a_{2t}, \dots, a_{jt})$  denotes the set of actions of all individuals in period  $t$ .

The payoff of taking action  $k$  for host  $j$  in period  $t$  is specified in Equation (13):

$$\begin{aligned} & \Pi_{jat}(a_{jt} = k) & (13) \\ & = \begin{cases} Revenue_{jkt} + ServiceCost_{jkt} + ImageCost_{jkt} + OperationCost_{jkt} + \varepsilon_{jkt} & \text{if } Active_{jt} \\ \varepsilon_{jkt} & \text{if } \neg Active_{jt} \end{cases} \end{aligned}$$

where  $\varepsilon_{jkt}$  refers to action-specific random shocks for individuals and is assumed to follow a Type-I extreme value distribution with  $\varepsilon_{jkt} \sim EV(\mu_\varepsilon, \sigma_\varepsilon)$ . Prior to taking an action, the host can only form an expectation on the payoff received in current period as the revenue is realized only at the end of the period. Let  $\tilde{\Pi}_{jkt} = E\left(\Pi_{jkt} \mid a_{jkt}, \left\{ \lambda_j^{effort}, \lambda_j^{operate}, X_{jt}, a_{jt} \right\}_{j=1}^J, \lambda^{HighImg}, \lambda^{MedImg}\right)$  denote the expected payoff, conditional on the set of individual-specific parameters  $\alpha_j, \lambda_j$  for all hosts, the market-share relevant

covariates  $X_{jt}$  for all properties, and the actions that her peers will take  $a_{jt}$  in current period. For identification's purpose, we normalized the mean payoff of 'snoozing the listing' to zero, then Equation (14) specifies the conditional expected payoff from taking action  $a_{jt}$  for  $j$  in period  $t$ . Note for simplicity, we use  $-j$  to denote the set of individuals excluding  $j$ .

$$\begin{aligned} & \tilde{\Pi}_{jkt}(a_{jt} | \{\lambda_j^{effort}, \lambda_j^{operate}, \lambda_{-j}^{effort}, \lambda_{-j}^{operate}, \lambda^{HighImg}, \lambda^{MedImg}, X_{jt}, X_{-jt}, a_{-jt}\}) \\ &= \begin{cases} ReserveDays_{jt}(a_{jt}, X_{jt}, X_{-jt}, a_{-jt}) \cdot (NightlyRate_{jt} + \lambda_j^{effort} I\{a_{jt}(HighEffort) = 1\}) \\ \quad + \lambda^{MedImg} I\{a_{jt}(MedImage) = 1\} \cdot I\{a_{jt-1}(MedImage) = 0\} \\ \quad + \lambda^{HighImg} I\{a_{jt}(HighImage) = 1\} \cdot I\{a_{jt-1}(HighImage) = 0\} \\ \quad + \lambda_j^{operate} & \text{if } Active_{jt} = 1 \\ 0 & \text{if } Active_{jt} = 0 \text{ (snooze listing)} \end{cases} \end{aligned} \quad (14)$$

where  $ReserveDays_{jt}(a_{jt}, X_{jt}, X_{-jt}, a_{-jt})$  denotes the number of booked days for  $j$  in period  $t$ , supposing she takes action  $a_{jt}$  with market-share relevant covariate  $X_{jt}$  and her peers take action  $a_{-jt}$  with covariate  $X_{-jt}$ . Recall that a property's market share is a function of her own and her peers' state (see Equation 8 and 9).  $I\{\cdot\}$  is an indicator function and  $a_{jt}(\cdot)$  refers to a specific activity in this action. For example,  $I\{a_{jt}(MedImage) = 1\}$  is 1 if the action of  $a_{jt}$  will result in a med-level property image for property  $j$  in period  $t$ , and is 0 if otherwise.

### 2.3.5 State Variables

This section defines state variables that affect an individual host's payoff over time and discusses the dynamics in the transition of individual states driven by hosts' actions.

The set of state variables for individual  $j$  at time  $t$  is  $s_{jt} = (MedImage_{jt-1}, HighImage_{jt-1}, NumReviews_{jt-1})$ . Similarly, let  $s_{-jt} = (MedImage_{-jt-1}, HighImage_{-jt-1}, NumReviews_{-jt-1})$  denote the set of state variables for  $j$ 's peers at time  $t$ . Note we write the  $t-1$  in the subscript for  $MedImage$ ,  $HighImage$ , and  $NumReviews$  to emphasize that the state of image and reviews in current period comes from the action and the outcome in the previous period. Lastly, the combination of the set of states for all individuals constitute state. Then, for  $j$  at time  $t$ , her strategy profile depends on  $s_t = (s_{jt}, s_{-jt})$ , as the not just her own state, but also the states of her peers affect her decisions and state transition.

## State Transitions

For each individual  $j$ , the evolution of her own state at time  $t$ ,  $s_{jt}$ , depends on  $s_{jt}$  itself and her choice  $a_{jt}$ . The individual type  $(\lambda_j^{effort}, \lambda_j^{operate})$  does not evolve over time. Below, we formalize the evolution of the dynamic states—namely images and number of reviews.

**Images.** The transition of images is governed by individual's action (choice of images) in every period. For example, if individual  $j$  chose med-level images in period  $t$ , i.e.,  $a_{jt}(MedImage_{jt}) = 1$ , then the outcome of this action leads to the state in next period. As a result, in period  $t+1$ , med-level images will be the individual's state, i.e.,  $s_{jt+1}(MedImage_{jt+1}) = 1$ . If a host chooses to snooze the listing for current period, then next period the image state remains the same. That is:

$$s_{jt+1}(Image_{jt+1}) = \begin{cases} a_{jt}(Image_{jt}) & \text{if } a_{jt}(Active_{jt}) = 1 \\ s_{jt}(Image_{jt}) & \text{if } a_{jt}(Active_{jt}) = 0 \end{cases}$$

Note that the evolution of image state of  $j$ 's peers depends on peers' choices of images in current period, regardless of  $j$ 's choice.

**Number of Reviews.** The evolution of number of reviews is straightforward. For each individual  $j$ , we have following transition rule for state *NumReviews*:

$$NumReview_{jt+1} | NumReview_{jt} = \begin{cases} NumReview_{jt} + NewReview_{jt} & \text{if } a_{jt}(Active_{jt}) = 1 \\ NumReview_{jt} & \text{if } a_{jt}(Active_{jt}) = 0 \end{cases}$$

where  $NumReview_{jt}$  denotes the number of reviews that  $j$  has accumulated till the beginning of period  $t$  and  $NewReview_{jt}$  indicates the number of new (added) review in period  $t$ . If a host chooses to snooze the listing for period  $t$ , then review state also 'snoozes' for time  $t$ , as there will be no booking at all.

In above equation,  $NewReview_{jt}$  is generated through guests in period  $t$  who, conditional on the transactions and their stays, are willing to leave a review. Hence,  $NewReview_{jt}$  is a function of  $\{X_{jt}, X_{-jt}, a_{jt}(MedImage), a_{-jt}(MedImage), a_{jt}(HighImage), a_{-jt}(HighImage)\}$ , which affects the current market share for property  $j$ ,  $s_{jt}$ . It is also a function of  $a_{jt}(MedImage), a_{jt}(HighImage), a_{jt}(HighEffort)$  which, combined with property  $j$ 's quality, affects the likelihood that guests, conditional on their stays, leave a review.

A host, before the property bookings and stay experiences are realized, he can only form an expectation on the number of generated new review, given the states and actions of her and her peers. We assume that hosts have learned the relationship between consumers' likelihood of writing reviews and the realized gap

(i.e., departure of property quality from image quality) and the quality of provided guest services. Specifically, we use the empirical relationship (see section 2.3.2 and Equation (5)) to compute the review-writing likelihood (conditional on one's property quality, improving image quality would reduce guest's post-consumption likelihood of writing reviews, see Table 15). To do so, we empirically identify the relationships from our data, i.e., the Law of Motion as shown in Equation (5):

$$\begin{aligned} WriteProb_{jt} = & \alpha_j + \alpha_{high}HighImage_{jt} + \alpha_{med}MedImage_{jt} + \alpha_{effort}Effort_{jt} \\ & + \alpha_{price}NightlyRate_{jt} + \alpha_4Control_{jt} + Period_t \end{aligned}$$

where  $\alpha_j$  is property fixed effect term, which is also a function of property quality. Using the Law-of-Motion as reported in Table 15, a host can approximate the likelihood that her guests will write reviews after their stays, given her  $\alpha_j$  as well as her choice of image quality and service effort.

Hence, the choice of images not only affect one's current market share, also impact the transition of reviews, as guests' post-consumption are affected by the 'gap' that arise when images create a high expectation. The choice of effort, though unobserved to consumers ex-ante and hence do not affect current market share, will actually impact one's long-term utilities through its control on the review evolution. At last, Equation (15) summarizes the review transition:

$$E(NumReview_{jt+1} | NumReview_{jt}, a_{jt}, a_{-jt}, s_{jt}, s_{-jt}, X_{jt}, X_{-jt}) = \quad (15)$$

$$\begin{aligned} & NumReview_{jt} + NumReservations_{jt} \cdot WriteProb_{jt} \\ & = NumReview_{jt} + \left( \frac{ResearvationDays_{jt}}{MinSatys_{jt}} \right) \cdot WriteProb_{jt} \\ & = NumReview_{jt} + (MarketShare_{jt} \cdot \frac{Marketsize}{MinSatys_{jt}}) \cdot WriteProb_{jt} \\ & = NumReview_{jt} + s_{jt}(X_{jt}, X_{-jt}, \mu_{jt}, \mu_{-jt}, s_j, s_{-jt}, \Sigma, \{v^i\}_{i=1}^I) \\ & \quad \cdot MarketSize / MinSatys_{jt} \\ & \quad \cdot WriteProb_{jt}(a_{jt}(Image_{jt}), a_{jt}(HighEffort_{jt}), \alpha_j) \end{aligned}$$

where  $\alpha_j$  is property fixed effect from Equation (5).

In summary, the whole state space for any individual  $j$  is at time  $t, s_{jt} = \{\alpha_j, MedImage_{jt}, HighImage_{jt}, NumReview_{jt}, MedImage_{-jt}, HighImage_{-jt}, NumReview_{-jt}\}$  all observed.

### 2.3.6 Individual's Optimization Problem

We model an individual host's choice of images and service effort as a dynamic optimization problem. On an infinite-time horizon, each individual  $j$  chooses an infinite sequence of actions  $a_{jt} = \{Image_{jt}, Effort_{jt}, Active_{jt}\}_{t=1}^{\infty}$  to maximize the sum of expected life-time payoff:

$$\max_{\{Image_{jt}, Effort_{jt}, Active_{jt}\}_{t=1}^{\infty}} E_{\{s'_{jt}, s'_{-jt}\}} \left\{ \sum_{t=0}^{\infty} \beta^t \cdot (\tilde{\Pi}_{jt}(Image_{jt}, Effort_{jt}, Active_{jt} | s_{jt}, s_{-jt})) \right\}$$

where  $s'_{jt}$ ,  $s'_{-jt}$  denotes the transitioned individual state in the next period for  $j$  and for her peers, respectively.  $\tilde{\Pi}_{jt}$  is the expected per-period payoff (expectation over the idiosyncratic payoff shocks  $\varepsilon_{jkt}$ , see Equation 13):

$$\begin{aligned} \tilde{\Pi}_{jt} = & (Active_{jt} == 0) * 0 + (Active_{jt} == 1) * \left\{ \int \frac{\exp(\mu_{jt} + X_{jt}v^i)}{1 + \sum_{l=1}^J \exp(\mu_{lt} + X_{lt}v^i)} \phi(v^i | \mathbf{0}, \Sigma) dv \right\} \\ & \cdot \frac{MarketSize}{MinSatys_{jt}} \cdot NightlyRate_{jt} + \lambda_j^{effort} \cdot (Effort_{jt} == HighEffort) + \lambda_j^{MedImg} \\ & \cdot (Image_{jt} == MedImage) \cdot (s_{jt}(MedImage) = \\ & = 0) + \lambda_j^{HighImg} \cdot (Image_{jt} == HighImage) \cdot (s_{jt}(HighImage) == 0) + \lambda_j^{operate} \end{aligned}$$

The specification of individual's per-period payoff and the state transition rule reveal the strategic interactions across peers—an individual's per-period payoff and decision is a function of her peers' states. That is, an individual need to approximate her peers' action in every period, given their states, as the peers' action affects one's current payoff and the evolution of the states. Moreover, this is a dynamic model in the sense that an individual's optimal decision change over time, as her and the peers' state evolve over time. Lastly, there are two interesting intertemporal trade-offs that worth emphasized. First, a host faces trade-off between posting high-level images to improve present property demand versus forging temporary revenue to grow the reviews and to improve future demand. Posting high-quality images improves the expected payoff for the consumers and thus improves a property's temporal market share. However, a high expectation may also induce a greater dissatisfaction as consumers will be happy about the 'negative gap' between the expectation and the realized property quality. As a result, their likelihood of writing review is reduced. In the long-run, this will hurt the host, as number of reviews plays a significant role in generating demand, particularly when one's peers are growing their review. Second, when choosing the amount of service effort, the host compares the effort cost for her (given her ability of investing effort) with the expected gain from an increased review (due to a better service to the guests) in the future. Providing a good service can improve the guests' post-consumption satisfaction and hence increase their likelihood of

writing reviews. As a result, the host can effectively grow the reviews and improve the future demand (and payoff). However, to do so, she must incur a present service cost.

### 2.3.7 A Dynamic Game and Equilibrium Concept

As mentioned in previous section, each individual host's decision is dependent on her own state and her peers' state. The hosts (properties), given the current states and each's private shock, make their decisions simultaneously and compete with each other in each period. A proper equilibrium concept for this dynamic game is Markov Perfect Equilibrium (Ericson and Pakes 1995, hereafter MPE).

An MPE is described as a profile of Markov strategies for each individual:  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_j)$ . For any individual  $j$ , her Markov strategy is operated as a function that maps the current state and  $j$ 's action-specific private shock into an action<sup>34</sup>:

$$\sigma_j: S \times \varepsilon_j \rightarrow A_j$$

where  $S$  denotes the state of all individuals and  $\varepsilon$  denotes the action-specific private shock that  $j$  received before making decision. Let  $V(s_{jt})$  denote the value function for individual  $j$  at time  $t$ :

$$V(s_{jt}) = \max_{\{Image_{jt}, Effort_{jt}, Active_{jt}\}_{t=1}^{\infty}} E_{\{s'_{jt}, s'_{-jt}, \varepsilon_t\}} \left\{ \sum_{t=0}^{\infty} \beta^t \cdot (\Pi_{jt}(Image_{jt}, Effort_{jt}, Active_{jt} | s_{jt}, s_{-jt})) \right\}$$

In an infinite-horizon optimization problem, the above equation can be solved through Bellman Equation (Bellman 1957):

$$\begin{aligned} V_j(s_j, s_{-j}; \sigma) &= E_{\varepsilon} [\Pi_j(\sigma(s_j, s_{-j}, \varepsilon), s_j, s_{-j}, \varepsilon_j)] \\ &+ \beta \int V_j(s'_j, s'_{-j}; \sigma) dP(s'_j, s'_{-j} | \sigma(s_j, s_{-j}), s_j, s_{-j}) | s_j, s_{-j} \end{aligned} \quad (16)$$

Note we dropped time subscript  $t$  as the Markov strategy does not dependent on time. In Equation (16),  $\beta$  is the common discount factor with  $0 \leq \beta < 1$ .  $s_j, s_{-j}$  denotes to the state of  $j$  and her peers  $-j$ , respectively.  $(s'_j, s'_{-j})$  denote to the states in the next period, conditional on the current state  $(s_j, s_{-j})$  and the actions that all individuals take, assuming their actions are governed by the Markov strategy.

In an MPE, every host  $j$ , given the Markov strategy profile  $\sigma$ , will choose the action that maximized the discounted life-time payoff. That is, for a profile  $\sigma$  to be an MPE, an individual would not choose an alternative strategy  $\sigma'_j$ , given that her peers follow  $\sigma_{-j}$ , that is,

$$V_j(s_j, s_{-j}; \sigma_j, \sigma_{-j}) \geq V_j(s_j, s_{-j}; \sigma'_j, \sigma_{-j})$$

---

<sup>34</sup> To be specific, we consider symmetric and anonymous pure strategy.

However, one challenge arises for computing MPE in our context. Specifically, because of the large number of individuals and the huge state space, solving for MPE is computationally intractable<sup>35</sup>. To resolve the issue of curse of dimensionality, Weintraub et al. (2008, 2010) proposed an approximation of MPE—Oblivious Equilibrium (OE). OE is developed for a market with large number of players (for example, see Huang et al. 2015 for an application of OE in the context of enterprise social media). The key notion is that in such a market, the simultaneous changes in each individual’s moves can be averaged out. As a result, the average industry state either remains stationary over time or can be tracked as a deterministic trajectory changing with a stationary (steady) pace (Weintraub et al. 2008, 2010). Thus, each individual does not need to track everyone’s state over time. Instead, it is sufficient for one to make a near-optimal decision by considering only her own state and the average industry state. As Farias et al. (2012) demonstrated, OE can approximate MPE very well, particularly if the market is not too concentrated and the number of individual players is not too few. Therefore, OE fits our setting and should give us a sufficiently good approximation to MPE for the following three reasons. First, we have a large number of individual Airbnb host, none of which is likely to dominate the market. Second, given the large number of hosts and the small role each host plays on the market, in the reality it is difficult for hosts to track all hosts’ states in every period. Third, in our data, we observe that the average state of images stay relatively constant over time, and the average number of reviews grow steadily over time (each period increases approximately 2 reviews). Thus, in our study we use OE to approximate MPE.

### 2.3.8 Unobserved Heterogeneity

In our data, individual hosts exhibit various responses/actions over time. For example, some hosts tend to invest more effort and provide a good service, while others frequently provide relative poor service. Some hosts choose high quality images to post on property page, while other tend to stay with relative poor images, even AirBnB was offering professional images for free. We hypothesize such different responses come through the heterogeneity in the consumers’ preference on revenues and in their ability of investing service effort. Following the stream of literature on hierarchical Bayesian framework (Ching et al. 2012, Rossi et al. 2005), we incorporate individual heterogeneity into our structural model by imposing a distributional assumption on the individual-specific parameters  $(\lambda_j^{effort}, \lambda_j^{operate})$ . Here  $\lambda_j^{effort}$  is individual  $j$ ’s marginal cost of investing high service effort and  $\lambda_j^{operate}$  is  $j$ ’s cost of operating an active Airbnb listing.

---

<sup>35</sup> Consider the number of individual’s state first. Each property at each period, has 3 possible states of image and 301 possible states of reviews (we truncate number of reviews at 300, as the observation with  $NumReviews > 300$  is less than 1%, hence reviews can vary between 0-300). Hence, for each individual, the number of her own state is  $3 \times 301 = 903$ . Then the whole state space (including one and her peers) has a dimension of  $(903)^{958}$ .

Specifically,  $(\lambda_j^{effort}, \lambda_j^{operate})$ . are assumed to be independent draws from a multivariate normal distribution (MVN) with mean  $\rho$  and covariance matrix  $\Sigma_\rho$ , i.e.,  $(\lambda_j^{effort}, \lambda_j^{operate}) \sim MVN(\bar{\lambda}, \Sigma_\lambda)$ .

The individual heterogeneities are time-persistent and unobserved to researchers. We assume that individuals know each other's  $(\lambda_j^{effort}, \lambda_j^{operate})$  and hence we estimate a complete information game with unobserved heterogeneity. Though  $(\lambda_j^{effort}, \lambda_j^{operate})$  is not explicitly specified by each host, individual hosts could learn or infer the distribution of  $(\lambda_{-j}^{effort}, \lambda_{-j}^{operate})$  through their own experience and relevant information of their peers such as host experiences, property locations, and property types etc.

## 2.4 Estimation Strategy and Identification

The model primitives (unknown parameters) include  $\{\{\theta_k^i\}_{k=1}^K, \{\mu_t, \eta_t, \xi_t\}_{t=1}^{12}, \delta, \Omega\}$  from the property market-share model (demand side) and  $\{\gamma_j^{effort}, \gamma_j^{operate}\}_{j=1}^{J=958}, \lambda_j^{MedImg}, \lambda_j^{HighImg}, \sigma_\varepsilon\}$  from the dynamic game model (supply side).

On the demand side,  $\theta_k^i$  refers to individual-level coefficient (preference) for simulated consumer  $i$  on the  $k^{\text{th}}$  characteristics, where  $k=1\dots 12$  indicates a series of dummy variables for the 12 months from January to December in a year (i.e., monthly fixed effects). For  $k=13\dots K$ , the corresponding characteristics for property  $j$  in period  $t$  are  $\{EntireHome_j, Apartment_j, Bedrooms_j, Bthrooms_j, Beds_j, MaxGuests_{jt}, MinimumStay_{jt}, NumReviews_{jt}, MedImage_{jt}, HighImage_{jt}, DriveTime_j, WalkScore_j, NightlyRate_{jt}\}$ .  $\mu_t = \{\mu_{jt}\}_{j=1}^J$  indicates a set of ‘mean utilities’ for properties  $j=1\dots J$  (mean utility is normalized to zero for outside option  $j=0$ ) in period  $t$ . Mean utility  $\mu_{jt}$  capture the overall preference in the population on property  $j$ .  $\eta_t = \{\eta_{jt}\}_{j=1}^J$  indicates a set of aggregate demand shocks to properties  $j=1\dots J$  in period  $t$ . Note that for each  $\eta_{jt}$  this demand shock is common across all consumers  $i=1\dots I$ .  $\xi_t = \{\xi_{jt}^{price}, \xi_{jt}^{image}, \xi_{jt}^{review}\}_{j=1}^J$  indicates a set of stochastic shock,  $\xi_{jt}$ , that is correlated with the both the aggregate demand common shock  $\eta_{jt}$  and the endogenous variables (i.e., property price, choice image quality, and number of reviews in our study).  $\delta = \{\delta^{price}, \delta^{image}, \delta^{review}\}$  relates endogenous price, image quality, and number of reviews, to their instruments, respectively, each through a linear regression.  $\Omega = \{\Omega^{price}, \Omega^{image}, \Omega^{review}\}$  specifies that endogeneity source—how demand shock  $\eta_t$  correlates with the stochastic shocks  $\xi_t$ —through a bivariate linear regression:  $\begin{pmatrix} \xi_{tj} \\ \eta_{jt} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}\right)$ .

Following Jiang et al. (2009), we employ a hierarchical Bayesian framework and impose distributional assumptions on the individual-level parameters. Specifically, we assume that  $\theta^i \sim MVN(\bar{\theta}, \Sigma)$ , where we allow that consumers preferences to be correlated, captured by the off-diagonal elements of  $K$  by  $K$  variance-covariance matrix  $\Sigma$ . Mean utility  $\mu_{jt} = X_{jt}\bar{\theta}$ . Equation (17) address the endogeneity concerns regarding property price, choice of image quality, and number of reviews, on the demand side, by describing the relationships among  $\xi_t$ ,  $\eta_t$ , and  $\delta$  with a set of bivariate normal distributions and a set of linear regressions where  $Z$  are the set of instruments for each endogenous variable:

$$\begin{cases} Price_{jt} = \delta^{price} Z^{price} \\ ImageQuality_{jt} = \delta^{image} Z^{image} \\ NumReview_{jt} = \delta^{review} Z^{review} \end{cases} \quad (17)$$

$$\begin{cases} \begin{pmatrix} \xi_{jt}^{price} \\ \eta_{jt} \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega^{price} = \begin{pmatrix} \Omega_{11}^{price} & \Omega_{12}^{price} \\ \Omega_{21}^{price} & \Omega_{22}^{price} \end{pmatrix} \right) \\ \begin{pmatrix} \xi_{jt}^{image} \\ \eta_{jt} \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega^{image} = \begin{pmatrix} \Omega_{11}^{image} & \Omega_{12}^{image} \\ \Omega_{21}^{image} & \Omega_{22}^{image} \end{pmatrix} \right) \\ \begin{pmatrix} \xi_{jt}^{review} \\ \eta_{jt} \end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega^{review} = \begin{pmatrix} \Omega_{11}^{review} & \Omega_{12}^{review} \\ \Omega_{21}^{review} & \Omega_{22}^{review} \end{pmatrix} \right) \end{cases}$$

Hence, the estimated parameters on the property market-share model are:  $\{\bar{\theta}, \Sigma, \Omega, \delta\}$ .

On the supply side,  $\lambda_j^{effort}$  indicates that cost of investing a high effort, relative to a low effort, in providing service to the guests for host  $j$ .  $\lambda_j^{operate}$  captures the cost, of operating an active AirBnB listing for host  $j$ . The operating cost includes costs related to listing page managing, social cost, and the potential oppournitiy cost (e.g., a property, if not listed on AirBnB, could be rented out via another home-sharing platform or long-term rental market). We assume individual-specific coefficients  $(\lambda_j^{effort}, \lambda_j^{operate})$  to be independent draws from a multivariate normal distribution i.e.,  $(\lambda_j^{effort}, \lambda_j^{operate}) \sim MVN(\bar{\lambda}, \Sigma_\lambda) = \begin{pmatrix} \Sigma_\lambda^{11} & \Sigma_\lambda^{12} \\ \Sigma_\lambda^{21} & \Sigma_\lambda^{22} \end{pmatrix}$ .  $\gamma_{med}$  and  $\gamma_{high}$  refers to the cost, common across hosts, of using (including preparing, making, and posting) medium-level and high-level quality of images, respectively. Lastly, we estimate the standard deviation in the distribution of the action-specific idiosyncratic shocks, where we assume  $\varepsilon_{jkt} \sim EV(\mu_\varepsilon, \sigma_\varepsilon)$ . Hence, the estimated parameters on the dynamic supply model are:  $\{\bar{\lambda}, \Sigma_\lambda, \lambda^{MedImg}, \lambda^{HighImg}, \sigma_\varepsilon\}$ .

### 2.4.1 Identification

On the demand side, the market-share model is identified by the variations across time. The intuition is that, different combinations of the parameter values would give different market equilibrium outcomes (observed market-share). For example, since the mean utility of outside option is normalized to be zero and constant over time, then the mean utility for property  $j$  at time  $t$  is identified through  $j$ 's market share at time  $t$ . Since a property  $j$  in period  $t$  is viewed as a bundle of property attributes  $X_{jt} = \{X^k\}_{k=1}^K$ , the coefficients for each  $X^k$  is then identified through the variations in market shares across different values of  $X^k$ . Similarly, the monthly fixed effects (i.e.,  $X^k$  for  $k=1,2, \dots,12$ ) is identified from changes in market shares across each month. For the endogenous variables price, the identification of property substitution patterns (i.e., the price self- and cross- elasticities) relies on the variation in the instrumental variables (lagged price and lagged number of reviews) that are assumed to be exogeneous to the aggregate demand shocks.

On the supply side, we first fix discount factor  $\beta$  to a constant between 0 and 1, as it cannot be jointly identified with the model primitives (Rust 1994, Magnac and Thesmar 2002). Since our model does not satisfy exclusion restriction and we are more interested in knowing how one's heterogenous coefficients would affect her choice of images and investment in service than in identifying  $\beta$ , we chose to fix  $\beta$  to 0.95 and identify other model primitives<sup>36</sup>.

We start with discussing the identification of photography cost and the service effort cost. Identifying the cost associated with medium-level images is straightforward. All properties must incur a cost of  $\lambda^{MedImg}$  if they update to a state of medium-level images, this is because Airbnb doesn't provide medium-quality images for free and hence hosts must pay on their own to have medium-quality images. As a result,  $\lambda^{MedImg}$  can be identified from the overall frequency that hosts if they transitioned from a quality level other than medium to level medium-level. Identifying  $\lambda^{HighImg}$  and  $\lambda_j^{effort}$  is a more complex. This is because there is an observational equivalence for two scenarios: 1) a host has a high cost of posting high-quality images and a low cost of investing service effort, and 2) a host has a low cost of posting high-quality images and a high cost of investing service effort. Both of the scenarios lead to the same observation: the host does not use high-quality images and as a result, he does not need to invest high-level effort in providing service. fact that helps us to separate the two scenarios is that Airbnb's professional photography program provides high-quality images to the same property for only once. For the case 1), the host would take the high-quality images for free (for the first time) and invest high-effort in service. When they do not qualify for the free service, they would not use high-quality images. For the case 2), we would not observe that the host's choice of image quality to vary a lot when they qualify versus not qualify for the free service.

---

<sup>36</sup> The main findings are insensitive to alternative discount factors we tested (0.9, 0.975, 0.995).

In the reality, we do not observe sufficient temporal variation in the image quality choices to help us to identify individual-specific photography cost (at least for the one-year panel of data). Hence, we identify a common  $\lambda^{HighImg}$  instead. However, the logic of separating  $\lambda^{HighImg}$  and  $\lambda_j^{effort}$  is the same: recall that approximately 30% of the properties in our sample had used the program by the time our observation started (i.e., by January 2016). Then for these properties if they updated to a state of high-level images, they must incur a cost of  $\lambda^{HighImg}$  as they cannot request a free photography service again. For other properties, the cost of posting high-level (professional) images are free, as they still qualify for requesting a free professional photography service from Airbnb. As a result,  $\lambda^{HighImg}$  can be identified from the overall frequency that hosts, who Airbnb will not provide images at quality level  $q$  for free, if they transitioned from a quality level  $q'$  to level  $q$  (where  $q' \neq q$  and  $q$  is high-level). Conditional on identified  $\lambda^{MedImg}$  and  $\lambda^{HighImg}$ , the variation in one's choice of service effort across periods with the same expected number of reviews in the next period helps us to identify another heterogeneous parameter  $\lambda_j^{effort}$ —the marginal cost of investing high service effort. To illustrate, recall that the probability that consumers write reviews depend on current number of reviews, expected 'gap' between chosen image quality and property quality, and the service effort. Hence, if two hosts have the same state and expected gap in a particular period, however one chose to invest high-effort and another invested low effort, then likely the latter has a high cost of investing service effort.

Lastly, conditional on identified  $\lambda^{MedImg}$ ,  $\lambda^{HighImg}$ , and  $\lambda_j^{effort}$ , the operating cost  $\lambda_j^{operate}$  is identified through the frequency that hosts observed to operate versus 'snooze' their listings, conditional on the expected revenue and costs. For example, with the same expected revenue and costs, if one host is observed to snooze the listing more often than another, then the former is likely to have a higher operation cost (or a higher value of outside option). The standard deviation of idiosyncratic shocks can be identified because we normalize the coefficient for revenue to 1.

#### 2.4.2 Estimating Demand-Side Model

Jiang et al. (2009) proposed a Bayesian approach of estimating an aggregated market share (BLP) model. The model is estimated using MCMC (Monte Carlo Markov Chain) algorithm.

Given the distributional assumptions on endogenous variable  $P$  and demand shock  $\eta$ , using Change-of-Variable Theorem, we derive the joint distribution of market share  $ms_t$  and  $P_t$  :

$$\pi(P_t, s_t | \bar{\theta}, \Sigma, \Omega, \delta) = \pi(\xi_t, \eta_t | \bar{\theta}, \Sigma, \Omega, \delta) J_{(\xi_t, \eta_t \rightarrow P_t, s_t)} = \pi(\xi_t, \eta_t | \bar{\theta}, \Sigma, \Omega, \delta) (J_{(P_t, s_t \rightarrow \xi_t, \eta_t)})^{-1}$$

where  $J_{(P_t, s_t \rightarrow \xi_t, \eta_t)} = \begin{vmatrix} \nabla_{\xi_t} P_t & \nabla_{\eta_t} P_t \\ \nabla_{\xi_t} s_t & \nabla_{\eta_t} s_t \end{vmatrix}$  is the Jacobian matrix  $J(s_t \rightarrow \eta_t) = \begin{vmatrix} \mathbf{I} & \mathbf{0} \\ \nabla_{\xi_t} s_t & \nabla_{\eta_t} s_t \end{vmatrix} = \|\nabla_{\eta_t} ms_t\|$ .

Furthermore, to ensure that the estimated covariance variance-matrix  $\Sigma$  is positive-definite, following the re-parameterization method used in Jiang et al. (2009), we use Cholesky decomposition and write:

$$\Sigma = U'U; U = \begin{bmatrix} e^{r_{11}} & r_{12} & \cdots & r_{1K} \\ 0 & e^{r_{22}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_{K-1,K} \\ 0 & \cdots & 0 & e^{r_{KK}} \end{bmatrix}$$

Lastly, given the priors on the parameters and likelihood function, the joint posterior distribution of the parameters is<sup>37</sup>:

$$\begin{aligned} \pi(\bar{\theta}, \Sigma, \Omega, \delta | \{P_t, s_t, X_t\}_{t=1}^T) &\propto L(\theta, r, \delta, \Omega) \times \pi(\bar{\theta}, r, \Omega, \delta) \\ &= \prod_{t=1}^T \left\{ \left( \begin{vmatrix} \nabla_{\xi_t} P_t & \nabla_{\eta_t} P_t \\ \nabla_{\xi_t} s_t & \nabla_{\eta_t} s_t \end{vmatrix} \right)^{-1} \times \phi \left( \begin{matrix} \xi_{jt} = P_{jt} - Z_{jt} \delta \\ \eta_{jt} = \mu_{jt} - X_{jt} \bar{\theta} \end{matrix} \middle| \begin{matrix} \mathbf{0} \\ \mathbf{0} \end{matrix}, \Omega \right) \right\} \\ &\times |V_{\bar{\theta}}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} (\bar{\theta} - \theta_0)' V_{\bar{\theta}}^{-1} (\bar{\theta} - \theta_0)\right\} \times \prod_{l=1}^K \exp\left\{-\frac{(r_{ll})^2}{2\sigma_{r_{ll}}^2}\right\} \\ &\times \prod_{l=1}^{K-1} \prod_{k=l+1}^K \exp\left\{-\frac{(r_{lk})^2}{2\sigma_{r_{off}}^2}\right\} \end{aligned} \quad (18)$$

where  $\pi(\bar{\theta}, r, \Omega, \delta)$  is specified priors on the parameters. Specifically, for variance-covariance matrix, we specify the priors on  $r = \{r_{lk}\}_{l,k=1\dots K, l \leq k}$  with  $r_{ll} \sim N(0, \sigma_{r_{ll}}^2)$ ,  $r_{lk} \sim N(0, \sigma_{r_{off}}^2)$  for the diagonal, and off-diagonal elements in matrix  $U$ , respectively. For the population mean for characteristics coefficients  $\bar{\theta}$ , as written in Equation (18), we specify a multivariate normal distribution prior:  $\bar{\theta} \sim MVN(\bar{\theta}_0, V_{\bar{\theta}})$ . We specify the following priors for  $\delta$  and  $\Omega$ :  $\delta \sim MVN(\bar{\delta}, V_{\delta})$ ,  $\Omega \sim IW(\nu_0, V_{\Omega})$ , where  $IW$  indicates an inverse Wishart distribution.

### MCMC Estimation Steps

The MCMC estimation steps follows a strategy of Gibbs sampling combined Metropolis steps (Jiang et al. 2009, Rossi et al. 2005). Briefly speaking, in each iteration of the MCMC, we first use Gibbs Sampler to draw the conditionals of  $\bar{\theta}, \delta, \Omega | r, \{ms_t, P_t, W_t, Z_t\}_{t=1}^T, \bar{\theta}_0, V_{\bar{\theta}}, \bar{\delta}, V_{\delta}, \nu_0, V_{\Omega}$  in a sequence. Then, conditional on updated  $\{\bar{\theta}, \delta, \Omega\}$ , data  $\{ms_t, P_t, W_t, Z_t\}_{t=1}^T$ , and priors  $(\sigma_{r_{ll}}^2, \sigma_{r_{off}}^2)$ , we update the variance-covariance matrix,  $\Sigma$ , by making draws of  $r$  through a Random-Walk (RW) Metropolis chain. Specifically, we draw a proposal of  $r$ , given the accepted  $r$  in the previous iteration:  $r_{new} = r_{old} + MVN(\mathbf{0}, \sigma^2 D_r)$ , where  $\sigma^2$  is one of  $(\sigma_{r_{ll}}^2, \sigma_{r_{off}}^2)$  depending on whether we're drawing a diagonal or off-diagonal element

<sup>37</sup> For the setup of hyper-parameters, we used diffuse priors. In appendix, we describe details on the choices of priors.

of  $D_r$  is a candidate covariance matrix.  $r_{new}$  is either accepted or rejected, based on ratio computed using Equation (18). The intuition is that, if conditional on data, priors, and other parameters updated in the Gibbs sampling step,  $\Sigma_{new}(r_{new})$ , relative to  $\Sigma_{old}(r_{old})$ , is closer the true posterior of  $\Sigma$ , then we should have  $(\bar{\theta}, \Sigma_{new}, \Omega, \delta|\{P_t, ms_t, X_t\}_{t=1}^T) > (\bar{\theta}, \Sigma_{old}, \Omega, \delta|\{P_t, ms_t, X_t\}_{t=1}^T)$ . In appendix we provide detailed technical notes of our estimation steps.

### 2.4.3 Estimating Supply-Side Model

Conditional on one's current state  $s = (s_{jt}, s_{-jt})$ , her wage and effort decisions can be described as sequentially solving a DP problem:

$$\{a_{jkt}\}_{t=0}^{\infty} = \underset{\{a_{jkt}\}_{t=0}^{\infty}}{\operatorname{argmax}} E_{\varepsilon_{jkt}} \left\{ \sum_{t=0}^{\infty} \cdot (\tilde{\Pi}_{jkt}(a_{jkt}|s_{jt}, s_{-jt}) + \varepsilon_{jkt}) \right\} \quad (19)$$

where  $\tilde{\Pi}_{jkt}$  is property  $j$ 's expected payoff from choosing action  $k$  in period  $t$  and  $\varepsilon_{jkt}$  is the random shock associated to action  $k$  that is received before  $j$  makes a decision.

As discussed in section 3.8., in such dynamic game with many players, computing an MPE is computationally infeasible, hence we use OE to approximate MPE. In a OE, the individual's conditional choice probability is a function of her own state  $s_{jt}$  only. The set of states of her peers,  $s_{-jt}$ , is captured by tracking an average industry state  $\bar{s}_t$ , which reflects the distribution of the number of the reviews across the properties. It can be seen one's action and payoff is influenced by her peers' state— $s_{-jt}$ , as it is the action of  $j$  and her peers and the subsequent state transitions that determine the average state in the next period. Then solving for an OE provides substantial computational advantage, as it converts a many-agent game problem into a problem similar to single-agent optimization, treating  $\bar{s}_t$  as a single state variable that is common across all individuals at time  $t$ . Thus, one can use any existing estimation method that can be applied to a single-agent discrete-choice dynamic programming (DDP) model to solve for an OE. Widely-used estimation strategy includes the nested fixed-point (NFXP) algorithm (Rust 1987) and conditional choice probability (CCP) based estimation (Hotz and Miller 1993, Aguirregabiria and Mira 2007).

In this paper, we use a Bayesian estimation strategy as this way we can flexibly incorporate individual heterogeneity—a key element in our model—in a hierarchical Bayesian framework (developed by Imai, Jain and Ching (2009), hereafter IJC). IJC algorithm allows estimating a heterogeneous model with a relatively low computational burden. In addition, it overcomes the problem of “curse of dimensionality”<sup>38</sup> when approximating the DP solution and avoids the complexity of searching for a global optimum in the

---

<sup>38</sup> The state space grows exponentially with the dimensionality of state variables, causing evaluating Bellman operator at every point in the state space infeasible.

space of the data likelihood function (IJC provides DP approximation that is comparable to state-of-the-art likelihood-based approaches, e.g., Keane and Wolpin (1994), Akerberg (2009). See Ching et al. (2012) for detailed discussions). The advantage of avoiding of searching in the parameter space, which usually requires the use of an optimization tool, is another reason we choose IJC algorithm. As we will discuss in section 4.4., Bayesian estimation approach can be easily combined with parallel computing and GPU computing techniques, without which it would be computationally infeasible given the large number of individuals and state space in our study.

### IJC Algorithm

We briefly introduce the logics and estimation procedure in IJC algorithm. In appendix we provide technical notes and details of implementing IJC.

IJC algorithm combines MCMC with DDP approximation, solving for the DP problem and making draws of structural parameters from the posterior distribution simultaneously. At each iteration  $m$  in the MCMC, IJC saves the simulated parameter vector  $\theta_{IJC}^{*m}$  and computes a corresponding pseudo-value function  $\tilde{W}^m(\theta_{IJC}^{*m})$ <sup>39</sup>. A total of the most recent  $N$  iterations of  $\{\theta_{IJC}^{*m}, \tilde{W}^m(\theta_{IJC}^{*m})\}$  are saved. When at new iteration  $m'$ , the simulated vector  $\theta_{IJC}^{*m'}$  is rejected or accepted by comparing the pseudo- posterior likelihood evaluated at the accepted parameters from the previous iteration,  $\theta_{IJC}^{*m'-1}$ , and at the proposed parameters at current iteration,  $\theta_{IJC}^{*m'}$ . When computing the pseudo-likelihood function, one needs to calculate the choice probability for each choice alternative. Recall that one solves for the DP problem by taking into account the value function (see Equation 19)), hence the likelihood function is also ‘pseudo-’ because the conditional choice probabilities are computed based on pseudo-value functions  $\{\tilde{W}^n(\theta_{IJC}^{*n})\}_{n=m'-N}^{n=m'-1}$  saved in the past  $N$  iterations. Specifically,  $\tilde{W}^{m'}(\theta_{IJC}^{*m'})$  is approximated by computing a (kernel-based) weighted average of the past  $N$  history draws of  $\{\tilde{W}^n(\theta_{IJC}^{*n}), \theta_{IJC}^{*n}\}_{n=m'-N}^{n=m'-1}$ , with the more weights attributed to history that have  $\theta_{IJC}^{*n}$  closer to current draw  $\theta_{IJC}^{*m'}$ . As Imai et al. (2009) proved, such an interactive steps of simulating parameter vector through the pseud0-Markov chain can effectively approach a steady state (after burn-in), where most of the structural parameters will be drawn from a distribution close to the true posterior distribution of the parameter vector.

---

<sup>39</sup> The pseudo-value function is obtained by applying the Bellman operator (i.e., solving for the value function) at the trial parameter vector. It is called ‘pseudo’ as the functions are evaluated at the simulated parameter vector not at the true parameter vector. Here the \* denotes that this is proposed parameter (regardless of whether it was accepted or rejected) at that iteration.

In summary, we use a Gibbs Sampler to sequentially simulate parameters of  $\left(\left\{\lambda_j^{effort}, \lambda_j^{operate}\right\}_{j=1}^J, \lambda, \Sigma_\lambda, \lambda^{MedImg}, \lambda^{HighImg}, \sigma_\varepsilon\right)$ , with  $\left(\lambda_j^{effort}, \lambda_j^{operate}\right) \sim MVN(\lambda, \Sigma_\lambda)$ . At each iteration  $m$ , we have the history of the drawn parameters and the associated pseudo-value functions:

$$\{\tilde{W}^n(\cdot; \left\{\lambda_j^{effort^{*n}}, \lambda_j^{operate^{*n}}\right\}_{j=1}^J, \gamma_{med}^{*n}, \gamma_{high}^{*n}, \sigma_\varepsilon^{*n})\}, \left\{\lambda_j^{effort^{*n}}, \lambda_j^{operate^{*n}}\right\}_{j=1}^J, \gamma_{med}^{*n}, \gamma_{high}^{*n}, \sigma_\varepsilon^{*n}, \bar{s}^{*n}\}_{n=m-1}^{n=m-N}$$

where  $\tilde{W}^n(\cdot; \cdot)$  indicates the pseudo-value functions at all possible state and  $\bar{s}^{*n}$  is proposed industry average state. Then each iteration  $m$  we simulate parameters of  $\left(\left\{\lambda_j^{effort}, \lambda_j^{operate}\right\}_{j=1}^J\right)$  using Metropolis-Hasting to draw a proposal, where we evaluate the pseudo-likelihood function given the observed hosts' choices and the choice-probabilities computed using  $\{\tilde{W}^n(\cdot; \cdot), \left\{\lambda_j^{effort}, \lambda_j^{operate}\right\}_{j=1}^J, \lambda, \Sigma_\lambda, \lambda^{MedImg}, \lambda^{HighImg}, \sigma_\varepsilon\}$ . In appendix, we provided technical details on the IJC estimation steps.

#### 2.4.4 Computation Challenges

Due to the large number of properties (900+) in our sample, estimating the demand and the supply model is computationally challenging. We solve this issue leveraging parallel GPU (Graphical Processing Unit) computing. Specifically, as Bayesian estimation does not involve searching optima in the parameter space, we easily implement the estimation (MCMC) leveraging GPU computing and distribute the computation for (independent) individuals to multiple cores for parallel processing. As GPU is specialized in vector/matrix (floating) operations, we were able to accelerate the estimation by as much as 4 to 60 times, with the more number of products, the more computation advantage of GPU computing.

### 2.5 Estimation Results

We report and discuss the estimation results, starting with the estimated coefficients of the demand side (property market-share estimation).

#### 2.5.1 Property Market Share Estimates (Airbnb Demand)

Table 16 presents the estimation results of the property market share model (see Equation 8). The model is estimated in a hierarchical Bayesian framework where Markov Chain Monte Carlo (MCMC) is used to make draws from the posterior distribution of the parameters  $\{\bar{\theta}, \Sigma, \Omega, \delta\}$ . Thus, we report the mean and the standard deviation of the posterior draws<sup>40</sup>. We performed MCMC diagnostics, by inspection of time-series plots with different starting points (Gilbride and Allenby 2004). The chains reached a common

<sup>40</sup> We run a total of 30000 MCMC iterations and drop the first 25000 for burn-in. H=1000 individual 'consumers' were simulated.

stationary convergence, suggesting that the chains have ‘forgotten’ the initial points and are drawn from the posterior distribution of  $\{\bar{\theta}, \Sigma, \Omega, \delta\}$ .

Several interesting findings in Table 16 note highlighting here. First, the mean coefficients on the month and area (neighborhood) fixed effect terms suggest significant dynamic patterns in the seasonality. In general, summer season (May-October) is peak season—February on average has the least demand (mean month effect=-3.803) with September has the highest demand (mean month effect=-1.106), and the area of East Harlem attracts the most demand (mean area effect=-1.516) while area of *Lower Manhattan* on average has the least popularity (mean area effect=-3.7675). Second, the positive and significant coefficient of number of reviews (mean=0.641, std.= 0.093) indicates that number of reviews helps to generate more bookings (demand). Third, the negative population mean (-0.623) of the coefficient for the driving (commute) to downtown area suggests that location plays a role when travelers choose lodging alternatives. Specifically, an average traveler to Manhattan prefers to stay somewhere close to the downtown area, possibly due a convenient public transit or concentration of attractions/restaurants/etc.—consistent with the positive coefficient of walk score (mean=0.486). Lastly, the positive coefficients of two image dummy variables confirm that good images, compared to low-quality images, can generate more bookings in current month. Specifically, as expected, the impact of high-level images (mean=0.959) is greater than the med-level images (mean=0.714).

**Table 16 Estimated Property Market Share Parameters**

VARIABLES #	Estimates <sup>+</sup>	
	Population Mean	Population Std. Dev.
<b>Preferences on Property Characteristics: <math>\bar{\theta}</math>, <math>diag(\Sigma)</math></b>		
<i>EntireHome</i>	0.5048 (0.1024)	0.2707 (0.0220)
<i>Apartment</i>	-0.8723 (0.1630)	0.3305 (0.0531)
<i>Bedrooms</i>	0.1945 (0.0969)	0.3977 (0.1166)
<i>Bathrooms</i>	0.0441 (0.1545)	0.3112 (0.0118)
<i>Beds</i>	-0.2037 (0.0704)	0.1118 (0.0204)
<i>MaxGuests</i>	-0.1062 (0.0451)	0.0497 (0.0054)

<i>MinStays</i>	-0.7346 (0.0567)	0.3227 (0.0148)
<i>NumReviews (scaled by 1/10)</i>	0.6407 (0.0934)	0.2113 (0.0317)
<i>MedImage</i>	0.7137 (0.1035)	0.3857 (0.0349)
<i>HighImage</i>	0.9592 (0.0997)	0.3164 (0.0289)
<i>DriveTime (100 mins)</i>	-0.62267 (0.2808)	0.2731 (0.0245)
<i>WalkScore (1/100)</i>	0.4855 (0.2971)	0.3468 (0.0132)
<i>log(NightlyRate)</i>	-4.004 (0.1043)	0.5015 (0.0140)

---

**Preferences on Monthly Effects (i.e., Seasonality):  $\bar{\theta}$ ,  $diag(\Sigma)$**

---

<i>January</i>	-2.9598 (0.2296)	0.1397 (0.0240)
<i>February</i>	-3.8027 (0.2276)	0.1382 (0.0762)
<i>March</i>	-2.4117 (0.2297)	0.1712 (0.0294)
<i>April</i>	-1.7310 (0.2303)	0.2008 (0.0067)
<i>May</i>	-1.3084 (0.2313)	0.1868 (0.054)
<i>June</i>	-1.7538 (0.23077)	0.1623 (0.0198)
<i>July</i>	-2.1646 (0.2316)	0.2048 (0.0226)
<i>August</i>	-1.9537 (0.2317)	0.1873 (0.0128)
<i>September</i>	-1.1059 (0.2373)	0.2146 (0.0061)
<i>October</i>	-1.5274	0.2018

	(0.2357)	(0.0381)
<i>November</i>	-2.5145	0.2077
	(0.2347)	(0.0091)
<i>December</i>	-2.0075	0.2549
	(0.23437)	(0.0064)

---

**Preferences on Area Effects (i.e., Neighborhood Popularity):  $\bar{\theta}$ ,  $diag(\Sigma)$**

---

<i>Central Harlem</i>	-3.0489	0.4122
	(0.33926)	(0.0581)
<i>Chelsea and Clinton</i>	-2.1104	0.5739
	(0.3478)	(0.0830)
<i>East Harlem</i>	-1.5159	0.3784
	(0.2479)	(0.0826)
<i>Gramercy Park and Murray Hill</i>	-2.5231	0.5767
	(0.3438)	(0.1592)
<i>Greenwich Village and Soho</i>	-2.2557	0.6069
	(0.3494)	(0.1083)
<i>Inwood and Washington Heights</i>	-3.7201	0.5028
	(0.2389)	(0.0872)
<i>Lower East Side</i>	-2.7561	0.6121
	(0.3449)	(0.1034)
<i>Lower Manhattan</i>	-3.7675	0.8407
	(0.4508)	(0.1226)
<i>Upper East Side</i>	-2.5014	0.4312
	(0.2480)	(0.0843)
<i>Upper West Side</i>	-2.6499	0.5957
	(0.2454)	(0.1183)

---

**Covariance Matrix of  $\xi_{jt}$  and  $\eta_{jt}$ :  $\Omega$**

---

$\Omega_{11}^{price}$	0.2193
	(0.0029)
$\Omega_{12}^{price} (\Omega_{21}^{price})$	0.0421
	(0.0126)
$\Omega_{22}^{price}$	13.6129
	(0.1792)

$\Omega_{11}^{image}$	0.1158 (0.0015)
$\Omega_{12}^{image}(\Omega_{21}^{image})$	0.0352 (0.0158)
$\Omega_{22}^{image}$	11.6165 (0.2793)
$\Omega_{11}^{review}$	0.3350 (0.0281)
$\Omega_{12}^{review}(\Omega_{21}^{review})$	0.0394 (0.0281)
$\Omega_{22}^{review}$	17.2284 (0.2532)
<hr/>	
Observations	8862
<hr/>	
+ Posterior means of estimates are computed on the 25000th~30000th draws.	
+ Table presents the posterior mean of each estimate, with the posterior standard deviation reported in the parenthesis below each posterior mean estimate.	
# Note due to limited space, for $\Sigma$ only diagonal elements (i.e., population variance of the coefficients) are presented.	
<hr/>	

### Traveler Preferences' Heterogeneity and Correlation

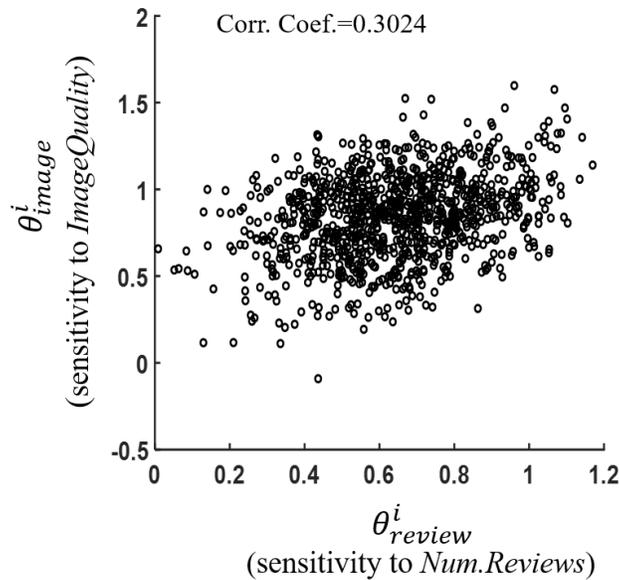
We explore heterogeneity and correlations in the identified individual preferences. Individual heterogeneity can be obtained from the population variance of the coefficients for each property characteristics, i.e., the diagonal of  $\Sigma$  (recall that individual-specific parameter  $\theta_i \sim N(\bar{\theta}, \Sigma)$ ). The correlation between every two population preferences is reflected through the off-diagonal of  $\Sigma$ . Since in our estimation, we simulated  $H=1000$  individual consumers over which we integrate to compute the market shares, we can simply use the identified coefficients of the  $H$  individuals to plot and exhibit the heterogeneity and correlation. Here we focused on the correlation between consumers' preferences (i.e., coefficients for each property attribute) on property image quality and number of guest reviews, as this correlation highlights an interesting trade-off problem for the hosts.

To examine the preference correlations, we look at how guests' taste sensitivity in property images is correlated with the sensitivity in number of reviews. In Figure 15 we provide a scatter plot between guests' review sensitivity (horizontal axis) and image sensitivity (vertical axis)<sup>41</sup>. The correlation

<sup>41</sup> In the model we have two dummy variables for property images: *MedImage*—indicating a med-level image, and *HighImage*—indicating a high-level image. The scatter plot uses an average of the coefficients for the two variables.

coefficient of the two sets of individual parameters is 0.302 (p-value <0.001), suggesting that consumers who give more weightages to image quality tend to be those who value the number reviews more. This positive correlation reveals an interesting ‘trade-off’ problem for AirBnB hosts—it is the same set of consumers who reward the hosts in the short-run that will punish him in the long-run. To see why this is happening, suppose a host is facing a pool of consumers who are ‘sensitive’ to the quality of property images (i.e., elasticity of image quality is high). Then posting high-quality images would be more effective in generating present bookings for this host than for other hosts. However, her consumers (potential guests) are those who also care about the number of reviews. As a result, the ‘penalty’ of unable to get more reviews due to images-induce high expectations would also be greater for her than for others if she is unable to meet guests’ expectations. Thus, the incentive of posting high-quality images may be reduced by the risk of penalty. As we discuss in the next section, it is the real quality of the property and host’s ability in investing service that determines what is the optimal choice of images.

**Figure 15 Scatter Plots of Individual Consumers’ Coefficients**



### 2.5.2 Host Choice of Images and Effort Estimates (AirBnB Supply)

In Table 17 we report the estimation results of the dynamic game model on the supply side (hosts’ optimization problem on choice of images and service effort, see Equation 14) and 16). Similar to the property market share estimation (section 5.1), we employ a hierarchical Bayesian framework and draw Markov chains from the posterior distribution of the parameters  $(\{\lambda_j^{operate}, \lambda_j^{effort}\}_{j=1}^J, \lambda_{med}, \lambda_{high})$ .

Further, we assume that individual’s operation cost,  $\lambda_j^{operate}$ , and on marginal service effort cost,  $\lambda_j^{effort}$ ,

are drawn from  $MVN(\bar{\lambda}, \Sigma_{\lambda})$ . Thus, in the table we report the population mean and population standard deviation of  $(\lambda_j^{operate}, \lambda_j^{effort})$ . We performed MCMC diagnostics and confirmed that common stationary convergence was achieved. We use the draws after having achieved convergence to compute the sample posterior means and standard deviations of the parameters<sup>42</sup>.

The negative coefficient of effort costs confirms our intuition that investing in service effort is costly for hosts. Particularly, if evaluating its impact in monetary-term, we obtain a marginal effort cost equivalent to be  $\$0.3244 * 100 = \$32.4/\text{day}$ <sup>43</sup>. That is, for an average host, if consider per-day return, she would be willing to invest a high effort in providing service to the guests only if she could charge an extra \$32.4 per day. Furthermore, the population standard deviation in the estimated effort cost suggests there exists heterogeneity in the consumers' ability of investing effort in providing guest service.

The estimated common photography costs suggest that posting good images, relative to posting low-quality images are costly. As expected, having high-level images cost more than having med-level images. The big cost may explain the observation that hosts, once not qualified for AirBnB's free photography program, rarely use high-quality photos on their own. Also, we should note that, the cost includes but is not limited to the fee for hiring a professional photographer. It includes all the costs associated with the preparation work. For example, one may need to spend time on searching photographers, communicating with them back-and-forth, scheduling a day, preparing the house (organizing and cleaning the place) etc. The estimated  $\lambda_j^{operate}$  suggests that the outside option value combined with the operation cost on Airbnb for an average AirBnB host in Manhattan is  $\$10.2715 / * 100 = \$1027.2/\text{month}$  in monetary term<sup>44</sup>. If we consider an average nightly rate of \$228, the result suggests that unless renting out her property for 6 days/month, one may prefer to leave AirBnB and choose an outside option (e.g., for long-term lease). The estimated  $\lambda_j^{operate}$  suggests that costly listing-managing and low property occupancy are may explain the high dropout rate on AirBnB. This is particularly a concern for AirBnB in areas where the hosts can receive high-value outside options (e.g., Manhattan has strong housing market).

---

<sup>42</sup> We run a total of 20000 MCMC iterations and drop the first 10000 for burn-in (the chains started to converge after about 10000 draws). We use a kernel bandwidth of 0.05 for the parameters and a bandwidth of 1 for the industry average state. We store N=500 of past pseudo-value functions for approximating  $E_{max}$  functions during the MCMC iterations.

<sup>43</sup> The property revenue in host's per-period payoff function is computed with  $ReservationDays \times NightlyRate/100$ , where property nightly rate was scaled by 1/100(i.e., the unit is 100 USD). So, the effect of effort cost to an average host equals  $0.3244 * 100 \text{ USD} = \$32.44 \text{ USD}$ .

<sup>44</sup> Since  $\gamma_{operate}$  captures two components: one is the cost of managing an active listing, the other is the outside option value (or opportunity cost)

**Table 17 Estimated Host Supply Model Parameters**

VARIABLES #	Estimates <sup>+</sup>	
	Posterior Mean	Population Std. Dev.
<b>Individual Parameters: <math>\rho, \Sigma_\rho</math></b>		
$\lambda_j^{effort}$	-0.3244 (0.0722)	0.1488 (0.0441)
$\lambda_j^{operate}$	-10.2715 (3.2015)	1.2866 (0.5877)
<b>Common Parameters</b>		
$\lambda^{MedImg}$	-3.018	--
$\lambda^{HighImg}$	-5.975	--
$\sigma_\varepsilon$	1.6271	--
Observations	11496	
<p>+ Posterior means and standard deviations of estimates are computed on the 10000th~20000th draws.            + Table presents the posterior mean of each estimate, with the posterior standard deviation reported in the parenthesis below each posterior mean estimate.</p>		

## 2.6 Policy Simulation

Our structural model on hosts' choice of images and effort, combined with property market share estimations, allows to assess the impact of AirBnB's photography policy on hosts' supply of lodgings as well as on its revenue. In this section, we conduct counterfactual analyses to analyze hosts' choice and their returns under simulated policies. We use the individual hosts in our sample and the use their estimated parameters to solve for the DP problem and 'observe' their behaviors over time, as they were making choices following the solutions of DP. Each policy was implemented for T=24 periods (i.e., two years), starting with their initial state assuming everyone just joined the platform. We implement each policy for 1000 runs and report the averaged outcomes over the 1000 runs.

### 2.6.1 Should AirBnB Provide Medium-Quality instead of High-Quality Images for Free?

The first policy experiment is motivated by the observation that a large number of hosts did not utilize AirBnB's professional photography program, which offered hosts high-quality property images for free. Our structural model suggests that hosts face temporal trade-off between short-term image-induced bookings and long-term review-induced market share. As a result, hosts may end up using low-quality images as high-quality images may hurt the host in the long-run (hosts may not use medium-quality images

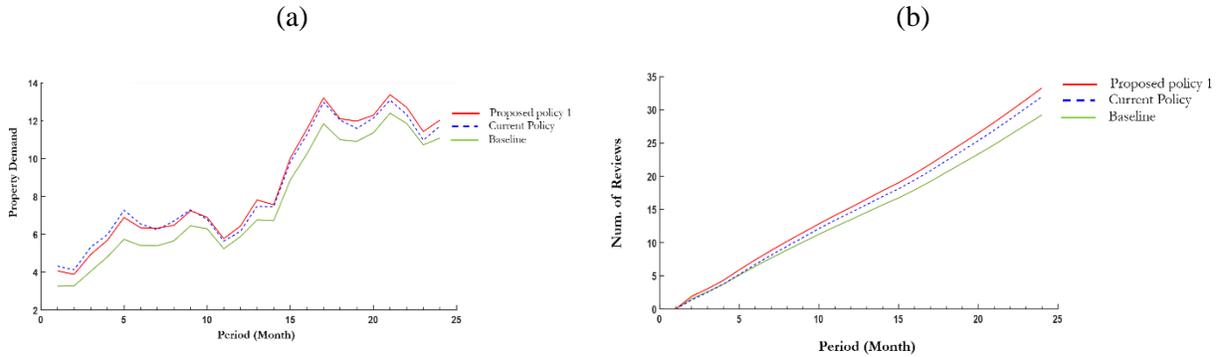
due to the cost and relative-low return on property bookings). Hence, we want to see can AirBnB do better by providing medium-quality images for free to all hosts. Specifically, we consider two options: 1) providing high-level images to all hosts for free—we refer this policy option as “current policy”, and 2) providing med-level images to all hosts for free—we refer this policy option as “proposed policy 1”. The baseline is where AirBnB does not offer any free photography-related program.

In practice, a home-sharing platform such as AirBnB can control the quality level of offered property images by guiding/training their photographers. In the simulation, the policies are implemented by reducing ones’ cost of posting high-level (med-level) images to 0 under the current policy (proposed policy 1). In Figure 16 we plot the average outcomes of the 958 properties (vertical axis) over time (horizontal axis) under the two alternative policies as well as the baseline. As expected, AirBnB would do better under both of the two image policies than under the baseline policy (where AirBnB does not implement any image policy).

Figure 16 (a) reports the average property demand (i.e., number of reserved days) across all properties in our sample. As can be seen, the current policy dominated proposed policy 1 in the short-run (for the first four periods). Interestingly, the advantage of current policy, relative the proposed policy 1, vanishes quickly over time. After seven to nine periods, the average property booking under proposed policy 1 is greater. In the long-run, proposed policy 1 outperformed the currently policy (1.3 additional reservation days/month vs 0.8 additional reservation days/month). The interpretation is that, medium-level images, compared to high-level images, despite forming a smaller expected utility for the consumers, has a greater effect on property demand in the long-run as they, with lower risks of creating a dissatisfactory gap, help hosts to obtain new reviews. Moreover, individual hosts who might end up using amateur (low-level) images to avoid the dissatisfactory gap under the current policy, now use free medium-level images to make more revenues under the proposed policy.

Such interpretation is supported by Figure 16 (b), where we plot the evolution of the average number of reviews across all properties. Clearly, the average number of review experienced a steady greater growth under proposed policy 1 than under the current policy. As properties accumulated consumer reviews, the significance of reviews started ‘cancelling out’ the relative negative effect of medium-quality images compared with high-quality images on property’s bookings. In the end, the properties are rented out more since the number of reviews is a key driver in consumers’ decisions on lodging options.

**Figure 16 Policy Simulation: The Impact of AirBnB’s Image Policies on  
Property Demand and Guest Reviews <sup>(\*)</sup>**



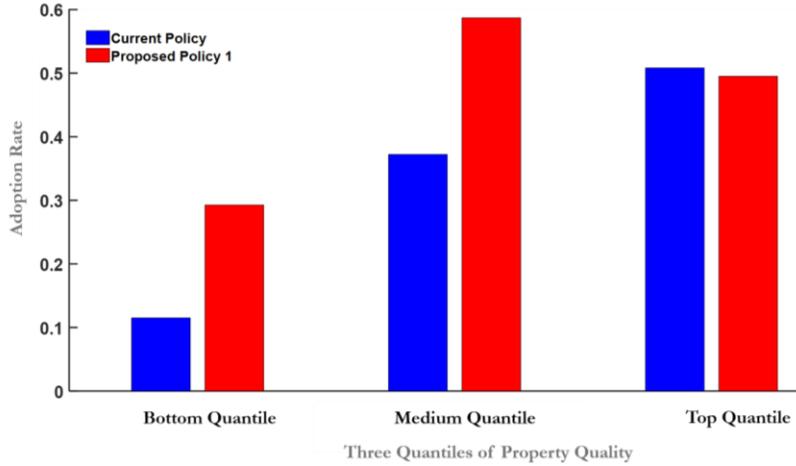
\* We plot starting the periods following the initial (entry) period.

### Understanding the Mechanism: Host Adoption of Current Policy versus Proposed Policy 1

To better understand the effect of proposed policy 1, here we discuss the adoption rates of Airbnb’s current policy versus the proposed policy 1 across the properties with different qualities. To do so, we categorize properties into three quantiles based on their property fixed effects,  $\alpha_j$ , identified from Law-of-Motion analysis (see Table 15 and Equation (5)). Note that though the quality of each property  $j$ ,  $PropertyQuality_j$ , is unobserved, the property fixed effect  $\alpha_j$  is an increasing function of  $PropertyQuality_j$ . As a result, the categorization gives us three quantile groups of properties whose quality fall into three levels—namely, high-level, medium-level, and low-level.

In Figure 17, we present the proportion of properties that adopted Airbnb’s current policy (in blue bar) and that adopted our proposed policy 1 (in red bar) across the three quantiles of property qualities. From the left to the right along the horizontal axis indicates the property quality quantiles vary from low-, medium-, to high- levels. The bars indicate that once medium-quality images are available for free (proposed policy 1), properties that fall into the low- and the medium- quantiles have significantly greater incentive to adopt, than adopting high-quality images (current policy). In contrast, the adoption rates among the properties in the high- quantile do not vary much when the policy shifts from current policy to proposed policy 1. This finding supports our interpretation above—that (some of) the properties in the low- and medium- quantiles would end up using low-quality images under the current policy, because they would face a risk of creating a big negative gap if they used the free high-quality images. Under the proposed policy, they have the motivation to adopt the medium-quality images as the risk of dissatisfaction gap would reduce. Hence, by making medium-quality images for free, Airbnb is able to improve the utilized capacity of the large number of low- and medium- tier properties.

**Figure 17 Comparing Adoption Rates of Policies Across Property Quality Quantiles**



### 2.6.2 Should Airbnb Discriminate Hosts on Offering Photography Policy

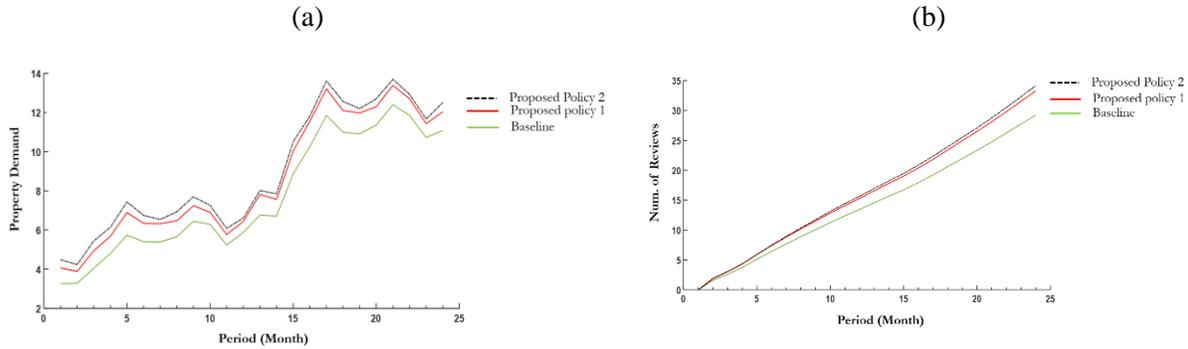
As shown in Figure 17, properties of which the qualities in low- and medium- quantile have strong incentive to adopt proposed policy 1, as medium-quality images, compared to high-quality images, are better representations (better matched to) of their properties. Since there exists a large degree of heterogeneity in the quality of Airbnb properties, and in hosts’ ability of providing good service (see Table 17), intuitively a better strategy is likely to be the one provides different qualities of images to different hosts. Since the true property quality and host’s type are unobserved, Airbnb in reality could implement a (second-degree) discrimination on providing the photography program.

Hence, we propose that Airbnb can offer a menu of image quality choices for free. The menu includes both high- and medium- level of property images (images examples are provided) and allow the hosts to self-select which program they want. We hypothesize that such a discrimination would incentive hosts with small  $\lambda_j^{effort}$  and/or high property quality to use high-level images as its relatively less costly for them to provide a good service that matches the high expectation. Hosts with big  $\lambda_j^{effort}$  and/or with lower-quality of properties, on the other hand, will self-select to use the medium-level images as providing a good service that matches high-level images is costly for them. We refer this alternative policy as ‘proposed policy 2’ and compare its impact to the proposed policy in simulation 1 (i.e., proposed policy 1’ in section 6.1).

As Figure 18 (a) shows, the proposed policy 2 (discriminative policy) improves the average property demand (vertical axis) in the long-run, by improving average property demand by 1.7 additional reservation days in a month (versus the long-term impact of 1.3 additional reservation days under the proposed policy 1). Hence, a photography program that provide both medium- and high- quality images for free can most effectively improve the property performance on Airbnb.

**Figure 18 Policy Simulation: The Impact of AirBnB’s Discriminative Image Policies on**

**Property Demand and Guest Reviews (\*)**



\* We plot starting the periods following the initial (entry) period.

**2.7 Discussion and Conclusion**

This paper focuses on investigating how the AirBnB hosts make decisions on the quality of photographs to post. Unlike property attributes that cannot be controlled by the hosts such as the type, size, and location of the property, images are in immediate control of the hosts. Previous study found a strong advertisement impact of property images on property booking, with professional images (high-quality), relative to amateur (low-quality) images, increase the present demand by 14.3%. The advertisement effect arises in the context of AirBnB because (1) there exists a large variation in offered properties and (2) consumers rely heavily on visual information in order to ease decision-making. Recognizing the importance of images, AirBnB in 2011 started offering highest quality professional images to all the hosts for free. To AirBnB’s surprise, only 30% of the hosts used the AirBnB professional photography program after its launch.

This study provides an explanation for hosts’ behavior. We posit that the host’s decision on the quality of images to post depends on the following factor: (1) the advertising impact of the photos on present demand, (2) the cost of photos, (3) the impact of the photos on the satisfaction level of the guest post consumption, and (4) the host’s cost (ability) of investing effort in providing god service to the guest. As suggested by the reference dependence literature, consumers’ satisfaction level in the post consumption depends not only on the actual outcome from consuming the product, but also on her reference point—the individual’s pre-consumption expectation. A decrease in the consumers’ satisfaction level would in turn adversely impact the future demand through consumer reviews. Since high quality photographs can create unrealistically high expectations, a host would be hesitant to post high quality photographs (even if they were free) if the actual property is not as good as portrayed in the images, especially if the hosts are unable to provide a high level of service to match the high expectations.

Our goal is to disentangle the aforementioned factors that influence the host's decision on the type of photographs to post, and to explore policies that platforms such as AirBnB can employ to improve the property performance. To achieve this goal, we model hosts' periodic (monthly) decisions on the quality of property images to post, and the quality of service to provide. The image decision affects the host's profits in the short run through the cost of posting (preparing, shooting etc.) of images and advertisement effect on present demand. And it affects the host's future profits through its impact on guests' dissatisfaction, which affects consumers' likelihood of writing reviews. The service decision impacts the host's profits in the short run through the service costs and in the long run through impacting the guests' satisfaction level.

We estimate our model exploring a unique panel data spanning one-year for 958 properties in Manhattan (New York City). We observe the dynamics in hosts' choice property images of and provided service quality (proxied with host responsiveness, see section 2.2). We find that guests who value professional images more, also value the number of reviews more, revealing an interesting trade-off problem for the hosts. Further, the estimation results highlight that hosts are heterogeneous in their ability (cost) in investing in service and that it is costly for hosts to post, on their own, images with above-average quality. Policy simulations suggests two proposed photography policies for AirBnB that outperforms its current policy (providing high-level images for free to the hosts) in the long-run. The first proposed policy provides medium-level of images for free to the hosts. Compared to the baseline where no policy is offered, the first proposed policy and the current policy improves the average property demand by 1.3 reservation days per month and 0.8 reservation days in the long-run, respectively. Interestingly, the proposed policy was dominated by the current policy through the first four periods. The interpretation is that, medium-level images, compared to high-level images, despite forming a smaller expected utility for the consumers, has a greater effect on property demand in the long-run as they, with lower risks of creating a dissatisfactory gap, help hosts to obtain new reviews. The second proposed policy offers both high- and medium- level images for free and allow the hosts to self-select to choose which program they want. We show that the second proposed policy performed the best in the long-run, by improving average property demand by 1.7 reservation days per month.

There are a few limitations to this study and directions for future research. First, we do not model hosts' decisions on pricing for model tractability. In addition, individual AirBnB hosts have very limited information to optimally set prices. In a context where prices are likely to be optimally determined (e.g., modeling firms' decisions), modeling pricing decision will help to identify the marginal (production) cost and allows for further implications. Second, due to the computational tractability, we look at only one market (Manhattan) across multiple period. With the future advances in computation power, one may

conduct a study across different markets (e.g., Chicago, Miami etc.), which may give interesting insights on potential impact of city and travelers' demographics on hosts' choices of posting images

## Reference

Ackerberg, D. A. (2009). A new use of importance sampling to reduce computational burden in simulation estimation". *Quantitative Marketing and Economics* 7(4): 343-376.

Aguirregabiria, V. & Mira, P. (2007). Sequential Estimation of Dynamic Discrete Games. *Econometrica* 75(1): 1-53.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica* 63(4), 841-890.

Bellman, R. (1957). Dynamic Programming. Princeton University Press.

Ching, A., Imai, S., Ishihara, M., & Jain, N. (2012). A practitioner's guide to Bayesian estimation of discrete choice dynamic programming models. [\*Quantitative Marketing and Economics\*](#) 10 (2):151–196.

Davis, P. 2006. Spatial competition in retail markets: Movie theaters. *RAND Journal of Economics* 37 (4): 964-982.

Dubé, J., Fox, J., & Su, C. (2012). Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation. *Econometrica* 80(5), 2231-2267.

Ericson, R., & Pakes, A. (1995). Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *The Review of Economic Studies* 62(1), 53-82.

Farias, V., Saure, D., Weintraub, G.Y. (2012). An approximate dynamic programming approach to solving dynamic oligopoly models. [\*The RAND Journal of Economics\*](#) 43(2):253–282.

Fradkin, A. & Holtz, D. (2018). The Determinants of Online Review Informativeness: Evidence from Field Experiments on Airbnb *Working Paper*.

Genesove, D. & Mayer, C. (2001). [Loss Aversion And Seller Behavior: Evidence From The Housing Market](#). *Quarterly Journal of Economics* 116(4): 233-1260.

Gilbride, T., & Allenby, G. (2004). A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules. *Marketing Science* 23(3), 391-406.

Gul, F. (1991). A Theory of Disappointment Aversion. *Econometrica* 59(3): 667-686.

- Huang, Y, Singh, P., Ghose, A. (2015). A Structural Model of Employee Behavioral Dynamics in Enterprise Social Media. *Management Science* 61(12):2825-2844.
- Hotz, V. & Miller, R. (1993). Conditional Choice Probabilities and the Estimation of Dynamic Models. *The Review of Economic Studies* 63(3): 497-529.
- Houde, JF. 2012. Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline. *American Economic Review* 102 (5): 2147-82.
- Imai, S., Jain, N., & Ching, A. (2009). Bayesian Estimation of Dynamic Discrete Choice Models. *Econometrica* 77 (6): 1865-1899.
- Jiang, R., Manchanda, P., & Rossi, P. (2009). Bayesian analysis of random coefficient logit models using aggregate data. *Journal of Econometrics* 149 (2): 136-148.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2) 263-291.
- Keane, M. P., & Wolpin, K. I. (1994). The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. *Review of Economics and Statistics* 76(4), 648–672.
- Kőszegi, B., & Rabin, M. (2006). A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics* 121(4), 1133-1165.
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems (NIPS)*
- Li, H. & Srinivasan, K. (2018). Competitive Dynamics in the Sharing Economy: An Analysis in the Context of Airbnb and Hotels. *CMU working paper*.
- Li, J., Moreno, A. & Zhang, D. (2016). Pros vs Joes: Agent Pricing Behavior in the Sharing Economy. *Ross School of Business Paper* No. 1298. Available at SSRN: <https://ssrn.com/abstract=2708279> .
- Mas, A. (2006). Pay, Reference Points, and Police Performance. *The Quarterly Journal of Economics* CXXI (3): 787-821.
- Nevo, A. 2000. Mergers with differentiated products: The case of the ready-to-eat cereal industry. *RAND Journal of Economics* 31(3): 395-421.
- Proserpio, D., Xu, W., & Zervas, G. (2017). You Get What You Give: Theory and Evidence of Reciprocity in the Sharing Economy. *Working Paper*

- Rossi, P., Allenby, G., & McCulloch, R. (2005). Bayesian Statistics and Marketing. *John Wiley & Sons*, Ltd.
- Rust, J. (1994). Structural Estimation of Markov Decision Processes. *in Handbook of Econometrics* Vol. 4, ed. by R. Engle and D. McFadden. Amsterda.
- Simonyan, K., A. Zisserman. 2015. Very deep convolutional networks for large-scale image recognition. *In International Conference on Learning Representations (ICLR)*
- [Shampanier](#) et al. (2007). Zero as a Special Price: The True Value of Free Products. *Marketing Science* 26 (6): 742-757.
- Sudhir, K., 2001. Competitive pricing behavior in the auto market: A structural analysis. *Marketing Science* 20(1): 42-60.
- Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics* 106(4), 1039-1061.
- Weintraub G., Benkard, C., & Van Roy B. (2008). Markov perfect industry dynamics with many firms. *Econometrica* 76(6):1375–1411.
- Weintraub, G., Bendard, C., & Van Roy, B. (2010). Computational Methods for Oblivious Equilibrium. *Operations Research* 58(4), pp. 1247–1265.
- Zhang, S., Lee, D., Singh, P., & Srinivasan, k. (2018). How Much Is an Image Worth? Airbnb Property Demand Estimation Leveraging Large Scale Image Analytics. *Working Paper* Available at SSRN: <https://ssrn.com/abstract=2976021>.
- Zervas, G., Proserpio, D., & Byers, J. (2015). A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average. *Working Paper*.

## Appendix to Chapter 2

### I. Technical Notes on Deep Learning-based Image Quality Classification

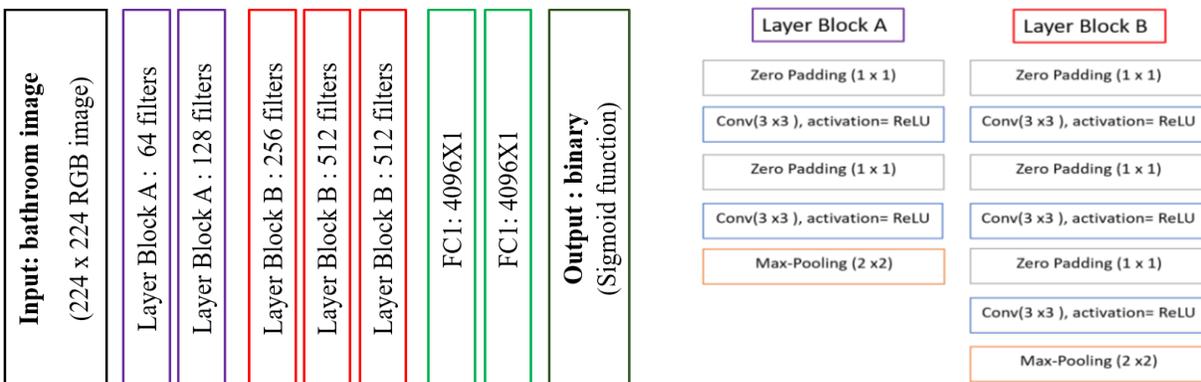
With a training data consists of images and the labels (in our study, the label is image quality), the task here is to build an image quality classifier that predicts, for any given input image. Given the large number of hosts in our sample, we leverage the advances in Convolutional Neural Networks (CNNs, an emerging deep learning framework, see Krizhevsky et al. 2012, Simonyan and Zisserman 2015) to build our classifier. Specifically, we use VGG16 model (Simonyan and Zisserman 2015).

#### 1. Convolutional Neural Networks and VGG16

A CNNs is a special kind of a deep learning model. As shown in Figure 19, a deep learning model consists of a sequence of layers, with each layer containing multiple neurons. Each layer is basically a multidimensional matrix, with each neuro ‘carrying’ a weight that represents the numeric value of each element. The number of layers that carry weight define the ‘depth’ of a deep learning model.

In a deep learning framework, high dimensional data such as images and texts is expressed as multidimensional matrices/arrays. Then the model processes the data through the neuron layers implementing matrix multiplication on the data. What defines a CNNs is a special layer—convolution layer, which operates dot productions on the input data (below we will describe operation of convolution layers). The CNN model processes data through matrix multiplication between the input image and the first layer of neurons. This operation generates an intermediate output (also represented by a multi-dimensional matrix), which can be viewed as ‘useful information’ extracted from the image and serve as the input for the next layer. Such implementations continue till the last layer of the model, i.e., the output layer that computes the probability distribution over the multiple labels. The probability distribution is then converted to labels.

**Figure 19 Description of Architecture and Layer Description of the VGG Model**



**Filters:** Indicates the number of convolution windows (i.e., # feature maps) on each convolution layer.

**Zero Padding:** pad the input with zeros on the edges. To control the spatial size of the output. Zero padding has no impact on the predicted output.

**Max-pooling:** subsampling method. A 2x2 window slides through (without overlap) each feature map at that layer, and then the maximum value in the window is picked as representation of the window. Reduces computation and provides translation invariance.

## 2. Operations of Key Layers

We describe convolution layer and pooling layer, which are the key layers in a CNNs.

### *Convolution Layer*

The convolution layer is the most important and unique layer in the CNN. A convolution layer consists of a stack of so called convolution filter or convolution kernel. A convolution filter is simply a matrix with each element representing a numeric value. For example, in a convolution block, a convolution layer with a size of 3X3 and hence consists of 9 such numeric values<sup>45</sup>. Such a matrix, treating an image or an intermediate input as a matrix, operates a dot production by ‘sliding’ through the input. Therefore, for an input with relative large size (e.g., 224X224), a 3X3convolution filter operates dot production for every 3X3 patch on that input matrix. The nice features of convolution operation are that: 1) it reduces the dimensionality of parameters, and 2) it well explores and reserves the (local) spatial relationships of the input. Particularly, an intuitive example of the second feature is that: if a convolution kernel extracts a particular oriented edge of an object, then operating this kernel on every small square (e.g., 3X3 and 1X1) on an image would extract all edges in that direct from the image. Many of such kernels that extract edges would extract edges in all directions—potentially constructing the contour of an object. As can be seen in **Figure 19**, each of the blocks consist of varying numbers of convolutional filters (e.g., 64, 128, 256, 512, 1024, and 2048filters). Hence, these kernels extract features from an input data, which represents the extracted features from the preceding layers. Towards the output layer in the CNN, the filters combined extract higher- and higher- level features. That is, the CNN is able to extract a hierarchical structure of features, that are related to predict the output labels.

### *Pooling Layer*

It’s a common practice in CNN to insert a pooling layer in-between the successive convolution layers. A pooling layer is a small square filter. In our model, the pooling filter is a 3X3 matrix. Similar to the operation of convolution filter, an average-pooling layer applies to every 3X3 square patch on an input data. The function of a pooling layer is to pick and using the average value in that 3X3 square. Adding pooling layers

---

<sup>45</sup> The size of a convolution layer is a choice of the model architecture. 3X3 is a widely-used choice. Another common choice is 5X5.

can reduce the spatial size of the intermediate features and the dimension of the trained parameters in the model. Particularly, it helps to efficiently prevent the problem of over-fitting.

### 3. Training Technical Notes

To effectively learn image features that have predicative power on image aesthetic quality labels, we leverage transfer learning and build our model on top of an existing deep learning model that was well-trained for a related task (Zhang et al. 2015). Specifically, we adopted the model of VGG16, with the output layer in the original VGG16 removed as it was specific to the original task (object classification). We then add three fully connected layers on top of that (dimension of 1), where the last layer is output layer.

To improve the training process, we initialize the model weights with the pre-trained weights of the original VGG16 and then fine-tune the parameters. For images, the extracted information is generic, to some extent, across various tasks (e.g., early layers in CNNs serve as edge and contour detectors). Hence, we were able to optimize our model starting from a point where it was already close to ‘optimum’. Hence, we efficiently improved the learning process of our model, with the initialized weights able to extract relevant features from the images. The added fully connected layers, without pre-trained weights available, were initialized with LeCun’s uniform scaled initiation method (LeCun et al. 1998, LeCun et al. 1998).

To improve the generalization power of the trained model, we employed a real-time data augmentation method, by randomly flipping, rescaling, and rotating the training samples during the training process (Krizhevsky et al. 2012). Specifically, we implement a real-time (i.e., during training) image transformation over each image in the training sample, by randomly 1) flipping input image horizontally, 2) rescaling input image within a scale of 1.2, 3) rotating the image within 20°. This method introduces random variation in the training sample, increasing the training set size and reducing the overfitting.

## II. Technical Notes on Estimation Strategy

### 1. Estimating Demand Model

Jiang et al. (2009) proposed a Bayesian approach of estimating an aggregated market share (BLP) model. The model is estimated using MCMC (Monte Carlo Markov Chain) algorithm.

As described above, a key step is to specify the distributional assumptions on endogenous price  $P$  and demand shock  $\eta$ , where endogenous variable  $P$  and exogenous variables  $Z$  satisfies:

$$\begin{cases} P_{jt} = Z_{jt}\delta + \varepsilon_{jt} \\ \mu_{jt} = X_{jt}\bar{\theta} + \eta_{jt} \\ \begin{pmatrix} \xi_{jt} \\ \eta_{jt} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega\right) \end{cases}$$

Using Change-of-Variable Theorem, we derive the joint distribution of market share  $s_t$  and price  $P_t$  :

$$\pi(P_t, s_t | \bar{\theta}, \Sigma, \Omega, \delta) = \pi(\xi_t, \eta_t | \bar{\theta}, \Sigma, \Omega, \delta) J_{(\xi_t, \eta_t \rightarrow P_t, s_t)} = \pi(\xi_t, \eta_t | \bar{\theta}, \Sigma, \Omega, \delta) (J_{(P_t, s_t \rightarrow \xi_t, \eta_t)})^{-1}$$

where  $J_{(P_t, s_t \rightarrow \xi_t, \eta_t)} = \left\| \begin{array}{cc} \nabla_{\xi_t} P_t & \nabla_{\eta_t} P_t \\ \nabla_{\xi_t} s_t & \nabla_{\eta_t} s_t \end{array} \right\|$  is the Jacobian matrix  $J(s_t \rightarrow \eta_t) = \left\| \begin{array}{cc} \mathbf{I} & \mathbf{0} \\ \nabla_{\xi_t} s_t & \nabla_{\eta_t} s_t \end{array} \right\| = \|\nabla_{\eta_t} s_t\|$ .

Next, we write the likelihood in Equation (20):

$$\begin{aligned} L(\bar{\theta}, \Sigma, \delta, \Omega) &= \prod_{t=1}^T \pi(P_t, m s_t | \bar{\theta}, \Sigma, \Omega, \delta) \\ &= \prod_{t=1}^T \left\{ \left( \left\| \begin{array}{cc} \nabla_{\xi_t} P_t & \nabla_{\eta_t} P_t \\ \nabla_{\xi_t} s_t & \nabla_{\eta_t} s_t \end{array} \right\| \right)^{-1} \times \phi \left( \begin{array}{c} \xi_{jt} = P_{jt} - Z_{jt} \delta \\ \eta_{jt} = \mu_{jt} - X_{jt} \bar{\theta} \end{array} \middle| \begin{array}{c} 0 \\ 0 \end{array}, \Omega \right) \right\} \end{aligned} \quad (20)$$

where in the normal distribution  $\phi(\cdot)$ ,  $\eta_{jt}$  relies on mean utility  $\mu_{jt}$ , which is numerically computed by the contraction mapping method proposed by Berry et al. (1995).

Furthermore, to ensure that the estimated covariance variance-matrix  $\Sigma$  is positive-definite, following the re-parameterization method used in Jiang et al. (2009), we use Cholesky decomposition and write:

$$\Sigma = U'U; U = \begin{bmatrix} e^{r_{11}} & r_{12} & \cdots & r_{1K} \\ 0 & e^{r_{22}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_{K-1,K} \\ 0 & \cdots & 0 & e^{r_{KK}} \end{bmatrix}$$

Hence, instead draw a whole variance-covariance matrix directly in each MCMC iteration, we draw parameters  $r = \{r_{lk}\}_{l,k=1 \dots K, l \leq k}$ . We rewrite Equation (20)

$$L(\bar{\theta}, r, \delta, \Omega) = \pi(\bar{\theta}, r, \Omega, \delta) \prod_{t=1}^T \left\{ \left( \left\| \begin{array}{cc} \nabla_{\xi_t} P_t & \nabla_{\eta_t} P_t \\ \nabla_{\xi_t} s_t & \nabla_{\eta_t} s_t \end{array} \right\| \right)^{-1} \times \phi \left( \begin{array}{c} \xi_{jt} = P_{jt} - Z_{jt} \delta \\ \eta_{jt} = \mu_{jt} - X_{jt} \bar{\theta} \end{array} \middle| \begin{array}{c} 0 \\ 0 \end{array}, \Omega \right) \right\}$$

Lastly, given the priors on the parameters and likelihood function, the joint posterior distribution of the parameters is<sup>46</sup>:

---

<sup>46</sup> For the setup of hyper-parameters, we used diffuse priors. In appendix, we describe details on the choices of priors.

$$\begin{aligned}
\pi(\bar{\theta}, \Sigma, \Omega, \delta | \{P_t, s_t, X_t\}_{t=1}^T) &\propto L(\theta, r, \delta, \Omega) \times \pi(\bar{\theta}, r, \Omega, \delta) \\
&= \prod_{t=1}^T \left\{ \left( \begin{vmatrix} \nabla_{\xi_t} P_t & \nabla_{\eta_t} P_t \\ \nabla_{\xi_t} s_t & \nabla_{\eta_t} s_t \end{vmatrix} \right)^{-1} \times \phi \left( \begin{pmatrix} \xi_{jt} = P_{jt} - Z_{jt} \delta \\ \eta_{jt} = \mu_{jt} - X_{jt} \bar{\theta} \end{pmatrix} \middle| \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Omega \right) \right\} \\
&\times |V_{\bar{\theta}}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\bar{\theta} - \theta_0)' V_{\bar{\theta}}^{-1} (\bar{\theta} - \theta_0) \right\} \times \prod_{l=1}^K \exp \left\{ -\frac{(r_{ll})^2}{2\sigma_{r_{ll}}^2} \right\} \\
&\times \prod_{l=1}^{K-1} \prod_{k=l+1}^K \exp \left\{ -\frac{(r_{lk})^2}{2\sigma_{r_{off}}^2} \right\}
\end{aligned} \tag{21}$$

where  $\pi(\bar{\theta}, r, \Omega, \delta)$  is specified priors on the parameters. Specifically, for variance-covariance matrix, we specify the priors on  $r = \{r_{lk}\}_{l,k=1\dots K, l \leq k}$  with  $r_{ll} \sim N(0, \sigma_{r_{ll}}^2)$ ,  $r_{lk} \sim N(0, \sigma_{r_{off}}^2)$  for the diagonal, and off-diagonal elements in matrix  $U$ , respectively. For the population mean for characteristics coefficients  $\bar{\theta}$ , as written in Equation (18), we specify a multivariate normal distribution prior:  $\bar{\theta} \sim MVN(\bar{\theta}_0, V_{\bar{\theta}})$ . Furthermore, for the instrumental variables related parameters  $\delta$  and  $\Omega$ , we specify the following priors:

$$\begin{aligned}
\delta &\sim MVN(\bar{\delta}, V_{\delta}) \\
\Omega &\sim IW(v_0, V_{\Omega})
\end{aligned}$$

where  $IW$  indicates an inverse Wishart distribution.

### MCMC Estimation Steps

The MCMC estimation steps follows a strategy of Gibbs sampling combined Metropolis steps (Jiang et al. 2009, Rossi et al. 2005). Briefly speaking, in each iteration of the MCMC, we first use Gibbs Sampler to draw the conditionals of  $\bar{\theta}, \delta, \Omega | r, \{ms_t, P_t, W_t, Z_t\}_{t=1}^T, \bar{\theta}_0, V_{\bar{\theta}}, \bar{\delta}, V_{\delta}, v_0, V_{\Omega}$  in a sequence. Then, conditional on updated  $\{\bar{\theta}, \delta, \Omega\}$ , data  $\{ms_t, P_t, W_t, Z_t\}_{t=1}^T$ , and priors  $(\sigma_{r_{ll}}^2, \sigma_{r_{off}}^2)$ , we update the variance-covariance matrix,  $\Sigma$ , by making draws of  $r$  through a Random-Walk (RW) Metropolis chain. Specifically, we draw a proposal of  $r$ , given the accepted  $r$  in the previous iteration:  $r_{new} = r_{old} + MVN(\mathbf{0}, \sigma^2 D_r)$ , where  $\sigma^2$  is one of  $(\sigma_{r_{ll}}^2, \sigma_{r_{off}}^2)$  depending on whether we're drawing a diagonal or off-diagonal element of  $\Sigma$ .  $D_r$  is a candidate covariance matrix.  $r_{new}$  is either accepted or rejected, based on ratio computed using Equation (18). The intuition is that, if conditional on data, priors, and other parameters updated in the Gibbs sampling step,  $\Sigma_{new}(r_{new})$ , relative to  $\Sigma_{old}(r_{old})$ , is closer the true posterior of  $\Sigma$ , then we should have  $(\bar{\theta}, \Sigma_{new}, \Omega, \delta | \{P_t, ms_t, X_t\}_{t=1}^T) > (\bar{\theta}, \Sigma_{old}, \Omega, \delta | \{P_t, ms_t, X_t\}_{t=1}^T)$ . In appendix we provide detailed technical notes of our estimation steps.

## 2. Estimating Supply Model

Conditional on one's current state  $s = (s_{jt}, s_{-jt})$ , her wage and effort decisions can be described as sequentially solving a DP problem:

$$\{a_{jkt}\}_{t=0}^{\infty} = \underset{\{a_{jkt}\}_{t=0}^{\infty}}{\operatorname{argmax}} E_{\varepsilon_{jkt}} \left\{ \sum_{t=0}^{\infty} \cdot (\tilde{U}_{jkt}(a_{jkt} | s_{jt}, s_{-jt}) + \varepsilon_{jkt}) \right\} \quad (22)$$

where  $\tilde{U}_{jkt}$  is property  $j$ 's expected utility from choosing action  $k$  in period  $t$  and  $\varepsilon_{jkt}$  is the random shock associated to action  $k$  that is received before  $j$  makes a decision.

As discussed in section 3.8., in such dynamic game with many players, computing an MPE is computationally infeasible, hence we use OE to approximate MPE. In a OE, the individual's conditional choice probability is a function of her own state  $s_{jt}$  only. The set of states of her peers,  $s_{-jt}$ , is captured by tracking an average industry state  $\bar{s}_t$ , which reflects the distribution of the number of the reviews across the properties. It can be seen one's action and utility is influenced by her peers' state— $s_{-jt}$ , as it is the action of  $j$  and her peers and the subsequent state transitions that determine the average state in the next period. Then solving for an OE provides substantial computational advantage, as it converts a many-agent game problem into a problem similar to single-agent optimization, treating  $\bar{s}_t$  as a single state variable that is common across all individuals at time  $t$ . Thus, one can use any existing estimation method that can be applied to a single-agent discrete-choice dynamic programming (DDP) model to solve for an OE. Widely-used estimation strategy includes the nested fixed-point (NFXP) algorithm (Rust 1987) and conditional choice probability (CCP) based estimation (Hotz and Miller 1993, Aguirregabiria and Mira 2007).

In this paper, we use a Bayesian estimation strategy as this way we can flexibly incorporate individual heterogeneity—a key element in our model—in a hierarchical Bayesian framework (developed by Imai, Jain and Ching (2009), hereafter IJC). IJC algorithm allows estimating a heterogeneous model with a relatively low computational burden. In addition, it overcomes the problem of “curse of dimensionality”<sup>47</sup> when approximating the DP solution and avoids the complexity of searching for a global optimum in the space of the data likelihood function (IJC provides DP approximation that is comparable to state-of-the-art likelihood-based approaches, e.g., Keane and Wolpin (1994), Akerberg (2009). See Ching et al. (2012) for detailed discussions). The advantage of avoiding of searching in the parameter space, which usually requires the use of an optimization tool, is another reason we choose IJC algorithm. As we will discuss in section 4.4., Bayesian estimation approach can be easily combined with parallel computing and GPU

---

<sup>47</sup> The state space grows exponentially with the dimensionality of state variables, causing evaluating Bellman operator at every point in the state space infeasible.

computing techniques, without which it would be computationally infeasible given the large number of individuals and state space in our study.

### *IJC Algorithm*

We briefly introduce the logics and estimation procedure in IJC algorithm. In appendix we provide technical notes and details of implementing IJC.

IJC algorithm combines MCMC with DDP approximation, solving for the DP problem and making draws of structural parameters from the posterior distribution simultaneously. At each iteration  $m$  in the MCMC, IJC saves the simulated parameter vector  $\theta_{IJC}^{*m}$  and computes a corresponding pseudo-value function  $\tilde{W}^m(\theta_{IJC}^{*m})$ <sup>48</sup>. A total of the most recent  $N$  iterations of  $\{\theta_{IJC}^{*m}, \tilde{W}^m(\theta_{IJC}^{*m})\}$  are saved. When at new iteration  $m'$ , the simulated vector  $\theta_{IJC}^{*m'}$  is rejected or accepted by comparing the pseudo- posterior likelihood evaluated at the accepted parameters from the previous iteration,  $\theta_{IJC}^{*m'-1}$ , and at the proposed parameters at current iteration,  $\theta_{IJC}^{*m'}$ . When computing the pseudo-likelihood function, one needs to calculate the choice probability for each choice alternative. Recall that one solves for the DP problem by taking into account the value function (see Equation 19)), hence the likelihood function is also ‘pseudo-’ because the conditional choice probabilities are computed based on pseudo-value functions  $\{\tilde{W}^n(\theta_{IJC}^{*n})\}_{n=m'-N}^{n=m'-1}$  saved in the past  $N$  iterations. Specifically,  $\tilde{W}^{m'}(\theta_{IJC}^{*m'})$  is approximated by computing a (kernel-based) weighted average of the past  $N$  history draws of  $\{\tilde{W}^n(\theta_{IJC}^{*n}), \theta_{IJC}^{*n}\}_{n=m'-N}^{n=m'-1}$ , with the more weights attributed to history that have  $\theta_{IJC}^{*n}$  closer to current draw  $\theta_{IJC}^{*m'}$ . Hence, IJC algorithm is efficient in providing a full solution to the DP problem as it keeps the simulated parameter draws and computed pseudo-value functions to approximate the current pseudo-value function, with Bellman operator evaluated exactly once at each interaction. As Imai et al. (2009) proved, such an interactive steps of simulating parameter vector through the pseudo-Markov chain can effectively approach a steady state (after burn-in), where most of the structural parameters will be drawn from a distribution close to the true posterior distribution of the parameter vector.

---

<sup>48</sup> The pseudo-value function is obtained by applying the Bellman operator (i.e., solving for the value function) at the trial parameter vector. It is called ‘pseudo’ as the functions are evaluated at the simulated parameter vector not at the true parameter vector. Here the \* denotes that this is proposed parameter (regardless of whether it was accepted or rejected) at that iteration.

In summary, we a Gibbs Sample to sequentially simulate parameters of  $\left(\left\{\lambda_j^{effort^{*n}}, \lambda_j^{operate^{*n}}\right\}_{j=1}^J, \lambda, \Sigma_\lambda, \lambda^{MedImg}, \lambda^{HighImg}, \sigma_\varepsilon\right)$ , with  $\left(\lambda_j^{effort}, \lambda_j^{operate}\right) \sim MVN(\lambda, \Sigma_\lambda)$ . At each iteration  $m$ , we have the history of the drawn parameters and the associated pseudo-value functions:

$$\left\{ \begin{array}{l} \tilde{W}^n \left( \cdot; \left\{\lambda_j^{effort^{*m}}, \lambda_j^{operate^{*m}}\right\}_{j=1}^J, \lambda^{MedImg^{*m}}, \lambda^{HighImg^{*m}}, \sigma_\varepsilon^{*m} \right), \quad n=m-1 \\ \left\{ \lambda_j^{effort^{*m}}, \lambda_j^{operate^{*m}} \right\}_{j=1}^J, \lambda^{MedImg^{*n}}, \lambda^{HighImg^{*n}}, \sigma_\varepsilon^{*n}, \bar{s}^{*n} \quad n=m-N \end{array} \right\}$$

where  $\tilde{W}^n(\cdot;)$  indicates the pseudo-value functions at all possible state and  $\bar{s}^{*n}$  is proposed industry average state. Then each iteration  $m$  consists the following steps (we pre-specify of  $M$  as the total number of iterations):

- 1) given  $\left\{\lambda_j^{effort^{m-1}}, \lambda_j^{operate^{m-1}}\right\}_{j=1}^J$ , generate the conditional posterior mean  $\bar{\lambda}^m$  from a multivariate normal distribution and the posterior variance-covariance matrix  $\Sigma_\lambda^m$  from an inverse-gamma distribution
- 2) let  $\rho_j = (\lambda_j^{effort}, \lambda_j^{operate})$  indicate the individual-specific parameters, we want to make a draw from its posterior distribution

$$f(\rho_j | \rho, \Sigma_\rho, \lambda^{MedImg}, \lambda^{HighImg}, \sigma_\varepsilon, b_j, data_j) \sim g(\rho_j | \rho, \Sigma_\rho) \cdot L_j(b_j | \rho_j, \lambda^{MedImg}, \lambda^{HighImg}, \sigma_\varepsilon; data_j)$$

where  $g(\rho_j | \rho, \Sigma_\rho)$  indicates is the probability density function of  $\rho_j$  given the population mean  $\rho$  and variance-covariance matrix  $\Sigma_\rho$ .  $L_j(b_j | \cdot; data_j)$  is the likelihood for individual  $j$  with the tracked action and data (covariates) across time, i.e.,  $b_j = \{a_{jkt}\}_{t=1}^T$ ,  $data_j = \{X_{jt}, S_{jt}, S_{-jt}\}_{t=1}^T$ , evaluated at the trial parameter vector. Since this posterior distribution  $f(\rho_j | \cdot)$  does not a closed-form from which we can easily make a draw, we use Metropolis-Hasting to draw a proposal,  $\rho_j^{*m} = N(\rho_j^{m-1}, \Upsilon)$ , which will be evaluated to be either accepted or rejected, determined by a computed acceptance ratio:

$$accept_j^m = \min\left\{1, \frac{g(\rho_j^{*m} | \rho^m, \Sigma_\rho^m) \cdot \tilde{L}_j^m(b_j | \rho_j^{*m}, \lambda^{MedImg^{m-1}}, \lambda^{HighImg^{m-1}}, \sigma_\varepsilon^{m-1}; data_j) \cdot N(\rho_j^{m-1} | \rho_j^{*r}, \Upsilon)}{g(\rho_j^{m-1} | \rho^m, \Sigma_\rho^m) \cdot \tilde{L}_j^m(b_j | \rho_j^{m-1}, \lambda^{MedImg^{m-1}}, \lambda^{HighImg^{m-1}}, \sigma_\varepsilon^{m-1}; data_j) \cdot N(\rho_j^{*m} | \rho_j^{m-1}, \Upsilon)}\right\}$$

Let  $\gamma_c = (\lambda^{MedImg}, \lambda^{HighImg}, \sigma_\varepsilon)$  indicate the common cost vector, we evaluate the pseudo-likelihood  $\tilde{L}_j^m$  at the corresponding parameters and approximated pseud-value functions

$$\begin{aligned} \tilde{E}_j^m[W(s'; \rho_j^{*m}, \theta_c^{m-1})|s, a] & \quad (23) \\ &= \sum_{n=m-N}^{m-1} \tilde{W}^n(s'; \rho_j^{*n}, \theta_c^{*n}) \\ & \quad * \frac{K_h(\rho_j^{*n}, \rho_j^{*m})K_h(\gamma_c^{*n}, \gamma_c^{m-1})K_h(\bar{s}^{*n}, \bar{s}')}{\sum_{l=m-N}^{m-1} K_h(\rho_j^{*l}, \rho_j^{*m})K_h(\gamma_c^{*l}, \gamma_c^{m-1})K_h(\bar{s}^{*l}, \bar{s}')} \end{aligned}$$

where  $K_h(\cdot)$  represents a multivariate Gaussian Kernel with bandwidth  $h$ .  $s'$  and  $\bar{s}'$  indicates the individual's state and the industry average state in the next period, conditional on action  $a$ . Note that the likelihood function obtained by computing the choice probability. In a DDP with Type-I EV distribution of idiosyncratic shocks, the probability of choosing alternative  $k$  is:

$$prob_{jk} = \frac{\exp(\tilde{U}_j(s, k) + \beta \tilde{W}^n(s'; \rho_j^{*n}, \gamma_c^{*n}|s, k))}{\sum_{k'=0}^J \exp(\tilde{U}_j(s, k') + \beta \tilde{W}^n(s'; \rho_j^{*n}, \gamma_c^{*n}|s, k'))} \quad (24)$$

- 3) conditional on the accepted  $\rho_j$ , we draw the common cost parameters  $\gamma_c = (\lambda^{MedImg}, \lambda^{HighImg}, \sigma_\varepsilon)$ :  $\gamma_c^{*m} \sim N(\gamma_c^{m-1}, \Upsilon)$ . Here similar to step 2), we use Metropolis-Hasting method to evaluate the drawn parameters by approximating the averaged pseudo-value functions, computing likelihood functions evaluated at corresponding parameters, and then computing the acceptance ratio.
- 4) with the simulated parameters at  $m^{\text{th}}$  iteration, we now update and store the pseudo-value function:

$$\begin{aligned} & \tilde{W}^m \left( \cdot ; \left\{ \lambda_j^{effort^{*m}}, \lambda_j^{operate^{*m}} \right\}_{j=1}^J, \lambda^{MedImg^{*m}}, \lambda^{HighImg^{*m}}, \sigma_\varepsilon^{*m} \right) \\ &= \left\{ \ln \left[ \sum_{k'=0}^J \exp(\tilde{U}_j(s, k') \right. \right. \\ & \quad \left. \left. + \beta \tilde{W}^n \left( s'; \lambda_j^{effort^{*m}}, \lambda_j^{operate^{*m}}, \lambda^{MedImg^{*m}}, \lambda^{HighImg^{*m}}, \sigma_\varepsilon^{*m} | s, k' \right) \right] \right\}_{j=1}^J \end{aligned}$$

- 5)  $m=m+1$  (iteration continues with step 1)-4) repeated, till  $m>M$ ).

## Chapter 3

# Demand Interactions in Sharing Economies: Evidence from a Natural Experiment Involving Airbnb and Uber/Lyft

### 3.1. Introduction

Every day, nearly 1 million people rent accommodations from Airbnb, a service that offers more than 3 million rooms in 65,000 cities in 191 countries (Airbnb 2017). However, these accommodations are provided by individuals rather than a hotel chain. All this is facilitated by a sharing economy platform. Airbnb, Lyft, Uber, SnapGoods, and TaskRabbit are some of the most prominent examples of the emerging “sharing economy” in which people rent beds, cars, boats, and other assets directly from each other through internet coordination (Sundararajan 2016). The sharing economy has witnessed remarkable growth in the last few years and is projected to cross \$335 billion in global revenues by 2025 (PwC 2015). A 2016 McKinsey report estimated that between 20 and 30% of the United States and European Union workforce works in the sharing economy (Manyika et al. 2016). This emerging model is now big and disruptive enough that it has introduced many unforeseen challenges for consumers, incumbent businesses, regulators, and policy makers (Zervas et al 2017, Cramer and Krueger 2016, Burtch et al. 2017, Calo and Rosenblat 2017, Greenwood and Watal 2017, Edelman and Geradin 2016, Miller 2016, Zhang et al. 2017).

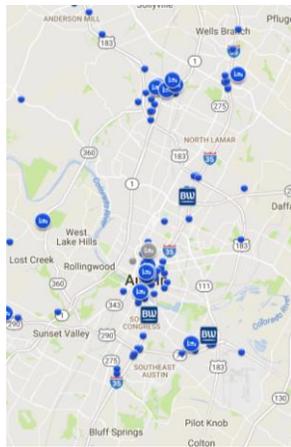
An emerging stream of research investigates the impacts of sharing economy platforms on the broader economy beyond the direct competitors (incumbent businesses). For example, studies have investigated the impact of Uber on local entrepreneurial activity (Burtch et al 2017), instances of drunk driving (Greenwood and Watal 2017), and sales of new cars in China (Gong et al. 2017). Similarly, researchers have investigated the impacts of Airbnb on apartment rental prices (Barron et al 2017), home values (Jefferson-Jones 2015), and neighborhoods (Levendis and Dicle 2016). However, the existing literature is silent on the interactions across sharing economies. In this work, we extend this stream of literature by developing an analysis of how Uber/Lyft affects Airbnb demand.

Faced with the disruptive nature of Airbnb and Uber/Lyft, regulators in many markets have introduced measures to supervise them. Such regulations include increasing taxes, imposing stricter conditions for participation in the sharing economy, introducing hefty fines for violations, or banning the service outright

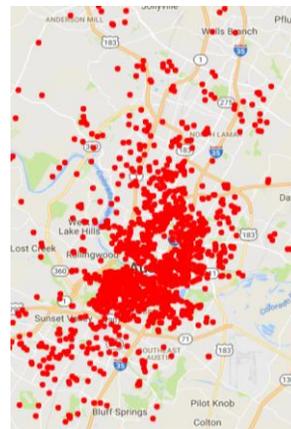
(Dobbins 2017, Manyika et al. 2016), which can shape the evolution of these platforms and their ability to penetrate markets. Thus, a deeper understanding of the impact of Uber/Lyft on Airbnb and vice versa is particularly important in view of changing regulations. Further, such an understanding may also potentially allow the platforms to find ways to leverage each other for future growth.

It may not be obvious a priori why Uber/Lyft may affect Airbnb demand. While Uber/Lyft makes travel easy, accessible, and often cheaper than a taxi, they do not discriminate between Airbnb and its direct competitor, hotels. Both Airbnb and hotel guests could benefit from Uber/Lyft. However, Airbnb differs from hotels in a subtle way. Hotels typically locate around travelers' main activities because easy access to transportation is a key driver in customers' lodging choices (Ellinger 1977, Wyckoff and Sasser 1981, Hotel Guest Satisfaction Survey 2015). Hence, hotels often geographically cluster in areas with convenient transportation (as shown in Figure 1a). In contrast, Airbnb properties (as shown in Figure 1b) are located all over a city. Particularly, 74% of Airbnb properties are located outside main hotel districts and potentially suffer from relatively poor transportation<sup>49</sup>. In this scenario, demand complementarities arise as Uber/Lyft, with their ease of access and price advantage relative to taxis<sup>50</sup>, may play a key role in alleviating a locational disadvantage of Airbnb properties (Picchi 2016).

**Figure 1 A Comparison of the Distributions of Hotels and Airbnb**



A map of the hotel distribution in Austin



A map of Airbnb properties' distribution in Austin

<sup>49</sup>74% of Airbnb properties are outside main hotel districts. <https://www.airbnb.com/about/about-us>.

<sup>50</sup>The price advantages of Uber relative to taxis depend on vehicle options, commute conditions, and cities. On average, UberX costs approximately 48%~67% of the average taxi fare. <http://www.businessinsider.com/uber-vs-taxi-pricing-by-city-2014-10>.

However, a key question is whether Uber/Lyft increases the overall demand for Airbnb properties or just simply redistribute the demand across Airbnb properties. On the one hand, the absence of Uber/Lyft may lead to *an overall decrease in demand* for Airbnb properties. As most of the Airbnb properties are in areas with poor access to transportation, travelers may not want to stay there when transportation costs increase in the absence of Uber/Lyft. Consequently, these properties will lose demand and travelers will switch to areas with good access to transportation, such as hotels in downtown locations. On the other hand, the exit of Uber/Lyft may lead to a *redistribution of demand* across Airbnb properties by shifting demand from Airbnb units with poor transportation access to units with good transportation access with the overall demand unchanged<sup>51</sup>. For example, if a traveler views Airbnb and hotels as horizontally differentiated and prefers Airbnb, then he/she would choose Airbnb properties in areas with good transportation when the transportation costs are high. Thus, it has been a main marketing effort of Airbnb to provide its customers with unique experiences that they would not get from hotels<sup>52</sup>. As a result, the exit of Uber/Lyft may simply redistribute the demand among Airbnb properties.

In this paper, we address this question empirically by 1) measuring the direction and quantifying the demand interactions between Airbnb and Uber/Lyft and 2) further investigating the mechanism behind the demand interactions. One obstacle to answering this question is the fact that causal inference is particularly difficult in this case due to endogeneity concerns. We leverage a natural experiment induced by the exit of Uber/Lyft from Austin on May 9, 2016 to estimate the causal effect of Uber/Lyft on Airbnb demand. The exit of Uber/Lyft was in response to regulations passed by the Austin city council that required a number of other changes, such as fingerprint-based background checks for Uber and Lyft drivers. Uber and Lyft claimed that these regulations made it harder for them to sign up drivers and, as a result, made it costly for them to operate in Austin.

The exit of Uber and Lyft, although not directly related to Airbnb, indirectly affected the Airbnb's demand because it introduced a significant *increase* in the transportation costs after May 2016 in Austin. By leveraging the Difference-in-Differences (DiD) methodology (with the "treatment" group composed of properties in Austin) based on a 9-month longitudinal panel dataset spanning 7,300 Airbnb properties across 7 U.S. cities, we are able to quantify the effects of Uber/Lyft's exit on Airbnb property demand and implement a set of robustness analyses.

---

<sup>51</sup> Our data shows the average occupancy rate at Airbnb is below 30%. Therefore, there is enough capacity for Airbnb properties to "absorb" the extra demand.

<sup>52</sup> <https://www.airbnb.com/livethere>.

Overall, we find that the exit of Uber/Lyft led to a drop of approximately 9.6% in Airbnb property demand, which is equivalent to a decrease of \$6,482 in the annual revenue for the average unit host. We present evidence to validate the assumption of the “parallel trend” in DiD analysis. The results are robust to an extensive set of robustness analyses, including balance checks on covariates, alternative model specifications (e.g., inclusion of additional time-varying control variables), model-independent matching estimations, robustness checks on the omitted variable test (e.g., the effect did not begin prior to the natural experiment), falsification checks on potential concerns regarding seasonality, and the random (shuffled) treatment test.

To further identify the nature of the effect of Uber/Lyft’s exit, we extend our analysis by decomposing the effect across Airbnb properties by their access to transportation and their luxuriousness. We obtain a consistent finding that Uber/Lyft’s exit results in a significant decrease in Airbnb property demand. Additionally, we find differential effects of Uber/Lyft’s exit across Airbnb properties. The non-uniform impact offers insights into the underlying mechanism of the effect of Uber/Lyft’s exit. *First*, our results suggest that the exit of Uber/Lyft from Austin mainly reduced the demand of properties in areas with poor access to public transportation. For example, for Airbnb properties in areas with poor access to transportation, the negative effect of Uber/Lyft’s exit on demand is 140% of the average effect due to the lack of substitutes for Uber/Lyft in these areas. This finding corroborates that Uber/Lyft plays a role in affecting Airbnb property demand through transportation costs. Further, Airbnb properties in areas with excellent access to public transportation will be (on average) booked 3.4% more frequently due to the extra demand flow from areas that lose demand. However, the extra demand was largely absorbed by other lodging alternatives (e.g., hotels) in the same areas, suggesting an insufficient differentiation in consumer’s perception of Airbnb and hotels.

*Second*, we find that low-end properties experience a greater negative impact on demand in the absence of Uber/Lyft. Particularly, in the areas with poor access to public transportation, low-end properties encountered 2.17 times the negative effect for high-end properties in the same neighborhoods. The differential effect arises because a surge in the transportation costs impacts the consumers of low-end properties (i.e., price-sensitive travelers) the most. When transportation costs increase, the more price-sensitive consumers are likely to prefer areas with low-price transportation services. *Third* and finally, we find a non-uniform impact of Uber/Lyft’s exit on properties with geographic differences, with properties farther away from downtown losing more demand. Specifically, every unit increase in commute (driving) time (in minutes) would lead to a decrease of approximately 0.21% in property booking. The results suggest that the exit of Uber/Lyft leads to a decrease in the geographic demand dispersion across the city and an increase in the demand concentration near hotel districts.

Our study makes the following three contributions. 1) The emerging literature has investigated the effects of various sharing economy platforms on incumbent industries and provided evidence that relates the sharing economy's unique advantages/disadvantages to these effects. However, the existing literature has not examined the interactions across sharing economies. We contribute to this literature by empirically quantifying the effect of Uber/Lyft on Airbnb demand while addressing the endogeneity issue in the relationship between transportation costs and Airbnb demand. 2) Our work provides implications for Airbnb and reveals how customers perceive and use Airbnb compared to hotels. Our quantification of the marginal effect of transportation costs on Airbnb host revenue suggests that, in the absence of Uber/Lyft, Airbnb would lose significant demand. We also find that Airbnb demand is geographically concentrated in the main hotel districts, with the properties farther from downtown losing the most demand. However, Airbnb's main marketing effort of "live like a local" is based on the fact that the majority of Airbnb properties are located far from conventional hotel districts and offer travelers a local experience that one would not get from hotels. Hence, our findings provide insights into a flipside of the key Airbnb features. In other words, the properties located in the main hotel districts, though they are less spatially differentiated from hotels and provide a less unique experience, are less vulnerable to demand shocks, such as the exit of Uber/Lyft. Furthermore, our work also gauges Airbnb's main marketing claims of horizontal differentiation from hotels. The results of significant losses in the overall property demand suggest that consumers' perception of Airbnb is not sufficiently (horizontally) differentiated from hotels to retain them within Airbnb under increasing transportation costs. 3) Finally, this paper provides certain policy implications. Despite the present debate on the negative externalities of sharing economies and the regulations against sharing economies (Malhotra and Van Alstyne 2014), our findings suggest that sharing economies may create positive externalities. As a result, consumer welfare may be affected in one economy when another economy is being regulated. Our results also indicate that sharing economies such as Uber/Lyft and Airbnb may benefit from mutual collaboration.

### **3.2. Literature Review**

Our work is related to the relatively new stream of studies on sharing economies. The impact of Airbnb on incumbent lodging industry, i.e., hotels, has been studied to some extent. Zervas et al. (2017) analyzed the impact of Airbnb's entry on hotels. The authors find an approximately 0.4% decrease in hotel room revenue in response to a 10% increase in Airbnb's market size. A few studies focus on Airbnb hosts from the supply side. In a recent paper, Edelman et al. (2017) found that discrimination occurs among Airbnb hosts, with the booking requests from guests with African American names being 16% less likely to be accepted compared to identical guests with white names. Examining listing prices and availability data, Li et al (2016) found that professional Airbnb hosts significantly outperform nonprofessional hosts. The authors further demonstrate that the gap can be partly explained by the pricing inefficiencies of nonprofessional hosts.

Other work has examined the platform design aspects, including the reputation system and the provided photography program. The existing work on Airbnb has examined various aspects of Airbnb that may affect its performance or its impacts on competitors. However, prior works have largely ignored an essential factor—the geographic location—of Airbnb properties, which is a key driver in guests' lodging choices and highlights a crucial difference between Airbnb and its competitors (hotels).

This paper is also related to the studies on the ride-sharing economy with respect to Uber and Lyft. The internet-based mobile technology has been shown to enable a higher utilization rate of ride-sharing services compared to taxi-cab services (Cramer and Krueger 2016). Using data on alcohol-related motor vehicle fatalities in California, Greenwood and Wattal (2017) suggest that the increased availability and the decreased cost of Uber significantly reduced the rate of fatalities. Prior work has focused on the supply side and studied the impact of Uber's entry on durable good purchases and entrepreneurial activity. Specifically, Gong et al. (2017), using data on new vehicle registrations in China, find that Uber is associated with an increase in car purchases due to the value enhancement effect. Burtch et al. (2017) find that Uber leads to a significant decrease in the activity of entrepreneurs by offering them stable employment. Prior literature has suggested that ride-sharing platforms can play a role through their improved efficiency (Cramer and Krueger 2016), increased availability (Greenwood and Wattal 2017), employment flexibility (Burtch et al. 2017), and enhanced value of goods (Gong et al. 2017). However, the prior work did not consider the effect of Uber/Lyft on changing consumers' perception of spatial distances. Specifically, Uber and Lyft change the relative spatial advantage/disadvantage by providing a cost-efficient option to move passengers between locations of their choice.

In this work, we consider how Uber/Lyft may influence travelers' choices of lodgings. As part of the travel industry, Uber/Lyft conveys travelers from their lodging places to local destinations. Plausibly, Uber and Lyft may impact the spatial competition across lodging options and thus impact the distribution of lodging demand. However, it remains unclear whether Uber and Lyft make average Airbnb properties and hotels more substitutable to each other by reducing location advantages of hotels relative to Airbnb properties. However, it may be that ride sharing platforms redistribute the demand distribution within Airbnb properties by changing the spatial differentiations across Airbnb properties. Our study fills this gap in the literature.

### **3.3. Research Context and Empirical Framework**

Our research context is Airbnb and Uber/Lyft. Airbnb is a home-sharing platform for users to list and discover lodgings. Airbnb has more than 3 million listings in over 65,000 cities and was recently valued at approximately \$31 billion. Airbnb hosts list their spare rooms on Airbnb.com and the guests choose properties that are available to be booked in the requested dates. For each transaction, Airbnb charges a

9~12% service fee from the guest and 3% from the host. Uber and Lyft are ride-sharing platforms that assign a registered owner-operator's vehicle to a user, which provides transportation to the user's intended destination. Uber is available in more than 600 cities worldwide and was recently valued at over \$6.5 billion. The trip fares vary across trips, vehicles, and locations, but in general they are much lower than what a taxi would cost. Uber takes 25% of the trip fare as a commission, and the driver takes the remaining 75%.

Our data on Airbnb properties include property bookings, property characteristics, and host characteristics. Additionally, based on each property's address, we collect information on its access to public transportation from Walkscore.com. The data involves the natural experiment of Uber/Lyft's exit of Austin in May 2016, which introduced an exogenous increase in the transportation costs in Austin.

### **3.3.1. Observational Data on Airbnb**

The panel data spans 9 months (January 2016 to September 2016) for 7,300 Airbnb properties in 7 cities in the United States, including Austin, Boston, Los Angeles, New York, San Diego, San Francisco, and Seattle. Hence, the panel is balanced with an equal length time window before and after the natural experiment. The data for each property consists of three parts that we describe below.

#### **Host Information**

For each host of the property in our data set, we collect information from the host page. The information includes whether the host has a verified Airbnb account, when the host joined Airbnb.com, and the total number of Airbnb properties he/she hosts.

#### **Property Time-Invariant Information**

We collect property characteristics that do not change over time: 1) property location (city, neighborhood name, zip code, and street address), 2) property capacity (number of bathrooms/bedrooms/beds and the maximum number of accommodated guests), 3) property type (e.g., house, apartment, flat), and 4) room type (entire place or shared place).

#### **Property Time-Varying Information**

In each month for each property in our sample, we collect property information that may change over time: 1) guest review data, including the number of accumulated guest reviews and the average review score, and 2) property's daily information (property booking and property daily price). We detail this below:

##### *Property Daily Information*

Our listing-level property booking data is obtained from a third-party company that specializes in collecting Airbnb data. The booking data includes (for each property on each day) whether the property is available (i.e., the property was available to be booked), unavailable (i.e., the property was booked/reserved by a

guest), or blocked (i.e., the property was marked as ‘unavailable’ by the host, without a real booking). We also have for each property on each day the property’s daily price and availability.

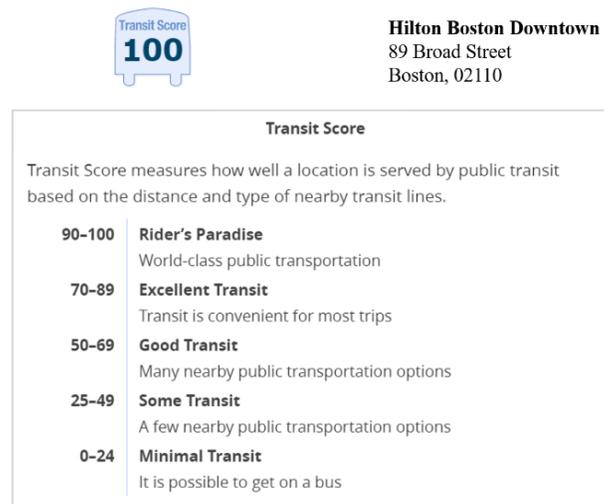
### 3.3.2. Access to Public Transportation: Walkscore.com

For each property, we collect information on its access to the public transportation, which is a key driver in guests’ lodging choices and a key difference between hotels and Airbnb properties. To capture the variation in transportation costs across Airbnb properties, we collect data from Walkscore.com, which (when given any address) provides real estate-related information regarding the nearby areas of the address. The most well-known feature provided by Walkscore.com is the transit score.

#### Transit Score

A numeric index (0–100) is assigned to a given address based on how well the address is served by public transportation (e.g., bus and subway). The algorithm awards the frequency of each transit route and penalizes the distance from the address to the nearest stop of the route (Hirsch et al. 2013). Figure 2 presents the transit score along with the associated address for a hotel (Hilton Boston Downtown) as an example.

**Figure 2 An Example of Transit Score for a Hotel in Downtown Boston**



### 3.3.3. Natural Experiment—Uber/Lyft’s Austin Exit

A unique feature of our data is a natural experiment—Austin voters rejected Austin’s Proposition 1, which would have replaced existing ordinances that required drivers of ride-sharing companies to undergo fingerprint background checks. Uber and Lyft refused to take fingerprint background checks and shut down

their services in Austin on 9 May 2016, following the failure of Proposition 1<sup>53</sup>. Due to the significant price advantages of Uber/Lyft relative to taxis, the transportation costs in Austin increased after the natural experiment (Picchi 2016). In contrast, the transportation costs in the other six cities in our sample stayed unchanged. Exploring the variation in property demands before and after the natural experiment across the different cities, we can identify the sharing effects of Uber/Lyft on Airbnb.

### 3.3.4. Definitions of Key Variables and Descriptive Statistics

#### Treatment and Control Group

The treatment is the increase in transportation costs caused by Uber/Lyft's exit. Airbnb properties in Austin form the treatment group and other properties form the control group. After removing missing data points, we were left with 880 properties in the treatment group and 6,420 properties in the control group. Let  $AUSTIN_i$  equal 1 (0) if property  $i$  is in Austin (in other 6 cities), and  $\{AFTER_t\}_{t=January}^{t=September}$  equal 1 (0) if period  $t$  is after (before) the Uber/Lyft's exit. Then, the treatment status indicator  $AUSTIN_i \cdot AFTER_t$ , equals 1 when  $t > May$  and  $i$  is in Austin.

#### Property Demand

We measure the demand of a property from its daily booking. It contains the three possible statuses of booked, available, and blocked. Since a property being 'blocked' does not reflect the real demand, we use only the days with bookings (i.e., reservations) to accurately account for the property demand (our dependent variable). Specifically, for each property  $i$  in each period  $t$ , property demand is measured as the ratio of the number of days a property is booked in that period. We further scale the ratio by 100. For example, if a property was booked (unavailable) for 10 days in April, then its demand in April is (10 days/30 days) \*100=33.3.

#### Property Price and its Instrument

For each property  $i$  in period  $t$ , we compute the mean property price,  $PRICE_{it}$ , by the averaging daily prices over the days in each period. Thus, the property's price in the current period are endogenous because of its correlation with random shocks that also affects the property's current demand. Following the previous literature (Villas-Boas and Winer 1999), we use the lagged price  $PRICE_{it-1}$ , which is unlikely to be correlated with the random shock period  $t$ , as a valid instrument variable for  $PRICE_{it}$ , to address the endogeneity concern.

---

<sup>53</sup> <https://www.wsj.com/articles/texas-capital-city-votes-to-keep-fingerprinting-for-uber-lyft-drivers-1462796972>.

Table 1 presents the summary statistics for the key variables at the group level. Above each variable is its short description. We report statistics for the pre-treatment period (April, the period before Uber/Lyft’s exit) and for the post-treatment period (June, the period after Uber/Lyft’s exit). Table 1 shows that the *DEMAND* for untreated properties increased from April (27.48) to June (30.52). However, it decreased from April (15.61) to June (11.30) for treated units. The different trends in *DEMAND* between the two groups suggest a negative treatment effect on the demand for properties in Austin. However, as shown in Table 1, the treated and untreated units are not comparable over the pre-treatment covariates. An imbalanced sample may lead to a violation of the critical assumption of “parallel trends” in Difference-in-Difference (DiD) analysis. In the next section, we discuss the DiD analysis and describe how we use the Propensity Score Weighting (PSW) method to address the concern of data imbalance.

**Table 1 Summary Statistics of Airbnb Properties**

	<i>Variables</i>	<b>Treated Units</b>		<b>Untreated Units</b>	
		(880 properties)		(6,420 properties)	
		<b>Mean</b>	<b>Std. Dev.</b>	<b>Mean</b>	<b>Std. Dev.</b>
	<i>DEMAND 100*the portion of days in a month a property was occupied</i>				
	<i>DEMAND (Scaled by 100)</i>	15.61	26.058	27.48	29.968
	<i>REVIEW_COUNT the number of accumulated guest reviews</i>				
<b>Pre-Treatment</b>	<i>REVIEW_COUNT</i>	9.89	21.117	20.38	31.773
(April)	<i>REVIEW_SCORE the overall review score rated by guests</i>				
	<i>REVIEW_SCORE</i>	58.87	46.447	79.11	33.797
	<i>PROPERTY_PRICE the average daily price of a property</i>				

	<i>PROPERTY_PRICE</i>	294.26	429.118	163.73	221.636
	<i>DEMAND</i>	11.30	23.588	30.52	31.720
<b>Post-Treatment</b>	<i>REVIEW_COUNT</i>	11.83	23.815	22.95	34.931
(June)	<i>REVIEW_SCORE</i>	60.83	45.955	79.81	34.834
	<i>PROPERTY_PRICE</i>	292.77	376.349	170.43	229.320
	<i>ACCOMMODATES the maximum number of guests that a property can accommodate</i>				
	<i>ACCOMMODATES</i>	4.32	2.645	2.98	1.952
	<i>BATHROOMS the number of bathrooms</i>				
	<i>BATHROOMS</i>	4.06	2.111	3.62	2.025
	<i>BEDROOMS the number of bedrooms</i>				
	<i>BEDROOMS</i>	3.11	1.507	2.47	1.261
<b>Time-invariant</b>	<i>BEDS the number of beds</i>				
	<i>BEDS</i>	2.12	1.567	1.61	1.153
	<i>PROPERTY_TYPE categorical variable=1, 2... for different property types such as: apartment, house, condo, flat, etc.</i>				
	<i>PROPERTY_TYPE</i>	7.55	4.854	4.02	4.699
	<i>ROOM_TYPE 1 (0) if room is an entire home (private room or shared place)</i>				
	<i>ROOM_TYPE</i>	0.68	0.467	0.57	0.495
	<i>TRANSIT_SCORE 0~100 rating nearby public transportation (by walkscore.com)</i>				
	<i>TRANSIT_SCORE</i>	43.81	13.881	78.77	22.343

### 3.4. Empirical Strategy and Results

#### 3.4.1. Difference-in-Difference (DiD) Model and Weighting Method

The Difference-in-Differences (DiD) approach (Heckman et al. 1997) is a widely applied strategy for evaluating the effect of an intervention or treatment (e.g., Uber/Lyft's exit) on an outcome variable of interest (e.g., Airbnb property demand). We exploit the natural experiment, which occurred only in Austin,

to estimate the treatment effect by comparing the difference in the changes before and after the natural experiment in Airbnb property demand between the treatment group (i.e., Austin) and the control group (i.e., the other six cities). Since Airbnb properties across cities are likely to be different (as shown in Table 1), we properly weigh our sample to construct comparable treatment and control groups.

### Propensity Score Weighting (PSW) Method

We compute the sample weights based on propensity score, which is the probability of receiving a treatment (Rosenbaum and Rubin 1983). True propensity scores are often unknown. In practice, we approximate unit  $i$ 's propensity score  $ps_i$  as a logistic function of a vector of variables  $X_i$  such that  $ps_i = g(X_i\beta)$ . Here,  $X_i$  includes the observed covariates presented in Table 1 and their higher order terms<sup>54</sup>.  $\beta$  is estimated by maximizing sample likelihood of treatment assignments. We compute sampling weights for each unit  $i$  with the Inverse Probability of Treatment Weighting (Austin and Stuart 2015):

$$\omega_i(\beta, \mathbf{X}_i) = \frac{T}{ps_i(\mathbf{X}_i, \beta)} I(\text{if } i \in \text{treatment group}) + \frac{1 - T}{1 - ps_i(\mathbf{X}_i, \beta)} I(\text{if } i \in \text{control group})$$

### DiD Model Specification and Results

We model  $DEMAND_{it}$ , the demand for Airbnb unit  $i$  in period  $t$ , using the following linear specification:

$$DEMAND_{it} = INTERCEPT + \alpha_1 AUSTIN_i + \alpha_2 AFTER_t + \alpha_3 (AUSTIN_i \cdot AFTER_t) + \varepsilon_{it}$$

where  $AUSTIN_i$  is 1 (0) if property  $i$  is in Austin (in the other six cities).  $AFTER_t$  is 1 (0) if period  $t$  is after (before) Uber/Lyft's exit.  $\varepsilon_{it}$  is an i.i.d. (normally distributed) random shock to  $DEMAND_{it}$ . To further account for the time-invariant factors that are specific to property and may affect property demand (e.g., property location), we incorporate a property fixed effect term,  $PROPERTY_i$ . Also included are time fixed effects,  $PERIOD_t$ , which capture seasonality in the trends of demand.  $AUSTIN_i$  and  $AFTER_t$  are absorbed by  $PROPERTY_i$  and  $PERIOD_t$ , respectively. Hence, we estimate the following demand model with a Weighted Least Squares regression combining PWS method:

$$DEMAND_{it} = INTERCEPT + \alpha_3 (AUSTIN_i \cdot AFTER_t) + PERIOD_t + PROPERTY_i + \varepsilon_{it} \quad (25)$$

where  $AUSTIN_i \cdot AFTER_t$  is a dummy variable indicating whether property  $i$  received the treatment of "Uber/Lyft's exit" in the period (1 for yes, and 0 for no). The key coefficient,  $\alpha_3$ , approximates the impact

---

<sup>54</sup>We started with linear terms of covariates and implemented a balance check. Interactions and higher order (square) terms are added for those variables on which the two groups were imbalanced. Then, we estimated propensity scores under new specification of  $X_i$  and implement a balance check. The steps are repeated until the samples are balanced.

of Uber/Lyft’s exit on the property demand of Airbnb properties. Note that, as a baseline DiD model for our main analysis, Equation (1) does not include any time-varying covariates. In the empirical extension analyses, we use the extended model incorporating the additional controls of guests’ reviews and property prices. As we will discuss later, our main results are consistent when including the additional controls.

Table 2 reports the estimation results of Equation (1). As seen, the coefficient of the key variable *AUSTIN · AFTER* is negative, suggesting that Uber/Lyft’s exit reduces the overall demand for Airbnb properties. Specifically, in the absence of Uber/Lyft, the Airbnb properties are on average 9.6% less frequently booked. The resulting effect is a decrease of (9.6%\*365 days) \*185 USD/day =6,482 USD in the annual revenue to the host of an average unit<sup>55</sup>.

**Table 2 Impact of Uber/Lyft’s Exit on Demand: DiD Model**

<b>VARIABLES</b>	<b>ESTIMATES</b>
<i>AUSTIN · AFTER</i>	-9.575*** (-9.57)
<i>INTERCEPT</i>	22.919*** (41.31)
Fixed Effect	Yes
Seasonality	Monthly
Num. Observations	67451
R-squared	0.699

The t statistics are in parentheses. \* p<0.05, \*\* p<0.01, and \*\*\* p<0.001.

### 3.4.2 Parallel Pre-Treatment Trends

The validity of the DiD approach (Equation (1)) relies on a critical assumption of pretreatment parallel trends. That is, the (weighted or matched) two groups should have parallel trends in their demands before the treatment (Angrist and Pischke 2008, Bertrand et al. 2004). In practice, there are two options of validating the parallel trend, which we discuss below.

#### Trends in Dependent Variable for Weighted Samples

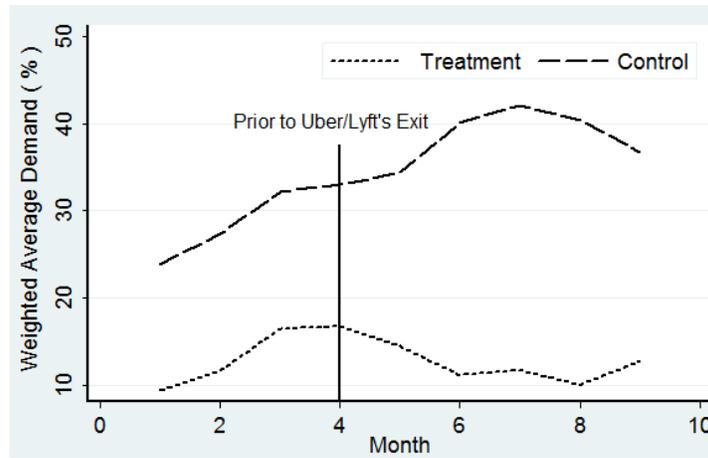
We want to show that after weighting/matching the treated and untreated units, there is no difference in the pre-treatment trends in the property demand between the groups (Athey and Imbens 2006). If the weighting

<sup>55</sup> The average property daily price in pre-treatment period (i.e., April) in our sample is 185 USD.

strategy properly balances the two groups over the covariates (which we further validate in the Robustness Checks section), their pre-treatment trends should be comparable.

In Figure 3, we plot the average demand for weighted samples (vertical) over the periods (horizontal). The plots confirm that the property demand of the two (weighted) groups followed the same trend until the intervention of Uber/Lyft's Austin exit. The plots further suggest that, compared to the properties in the control group (dash), the properties in Austin (dotted) received a negative treatment effect on their demand after Uber/Lyft's exit.

**Figure 3 Parallel Trends in Property Demand in the Weighted Sample**



### Relative Time Model

Another method of validating the parallel pre-treatment trends assumption is the relative time model with the inclusion of the leads and lags in periods (Autor 2003). Following the extant literature (Agrawal and Goldfarb 2008), we implement the model by adding a series of period dummy variables within a time window prior to the treatment. The coefficients of the period lags will test whether the estimated treatment effect began prior to the exit of Uber/Lyft. Specifically, we implement the relative time model with the following two alternative specifications:

$$\begin{aligned}
 DEMAND_{it} = & INTERCEPT + \alpha_3(AUSTIN_i \cdot AFTER_t) + \sum_j \beta_j(PRE_{it}(j) \cdot AUSTIN_i) \quad (26) \\
 & + PERIOD_t + PROPERTY_i + \varepsilon_{it}
 \end{aligned}$$

$$\begin{aligned}
 DEMAND_{it} = & INTERCEPT + \sum_j \beta_j(PRE_{it}(j) \cdot AUSTIN_i) + \sum_k \beta_k(POST_{it}(k) \cdot AUSTIN_i) \quad (27) \\
 & + PERIOD_t + PROPERTY_i + \varepsilon_{it}
 \end{aligned}$$

where the newly added terms  $\sum_j \beta_j (PRE_{it}(j) \cdot AUSTIN_i)$  allow us to examine possible false significant treatment effects prior to the treatment.  $PRE_{it}(j)$  is an indicator function that equals 1 if the period  $t$  is  $j$  months prior to the natural experiment. Similarly, item  $POST_{it}(k)$  in Equation (27) is an indicator function that equals 1 if the period  $t$  is  $k$  months after the month of Uber/Lyft's exit. Hence, the coefficient  $\beta_j$  for  $j=-J, -J-1, \dots, -1, 0$  captures the pre-treatment trend of the impact of Uber/Lyft on Airbnb property demand. If  $\beta_j$  is negative and significant, then it indicates that the trend of decreasing demand in Austin, relative to the control group, already existed prior to the exit of Uber/Lyft, suggesting false significance. The coefficient  $\beta_k$  for  $k=1, 2, \dots, K$  captures the effect of Uber/Lyft's exit in each post-treatment period.

**Table 3 Relative Time Model of the Effect of Uber/Lyft's Exit on Airbnb Property Demand**

VARIABLES	Estimates	
	Equation (2)	Equation (3)
<i>AUSTIN · AFTER</i>	-9.187*** (-6.25)	
<i>PRE_TREATMENT (-4)</i>	2.632 (1.28)	2.632 (1.28)
<i>PRE_TREATMENT (-3)</i>	4.243* (2.01)	4.243* (2.01)
<i>PRE_TREATMENT (-2)</i>	-0.278 (-0.12)	-0.278 (-0.12)
<i>PRE_TREATMENT (-1)</i>	omitted	omitted
<i>TREATMENT_MONTH (0)</i>	-4.72*** (-2.11)	-4.72*** (-2.11)
<i>POST_TREATMENT (1)</i>		-7.718*** (-4.16)
<i>POST_TREATMENT (2)</i>		-10.784*** (-6.04)
<i>POST_TREATMENT (3)</i>		-12.541*** (-6.78)
<i>POST_TREATMENT (4)</i>		-5.477** (-2.86)
<i>INTERCEPT</i>	22.039*** (44.17)	22.04*** (44.16)

Fixed Effect	Yes	Yes
Seasonality	Monthly	Monthly
Num. Observations	67451	67451
R-squared	0.700	0.702

The t statistics are in parentheses. \* p<0.05, \*\* p<0.01, and \*\*\* p<0.001.

Following prior work (Agrawal and Goldfarb 2008), we set the period prior to the month of Uber/Lyft’s exit (i.e., April) as the reference period (by normalizing its coefficient to zero) and consider a three-period interval prior to the reference period for better interpretability. Table 3 reports the results from estimating equations (2) and (3). The coefficients of the pre-treatment indicators are either statistically insignificant or positive. This suggests that 1) there was no pre-existing trend towards a decrease in the demand for properties in Austin relative to the control group, and 2) the estimated impact of Uber/Lyft’s exit (presented in Table 2) was not due to a false impact that began prior to the natural experiment. Furthermore, the estimated coefficient of the treatment indicator  $AUSTIN \cdot AFTER$  remains negative and significant, which is consistent with the finding of the negative impact of Uber/Lyft’s exit.

### 3.5. Robustness Checks

We implement an extensive set of analyses to verify the robustness of our main results. We begin with the validation on our PSW weighting strategy. It is followed by the free-form matching estimation, the seasonality examination (using prior years’ data), the random (shuffled) treatment test, the inclusion of additional control variables, and the exclusion of alternative explanations.

#### 3.5.1. Validating PSW Strategy

This is a critical step of constructing comparable treatment and control groups in the weighting strategy. To validate the PSW strategy, we implement balance checks through the standardized difference in means (Rubin 2001, Stuart 2010) that compares, over the  $M$ -dimensional covariates, the weighted means of the treatment group,  $\bar{X}_{treatment} = \frac{\sum_{i \in treatment} \omega_i X_i}{\sum_{i \in treatment} \omega_i}$ , and the control group,  $\bar{X}_{control} = \frac{\sum_{i \in control} \omega_i X_i}{\sum_{i \in control} \omega_i}$ , which are weighted based on their propensity score weighting  $\omega_i$ . Then, for variable  $X^m$  ( $m=1, 2 \dots M$ ), we use the absolute difference in the means, normalized by weighted sample variance,  $s_{treatment}^2$  and  $s_{control}^2$ :

$$d^m = \frac{\left| \bar{X}_{treatment}^m - \bar{X}_{control}^m \right|}{\sqrt{\frac{s_{treatment}^2 + s_{control}^2}{2}}}$$

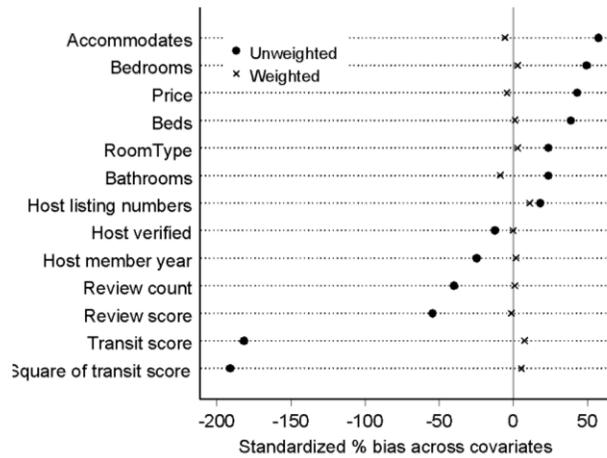
where the weighted sample variances are:

$$s_{treatment}^2 = \frac{\sum_i \omega_i}{(\sum_i \omega_i)^2 - \sum_i (\omega_i)^2} \sum_i \omega_i (X_i^m - \bar{X}_{treatment}^m) \quad \text{if } i \text{ in treatment group}$$

$$s_{control}^2 = \frac{\sum_i \omega_i}{(\sum_i \omega_i)^2 - \sum_i (\omega_i)^2} \sum_i \omega_i (X_i^m - \bar{X}_{control}^m) \quad \text{if } i \text{ in control group}$$

An absolute standardized difference below 10% is often considered as an indication of a negligible sample imbalance (Austin and Stuart 2015). Figure 4 shows that the PSW strategy eliminated the significant imbalances (existing in unweighted samples) from the weighted samples.

**Figure 4 Absolute Standardized Differences in Means of Covariates**



### 3.5.2. Matching Estimator of the Effect of Uber/Lyft’s Exit

One concern on the estimators with DiD analysis is that the specification of the demand model is restricted to an assumed functional form. Specifically, it assumes a linear relationship between the dependent variable and the independent variables. A violation of the assumed functional form may lead to a biased estimator due to model misspecification. To address the issue, we compare the estimation obtained from our main analysis to the matching estimator obtained from a standard (free-form) matching analysis. Specifically, we employ a one-to-one exact matching that matches each treated property in the treatment group with an untreated property in the control group. The average treatment effect is then computed by contrasting the difference between each pair of treated and untreated units before and after the treatment.

Table 4 presents the average treatment effect. In the matched sample, the effect of Uber/Lyft’s exit is positive and statistically significant, consistent with the results obtained from our main analysis. Hence, our estimation of the effect of Uber/Lyft’s exit is robust to the model’s specification. It is interesting to note that the difference in the demand between the treated and untreated groups is greater in the unmatched

sample (-18.91) than in the matched sample (-7.24). The results suggest that the difference in the unmatched sample captures possible seasonality between the treated and the untreated groups and that the matching method corrects this potential bias. Of course, a full assessment of the potential false significance caused by seasonality would require replicating the main DiD analysis using the prior year’s data (i.e., 2015), which we describe in the next section.

**Table 4 Robustness Check on Treatment Effect — A Standard Matching Analysis**

<b>OUTCOME</b>	<b>SAMPLE</b>	<b>TREATED</b>	<b>UNTREATED</b>	<b>DIFFERENCE</b>	<b>S.E.</b>
<i>DEMAND</i> (September)	Unmatched	12.723	31.632	-18.908	1.571
	<b>Matched</b>	<b>12.723</b>	<b>19.962</b>	<b>-7.239</b> <b>(-4.64)</b>	<b>1.561</b>

The t statistics are in parentheses.

### 3.5.3. Seasonality Examination (on Prior Years’ Data)

The DiD estimation relies on the parallel pre-treatment trends assumption that we have verified in section 4.2. However, one concern is that in the post-treatment periods, there was an idiosyncrasy associated with Austin but not with the cities in the control group. It would bias our treatment effect estimation if, for example, the two groups of cities have similar seasonal trends in January–May. However, compared to the other six cities, the number of visitors to Austin starts to decrease following May. To establish the robustness of our main results, we re-estimate our DiD model using the 2015 data:

$$\begin{aligned}
 DEMAND_{it} = & INTERCEPT + \alpha_3(AUSTIN_i \cdot AFTER_t^{2015}) + PERIOD_t + PROPERTY_i \quad (28) \\
 & + \varepsilon_{it}
 \end{aligned}$$

where the definitions of the variables are the same as in Equation (1), except 1) subscript  $t$  now indicates the corresponding periods in 2015, and 2)  $AFTER_t^{2015}$  is a dummy variable that equals 1 if  $t$  is a period after May in 2015. Hence, the coefficient of  $AUSTIN_i \cdot AFTER_t^{2015}$  captures the impact on the property

demands in Austin after May in 2015. Ideally, there should be no impact since the treatment of Uber/Lyft's exit did not occur in Austin in 2015.

We present the results in Table 5. As seen, the estimated coefficient of the key variable  $AUSTIN_i \cdot AFTER_t^{2015}$ , is not significantly different from zero (coefficient of -2.10, p-value=0.745). The insignificant estimation suggests that the estimated treatment effect in our main analysis is unlikely to be driven by false significance caused by an idiosyncrasy, such as the seasonality associated with Austin in the periods after May.

**Table 5 Examining Seasonality in Austin: Using Data on Prior Year**

<b>VARIABLES</b>	<b>ESTIMATES</b>
<i>AUSTIN · AFTER</i>	-2.102 (-0.33)
<i>INTERCEPT</i>	25.671*** (32.74)
Fixed Effect	Yes
Seasonality	Monthly
Num. Observations	64136
R-squared	0.531

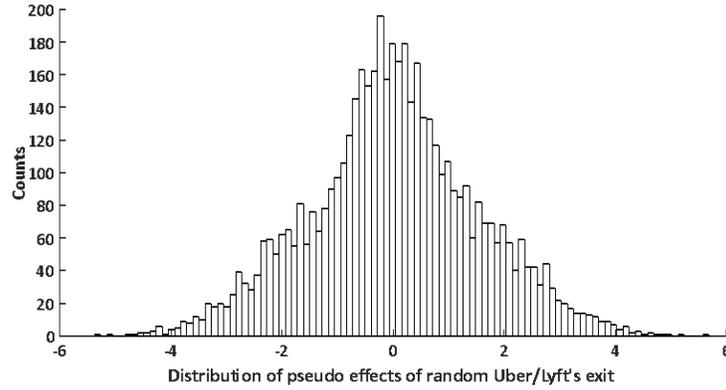
The t statistics are in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ .

### 3.5.4. Random (Shuffled) Treatment Test

In addition to the falsification test on the possible seasonality factor, we implement a random treatment test to examine the robustness of our main results to the possible false significance caused by serial correlation in the dependent variables (Bertrand et al. 2004). Specifically, we shuffle the treatment indicators  $AUSTIN_i \cdot AFTER_t$ , and then (randomly) reassign treatment indicators to the units in our sample (following the random treatment test in Greenwood and Agarwal (2016) and Burtch et al. (2017)). As a result, for property  $i$  in Austin in period  $t > \text{May}$  (i.e., the true treatment indicator  $AUSTIN_i \cdot AFTER_t$  equals 1 in the DiD demand model (1)), it could be assigned a pseudo treatment indicator that equals 0 in our random treatment test. Then, we estimate the DiD model with shuffled treatment indicators. We replicate the procedure 5,000 times and store the 5,000 sets of estimations. We expect the mean of the estimation results to be insignificant in order to reject a high probability of false significance.

In Figure 5, we plot the distribution of the estimated coefficient of the pseudo (shuffled) treatment indicators from the 5,000 replications. As seen, the estimations are centered on 0. In fact, we cannot reject the null hypothesis that the mean estimated pseudo effect is not different from 0 at a significant level (mean=0.022, standard deviation=1.532, and t-statistic=1.015). Furthermore, the distribution of the effects shows that the probability that we obtained the significant effect of Uber/Lyft’s exit in our main analysis is less than 0.1% ( $p\text{-value} < 0.001$  for rejecting the null that the mean of the distribution is significantly different from -9.6).

**Figure 5 Shuffled Treatment Test: Distribution of Pseudo Effects**



### 3.5.5. Inclusion of Additional Controls

One concern about our main analysis on the effect of Uber/Lyft's exit through analyzing Equation (1) is that there may be some omitted time-varying variables that are correlated with Austin's property demands. Following prior work (Agrawal and Goldfarb 2008), we address this concern by including additional covariates in the demand model to control for observable changes over time. Specifically, we incorporate the property price, the number of guests' reviews, and the average review score as the three variables that may affect a property's demand. We estimate the following demand model:

$$\begin{aligned}
 DEMAND_{it} = & INTERCEPT + \alpha_3(AUSTIN_i \cdot AFTER_t) + \gamma_1 REVIEW\_COUNT_{it} \\
 & + \gamma_2 REVIEW\_SCORE_{it} + \gamma_3 PRICE_{it} + PERIOD_t + PROPERTY_i + \varepsilon_{it}
 \end{aligned} \tag{29}$$

In Table 6 we present the results. The coefficients of the additional controls *REVIEW\_COUNT*, *REVIEW\_SCORE*, and *PRICE* are all significant, indicating that some of the variation in property demand is explained by these time-varying control variables. However, the relationship between property demand and the exit of Uber/Lyft remains statistically and economically significant. Specifically, the property booking will decrease by approximately 9.7 points if Uber/Lyft exits from a city. The consistent estimation result after including time-varying control variables indicates that our main results are robust. We use the specification with additional controls (i.e., Equation (5)) for the analyses in our empirical extension section.

**Table 6 Difference-in-Difference Model: Including Additional Controls**

<b>VARIABLES</b>	<b>ESTIMATES</b>
<i>AUSTIN · AFTER</i>	-9.653*** (-9.87)
<i>REVIEW_COUNT</i>	0.131*** (3.79)
<i>REVIEW_SCORE</i>	0.226*** (3.89)
<i>PRICE</i> <i>(instrumented)</i>	-0.107* (-2.29)
<i>INTERCEPT</i>	25.619*** (7.19)
Fixed Effect	Yes
Seasonality	Monthly
Num. Observations	67451
R-squared	0.701

The t statistics are in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ .

### **3.5.6. Excluding Alternative Explanations**

So far, we have tested the robustness of our main results with an extensive set of analyses. However, there are two alternative explanations for the negative impact of Uber/Lyft's exit. The first explanation relates to the concern of simultaneous regulations against Airbnb in Austin. This is possible if, for example, Austin executed a series of regulations on sharing economy platforms in 2016, with the one against Airbnb occurring to begin along with the exit of Uber and Lyft. After careful examination, we exclude this alternative explanation that the exit of Uber/Lyft was purely caused by Austin's policy on their background checks and was unrelated to other sharing economies. There was no policy debate on regulating/restricting Airbnb in Austin during the time when Uber/Lyft was absent.

The other alternative explanation is that the popularity of Austin as a travel destination decreased starting in May 2016. That is, the estimated effect of Uber/Lyft's exit was driven by the occurrence of fewer people visiting Austin in the periods when Uber/Lyft was absent. To address this concern, we collected information on Austin's tourism. Two organizations, [downtownaustin.com](http://downtownaustin.com) and [austintexas.org](http://austintexas.org), collect and report the tourism market performance in Austin, which includes the total number of visitors in each month to Austin.

We also obtain passenger boarding (enplanement) and all cargo data for the Austin airport from the Federal Aviation Administration<sup>56</sup>. The data from the three sources all indicate that Austin did not experience a decline in the travel market following May 2016 in terms of the total number of visitors.

### 3.6. Empirical Extensions

Our main analyses have shown consistently negative and significant impacts of the exit of Uber/Lyft on Airbnb property demand. The results are robust to a series of falsification checks. Next, we extend our analyses to identify the nature of the treatment effect, which is the mechanism through which Uber/Lyft plays a role in affecting Airbnb property demand. Exploring possible heterogeneity in the treatment effect helps us to identify the underlying mechanism. Specifically, we examine three key dimensions along which the treatment effect may vary across properties.

#### 3.6.1. Effect by Access to Transportation

We investigate how the transportation costs moderate the effect of Uber/Lyft exit on Airbnb property demand. The transportation cost from a property is captured by the transit score obtained from walkscore.com. A low transit score implies that a property has negligible access to public transportation and a guest would need to arrange a cab or a ride sharing service. The lack of access to public transportation implies that the transportation costs from a location are high. To capture the moderating effect of transit scores on the treatment effect of Uber/Lyft's exit, we include the interaction term of the transit score and the treatment indicator using the model specification in Equation (5). Then, we estimate the following demand equation

$$\begin{aligned}
 DEMAND_{it} = & INTERCEPT + \alpha_3(AUSTIN_i \cdot AFTER_t) \\
 & + \rho(AUSTIN_i \cdot AFTER_t \cdot TRANSIT\_SCORE_i) + \gamma_1 REVIEW\_COUNT_{it} \\
 & + \gamma_2 REVIEW\_SCORE_{it} + \gamma_3 PRICE_{it} + PERIOD_t + PROPERTY_i + \varepsilon_{it}
 \end{aligned} \tag{30}$$

Note that *TRANSIT\_SCORE* is time-invariant (not affected by the exit of Uber/Lyft) and hence is absorbed by the property fixed effect term. The key coefficient,  $\alpha_3$ , captures the marginal effect of having an extra transit score, i.e., having better access to the public transportation, on property demand in the absence of Uber/Lyft. Table 7 reports the estimation results from model (6). The positive coefficient of *AUSTIN·AFTER·TRANSIT\_SCORE* suggests that the transit score positively moderates the treatment effect. For example, the treatment effect for the properties having a transit score of 0 is -12.93, indicating an approximate loss of 13% in demand in the absence of Uber/Lyft. The effect is due to the lack of substitutes

---

<sup>56</sup> Data from Austin–Bergstrom International Airport shows the following respective number of visitors for January through September in 2016: 873,560, 839,213, 1,066,146, 1,028,337, 1,081,450, 1,135,796, 1,133,641, 1,054,496 and 1,018,292.

for Uber/Lyft in areas with poor public transportation access. However, for Airbnb properties located in a neighborhood with a transit score of 100, the maximum transit score sample would experience a positive effect ( $-12.93\% + 0.163\% \cdot 100 = 3.4\%$ ). The differential effects on the property demand arise because the access to public transportation becomes a key factor that mediates travelers' choices of lodging alternatives. Hence, when Uber and Lyft exit a city, the increased transportation costs would cause some travelers to switch to areas with better transportation. As a result, the lodgings in the areas with excellent access to public transportation services (e.g., main hotel districts) will receive extra demand.

Although Airbnb properties in the areas with good access to transportation will benefit from the exit of Uber/Lyft, there are only a few of them. For the majority of Airbnb properties, they will be worse off in the absence of Uber/Lyft. Together, these results indicate that Uber/Lyft's exit may significantly reduce the overall demand of Airbnb properties. Further, the geographic demand dispersion in Airbnb will significantly decrease with respect to properties that rely on Uber/Lyft to reduce transportation costs, thus losing a significant amount of demand.

In Figure 6, we plot the distribution of transit scores associated with the properties and indicate the regions where the exit of Uber/Lyft leads to a negative (bins coded in red) or positive (bins coded in green) effect on the property demand.

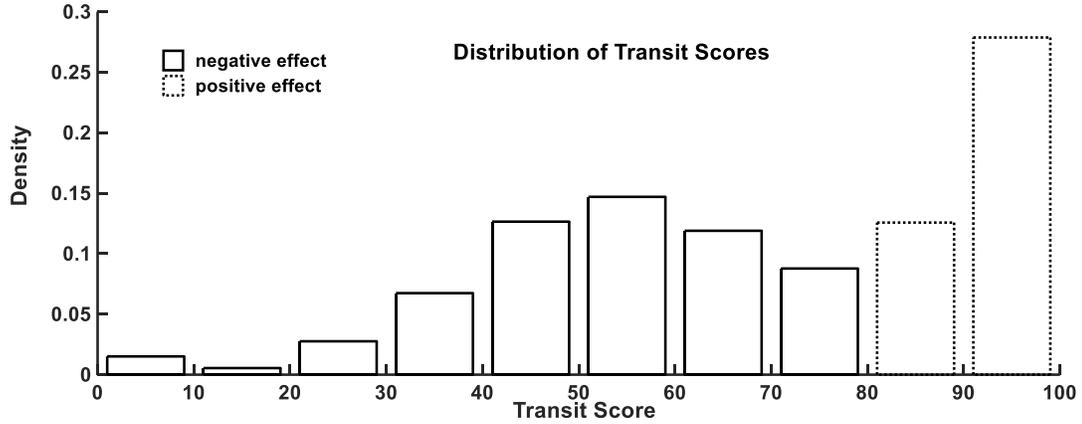
**Table 7 Heterogeneous Effect of Uber/Lyft's Exit on Demand: Access to Transportation**

<b>VARIABLES</b>	<b>ESTIMATES</b>
<i>AUSTIN · AFTER</i>	-12.934*** (-14.53)
<i>AUSTIN · AFTER · TRANSIT_SCORE</i>	0.163*** (11.10)
<i>REVIEW_COUNT</i>	0.110*** (6.69)
<i>REVIEW_SCORE</i>	0.133** (3.13)
<i>PRICE</i> (instrumented)	-0.106*** (-9.49)
<i>INTERCEPT</i>	28.901*** (29.96)
Fixed Effect	Yes
Seasonality	Monthly

Num. Observations	49658
R-squared	0.77

The t statistics are in parentheses. \* p<0.05, \*\* p<0.01, and \*\*\* p<0.001.

**Figure 6 The Effect of Uber/Lyft's Exit on Demand: Distribution of Transit Scores**



### 3.6.2. Effect by Property Luxuriousness

To further understand the underlying mechanism of the treatment effect, we investigate the heterogeneity in the effects of Uber/Lyft's exit across Airbnb properties' luxuriousness levels. To do so, we construct a dummy variable that reflects whether an Airbnb property's average daily price falls into the top or bottom half of the property price distribution in the city. Then, we interact the treatment indicator  $AUSTIN \cdot AFTER$  with the dummy reflecting the top half, i.e., high end. The differential treatment effects across property luxuriousness levels are captured by  $\delta$ , which is the coefficient of the interaction term  $AUSTIN \cdot AFTER \cdot HIGHEND$ .

$$\begin{aligned}
DEMAND_{it} = & INTERCEPT + \alpha_3(AUSTIN_i \cdot AFTER_t) & (31) \\
& + \delta(AUSTIN_i \cdot AFTER_t \cdot HIGHEND_{it}) + \gamma_1 REVIEW\_COUNT_{it} \\
& + \gamma_2 REVIEW\_SCORE_{it} + \tau HIGHEND_{it} + \gamma_3 PRICE_{it} + PERIOD_t \\
& + PROPERTY_i + \varepsilon_{it}
\end{aligned}$$

Column (1) in Table 8 reports the estimates from model (7) where we use all properties in our sample. The positive coefficient of  $AUSTIN \cdot AFTER \cdot HIGHEND$  suggests that a low-end property, compared to high-end property, received a greater negative treatment effect. The interpretation is that the exit of Uber/Lyft would most impact the target consumers of low-end properties, which are guests with relatively tighter travel budgets. These guests choose alternatives (e.g., hotels and low-end properties in downtown) to substitute for the low-end Airbnb properties due to the increased transportation costs.

Next, we specifically investigate properties in neighborhoods with poor transit scores. Following walkscore.com’s classification of transit score (Figure 2), we label that a property is in a “Low Transit Zone” if its transit score is below 50. This gives us the bottom quartile of Airbnb properties that fall into the ‘low transit’ category. Column (2) reports the results from estimating properties in Low Transit Zones. The positive coefficients of *AUSTIN·AFTER·HIGHEND* are consistent with the results in Column 1. Specifically, in the low transit areas, the heterogeneity in the treatment effect across Airbnb properties is even more significant. Specifically, high-end properties received only  $(-12.1+6.528)/(-12.1)=46.1\%$  of the negative impact observed for low-end properties in the same low transit zone. This is because the increased transportation costs due to Uber/Lyft’s exit are particularly a concern if travelers stay in low-transit areas. Considering that guests going for low-end properties are likely to be more price-sensitive and more sensitive to any increase in the transportation costs, they would switch to areas with cheaper transportation services.

**Table 8 Heterogeneous Effect of Uber/Lyft’s Exit on Demand: Property Luxuriousness**

VARIABLES	ESTIMATES	
	(1)	(2)
	All Zones	Low Transit Score Zones
<i>AUSTIN · AFTER</i>	-10.475*** (-30.51)	-12.100*** (-18.82)
<i>AUSTIN · AFTER · HIGHEND</i>	3.063*** (5.51)	6.528*** (7.55)
<i>REVIEW_COUNT</i>	0.185*** (11.19)	0.225*** (7.14)
<i>REVIEW_SCORE</i>	0.215*** (5.25)	0.231*** (4.55)
<i>HIGHEND</i>	1.403** (2.65)	0.058 (0.05)
<i>PRICE</i> (instrumented)	-0.118*** (-10.99)	-0.183*** (-10.12)
<i>INTERCEPT</i>	25.121*** (24.66)	28.940*** (16.39)
Fixed Effect	Yes	Yes
Seasonality	Monthly	Monthly
Num. Observations	65154	15695
R-squared	0.706	0.702

---

t statistics in parentheses. \* p<0.05 \*\* p<0.01 \*\*\* p<0.001

### 3.6.3. Effect by Geographic Locations

We incorporate a property’s geographic information to directly estimate the differential treatment effects across properties with varying commute times to downtown areas. Downtown is a good proxy for the main destination of travelers since these areas involve concentrated business/tourist activities. From walkscore.com, we collect data on the (driving) commute time from each Airbnb property’s address to the downtown area in that city, and then use the commute times to approximate transportation costs. Let  $COMMUTE_i$  denote the average time (in minutes) it takes to drive a car from property  $i$  to the downtown of the city. The following model identifies the differential effects across commute times through coefficient  $\eta$ .

$$\begin{aligned}
 DEMAND_{it} = & INTERCEPT + \alpha_3(AUSTIN_i \cdot AFTER_t) + \eta(AUSTIN_i \cdot AFTER_t \cdot \\
 & COMMUTE_i) + \gamma_1REVIEW\_COUNT_{it} + \gamma_2REVIEW\_SCORE_{it} + \gamma_3PRICE_{it} + \\
 & PERIOD_t + PROPERTY_i + \varepsilon_{it}
 \end{aligned} \tag{32}$$

Note that although the commute times vary depending on the time of day (e.g., rush hours) and dates (i.e., weekends), we use the commute time averaged in a year provided by walkscore.com. Therefore, in Equation (8),  $COMMUTE$  is time-invariant (not affected by the exit of Uber/Lyft) and is absorbed by the property fixed effect term.

Table 9 reports the estimates from the model in Equation (8). As anticipated,  $\eta$  is negative, which suggests that properties farther from downtown receive greater negative shocks to their demands in the absence of Uber/Lyft. Specifically, every unit increase in commute time (minutes) would lead to a decrease of approximately 0.208% in the property booking. Considering the \$0.93 UberX fare relative to the \$1 taxi fare per minute travel in Austin on average<sup>57</sup>, if transportation costs increased by \$1, the properties would be 0.208%/0.93=0.22% less frequently booked. The resulting effect is a drop of (0.22%\*\$185/day\*365days=\$148.5) in the average annual revenue to the host of an average unit for every \$1 increase in transportation costs.

**Table 9 Economic Impact of Commute Time to Downtown on Airbnb Property Demand**

VARIABLES	ESTIMATES
-----------	-----------

---

<sup>57</sup> We estimate the average fares of taxis and Uber from the following sources: 1) <https://www.taxifarefinder.com>, 2) <http://uberestimate.com/prices>, and. 3) <http://uber-rates-austin-tx-us.uber-fare-estimator.com/>. Taxi fares include an additional 15% tip.

<i>AUSTIN · AFTER</i>	0.841*
	(2.15)
<i>AUSTIN · AFTER · COMMUTE</i>	-0.208***
	(-8.98)
<i>REVIEW_COUNT</i>	0.082***
	(4.82)
<i>REVIEW_SCORE</i>	0.131***
	(2.96)
<i>PRICE</i>	-0.081***
<i>(instrumented)</i>	(-6.93)
<i>INTERCEPT</i>	29.351***
	(25.97)
Fixed Effect	Yes
Seasonality	Monthly
Num. Observations	64391
R-squared	0.703

The t statistics are in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ .

### 3.7. Conclusion

The internet-based peer-to-peer platforms have gained momentum across industries and emerged as the means for individual users to monetarize their excessive capacities. In this study, we investigate the demand interactions across different sharing economies, specifically between ride-sharing economy (Uber/Lyft) and the home-sharing economy (Airbnb). The interaction arises because the advantage of Uber/Lyft in reducing transportation costs can alleviate a major disadvantage that Airbnb properties face caused by their poor locations relative to hotels. That is, Uber/Lyft solves Airbnb customers' problem of significantly higher transportation costs compared to staying in hotels. Uber/Lyft influences Airbnb property demand by 1) dispersing hotels' demand to Airbnb properties and 2) redistributing lodging demand within Airbnb properties. The nature of Uber/Lyft's effect on Airbnb depends on the extent to which travelers think Airbnb is horizontally differentiated from hotels. We document evidence supporting that Airbnb may not be horizontally differentiated as marketed.

To examine the effect of Uber/Lyft on Airbnb, we exploit a natural experiment that significantly increased the transportation costs in Austin due to the exit of Uber and Lyft, resulting in shifts in consumers' lodging choices. We leverage a Difference-in-Difference methodology to quantify the effect of Uber/Lyft's exit on the demand of Airbnb properties. We test the robustness of our models and results with an extensive

set of robustness analyses, including the validation of the pre-treatment parallel trend assumption, the seasonality examination, the model-free matching estimation, the random treatment test, and the inclusion of additional controls.

Several notable findings are drawn from our analyses. *First*, we find that Airbnb properties are, on average, 9.6% less frequently booked in the absence of Uber/Lyft. The resulting effect is a decrease of 6,482 USD in annual revenue to the Airbnb host of an average unit. Furthermore, the effect varies across the transit scores (amount of public transportations) near a property, with the properties having poor transit scores being the most affected. The findings reveal the mechanism of the demand interaction in which Uber/Lyft plays a role in affecting the demand of Airbnb properties through transportation. *Second*, the heterogeneous effects suggest that Airbnb properties in areas with excellent access to public transportation will be, on average, 3.4% more frequently booked. However, since the majority of the Airbnb properties are in locations with relatively poor access to transportation, the exit of Uber/Lyft will lead to a drop in the overall Airbnb property demand. *Third*, we find that in areas with poor transportation, the exit of Uber/Lyft leads to a decrease of approximately 12.1% in the demand for low-end properties. The effect is more than double the impact for high-end properties in the same areas. The differential effects arise because Uber/Lyft's exit more impacts consumers of low-end properties. Those customers choose alternatives in areas with better transit scores (e.g., hotels and Airbnb properties in downtown locations) to avoid the increased transportation costs. *Fourth* and finally, we find a decrease of 0.21% in property demand or a decrease of \$148.5 in the average host's annual revenue in response to each one-dollar increase in commute costs. The results show that the exit of Uber/Lyft causes a decrease in the geographic dispersion of demand throughout the city and an increase in the concentration of demand in conventional hotel districts.

Altogether, this paper quantifies the demand interaction of Uber/Lyft and Airbnb and provides insights into the mechanism behind the interactions. Our results revealed that a significant portion of Airbnb demand depends on easy access to transportation services. While ride-sharing services take the guests to areas that are underserved by public transportation and traditional cab services, this also highlights a vulnerability of Airbnb. The fact that high-end properties in areas with poor transit scores remain largely unaffected by the exit of Uber/Lyft suggests that Airbnb should try to attract more high-end properties in such areas to reduce its vulnerability to changes in transportation costs.

Finally, we note that the rise of sharing economies has drawn massive attention from academia and led to policy debates. However, prior studies have largely focused on the impact of one sharing economy on incumbent industries while ignoring the interactions among sharing economies. Our research effort is the first step toward understanding the externalities between sharing economies.

## References

- Agrawal, A., and Goldfarb, A. 2008. "Restructuring Research: Communication Costs and the Democratization of University Innovation." *American Economic Review*, 98(4): 1578-90.
- Airbnb. 2017. "The Economic Impacts of Home Sharing in cities around the world". <https://www.airbnb.com/economic-impact>; <https://www.airbnb.com/about/about-us>.
- Angrist, J. D. and Pischke, J.-S. 2008. "Mostly harmless econometrics: An empiricist's companion", Princeton university press.
- Athey, S. and Imbens, G. W. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models.", *Econometrica* 74(2): 431–497.
- Austin, P.C., Stuart, E.A. 2015. "Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies.", *Statistics in Medicine* 34 (28): 3661–3679.
- Autor, D.H. 2003. Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing. *Journal of Labor Economics* 21(1): 1-42.
- Bertrand, M., Duflo, E., Mullainathan, S. 2004. "How Much Should We Trust Differences-in-Differences Estimates?", *The Quarterly Journal of Economics* 119(1): 249–275.
- Barron, K., Kung, E., Proserpio, D. 2017. "The Sharing Economy and Housing Affordability: Evidence from Airbnb.", Available at SSRN: <https://ssrn.com/abstract=3006832>.
- Burtch, G., Carnahan, S., Greenwood, B.N. 2017. "Can You Gig It? An Empirical Examination of the Gig Economy and Entrepreneurial Activity", *Management Science* Forthcoming. Available at SSRN: <https://ssrn.com/abstract=2744352>.
- Calo, R., Rosenblat, A. 2017. "The Taking Economy: Uber, Information, and Power.", *Columbia Law Review* Vol. 117. Available at SSRN: <https://ssrn.com/abstract=2929643>.
- Cramer, J., and Krueger, A.B. 2016. "Disruptive Change in the Taxi Business: The Case of Uber." *American Economic Review* 106(5): 177–82.
- Dobbins, J. 2017. "Making a Living With Airbnb", Available at <https://www.nytimes.com/2017/04/07/realestate/making-a-living-with-airbnb.html>.

- Edelman, B., Luca, M., Svirsky, D. 2017. "Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment.", *American Economic Journal: Applied Economics* 9(2): 1–22.
- Edelman, B., Geradin, D. 2016. "Efficiencies and Regulatory Shortcuts: How Should We Regulate Companies like Airbnb and Uber?", *Stanford Technology Law Review* 19(2): 293–328.
- Ellinger, R.1977. "Industrial Location Behavior and Spatial Evolution.", *Journal of Industrial Economics*, 25 295–312.
- Gong, J., Greenwood, B. N., Song, Y. 2017. "Uber Might Buy Me a Mercedes Benz: An Empirical Investigation of the Sharing Economy and Durable Goods Purchase.", Available at SSRN: <https://ssrn.com/abstract=2971072>.
- Greenwood, B.N., Wattal, S. 2017. "Show Me the Way to Go Home: An Empirical Investigation of Ride Sharing and Alcohol Related Motor Vehicle Fatalities.", *MIS Quarterly* 41(1): 163–187.
- Heckman, J., Ichimura, H., Todd, P.E. 1997. "Matching As an Econometric. Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies* 64(4): 605–654.
- Hirsch, JA, Moore, K.A., Evenson, K.R., Rodriguez, D.A., Diez Roux, A.V. 2013. Walk Score and Transit Score and Walking in the Multi-Ethnic Study of Atherosclerosis. *American journal of preventive medicine* 45(2):158-166.
- J.D. Power. 2015. "2015 North America Hotel Guest Satisfaction Index Study". Available at <http://www.jdpower.com/press-releases/2015-north-america-hotel-guest-satisfaction-index-study>.
- Jefferson-Jones, Jamila. 2015. "Can Short-Term Rental Arrangements Increase Home Values?: A Case for Airbnb and Other Home Sharing Arrangements.", *The Cornell Real Estate Review* Vol. 13, June 2015. Available at SSRN: <https://ssrn.com/abstract=2714051>.
- Levendis, J., Dicle, M.F. 2016. "The Neighborhood Impact of Airbnb on New Orleans.", Available at SSRN: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2856771](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2856771).
- Li, J., Moreno, A., Zhang, D.J. 2016. "Pros vs Joes: Agent Pricing Behavior in the Sharing Economy" (August 28, 2016). Ross School of Business Paper No. 1298. Available at SSRN: <https://ssrn.com/abstract=2708279>.

Malhotra, A., Van Alstyne, M. 2014. The Dark Side of the Sharing Economy... and How to Lighten It. *Communications of the ACM* 57(11): 24–27.

Manyika, J., Lund, S., Bughin, J., Robinson, K., Mischke, J., Mahajan, D. 2016. “Independent work: Choice, necessity, and the gig economy.”, Available at <https://www.mckinsey.com/global-themes/employment-and-growth/independent-work-choice-necessity-and-the-gig-economy>.

Miller, S.R. 2016. “First Principles for Regulating the Sharing Economy.”<sup>53</sup> *Harvard Journal on Legislation* 147 (2016). Available at SSRN: <https://ssrn.com/abstract=2568016>.

Picchi, A. 2016. “Uber vs. Taxi: Which Is Cheaper?”, *Consumer Report*. June 10, 2016. <http://www.consumerreports.org/personal-finance/uber-vs-taxi-which-is-cheaper/>.

PwC Report. 2015. “Consumer Intelligence Series: The sharing economy”, <https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/sharing-economy.html>.

Rosenbaum, P., Rubin, D.B. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.”, *Biometrik* 70(1): 41–55.

Rubin, D.B. 2001. “Using propensity scores to help design observational studies: application to the tobacco litigation.”, *Health Services & Outcomes Research Methodology* 2:169–188.

Stuart, E.A. 2010. “Matching methods for causal inference: A review and a look forward.”, *Stat Sci.* 25(1): 1–21. doi: 10.1214/09-STS313.

Sundararajan, A. 2016. “The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism”, The MIT Press.

Villas-Boas, J., Winer, R. 1999. “Endogeneity in brand choice models.”, *Management Science* 45(10):1324–1338.

Wyckoff, D. D., Sasser, W.E.1981. “The U.S. Lodging Industry”. Lexington, MA: D.C. Heath.

Zervas, G., Proserpio, D., Byers, J.W. 2017. “The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry.”, *Journal of Marketing Research* 54(5): 687–705.

Zhang, S., Lee, D., Singh, P., Srinivasan, K. 2017. "How Much Is an Image Worth? Airbnb Property Demand Estimation Leveraging Large Scale Image Analytics.", Available at SSRN: <https://ssrn.com/abstract=2976021>