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

ABSTRACT

We investigate the economic impact of images and lower-level image factors that influence property demand in Airbnb. Using Difference-in-Difference analyses on a 16-month Airbnb panel dataset spanning 7,711 properties, we find that units with *verified* photos (taken by Airbnb's photographers) generate additional revenue of $2,521 per year on average. For an average Airbnb property (booked for 21.057% of the days per month), this corresponds to 17.51% increase in demand due to *verified* photos. Leveraging computer vision techniques to classify the image quality of more than 510,000 photos, we show that 58.83% of this effect comes from the high image quality of *verified* photos.

Next, we identify 12 interpretable image attributes from photography and marketing literature relevant for real estate photography that capture image quality as well as consumer taste. We quantify (using computer vision algorithms) and characterize unit images to evaluate the economic impact of these human-interpretable attributes. The results reveal that verified images not only differ significantly from low-quality unverified photos, but also from high-quality unverified photos on most of these features. The treatment effect of verified photos becomes insignificant once we control for these 12 attributes, indicating that Airbnb's photographers not only improve the quality of the image but also align it with the taste of potential consumers.

This suggests there is significant value in optimizing images in e-commerce settings on these attributes. From an academic standpoint, we provide one of the first large-scale empirical evidence that directly connects systematic lower-level and interpretable image attributes to 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.

Keyword: sharing economy, Airbnb, economic impact of images, photography, computer vision, deep learning, image quality classification, image feature extraction, treatment effect, image attribute analysis

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

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.[1] 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 drivers of a good image for an Airbnb property? 4) Finally, even if the drivers can be identified, 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 drivers in image attributes that 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 systematic manner to connect directly to economic outcomes 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

[1] http://airhostsforum.com/t/professional-photography/3675/35.

literature in psychology and marketing. Twelve dimensions of image attributes (see Table 6) 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 17.51% more frequently booked by having verified photos[2]. The effect of verified images is positive and significant even after controlling for other sources of information such as property reviews. Second, we separate the effect of higher quality images from the additional trust arising from verified images. We find that 58.83% of treatment effect of verified photos comes from high quality of those 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 treatment effect of verification becomes statistically insignificant. The results suggest that most of the effect of the verified photos is coming from these 12 image attributes which together capture not only quality but also taste. In comparison to amateurs as well as external professionals, Airbnb professional photographers better capture the attributes that matter for Airbnb property demand. Fourth, results further show that color is the most important component affecting Airbnb property demand. Improving color attributes by one standard deviation could bring extra revenue of approximately $6,484 per year, followed by $3,485 for figure-ground relationship attributes and $2,432 for attributes under composition component, with similar improvements (one standard deviation)[3].

Our research effort makes several key contributions. *First*, this is among the first papers to connect image attributes 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 and causally test the impacts of three key interpretable image components on product demand. We validate that color is the most important image component that impacts property

[2] On average a property without verified photos is booked 21.057% of the days in pre-treatment period (i.e., January 2016) in our sample. If this property gets verified photos, it will be booked (21.057+3.687)% of the days in a month. This corresponds to a 17.51% increase in demand. For a property in 365 days a year, this corresponds to 13.46 days of additional booked days in a year.
[3] The extra revenue generated from the extra booking is computed by multiplying the extra booked days with the average property daily price. We use the price for an average unit (without verified photos) in the pre-treatment period, i.e., 187.30 USD/day (see Table 1).

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. On average our image analytics algorithm efficiently computes an extensive set of image attributes in 1.06 seconds per image, on a 2 Intel Haswell (E5-2695 v3) CPU. The analytics step can be easily 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.

2. Relevant Literature and Theoretical Framework

2.1 Background 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 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 identify the direct 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.

2.2 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.

2.2.1 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.

<u>Attribute 1.</u> ***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 2, the image on the right is more diagonally dominant that 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 2 Compare Images on Diagonal Dominance

Image Without Diagonal Dominance

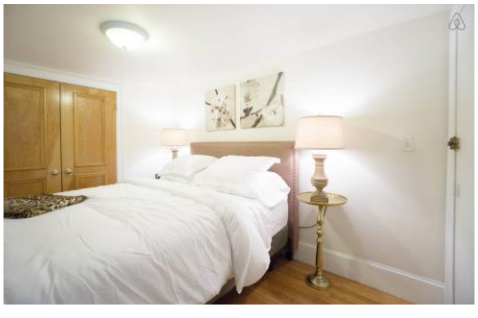


Image With Diagonal Dominance

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 3, 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 points and the important 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 3 Compare Images on Rule of Thirds

Image (Relatively) Doesn't Follow Rule of Thirds

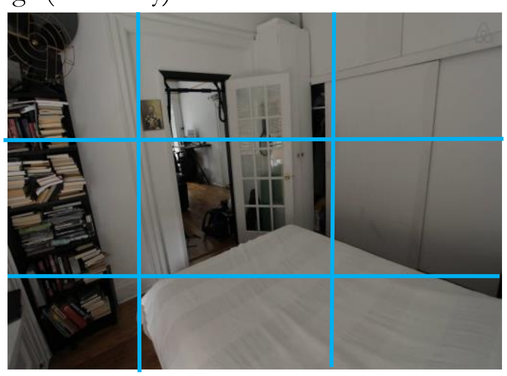


Image (Relatively) Follows Rule of Thirds

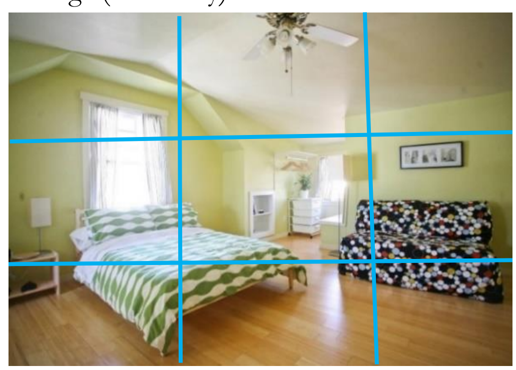


Figure 4 Compare Images on Visual Balance

Image without Visual Balance

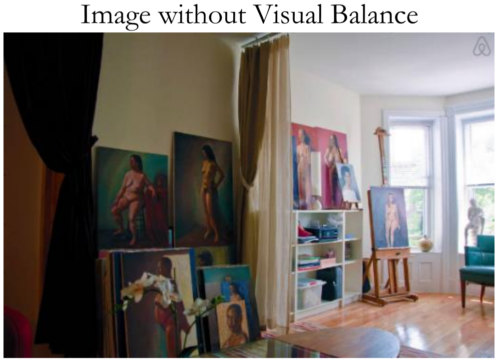


Image with Visual Balance

Attributes 3 and 4. ***Visual Balance Intensity and Visual Balance Color.*** Visual Balance relates to the distribution and arrangement of visual elements (Krages 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 4 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.

2.2.2 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 5 we present a cool photo of a living room on the left and a warm photo of a living room on the right.

Figure 5 Compare Images on Hue (Cool Color vs. Warm Color)

Image with Cool Color

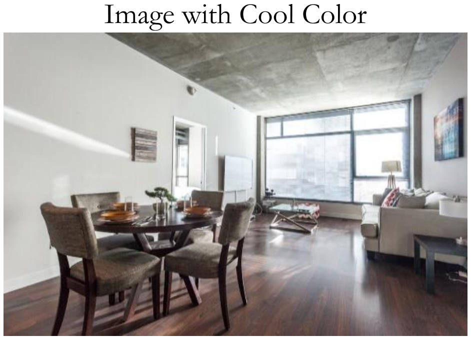


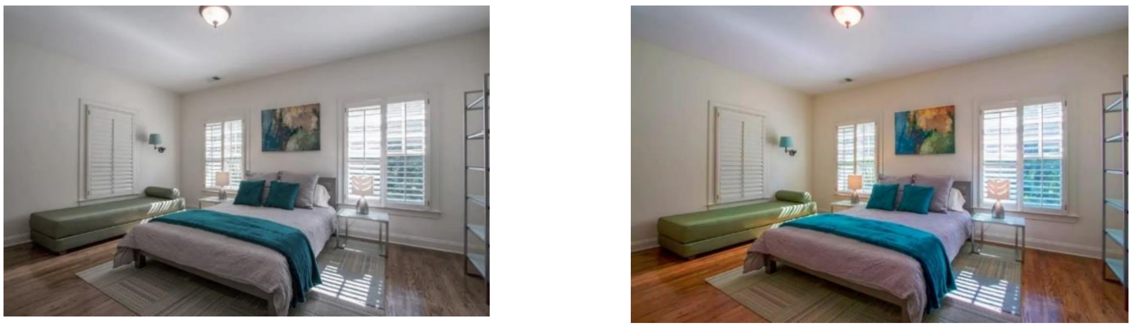
Image with Warm Color

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 6. The only difference between the images is that the image on the right is more saturated than the one on the left.

Figure 6 Compare Images on Saturation

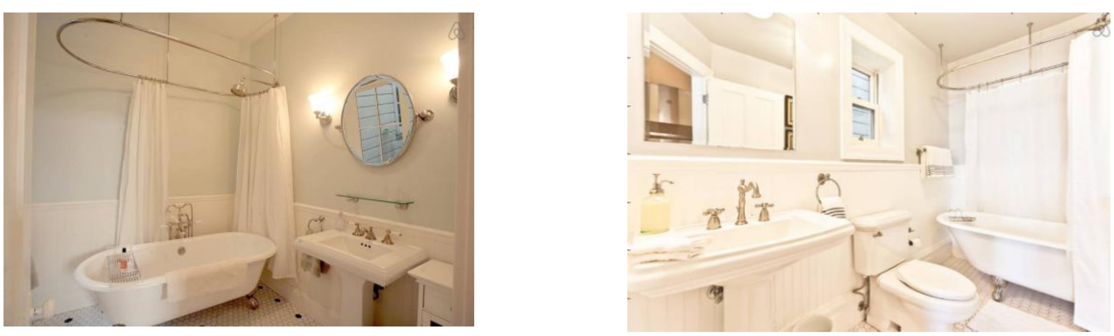
Image with Low Saturation | Image with High Saturation



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 7, 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 7 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 8 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.

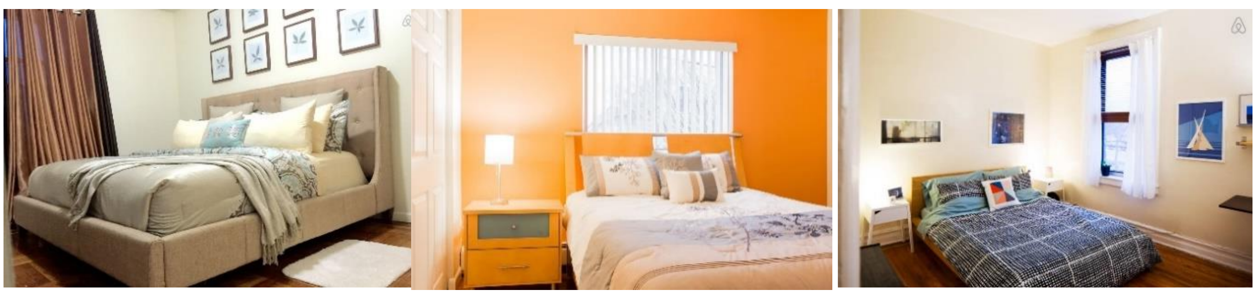
Figure 8 Compare Images on Clarity

Image with Dull Color

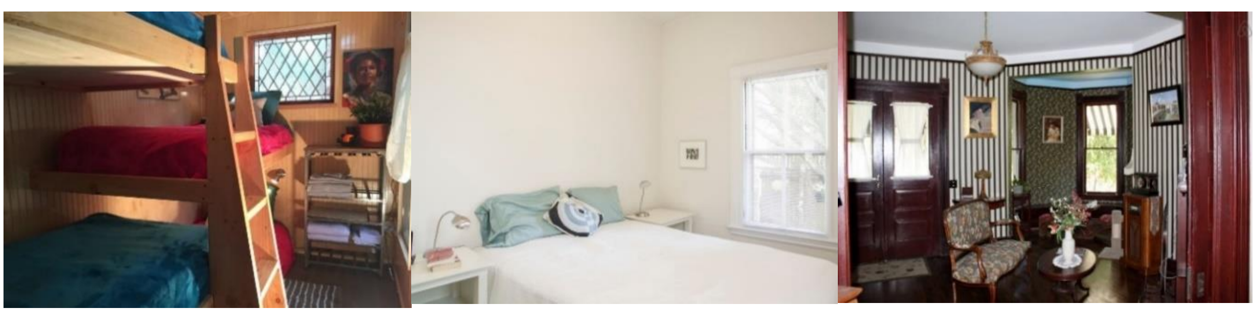
Image with Clear Color

Figure 9 Compare Images on Figure-Ground Relationships

9a. Clear Separation of Figure from Ground



9b. Unclear Separation of Figure from Ground



2.2.3 Component: Figure-Ground Relationship

<u>Attributes 10, 11, and 12.</u> ***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 9, we present one set of images where the figure is clearly separable from the ground and another set where the separation is not obvious.

3. Empirical Framework

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, neighborhood), property type (e.g., house vs. apartment), property size (i.e., the number of bedrooms and beds), and capacity (number of beds, 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.

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 7,711 unique properties that did not have a verified photo by January 2016. Out of these properties, 224 had verified photos by the end of April 2017 (constituting the treatment group), and the remaining 7,487 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 property booking data from a third-party company that sells Airbnb property demand data. The booking data includes the number of days in a month that a property is available (i.e., the property was available to be booked), booked (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). 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. For each property *i* in each period *t*, property demand $DEMAND_{it}$ is measured as the fraction of days that a property is booked in that period. We further scale the fraction by 100. For example, if a property in March and was booked for 6 days, then the property's demand in that month is (6 days/31 days) * 100=19.35.

Property Price and its Instrument

Property price for a property *i* in period *t*, $PRICE_{it}$, refers to the average over property *i*'s daily price for days in period *t*. Property price is endogenous because it is correlated with random shock in the current period that also affects property demand. To address the endogeneity concern, we use lagged price—the property price in the previous period, $PRICE_{it-1}$—as the instrument variable for price. Lagged price can be a valid instrument for current price because it is correlated with current price, $PRICE_{it}$; however, it is uncorrelated with random shock in the current period. In marketing research, lagged prices often serve as an instrument for endogenous price, due to their ready availability to researchers (Villas-Boas and Winer 1999).

Property Photos

Property photos refer to the set of photos posted on the property webpage in a period. Two variables are measured from the data on property photos: the photo quantity and the photo quality. 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),[4] 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

[4] 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.

property i had 10 image 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.

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.

Table 1 Summary Statistics of Airbnb Properties

	Variables	**Treatment Group** (224 properties)		**Control Group** (7487 properties)	
		Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Pre-treatment (January 2016)	*DEMAND indicates the occupancy rate of a property in a period*				
	DEMAND (%)	20.696	28.559	21.057	28.470
	IMAGE_QUALITY indicates the average quality over all property images observed in each period				
	IMAGE_QUALITY	0.284	0.255	0.282	0.258
	IMAGE_COUNT indicates the number of property images observed in each period				
	IMAGE_COUNT	15.040	9.345	12.662	9.613
	REVIEW_COUNT indicates the total number of guest reviews accumulated by the beginning of each period				
	REVIEW_COUNT	19.522	22.841	18.175	29.744
	REVIEW_SCORE indicates the overall review score rated by guests observed in each period				
	REVIEW_SCORE	87.022	25.024	75.584	37.110
	PROPERTY_PRICE indicates the average daily price of a property in each period				
	PROPERTY_PRICE	162.491	185.454	187.295	259.681
Post-treatment (April 2017)	*DEMAND (%)*	33.515	33.183	25.997	28.997
	IMAGE_QUALITY	0.751	0.220	0.283	0.257
	IMAGE_COUNT	22.821	12.811	13.938	11.724
	REVIEW_COUNT	45.293	43.939	39.357	51.827
	REVIEW_SCORE	92.619	12.114	79.387	32.923
	PROPERTY_PRICE	217.373	154.807	221.186	260.981
Time-invariant	*NUM_BEDROOMS indicates the number of bedrooms*				
	NUM_ BEDROOMS	1.308	0.852	1.274	0.844
	NUM_BEDS indicates the number of beds				
	NUM_BEDS	1.772	1.208	1.720	1.249
	PROPERTY_TYPE categorical variable=1, 2… for different property types such as: apartment, house, condo, flat, etc.				
	PROPERTY_TYPE	6.407	7.570	6.421	7.682
	ROOM_TYPE categorical variable =1 (indicating entire place) or 2 (indicating shared) place)				
	ROOM_TYPE	1.473	0.578	1.439	0.553

As shown in Table 1, for units in the control group, the image quality stayed approximately the same, from 0.282 in January 2016 to 0.283 in April 2017. However, for the units in the treatment group, the image quality drastically improved from 0.284 in January 2016 to 0.751 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.

3.3 Analysis on Property Images

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 the 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

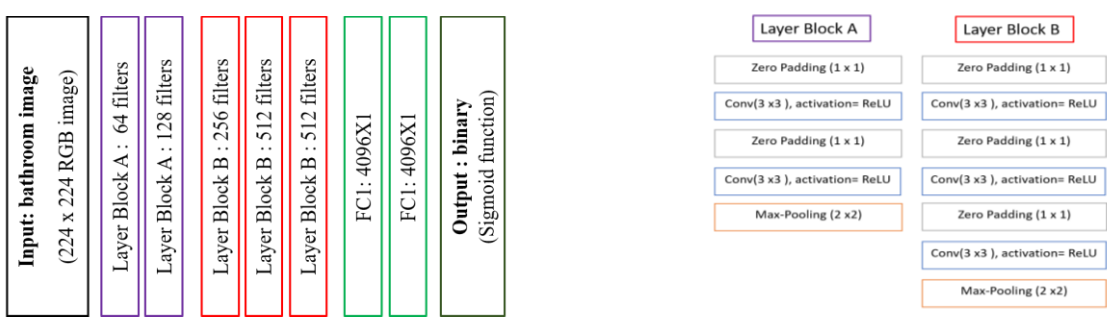
Convolutional Neural Networks (CNN) Approach: We apply Convolutional Neural Networks (CNN), an emerging deep learning framework widely applied in the field of computer vision that has shown breakthrough performance tasks including object recognition and image classification (Krizhevsky et al. 2012, Simonyan and Zisserman 2015). Our CNN image quality classifier, as shown in Figure 10, 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°. 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 10) and assigns "1" to high-quality images and "0" to low-quality images.

Figure 10 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 2×2 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.

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).

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 224 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 7,487 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 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.

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 a logit model as a function of observed covariates. That is, the propensity score of unit i, ps_i, is computed as follows:

$$ps_i = f(\boldsymbol{X_i\beta})$$

where $\boldsymbol{X_i}$ is a $1 * M$ dimensional vector of pre-treatment observed covariates of unit i and $\boldsymbol{\beta}$ is an $M * 1$ dimensional vector of parameter for 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 $\boldsymbol{X}$ is based on performing a covariates balance check. That is, the differences in the means of covariates between treatment and control groups should be minimized. With parameter vector $\boldsymbol{\beta}$ estimated, we approximate for each unit *i* (with observed covariates $\boldsymbol{X_i}$) the propensity score $\widehat{ps_i(\boldsymbol{X_i})}$, 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 inversing its propensity score. Specifically, the weight for unit i is defined as

$$\omega_i(T, \boldsymbol{X_i}) = \frac{T}{\widehat{ps_i(\boldsymbol{X_i})}} + \frac{1-T}{1-\widehat{ps_i(\boldsymbol{X_i})}}$$

where $\widehat{ps_i(\boldsymbol{X_i})}$ is the estimated propensity score of unit i computed with its observed covariates $\boldsymbol{X_i}$. T is a dummy variable that equals 1 if i is in treatment group and is 0 if otherwise. Having obtained sampling weights from propensity scores, we implement DiD analysis on the re-weighted sample, i.e., run a weighted regression.

Validating the Propensity Score Weighting Method: Balance Check on Covariates

To ensure that our PSW strategy is valid, we need to show that the propensity-scores-based sampling weights properly balance the two groups. We implement a balance check through the standardized difference in means (Rubin 2001, Stuart 2010), which compares, over the covariates, the weighted means of treatment group, $\boldsymbol{\bar{X}_{treatment}} = \frac{\sum_{i\epsilon treatment} \omega_i \boldsymbol{X_i}}{\sum_{i\epsilon treatment} \omega_i}$, and of control group, $\boldsymbol{\bar{X}_{control}} = \frac{\sum_{i\epsilon control} \omega_i \boldsymbol{X_i}}{\sum_{i\epsilon control} \omega_i}$. Here $\boldsymbol{X_i}$ is a 1*M dimensional vector of pre-treatment observed covariates of unit i and ω_i is the sample weight for unit i, computed based on the estimated propensity scores.

For variable X^m (m=1, 2, …, M), the standardized differences in means are calculated by the following

$$d^m = \frac{\bar{X}^m_{treatment} - \bar{X}^m_{control}}{\sqrt{\frac{s^2_{treatment} + s^2_{control}}{2}}}$$

where $s^2_{treatment} = \frac{\sum_i \omega_i}{(\sum_i \omega_i)^2 - \sum_i (\omega_i)^2} \sum_i \omega_i (X_i^m - \bar{X}^m_{treatment})$ with $i \in treatment$ and $s^2_{control} = \frac{\sum_i \omega_i}{(\sum_i \omega_i)^2 - \sum_i (\omega_i)^2} \sum_i \omega_i (X_i^m - \bar{X}^m_{control})$ with $i \in control$ indicate the weighted sample variance in treatment and control group, respectively.

The balance check is then implemented by observing the absolute values of standardized differences in means, $|\mathrm{d^m}|$. An absolute standardized difference below 10% (i.e., 0.1) is often considered an indication of negligible sample imbalance (Austin and Stuart 2015). Results in Table 2 show that the imbalances from weighted samples are negligible after PSW method.

Table 2 Propensity Score Weighting Validation: Covariates Balance Check

VARIABLES	Weighted Mean in Treatment	Weighted Mean in Control	Standardized Difference
IMAGE_QUALITY	0.28	0.27	0.051
IMAGE_COUNT	14.19	13.99	0.022
PRICE	176.55	179.98	-0.018
MAX_GUESTS	3.38	3.33	0.023
REVIEW_SCORE	81.46	79.54	0.055
REVIEW_COUNT	18.26	17.65	0.031
MINIMUM_STAY	2.41	2.47	-0.04
RESPONSE_RATE	92.26	93.41	-0.075
RESPONSE_TIME (seconds)	18176.17	16696.48	0.074
CANCELATION_POLICY	2.44	2.44	-0.003
NUM_BEDROOMS	1.24	1.26	-0.047
NUM_BEDS	1.75	1.76	-0.005
PROPERTY_TYPE	6.98	6.82	0.021
ROOM_TYPE	1.41	1.39	0.026

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 $DEMAND_{it}$ denote the demand for property i in period t, which can be modeled as

$$\begin{aligned} DEMAND_{it} = {} & INTERCEPT + \alpha_1 TREAT_i + \alpha_2 AFTER_{it} + \alpha_3 (TREAT_i \cdot AFTER_{it}) \\ & + \rho PRICE_{it} + \lambda CONTROLS_{it} + PERIOD_t + PROPERTY_i + \varepsilon_{it} \end{aligned} \quad (1)$$

where $TREAT_i$ equals 1 (0) if property i belongs to the treatment group (control group). $AFTER_{it}$ equals 1(0) if period t is after (before) the period when property i was first observed to be treated. Let the interaction term $TREAT_i \cdot AFTER_{it} = TREATIND_{it}$, then the treatment status indicator, $TREATIND_{it}$, equals 1 if property i has received treatment in period t, and equals 0 if otherwise. The key coefficient α_3 estimates the percentage change in property demand caused by having verified photos, compared to property demand without verified photos. ε_{it} is a random shock in period t on property i's demand, assumed to follow an i.i.d. normal distribution. The vector $CONTROLS$ represents a set of control variables that may be correlated to property demand, e.g., the property rules and the consumer reviews [5]. We include the property fixed effect

[5] 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 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_POLICY*, 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

term, $PROPERTY_i$, 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 $PERIOD_t$, which captures any seasonality pattern in property demands. Note that after adding the two fixed effect terms $PROPERTY_i$ and $PERIOD_t$ to the demand equation, the coefficients α_1 and α_2 are absorbed by the fixed effect terms and become unidentifiable. Thus, we rewrite the main model with fixed effect terms

$$DEMAND_{it} = INTERCEPT + \alpha_3 TREATIND_{it} + \rho PRICE_{it} + \lambda CONTROLS_{it} + PERIOD_t + PROPERTY_i + \varepsilon_{it} \quad (2)$$

Table 3 Difference-in-Difference Model: The Impact of Verified Photos on Property Demand

VARIABLES	Main DiD Model (Equation 2)	
	ESTIMATES	**t-Statistics**
TREATIND	**3.687*****	**13.95**
REVIEW_COUNT	0.152***	29.22
REVIEW_SCORE	0.0978***	34.17
MAX_GUESTS	-0.334	-1.83
RESPONSE_RATE	0.0147	1.67
RESPONSE_TIME (minutes)	-0.00446***	-11.51
MIN_STAY	-0.223***	-9.91
SECURITY_DEPOSIT	-0.00148***	-7.31
CANCELATION_POLICY	-0.0203	-0.21
SUPER_HOST	0.808*	2.49
BUSINESS_READY	0.742	1.30
DESCRIPTION_LENGTH	0.0144	0.61
PRICE (instrumented)	-0.0522**	-3.22
INTERCEPT	15.00***	8.97
Fixed Effect	Yes	
Seasonality	Monthly	
Num. Observations	122707	
R-squared	0.5842	

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

As shown in Table 3, the estimated coefficient of the key variable *TREATIND*, as anticipated, suggests a positive significant treatment effect. Specifically, under this model, having verified photos would lead to an extra property booking of 3.687 points. Since, on average an untreated property is booked 21.057% of the days in a month in the pre-treatment period (see Table 1), if getting verified photos, on average this property would be booked (21.057+3.687)% of the time. This corresponds to an increase of 3.687%/21.057%=17.51% in demand as a result of getting verified photos on average. In terms of number of days, this corresponds to 3.687/100*365 = 13.46 days additional booking in a year. In terms of revenue, treatment brings in

has (1) business-related amenities or not (0), and *DESCRIPTION_LENGTH*, the number of words in the (host-created) description of the property that appears on the listing page.

\$187.30/day*13.46days = \$2521.1 more per year in revenue for an average untreated property priced at \$187.30 per night on average.

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.

Falsification Checks on the Pre-treatment Trends

The validity of the causality of the DiD approach (Equation (2)) relies on a critical assumption of pretreatment 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 the pre-treatment trends assumption is the relative time model, with the inclusion of pre-treatment periods. Following the extant literature (e.g., Wang and Goldfarb 2017), we implement the falsification checks by decomposing the pretreatment periods into a series of dummies of the periods prior to the treatment-- $PRE_{it}(j)$:

$$\begin{aligned} DEMAND_{it} = {} & INTERCEPT + \alpha_3 TREATIND_{it} + \sum_j \beta_j (PRE_{it}(j) \cdot TREAT_i) \\ & + \rho PRICE_{it} + \lambda CONTROLS_{it} + PERIOD_t + PROPERTY_i + \varepsilon_{it} \end{aligned} \tag{3}$$

We set the period prior to the treatment month as the reference period (by 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 β_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 β_j cannot be positive and significant for validating the DiD model.

Table 4 reports the estimated results for the falsification test. As can be seen, the coefficients for the period dummies β_j are statistically not significantly different from zero. The set of β_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 4 Falsification Checks on Pre-treatment Trends: Relative-time Model

VARIABLES	Main DiD Model (Equation 2)	
	ESTIMATES	**t-Statistics**
*PRE (-4) * TREAT*	-0.137	-0.31
*PRE (-3) * TREAT*	0.176	0.31
*PRE (-2) * TREAT*	-0.052	-0.10
*PRE (-1) * TREAT* (Reference Month)	--	--
TREATIND	3.742***	9.78
REVIEW_COUNT	0.152***	29.37
REVIEW_SCORE	0.0976***	38.95
MAX_GUESTS	-0.337	-1.85
RESPONSE_RATE	0.0148	1.67
RESPONSE_TIME (minutes)	-0.00447***	-11.52
MIN_STAY	-0.223***	-9.93
SECURITY_DEPOSIT	-0.00149***	-7.30
CANCELATION_POLICY	-0.0196	-0.20
SUPER_HOST	0.806*	2.48
BUSINESS_READY	0.737	1.29
DESCRIPTION_LENGTH	0.0126	0.51
PRICE (instrumented)	-0.0514**	-3.16
INTERCEPT	14.95***	8.92
Fixed Effect	Yes	
Seasonality	Monthly	
Num. Observations	122707	
R-squared	0.5842	

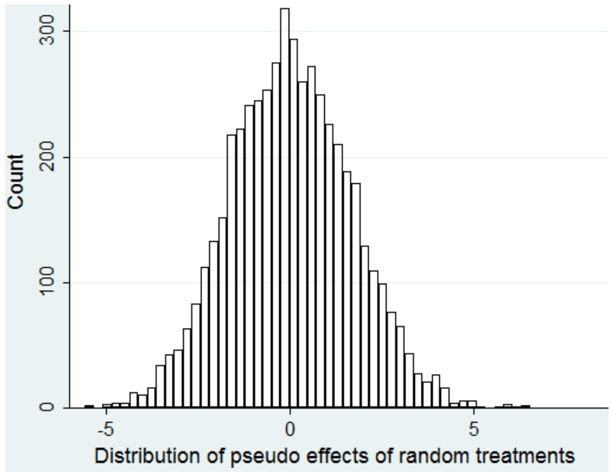
* p<0.05 ** p<0.01 *** p<0.001

Random (Shuffled) Treatment Test

We implement a placebo test through random treatment analysis 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 and then (randomly) reassign the indicators to the units in our sample. 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 to expect the mean of the estimation results to be insignificant.

As can be seen in Figure 11, the distribution of the estimated coefficients of the pseudo (shuffled) treatment indicators suggest that we can reject, at the 95% confidence level, the hypothesis that the estimated effect of the Airbnb professional photos on property demand was driven by the serial correlation in demand or spurious effect (Bertrand et al. 2004).

Figure 11 Shuffled Treatment Test: Distribution of Pseudo Effects



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).

The logic of Rosenbaum bounds analysis is as follows. Suppose the participation probability of unit i is $P_i(treated_i = 1|\boldsymbol{x_i}, \boldsymbol{u_i}) = f(\boldsymbol{\beta}\mathbf{x_i} + \boldsymbol{\gamma u_i})$, where $\mathbf{x_i}$ and $\mathbf{u_i}$ are vectors of observed and unobserved variables, respectively. γ is 0 if there are no unobserved variables affecting treatment selection process. For two units i and j with $\mathbf{x_i} = \mathbf{x_j}$, they have the same probability of receiving the treatment if and only if $\gamma(\boldsymbol{u_i} - \boldsymbol{u_j}) = 0$. 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 provided in an Appendix to this paper), 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 50%. 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.

4.5 The Impact of Verified Photos on Property Demand: Impact of Image Quality

To identify the contribution of image quality to the effect of verified photos, we include predicted image quality and estimate the following demand equation:

$$DEMAND_{it} = INTERCEPT + \alpha_3 TREATIND_{it} + \gamma_1 IMAGE\_QUATLITY + \gamma_2 IMAGE\_COUNT + \rho PRICE_{it} + \lambda CONTROLS_{it} + PERIOD_t + PROPERTY_i + \varepsilon_{it} \quad (4)$$

Table 5 reports the estimates from the model where we have added *IMAGE_QUALITY* and *IMAGE_COUNT* as additional control variables. The decrease in the magnitude of estimated treatment effect, combined with the positive significant coefficient of variable *IMAGE_QUALITY*, suggests that one source of the treatment effect is the high quality of the verified property photos. Specifically, the coefficient for *IMAGE_QUALITY* is 3.016. This indicates that if a property with all low-quality images replaces the images with all high quality images from external professionals, it will be booked 3.016% of the days more. In terms of revenue, this corresponds to 3.016%*365days*$187.30/day= $2,061.90 additional revenue per year, for a property in a year and is priced at $187.30 per night on average.

Table 5 Difference-in-Difference Model: The Impact of Verified Photos on Property Demand

VARIABLES	Model Controlling for Image Quality (Equation 4)	
	ESTIMATES	**t-Statistics**
TREATIND	**1.518*****	**4.16**
IMAGE_COUNT	0.0711***	6.07
IMAGE_QUALITY	**3.016*****	**5.91**
REVIEW_COUNT	0.154***	29.37
REVIEW_SOCRE	0.0946***	32.78
MAX_GUESTS	-0.332	-1.82
RESPONSE_RATE	0.0165	1.86
RESPONSE_TIME (minutes)	-0.00434***	-11.18
MIN_STAY	-0.214***	-9.50
SECURITY_DEPOSIT	-0.00161***	-7.91
CANCELATION_POLICY	-0.0123	-0.13
SUPER_HOST	0.867**	2.67
BUSINESS_READY	0.976	1.71
DESCRIPTION_LENGTH	0.00763	0.32
PRICE (instrumented)	-0.0479**	-2.96
INTERCEPT	12.81***	7.58
Fixed Effect	Yes	
Seasonality	Monthly	
Num. Observations	122707	
R-squared	0.5845	

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

From Table 3, the coefficient of treatment status indicator *TREATIND* is 3.687. In Table 5 where we control for image quality, it reduces to 1.518. This implies that (3.687-1.518)/3.687 = 58.83% of the treatment effect is due to the high image quality of verified photos. However, there is a significant residual impact of the treatment even after controlling for image quality.

4.6. What Makes a Good Airbnb Image?

In this section, we quantify the economic impact of the 12 interpretable lower level image attributes. This analysis would also help us identify the cause of the residual impact of treatment indicator after controlling for image quality. If the residual impact of treatment indicator is due to the presence of the verification seal, it should stay significant even after we control for the 12 interpretable image attributes. In contrast, if the residual impact is due to systematic difference along the 12 interpretable image attributes in images taken by Airbnb photographers versus high quality images taken by external photographers, then the treatment effect should become insignificant after we control for these attributes.

4.6.1 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 6 summarizes the list of 12 key attributes and the brief description for each.

Table 6 List of 12 Image Attributes and the Descriptions

COMPONENT	ATTRIBUTE		DESCRIPTION
Composition	1	Diagonal Dominance	Evaluates how close the key region in an image is placed to the diagonals.
	2	Visual Balance Intensity	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.
Color	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.
	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.

Figure-Ground Relationship	10	Size Difference	Difference in area between image's figure and ground.
	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 7 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 7 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.	**HQ_UN** High Quality Unverified 69,380 Obs.	**HQ_V** High Quality Verified 72,608 Obs.	**HQ_V V.S. HQ_UN**
		Mean (Std. Dev.)	Mean (Std. Dev.)	Mean (Std. Dev.)	Difference (t-statistic)
Composition[6]	Diagonal Dominance	-0.342	-0.281	-0.236	**0.045*****
		(0.160)	(0.109)	(0.081)	(88.56)
	Visual Balance Intensity	-0.865	-0.774	-0.757	**0.017*****
		(0. 110)	(0.103)	(0.105)	(30.78)
	Visual Balance Color	-59.281	-53.093	-50.096	**2.997*****
		(19.460)	(15.509)	(15.070)	(36.93)
	Rule of Thirds	-0.147	-0.089	-0.089	0.0003
		(0.082)	(0.045)	(0.047)	(1.23)
	Warm Hue	0.738	0.751	0.789	**0.038*****
		(0.230)	(0.208)	(0.181)	(36.77)

[6] 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.

Color	Saturation	59.023	73.942	73.683	-0.259
		(37.528)	(31.300)	(26.929)	(0.87)
	Brightness	136.029	154.212	175.802	**21.590*** **
		(32.488)	(27.558)	(22.593)	(161.75)
	Contrast of Brightness	60.601	58.029	53.996	**-4.033*** **
		(13.628)	(13.056)	(12.990)	(58.33)
	Image Clarity	0.324	0.413	0.595	**0.182*** **
		(0.232)	(0.217)	(0.195)	(166.38)
Figure-Ground Relationship	Size Difference	-0.405	-0.181	-0.140	**0.041*** **
		(0.181)	(0.188)	(0.153)	(45.16)
	Color Difference	23.090	33.054	39.063	**6.009*** **
		(20.056)	(17.552)	(15.580)	(68.29)
	Texture Difference	0.043	0.057	0.059	**0.002*** **
		(0.033)	(0.026)	(0.018)	(16.92)

Standard deviations in parentheses (for the rightmost column: t statistics in parentheses)
* p<0.05   ** p<0.01   *** p<0.001

Table 7 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.

4.6.2 Impact of Interpretable Image Attributes on Property Demand

This section investigates the impact of interpretable image features on 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. A DiD analysis on Airbnb property demand identifies the impact of image attributes via the following model:

$$\begin{aligned} DEMAND_{it} = {} & INTERCEPT + \alpha_3 TREATIND_{it} + \eta_1 IMAGE\_ATTRIBUTES_{it} \\ & + \eta_2 IMAGE\_COUNT + \rho PRICE_{it} + \lambda CONTROLS_{it} + PERIOD_t \\ & + PROPERTY_i + \varepsilon_{it} \end{aligned} \tag{5}$$

Table 8 presents the results from estimating Equation (5).[7] Noticeably, with the effects of image attributes being teased out, the coefficient of the key variable *TREATIND* reduces to 0.351 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 5 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. The coefficient of *TREATIND* can be interpreted as the remaining effect of having verified photos that cannot be explained by image attributes—that is, the effect of the verification. One potential reason as to why verification seal has insignificant effect is that most guests are unaware of the verification seal.

[7] For ease of understanding, we use standardized values for image attributes (variables are normalized to zero-mean and unit-variance).

Table 8 Impact of Image Attributes on Property Demand (Equation 5)

COMPONENT	Image Attribute	VARIABLES	ESTIMATES	t-Statistics
		TREATIND	**0.351**	**1.41**
Composition	1	*DIAGONAL_DOMINANCE*	1.245***	7.05
	2	*VISUAL_BALANCE_INTENSITY*	0.561*	2.26
	3	*VISUAL_BALANCE_COLOR*	0.729**	2.69
	4	*RULE_OF_THIRDS*	0.891***	3.87
Color	5	*WARM_HUE*	0.824**	3.11
	6	*SATURATION*	0.908 ***	5.79
	7	*BRIGHTNESS*	1.592 ***	6.17
	8	*CONTRAST_OF_BRIGHTNESS*	-2.721***	-10.04
	9	*IMAGE_CLARITY*	3.440***	15.21
Figure-Ground Relationship	10	*SIZE_DIFFERENCE*	2.624***	10.22
	11	*COLOR_DIFFERENCT*	2.201***	9.97
	12	*TEXTURE_DIFFERENCE*	0.272*	2.13
Control Variables		*REVIEW_COUNT*	0.172***	31.08
		REVIEW_SCORE	0.109***	36.32
		IMAGE_COUNT	0.0820***	9.10
		MAX_GUESTS	-0.321	-1.66
		RESPONSE_RATE	0.0136	1.55
		RESPONSE_TIME (minutes)	-0.00448***	-12.04
		MIN_STAY	-0.211***	-9.81
		SECURITY_DEPOSIT	-0.00149***	-7.33
		CANCELATION_POLICY	-0.0244	-0.39
		SUPER_HOST	0.812*	2.50
		BUSINESS_READY	0.764	1.34
		DESCRIPTION_LENGTH	0.0144	0.60
		PROPERTY_PRICE (instrumented)	-0.0510**	-3.12
		INTERCEPT	12.29***	7.90
Fixed Effect			Yes	
Seasonality			Monthly	
Observations			122707	
R-squared			0.6027	

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

All image attributes in composition component play significant roles in affecting property demand, with the rule of thirds and diagonal dominance being the most important attributes. For example, the coefficient of variable *DIAGONAL_DOMINANCE* suggests that one standard deviation of improvement in the average image diagonal dominance would lead to an extra annual revenue of roughly (1.245%*365days)*$187.30 USD/day=$851.14 to the host of an average property unit. The coefficient of variable *RULE_OF_THIRDS* suggests an extra property booking of 0.891 points with similar improvement in the rule of thirds. Though less influential compared to diagonal dominance and the rule of thirds, the two visual balance attributes also significantly affect property demand.

In color component, the most potent attribute is *IMAGE_CLARITY*, with one standard deviation of attribute improvement bringing an extra annual revenue of (3.440%*365days)*$162.49/day=$2,040.44 to the host of an average unit. Enhancing the "warmth" of property images with the same improvement will, on average, lead to an increase of 0.824%/21.057%=3.91% in property demand for an average property (unit that is on average booked 21.057% of the days). As for saturation, the coefficient suggests a contribution of 0.908 and suggests an effect of 0.908%/21.057%=4.31% increase in the property booking with the same improvement in the image saturation. Image brightness, as anticipated, positively impacts property demand. Being the only image attribute that negatively affects property demand, 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. Should the contrast of brightness of property images decrease by 1 standard deviation, the property would receive an extra booking of 2.721%*365 days=9.93 days in a year.

Figure-Ground (F-G) relationship also plays a significant role in affecting property demand. The statistically significant coefficients of the three attribute measurements confirm that a good property image should have a clearly distinguishable figure from its ground. One standard deviation of improvement in the 3 attributes (area, color and texture) will lead to an increase in the annual property booking of 9.58 days, 8.03 days, and 0.99 days, respectively, for 365 days in a year.

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 isolate and identify the effect of image quality from the effect of photo verification. 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 an increase of $2,061.90 in the annual revenue of an average property unit priced at $187.30 per night if the host replaces all low-quality (unverified) photos with high-quality (unverified) photos without joining the photography program. 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 identify the contribution of each attribute to property demand. 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 being the most decisive image 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 investigates the differential effects of property images on property demand in both high-level and lower-level dimensions. 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[8]. 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.

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.

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.

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.

[8] 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.

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-Guptill.

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

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.

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.

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.

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)*

Stuart, Elizabeth A. 2010. Matching methods for causal inference: A review and a look forward. *Stat Sci.* **25**(1) 1–21. doi:10.1214/09-STS313.

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.