Cleanliness is Next to Income: The Impact of COVID-19 on Short-Term Rentals *

October 30, 2021

Abstract

The short-term rental market provides a close to real time signal of how events of regional and national importance can affect the demand for housing. We use Airbnb data from Austin, Texas to empirically investigate the impact of the onset of Corona Virus Disease 2019 (COVID-19) on the short-term rental market. Specifically, we employ a machine learning algorithm to create an extensive cleanliness dictionary to detect whether an Airbnb unit is clean. We use a difference-in-difference specification to value the change in income related to reviewer perceived cleanliness during the COVID-19 pandemic. We find the following results: First, available listings declined by 25% once the pandemic hit and those that remained lost 22% of their income and had occupancy decrease by 20%. Second, properties that were perceived to be clean increased their income by 17.5% and their occupancy by 16.5%, mitigating the negative shock due to COVID-19. Third, rental prices for clean Airbnb listings did not increase after COVID-19. In addition, we study the interaction of Airbnb supply on the long-term rental market during a market decline.

Keywords: COVID-19, Pandemic, Airbnb, Short-Term Rental, Machine Learning, Small Business, Long-Term Rental

JEL codes: *R31*

*Lily Shen, Wilbur O. and Ann Powers College of Business, Clemson University, lilyshencal@gmail.com; Sean Wilkoff, University of Nevada, Reno, swilkoff@unr.edu.

We sincerely thank the helpful comments made by our anonymous referees. Special thanks to Sean Larkin, Max Koch, Nguyen Bao Lam, and Alex Han for their excellent RA work, and the Clemson Creative Inquiry for the generous research funding. All errors are our own.

1 Introduction

On January 31, 2020, Coronavirus Disease 2019 (COVID-19) was declared a public health emergency and deemed a national emergency in the United States (U.S.) in March 2020[1]. For the first time in decades, the federal and state governments had to mitigate the effects of a devastating pandemic. By October 2021, COVID-19 would lead to the death of over 743,000 people in the U.S. with no clear end in sight[2].

COVID-19 has impacted every aspect of the U.S. economy and the short-term rental market is no exception. Even though the short-term rental market had $87 billion in revenue in 2019[3], it is one sector of the real estate market that has mostly been left out of the growing literature documenting the impacts of COVID-19. While Alyakoob and Rahman (2018), Barron et al. (2021), Àngel Garcia-López et al. (2020), and Xu and Xu (2021) study the spillover effect of Airbnb to the regional economy, this paper aims to fill that gap and analyze the impact of COVID-19 on Airbnb in order to highlight the value of cleanliness and its impact on the short-term rental market during the pandemic.

Given the speed at which COVID-19 spread, the short-term rental market is an ideal laboratory to study the value of cleanliness in the real estate market. While the effects of COVID-19 on long-term rentals and housing purchases may have a delayed and dispersed effect on their respective markets, the short-term rental market has little contractual overhead and instantaneous transactions, which make it easy to directly relate market adjustments to events of regional and national importance like the pandemic. Instantaneous transactions include the ability to book or cancel a property. During COVID-19, the ability to cancel an Airbnb reservation became even less costly for occupants. They were able to cancel any

[1]Five months in: A timeline of how COVID-19 has unfolded in the US, USA Today, Written 6/23/2020 https://www.usatoday.com/in-depth/news/nation/2020/04/21/coronavirus-updates-how-covid-19-unfolded-u-s-timeline/2990956001/, Last checked 10/30/2020.

[2]CDC COVID Data Tracker. https://covid.cdc.gov/covid-data-tracker/##cases_casesper100klast7days, Last checked 10/30/2021.

[3]Vacation Rental Industry Statistics. https://ipropertymanagement.com/research/vacation-rental-industry-statistics

reservations on or after March 14, 2020 at anytime before check-in without penalty[4]. This is in contrast to the long-term market rentals or real estate purchases, where the parties cannot cancel contracts or transactions without added fees. Herein, we define the short-term rental market as furnished properties or spaces that are rented for short periods of time, usually by the day. Furthermore, Airbnb rentals are preferable proxies for the short-term rental market compared to hotels. While hotel demand is subject to brand recognition, credit card rewards, and other potentially confounding variables, Airbnb property ownership is more distributed and there is more variation in cleanliness from property to property.

The main variable that we consider in this study is the perceived cleanliness of the Airbnb property. We focus on the impact of perceived cleanliness at the beginning of COVID-19 because there was uncertainty around exactly how COVID-19 was transmitted. There was certainly concern over contracting COVID-19 by touching surfaces covered with the virus or breathing air containing the virus. Therefore, we hypothesize that guests would make choices of where they are staying based on perceived cleanliness during the pandemic. In this study, perceived cleanliness is based on the reviews left by prior occupants. We find perceived cleanliness to be a better proxy for information used by future guests compared to a host's own description of their property. We are able to use the variation in perceived cleanliness between Airbnb listings to estimate the effect of perceived cleanliness on the income and the occupancy of short-term rentals. Using Airbnb data, we estimate the change in income for a short-term rental during the pandemic due to cleanliness and measure the overall impact of COVID-19 on the short-term rental market.

Specifically, we examine Airbnb property data in Austin, Texas from July 2018 to July 2020. By analyzing Airbnb performance between July 2018 to July 2020, we have a clean dataset with variation between perceived cleanliness before and during the pandemic. We choose July 2018 as the start of our time frame because it allows us to have consistent monthly data and it provides a full two years of data. This allows us to examine any

[4] https://www.airbnb.com/help/article/2701/extenuating-circumstances-policy-and-the-coronavirus-covid19, Last checked 10/30/2021.

possible seasonal effects[5]. We do not use any data after July 2020 because Airbnb mandated a regulation requiring a minimum cleaning standard for Airbnb properties, which might reduce the variation in perceived cleanliness, related to COVID-19, between properties[6].

In order to measure the COVID-19 impact on the short-term real estate market, we needed to find a location that was operating with limited policy impact during the pandemic. Austin is an ideal laboratory since the local authorities did not impose any mandatory travel restrictions after COVID-19, allowing Austin to be open to visitors[7]. By far, Austin is the biggest Airbnb market in Texas, accounting for almost 30% of the 1.5 million Airbnb guest arrivals in the state as of 2017. Austin, Texas also represents a fast growing market in short-term rentals. Austin Airbnb rentals witnessed 449,200 guest arrivals in 2017 compared to 325,400 arrivals in 2016, representing a 38% annual growth. Additionally, Austin is considered one of the fastest growing economies in the US that offers many job and business opportunities and is home to a university that attracts local and international students[8]. Therefore, the COVID-19 impact on Austin's regional economics is of interest to many researchers and policymakers.

We employ a difference-in-difference approach to compare Airbnb listings that have and do not have clean reviews before and after the onset of the COVID-19 pandemic. We rely on a novel natural language processing (NLP) method to construct a "cleanliness" dictionary based on reviews left by past Airbnb occupants. Lawani et al. (2019) show Airbnb prices are influenced by the user-generated reviews. We differentiate between clean and not clean

[5]We control for any possible seasonal effects, including the peak and off-peak seasons and UT Austin's commencement in May. The cancellation of commencement the week of May 23rd, 2020 by UT Austin does not drive our results. We find our results are not driven by the cancellation of graduation or the spring and summer classes in 2020.

[6]Our Commitment to Safer Travel: New Health and Safety Mandate, Airbnbś News, written 10/7/2020 https://news.airbnb.com/health-and-safety-mandate/, Last checked 10/30/2021.

[7]Austin Texas Governor's orders, government website https://www.austintexas.gov/department/covid-19-information/orders-rules, Last checked 10/30/2021.

[8]Should You Invest in Traditional Rentals or Airbnb Rentals in the Austin Real Estate Market in 2018, written 7/28/2018 https://www.mashvisor.com/blog/austin-real-estate-market/, Last checked 10/30/2021. According to the Bureau of Labor Statistics, Austin has a location quotient of 0.93 for higher education. This means Austin has other industries besides the University that provide reasons that attract travellers. https://www.bls.gov/regions/southwest/news-release/2021/occupationalemploymentandwages_austin_20210625.htm, Last checked 10/30/2021.

listings indexed by the language pertaining to "cleanliness". We refer to Airbnb properties with a clean review as "perceived clean" and properties without a clean review in a given month as "not perceived clean." The difference-in-difference regression allows us to estimate the value of "cleanliness" in the short-term rental market during the COVID-19 pandemic.

We find the following results: First, available listings declined by approximately 25% once the pandemic hit. Second, there was a 22% decline in income and a 20% decline in occupancy for Airbnb properties due to the pandemic. Economically, this translates into a \$502,113 per month decrease in income[9]. Third, properties that were perceived to be clean increased their income by 17.5% and their occupancy by 16.5%, significantly mitigating the negative impact of COVID-19. We find that being perceived as a clean property increases income relative to not being perceived as clean even more during COVID-19. Fourth, prices for clean units did not increase after COVID-19. Fifth, there is a positive relationship between the number of Airbnb units and the long-term rental market prices in the city, which is consistent with findings in Barron et al. (2021).

These results suggest that Airbnb hosts can increase their revenue by keeping their listing clean during a pandemic. A listing that is perceived clean will encourage more occupancy and lead to higher revenue. Additionally, we examine alternative theories as to what this change might represent. We find that during the COVID-19 period, prices for "clean" listings did not increase, suggesting the change in income is due to the increase in occupancy. Based on prior reviews, occupants choose to stay in places they perceive to be clean. Our results contribute to the literature on health effects as a priced real estate factor. Sun et al. (2019) and Chay and Greenstone (2005) show that improvement in air quality is likely to stimulate demand for housing in urban cities. We add on to this literature by examining the impact of perceived cleanliness on Airbnb performance during a pandemic.

Our paper also contributes to the growing literature on the impact of COVID-19 on

[9]The decrease in income is a short-term effect due to the reduction in occupancy. We find that the price of Airbnb rentals has not significantly changed over our sample period. We do not know the long term impact of the pandemic on Airbnb rental incomes because the pandemic is still ongoing. In the long run prices and quantity of Airbnb units may readjust to a new equilibrium.

the local economy (Coulson et al., 2021, Glaeser et al., 2021, Brough et al., 2021, Agarwal et al., 2020, Ascani et al., 2020, Chen et al., 2020, and Ling et al., 2020). In addition, our results are inline with extant findings on the impact of Airbnb on long-term rentals. Barron et al. (2021) find that an increase in Airbnb listings leads to an increase in rents of 0.018%, which translates to a $9 increase in monthly rent. We examine the impact of Airbnb on the long-term markets during a negative shock to Airbnb supply. We find that a decrease in Airbnb listings corresponds to a 0.7% decrease in monthly long-term rents, which translates to a monthly decrease of $9.8. We also contribute to the large literature on big data by using machine learning to create an extensive dictionary based on hundreds of thousands of Airbnb reviews, which allows us to identify properties perceived as clean.

The paper is structured as follows. Section 2 provides a detailed description of the data, which includes summary statistics. There is a discussion about the empirical methodology in Section 3. The main results are presented in Section 4. Robustness tests are discussed in Section 5. A discussion of the interaction of the short-term rental market and long-term rental market is provided in Section 6. Section 7 concludes the paper.

2 Data and Descriptive Statistics

2.1 InsideAirbnb

Our paper uses Airbnb data, provided by InsideAirbnb, for short-term rentals occurring between July 1, 2018 and July 31, 2020 in Travis County Austin, Texas. The majority of Austin, Texas sits in Travis County. The Airbnb data includes more than 98% of Airbnb listings in the city of Austin. As previously mentioned, this time period and location allows us to directly examine the short-term rental market without COVID-19 related minimum cleaning mandates from Airbnb or mandatory travel restrictions from the locality.

According to InsideAirbnb's website, their data comes from "public information compiled from the Airbnb website and the reviews for each listing." We use two file sets extracted by

InsideAirbnb each month: Airbnb listings and reviews. We cleansed, verified, analyzed and aggregated the data.

The Airbnb listing data includes key property and host characteristics that allow us to separate the effect of COVID-19 on Airbnb host income from other factors. The variables we control for include: property location, number of bedrooms, number of bathrooms, property type, type of space, Superhost status, allowable number of guests the property accommodates, speed of booking, verified host, cancellation policy, and Airbnb Scores. Summary statistics of these variables are provided in Section 2.2. Detailed definitions for income and other variables can be found in Appendix: Table 1. Overall, the final data set includes 97,092 year-month listings of 6,460 unique properties.

The Airbnb reviews data includes each unique timestamped review for each Airbnb Listing. The data includes the listing ID, a reviewer ID, the date of the review, and the exact comment left by the occupant. We use the comments left by occupants to create an indicator for whether or not a listing is viewed as clean in a given month[10].

For this analysis, we filter down our dataset and use only active listings. We follow Zervas et al. (2017), which requires the following criteria to qualify as active: 1) a property listing exists in the listing data and 2) a review must exist during the sample period. Thus, we only include Austin property listings in our dataset with at least one review in the July 2018 - July 2020 time period. Our sample consists of houses and apartments.

2.2 Descriptive Statistics

Figure 1 shows the geographical distribution of active Airbnb property listings spread throughout Austin. There are 6,460 unique properties on the map that make up a total of 97,092

[10]Only occupants that stayed at the Airbnb can leave a review but not every occupant does leave a review. Although, in an unreported analysis, we find there is no shift in the tendency of guests to leave a review from before the pandemic to during the pandemic, there might be a sample selection bias in the form of the type of person that leaves a review overall. We follow the current literature including such papers as Barron et al. (2021) and Chen et al. (2020) in assuming no sample selection bias in the data; however, any potential sample selection bias would skew the data away from us finding a result. Overall, even with these potential biases and conservative assumptions, we find statistically significant results.

year-month listing observations. The properties are spread throughout the city with the largest density of properties centered around downtown Austin and the University of Texas at Austin.

Table 2 provides the descriptive statistics of all the Airbnb listings in our sample. The average property has approximately two bedrooms, one and a half baths, and accommodates approximately five people. Over 78% of the properties are entire space listings as opposed to the 22% which are shared space listings. Houses make up 62% of the sample. About half of the observations belong to Superhosts. A Superhost is an Airbnb designation that denotes a host with good response time and good ratings[11]. The average price is $215 with a standard deviation of $273. The average income per month is $754, which is estimated monthly by multiplying the average daily price by the occupancy. Occupancy is calculated as one night plus twice the number of reviews received. This estimate is based on statistics provided by InsideAirbnb, where 50% of bookings in Austin provide a review[12]. Every listing is active, so we assume that every listing had at least one booking. This also accounts for listings that had no reviews, but still may have had a booking that did not leave a review. Additionally, to provide a conservative estimate on income, we assume each booking is for one night[13]. The average property gets 1.6 comments per month, which translates into each property being occupied for an average of 3.2 days[14].

We divided our year-month sample into pre- and during COVID-19 cohorts. Approximately 17% of the sample listings occur during COVID-19, which is in line with expectations; the COVID-19 period takes up 20% of the sample's time period. Using a data dictionary on customer reviews (described in greater detail in Section 3.1), we labeled perceived clean and not perceived clean properties based on monthly property reviews. We find that 28%

[11]https://www.airbnb.com/help/article/828/what-is-a-superhost, Last checked 10/30/2021.

[12]http://insideairbnb.com/austin/, Last Checked 10/30/2021.

[13]Unfortunately, the actual length of stay of each guest is confidential data that Airbnb does not release. We follow similar methods common in the literature: Barron et al. (2021), Valentin (2021), Àngel Garcia-López et al. (2020), Chen et al. (2020), and Wyman et al. (2020). These papers assume each review translates to a stay of a minimum of three nights. We are more conservative and only assume one night per review.

[14]See Appendix: Table 1 for variable definitions.

of the observations were year-month listings that were described as clean by reviewers. For future reference, we shall refer to properties that were classified as perceived clean as "clean properties" and properties that were not perceived clean as "not-clean properties". As will be discussed in detail, not-clean properties do not necessarily mean the properties were reviewed as unclean. They should be viewed as the set of properties where the algorithm classification was outside the set of clean properties because "cleanliness" was not indicated in the reviews. Our variable of interest in the regressions is COVID×Clean, which makes up only 2.3% of the properties. This implies that not many properties were considered clean during the COVID-19 time period.

Table 3 provides the summary statistics of sample listings segmented by pre/during COVID-19 and by cleanliness classification. In Section 4.1, we discuss the similarities between perceived clean properties and not perceived clean properties. In terms of variables, we see in Table 3 that on average the two groups, both pre- and during COVID-19, have similar property and booking characteristics (e.g. bedrooms, bathrooms, accommodation, space, house, and cancellation policy). One main difference between perceived clean and not perceived clean cohorts is in the percentage of overall Superhosts, which not surprisingly are a larger fraction of the perceived clean properties. Furthermore, the properties with clean reviews also receive more reviews on average both before COVID-19 and during. We provide a comparison of income, occupancy, and price in Figures 3, 4, and 5, respectively, and provide a detailed analysis of their trends pre-COVID-19 in Section 4.1.

Figure 2 shows the number of active Airbnb listings available each month during our sample time period. The number of properties listed started to decline around March 2019. The COVID-19 pandemic drastically decreased the number of listed properties on Airbnb in Austin by over 25%. The active listings went from a peak of approximately 4,422 in March 2019 to 2,831 in July 2020.

3 Empirical Methods

In this paper we examine the impact of COVID-19 on the short-term rental market using a difference-in-difference framework. COVID-19 provides an exogenous shock to both regional and national short-term rental markets. The first case of COVID-19 in the U.S. was confirmed on January 21, 2020, but it was not until March 13, 2020 that America declared a national emergency[15]. Since COVID-19 was downplayed in the U.S. until March, we use March 2020 as the month the event began. All observations prior to March 2020 are pre-COVID-19, and any observations during or after March 2020 are during COVID-19.

Our research design follows three steps. First, we use a machine learning (ML) natural language processing (NLP) algorithm to build a Clean dictionary of words that appeared in Airbnb reviews. Second, based on the appearance of "clean" or "dirty" words used in each Airbnb review, we create an indicator variable (Clean) for each Airbnb listing each month. Finally, we estimate the impact of perceived cleanliness on Airbnb operational outcomes during COVID-19 using a hedonic framework.

In order to estimate the effect of COVID-19 on the short-term rental market, we fit the following difference-in-difference model:

$$ln(Outcome_{it}) = \alpha + \gamma_1 Clean_i + \delta COVID_t + \beta_{it}(COVID_t \times Clean_{i,t}) + X_{it}'\theta + \epsilon_{it} \quad (1)$$

where $Outcome_{it}$ is one of three different outcome variables we examine for property i in month t including: *Income*, *Occupancy*, and *Price*, all of which are defined in more detail in Appendix: Table 1. $X_{i,t}$ is a matrix of controls that include physical characteristics of the properties and booking characteristics along with year and location (zip code) fixed effects, θ is a vector of the corresponding coefficients, and ϵ_{it} is a property month error term. We cluster standard errors by zip code to control for location unobservables.

[15]Five months in: A timeline of how COVID-19 has unfolded in the US, USA Today, Written 6/23/2020 https://www.usatoday.com/in-depth/news/nation/2020/04/21/coronavirus-updates-how-covid-19-unfolded-u-s-timeline/2990956001/, Last checked 10/30/2021.

Our results focus on three variables: $COVID_t$, $Clean_{i,t}$, and $COVID_t \times Clean_{i,t}$. $COVID_t$ takes on the value of 0 if the observation is pre-COVID-19 and a value of 1 if the observation occurs during COVID-19. $Clean_i$ takes on the value 1 if the listing has a cleanliness Score (Score) of 1 or greater based on reviews left in month t-1 and 0 otherwise[16]. $COVID_t \times Clean_{i,t}$ is the interaction between COVID and Clean. A listing that has a positive Score after COVID-19 would be 1 and 0 otherwise. The coefficient on $COVID_t \times Clean_{i,t}$ is the impact of a listing being perceived as clean on the outcome variable. $COVID_t \times Clean_{i,t}$ is the variable of interest in this study. Section 3.1 discusses the detailed construction of the machine learning driven cleanliness indicator (Clean) based on guest reviews.

There are numerous variables besides cleanliness that affect income, occupancy, and price for Airbnb rentals. Therefore, we include a matrix of controls X_{it} that include physical characteristics such as: bedrooms, bathrooms, capacity, property type, space type, and booking characteristics such as: instant booking, Superhost, verified host, the cancellation policy, and Airbnb Scores. These variables are all defined in Appendix: Table 1.

One condition for the validity of a difference-in-difference model is that each dependent variable has a parallel trend between the treatment and control group prior to the event date. We test income, occupancy, and price to make sure we have parallel trends prior to COVID-19. We use a placebo test based on the same difference-in-difference specification above, but instead of the full sample we use observations from July 2018 through July 2019 and treat March 2019 as the event date.

3.1 Clean Dictionary

Our ability to measure the impact of cleanliness on income, occupancy, or price will be determined by our cleanliness measure. The cleanliness indicator we create is crucial to our identification strategy to identify the impact of COVID-19 on the short-term rental market. Therefore, it is important that we can identify which listings would be perceived as clean by

[16]Score is defined in Appendix: Table 1

occupants. In order to create an indicator for whether a listing is perceived as clean, we first need to identify an extensive list of words that imply clean or dirty in the reviews. Typically, building a customized dictionary from a large textual database is labor-intensive and time consuming (Loughran and McDonald, 2011).

In recent years, the advancements in machine learning natural language processing algorithms have made the timely construction of an accurate and customized dictionary possible. Natural language processing (NLP) algorithms use quantitative approaches to process human language. In this paper, we rely on an NLP algorithm to build a dictionary of words and phrases that are used in Airbnb reviews to indicate that a unit is clean. We follow the method introduced in Mikolov et al. (2013) and Li et al. (2020) to create a cleanliness dictionary based on reviews in our data. Mikolov et al. (2013) introduce a neural network model that measures associations between words using the context in which they appear. For every seed word provided by the researchers, the algorithm is able to return an extensive list of synonyms that appeared in the textual database. The mathematical derivation of such an NLP model and its relevance to the widely used maximum likelihood model in economics are detailed in Shen and Ross (2020) and Ambrose et al. (2020).

Using this method and all guest reviews in our data, we create two dictionaries using two seed words: clean and dirty. Each seed word provides us with a list from which we select the most relevant and related words. These lists are comprised of singular words, two word (bigrams), and three words (trigrams). The ultimate word lists and their frequency in the comments is presented in Table 1. The singular words for "clean" includes derivations of clean, spotless, neat, immaculate, tidy, pristine, sanitiz, uncluttered, disinfect, and declutter. The singular words selected from the list generated by dirty are derivations of dirt, dust, stain, mold, cockroach, filth, clutter, gross, mess, disgusting, messy, messed, grimy. These singular words and all their derivations and negations are used when searching through a review.

Table 1 provides the singular words used and the list of negated clean words that are used to create the Score variable. Table 1 also provides the number of reviews that contain the dictionary word or a derivative of them. "Clean" appears in the most reviews with a total of 40,351 reviews. The next two largest are "spotless" and "neat" with 1,528 and 866, respectively. "Dirt" appears in the most reviews of the dirty words with a total of 863 reviews. Out of a total 140,601 reviews left by occupants, 44,509 reviews have a clean word, 2,462 have a dirty word, and 437 have a negated clean word. The remaining reviews do not have any clean or dirty words.

Appendix: Table 2 provides examples of reviews and the score they would receive based on our method. One example of a clean review reads:

"She was very welcoming! It was a last-minute trip but she was very helpful and accommodating! The place was very ***clean****."*

Because the word clean appears once in this review and there are no negations or dirty words, this review would have a Score value of 1. In contrast, a dirty word example reads:

"Great location, ***not clean****."*

This review would have a Score of -1 because of the words "not clean". Additional examples and their corresponding Score can be found in Appendix: Table 2.

We create a cleanliness indicator to proxy for the perceived cleanliness of the listing during a given month. The clean and dirty dictionaries allow us to count the number of clean and dirty words received by a listing on a monthly basis. We use all the reviews of a listing in month t-1 to calculate the cleanliness indicator for month t. For a given listing and month, if a word from the clean word dictionary appeared in a review, then the Score increases by one. If a word from the dirty dictionary appeared in a review for that listing, then the Score decreases by one. The clean indicator is defined as one if Score is positive and zero otherwise[17].

[17] We recognize that there are alternative rules to create the Clean indicator, such as counting the number of clean/dirty reviews instead of words. Our results are robust to those alternative measures.

4 Empirical Results

4.1 Parallel Trends

One of the underlying assumptions when employing a difference-in-difference model is that the treatment and control group follow similar trends before the treatment. To validate our difference-in-difference model, we demonstrate that properties perceived clean and properties not perceived clean were similar in all observable variables, including our dependent variables income, occupancy, and price.

Figures 3, 4, and 5 show the median monthly income, occupancy, and price of properties perceived as clean compared to properties not perceived as clean without controlling for any other characteristics. The two groups share the same trend before COVID-19. Figure 3 displays the monthly income. The solid line represents properties perceived as clean, while the dashed line represents the properties not perceived as clean. Before COVID-19 the two lines have the same trend, and they both decrease with COVID-19, but properties perceived as clean rebound significantly more than properties not perceived as clean. The line representing properties perceived as clean increases sharply, while the not-clean properties line remains flatter. Figure 4 outlines the average number of nights Airbnb units perceived as clean were booked compared to the properties not perceived as clean. We find the results for the average number of nights booked, which factors into income, are similar to the results for income. The number of bookings for clean properties increases sharply after COVID-19, and the line for properties without clean reviews remains flat. We examine the trends in pricing in Figure 5 and find that prices, the other factor in income, follow the same trend up until the beginning of COVID-19. The prices appear to be closer after COVID-19, unlike the separation that occurs between income and occupancy. However, these results do not control for any property characteristics, except if a property was perceived as clean or not. The results in Figures 3, 4, and 5 confirm that a difference-in-difference model will provide valid results to examine the impact of COVID-19, while controlling for additional observable and unobservable factors.

4.2 Difference-in-Difference

We use the empirical specification outlined in equation (1) to examine the impact of COVID-19 on short-term rental markets. Specifically, we focus on the impact of COVID-19 and the marginal impact that being perceived as clean has on income, occupancy, and price.

In Table 4 we report our empirical results examining the impact of COVID-19 on Airbnb host's income and occupancy. The dependent variable in Column (1) and Column (4) are income and occupancy, respectively. This baseline specification includes location and year fixed effects as well as property characteristics. Columns (2) and (5) include the same controls as well as additional booking characteristics. The booking controls include: if the property can be instantly booked, if the host has been verified, and if the cancellation policy is strict or moderate. Our explanatory power (R^2) increases as we add in booking controls. Columns (3) and (6) include the same controls as Columns (2) and (5) and adds Airbnb Scores for location, communication, check-in, and listing accuracy. Further detailed descriptions of all variables can be found in Appendix: Table 1. We find that the results are consistent under specifications in Columns (1)–(3) and (4)–(6). Therefore, the rest of the section will focus on interpreting the results in Columns (3) and (6).

Our results indicate that the onset of COVID-19 caused a decrease in income in the short-term rental market. The coefficient estimate for the COVID indicator in Column (3) of Table 4 suggests that during COVID-19, income of Airbnb properties in Austin decreased by 22%. Since our dependent variable is the natural log of income, the coefficients can be interpreted as a percentage change. However, the coefficient first has to be transformed with the following transformation $(e^{coefficient}-1)\times 100$. The 22% decrease is both statistically and economically significant. The average monthly income per listing is \$754, so economically the impact of a 22% decrease in income is \$167 per month per listing. There are approximately 3,000 active listings after COVID-19 began. A conservative estimate that does not account for lost income to the Airbnb listings that stopped being active would be to multiply the $167 per listing by the 3,000 active listings. This would yield an impact on the Austin Airbnb market of $502,113 per month.

Table 4 Column (3) reports a statistically significant coefficient estimate of 1.223 on Clean, which implies that properties perceived as clean have on average a 239% higher income than properties not perceived as clean. Given that properties perceived as clean have a higher average income, the result we are most interested in is the marginal impact of being perceived as clean during the pandemic. The statistically significant coefficient on the COVID×Clean interaction term of 0.162 provides the result. The coefficient on COVID×Clean is statistically and economically significant and translates to an extra 17.5% in income or in dollar terms an extra $132 per month for properties perceived as clean during COVID-19.

Overall, our results suggest that during the pandemic all properties lost income on average. Properties not perceived as clean lost $167 per month while properties perceived as clean saw their income decrease by $35 ($167 minus $132).

Further examination of column (3) shows that the other control variables are all significant except for Airbnb Score. While standard controls help to determine the monthly income, Airbnb Score does not.

After finding that properties perceived as clean were not impacted as greatly as those not perceived as clean during COVID-19, we next explored the channel through which the perceived properties could earn more income than properties not perceived to be clean. Income is calculated by multiplying occupancy by the price, which means that there are only a few ways that income can be affected. The price for perceived clean properties could increase during the pandemic, and occupancy could stay the same. Occupancy could increase, and the price could remain the same, or both occupancy and price could increase. We explore both channels using a difference-in-difference approach with occupancy and price as the dependent variable. The results for occupancy are reported in Table 4 columns (4)-(6)

while the results for price are reported in Table 5.

Table 4 Column (6) reports a statistically significant coefficient of -0.227 during COVID-19, which implies that all properties saw an average decline of 20% in occupancy. We also find that properties perceived as clean have on average 287% larger occupancy compared to the rest of the units based on the Clean indicator coefficient estimate of 1.352. A statistically significant coefficient of 0.153 on the (COVID$\times$Clean) interaction reports the marginal impact of being perceived as clean during COVID-19 on occupancy. This translates to an increase of 16.5% in occupancy for properties perceived as clean over properties not perceived as clean. The result implies that occupants are even more likely to choose properties perceived as clean than they were before COVID-19. This suggests that the marginal increase in income to properties perceived as clean came at least partially from the increased occupancy. Next, we turn to the analysis of Airbnb prices to investigate if occupancy alone is responsible for the marginal increase in income during COVID-19 for properties perceived as clean.

In Table 5 columns (1)-(3) we examine the results from a model with the same specifications as in Table 4 but with price as the dependent variable. Similar to Table 5, we find that the different model specifications across columns (1)-(3) do not affect the significance of our results. Overall, Airbnb prices in Austin decreased. The coefficient on the COVID indicator is -0.024 and is statistically significant, which implies that Airbnb prices decreased by 2.37%. However, we find that the coefficient on COVID$\times$Clean, 0.009, is not statistically significant. Our results suggest that prices went down on average for all Airbnb properties in Austin and being perceived as clean did not impact the price the property charged. This implies that the increase in income to properties perceived as clean is not due to a marginal difference in price, but is a result of the marginal increase in occupancy. Therefore, we find that during the pandemic, occupants worried more about their exposure to COVID-19 and chose to stay in properties viewed as clean instead of taking a risk and staying in not-clean properties.

Furthermore, examining the coefficient on the wide range of property controls shows our price estimates are consistent with those reported in Lawani et al. (2019). For example, Table 5 shows the coefficient for an additional bedroom is 0.163 and is statistically significant at the 1% level. The hedonic price coefficient reported in Lawani et al. (2019) is 0.168 and is statistically significant at the 1% level. Chen et al. (2020), which examines the impact of COVID-19 and lockdown on Airbnb hosts in Sydney, Australia, find that Airbnb host income decreased by 89.5%, between January 2020 and August 2020, and the number of Airbnb properties actively listed declined by 82%. The authors focus on Sydney which was under lockdown during their sample period. In contrast, we examine the impact of COVID-19 on the Austin, Texas Airbnb market. As Austin, Texas did not have any lockdown or other restrictions, we measure the impact of COVID-19 without government restrictions on Airbnb properties. Additionally, we show that during the pandemic, prices for properties did not increase in the short-term rental market.

5 Robustness

In Section 4.2, we document that the positive and statistically significant interaction effect (COVID×Clean) on monthly income and occupancy is due to the reduced risk to contract COVID-19. In this section, we implement several robustness exercises to address the possible concerns that our main findings might have been driven by other confounding factors.

5.1 Placebo Test

In order to eliminate the possibility that our results are a product of confounding seasonal factors that coincided with COVID-19, we conduct a placebo test to examine the data over the same months but a year prior. Specifically, we perform three exercises to estimate our primary specifications on observations between July 2018 to July 2019 and change the beginning of COVID-19 to March 2019. We use the period July 2018 to July 2019 as a placebo time frame for COVID-19 so that the placebo event takes place at the same time of year as our model regression. We use the same time frame in the placebo regression to account for possible seasonal effects. We choose this time frame instead of earlier years because by examining the change in prices from March 2019 through July 2019, we are able to replicate the same change in months as our standard specifications except for one year prior. This eliminates the idea that the change in income was due to a seasonal effect that would have happened in March 2020 regardless of COVID-19. In addition, if we find no statistically significant result for the placebo test, it is further support that we meet the parallel trends criteria for our difference-in-difference model.

Table 6 displays the placebo test results. The model specifications for Columns (1)-(3) in Table 6 correspond to those in Column (3) and Column (6) in Table 4 and Column (3) in Table 5, respectively, except that we changed the beginning of COVID-19 and the data sample period. Columns (1), (2), and (3) display results using a subsample of the Airbnb data from a time (July 2018 to July 2019) that predates COVID-19. The coefficients that are displayed correspond to the placebo COVID indicator for March 2019, the Clean indicator, and the variable of interest COVID×Clean. The control variables have been suppressed to focus on our results.

Column (1) in Table 6 displays the results for Ln(Income) as the dependent variable. We find that the coefficient on the Placebo COVID indicator is significant and positive, implying that the income increased from July 2018 through July 2019. We also find a positive significant coefficient for clean properties which confirms that properties perceived as clean receive more income than properties that are not perceived as clean on average. The coefficient on COVID×Clean is not statistically significant. This implies that there is no marginal benefit to being perceived as clean between March 2019 and July 2019, relative to not being perceived clean from July 2018 to March 2019. Additionally, this suggests that the income of properties perceived as clean and properties not perceived as clean follow the same trend from July 2018 to March 2020. This provides further support for the validity of our difference-in-difference results with income as the dependent variable in Section 4.

The results of the additional regressions with the same specification for occupancy and price are displayed in Columns (2) and (3), respectively. Similar to Column (1) we find that the coefficient on the Placebo COVID indicator in Column (2) is positive and significant. This implies that occupancy for Airbnb properties in Austin increased between July 2018 and July 2019. Based on the coefficient for the Clean indicator, we also find that properties perceived as clean have more nights booked on average than properties not perceived as clean. The coefficient on the COVID×Clean indicator is not statistically significant. This finding implies that the difference between the occupancy for properties perceived as clean and properties not perceived as clean stayed the same statistically from July 2018 through July 2019. The prior result supports the validity of our occupancy analysis in Section 4.

Our results reported in Column (3) are similar to our results in Columns (1) and (2). The coefficient estimate on the Placebo COVID indicator in Column (3) is positive and significant but only at the 10% level. All previous coefficients described were significant at the 1% level. We find that the coefficient estimate on the COVID×Clean interaction is not statistically significant. This implies that the difference between the price of properties perceived as clean and properties not perceived as clean stayed the same from July 2018 through July 2019. The prior result provides support for the validity of our difference-in-difference regression with price as the dependent variable in our main results section.

The takeaway message from the statistical insignificance of the placebo interaction term (COVID×Clean) in Columns (1)-(3) is that our main results are not from a seasonal effect that occurs between March and July every year. If that was the explanation, then we would expect to find a statistically significant result for our placebo test. Our results are consistent with our explanation that properties perceived as clean had an increase in their income during COVID-19. Overall, our robustness results provide further support that our results are valid and that the results are not caused by some seasonal feature found between March and July.

5.2 Alternative Channels

In this section we provide analyses to address two potential factors that might be driving our results: the impact of UT Austin cancelling commencement ceremonies in 2020 and hotel vacancy.

5.2.1 UT Austin Graduation

In May 2020, UT Austin cancelled its graduation, which may have impacted the demand and price for Airbnb units and ultimately the income derived from rentals in that month. It is worth mentioning that although UT Austin is a large university, the city of Austin is much more than a university town. The city is home to many businesses across numerous industries and is an entertainment destination for many visitors. According to the Bureau of Labor Statistics, Austin has a location quotient of less than one (0.93) for higher education, indicating that travellers visit Austin for more than just the university[18]. However, to formally eliminate the concern that our results are driven by the cancellation of UT Austin's graduation in May 2020, we conduct several robustness tests to control for the graduation effect. Specifically, we perform three alternative regression specifications for each of our outcome variables: income, occupancy, and price.

Table 8 shows the results for the alternative specifications for monthly income and occupancy. Our results show that while there is a weakly significant premium associated with UT Austin's graduation month, this premium does not affect our results. The regression specification in columns (1) and (5) of Table 8 correspond to the regression specification in columns (3) and (6) in Table 4, respectively. They represent the difference-in-difference regression specifications without controlling for UT Austin's graduation. In order to control for a possible graduation effect, the first alternative regression specification shown in column (2) and (6), for dependent variables of income and occupancy, respectively, include an

[18]https://www.bls.gov/regions/southwest/news-release/2021/occupationalemploymentandwages_austin_20210625.htm, Last checked 10/30/2021.

additional control variable, a dummy for UT Austin's graduation month. The graduation dummy variable takes on the value of 1 if the observation comes from May of 2019 or May of 2020, and 0 otherwise[19]. We find that the coefficients on (COVID$\times$Clean) that represent the marginal impact of being perceived as clean during COVID-19 on income and occupancy are 0.163 and 0.153, respectively, and are statistically significant. These coefficients are the same magnitude and statistical significance of our results without the inclusion of a graduation dummy.

The second specification applied to income and occupancy are represented by columns (3) and (7), respectively. The second specification is our original specification but excludes observations that occurred during May 2019 or May 2020, reducing the number of observations in the sample. Again, our primary results are robust, suggesting that the cleanliness premium is not driven by graduation months.

The third specification, represented in columns (4) and (8) for income and occupancy, respectively, is the same as the regression specification in columns (3) and (6) in Table 4, respectively with the exclusion of observations that occurred during the cancelled graduation in May 2020. Our results remain robust to this third specification[20].

The same three alternative specifications, described above, controlling for the graduation effect are applied to the price outcome variable. The results are shown in Table 9. Our result that price is not impacted during COVID-19 holds even when controlling for a graduation effect. The magnitude of the coefficient on (COVID$\times$Clean) in all columns of Table 9 are the same and all are statistically insignificant.

In all three alternative specifications, we find that our baseline results hold. The coefficients of interest are highly statistically significant and suggest that the increase in premium for cleanliness during COVID-19 is not caused by UT Austin's graduation.

[19]We use UT Austin's graduation month and not week due to the fact that our analysis is run at the monthly level.

[20]We cannot use a dummy in this case because May 2020 is during COVID-19 and therefore is colinear with our COVID dummy.

5.2.2 Hotel Vacancy

As an additional robustness check, we consider the possibility that hotel vacancies may impact Airbnb host income, occupancy, and price. In order to rule out that hotel occupancy drives our result we run additional analyses. We first examine the trend in hotel occupancy over our sample period. Figure 6 shows that the trend in hotel occupancy rate decreases in response to COVID-19. We see a similar pattern in Airbnb occupancy over the same period as shown in Figure 4. Table 7 provides descriptive statistics of hotel occupancy and natural log of hotel occupancy over our sample period. Hotel occupancy is reflected as the percentage of hotel rooms that are occupied in a given month. On average, hotel occupancy was 59.26% over our sample period from August 2018 through November 2020 but reached a low in April 2020 of 24.50%. While hotel occupancy did increase after April, the rate stayed below pre-COVID-19 rates.

The results of our additional analyses including hotel occupancy can be found in Tables 10 and 11. The analysis in Table 10 includes two specifications in addition to our baseline specifications in column (1) and (4) of Table 10 taken from columns (3) and (6) of Table 4. Specifically, we perform two alternative regression specifications for each of the outcome variables: income, occupancy, and price. The first specification takes our standard framework and includes a variable measuring the hotel occupancy rate. We find that the coefficient on hotel occupancy is statistically significant and 0.011 and 0.010 for income and occupancy, respectively. We also find that the coefficients on (COVID$\times$Clean) in columns (2) and (5), which represent the marginal impact of being perceived as clean during COVID-19, on income and occupancy are 0.137 and 0.133, respectively, and statistically significant. Even though we find that hotel occupancy impacts Airbnb income and occupancy, hotel occupancy does not affect the magnitude or significance of the "Clean" premium. The results show that the premium on COVID-19 is insignificant. This is due to the fact that the decline in hotel occupancy is correlated with COVID-19. We also find that there is still no impact on Airbnb price after controlling for hotel occupancy in column (2) of Table 11. The second hotel occupancy specification replaces hotel occupancy rate with the natural log of the hotel occupancy rate.

Our results in Table 10 and Table 11 show that hotel occupancy or the natural log of the hotel occupancy rate are related to Airbnb bookings, host income, and Airbnb price. In all specifications, hotel occupancy and the natural log of hotel occupancy are positive and statistically significant. However, we note that our main findings remain qualitatively and quantitatively similar with the inclusion of the hotel occupancy variables. Our results align with the literature that examines the connection between hotel occupancy and Airbnb overall.

The takeaway message from our hotel vacancy robustness check is that our main results are not due to or impacted by hotel vacancy. If hotel vacancy was responsible for our "Clean" premium, then we would expect to find a statistically insignificant or different magnitude result for our income and occupancy regressions. Our results are consistent with our explanation that properties perceived as clean had an increase in their income during COVID-19.

6 Discussion: Airbnb and Long-Term Rent

Since Airbnb was founded, research has sought to examine the impact of Airbnb on the long-term rental market (See e.g., Horn and Merante, 2017, Àngel Garcia-López et al., 2020, and Barron et al., 2021). The extant literature finds that an increase in available Airbnb units increases long-term rents across a variety of different markets. For example, in the most comprehensive study spanning the entirety of the U.S., Barron et al. (2021) find, at the median owner-occupancy rate zip code, that an increase in the number of Airbnb listings by 1% corresponds to a 0.018% increase in long-term rents. Similarly, Coulson et al. (2020) find that reduced supply of lont-term rental units will eventually increase local rent prices. Further, extant studies examine the impact of short-term rentals on the long-term rental market during periods in which the number of available Airbnb units were increasing. In contrast, our study allows us to briefly examine what happens to the long-term rental market as the supply of Airbnb properties declines, especially following the onset of COVID-19. Thus, we provide some insight into the relationship between short-term rentals and the long-term rental market during a period in which the short-term rental market experienced a negative shock.

Table 7 provides descriptive statistics of the number of Airbnb units, monthly percentage change in Airbnb, monthly long-term rents in Austin, and the monthly percentage change in the long-term rents over our entire sample. We used the Zillow rent index for Austin as the source of our long-term rents[21]. The zillow rent index is the same data used in the literature examining Airbnb impact on long-term rents. We see that, on average, the supply of Airbnb properties decreased by 1% per month and had a maximum monthly decrease of 12% over the sample period. Monthly rent in Austin is an average of $1,409 over our sample period. The largest percentage drop in Austin rents is 14% over our sample period.

In order to estimate the effect of Airbnb supply on the long-term rental market, we fit the following regression model:

$$\Delta Rent_t = \alpha + \beta_t \Delta Units_t + X_t' \theta + \epsilon_t \tag{2}$$

where $\Delta Rent_t$ is the monthly percentage change in long-term rent for Austin, Texas in month t. $\Delta Units_t$ is the monthly percentage change in Airbnb units in month t. X_t is a matrix of controls that include year fixed effects, θ is a vector of the corresponding coefficients, and ϵ_t is the error term. We cluster standard errors by year to control for yearly unobservables.

Figure 7 displays the fitted regression lines between the change in rent (%) and the change in local Airbnb supply with a 95% confidence interval. The top three panels display the yearly breakdown of the regression, and the bottom panel display analysis employing data for the entire sample period. All of the four linear fits show that as the number of

[21] https: //www.zillow.com/research/zillow-rent-index-methodology-2393/, Last checked 10/30/2021.

Airbnb units decreases, the rental prices also decline.

Our analysis results for Equation 2 can be found in Table 12. Table 12 shows four different specifications for testing the effect of Airbnb listings on long-term rents. The first two specifications do not control for year fixed effects. We do find a statistically significant coefficient of 0.033 on the monthly change in Airbnb units. However, these coefficient estimates do not control for any yearly unobservables. Therefore, we focus our analysis on specifications in columns (3) and (4) that add on year fixed effects. We find that no matter how we cluster standard errors the power of our variable of interest stays statistically significant. We find the coefficient on our variable of interest is 0.07% and inline with Barron et al. (2021). Barron et al. (2021) find that an increase in Airbnb listings leads to an increase in rents of 0.018%, which translates to a $9 increase in monthly rent[22]. Our results suggest that for every 1% change in monthly Airbnb units we see a 0.7% increase in monthly long-term rents. Given that the average monthly rent in Austin is $1,409, an increase of 0.7% translates to a $9.8 increase in long-term rents. The positive relationship also indicates that as the Airbnb supply decreased during our sample period, rents in the long-term rental market also decreased.

These results represent a first pass at analyzing the interaction between the short-term rental market and the long-term rental market during a market decline. Given that the pandemic is still ongoing at the time of this paper, we cannot say where the new rental equilibrium will be. This provides a great avenue for further research as more data become available.

7 Conclusion

As of the writing of this paper, there have already been over 45 million cases of COVID-19 in the U.S. with new cases emerging everyday[23]. People continue to travel and the U.S. is more

[22]We conduct four specifications that control for different clustering methodologies and year fixed effects. Our results remain qualitatively similar regardless of the specification.

[23]https://covid.cdc.gov/covid-data-tracker/##cases_casesper100klast7days, Last checked 10/30/2021

open than it was at the beginning of the COVID-19 pandemic. Airbnb has acknowledged the severity of the spread of COVID-19 and mandated that all Airbnb hosts follow an enhanced clean protocol by October 2020, emphasizing the importance of cleanliness in the short-term rental market[24].

We document a negative impact of COVID-19 on the short-term rental market and the importance of cleanliness during a pandemic. We use a natural language processing algorithm to create a novel dictionary to detect Airbnb cleanliness based on guest reviews. On average, we find a statistically significant negative 22% difference in income between properties prior to COVID-19 and properties during COVID-19. Economically, this translates to approximately $167 a month per listing or approximately a $502,113 reduction in total monthly income. This reduction in income is not uniform across properties. Properties that are not perceived as clean have on average a $167 per month decrease in income. In contrast, properties that are perceived as clean have a decrease of $35. The reduction in income loss for properties perceived to be clean comes from a marginal increase in occupancy relative to properties not perceived to be clean. We also confirm earlier findings in Barron et al. (2021) about the positive relationship between the supply of Airbnb units and the long-term rental market rents.

[24]Our Commitment to Safer Travel: New Health and Safety Mandate written October 7, 2020 by Airbnb. https://news.airbnb.com/health-and-safety-mandate/, Last checked 10/30/2021.

References

Agarwal, S., B. W. Ambrose, L. A. Lopez, and X. Xiao (2020). Pandemic risk factors and the role of government intervention: Evidence from covid-19 and cmbs mortgage performance. *Available at SSRN 3674960*.

Alyakoob, M. and M. Rahman (2018). The sharing economy as a local economic engine: The heterogeneous impact of airbnb on restaurant employment growth.

Ambrose, B., Y. Han, S. Korgaonkar, and L. Shen (2020). Contractual completeness in the cmbs market: Insights from machine learnin. *Working Paper*.

Ascani, A., A. Faggian, S. Montresor, et al. (2020). The geography of covid-19 and the structure of local economies: the case of italy. *Journal of Regional Science*.

Barron, K., E. Kung, and D. Proserpio (2021). The effect of home-sharing on house prices and rents: Evidence from airbnb. *Marketing Science 40*(1), 23–47.

Brough, R., M. Freedman, and D. C. Phillips (2021). Understanding socioeconomic disparities in travel behavior during the covid-19 pandemic. *Journal of Regional Science 61*(4), 753–774.

Chay, K. and M. Greenstone (2005). Does air quality matter? evidence from the housing market. *Journal of Political Economy 113*(2), 376–424.

Chen, G., M. Cheng, D. Edwards, and L. Xu (2020). Covid-19 pandemic exposes the vulnerability of the sharing economy: a novel accounting framework. *Journal of Sustainable Tourism*.

Coulson, E., S. Brakman, A. Faggian, and Y. Murata (2021). Covid-19 and regional economies: An introduction to the special issue. *Journal of Regional Science 61*(4), 691–695.

Coulson, N. E., T. Le, and L. Shen (2020). Tenant rights, eviction, and rent affordability. *Eviction, and Rent Affordability (July 4, 2020)*.

Glaeser, E. L., G. Z. Jin, B. T. Leyden, and M. Luca (2021). Learning from deregulation: The asymmetric impact of lockdown and reopening on risky behavior during covid-19. *Journal of Regional Science 61*(4), 696–709.

Horn, K. and M. Merante (2017). Is home sharing driving up rents? evidence from airbnb in boston. *Journal of Housing Economics 38*, 14–24.

Lawani, A., M. R. Reed, T. Mark, and Y. Zheng (2019, 3). Reviews and price on online platforms: Evidence from sentiment analysis of airbnb reviews in boston. *Regional Science and Urban Economics 75*, 22–34.

Li, K., F. Mai, R. Shen, and X. Yan (2020, 07). Measuring Corporate Culture Using Machine Learning. *The Review of Financial Studies*.

Ling, D. C., C. Wang, and T. Zhou (2020, 9). A first look at the impact of covid-19 on commercial real estate prices: Asset-level evidence. *The Review of Asset Pricing Studies*.

Loughran, T. and B. McDonald (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance 66*(1), 35–65.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pp. 3111–3119.

Shen, L. and S. Ross (2020). Information value of property description: A machine learning approach. *Journal of Urban Economics*, 103299.

Sun, C., S. Zheng, J. Wang, and M. E. Kahn (2019). Does clean air increase the demand for the consumer city? evidence from beijing. *Journal of Regional Science 59*(3), 409–434.

Valentin, M. (2021, March). Regulating short-term rental housing: Evidence from New Orleans. *Real Estate Economics 49*(1), 152–186.

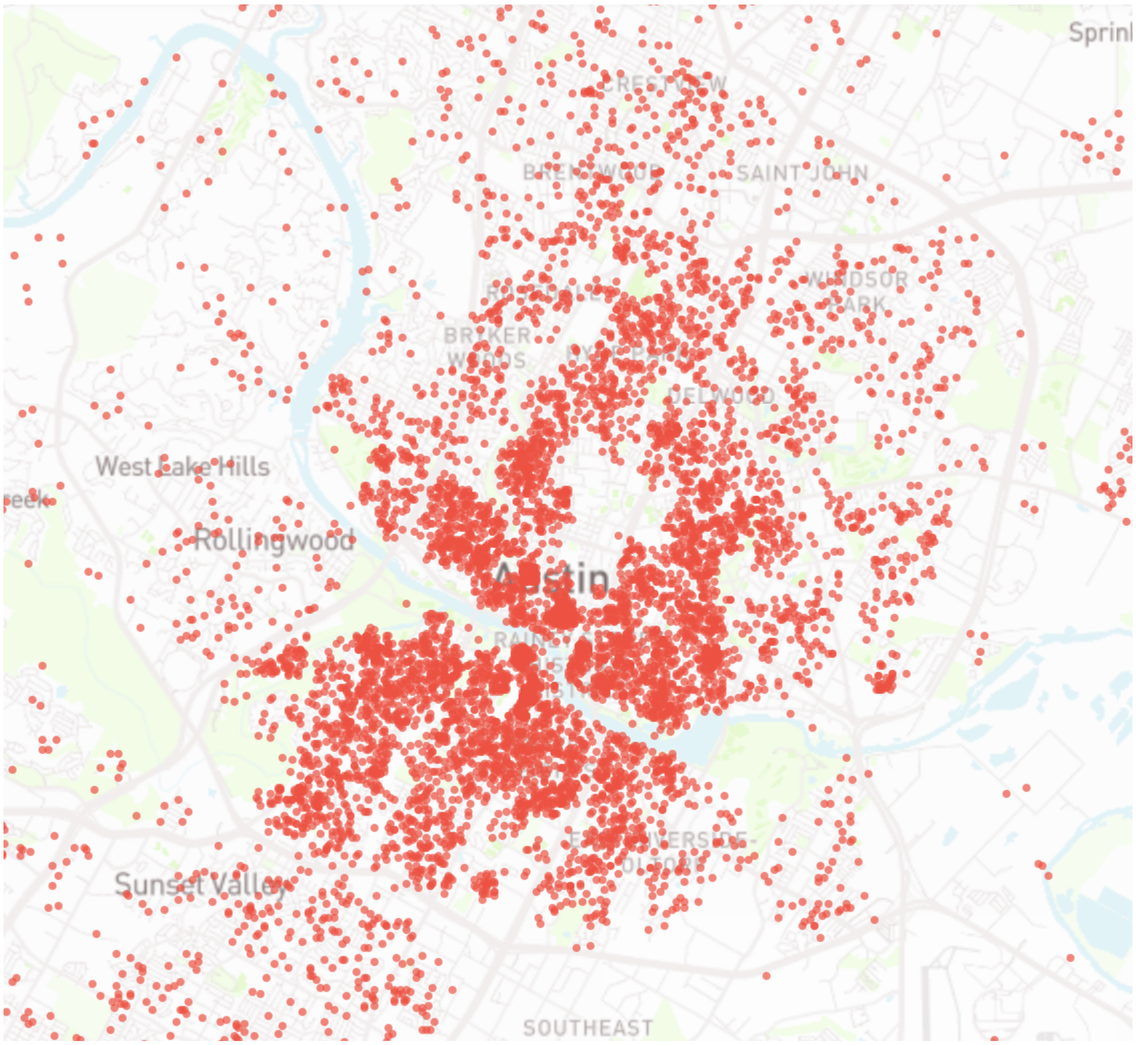
Wyman, D., C. Mothorpe, and B. McLeod (2020). Airbnb and vrbo: the impact of short-term tourist rentals on residential property pricing. *Current Issues in Tourism 0*(0), 1–12.

Xu, M. and Y. Xu (2021). What happens when airbnb comes to the neighborhood: The impact of home-sharing on neighborhood investment. *Regional Science and Urban Economics 88*, 103670.

Zervas, G., D. Proserpio, and J. W. Byers (2017, 10). The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry. *Journal of Marketing Research 54*, 687–705.

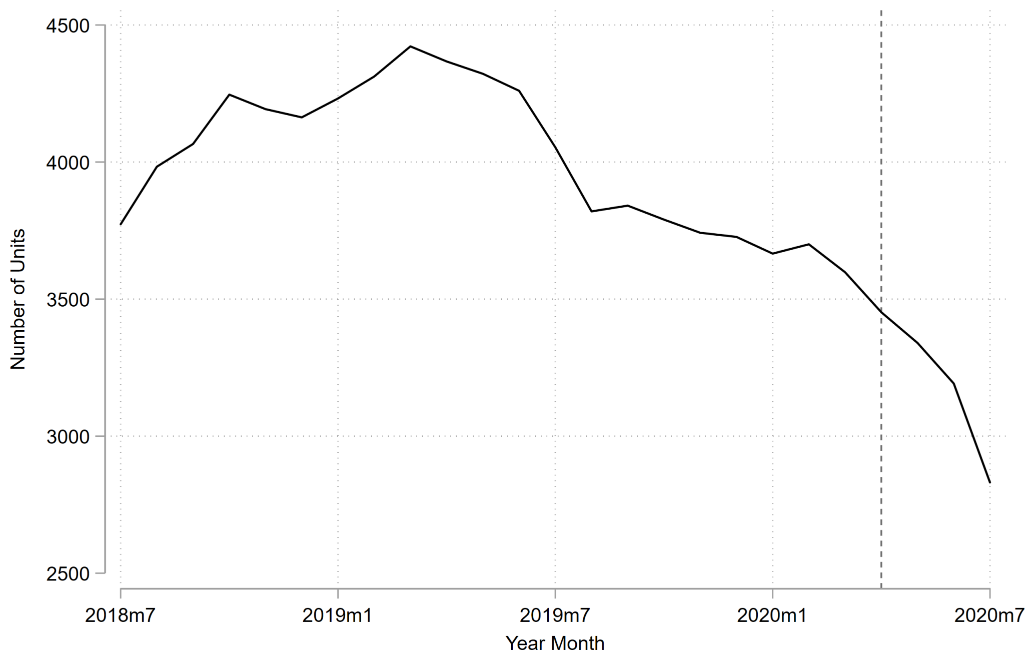
Àngel Garcia-López, M., J. Jofre-Monseny, R. Martínez-Mazza, and M. Segú (2020). Do short-term rental platforms affect housing markets? evidence from airbnb in barcelona. *Journal of Urban Economics 119*, 103278.

Figure 1: Geographical Distribution of the Airbnb Sample



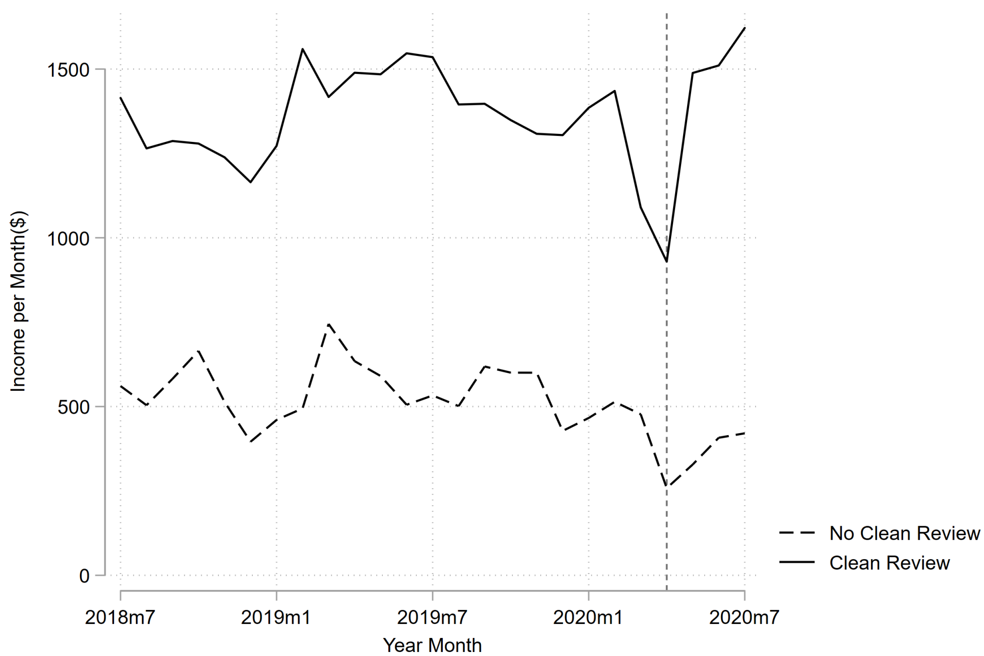
Notes: This figure displays an overview of the geographical distribution of the 6,460 Airbnb listings analyzed in this study. Each dot represents an Airbnb in Travis County, Austin. The sample consists of conventional Airbnb units that are likely run by non-corporate hosts, including houses and apartments. We exclude hotel properties on the Airbnb platform. The center of the map is downtown Austin and the University of Texas at Austin.

Figure 2: Units of Airbnb Available



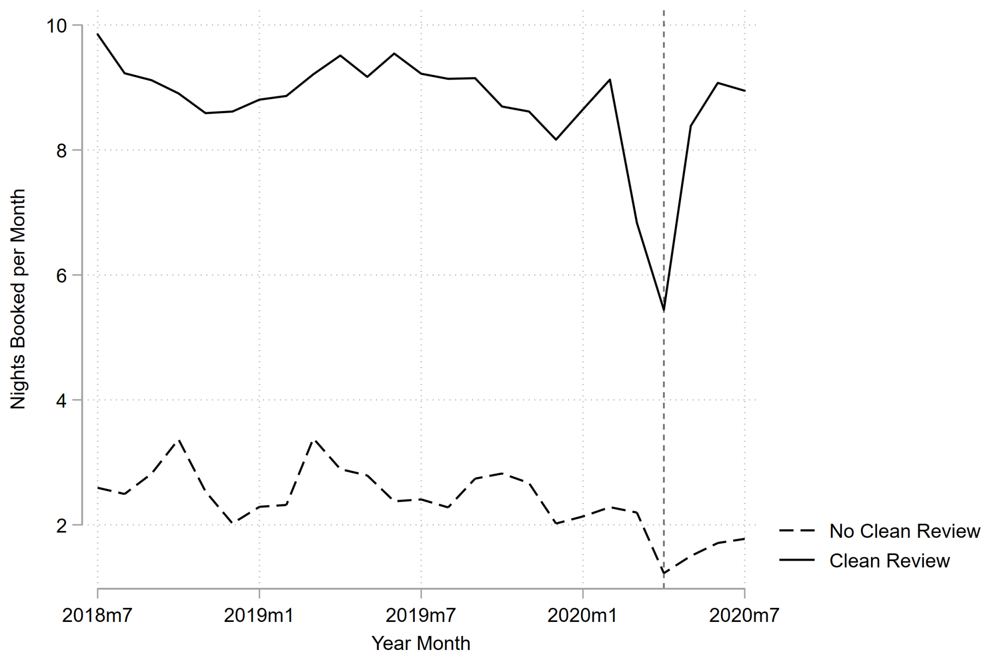
Notes: This figure plots the number of active Airbnb listings each month in Austin from July 2018 to July 2020. The sample consists of conventional Airbnb units that are likely run by non-corporate hosts, including houses and apartments. We exclude hotel properties that are listed on the Airbnb platform. We only include active units, defined as units that have at least one review during our sample period (July 2018 to July 2020). The dashed vertical line represents the month COVID-19 began in the U.S., March 2020.

Figure 3: Compare Monthly Income



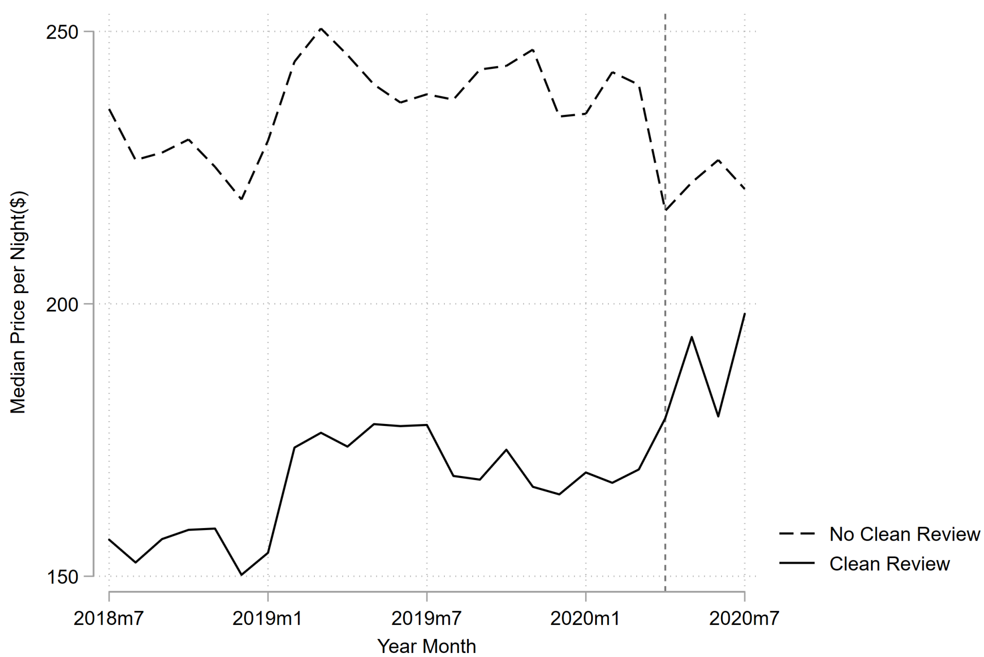
Notes: This figure compares monthly income by whether an Airbnb unit has any clean reviews. The monthly income is calculated using the nightly price and the nights booked in the given month. The dashed vertical line represents the month COVID-19 began in the U.S., March 2020.

Figure 4: Compare Nights of Booking



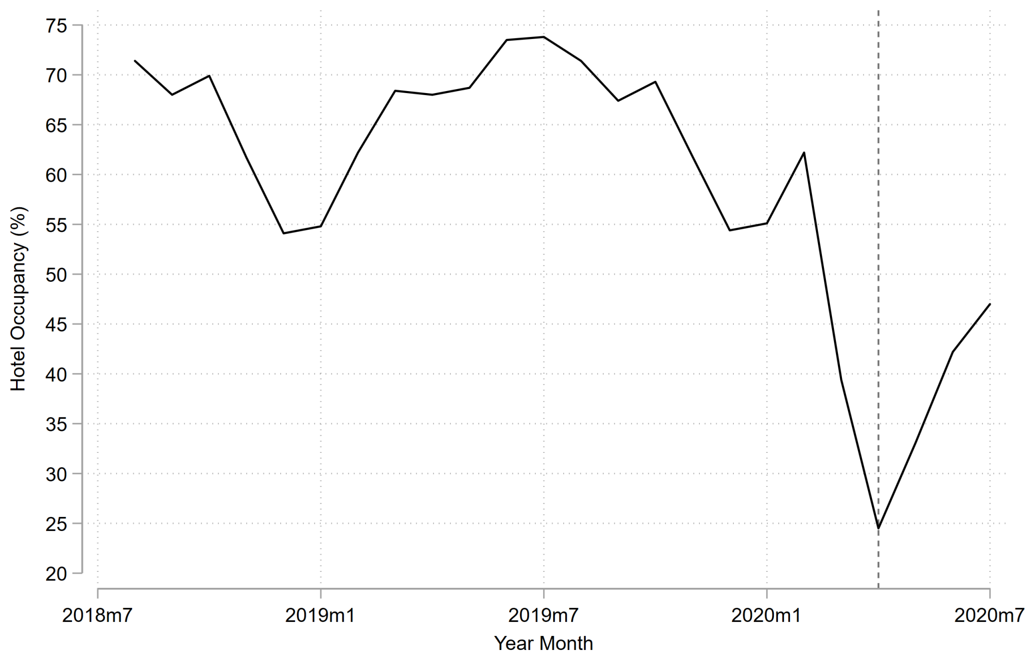
Notes: This figure compares the total nights booked in a month by whether an Airbnb unit has any clean reviews. The total nights booked is estimated following InsideAirbnb statistics that 50 percent of guests in Austin leave reviews. We take a conservative approach and set the duration of the stay to one-night for each review. The dashed vertical line represents the month COVID-19 began in the U.S., March 2020.

Figure 5: Compare Price



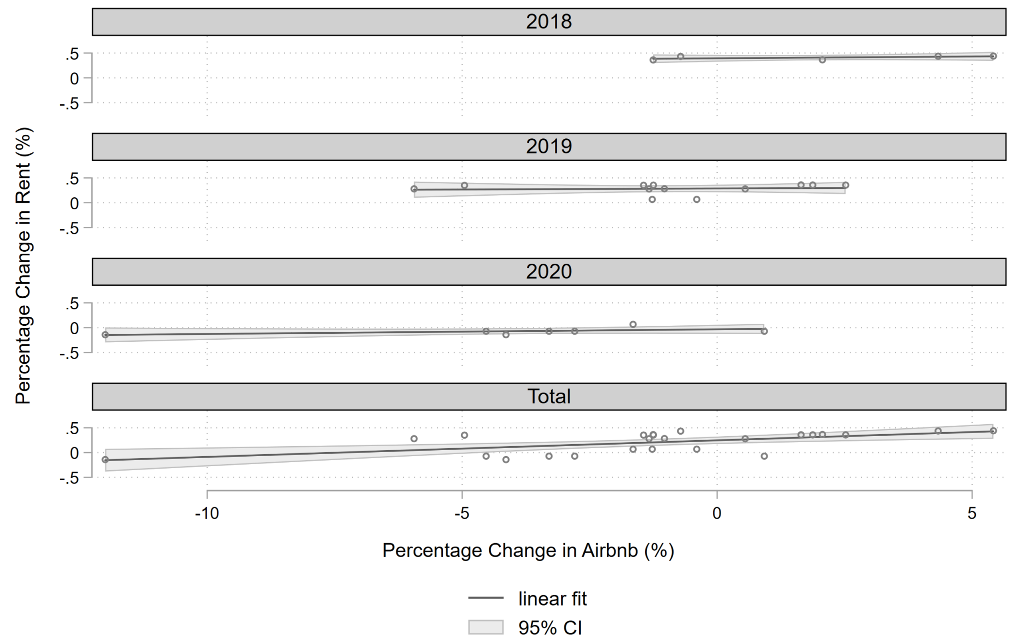
Notes: This figure displays the average daily price for a given month by whether an Airbnb unit has any clean reviews. The dashed vertical line represents the month COVID-19 began in the U.S., March 2020.

Figure 6: Monthly Hotel Occupancy



Notes: This figure displays the percentage of hotel rooms that are occupied by month in the U.S. from August 2018 through November 2020. The data comes from STR. We use nationwide data due to data limitations. We compared the limited Austin hotel occupancy available for April and May 2020 and found it similar to the nationwide occupancy measure. The dashed vertical line represents the month COVID-19 began in the U.S., March 2020.

Figure 7: Relationship between Airbnb Listings and Long-term Rental Market Rents



Notes: This figure displays the fitted regression lines between the change in rent (%) and the change (%) in local Airbnb supply with a 95% confidence interval. The top three panels display the yearly breakdown of the regression, and the bottom panel display analysis employing data for the entire sample period. All of the four linear fits show that as the number of Airbnb decreases, the rental prices also decline.

Table 1: Cleanliness Dictionary

Clean Words	# of Reviews	Dirty Words	# of Reviews	Negated Clean Words (Dirty)	# of Reviews
clean	40531	dirt	863	not clean	189
spotless	1528	dust	338	unclean	60
neat	866	stain	319	not cleaned	56
immaculate	705	mold	201	not been cleaned	41
tidy	660	cockroach	179	little cleaner	30
pristine	96	filth	126	bit cleaner	19
sanitiz	59	clutter	102	hadnt been cleaned	17
uncluttered	30	gross	94	more thorough cleaning	5
disinfect	30	mess	93	not spotless	4
declutter	4	disgusting	89	needs cleaning	3
		messy	28	not disinfect	2
		messed	17	not sanitiz	2
		grimy	13	no vacuum cleaner	1
				cleaned badly	1
				yet been cleaned	1
				hadnt cleaned	1
				needs deep cleaning	1
				need deep cleaning	1
				need cleaning	1
				not immaculate	1
				no disinfect	1
				needs thorough cleaning	0
				need thorough cleaning	0
				special cleaning challenges	0
				no serious cleaning	0
				cleaners hadnt	0
				cleaners havent	0
				complain about cleanness	0
				cleanliness disgusting	0
				overlooked cleaning	0
				overlook cleaning	0
				overlooks cleaning	0
				lack cleaning	0
				lack cleanliness	0
				lacks cleaning	0
				lacks cleanliness	0
				not neat	0
				not tidy	0
				not pristine	0
Total	44509	Total	2462	Total	437

Notes: This table provides a dictionary of the word we used in creating our measure of cleanliness. If a word on the Clean list or any of its derivatives appeared in a review we would keep count. For example, if clean, cleans, cleaner, cleaning, cleanliness, cleaned appear in a review our score for that review would increase by one for each instance that appeared. If a word on the Dirty list of the Negated Clean list or any of its derivatives appeared in a review we would keep count. For example, if dirt or dirty appear in a review our score for that review would decrease by one for each instance that appeared. mess and messy are on the dirty list because it is actually mess with a space. We did not want to pick up anytime the word message appeared which would have happened without adding a space to mess. We therefore had to add messy to pick up that dirty word. After subtracting the number of clean words from the number of dirty or negated clean words that appear the listing was perceived as clean if the total was positive and not perceived as clean otherwise.

Table 2: Property Descriptions: Full Sample

	Mean (1)	Std. Dev. (2)
Number of Bedrooms (Bedroom)	1.811	1.158
Number of Bathrooms (Bathroom)	1.513	0.752
Unit Capacity (Accommodation)	4.748	2.980
Indicator if Entire Space (Space)	0.784	0.411
Indicator if House (House)	0.624	0.484
Price Per Night (Price)	215.3	271.1
Income Per Month (Income)	754.1	1,294
Nights Booked per Month(Nights)	4.232	4.614
Ln(Price)	4.966	0.838
Ln(Income)	5.900	1.218
Ln(Nights)	0.934	0.992
Comments Per Month (Comment)	1.616	2.307
Indicator if COVID-19 Period (COVID)	0.169	0.375
Indicator if Clean Review (Clean)	0.284	0.451
Interaction term (COVID×Clean)	0.0235	0.151
Indicator if Instant Booking (Instant)	0.490	0.500
Indicator if Superhost (Superhost)	0.504	0.500
Indicator if Verified Host (Verified Host)	0.575	0.494
Indicator if Strict Cancel (Strict)	0.479	0.500
Indicator if Moderate Cancel (Moderate)	0.333	0.471
Airbnb Score(Rate Accuracy)	9.850	0.488
Airbnb Score (Rate Check-in)	9.913	0.431
Airbnb Score (Rate Communication)	9.910	0.416
Airbnb Score (Rate Location)	9.813	0.495
Airbnb Score (Rate Value)	9.685	0.616
# Observations	97,092	
# Airbnb Property	6,460	

Notes: This table provides the descriptive statistics of the full Airbnb sample employed in this study. There are 6,460 unique properties on the map that altogether make up a total of 97,092 year-month listing observations. See Appendix: Table 1 for variable definition.

Table 3: Descriptive Statistics by Group

	Before COVID-19				During COVID-19			
	No Clean		Clean		No Clean		Clean	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Bedrooms (Bedroom)	1.87	1.18	1.64	1.08	1.87	1.16	1.80	1.16
Number of Bathrooms (Bathroom)	1.56	0.79	1.39	0.63	1.55	0.77	1.47	0.70
Unit Capacity (Accommodation)	4.79	2.99	4.60	2.89	4.79	3.04	5.15	3.38
Indicator if Entire Space (Space)	0.77	0.42	0.81	0.39	0.79	0.40	0.89	0.31
Indicator if House (House)	0.64	0.48	0.58	0.49	0.64	0.48	0.59	0.49
Price Per Night (Price)	236	297	166	195	225	276	181.97	206
Income Per Month (Income)	540	926	1,382	1,835	371	769	1,341	1,644
Nights Booked per Month(Nights)	2.54	2.72	9.04	5.20	1.66	1.80	7.90	5.15
Ln(Price)	5.03	0.88	4.81	0.72	5.00	0.86	4.89	0.73
Ln(Income)	2.69	3.58	7.79	0.90	1.33	2.83	7.66	0.98
Ln(Nights)	0.56	0.79	2.05	0.57	0.26	0.58	1.88	0.60
Comments Per Month (Comment)	0.77	1.36	4.02	2.60	0.33	0.90	3.45	2.58
Indicator if Instant Booking (Instant)	0.46	0.50	0.56	0.50	0.46	0.50	0.58	0.49
Indicator if Superhost (Superhost)	0.42	0.49	0.66	0.47	0.52	0.50	0.70	0.46
Indicator if Verified Host (Verified Host)	0.58	0.49	0.57	0.49	0.56	0.50	0.51	0.50
Indicator if Strict Cancel (Strict)	0.49	0.50	0.44	0.50	0.49	0.50	0.49	0.50
Indicator if Moderate Cancel (Moderate)	0.31	0.46	0.39	0.49	0.32	0.47	0.36	0.48
Airbnb Score (Rate Accuracy)	9.82	0.54	9.91	0.33	9.84	0.53	9.88	0.41
Airbnb Score (Rate Check-in)	9.90	0.48	9.95	0.27	9.90	0.48	9.94	0.29
Airbnb Score (Rate Communication)	9.90	0.46	9.94	0.26	9.89	0.46	9.93	0.35
Airbnb Score (Rate Location)	9.78	0.54	9.85	0.39	9.85	0.49	9.89	0.36
Airbnb Score (Rate Value)	9.64	0.67	9.78	0.45	9.66	0.65	9.74	0.52
# Observations	55,405		25,274		14,136		2,277	

Note: This table provides the summary statistics of Airbnb listings grouped by pre/duringCOVID-19 and by cleanliness classification. There are 6,460 unique properties on the map that altogether make up a totalof 97,092 year-month listing observations. See Appendix: Table 1 for variable definition.

Table 4: Effect on Monthly Income and Occupancy

	Host Monthly Income			**Monthly Occupancy**		
	Ln(Income)	Ln(Income)	Ln(Income)	Ln(Nights)	Ln(Nights)	Ln(Nights)
	(1)	(2)	(3)	(4)	(5)	(6)
Indicator if COVID-19 Period (COVID)	-0.239***	-0.244***	-0.251***	-0.209***	-0.219***	-0.227***
	(0.029)	(0.029)	(0.028)	(0.024)	(0.025)	(0.026)
Indicator if Clean Review (Clean)	1.286***	1.227***	1.223***	1.441***	1.361***	1.352***
	(0.019)	(0.019)	(0.019)	(0.010)	(0.010)	(0.011)
Interaction term (COVID×Clean)	0.150***	0.157***	0.162***	0.135***	0.147***	0.153***
	(0.025)	(0.026)	(0.025)	(0.021)	(0.020)	(0.020)
Number of Bedrooms (Bedroom)	0.075***	0.085***	0.082***	-0.088***	-0.074***	-0.080***
	(0.020)	(0.019)	(0.019)	(0.017)	(0.015)	(0.014)
Number of Bathrooms (Bathroom)	0.125***	0.143***	0.150***	-0.118***	-0.095***	-0.092***
	(0.028)	(0.029)	(0.028)	(0.014)	(0.011)	(0.011)
Indicator if House (House)	0.061**	0.046	0.044	0.013	-0.007	-0.016
	(0.029)	(0.030)	(0.029)	(0.025)	(0.021)	(0.021)
Unit Capacity (Accommodation)	0.104***	0.094***	0.097***	0.054***	0.045***	0.048***
	(0.010)	(0.010)	(0.010)	(0.003)	(0.004)	(0.004)
Indicator if Entire Space (Space)	0.613***	0.594***	0.594***	0.094***	0.079***	0.084***
	(0.038)	(0.038)	(0.037)	(0.027)	(0.025)	(0.024)
Indicator if Superhost (Superhost)		0.195***	0.188***		0.277***	0.262***
		(0.013)	(0.015)		(0.019)	(0.018)
Indicator if Strict Cancel (Strict)		0.095***	0.095***		0.032**	0.037**
		(0.028)	(0.027)		(0.013)	(0.015)
Airbnb Score (Value)			-0.006			0.035***
			(0.016)			(0.008)
Observations	97,092	97,092	93,173	97,092	97,092	93,173
R-squared	0.578	0.589	0.596	0.515	0.540	0.543
Physical Char	Yes	Yes	Yes	Yes	Yes	Yes
Booking Char	No	Yes	Yes	No	Yes	Yes
Airbnb Score	No	No	Yes	No	No	Yes
Location-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Zipcode	Zipcode	Zipcode	Zipcode	Zipcode	Zipcode

Note: This table reports the results of our difference-in-difference regression examining the impact of COVID-19 on Airbnb monthly income and occupancy. Only selected coefficient estimates of property characteristics controls, booking characteristics controls, and Airbnb Scores are displayed in the table. All models include Zip code and Year fixed effects. Robust standard errors are clustered at the Zipcode level, shown in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

Table 5: Effect on Airbnb Price

Dependent Variables: Ln(Price)	(1)	(2)	(3)
Indicator if COVID-19 Period (COVID)	-0.030**	-0.025**	-0.024**
	(0.013)	(0.011)	(0.010)
Indicator if Clean Review (Clean)	-0.155***	-0.134***	-0.129***
	(0.020)	(0.017)	(0.014)
Interaction term (COVID×Clean)	0.015	0.010	0.009
	(0.011)	(0.011)	(0.012)
Number of Bedrooms (Bedroom)	0.163***	0.159***	0.162***
	(0.024)	(0.023)	(0.023)
Number of Bathrooms (Bathroom)	0.243***	0.239***	0.241***
	(0.025)	(0.026)	(0.026)
Indicator if House (House)	0.048	0.053	0.060
	(0.044)	(0.041)	(0.038)
Unit Capacity (Accommodation)	0.050***	0.049***	0.049***
	(0.010)	(0.010)	(0.010)
Indicator if Entire Space (Space)	0.518***	0.515***	0.510***
	(0.024)	(0.025)	(0.026)
Indicator if Superhost (Superhost)		-0.082***	-0.073***
		(0.019)	(0.017)
Indicator if Strict Cancel (Strict)		0.063***	0.058***
		(0.023)	(0.020)
Airbnb Score (Value)			-0.041***
			(0.010)
Observations	97,092	97,092	93,173
R-squared	0.643	0.648	0.656
Physical Char	Yes	Yes	Yes
Booking Char	No	Yes	Yes
Airbnb Score	No	No	Yes
Location-Year FE	Yes	Yes	Yes
Cluster SE	Zipcode	Zipcode	Zipcode

Note: This table reports the results of our difference-in-difference regression examining the impact of COVID-19 on Airbnb price. Only selected coefficient estimates of property characteristics controls, booking characteristics controls, and Airbnb Scores are displayed in the table. All models include Zip code and Year fixed effects. Robust standard errors are clustered at the Zip code level, shown in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

Table 6: Robustness: Placebo Test

Dependent Variables:	Ln(Income) (1)	Ln(Nights) (2)	Ln(Price) (3)
March to July (Placebo COVID)	0.069***	0.037***	0.032*
	(0.024)	(0.013)	(0.016)
Indicator if Clean Review (Clean)	1.179***	1.315***	-0.136***
	(0.020)	(0.015)	(0.011)
Placebo COVID×Clean	0.023	0.004	0.019
	(0.021)	(0.016)	(0.012)
Observations	51,475	51,475	51,475
R-squared	0.587	0.520	0.673
Physical Char	Yes	Yes	Yes
Booking Char	Yes	Yes	Yes
Airbnb Score	Yes	Yes	Yes
Location FE	Yes	Yes	Yes
Cluster SE	Zipcode	Zipcode	Zipcode

Note: The table shows the placebo analysis results. We estimate our primary specifications on observations between July 2018 to July 2019 and change the beginning of COVID-19 to March 2019. We use the period July 2018 to July 2019 as a placebo time frame to account for any possible seasonal effects. All models include Zip code fixed effects. Robust standard errors are clustered at the Zip code level, shown in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

Table 7: Descriptive Statistics of Airbnb Units, Long-Term Rent, and Hotel Occupancy

	N (1)	Mean (2)	Std. Dev. (3)	Min (4)	Max (5)
Airbnb Units (#)	24	3,888	408.0	2,831	4,422
Monthly Change in Airbnb Units (%)	24	-1.197	3.642	-12.00	5.416
Monthly Austin Rent ($)	24	1,409	23.34	1,361	1,432
Monthly Change in Rent (%)	24	0.207	0.206	-0.140	0.442
Hotel Occupancy (%)	24	59.26	13.45	24.50	73.80
Ln(Hotel Occupancy)	24	4.050	0.278	3.199	4.301

Note: This table provides the descriptive statistics of the analysis data employed in this section. The monthly rents variables are obstained from the Zillow rent index for Austin, Texas during our sample period. https://www.zillow.com/research/zillow-rent-index-methodology-2393/ The hotel occupancy variables comes from STR data.

Table 8: Robustness: Effect of UT Austin Graduation on Monthly Income and Occupancy

	Host Monthly Income				Monthly Occupancy			
	Ln(Income) (1)	Ln(Income) (2)	Ln(Income) (3)	Ln(Income) (4)	Ln(Nights) (5)	Ln(Nights) (6)	Ln(Nights) (7)	Ln(Nights) (8)
Indicator if COVID-19 Period (COVID)	-0.251***	-0.255***	-0.233***	-0.234***	-0.227***	-0.229***	-0.210***	-0.211***
	(0.028)	(0.028)	(0.026)	(0.026)	(0.026)	(0.025)	(0.023)	(0.023)
Indicator if Clean Review (Clean)	1.223***	1.223***	1.225***	1.221***	1.352***	1.352***	1.354***	1.349***
	(0.019)	(0.019)	(0.018)	(0.019)	(0.011)	(0.011)	(0.012)	(0.012)
Interaction term (COVID×Clean)	0.162***	0.163***	0.147***	0.150***	0.153***	0.154***	0.133***	0.137***
	(0.025)	(0.024)	(0.026)	(0.026)	(0.020)	(0.020)	(0.021)	(0.021)
Graduation Dummy		0.021**				0.012		
		(0.010)				(0.009)		
Observations	93,173	93,173	85,727	89,923	93,173	93,173	85,727	89,923
R-squared	0.596	0.596	0.593	0.593	0.543	0.543	0.539	0.538
Physical Char	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Booking Char	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Airbnb Score	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Zipcode	Zipcode	Zipcode	Zipcode	Zipcode	Zipcode	Zipcode	Zipcode

Note: This table reports the results of our difference-in-difference regression examining the impact of COVID-19 on Airbnb monthly income and occupancy. Only selected coefficient estimates of property characteristics controls, booking characteristics controls, and Airbnb Scores are displayed in the table. All models include Zip code and Year fixed effects. The regression specification in (1) and (5) correspond to the regression specification in (3) and (6) in Table 4, respectively. They represent the difference-in-difference regression specifications without controlling for UT Austin's graduation. Regressions (2) and (6) are the same as regressions (1) and (5), respectively, with the inclusion of a dummy variable for the graduation month of May. The graduation dummy takes on a value of one if the observation is in May 2019 or May 2020 and zero otherwise. Regressions (3) and (7) are the same as regressions (1) and (5), respectively, except any observation occurring in May 2019 or May 2020 has been excluded. Regressions (4) and (8) are the same as regressions (1) and (5), respectively, except any observation occurring in May 2020 has been excluded.Robust standard errors are clustered at the Zip code level, shown in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

Table 9: Robustness: Effect of UT Austin Graduation on Airbnb Price

Dependent Variables: Ln(Price)	(1)	(2)	(3)	(4)
Indicator if COVID-19 Period (COVID)	-0.024**	-0.026**	-0.023**	-0.023**
	(0.010)	(0.010)	(0.009)	(0.009)
Indicator if Clean Review (Clean)	-0.129***	-0.129***	-0.129***	-0.129***
	(0.014)	(0.014)	(0.014)	(0.014)
Interaction term (COVID×Clean)	0.009	0.010	0.014	0.013
	(0.012)	(0.012)	(0.011)	(0.011)
Graduation Dummy		0.009***		
		(0.002)		
Observations	93,173	93,173	85,727	89,923
R-squared	0.656	0.656	0.657	0.657
Physical Char	Yes	Yes	Yes	Yes
Booking Char	Yes	Yes	Yes	Yes
Airbnb Score	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Cluster SE	Zipcode	Zipcode	Zipcode	Zipcode

Note: This table reports the results of our difference-in-difference regression examining the impact of COVID-19 on Airbnb price. Only selected coefficient estimates of property characteristics controls, booking characteristics controls, and Airbnb Scores are displayed in the table. All models include Zip code and Year fixed effects. Regression (1) is the same as regression (3) in Table 5. It represent the difference-in-difference regression specifications without controlling for UT Austin's graduation. Regression (2) is the same as regression (1) with the inclusion of a dummy variable for the graduation month of May. The graduation dummy takes on a value of one if the observation is in May 2019 or May 2020 and zero otherwise. Regression (3) is the same as regression (1) except any observation occurring in May 2019 or May 2020 has been excluded. Regression (4) is the same as regression (1) except any observation occurring in May 2020 has been excluded. Robust standard errors are clustered at the Zip code level, shown in parentheses (*** p<0.01, **p<0.05, * p<0.1).

Table 10: Robustness: Effect of Hotel Occupancy on Monthly Income and Occupancy

	Host Monthly Income			**Monthly Occupancy**		
	Ln(Income)	Ln(Income)	Ln(Income)	Ln(Nights)	Ln(Nights)	Ln(Nights)
	(1)	(2)	(3)	(4)	(5)	(6)
Indicator if COVID-19 Period (COVID)	-0.251***	-0.002	0.035	-0.228***	-0.008	0.027
	(0.028)	(0.028)	(0.027)	(0.026)	(0.025)	(0.023)
Indicator if Clean Review (Clean)	1.223***	1.211***	1.213***	1.349***	1.339***	1.340***
	(0.019)	(0.019)	(0.019)	(0.011)	(0.011)	(0.011)
Interaction term (COVID×Clean)	0.162***	0.137***	0.115***	0.156***	0.133***	0.114***
	(0.025)	(0.024)	(0.023)	(0.020)	(0.019)	(0.018)
Hotel Occupancy		0.011***			0.010***	
		(0.001)			(0.001)	
Ln(Hotel Occupancy)			0.575***			0.511***
			(0.025)			(0.027)
Observations	89,637	89,637	89,637	89,637	89,637	89,637
R-squared	0.596	0.599	0.599	0.543	0.547	0.547
Physical Char	Yes	Yes	Yes	Yes	Yes	Yes
Booking Char	Yes	Yes	Yes	Yes	Yes	Yes
Airbnb Score	Yes	Yes	Yes	Yes	Yes	Yes
Location FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Zipcode	Zipcode	Zipcode	Zipcode	Zipcode	Zipcode

Note: This table reports the results of our difference-in-difference regression examining the impact of COVID-19 on Airbnb monthly income and occupancy. Only selected coefficient estimates of property characteristics controls, booking characteristics controls, and Airbnb Scores are displayed in the table. All models include Zip code and Year fixed effects. Regressions (1) and (4) are the same as regressions (3) and (6) in Table 4. They represent the difference-in-difference regression specifications without controlling for hotel occupancy. Regressions (2) and (5) are the same as regressions (1) and (4), respectively, with the inclusion of a hotel occupancy variable that captures the percentage of hotel rooms in Austin, Texas that are occupied monthly. Regressions (3) and (6) are the same as regressions (1) and (4), respectively) with the inclusion of the natural log of hotel occupancy variable that captures the percentage of hotel rooms in Austin, Texas that are occupied monthly. Robust standard errors are clustered at the Zip code level, shown in parentheses (*** p<0.01, **p<0.05, * p<0.1).

Table 11: Robustness: Effect of Hotel Occupancy on Airbnb Price

Dependent Variables: Ln(Price)	(1)	(2)	(3)
Indicator if COVID-19 Period (COVID)	-0.023**	0.007	0.008
	(0.010)	(0.006)	(0.006)
Indicator if Clean Review (Clean)	-0.126***	-0.127***	-0.127***
	(0.014)	(0.014)	(0.014)
Interaction term (COVID×Clean)	0.007	0.004	0.002
	(0.012)	(0.012)	(0.013)
Hotel Occupancy		0.001***	
		(0.000)	
Ln(Hotel Occupancy)			0.063***
			(0.015)
Observations	89,637	89,637	89,637
R-squared	0.655	0.655	0.655
Physical Char	Yes	Yes	Yes
Booking Char	Yes	Yes	Yes
Airbnb Score	Yes	Yes	Yes
Location FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Cluster SE	Zipcode	Zipcode	Zipcode

This table reports the results of our difference-in-difference regression examining the impact of COVID-19 on Airbnb price. Only selected coefficient estimates of property characteristics controls, booking characteristics controls, and Airbnb Scores are displayed in the table. All models include Zip code and Year fixed effects. Regression (1) is the same as regression (3) in Table 5. It represents the difference-in-difference regression specification without controlling for hotel occupancy. Regression (2) is the same as regression (1) with the inclusion of a hotel occupancy variable that captures the percentage of hotel rooms in Austin, Texas that are occupied monthly. Regression (3) is the same as regression (1) with the inclusion of the natural log of hotel occupancy variable that captures the percentage of hotel rooms in Austin, Texas that are occupied monthly. Robust standard errors are clustered at the Zip code level, shown in parentheses (*** p<0.01, **p<0.05, * p<0.1).

Table 12: Robustness: Effect of Airbnb Listings on Long-Term Rental Market Rents

Dependent Variables: Rents	(1)	(2)	(3)	(4)
Monthly Change in Airbnb Units (%)	0.033***	0.033*	0.007*	0.007*
	(0.007)	(0.009)	(0.004)	(0.002)
Observations	24	24	24	24
Adjusted R-squared	0.321	0.321	0.823	0.823
Year FE	No	No	Yes	Yes
Cluster SE	Robust	Year	Robust	Year

Note: This table reports the results of our regression examining the impact of Airbnb rental supply on rents of long-term housing. Models differ by their inclusion of year fixed effects and the type of clustering. The low statistical power is a result of the small sample size. Figure 7 shows that there is a positive trend by year and overall for the relationship between Airbnb supply and long-term rents. Robust standard errors are clustered at the Zip code level, shown in parentheses (*** p<0.01, **p<0.05, * p<0.1).

Table Appendix: Table 1: Variable Definition

Variable	Definition
Bedrooms	Number of bedrooms in the listing
Bathrooms	Number of bathrooms in the listing
Unit capacity	Number of people the listing can accommodate
Entire Space	This is an indicator variable with value 1 if the listing is for the entire property or 0 if the listing is for a shared space.
House	This defines property type and takes the value 1 if the listing is a house and a zero if the listing is an apartment.
Price Per Night	The average daily price of the listing per night
Income Per Month	The price per night multiplied by the occupancy.
Occupancy	The number of nights a listing was booked during the month. Airbnb reports that half of guests leave a review. Therefore, occupancy is two times the number of reviews. If there are no reviews for a listing then occupancy is 1 because there is a 50% chance the host still had one booking and the guest did not leave a comment.
Reviews	An occupant can leave a public reviews regarding their stay and impression of the Airbnb listing for others to see.
COVID-19	An indicator variable that is 1 if the observation is after or on March 1, 2020 and 0 otherwise.
Score	calculated as the difference between number of clean words appearing in the reviews of that single month and number of dirty or negated clean words appearing in the reviews of that single month.
Clean Review	An indicator variable that is defined as 1 if a listing has a positive for a given month and 0 Otherwise.
Instant Booking	An indicator variable that is 1 if an Airbnb allows a guest to book their property instantly without the host reviewing the booking and making a decision.
Superhost	An indicator variable that is 1 if the host has achieved SuperHost status and 0 otherwise. A host becomes a super host by having hosted 10 hosts or completed 3 reservations that total at least 100 nights, maintained a 90% response rate or higher, maintained a 1% cancellation rate or lower, and maintained a 4.8 overall rating.
Verified host	An indicator that is 1 if the host has had their identity verified and 0 otherwise.
Strict Cancel	An indicator that is 1 if the host has a strict cancellation policy and 0 otherwise. A strict cancellation policy is defined as: Free cancellation for 48 hours, as long as the guest cancels at least 14 days before check-in. After that, guests can cancel up to 7 days before check-in and get a 50% refund of the nightly rate, and the cleaning fee, but not the service fee.
Moderate Cancel	An indicator that is 1 if the host has a moderate cancellation policy and 0 otherwise A moderate cancellation policy is defined as: Free cancellation until 5 days before check-in. After that, cancel before check-in and get a 50% refund, minus the first night and service fee.
Rate Accuracy	Rating of the accuracy of the description of the listing on the Airbnb platform.
Rate Check-in	Rating of how welcome the guest felt when he/she first arrived.
Rate Communication	Rating by the guest of an evaluation of how long it takes the host to respond and the accuracy and usefulness of the host's responses.
Rate Location	Rating by the guest of how satisfied he/she is about the location of the property in the neighborhood and its proximity to amenities.
Rate Value	Rating of guest satisfaction with paying the room rate for the service received.
Zipcode	The zip code the listing is in

Table Appendix: Table 2: Sample of Reviews with Cleanliness scores

Reviews	Score
Clean Reviews	
Sylvia was an incredible host. Her space was very clean and tidy. Looked exactly like the photos. Location was fantastic and central to all the things we wanted to experience including Rainey street, South Congress and downtown. Thank you for a great stay. I would highly recommend!	2
A must stay! Elizabeth's place is wonderful and absolutely neat and clean. She responds to messages quickly and makes sure you have whatever you need. Check in process is straightforward. The place is 20-25 mins away from Downtown, Austin. Plenty of restuarants nearby. Highly recommend staying at this Airbnb.	2
She was very welcoming! It was a last-minute trip but she was very helpful and accommodating! The place was very clean.	1
Susan and Anthony are truly superhosts. The place was super clean and efficient, with a spacious bathroom. Best value in Austin that we have found, and we stay in Austin frequently.	1
Kerry is SUPER responsive and easy to communicate with. Place is VERY CLEAN and feels sanitized, a huge plus with Corona going on. Super cute patio facing the street and a nice encolosed backyard. Also is a small park/trail right across the street. Lovely!	2
Dirty Reviews	
I'm afraid this was a very negative experience. Among other things, the bed was dirty, and the bathroom smelled strongly of mold/mildew. We left early due to health concerns.	-2
Great location, not clean.	-1
The location is unbeatable and the apartment building is wonderful. However, the apartment was not clean when I arrived and I would be hesitant to book with Lux Haus again.	-1
No Dictionary Words Reviews	
Great spot in a very convenient location!	0
Tom was a great host. House was excellent! Highly recommend!	0
This place is great if you plan to travel through the heart of Austin. Everything was great :)	0

Notes: This table displays several examples of reviews with cleanliness scores. Total Reviews with Clean Words: 44,509; Total Reviews with Dirty Words: 2,462; Total Reviews with Negated Clean Words: 437; Total Reviews with no Clean or Dirty Words: 98,991; Total Reviews left by Occupants: 140,601. The reviews are presented as written on Airbnb; we did not correct the typos.