Safety Reviews on Airbnb: An Information Tale*

Aron Culotta[†] Ginger Zhe Jin[‡] Yidan Sun[§] Liad Wagman[¶]

March 2023

Abstract

Many online platforms facilitate and publish user reviews in order to build trust between anonymous buyers and sellers. At the same time, platforms can also monitor, filter, and remove certain user reviews, even if they reflect the true experiences of buyers. Using Airbnb and official crime data in five major US cities, we study a December-2019 Airbnb policy that has a potential to remove and discourage reviews about the safety of a listing's vicinity. It is unclear how Airbnb implements this policy, but counterfactual simulation suggests that a complete removal of vicinity safety reviews would hurt guests and increase revenues from reservations on Airbnb, with positive sorting towards listings with such vicinity safety reviews. Conversely, incorporating vicinity safety reviews in a listing's overall ratings or highlighting vicinity safety reviews as if the guest had written a vicinity safety review out of her previous experience would generate opposite effects. Because vicinity safety reviews are more closely correlated with official crime statistics in low-income and minority neighborhoods, our findings suggest that suppressing or highlighting vicinity safety reviews would have different effects on different neighborhoods.

Keywords: Airbnb, safety review, crime, information design, online platform

---

*We are grateful to AirDNA for providing the data and to our home universities for financial support. Marshall Van Alstyne, Matthias Hunold, Xiang Hui, Meng Liu, Peter Coles, Francine Lafontaine, Ying Fan and participants at the Luohan Academy Webinar, Washington University at St. Louis, Boston University, University of Michigan, the 2022 MaCCI annual conference, the 2022 IIOC annual conference, and the 2022 INFORMS Marketing Science virtual conference have provided constructive comments. Tejas Nazare, Nour Ben Ltaifa, and Hunter Petrik provided excellent research assistance. The content and analyses in this paper reflect the authors' own work and do not relate to any institution or organization with whom the authors are affiliated. None of us has a significant financial relationship with Airbnb or any competing short term rental platforms. All errors are our own.

[†]Tulane University. Email: aculotta@tulane.edu.

[‡]University of Maryland & NBER. Email: ginger@umd.edu.

[§]Illinois Institute of Technology. Email: ysun100@hawk.iit.edu.

[¶]Illinois Institute of Technology. Email: lwagman@stuart.iit.edu.

1 Introduction

Information design is crucial for online platforms. Because online platforms derive revenue from the trade they intermediate, they often adopt information mechanisms that allow buyers to discern high- and low-quality sellers. For example, consumer feedback, product recommendations, quality certification, and advertising are widely used by online platforms. This is consistent with the vast economic literature about asymmetric information and market efficiency. Arguably, such mechanisms for conveying information can be even more effective online than offline, because online platforms face fewer physical constraints in space, have a lower barrier to entry, and allow online users to access, search, and utilize the vast pool of information collected by the platform (Einav et al., 2016).

However, there are times when online platforms explicitly choose to *limit* the flow of information, even if that information is likely to be authentic and reflects users' true experiences. For example, in a recent policy change effective December 11, 2019, Airbnb announced that, going forward, guest reviews about listings that include "content that refers to circumstances entirely outside of another's control" may be removed by the platform.[1] This policy change implies that reviews about the safety of a listing's vicinity (henceforth, "vicinity safety reviews") are discouraged and may be subject to deletion by the platform, while guest reviews regarding safety issues within the listed property are still permitted (henceforth, "listing safety review").

It is not obvious whether this limit on vicinity safety reviews is beneficial or detrimental to players on the platform. On the one hand, location is a fixed attribute of any specific listing, and the safety of the vicinity of a listing is usually out of the control of the host. Limiting guest feedback to within-listing safety may motivate hosts to focus on the dimensions they can control and improve. On the other hand, future guests may care about both listing and vicinity safety when they choose which listing to reserve, regardless of whether these safety issues are under the control of hosts or not.

To better understand the impact of this limit on information disclosure, we study detailed data of all Airbnb listings and their reviews in five major US cities (Atlanta, Chicago, Los Angeles, New Orleans, and New York City), from May 2015 to December 2019. This data is collected by AirDNA, a third party that tracks Airbnb listings and listing-specific feedback

[1] See, e.g., `https://rb.gy/0pu5ck` and `https://rb.gy/9y6bum` .

across the US. We use a Lexicon approach to identify vicinity safety reviews and listing safety reviews posted by Airbnb guests. Because our data ends in December 2019, our data largely precede the new Airbnb policy that discourages vicinity safety reviews. As detailed below, we find that 0.51% of guest reviews are related to safety, of which 48.08% are about vicinity safety rather than listing safety. This implies that safety concerns are noticed regardless of whether they relate to the inside of the particular property or its nearby surroundings.

Since guest feedback may reflect guests' subjective opinion of their stay experience, we also obtain (local) government-reported crime statistics for these five cities, by zip code and month during the same period. The data suggest that, as vicinity safety reviews accumulate slowly on Airbnb, the rank correlation between the normalized total count of vicinity safety reviews in a zip code up to a month t and the normalized official crime statistics of that zip code-month is increasing over time. For low-income or minority neighborhoods, the rank correlation can be as high as 0.75 by the end of our sample period (December 2019).

One key question is how much impact these safety reviews have on consumer choice as far as which Airbnb listings to book. If prospective guests do not read or do not care about vicinity safety reviews, it does not matter whether the platform puts any limit on vicinity safety reviews. Our findings show that, within the same listing, having any vicinity safety review is associated with a 1.82% reduction in the listing's monthly occupancy rate and a 1.48% reduction in its average paid price per night. The association with listing safety review is even stronger: having any listing safety review is associated with a 2.58% drop in occupancy and 1.52% in price. These findings, all significant at 99% confidence, suggest that prospective guests are concerned about both listing and vicinity safety, and have different sensitivities to changes in these two types of safety reviews.

Another way to understand consumer sensitivity to vicinity safety is examining whether the guests that wrote a vicinity safety review on Airbnb have changed their subsequent Airbnb activities after having experienced the safety issue mentioned in their review. Arguably, the effect of self-experience is more direct and salient than reading vicinity safety reviews written by anonymous strangers. Indeed, our analysis supports this intuition: guests that wrote a vicinity safety review on Airbnb are less likely to book future stays on Airbnb after posting the safety review, and when they do book future stays on Airbnb, they tend to book in areas with fewer official crimes, fewer overall safety reviews, and a lower percentage of listings with vicinity safety reviews. These findings are compared to guests that have used Airbnb with similar frequencies and booked similar listings in terms of crime and vicinity safety reviews but never write any vicinity safety review within our sample period.

Such guest sensitivity to vicinity safety reviews suggests that omitting them on Airbnb could make future guests worse off, as they may mistakenly book listings in potentially unsafe locales. To gauge the potential loss of guest welfare, we obtain a dataset of competing entire-home VRBO listings and use a discrete choice model to estimate consumer utility from Airbnb entire-home listings, while treating VRBO listings in the same zip code-month as the outside good. We then use the structural estimates to quantify consumer surplus under the status quo of our sample (i.e., vicinity safety reviews are largely permitted) versus three counterfactual scenarios: eliminating all vicinity safety reviews, adjusting the rating of each listing to account for the number of vicinity safety reviews of the listing itself and nearby listings, or alerting all guests to the existing vicinity safety reviews and making them as sensitive as those that have written safety reviews themselves.

Compared to the status quo, we find that the scenario that eliminates all vicinity safety reviews on Airbnb decreases consumer surplus by around 3.12% (without price change), because vicinity safety reviews help guests substitute Airbnb listings with such reviews for Airbnb listings without them or for ones located elsewhere, or for listings on a competing platform (VRBO). In comparison, adjusting a listing's (overall) rating to account for the number of vicinity safety reviews of the listing and nearby listings can increase consumer surplus relative to the status quo, by 1.03% (without price change). Whereas making all guests as alert as those that have written vicinity safety reviews themselves can further increase consumer surplus relative to the status quo by 18.56% (without price change). These effects are slightly reduced if we allow vicinity safety listings to respond to the information regime by adjusting price in 1%. All these estimates are conservative, in part because our conservative definition identifies only 0.25% of all Airbnb reviews as vicinity safety reviews and only 4.43% of listings ever had any vicinity safety reviews in our 2015-2019 sample.

These counterfactual scenarios have different implications for Airbnb hosts and Airbnb as a platform. Because vicinity safety reviews make guests more hesitant to book certain Airbnb listings, posting or emphasizing them have a sorting effect among Airbnb listings, and a demand shrinkage effect on Airbnb listings with vicinity safety reviews. In total, we find that eliminating vicinity safety reviews from Airbnb listings generates 0.52% more gross booking value (GBV), or revenue from reservations, for listings on Airbnb in our sample than the status quo. Adjusting each listing's rating to account for the number of vicinity safety reviews of the listing and nearby listings could decrease Airbnb's GBV by 0.13%. Making all guests as alert as those that have written vicinity safety reviews could decrease Airbnb's GBV by 4.06%. These calculations highlight the diverging interests between Airbnb guests, Airbnb hosts located in areas with different vicinity safety, and the entire Airbnb platform.

As we detail in the next section, our work contributes to the empirical literature of online feedback and seller reputation, and the rising literature of information design in online platforms. As an information intermediary, online platforms have more incentives than a traditional seller to alleviate information asymmetries between buyers and sellers. However, online platforms are still inherently different from a social planner, because they may put more weight on their own business interests than on the welfare of buyers and sellers on the platform, and they may not fully internalize the impact of their policies on competing platforms. Our empirical findings highlight these differences, and quantify the potential impact of a policy that limits the release of information for different economic stakeholders. We also document how the impact of the policy may vary for neighborhoods of different income or different minority populations, as being inclusive could be important for the platform or the social planner. These findings can help facilitate ongoing discussions as to what role and responsibility digital platforms should have as far as collecting and disseminating quality-related information online.

The remainder of the paper is organized as follows. Section 2 reviews the related literature. Section 3 provides some background regarding Airbnb's review system. Section 4 describes the dataset and provides summary statistics. Section 5 reports our empirical findings and implications. Section 6 provides back-of-envelope calculations on how listings' GBV and consumer surplus would change under three counterfactual scenarios. Section 7 offers a discussion, and Section 8 concludes.

2 Related Literature

Safety reviews are a type of buyer-to-seller feedback; thus, our study is directly related to the literature of online feedback and seller reputation.

Arguably, buyer-to-seller feedback is more important for marketplace efficiency than seller-to-buyer feedback, because the key information asymmetry is sellers' private information of their own product or service quality. By the law of large numbers, a sufficient volume of authentic buyer feedback would eventually reveal hidden information regarding seller quality. However, not all buyers are willing to provide feedback, partly because reviewers are not compensated for submitting reviews. For example, 64% of eBay transactions are rated by buyers in the sample studied by (Hui et al., 2021), and 73.5% of New York City UberX trips are rated by passengers (Liu et al., 2021). In comparison, 44.6% of Airbnb trips in our sample have received feedback from guests, which is in line with the guest review rate reported by Fradkin et al. (2021) based on earlier Airbnb data in 2014.

Since accurate feedback is a public good subject to under-provision, many platforms attempt to encourage buyer feedback by offering status, coupons, and merchandise discounts (Li and Xiao, 2014; Cabral and Li, 2015; Li et al., 2020; Fradkin et al., 2015). In contrast, the policy studied in this paper aims to discourage buyers from giving feedback on a particular dimension of quality, which could exacerbate the public good problem in online feedback.

The imperfect review rate is particularly problematic as far as negative feedback is concerned. Studies have shown that buyers tend to under-report bad experiences, with potential explanations that include fear of retaliation, preference to leave the platform after bad experience (Nosko and Tadelis, 2015), pressure to give above-average ratings (Barach et al., 2020), and social connection to the rated seller (Fradkin et al., 2015). For arguably rare, negative events such as safety, the probability of observing pertinent feedback from prior buyers could be further reduced, simply because the chance of experiencing a safety issue is small in absolute terms, even if a neighborhood has safety risks. A policy that discourages vicinity safety reviews could reinforce an existing bias against negative feedback. In fact, perhaps in part due to such a bias against negative feedback, Chakravarty et al. (2010) finds that consumers are more responsive to negative feedback than to positive feedback. This pattern is confirmed in our study: the marginal effect of having any vicinity safety review on a listing's occupancy rate is comparable to that of a 70.18% reduction in the listing's average guest rating.

Another consequence of the bias against negative feedback is that safety reviews on any Airbnb listing can only accumulate slowly over time. This could affect the overall informativeness of safety reviews. Between 2015 and 2019, we observe a growing rank correlation between a zip code's normalized cumulative safety review count and that zip code's normalized official crime statistics in low income and minority areas. This suggests that cumulative safety reviews do contain useful information regarding a zip code's actual safety status, with informativeness that may have increased over time. In comparison, a few studies argue that the online feedback systems may become less informative over time because of the feedback bias reasons cited above (Barach et al., 2020; Klein et al., 2009; Hui et al., 2021). Most of these studies infer feedback informativeness from the content of feedback or policy variations within the feedback system. Our approach is different, as we compare online feedback with a completely independent data source.

More broadly, our study contributes to the growing literature of information design in online platforms. Because feedback is under-provided and there is a selection against negative feedback, researchers have studied the design of feedback systems in terms of who is allowed to provide feedback (Klein et al., 2016; Mayzlin et al., 2014; Zervas et al., 2021), how to improve the authenticity of feedback (Wagman and Conitzer, 2008; Conitzer et al., 2010; Conitzer and Wagman, 2014), when the feedback is revealed to the public (Bolton et al., 2013; Fradkin et al., 2021), what kind of feedback is shown to the public, and how to aggregate historical feedback (Staats et al., 2017; Dai et al., 2018).

Interestingly, some platforms highlight negative consumer feedback, so that future consumers are aware of potential risks associated with the target seller or target product (Pan and Zhang, 2011). An economic reason to do so is that many consumers on online platforms tend to be more responsive to negative feedback than to positive feedback (Chakravarty et al., 2010). Highlighting negative feedback may hurt the sellers with negative feedback but divert buyers towards other sellers on the same platform with zero or not as much negative feedback. If this sorting effect generates more revenue for the platform or reinforces the platform's reputation as far as honesty and transparency, the platform would have an incentive to highlight negative feedback.

In our setting, we observe a counterexample where the platform's policy may discourage buyers from providing a specific type of negative feedback. This is similar to a platform hiding, obfuscating, or deleting negative feedback. To be clear, there are legitimate reasons to do so in some situations: for example, a platform may find certain feedback fake, abusive, or misleading ex post; omitting such feedback could make the information system more authentic and informative for both buyers and sellers (Luca and Zervas, 2016; Chevalier and Mayzlin, 2006).

At the same time, studies have shown that platforms may be strategically motivated to omit certain information, including negative feedback. For instance, Kovbasyuk and Spagnolo (2018) explain why sometimes platforms seek to erase some historical bad records of sellers, in order to increase matching rates. Romanyuk and Smolin (2019) show that platforms such as Uber may seek to hide some buyer information (say, destination) prior to completing a buyer-seller match, because doing so would avoid sellers waiting for a specific type of next buyer which would reduce the overall matching rate on the platform. These two papers differ in the direction of information withholding: the former withholds seller-relevant information from future buyers, while the latter withholds buyer-relevant information from future sellers. Both suggest that the party from whom the information is kept hidden may be worse off and the platform has an incentive to trade off their welfare loss against the welfare gain of the other side of the platform and the platform's overall matching rate. Also, Lewis (2011) shows that online disclosures are important price determinants, and the disclosure costs impact both the level of disclosure and prices.

Airbnb's new policy regarding vicinity safety reviews is an example of withholding or discouraging seller-related information from prospective buyers. As shown in our counterfactual analysis, Airbnb may have economic incentives to downplay vicinity safety reviews, because the more guests are alerted about vicinity safety, the lower the matching rate for the whole platform. In theory, such incentives could be dominated by a sorting effect, if posting or highlighting vicinity safety reviews could direct buyers towards safer listings on the same platform and motivate the safer listings to increase their prices sufficiently high to compensate for the platform's loss from a lower matching rate. Our back-of-envelope calculation suggests that this is not the case.

One welfare aspect that is difficult to quantify but may be relevant for Airbnb is the long-run entry and exit of users. As shown in our back-of-envelope calculations, a policy that encourages and highlights vicinity safety reviews could disproportionately hurt Airbnb hosts in relatively unsafe neighborhoods. In the long run, this could lead to a smaller choice set for guests, drive away some types of hosts and guests, and affect the economic parity across different neighborhoods. How important these long-run considerations are for Airbnb and the social planner depends on how they weigh the welfare of different users. Unfortunately, such a policy was introduced by Airbnb at the end of our data period and just prior to the emergence of the COVID-19 pandemic. It is unclear whether and how Airbnb has enforced this policy. Thus, aside from estimating counterfactuals, we cannot observe the de-facto changes in seller entry and exit because of this policy change.

Another aspect that is worth highlighting in our setting is the potential spillover effect of vicinity safety reviews across listings. As stated before, all buyer-to-seller feedback is a public good that provides little economic return to the reviewer but could benefit many future prospective buyers. In addition to this common feature, vicinity safety reviews could also generate spillovers among listings in nearby geographies, should guests infer the overall safety of the vicinity from multiple nearby listings. We find some suggestive evidence for such negative spillovers: for a focal listing, a higher percentage of other nearby listings with vicinity safety reviews within a 0.3-mile radius is negatively associated with the focal listing's occupancy rate, as well as its price. Among listings within a 0.3-mile radius area, hosts may seek to minimize such a negative externality; but from a prospective guest's perspective, this is a positive information externality that could help guests make more informed choices ex ante. Hence the information design optimal to the platform can be different from the information design optimal to guests.

Our work in part hinges on guests' reactions to safety reviews, and thus relates to the literature on the role of information disclosure in online user behavior. Researchers have shown that product attributes, seller attributes, seller-buyer interactions, and the way in which an online platform aggregates and presents such information (e.g. search ranking, product recommendations, price, consumer ratings, and images of the property) are all important elements in consumers' decisions in e-commerce and the sharing economy (Tadelis, 2016; Ert et al., 2016; Tussyadiah and Park, 2018; Ursu, 2018; Reimers and Waldfogel, 2021; Xu et al., 2021; Zhang et al., 2021c,b). Furthermore, consumer decisions depend on the quality, quantity, resource, and accuracy of the disclosed information (De Pelsmacker and Janssens, 2007; Liu et al., 2017; Munzel, 2016), which in turn depend on the extent to which the users of a platform are willing to disclose such information (Hao and Tan, 2019; Liang et al., 2019; Morosan, 2018; Morosan and DeFranco, 2015; Moon, 2000). In this paper, we focus on guests' reactions to safety reviews, while taking the existence of each historical safety review as given. To ensure the response is specific to safety reviews, we control for listing and host attributes along with listing fixed effects.

We are not the first to study safety issues regarding online short-term rental platforms. Suess et al. (2020) find that non-hosting residents with a higher emotional solidarity with Airbnb visitors are more supportive of Airbnb hosts, and residents hold different views about safety ("stranger danger") and Airbnb depending on whether they have children in the household. Local planners pay attention to the impact of online short-term rentals on neighborhood noise, congestion, safety, and local housing markets (Gurran and Phibbs, 2017; Nieuwland and Van Melik, 2020; Kim et al., 2017). Zhang et al. (2021a) shows that regulations that negatively affect Uber/Lyft services may also negatively affect the demand for Airbnb. Han and Wang (2019) document a positive association between commercial house-sharing and the rise of crime rate in a city, while non-commercial house-sharing does not have this association. A number of studies find that an increase in Airbnb listings — but not reviews — relates to more neighborhood crimes in later years (Xu et al., 2019; Maldonado-Guzmán, 2020; Roth, 2021; Han et al., 2020; Filieri et al., 2021). More specifically, Airbnb clusters are found to correlate positively with property crimes such as robbery and motor vehicle theft, but negatively with violent crimes such as murder and rape. Also, Airbnb listings of the type in which guests may share a room with other unrelated guests are found to be more related to crimes (Xu et al., 2019; Maldonado-Guzmán, 2020) and to skirting local regulations (Jia and Wagman, 2020).

Our study complements this growing literature, by highlighting safety reviews, distinguishing vicinity and listing safety reviews, and documenting consumer responses to safety reviews or experiencing safety issues. Although we cannot identify the effect of Airbnb on local crime rates, our work helps quantify guest preferences regarding safety, and clarify how the interests of guests, different hosts and the platform diverge with respect to the disclosure of safety reviews. As shown in our counterfactuals, disclosing and highlighting vicinity safety reviews could encourage guests to shy away from potentially unsafe listings and disproportionately affect hosts in certain areas.

3 Background of Airbnb's Review System

Over the past decade, short-term vacation rental markets have quickly expanded worldwide. Airbnb, the leading home-sharing marketplace, now offers 6.6 million active listings from over 4 million hosts in more than 220 countries and regions.[2] As with any lodging accommodation, the specific location of a listing can affect the experience of its guests. For instance, if a property is located in a relatively unsafe area, crimes such as carjacking or burglary may be more likely. In Los Angeles, the number of victims to crimes such as theft or burglary at short-term rental lodgings reportedly increased by 555% in 2017-2019.[3] As is common in the lodging industry, guests, who may be traveling outside their home towns and are therefore less familiar with local neighborhoods, are responsible for their own safety in the areas in which they choose to stay. In particular, as with hotels, guests receive little to no protection from rental platforms as far as crimes they may experience in a listing's vicinity.[4]

However, prior to making a reservation, potential guests may refer to a number of sources to gauge the safety of a listing's area — these sources include local news, crime maps, websites that summarize neighborhoods[5], and perhaps most readily linked to each listing, the listing's reviews from prior guests.[6] Airbnb enables guests and hosts to blindly review each other after a guest's stay.[7] In an effort to appease hosts, and perhaps to encourage more listings across a larger number and variety of neighborhoods, a recent Airbnb policy effective December 11 2019 announced that, going forward, guest reviews about a listing that include "content that refers to circumstances entirely outside of another's control" may be irrelevant and subject to removal.[8] This policy change implies that reviews about the safety of a listing's vicinity ("vicinity safety reviews") may be deemed irrelevant and subject to removal, since such a safety aspect is outside the control of the host. Due to data limitations, we do not know how

[2]See Airbnb's official statistics as of December 31, 2022 available at `https://news.airbnb.com/about-us/#:~:text=Airbnb%20was%20born%20in%202007,every%20country%20across%20the%20globe.`

[3]See, e.g., `https://rb.gy/1eohbw` .

[4]See, e.g., `https://rb.gy/nwetrv` and `https://rb.gy/wrqvy4` .

[5]See, e.g., `https://www.neighborhoodscout.com/`.

[6]Reviews have been well established as having a potential effect on buyer decisions and sellers' reputations, particularly in the tourism industry (Schuckert et al., 2015). The literature also suggests that negative information in reviews in particular can have an effect on guest decisions and be useful to platforms in distinguishing seller and product quality (Jia et al., 2021).

[7]If one side does not review the other, the other's review becomes visible after 14 days.

[8]See, for example, `https://rb.gy/0pu5ck` and `https://rb.gy/9y6bum` .

Airbnb enforces this policy as far as vicinity safety reviews, but anecdotes suggest that some reviews that touched on neighborhood safety have been removed.[9] The policy does not apply to "listing safety reviews," because these reviews are about the safety within the listed property, which presumably can be more readily controlled and improved by the listing's host.

It is difficult to pin down exactly why Airbnb adopted this new review policy in December 2019. If Airbnb believes that the main role of online reviews is to motivate hosts to provide high-quality services to guests, review content regarding something outside the host's control may not help in that regard. Anecdotes suggest that hosts have complained about the harm they suffer from "irrelevant" reviews about the vicinity of their listings,[10] and this policy change could be a way to address these complaints. Another reason might be the concern of review accuracy: arguably, vicinity safety is a subjective feeling subject to the reviewer's prior and interpretation, and it is often difficult to prove correct or wrong. However, similar accuracy concerns could apply to other review content, though the degree of objectiveness may vary. A third reason may have something to do with the aspiration of being inclusive. Airbnb has advocated for inclusive design, which is defined as "consciously designing products, services, and environments that don't create barriers to belonging."[11] The same aspiration may have motivated Airbnb to adopt an anti-discrimination policy, establish a permanent anti-discrimination team, and encourage designs and services friendly to users with disabilities. To the extent that vicinity safety reviews are more present in low-income or minority neighborhoods, the new review policy could be another effort to make the platform more friendly to hosts in economically disadvantaged neighborhoods. The key question we address in this paper is how the new policy, if fully implemented as far as vicinity safety reviews, would redistribute the economic benefits and costs among hosts, guests, and the platform.

---

[9] For example, on Jan. 27, 2020, a tweet from "PatrickR0820" wrote "I used @Airbnb when we went to Atlanta for the Panthers game. In my review I left numerous things that could be fixed as well as 'the area that it is located in, is pretty sketchy.' My review and 4 other similar recent reviews were deleted because it wasn't relevant." Another tweet by "AveryBrii" on May 18, 2021 stated: "@Airbnb is such a joke!!! we literally had a car stolen at the place we stayed at, didn't get refunded (which wahtever) & then i try to leave a review to inform others that it clearly was not a safe area (cops told us this & other info that i tried to include) & they didn't post." A journalist also describes his experience on Bloomberg Opinion: "Airbnb Took Down My Negative Review. Why?" (May 26, 2021 by Timothy L. O'Brien), accessed at `https://rb.gy/dxfkxw` , on November 26, 2021.

[10] Nina Medvedeva, "Airbnb's Location Ratings as Anti-Black Spatial Disinvestment in Washington D.C." Platypus: The CASTAC Blog (March 16, 2021) accessed at `https://rb.gy/ottzf9` .

[11] See, e.g., `https://rb.gy/eq7ltv` .

To be clear, Airbnb has adopted other methods to address neighborhood safety directly. For example, Airbnb introduced a neighborhood support hotline in December 2019[12], around the same time as Airbnb adopted the new review policy. This hotline is primarily intended to be a means for neighbors of Airbnb listings to contact the platform in certain situations (e.g., in the event of a party taking place at a listed property). In addition, since our dataset ends in December 2019 and we do not know how many guests that left vicinity safety reviews in our sample would have used the hotline should the hotline exist at the time of the review, we cannot predict how the hotline could counter some of the effects shown in our analysis. That being said, hotline usage is ex post and is not visible to future guests, hence its impact on guests can be fundamentally different from the impact of reviews visible under each listing on Airbnb.

Airbnb's review system also allows guests to leave a 1-5 star rating by specific categories (cleanliness, accuracy, check-in, communication, location, and value), in addition to leaving an overall rating and detailed review. According to Airbnb's response to a host's question, location rating is meant to "help future guests get a sense of the area and tends to reflect proximity to nearby destinations."[13] Hence, location rating could capture many location-specific aspects such as local transit, nearby stores, neighborhood walkability and noise, and may not be directly related to vicinity safety.

4 Data

The first dataset we use has information on the set of short-term rental listings that had been advertised on Airbnb from May 2015 to December 2019, and on VRBO from June 2017 to December 2019, in five US cities (Atlanta, Chicago, Los Angeles, New Orleans, and New York). The data was acquired from AirDNA, a company that specializes in collecting Airbnb and VRBO data. For Airbnb listings, this dataset includes the textual contents of all Airbnb listing reviews in those cities.

Each listing is identified by a unique property ID and comes with time-invariant characteristics such as the listing zip code, listing's property type (entire home, private room, shared

[12]See, e.g., `https://rb.gy/sykoim` .
[13]See, e.g., `https://rb.gy/qs13gh` .

room, or hotel room) as well as the host's unique identifier. Listings also have time-variant characteristics, including average daily rate,[14] the number of reservations, days that are reserved by guests, occupancy rate,[15] number of reviews, overall rating scores,[16] the listing's Superhost status,[17] the listing's guest-facing cancellation policy,[18] the average number of words in the listing's reviews, the number of listings in the same zip code, and whether the listing is cross-listed on VRBO.[19]

Our unit of observation is listing-month. We focus on "active listings" (listings whose calendars are not indicated as 'blocked' in the dataset for an entire month), and exclude observations with ADR over $1000, as some hosts may set their rates prohibitively high in lieu of blocking their calendars. We use regular monthly scrapes between May 2015 and December 2019 on Airbnb (July 2017 to December 2019 for VRBO). In total, the sample comprises 2,866,238 listing-months observations on Airbnb, and 201,718 listing-months observations on VRBO.

We define two different types of safety reviews — listing safety reviews and vicinity safety reviews. Listing safety reviews are those reviews that describe issues pertaining to safety within a listing (e.g., "the listing is unsafe because there are fire hazards", "the listing is unsafe because of the slippery tub", or "we saw mice in the kitchen three times during our stay"). Vicinity safety reviews contain information pertaining to the safety of the nearby vicinity or neighborhood of the listing (e.g., "the neighborhood is not safe", "shady neighborhood", or "unsafe area"). While there is considerable research regarding the use of machine learning for automated content analysis, these methods typically require a large number of hand-labeled examples for training. We instead use a lexicon approach due to its simplicity and transparency. Lexicons are also found to have high levels of precision as compared to machine learning approaches (Zhang et al.,

[14]Average daily rate (ADR) is calculated by dividing the total revenue, including both nightly rates and cleaning fees, earned by the host from reservations over a given month by the total number of nights in that month's reservations.

[15]Occupancy rate is calculated by dividing the number of booked nights by the sum of the available nights and booked nights.

[16]Overall rating scores are normalized to 0-10 range. Our dataset also includes location star ratings. Adding it as an extra control variable does not change our main results, so we do not report it in this paper. Results are available upon request.

[17]Superhost refers to a status badge related to metrics concerning a listing's performance. Hosts who meet the following criteria, evaluated quarterly, receive a Superhost designation: (i) Completed at least 10 reservations in the past 12 months; (ii) maintained a high response rate and low response time; (iii) received primarily 5-star reviews; (iv) did not cancel guest reservations in the past 12 months.

[18]Cancellation policy could be strict, moderate, flexible. For simplicity, we use a dummy variable to indicate whether a listing's cancellation policy is strict or not.

[19]Only listings with entire home that could be both listed on Airbnb and VRBO.

2014; Hutto and Gilbert, 2014), and have been used extensively in the literature (Monroe et al., 2008; Dhaoui et al., 2017).

To identify a suitable set of keywords, we use an iterative approach, starting with terms such as "unsafe," "dangerous," and "scary" and all of their synonyms, to obtain an initial keyword set; next, we manually inspect reviews containing such keywords so as to identify additional keywords. We then select keywords based on the accuracy of safety reviews. More specifically, we conduct two iterations of manual labeling. In the first iteration, three research assistants (comprising both male and female and different races) labeled 1.4K reviews that were generated from the Lexicon approach algorithm with the initial keyword set for both listing safety reviews and vicinity safety reviews. While labeling, for each review the reviewers identified (i) whether the review pertains to neighborhood and/or listing safety, (ii) whether the review has a negative sentiment with respect to neighborhood and/or listing safety, and (iii) the three keywords that supported the reviewer's decision in (i) and (ii). With these human-labeled keywords, we obtain an updated list of vicinity and listing safety keywords such that the percentage of negative vicinity safety (listing safety) reviews in the 1.3K sample with such a human-selected keyword is greater than 0% (10%). In the second iteration of labeling, two research assistants (male and female) of different races labeled 3.1K reviews that were generated from the Lexicon approach algorithm with the updated keyword set for both listing safety reviews and vicinity safety reviews, such that 5 reviews associated with each keyword were randomly selected. In this iteration, the reviewers labeled whether each review pertains to negative sentiment about vicinity safety and/or listing safety. The final set of keywords is the one where each vicinity safety (listing safety) keyword has a percentage of negative-sentiment vicinity safety (listing safety) reviews greater than or equal to 60% from both reviewers' second-iteration labeling results. After two iterations, we expanded the list to 41 vicinity safety keywords and 50 listing safety keywords, as delineated in Table 1.[20]

The keyword lists developed above are not the only inputs we use to define vicinity or listing safety reviews. As far as vicinity safety reviews, to improve precision and to ensure that the text is indeed describing issues pertaining to the safety of a listing's vicinity and not other aspects

[20]Most of the keywords appear relatively infrequently, and removing any one of them alone has little effect on the results. For example, one may argue that "government housing" suggests a low-income area rather than vicinity safety issues. Including it in our vicinity safety keyword list would only identify three more vicinity safety reviews and removing the keyword has no qualitative impact on the results.

of a listing, we identified a list of 24 location keywords that tend to indicate a statement about the surrounding area (e.g., "neighborhood", "area", "outside") in Table 1. We then categorized the matching reviews into those in which the vicinity safety keyword occurred within 20 words of a location keyword as vicinity safety reviews, and those in which the listing safety keyword occurred outside of the 20-word context as listing safety reviews.[21] Next, we selected 13 negative keywords, and filtered out double-negative reviews where the negative keyword occurs within 5 words of a safety keyword.

Overall, our approach resulted in 11.8k matched vicinity safety reviews (VSRs) and 12.8k matched listing safety reviews (LSRs) across the 5 sample cities. In total, they account for 0.25% and 0.27% of all the observed Airbnb reviews respectively. From May 2015 to December 2019, only 4.43% of listings ever had any VSRs, and only 8.49% of listing ever had any safety reviews (VSRs or LSRs).

As shown in Figure 1 and Figure 2, the top matching vicinity safety keywords are "unsafe" (4,519), "homeless" (3,398), "yelling" (854), and "uneasy" (733), and the top matching listing safety keywords are "worst" (1,803), "mold" (1,350), "stained" (1,172), and "filthy" (1,135). As an additional validation check, we sampled several thousand matches at random, and manually labeled them as relevant or not, finding 78.21% and 75.64% accuracy for vicinity safety keywords and listing safety keywords, respectively.[22] The mislabeled data often used figurative language ("scary how perfect this neighborhood is") or used safety words in other contexts (e.g., "watched a scary movie on Netflix"). While any such method will be imperfect, we did not find any evidence suggesting that the error rates were systematically biased for some neighborhoods over others. However, we did restrict our keywords to English, so the method will be less effective in areas with many non-English reviews.

A second dataset we collect covers official crime records from databases tracking crimes in

[21] While the 20-word window is arbitrary, a sensitivity analysis suggests no qualitative difference when using a slightly longer or shorter window. Moreover, the average review had roughly 50 words, so this seemed to restrict to the 1-2 sentences around the keyword match.

[22] This indicates a 21.79% false-positive error rate for vicinity safety reviews (24.36% for listing safety reviews). Since our lexicon approach aims to minimize the false-positive rate while allowing false negatives, the safety reviews identified by this approach tends to make the estimated impact of safety reviews more conservative than the true effect.

Chicago[23], New Orleans[24], New York City[25], Atlanta[26], and Los Angeles.[27] These databases cover different types of crimes, including property-related crimes and violent crimes. In terms of the geographical granularity of crimes, we consider crime events at the zip code level. We also obtain median income and other demographic information at the zip code level from 2014, one year before our Airbnb sample period begins, from the United States Census Bureau[28]. We make the assumption that the income and demographic information did not change significantly over our sample period. Throughout the paper, we refer to a zip code as high-income (H) or low-income (L) according to whether its average income is above or below the median of the city it locates in. Similarly, we refer to a zip code as minority (M) or white (W) according to whether its percentage of minorities in population is below or above the city median.

Table 2 summarizes the data at the listing-month level, where vicinity safety (VS) Airbnb listings are defined as observations that have at least one vicinity safety review (VSR) before the reporting month, while "normal" Airbnb listings do not have any VSR before the reporting month. As the table indicates, about 4% of the total observations are VS listings. On average, VS listings have higher occupancy rates, a higher number of reservations, a higher fraction of Superhosts, and a higher number of reviews than normal listings. In contrast, the nightly rates and overall rating of VS listings are lower on average than normal listings.

The mean number of cumulative VSRs (aggregated up to the reporting month) is 0.06 across all Airbnb listings, and the mean number of cumulative listing safety reviews (LSR) is 0.06. The monthly trends for the percentages of VSRs and LSRs are depicted in Figure 3, where the percentage of VSRs increases at a faster rate relative to the percentage of LSRs over time. Figure 4 and 5 demonstrate the distribution of VS keywords for four groups of zip codes (high-income, low-income, white, and minority). Comparing high-income with low-income (and white with minority) groups, it appears that the low-income (minority) group dominates the volume of VSRs.

We also test the rank correlation between the official crime records and VSRs. Specifically, we use the percentile rank of normalized crime records in each zip code-month within each city

[23] Official crime data in Chicago: `https://rb.gy/atjsss` .
[24] Official crime data in New Orleans: `https://rb.gy/4vue82` .
[25] Official crime data in New York City: `https://rb.gy/iwrwp2` .
[26] Official crime data in Atlanta: `https://rb.gy/96txbl` .
[27] Official crime data in Los Angeles: `https://rb.gy/tebnla` .
[28] See, e.g., `https://www.census.gov/data.html`.

— calculated as the number of reported crime cases in a month, divided by the size of the population in that zip code. For each month, we rank the normalized crime data within each city, and determine the percentile crime rank of the zip code for that month. For VSRs, we use the percentile rank of the number of cumulative VSRs in the zip code up to the reporting month.[29] We then test the percentile rank correlation index between the crime records and VSRs in each month, resulting in the time-series correlation trends depicted in Figure 6, which illustrates the correlation trends for the four different groups of zip codes (high-income, low-income, white, and minority). Figure 6 indicates that the correlation in low-income and minority groups exhibits an increasing trend, suggesting that the percentile rank of VSRs in a zip code is more likely to reflect the actual crime reports in the zip code over time in these areas.

5 Empirical Analysis

5.1 Effects of Safety Reviews

We begin by assessing the effects of VSRs and LSRs. Our hypothesis is that if potential guests view VSRs and LSRs as a proxy for safety around or within a listing, such reviews would reduce the guests' willingness to book the listing. Our base specification is given by:

$$y_{i,t} = \alpha_i + \alpha_{k,t} + \delta X_{i,t} + \beta_1 Crime_{i,t-1} + \beta_2 LSR_{i,t-1} + \beta_3 VSR_{i,t-1} + \beta_4 VSRADIUS_{i,t-1} + \epsilon_{i,t}, \quad (1)$$

where i denotes a listing i-month t observation, $Crime_{i,t-1}$ is a log transformed variable that indicates the normalized number of cumulative official crime reports since the start of the sample period for the zip code where listing i is located, $LSR_{i,t-1}$ and $VSR_{i,t-1}$ are two dummy variables that equal 1 if the listing has at least one LSR and VSR, respectively, before month t, $VSRADIUS_{i,t-1}$ is the percentage of listings that have at least one VSR within a 0.3-mile radius of listing i prior to month t, $X_{i,t}$ are listing-level controls (logged except for dummy variables), including the number of reviews, overall ratings, cancellation policy, number of listing in the same zip code, cross-listing status (i.e., whether the listing is also listed on VRBO), and whether

[29] Due to data limitations, we assume that both records begin with clean slate (0 records) as of the beginning of our dataset.

the listing is hosted by a Superhost. The dependent variable $y_{i,t}$ is either the log of listing i's average daily rate (ADR) in month t, or the log of listing i's monthly occupancy rate (calculated as log of 1 plus the occupancy rate).[30] Listing and City–year-month fixed effects are denoted by α_i and $\alpha_{k,t}$, respectively, where the city of listing i is denoted by k. Standard errors are clustered by Airbnb property ID. The primary assumption is that, within a listing, the presence and timing of safety reviews are correlated with the true safety condition around or inside the listing and do not reflect selective reporting, fake reviews, or other strategic reasons once we control for other time-varying listing attributes.

Our main specifications in Table 3 indicate that both VSRs and LSRs significantly decrease a listing's price (ADR) and occupancy. Specifically, for an average Airbnb listing in our sample, having any VSR is associated with a 1.82% reduction in the listing's monthly occupancy rate and a 1.48% reduction in its average price per reserved night; having an LSR is associated with a 2.58% drop in occupancy and 1.52% in price. LSRs thus have a larger effect on price and occupancy than VSRs, possibly because some prospective guests have a specific geographic area (e.g., neighborhood) in mind, regardless of safety issues concerning that area, whereas LSRs describe safety issues that pertain to the listing itself. The percentage of listings with VSRs within a 0.3-mile radius is associated with lower prices and lower occupancy, suggesting that guests may also infer vicinity safety from the VSRs of nearby listings.

In contrast, normalized official crime records is associated with lower prices but higher occupancy. A potential explanation is that hosts are aware of safety issues in the areas of their listings, and proactively lower their rates when their listings are located in relatively unsafe areas. These lower prices attract more guest bookings, perhaps either because guests tend not to seek information about crimes in the neighborhood or because they prioritize price. In particular, for the average Airbnb listing in our sample, given a 1% increase in the normalized official crime records, the daily rate is 0.05% lower whereas the occupancy rate is 0.07% higher.

[30]Some listing-month observations have an occupancy rate of 0 and consequently are missing an average reserved daily rate in the dataset for those months, though the dataset does offer a separate "listing price" (i.e., a base rate) for those listings. To extrapolate the ADR of these listings in the months in which they are missing, we calculate the mean ratio of their ADR to their listing price in the months in which they are available, and multiply this average by the listing price in the missing months (if available, or by using the listing price from the nearest month in which it is available).

5.2 Robustness

Column 1 of Table 4 considers as the dependent variable a dummy that equals 1 when a listing's occupancy rate is positive and 0 otherwise. It reports a positive coefficient on $Crime_{i,t-1}$, suggesting that the variable $Crime_{i,t-1}$ not only describes the relative crime status of a zip code, but may also capture the relative guest traffic to the area, where areas with relatively high guest traffic (e.g., downtown areas) tend to have a higher number of reported (normalized) crimes. Comparing the coefficients on VSRs and LSRs for the whole-sample specifications (Table 3) to the conditional sample with positive occupancy rates (Columns 2 and 6 of Table 4), we find that the coefficients are similar but have somewhat higher magnitudes for the whole sample.

We conduct a number of additional checks. First, we split the sample by whether a listing has an above- or below-median number of reviews in a given month (median is 12), as a proxy for whether the listing is in its early or later "stage" of taking guest reservations, since only staying guests can post a review.[31] Another motivation for this partition is that prospective guests are more likely to notice safety reviews (both VSRs and LSRs) when listings have a lower number of reviews. Indeed, Columns 3 and 4 of Table 4 report that in the subsample of listings with 13 or fewer reviews, the negative effects of having any VSR and LSR on occupancy rate (2.12% for VSR and 3.62% for LSR) are higher than the corresponding negative effects for listings with more than 13 reviews (1.00% for VSR and 1.73% for LSR). However, Columns 7 and 8 indicate that as far as listings' daily rates are concerned, this comparison is reversed, possibly because hosts of newer listings may still be in the process of identifying their pricing for those listings.

Second, we add additional controls for the average word count of a listing's reviews. [32] As Columns 5 and 9 of Table 4 indicate, the results do not qualitatively change from our main specifications when incorporating the additional control.

5.3 Heterogeneous Effects

We next explore a number of heterogeneous effects. Table 5 provides summary statistics based on the type or area of a listing. In particular, the table reports different normalized zip code

[31] To be clear, the same listing may be in both subsamples over time, but belong to only one of the subsamples in any given month.

[32] Host responses to safety reviews are not observed in our data

crime levels for listings in these categories. We proceed with a similar empirical methodology as in (Equation 1), but with different subsamples.

We begin by analyzing four groups of zip codes separately (high-income, low-income, white, and minority). Table 6 shows that VSRs have negative effects on occupancy rates across all four subsamples. The negative effects of having any VSR on occupancy rates have higher magnitudes in high-income and white zip codes (1.76% and 1.89%) than in low-income and minority zip codes (1.72% and 1.75%). A similar comparison holds for LSRs. One potential explanation is that guests may have different prior beliefs and different sensitivities to safety issues, and perhaps more so if their search targets a specific area that they believe is relatively safe. Hosts in different areas may also react differently to VSRs and LSRs, based on how they gauge guest perception and guest preferences.

We next consider subsamples comprising different listing types (entire home, private room, shared room, and hotel room). Additional heterogeneous effects may arise here because, for instance, for guests who seek partial spaces (private room, shared space) within a dwelling, safety issues may be more salient. The results in Table 7 indeed show that the magnitude of the negative effects from having any VSRs and LSRs on occupancy are larger for private rooms and shared spaces (2.10% and 3.01% for VSR and 3.08% and 2.89% for LSR, respectively) in comparison with entire-home listings (1.61% for VSR and 2.36% for LSR).

5.4 Safety Experience and Future Guest Activity on Airbnb

We conduct user-level analyses to test whether guests who leave any vicinity safety reviews (henceforth, VS guests) act differently before and after they post their first VSR in comparison to otherwise similar guests who did not leave any VSRs. To that end, we assume that the first VSR that a VS guest posts for one of the listings in our sample (i.e., covering Airbnb listings in the five cities we consider, with reviews beginning in May 2015) is the first VSR that this guest posted. To reiterate, any such guests who have ever posted VSRs in our sample are considered VS guests; otherwise, they are treated as 'normal' users. To ensure that the VS users have had some experience on Airbnb prior to leaving their first VSR, we focus on the subset of VS users that left at least two reviews in the five sample cities before leaving their first VSR.

In order to match VS users with normal users, we use a K-nearest neighbor (KNN) method to select the two most similar control (normal) users for each treatment (VS) user. The user characteristics used in the KNN method (as of the time of the treatment user's first VSR) are the user's number of prior reviews, the average normalized crime reports in the cities in which the user stayed (based on their prior reviews), the average number of VSRs for listings for which the user left reviews, the average percentage of overall VS listings in the same zip codes as well as in the 0.3-mile radius area as listings for which the user had previously left reviews, and the average number of words for the reviews that the user posted before. The matching is done for each month (i.e., based on new treatment users in each month). The same "treatment month" is applied (hypothetically) to control users that are matched with a treatment (VS) user, based on the latter's timing of their first VSR.

In order to assess if the treatment and control users have the same tendency to post VSRs, we also calculate the propensity score for each user in our matched sample. In particular, we regress the dummy of a user being treated (i.e. being a VS user) on the number of reservations she had made on Airbnb before the treatment time, the average zip code-wide crime rate of these reservations at the time of reservation, the average number of VSRs in these reservations, the percent of listings with any VSR in the zip code as well as in the 0.3-mile radius area of these reservations, and the average number of words for the reviews that the user posted before. For a treated user, the treatment time is when she wrote her first VSR in our sample. For a control user, the treatment time is when the treatment user she is paired with wrote her first VSR in our sample. Table 8 reports that the treatment and control users are similar as far as the characteristics considered in the KNN method; the two user groups also have similar propensity scores, as shown in Figure 7 and 8.

We first test whether VS users behave differently in terms of subsequent reservations on Airbnb after their first VSR (as exhibited by their subsequent listing reviews). We use a difference-in-differences methodology (DID) as follows:

$$y_{i,t} = \alpha_t + \alpha_p + \beta \cdot treat_i + \gamma \cdot treat_i \times post_t + \epsilon_{i,t}, \tag{2}$$

where the subscript p denotes the treatment-control pair we have identified in the sample construction and α_p is pair fixed effects. We have constructed several measures for the dependent variable $y_{i,t}$: the first is the number of reviews that user i wrote in month t. We use it as a proxy of user i's Airbnb reservations in t, which can be zero. Because it is a count variable, we use a Poisson regression instead of ordinary least squares. The second measure is the normalized cumulative count of officially reported crimes in the zip codes of user i-reviewed listings in month t. The other measures are the number of VSRs in the reserved listings, the percentage of VS listings in the zip codes as well as in the 0.3-radius area of the reserved listings, and whether the reserved listings have any VSR. The dummy variable $treat_i$ equals 1 for VS users and 0 otherwise, and the dummy variable $post_t$ equals 1 if t is after the time of the first VSR of user i. Treatment-control pairs fixed effects are denoted by α_t, standard errors are robust and clustered by treatment-control pairs.

Column 1 of Table 9 reports results from a Poisson model based on an unbalanced monthly panel data, indicating that VS users tend to book fewer reservations (as evidenced by subsequent reviews) after posting their first VSR. In particular, the average monthly number of subsequent reviews is expected to be 60.07% lower for VS users in comparison with normal users.[33]

We also assess whether VS users are more sensitive to safety information when booking subsequent Airbnb listings after posting their first VSR. In order to test this hypothesis, we use the booked listings' characteristics as the dependent variables. Columns 2-6 of Table 9 suggest that the subsequent listings chosen by VS users exhibit the following characteristics: They tend to be located in zip codes that have fewer normalized crime reports, they are less likely to have VSRs, and they are less likely to be located in zip codes that have a higher overall percentage of VSRs or a higher percentage of other listings with VSRs. This suggests that VS users, relative to normal users, are more sensitive to safety information after posting their first VSR.

We further examine whether VS users subsequently act differently as a function of the area (high-income, low-income, minority or white) in which they posted their first VSR. To do so, we group VS users according to the zip code of the listing for which they posted their first VSR, and proceed to conduct the DID analysis separately for each of the four subsamples.

Table 10 reports that VS users tend to book subsequent stays (as proxied by their subsequent listing reviews) in areas that are the opposite of where they posted their first VS review. That is, VS users whose first VSRs are posted in high-income or white areas tend to book fewer subsequent stays in those areas but more in low-income or minority areas, and vice versa. One

[33] This is not the coefficient of the treatment dummy (-0.918) because we use a Poisson model for this regression, i.e., the applicable percentage is $1 - e^{-.918}$.

possible explanation for the former direction is that VS users expected a higher level of safety in high-income or white areas, and when they encountered the opposite, they preferred to pay the average lower rates for listings in low-income and minority areas in subsequent stays. A potential explanation for the latter direction is that VS users associate safety issues with low-income or minority areas, and tend to avoid such areas in future bookings.

From the interaction term in Table 10, it is apparent that VS users exhibit a positive effect on subsequent reservations in opposite zip codes (Columns 2 and 4) and a negative effect in the same zip codes (Columns 1 and 3). The number of VS users who post their first VSR in high-income or white areas is fewer than those in the low-income or minority areas, hence those that shy away from high-income or white areas are fewer than those that switch to such areas because of their safety experience in low-income or minority areas. As a result, the overall effect on subsequent bookings (as proxied by the total number of subsequent reviews) is positive in high-income or white areas but negative in low-income or minority areas.

6 Back-of-the-Envelope Calculations

6.1 Airbnb's Gain and Loss

So far, our analysis has shown that (1) VSRs have negative effects on Airbnb listings' prices and occupancy, and (2) guests who have posted any VSRs appear to book fewer subsequent reservations and are more sensitive to safety information than other guests that never posted any VSRs.

These findings suggest that the status quo (in our data period, 2015-2019) has disclosed vicinity safety reviews but they are not as salient as they could be. For some listings, prospective guests can find vicinity safety information in prior consumer reviews, but the guests that incurred a vicinity safety issue during their own stay at Airbnb listings become more alert to safety information than other users, likely because self experience is more salient than safety reviews from other anonymous users. The status quo also implies diverging interests in the information value of vicinity safety reviews: while guests view VSRs as a negative but informative attribute of a listing, the host of VS listings may perceive VSRs as a harm to future business. In contrast, the hosts of normal listings may consider their lack of VSRs as a competitive advantage over VS listings.

To highlight these diverging interests, we run back-of-envelope comparisons with respect to the revenues of VS and normal listings, under four information regimes: (i) the status quo; (ii) no disclosure (where all VSRs are removed); (iii) adjusted-rating (where each listing's overall rating is adjusted to account for the number of VSRs of the listing itself as well as listings in a 0.3-mile radius area); in particular, we compute a safety score for each observation by using the reversed percentile of the number of VSRs of the listing itself and listings in a 0.3-mile radius area for each city-month, normalizing it on a range from 0 to 10 with a uniform distribution, and then adjusting the new overall rating as a weighted average of the overall rating and the safety score, where the overall rating has a weight of 6/7, to account for the 6 ratings originally included by Airbnb of communication, accuracy, cleanliness, check-in, location, and value; and (iii) high alert (where all users react to VSRs as much as VS users react to their own reported VSR). To understand heterogeneity across areas, we run these back-of-envelope calculations for high-income, low-income, minority, and white areas separately.

For the no-disclosure counterfactual, we set vicinity safety reviews to zero while holding everything else equal. This implies that guests take zero reviews as literally zero and would not reinterpret the lack of vicinity safety reviews in the counterfactual. This assumption is reasonable because vicinity safety reviews are rare (only 0.25% of reviews are identified as vicinity safety reviews and only 8.49% of listings have ever had any safety reviews), many guests are casual users that are likely inattentive to Airbnb policy changes, and most guests may not know a listing could have no vicinity safety reviews because they are removed by the platform.[34] To run the counterfactual, we use the results in Table 6, which capture the effects of VSRs and the percentage of VS listings within a 0.3-mile radius on price and occupancy, in high-income, low-income, white, and minority areas. We next collect the number of Airbnb observations, average occupancy in days, average price (ADR), and average percentage of VS listings within a 0.3-mile radius area for both VS and normal listings in these four areas. The gain/loss of VS listings from the no-disclosure regime (from May 2015 to December 2019) is calculated using the change in occupancy rate and price, had their VSRs and the VSRs of other

[34]In theory, if most guests are fully aware of the no-disclosure policy and correctly predict the extent of vicinity safety reviews suppressed by the policy, they could readjust their safety perceptions of all listings. It is difficult to derive the new equilibrium because the adjustments in safety perceptions depend on whether and how guests may search for vicinity safety information elsewhere.

VS listings within a 0.3-mile radius area been removed. The gain/loss of normal listings from the no-disclosure regime is calculated as the change in occupancy and price had there been no VS listings in the 0.3-mile radius area.

For the high-alert counterfactual, where all prospective guests behave as VS users, we need to extend the DID results from Table 10. to all users. To do so, our raw data contains the number of reservations (as proxied by the number of reviews) and average days per reservation of all VS users for both VS and normal listings in each of the four demographic areas. These are the observed bookings in the status quo. The DID estimates tell us how VS users' bookings react to their own experience of VSR. Appendix A.1 provides formulas to compute changes in the number of bookings in the four demographic areas, assuming the same DID estimates apply to all users. Based on the computed booking changes (and assuming prices do not change), we compute the gain/loss in revenue.

The calculations are summarized in Table 11. The results suggest that the switch from the status quo to no disclosure leads to gains for VS listings in all four demographic areas. The revenues of VS listings in low-income and minority zip codes increase by 6.20% and 6.29%, respectively, which is lower than the corresponding increases of VS listings in high-income and white zip codes (6.32% and 6.70%, respectively). One possibility is that users are more sensitive to VSRs in high-income and white zip codes, hence they stand to benefit more when such reviews are eliminated. Normal listings in low-income and minority zip codes benefit more when no VSRs are available, with revenues increasing 0.21% and 0.35%, compared to 0.17% and 0.13% for normal listings in high-income and white zip codes, respectively. A potential explanation is that normal listings in low-income and minority zip codes have a higher percentage of VS listings within 0.3-mile radius area, and thus suffer from a higher negative spillover effect on daily rates under the status quo. When VSRs are unavailable, they stand to benefit more from higher prices and occupancy rates.

Under the scenario in which we incorporate each listing's own and nearby VSRs in its overall rating, Table 11 shows that both VS listings and normal listings are harmed from having the revised overall rating, where the revenue decrease for listings in the low-income (0.24% and 0.17% for VS listings and normal listings, respectively) and minority areas (0.32% and 0.24% for VS listings and normal listings, respectively) are higher than for listings in high-income (0.14% and 0.09% for VS listings and normal listings, respectively) and white (0.11% and

0.06% for VS listings and normal listings, respectively) areas. This implies that listings in low-income and minority areas have more VSRs, which lowers the adjusted overall ratings of listings in those areas more so than listings in high-income and white areas.

Table 11 also reports results for the high-alert counterfactual, where prospective guests behave as VS guests. Both VS and normal listings in low-income and minority zip codes stand to lose revenues in the high-alert regime as compared to the status quo, with 4.49% and 4.84% declines for VS listings and 4.34% and 4.87% declines for normal listings. And, their counterparts' VSL in high-income and white zip codes also lose revenues if the regime is changed from the status quo to high alert. This is because, after guests become more alert to VSRs, the guest switches from low-income and minority areas into high-income and white areas dominate those that switch away in the other direction, given the fact that VSRs are less likely to occur in high-income and white zip codes.

From the platform's perspective, the overall revenue or GBV sums up the revenue gains and losses across all areas. As shown in Panel C of Table 11, shifting from the status quo to the no-disclosure regime will increase Airbnb's GBV by 0.52%. In comparison, a shift from the status quo to the adjusted-rating and high-alert regimes will reduce Airbnb's GBV by 0.13% and 4.06%, respectively.

6.2 Consumer Surplus

So far, the back-of-the-envelope calculations have focused on listing revenues under the three disclosure regimes. We now aim to do the same on the guest side.

To do so, we define the market as online short-term entire-home rentals in each zip code-month, where Airbnb and VRBO are assumed to be the only two platforms that supply this market. Each guest chooses among all Airbnb entire-home listings available in the target zip code-month, with the pool of VRBO-only listings in the same zip code-month as the outside good.[35] We focus on entire-home listings because only entire-home listings are available on VRBO. Since our VRBO data period is from June 2017 to December 2019, our analysis in this subsection considers Airbnb entire-home listings from June 2017 to December 2019 only.

Following Berry (1994), we assume that each prospective guest chooses an Airbnb entire-

[35]Listings that co-list on Airbnb and VRBO are treated as Airbnb listings, thus inside goods.

home listing or the outside good (VRBO) so as to maximize her utility from the listing, where the utility associated with an Airbnb listing i in zip code z of city k and month t can be written as:

$$\begin{aligned} U_{i,t} &= EU_{i,t} + \epsilon_{i,t} \\ &= \alpha_i + \alpha_{k,t} + \delta \cdot X_{i,t} + \beta_0 \cdot log(ADR_{i,t}) + \beta_1 \cdot Crime_{z,t-1} \\ &+ \beta_2 \cdot LSR_{i,t-1} + \beta_3 \cdot VSR_{i,t-1} + \beta_4 \cdot VSRADIUS_{i,t-1} + \epsilon_{i,t}. \end{aligned}$$

If $\epsilon_{i,t}$ conforms to the logistic distribution, we can express the market share of listing i at time t as $s_{i,t} = \frac{exp(EU_{i,t})}{1+\sum_j exp(EU_{j,t})}$. Thus:

$$ln(s_{i,t}) - ln(s_{0,t}) = EU_{i,t} \tag{3}$$

This is equivalent to regressing the difference of log market share between listing i and the outside good ($ln(s_{i,t}) - ln(s_{0,t})$) on the attributes of listing i in month t. The right-hand side of Equation 3 is similar to Equation 1 except for two changes: first, we exclude the number of Airbnb listings in the zip code-month because the discrete choice model already accounts for the size of the choice set; second, we include the log of the listing's ADR (i.e. price). To the extent that $log(ADR)$ might be endogenous, we instrument it by using the average attributes of entire-home listings within a 0.3-mile radius of the focal listing in the same zip code-month, following Berry et al. (1995) . The underlying assumption is that these so-called "BLP" instruments are correlated with price because of horizontal competition (whereby competitors' attributes affect margins) but are excluded because they do not affect the focal listing's utility directly. As shown in the first column of Table 12, the instrument is strongly correlated with $log(ADR)$, and the first stage F-statistics is high (288.5). The OLS and IV estimation results of the utility function are reported in the last two columns of Table 12. The results suggest that guest reservations are sensitive to price, and guests dislike listings with any VSRs or LSRs, everything else being equal. Based on the IV estimates, the guest's dis-utility from a listing with any VSR (as compared to no VSR) is equivalent to 2.2% of average daily rate ($164.7).

Using the IV results in Column 4 of Table 12, we then calculate $EU_{i,t}$ for each Airbnb listing-month under the status quo, and normalize it into US dollars.[36] The sum of guest utility weighted by the simulated market shares give us the total consumer surplus under the status quo.

For the counterfactual of no-disclosure, we set all VSRs as zero in the utility function, recompute $EU_{i,t}$ for each Airbnb entire-home listing, and simulate its market share. This calculation assumes everything else remains the same when the platform removes all VSRs. It could be violated if listings adjust prices after the regime shift. Unfortunately, the vast majority of our data precede Airbnb's new review policy, so we cannot observe such price adjustments directly. The reduced-form regressions in Table 6 describe the relative price difference between VS and normal listings in the four demographic areas (under the status quo). In an alternative calculation, we assume the no-disclosure regime would erase the price discounts of VS listings while the pricing of normal listings remains unchanged. This gives us a comparison between no disclosure with price changes versus no disclosure without price changes.

Under the adjusted-rating counterfactual, we change the overall rating in the utility function, where each listing's overall rating is adjusted to account for the number of VSRs of the listing itself as well as listings in a 0.3-mile radius area. This calculation assumes the platform has one additional safety rating dimension in addition to the existing 6 rating dimensions (cleanliness, accuracy, check-in, communication, location, and value). Since we do not know how much prices would adjust with such a rating change, we assume an ad-hoc price change (-1% for VS listings) and simulate market shares with and without price changes under the adjusted-rating counterfactual.

To consider the high-alert counterfactual, we use the treatment effect estimated in Column 1 of Table 9 to adjust the coefficient on VSR in the utility equation. In particular, Column 1 of Table 9 estimates the coefficient of the interaction between treated and post as -0.918, suggesting that VS users would reduce their average monthly Airbnb reservations by 60.07% after they posted their first VSR on airbnb. Assuming this effect is completely driven by the coefficient on the VSR dummy in the utility equation, we calibrate how much this coefficient has to decline (i.e. become more negative) in order to generate the same decline as estimated

[36]Normalized $EU_{i,t} = EU_{i,t} \cdot ADR_{i,t}/|\beta_0| + constant$ where β_0 is the estimated coefficient of $log(ADR)$ in Equation 3 and the constant is chosen such that the normalized EU is always positive. Since we use the same constant when we compute utility in different scenarios, the value of the constant does not affect any comparison between scenarios.

in Table 9. Following the detailed calibration procedure in Appendix A.2, we find that the coefficient on the VSR dummy in the utility must decline -2.17 to fit the observed switching behavior of VS users, suggesting that VS users are nearly 15 times more sensitive than a typical user in the status quo. We then use this calibrated coefficient (and all the other coefficients estimated in Table 12) to simulate market shares under the high-alert regime without price changes. Since we do not know how much prices would adjust as a result of the shift to the high-alert regime, in an alternative scenario we assume an ad-hoc price change (-1% for VS listings), to illustrate how price changes may alleviate the impact of making all users highly alert regarding VSRs. The resulting simulation is presented as high-alert with price changes.

Table 13 reports the consumer surplus results under the above six counterfactual scenarios (no disclosure with and without price changes, adjusted-rating with and without price changes, high alert with and without price changes), separating VS listings (on Airbnb), normal listings (on Airbnb), and VRBO-only listings.[37] A particular element in the consumer surplus calculations is worth mentioning: because guest perception of safety can be different from guests' real experience of safety, the realized consumer surplus should use the utilities that represent guests' *realized* utilities and the simulated choice of market shares based on their *perceived* utilities as shown in Table 9. More specifically, we assume the utility function described above for each scenario represents the perceived utility, and a guest's realized utility is represented by her utility when the coefficient on VSR in her utility function is the same as the coefficient we have calibrated for VS users.

Table 13 indicates that, under the regime of high alert without price changes, consumer surplus from VS listings would decrease by 84.22% in comparison to the status quo, mostly because highly-alert guests would switch away from VS listings towards normal and VRBO listings. A hypothetical 1% price drop for VS listings (in the regime of high alert with price changes) may partially compensate the loss, leading to a smaller decline of consumer surplus from VS listings (83.22%) in comparison to the status quo. At the same time, consumer surplus from normal and VRBO listings (under high alert without price changes) increases by 5.42% and 5.30%, respectively, than the status quo, and by 5.36% and 5.22% if we incorporate the hypothetical 1% price drop of VS listings. Overall, consumer surplus under the high-alert

[37]The consumer surplus reported in Table 13 is for an average user in an average reservation day across all 9,940 zipcode-months.

counterfactual increases relative to the status quo 18.56% without price changes and 18.48% with price changes, because the high-alert regime helps guests to reduce stays in relatively unsafe listings.

For the same reason, consumer surplus under the no-disclosure counterfactual declines as compared to the status quo (by 3.12% without price change and 2.59% with price changes) because consumers cannot use VSRs as an information source to sort between VS, normal and VRBO listings. Consumer surplus under the adjusted-rating counterfactual increases slightly as compared to the status quo (by 1.03% without price change and 0.56% with price changes), because the adjusted rating has incorporated vicinity safety reviews at the listing level, though this change is much milder than the highlight in the high-alert counterfactual.

These estimated changes in consumer surplus are conservative, in part because our definition identifies only 0.25% of all Airbnb reviews as vicinity safety reviews, and only 4.43% of listings ever had any vicinity safety reviews in our 2015-2019 sample. Because of this, the no disclosure counterfactual only moves 0.74% of market share from VRBO and normal Airbnb listings to VS listings (before we take into account any price change), and the adjusted ratings incorporating VSRs only moves 0.32% of market shares away from VS listings. In comparison, the dramatic "high-alert" counterfactual would move 5.05% of market share away from VS listings, leaving less than 1% of users choosing VS listings (with or without price change).

7 Conclusion

Examining the effects of vicinity safety reviews and listing safety reviews on listing performance, we find that they both negatively affect occupancy and price, and the effect from listing safety reviews is stronger. We also demonstrate that for guests that post about vicinity safety issues, concerns about vicinity safety appear to be more salient, such that they are less likely to book further stays on Airbnb, and when they do book, they tend to book in areas with fewer official crime reports and fewer vicinity safety reviews. Using back-of-the-envelope calculations, we show that expanding the disclosure of vicinity safety issues may disproportionately affect hosts in low income and minority areas, and that a GBV-centric platform may prefer to limit the disclosure of safety information about the vicinity of listings altogether, even though the aggregate surplus of guests appears to increase when the related safety reviews are instead emphasized to alert prospective guests.

Combined, our findings suggest that the platform faces a tradeoff. On the one hand, the results suggest that all Airbnb listings, and thus the platform, stand to lose revenue under the high-alert regime. Moreover, listings in low-income and minority zip codes stand to lose a disproportionate share of their revenues than their counterparts in high-income and white zip codes. On the other hand, consumer surplus under the high-alert regime is higher than under the status quo and the no-disclosure regimes. The platform thus faces a tradeoff of generating higher revenues and attracting hosts in low-income and minority areas on the one hand, and providing additional value to its buyers on the other.

To the extent that being inclusive is one motivation behind Airbnb's new review policy, our findings suggest that the policy, if fully implemented, may have some unintended consequences on consumers and listings without safety reviews. How to balance the economic interests of all users is a challenge to platforms and policy makers that strive to maximize social welfare. One potential solution is that the platform may import external information about vicinity safety and present it as an alternative to vicinity safety reviews for each listing. Unfortunately, not all cities publish official crime statistics as the five cities in our sample do, and crime statistics may not fully capture all of the safety concerns a guest may have in mind at the time of booking. How to overcome these data difficulties and how to design an objective, universal, and user-friendly metric of vicinity safety certainly merits future research.

There are a number of limitations to our analyses due to the limitations of the data. First, the listing reviews in our data do not include potential responses from hosts. On Airbnb, hosts can reply to guests' reviews, which may also play a role in prospective guests' decisions. Second, in the user-level analysis, we only observe a user's reservation provided that they post a review. It is unclear whether guests are more, less or equally likely to post subsequent reviews after posting their first VS review. More specifically, if VS users are more vocal and thus more likely to post subsequent reviews after their first VS review, then our findings underestimate the magnitude of the effects on their subsequent booking activity; if, however, VS users are less likely to post subsequent reviews, then our findings overestimate the effects. Third, the users in our user-level analysis are limited to those users who have ever made reservations in the five major US cities we consider. Fourth, we do not have listing reviews for VRBO listings nor did we consider hotels as an outside option in our utility estimation.[38] Fifth, our data analysis ends in December 2019, the same month when Airbnb announced its new review policy. Because we do not know exactly how Airbnb implements its new policy on "irrelevant" reviews, our simulations about no-disclosure and high-alert counterfactuals are hypothetical, and do not account for other changes in which Airbnb guests and hosts may engage should these counterfactual regimes happen in reality. In particular, we do not know how guests may readjust their beliefs regarding vicinity safety for all listings if they are fully aware that the lack of vicinity safety reviews is driven by a platform policy rather than user experience. In that case, they may seek safety information from alternative sources, and adjust their perspective regarding the vicinity safety of all listings on the platform.

These limitations suggest directions for future work. In particular, VRBO does not have a policy of discouraging reviews about the vicinity of listings, as Airbnb introduced in December 2019. This may facilitate an interesting comparison between VRBO and Airbnb listings in the same locales, given a sample period that encompasses Airbnb's implementation of its new review policy.

More broadly, the tradeoff we observe on Airbnb are becoming more common. YouTube, for instance, has recently adopted a policy of hiding dislike counts on shared videos,[39] and Instagram has considered giving users the option of hiding likes.[40] As digital platforms expand, tradeoffs regarding information disclosure are likely to attract more attention from researchers, user groups, and policymakers.

References

Barach, M. A., J. M. Golden, and J. J. Horton (2020): "Steering in online markets: the role of platform incentives and credibility," *Management Science*, 66, 4047–4070.

Berry, S. (1994): "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242–262.

[38] Hotels, in particular, may offer enhanced safety measures to their guests through security arrangements and by having door and security staff.

[39] See, e.g., `https://rb.gy/xhhqnd` .

[40] See, e.g., `https://rb.gy/tacuj5` .

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.

Bolton, G., B. Greiner, and A. Ockenfels (2013): "Engineering trust: reciprocity in the production of reputation information," *Management science*, 59, 265–285.

Cabral, L. and L. Li (2015): "A dollar for your thoughts: Feedback-conditional rebates on eBay," *Management Science*, 61, 2052–2063.

Chakravarty, A., Y. Liu, and T. Mazumdar (2010): "The differential effects of online word-of-mouth and critics' reviews on pre-release movie evaluation," *Journal of Interactive Marketing*, 24, 185–197.

Chevalier, J. A. and D. Mayzlin (2006): "The effect of word of mouth on sales: Online book reviews," *Journal of marketing research*, 43, 345–354.

Conitzer, V., N. Immorlica, J. Letchford, K. Munagala, and L. Wagman (2010): "False-name-proofness in social networks," in *International Workshop on Internet and Network Economics*, Springer, 209–221.

Conitzer, V. and L. Wagman (2014): "False-name-proof voting over two alternatives," *International Journal of Game Theory*, 43, 599–618.

Dai, W. D., G. Jin, J. Lee, and M. Luca (2018): "Aggregation of consumer ratings: an application to Yelp. com," *Quantitative Marketing and Economics*, 16, 289–339.

De Pelsmacker, P. and W. Janssens (2007): "A model for fair trade buying behaviour: The role of perceived quantity and quality of information and of product-specific attitudes," *Journal of business ethics*, 75, 361–380.

Dhaoui, C., C. M. Webster, and L. P. Tan (2017): "Social media sentiment analysis: lexicon versus machine learning," *Journal of Consumer Marketing*.

Einav, L., C. Farronato, and J. Levin (2016): "Peer-to-peer markets," *Annual Review of Economics*, 8, 615–635.

Ert, E., A. Fleischer, and N. Magen (2016): "Trust and reputation in the sharing economy: The role of personal photos in Airbnb," *Tourism management*, 55, 62–73.

Filieri, R., E. Raguseo, and C. Vitari (2021): "Extremely negative ratings and online consumer review helpfulness: the moderating role of product quality signals," *Journal of Travel Research*, 60, 699–717.

Fradkin, A., E. Grewal, and D. Holtz (2021): "Reciprocity and unveiling in two-sided reputation systems: Evidence from an experiment on airbnb," *Marketing Science*.

Fradkin, A., E. Grewal, D. Holtz, and M. Pearson (2015): "Bias and Reciprocity in Online Reviews: Evidence From Field Experiments on Airbnb." *EC*, 15, 15–19.

Gurran, N. and P. Phibbs (2017): "When tourists move in: how should urban planners respond to Airbnb?" *Journal of the American planning association*, 83, 80–92.

Han, W. and X. Wang (2019): "Does Home Sharing Impact Crime Rate? A Tale of Two Cities." in *ICIS*.

Han, W., X. Wang, M. Ahsen, and S. Wattal (2020): "Does Home Sharing Impact Crime Rate? An Empirical Investigation," *An Empirical Investigation (January 16, 2020)*.

Hao, L. and Y. Tan (2019): "Who wants consumers to be informed? Facilitating information disclosure in a distribution channel," *Information Systems Research*, 30, 34–49.

Hui, X., T. J. Klein, K. Stahl, et al. (2021): "When and Why Do Buyers Rate in Online Markets?" Tech. rep., University of Bonn and University of Mannheim, Germany.

Hutto, C. and E. Gilbert (2014): "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8.

Jia, J., G. Z. Jin, and L. Wagman (2021): "Platform as a Rule Maker: Evidence from Airbnb's Cancellation Policies," Working Paper.

Jia, J. and L. Wagman (2020): "Platform, Anonymity, and Illegal Actors: Evidence of Whac-a-Mole Enforcement from Airbnb," *The Journal of Law and Economics*, 63, 729–761.

Kim, J.-H., T. C. Leung, and L. Wagman (2017): "Can restricting property use be value enhancing? Evidence from short-term rental regulation," *The Journal of Law and Economics*, 60, 309–334.

Klein, T. J., C. Lambertz, G. Spagnolo, and K. O. Stahl (2009): "The actual structure of eBay's feedback mechanism and early evidence on the effects of recent changes," *International Journal of Electronic Business*, 7, 301–320.

Klein, T. J., C. Lambertz, and K. O. Stahl (2016): "Market transparency, adverse selection, and moral hazard," *Journal of Political Economy*, 124, 1677–1713.

Kovbasyuk, S. and G. Spagnolo (2018): "Memory and markets," *Available at SSRN 2756540*.

Lewis, G. (2011): "Asymmetric information, adverse selection and online disclosure: The case of eBay motors," *American Economic Review*, 101, 1535–46.

Li, L., S. Tadelis, and X. Zhou (2020): "Buying reputation as a signal of quality: Evidence from an online marketplace," *The RAND Journal of Economics*, 51, 965–988.

Li, L. and E. Xiao (2014): "Money talks: Rebate mechanisms in reputation system design," *Management Science*, 60, 2054–2072.

Liang, S., H. Li, X. Liu, and M. Schuckert (2019): "Motivators behind information disclosure: Evidence from Airbnb hosts," *Annals of Tourism Research*, 76, 305–319.

Liu, M., E. Brynjolfsson, and J. Dowlatabadi (2021): "Do digital platforms reduce moral hazard? The case of Uber and taxis," *Management Science*.

Liu, Y., J. Feng, and X. Liao (2017): "When online reviews meet sales volume information: is more or accurate information always better?" *Information Systems Research*, 28, 723–743.

Luca, M. and G. Zervas (2016): "Fake it till you make it: Reputation, competition, and Yelp review fraud," *Management Science*, 62, 3412–3427.

Maldonado-Guzmán, D. J. (2020): "Airbnb and crime in Barcelona (Spain): testing the relationship using a geographically weighted regression," *Annals of GIS*, 1–14.

Mayzlin, D., Y. Dover, and J. Chevalier (2014): "Promotional reviews: An empirical investigation of online review manipulation," *American Economic Review*, 104, 2421–55.

Monroe, B. L., M. P. Colaresi, and K. M. Quinn (2008): "Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict," *Political Analysis*, 16, 372–403.

Moon, Y. (2000): "Intimate exchanges: Using computers to elicit self-disclosure from consumers," *Journal of consumer research*, 26, 323–339.

Morosan, C. (2018): "Information disclosure to biometric e-gates: The roles of perceived security, benefits, and emotions," *Journal of Travel Research*, 57, 644–657.

Morosan, C. and A. DeFranco (2015): "Disclosing personal information via hotel apps: A privacy calculus perspective," *International Journal of Hospitality Management*, 47, 120–130.

Munzel, A. (2016): "Assisting consumers in detecting fake reviews: The role of identity information disclosure and consensus," *Journal of Retailing and Consumer Services*, 32, 96–108.

Nieuwland, S. and R. Van Melik (2020): "Regulating Airbnb: how cities deal with perceived negative externalities of short-term rentals," *Current Issues in Tourism*, 23, 811–825.

Nosko, C. and S. Tadelis (2015): "The limits of reputation in platform markets: An empirical analysis and field experiment," Tech. rep., National Bureau of Economic Research.

Pan, Y. and J. Q. Zhang (2011): "Born unequal: a study of the helpfulness of user-generated product reviews," *Journal of retailing*, 87, 598–612.

Reimers, I. and J. Waldfogel (2021): "Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings," *American Economic Review*, 111, 1944–71.

Romanyuk, G. and A. Smolin (2019): "Cream skimming and information design in matching markets," *American Economic Journal: Microeconomics*, 11, 250–76.

Roth, J. J. (2021): "Home Sharing and Crime Across Neighborhoods: An Analysis of Austin, Texas," *Criminal justice review*, 46, 40–52.

Schuckert, M., X. Liu, and R. Law (2015): "Hospitality and tourism online reviews: Recent trends and future directions," *Journal of Travel & Tourism Marketing*, 32, 608–621.

Staats, B. R., H. Dai, D. Hofmann, and K. L. Milkman (2017): "Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare," *Management Science*, 63, 1563–1585.

Suess, C., K. M. Woosnam, and E. Erul (2020): "Stranger-danger? Understanding the moderating effects of children in the household on non-hosting residents' emotional solidarity with Airbnb visitors, feeling safe, and support for Airbnb," *Tourism Management*, 77, 103952.

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

Tussyadiah, I. P. and S. Park (2018): "When guests trust hosts for their words: Host description and trust in sharing economy," *Tourism Management*, 67, 261–272.

Ursu, R. M. (2018): "The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions," *Marketing Science*, 37, 530–552.

Wagman, L. and V. Conitzer (2008): "Optimal False-Name-Proof Voting Rules with Costly Voting." in *AAAI*, 190–195.

Xu, X., S. Zeng, and Y. He (2021): "The impact of information disclosure on consumer purchase behavior on sharing economy platform Airbnb," *International Journal of Production Economics*, 231, 107846.

Xu, Y.-H., L. Pennington-Gray, and J. Kim (2019): "The sharing economy: A geographically weighted regression approach to examine crime and the shared lodging sector," *Journal of travel research*, 58, 1193–1208.

Zervas, G., D. Proserpio, and J. W. Byers (2021): "A first look at online reputation on Airbnb, where every stay is above average," *Marketing Letters*, 32, 1–16.

Zhang, H., W. Gan, and B. Jiang (2014): "Machine learning and lexicon based methods for sentiment classification: A survey," in *2014 11th web information system and application conference*, IEEE, 262–265.

Zhang, S., D. Lee, P. Singh, and T. Mukhopadhyay (2021a): "EXPRESS: Demand Interactions in Sharing Economies: Evidence from a Natural Experiment Involving Airbnb and Uber/Lyft," *Journal of Marketing Research*, 00222437211062172.

Zhang, S., D. Lee, P. V. Singh, and K. Srinivasan (2021b): "What Makes a Good Image? Airbnb Demand Analytics Leveraging Interpretable Image Features," *Management Science*.

Zhang, S., N. Mehta, P. V. Singh, and K. Srinivasan (2021c): "Can an AI Algorithm Mitigate Racial Economic Inequality? An Analysis in the Context of Airbnb," *An Analysis in the Context of Airbnb (January 21, 2021)*.

Appendix

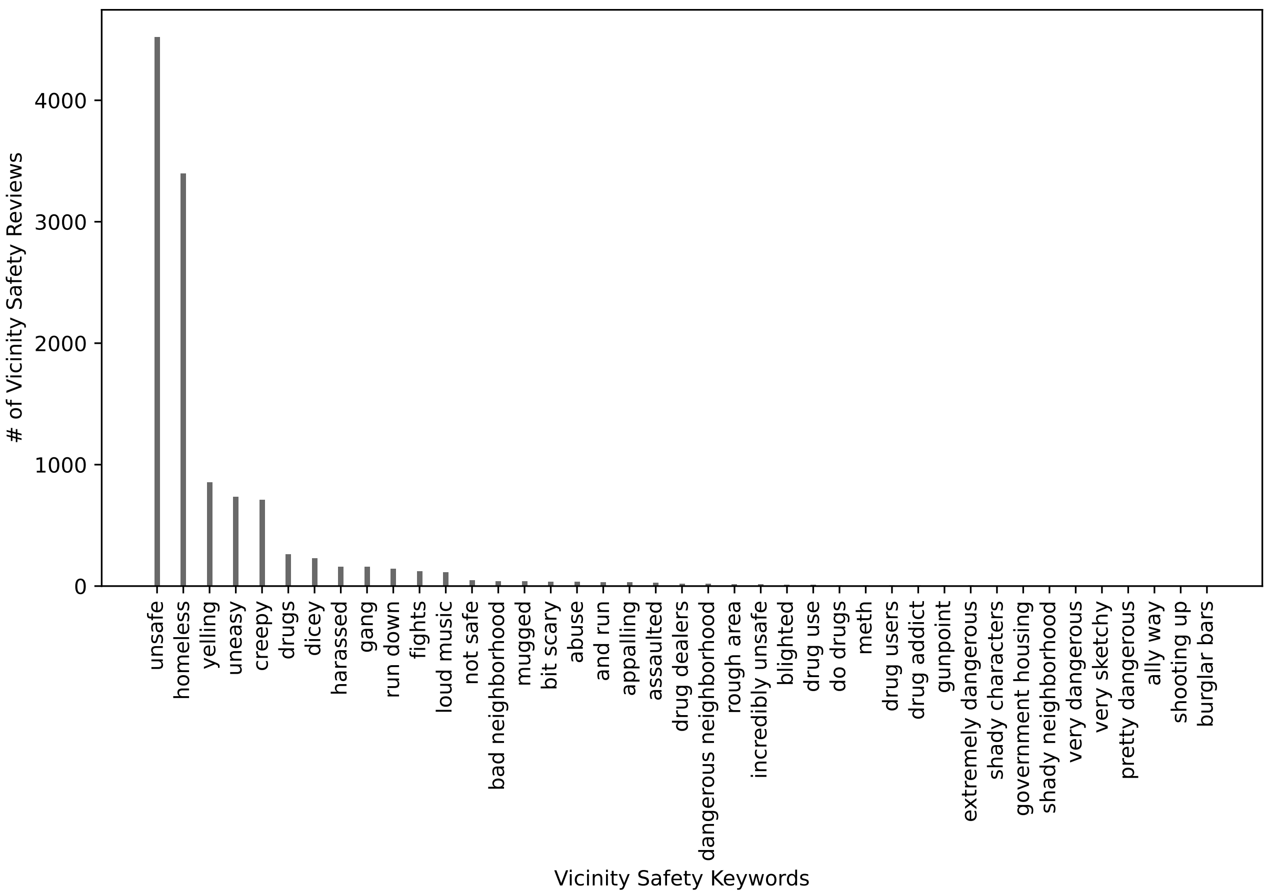

Figure 1: Distribution for keywords of vicinity safety review

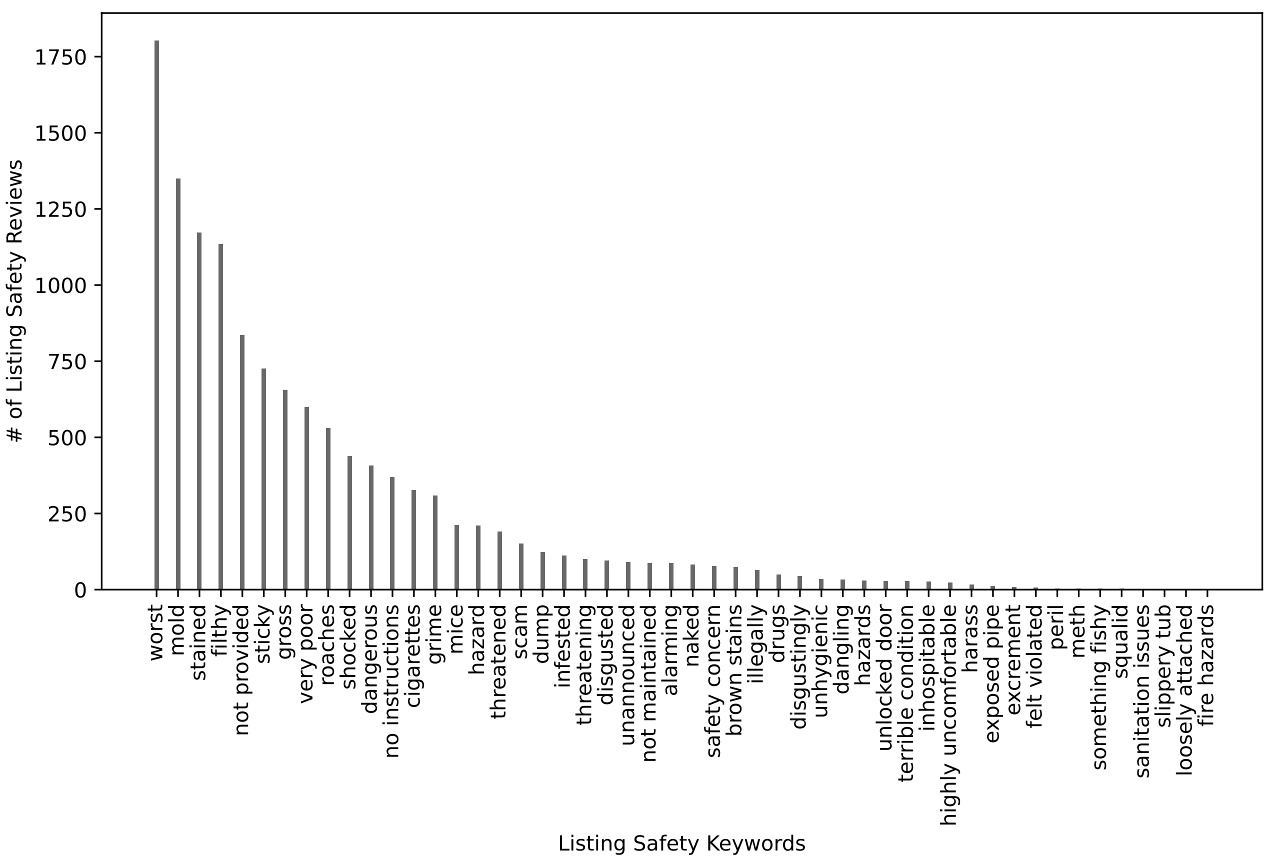

Figure 2: Distribution for keywords of listing safety review

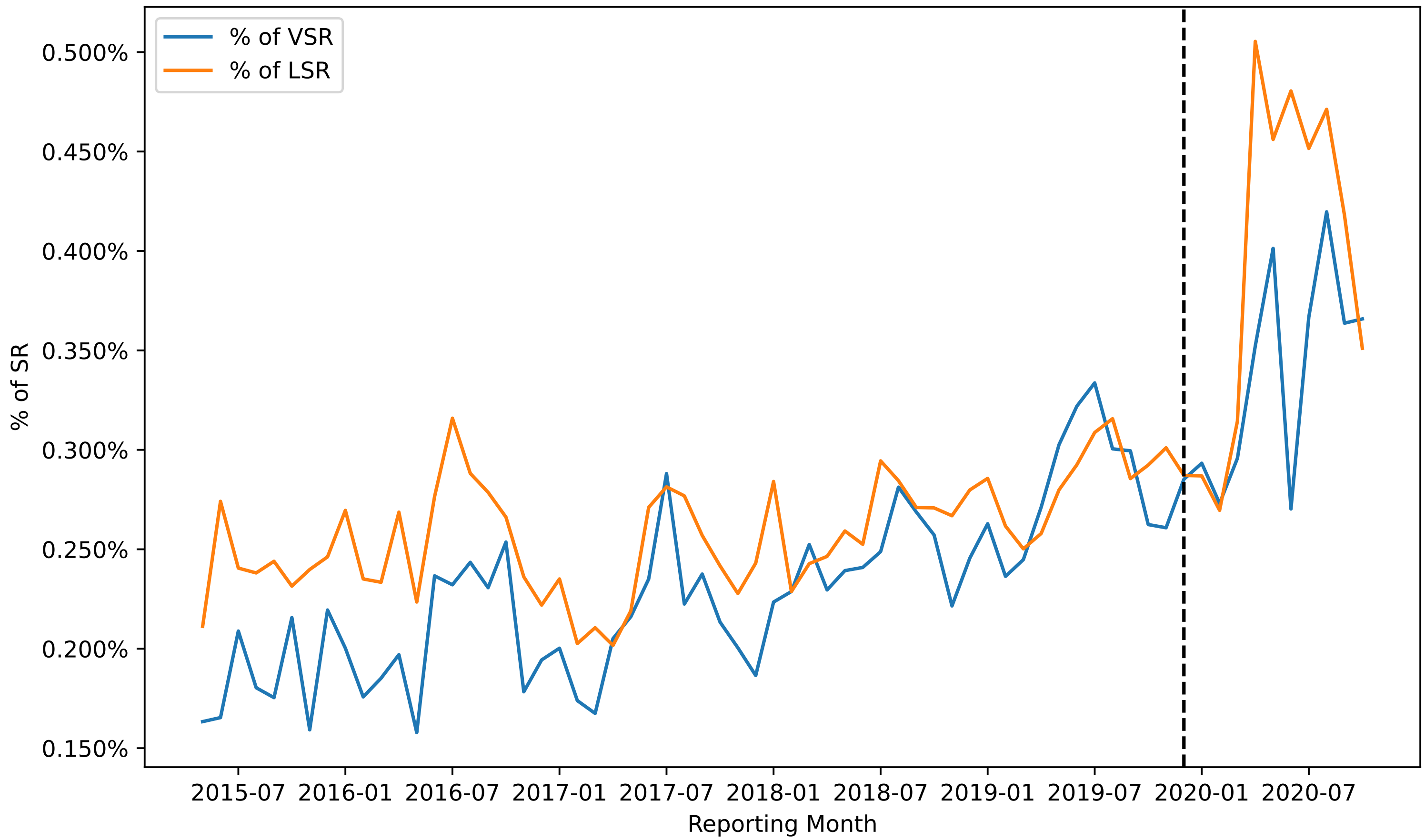

Figure 3: Percentage of safety review over time

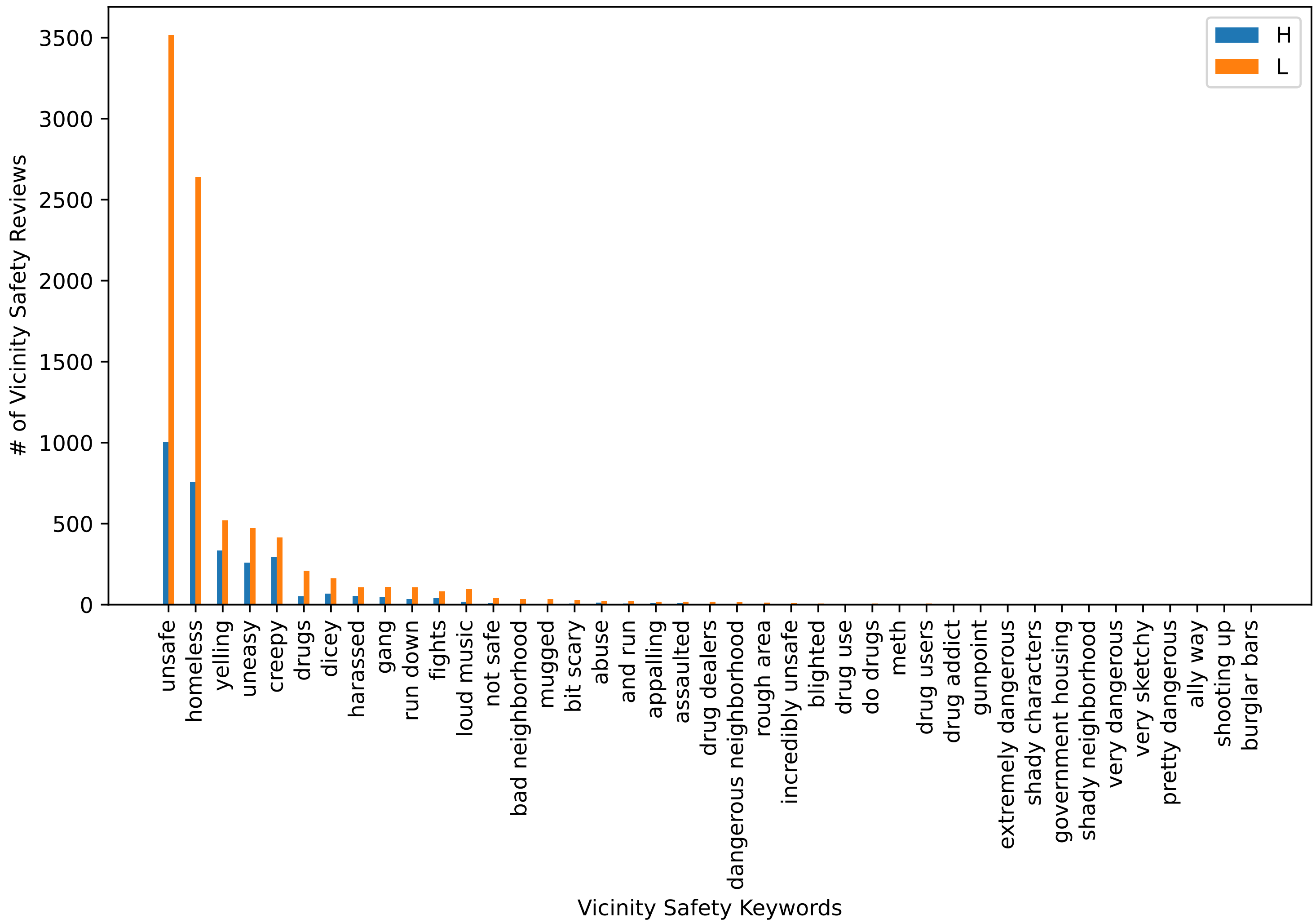

Figure 4: Distribution for keywords of vicinity safety review in H & L zip codes

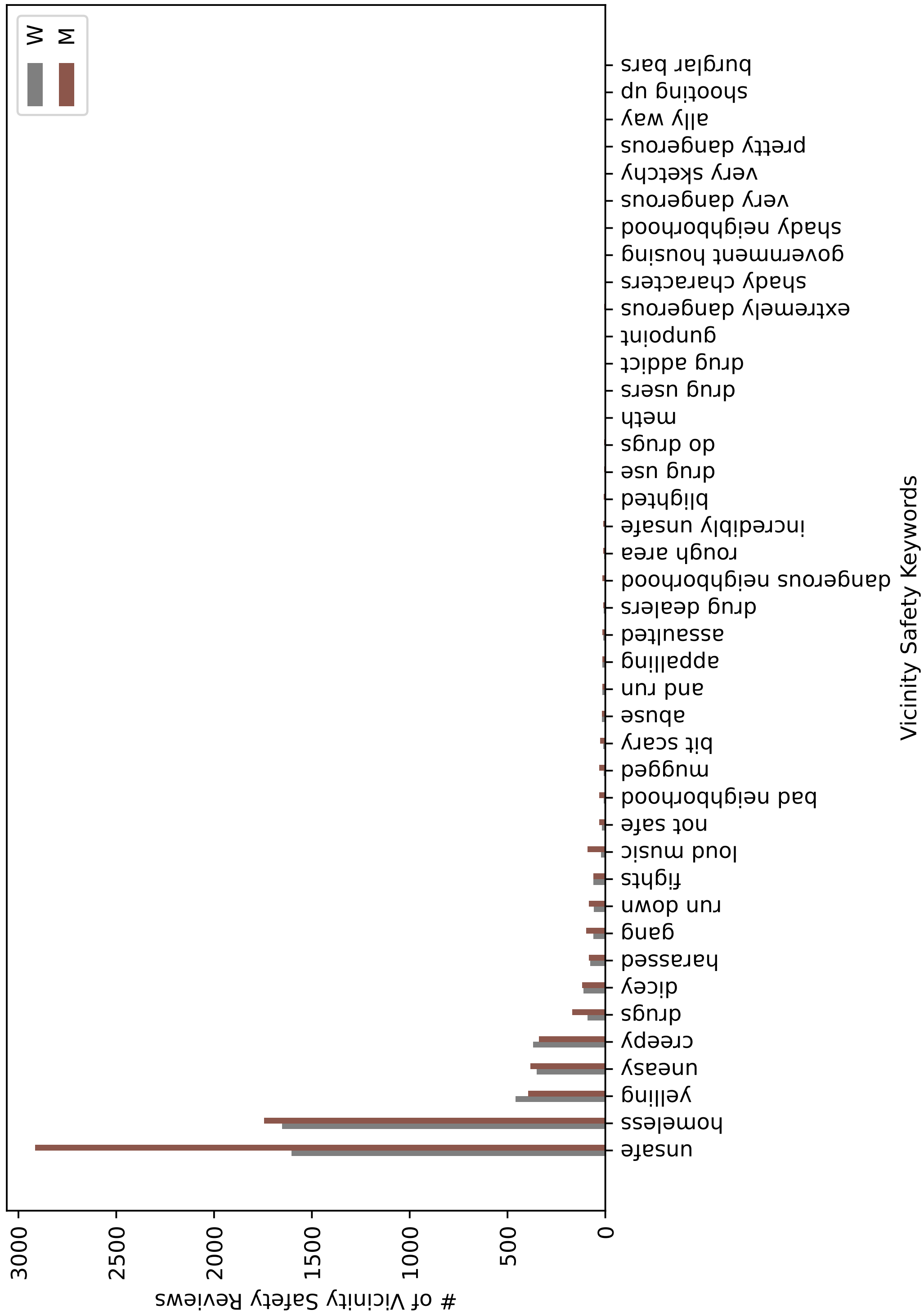

Figure 5: Distribution for keywords of vicinity safety review in W & M zip codes

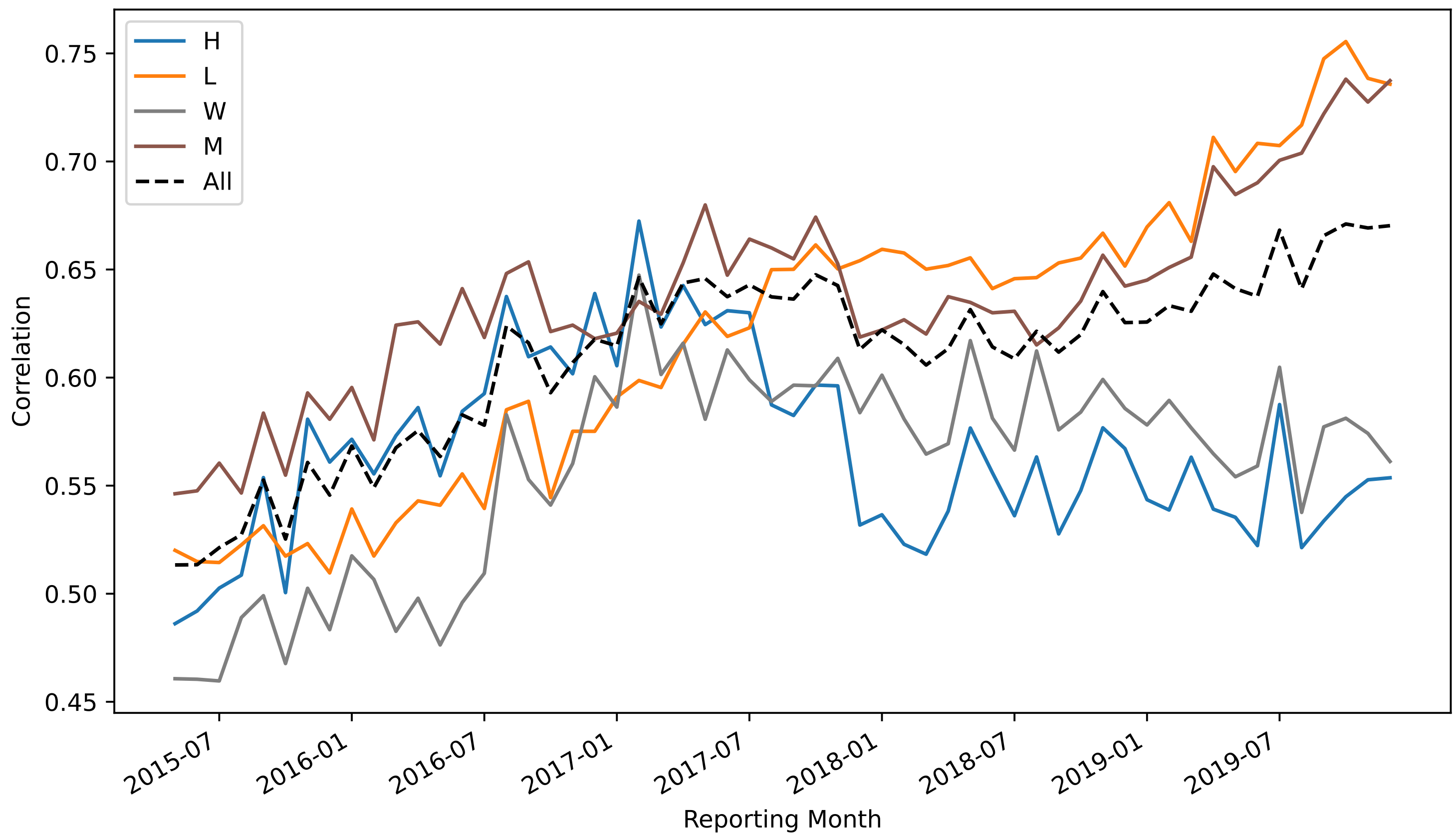

Figure 6: Correlation between the rank of normalized crime flow and the normalized total vicinity safety reviews

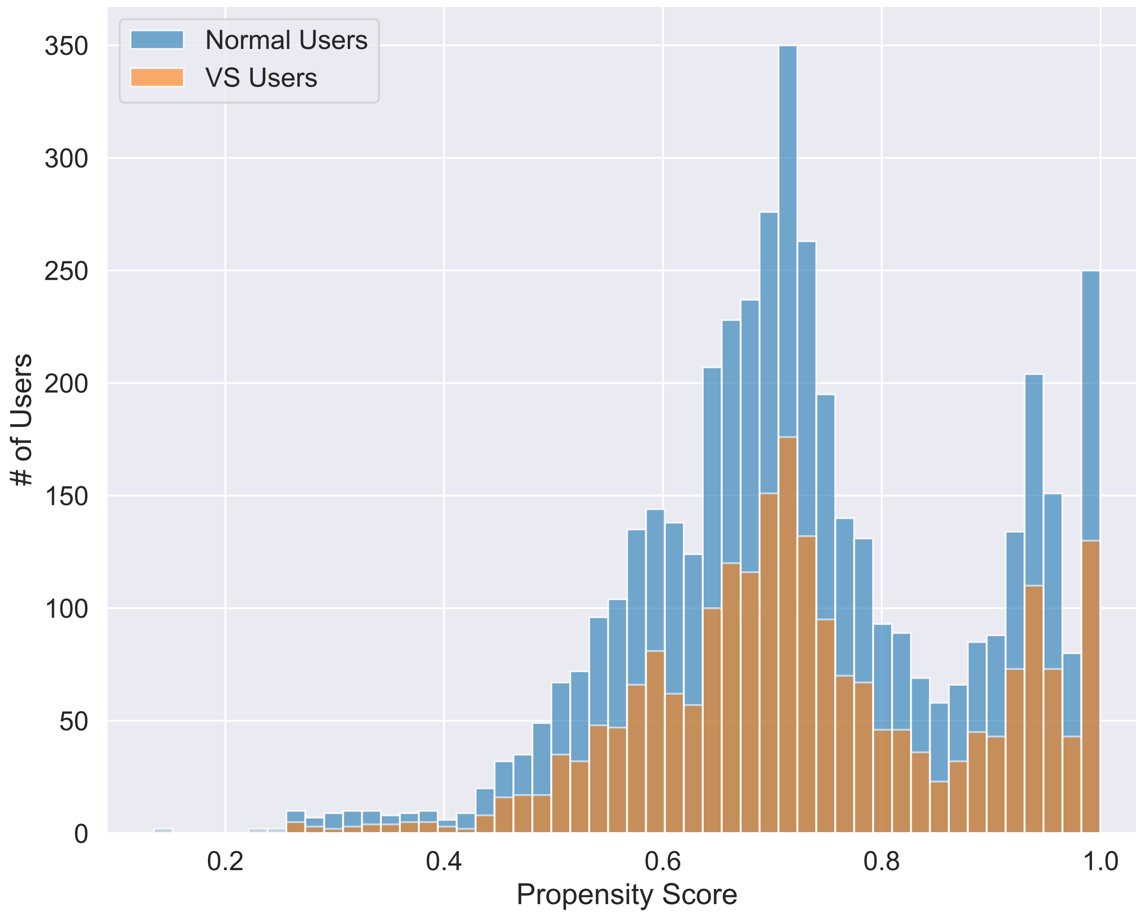

Figure 7: Distribution of Propensity Score for Control and Treatment Group

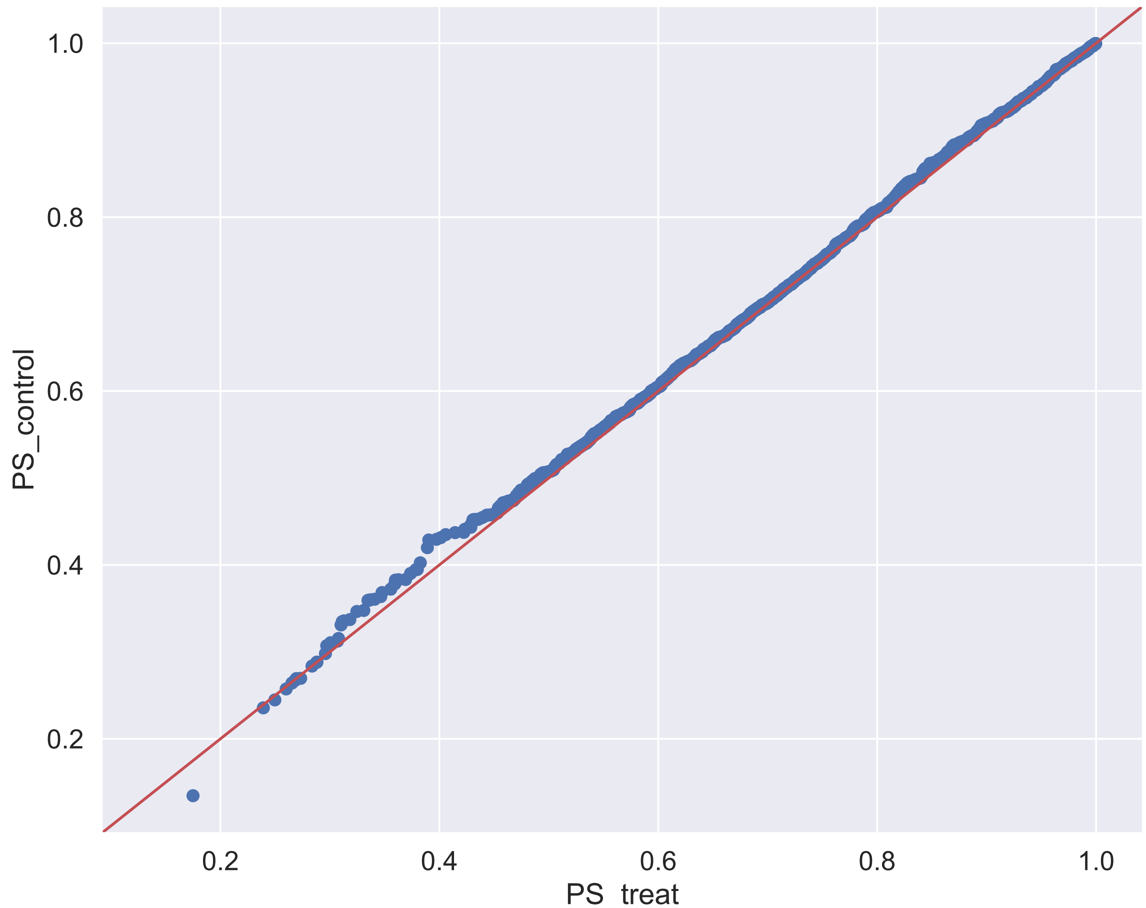

Figure 8: Quantile – Quantile Plot of Propensity Score for Control and Treatment Group (cont.)

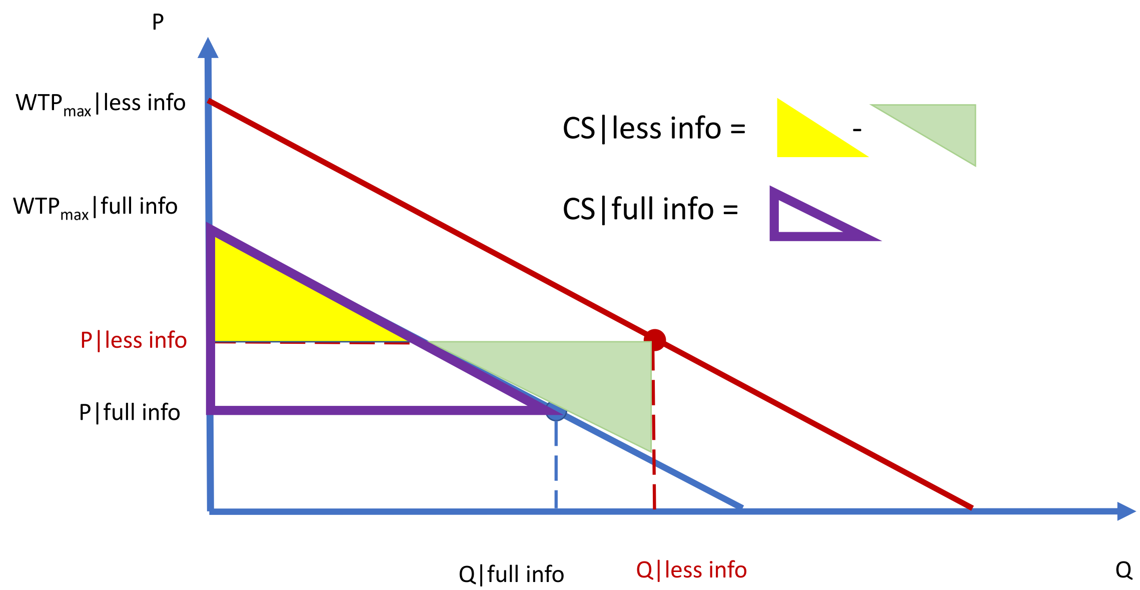

Figure 9: Consumer Surplus

Vicinity safety keywords:	‘abuse’, ‘ally way’, ‘and run’, ‘appalling’, ‘assaulted’, ‘bad neighborhood’, ‘bit scary’, ‘blighted’, ‘burglar bars’, ‘creepy’, ‘dangerous neighborhood’, ‘not safe’, ‘dicey’, ‘do drugs’, ‘drug addict’, ‘drug dealers’, ‘drug use’, ‘drug users’, ‘drugs’, ‘extremely dangerous’, ‘fights’, ‘gang’, ‘government housing’, ‘gunpoint’, ‘harassed’, ‘homeless’, ‘incredibly unsafe’, ‘loud music’, ‘meth’, ‘mugged’, ‘pretty dangerous’, ‘rough area’, ‘run down’, ‘shady characters’, ‘shady neighborhood’, ‘shooting up’, ‘tenement area’, ‘uneasy’, ‘unsafe’, ‘very sketchy’, ‘yelling’
Listing safety keywords:	‘alarming’, ‘threatening’, ‘brown stains’, ‘cigarettes’, ‘dangerous’, ‘dangling’, ‘peril’, ‘disgusted’, ‘disgustingly’, ‘drugs’, ‘dump’, ‘excrement’, ‘exposed pipe’, ‘felt violated’, ‘filthy’, ‘fire hazards’, ‘something fishy’, ‘very poor’, ‘mold’, ‘grime’, ‘not maintained’, ‘gross’, ‘harass’ ‘hazard’, ‘hazards’, ‘highly uncomfortable’, ‘safety concern’, ‘illegally’, ‘infested’, ‘inhospitable’, ‘loosely attached’, ‘meth’, ‘mice’, ‘naked’, ‘no instructions’, ‘not provided’, ‘scam’, ‘unhygienic’, ‘roaches’, ‘sanitation issues’, ‘shocked’, ‘slippery tub’, ‘squalid’, ‘stained’, ‘sticky’, ‘terrible condition’, ‘threatened’, ‘unannounced’, ‘unlocked door’, ‘worst’,
Vicinity location keywords:	‘neighborhood’, ‘area’, ‘feel’, ‘felt’, ‘night’, ‘location’, ‘walking’, ‘people’, ‘seemed’, ‘outside’, ‘looked’, ‘looks’, ‘late’, ‘surrounding’, ‘located’, ‘neighbourhood’, ‘walked’, ‘areas’, ‘feeling’, ‘streets’, ‘street’, ‘outside’, ‘parking’, ‘neighbors’
Negative keywords:	‘hardly’, ‘never’, ‘scarcely’, ‘seldom’, ‘barely’, ‘no’, ‘not’, ‘without’, ‘nothing’, ‘nobody’, ‘neither’, ‘nor’, ‘none’

Table 1: Vicinity and listing safety review keywords

	Panel A: All listings				Panel B: VS listings				Panel C: Normal listings			
VARIABLES	mean	p50	sd	N	mean	p50	sd	N	mean	p50	sd	N
occupancyrate	0.56	0.64	0.36	2,866,238	0.68	0.78	0.30	126,868	0.56	0.64	0.36	2,739,370
adrusd	164.69	125.51	173.53	2,866,238	134.15	106.31	107.29	126,868	166.10	126.67	175.87	2,739,370
numberofreservations	3.77	3.00	3.35	2,866,238	5.76	5.00	3.76	126,868	3.68	3.00	3.31	2,739,370
reservationdays	14.16	14.00	10.41	2,866,238	18.56	21.00	9.29	126,868	13.95	14.00	10.42	2,739,370
ratingoverall	9.18	9.60	1.50	2,866,238	9.09	9.20	0.73	126,868	9.18	9.60	1.53	2,739,370
ratingcommunication	9.55	10.00	1.50	2,866,238	9.61	10.00	0.68	126,868	9.54	10.00	1.52	2,739,370
ratingaccuracy	9.38	10.00	1.55	2,866,238	9.41	10.00	0.76	126,868	9.38	10.00	1.58	2,739,370
ratingcleanliness	9.16	10.00	1.61	2,866,238	9.17	9.00	0.89	126,868	9.16	10.00	1.64	2,739,370
ratingcheckin	9.53	10.00	1.51	2,866,238	9.64	10.00	0.66	126,868	9.53	10.00	1.54	2,739,370
ratinglocation	9.31	10.00	1.53	2,866,238	9.03	9.00	0.89	126,868	9.33	10.00	1.55	2,739,370
ratingvalue	9.20	10.00	1.56	2,866,238	9.17	9.00	0.76	126,868	9.20	10.00	1.59	2,739,370
median_income_zipcode	57,187.40	50,943.00	26,179.42	2,866,238	42,644.64	34,432.00	21,641.93	126,868	57,860.92	51,427.00	26,175.62	2,739,370
population_zipcode	48,157.80	45,747.00	24,505.14	2,866,238	42,514.47	36,654.00	25,184.66	126,868	48,419.16	46,025.00	24,441.66	2,739,370
white_percent_zipcode	0.53	0.59	0.24	2,866,238	0.41	0.38	0.22	126,868	0.53	0.60	0.24	2,739,370
h_zip	0.52	1.00	0.50	2,866,238	0.29	0.00	0.45	126,868	0.53	1.00	0.50	2,739,370
w_zip	0.60	1.00	0.49	2,866,238	0.44	0.00	0.50	126,868	0.61	1.00	0.49	2,739,370
crimewhole_cumu	19,435.43	9,650.00	34,638.89	2,866,238	31,230.05	14,205.00	45,889.33	126,868	18,889.18	9,475.00	33,928.66	2,739,370
review_utd	33.71	15.00	49.49	2,866,238	93.02	70.00	83.73	126,868	30.96	14.00	45.47	2,739,370
listing_size_zip	540.67	449.00	416.46	2,866,238	554.66	481.00	388.87	126,868	540.02	447.00	417.69	2,739,370
cross_listing	0.02	0.00	0.15	2,866,238	0.03	0.00	0.16	126,868	0.02	0.00	0.15	2,739,370
occ_dummy	0.85	1.00	0.36	2,866,238	0.95	1.00	0.21	126,868	0.85	1.00	0.36	2,739,370
lag_vicinity_sr_cumu_dummy	0.04	0.00	0.21	2,866,238	1.00	1.00	0.00	126,868	0.00	0.00	0.00	2,739,370
lag_listing_sr_cumu_dummy	0.05	0.00	0.22	2,866,238	0.20	0.00	0.40	126,868	0.04	0.00	0.20	2,739,370
lag_vicinity_listing_radius_pct	0.07	0.04	0.11	2,866,238	0.10	0.07	0.11	126,868	0.07	0.03	0.11	2,739,370
superhost_dummy	0.23	0.00	0.42	2,866,238	0.26	0.00	0.44	126,868	0.23	0.00	0.42	2,739,370
strict_cp	0.50	0.00	0.50	2,866,238	0.58	1.00	0.49	126,868	0.49	0.00	0.50	2,739,370
ave_review_word_count_cumu	53.83	50.43	30.65	2,866,238	57.49	53.91	22.43	126,868	53.66	50.20	30.97	2,739,370
lag_vicinity_sr_cumu	0.06	0.00	0.34	2,866,238	1.34	1.00	0.96	126,868	0.00	0.00	0.00	2,739,370
lag_listing_sr_cumu	0.06	0.00	0.28	2,866,238	0.26	0.00	0.60	126,868	0.05	0.00	0.25	2,739,370
safety_score	4.96	5.09	2.81	2,866,238	2.83	2.33	2.25	126,868	5.06	5.23	2.79	2,739,370

Table 2: Data summary statistics – listing characteristics at listing-month level

	(1)	(2)
VARIABLES	log_occupancy_rate	log_adr
lag_vicinity_sr_cumu_dummy	-0.0182***	-0.0148***
	(0.00140)	(0.00219)
lag_listing_sr_cumu_dummy	-0.0258***	-0.0152***
	(0.00135)	(0.00210)
lag_vicinity_listing_radius_pct	-0.00859***	-0.00872**
	(0.00253)	(0.00390)
lag_log_crimewhole_cumu_norm	0.0693***	-0.0508***
	(0.00826)	(0.0130)
lag_log_review_utd	0.00420***	0.0117***
	(0.000415)	(0.000678)
log_listing_size_zip	-0.0212***	0.0146***
	(0.00185)	(0.00289)
log_rating_overall	0.0257***	-0.00240
	(0.00128)	(0.00200)
superhost_dummy	0.0175***	0.00817***
	(0.000586)	(0.000845)
cross_listing	0.0311***	-0.00564
	(0.00278)	(0.00384)
strict_cp	0.000601	0.0123***
	(0.000803)	(0.00126)
Constant	0.412***	4.762***
	(0.0119)	(0.0185)
Observations	2,866,238	2,866,238
R-squared	0.559	0.928
Time*City FE	Yes	Yes
PropertyID FE	Yes	Yes
CSEs	PropertyID	PropertyID
Sample	whole	whole

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Listing level: Regressions with all Airbnb listings, using the main specification.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	occ_dummy	log occupancy_rate	log occupancy_rate	log occupancy_rate	log occupancy_rate	log adr	log adr	log adr	log adr
lag_vicinity_sr_cumu_dummy	-0.0132***	-0.0129***	-0.0212***	-0.0100***	-0.0180***	-0.0126***	-0.00411	-0.0110***	-0.0150***
	(0.00155)	(0.00119)	(0.00541)	(0.00146)	(0.00140)	(0.00201)	(0.00726)	(0.00231)	(0.00219)
lag_listing_sr_cumu_dummy	-0.0131***	-0.0213***	-0.0362***	-0.0173***	-0.0251***	-0.0112***	-0.00152	-0.0124***	-0.0158***
	(0.00153)	(0.00114)	(0.00458)	(0.00142)	(0.00135)	(0.00189)	(0.00705)	(0.00218)	(0.00210)
lag_vicinity_listing_radius_pct	-0.0100**	-0.00575**	-0.00760**	-0.00334	-0.00864***	-0.00848**	0.00263	-0.0140***	-0.00868**
	(0.00416)	(0.00238)	(0.00378)	(0.00342)	(0.00253)	(0.00337)	(0.00620)	(0.00478)	(0.00390)
lag_log_crimewhole_cumu_norr	0.180***	-0.0167**	0.219***	0.0118	0.0693***	-0.000974	-0.0600**	-0.0242	-0.0508***
	(0.0123)	(0.00734)	(0.0150)	(0.0106)	(0.00826)	(0.0120)	(0.0240)	(0.0168)	(0.0130)
lag_log_review_utd	-0.00881***	0.00582***	-0.00433***	-0.0111***	0.00474***	0.0170***	0.0113***	0.0163***	0.0113***
	(0.000592)	(0.000341)	(0.000739)	(0.000754)	(0.000418)	(0.000557)	(0.00119)	(0.00134)	(0.000687)
log_listing_size_zip	-0.0284***	-0.00953***	-0.0233***	-0.0169***	-0.0212***	0.00569**	0.0287***	0.00899***	0.0146***
	(0.00279)	(0.00160)	(0.00328)	(0.00230)	(0.00185)	(0.00251)	(0.00578)	(0.00338)	(0.00289)
log_rating_overall	0.0390***	0.0162***	0.0182***	0.0321***	0.0258***	-0.00613***	-0.0136***	0.00377	-0.00244
	(0.00242)	(0.00112)	(0.00137)	(0.00339)	(0.00128)	(0.00162)	(0.00214)	(0.00375)	(0.00200)
superhost_dummy	0.0161***	0.0123***	0.0109***	0.0158***	0.0172***	0.0146***	0.00608***	0.00723***	0.00835***
	(0.000809)	(0.000482)	(0.00139)	(0.000674)	(0.000586)	(0.000752)	(0.00195)	(0.000970)	(0.000845)
cross_listing	0.0233***	0.0235***	0.0435***	0.0225***	0.0310***	-0.000407	-0.0293***	0.00335	-0.00554
	(0.00382)	(0.00220)	(0.00541)	(0.00328)	(0.00278)	(0.00355)	(0.00741)	(0.00482)	(0.00384)
strict_cp	0.00324***	-0.00126*	0.00654***	-0.00670***	0.000749	0.00480***	0.0121***	0.0108***	0.0122***
	(0.00118)	(0.000645)	(0.00141)	(0.000994)	(0.000803)	(0.00105)	(0.00233)	(0.00151)	(0.00126)
lag_log_ave_r_wct_cumu					-0.00890***				0.00645***
					(0.000744)				(0.00106)
Constant	0.904***	0.459***	0.352***	0.509***	0.443***	4.783***	4.719***	4.720***	4.739***
	(0.0180)	(0.0103)	(0.0205)	(0.0168)	(0.0122)	(0.0160)	(0.0358)	(0.0237)	(0.0189)
Observations	2,866,238	2,441,566	1,370,655	1,495,583	2,866,238	2,441,566	1,370,655	1,495,583	2,866,238
R-squared	0.420	0.499	0.565	0.522	0.559	0.943	0.931	0.937	0.928
Time*City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PropertyID FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CSEs	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID
Sample	whole	occ>0	review_utd<=13	review_utd>13	whole	occ>0	review_utd<=13	review_utd>13	whole

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Listing Level: Regressions for robustness check

	Panel A: H				Panel B: L				Panel C: W				Panel D: M			
VARIABLES	mean	p50	sd	N	mean	p50	sd	N	mean	p50	sd	N	mean	p50	sd	N
occupancyrate	0.56	0.64	0.36	1,484,474	0.57	0.65	0.36	1,381,764	0.56	0.64	0.36	1,716,774	0.57	0.65	0.36	1,149,464
adrusd	190.38	149.00	184.23	1,484,474	137.08	103.62	156.63	1,381,764	188.33	147.48	181.88	1,716,774	129.37	98.07	153.63	1,149,464
numberofreservations	3.65	3.00	3.31	1,484,474	3.91	3.00	3.40	1,381,764	3.71	3.00	3.30	1,716,774	3.87	3.00	3.43	1,149,464
reservationdays	13.86	14.00	10.40	1,484,474	14.48	15.00	10.42	1,381,764	13.99	14.00	10.37	1,716,774	14.41	15.00	10.48	1,149,464
ratingoverall	9.23	9.60	1.49	1,484,474	9.13	9.50	1.52	1,381,764	9.23	9.60	1.48	1,716,774	9.11	9.50	1.53	1,149,464
ratingcommunication	9.57	10.00	1.49	1,484,474	9.52	10.00	1.51	1,381,764	9.57	10.00	1.48	1,716,774	9.52	10.00	1.52	1,149,464
ratingaccuracy	9.41	10.00	1.54	1,484,474	9.35	10.00	1.56	1,381,764	9.41	10.00	1.53	1,716,774	9.34	10.00	1.58	1,149,464
ratingcleanliness	9.21	10.00	1.59	1,484,474	9.11	10.00	1.63	1,381,764	9.21	10.00	1.58	1,716,774	9.09	10.00	1.65	1,149,464
ratingcheckin	9.55	10.00	1.50	1,484,474	9.51	10.00	1.52	1,381,764	9.55	10.00	1.50	1,716,774	9.50	10.00	1.53	1,149,464
ratinglocation	9.53	10.00	1.47	1,484,474	9.08	9.00	1.56	1,381,764	9.52	10.00	1.47	1,716,774	9.01	9.00	1.58	1,149,464
ratingvalue	9.22	10.00	1.55	1,484,474	9.17	10.00	1.57	1,381,764	9.22	10.00	1.55	1,716,774	9.16	10.00	1.59	1,149,464
median_income_zipcode	75,865.21	71,278.00	22,763.12	1,484,474	37,121.22	35,112.00	9,356.77	1,381,764	69,745.17	68,346.00	25,639.20	1,716,774	38,431.83	37,116.00	11,826.75	1,149,464
population_zipcode	43,534.94	41,453.00	20,778.30	1,484,474	53,124.28	51,791.00	27,095.86	1,381,764	44,706.36	38,752.00	24,110.83	1,716,774	53,312.67	54,440.00	24,181.98	1,149,464
white_percent_zipcode	0.68	0.72	0.16	1,484,474	0.37	0.32	0.20	1,381,764	0.69	0.72	0.13	1,716,774	0.28	0.30	0.13	1,149,464
h_zip	1.00	1.00	0.00	1,484,474	0.00	0.00	0.00	1,381,764	0.78	1.00	0.42	1,716,774	0.13	0.00	0.33	1,149,464
w_zip	0.90	1.00	0.30	1,484,474	0.27	0.00	0.45	1,381,764	1.00	1.00	0.00	1,716,774	0.00	0.00	0.00	1,149,464
crimewhole_cumu	14,737.33	7,735.00	26,575.65	1,484,474	24,482.74	12,205.00	40,999.30	1,381,764	20,925.04	8,347.00	40,906.05	1,716,774	17,210.63	11,637.00	22,010.67	1,149,464
review_utd	33.07	14.00	48.89	1,484,474	34.40	15.00	50.10	1,381,764	33.65	15.00	49.50	1,716,774	33.80	15.00	49.46	1,149,464
listing_size_zip	502.69	463.00	306.34	1,484,474	581.47	428.00	505.71	1,381,764	609.92	512.00	457.86	1,716,774	437.24	372.00	318.63	1,149,464
cross_listing	0.02	0.00	0.16	1,484,474	0.02	0.00	0.14	1,381,764	0.03	0.00	0.16	1,716,774	0.02	0.00	0.13	1,149,464
occ_dummy	0.84	1.00	0.36	1,484,474	0.86	1.00	0.35	1,381,764	0.85	1.00	0.36	1,716,774	0.85	1.00	0.35	1,149,464
lag_vicinity_sr_cumu_dummy	0.02	0.00	0.15	1,484,474	0.07	0.00	0.25	1,381,764	0.03	0.00	0.18	1,716,774	0.06	0.00	0.24	1,149,464
lag_listing_sr_cumu_dummy	0.05	0.00	0.21	1,484,474	0.05	0.00	0.22	1,381,764	0.05	0.00	0.22	1,716,774	0.05	0.00	0.22	1,149,464
lag_vicinity_listing_radius_pct	0.05	0.02	0.10	1,484,474	0.09	0.05	0.12	1,381,764	0.05	0.03	0.09	1,716,774	0.09	0.05	0.13	1,149,464
superhost_dummy	0.25	0.00	0.43	1,484,474	0.22	0.00	0.42	1,381,764	0.24	0.00	0.43	1,716,774	0.22	0.00	0.42	1,149,464
strict_cp	0.50	0.00	0.50	1,484,474	0.49	0.00	0.50	1,381,764	0.51	1.00	0.50	1,716,774	0.48	0.00	0.50	1,149,464
ave_review_word_count_cumu	53.93	50.67	30.62	1,484,474	53.72	50.15	30.68	1,381,764	54.40	51.18	30.47	1,716,774	52.98	49.25	30.89	1,149,464
lag_vicinity_sr_cumu	0.03	0.00	0.20	1,484,474	0.09	0.00	0.44	1,381,764	0.04	0.00	0.29	1,716,774	0.08	0.00	0.40	1,149,464
lag_listing_sr_cumu	0.05	0.00	0.27	1,484,474	0.06	0.00	0.28	1,381,764	0.06	0.00	0.28	1,716,774	0.06	0.00	0.27	1,149,464
safety_score	5.62	6.15	2.60	1,484,474	4.26	3.67	2.85	1,381,764	5.31	5.76	2.75	1,716,774	4.45	3.87	2.82	1,149,464

Table 5: Data summary statistics – listing characteristics at listing–month level for subgroups

VARIABLES	Panel E: Entire Home mean	p50	sd	N	Panel F: Private Room mean	p50	sd	N	Panel G: Shared Room mean	p50	sd	N	Panel H: Hotel Room mean	p50	sd	N
occupancyrate	0.58	0.67	0.35	1,745,432	0.55	0.63	0.38	1,016,553	0.44	0.41	0.37	94,722	0.46	0.43	0.33	9,531
adrusd	212.81	170.46	175.04	1,745,432	91.67	76.25	140.69	1,016,553	58.23	39.36	128.74	94,722	197.16	153.87	250.23	9,531
numberofreservations	3.86	3.00	3.25	1,745,432	3.65	3.00	3.49	1,016,553	3.31	2.00	3.46	94,722	5.79	5.00	4.79	9,531
reservationdays	14.45	15.00	10.17	1,745,432	13.93	14.00	10.79	1,016,553	11.44	9.00	10.31	94,722	12.53	11.00	9.59	9,531
ratingoverall	9.26	9.60	1.36	1,745,432	9.09	9.50	1.67	1,016,553	8.74	9.30	1.94	94,722	9.03	9.40	1.35	9,531
ratingcommunication	9.61	10.00	1.35	1,745,432	9.48	10.00	1.67	1,016,553	9.18	10.00	1.96	94,722	9.38	10.00	1.35	9,531
ratingaccuracy	9.47	10.00	1.40	1,745,432	9.29	10.00	1.72	1,016,553	8.91	10.00	2.01	94,722	9.26	10.00	1.42	9,531
ratingcleanliness	9.27	10.00	1.45	1,745,432	9.02	10.00	1.79	1,016,553	8.64	9.00	2.05	94,722	9.24	10.00	1.34	9,531
ratingcheckin	9.59	10.00	1.37	1,745,432	9.47	10.00	1.69	1,016,553	9.16	10.00	1.98	94,722	9.47	10.00	1.24	9,531
ratinglocation	9.43	10.00	1.38	1,745,432	9.15	10.00	1.71	1,016,553	8.87	9.00	1.97	94,722	9.46	10.00	1.11	9,531
ratingvalue	9.25	10.00	1.42	1,745,432	9.14	10.00	1.73	1,016,553	8.86	9.00	2.00	94,722	9.01	9.00	1.49	9,531
median_income_zipcode	59,726.31	54,023.00	27,359.80	1,745,432	53,568.49	47,050.00	23,576.57	1,016,553	48,929.01	40,873.00	23,905.81	94,722	60,290.69	56,337.00	28,145.90	9,531
population_zipcode	44,465.10	38,752.00	23,812.97	1,745,432	54,260.35	54,440.00	24,480.23	1,016,553	52,002.62	48,852.00	23,587.40	94,722	35,315.05	30,648.00	22,650.94	9,531
white_percent_zipcode	0.56	0.61	0.23	1,745,432	0.48	0.46	0.24	1,016,553	0.45	0.44	0.23	94,722	0.55	0.60	0.20	9,531
h_zip	0.59	1.00	0.49	1,745,432	0.41	0.00	0.49	1,016,553	0.37	0.00	0.48	94,722	0.59	1.00	0.49	9,531
w_zip	0.68	1.00	0.47	1,745,432	0.48	0.00	0.50	1,016,553	0.43	0.00	0.49	94,722	0.74	1.00	0.44	9,531
crimewhole_cumu	22,154.28	9,569.00	39,746.76	1,745,432	15,269.48	9,806.00	24,315.66	1,016,553	13,058.37	9,260.00	17,514.01	94,722	29,232.21	13,756.00	41,253.67	9,531
review_utd	34.28	16.00	48.41	1,745,432	34.20	14.00	52.47	1,016,553	19.24	8.00	31.39	94,722	21.77	7.00	40.36	9,531
listing_size_zip	562.51	481.00	406.28	1,745,432	513.44	401.00	435.57	1,016,553	433.98	339.00	364.30	94,722	504.51	449.00	309.40	9,531
cross_listing	0.04	0.00	0.19	1,745,432	0.00	0.00	0.00	1,016,553	0.00	0.00	0.00	94,722	0.00	0.00	0.00	9,531
occ_dummy	0.87	1.00	0.34	1,745,432	0.83	1.00	0.38	1,016,553	0.77	1.00	0.42	94,722	0.87	1.00	0.34	9,531
lag_vicinity_sr_cumu_dummy	0.05	0.00	0.21	1,745,432	0.04	0.00	0.20	1,016,553	0.04	0.00	0.19	94,722	0.06	0.00	0.23	9,531
lag_listing_sr_cumu_dummy	0.06	0.00	0.23	1,745,432	0.04	0.00	0.20	1,016,553	0.03	0.00	0.16	94,722	0.03	0.00	0.18	9,531
lag_vicinity_listing_radius_pct	0.06	0.03	0.11	1,745,432	0.07	0.04	0.12	1,016,553	0.08	0.05	0.11	94,722	0.06	0.04	0.07	9,531
superhost_dummy	0.25	0.00	0.43	1,745,432	0.22	0.00	0.42	1,016,553	0.11	0.00	0.32	94,722	0.13	0.00	0.34	9,531
strict_cp	0.53	1.00	0.50	1,745,432	0.43	0.00	0.50	1,016,553	0.52	1.00	0.50	94,722	0.41	0.00	0.49	9,531
ave_review_word_count_cumu	55.03	51.60	30.24	1,745,432	52.94	49.48	30.99	1,016,553	42.84	38.70	31.52	94,722	37.27	33.14	29.07	9,531
lag_vicinity_sr_cumu	0.06	0.00	0.35	1,745,432	0.06	0.00	0.31	1,016,553	0.06	0.00	0.47	94,722	0.11	0.00	0.57	9,531
lag_listing_sr_cumu	0.07	0.00	0.30	1,745,432	0.05	0.00	0.24	1,016,553	0.03	0.00	0.17	94,722	0.04	0.00	0.20	9,531
safety_score	4.90	5.05	2.79	1,745,432	5.10	5.23	2.81	1,016,553	4.83	4.87	2.92	94,722	4.52	4.38	2.78	9,531

Table 5: Data summary statistics – listing characteristics at listing–month level for subgroups(cont.)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	log occupancy_rate	log occupancy_rate	log occupancy_rate	log occupancy_rate	log adr	log adr	log adr	log adr
lag_vicinity_sr_cumu_dummy	-0.0176***	-0.0172***	-0.0189***	-0.0175***	-0.0163***	-0.0138***	-0.0153***	-0.0136***
	(0.00257)	(0.00168)	(0.00215)	(0.00185)	(0.00389)	(0.00267)	(0.00330)	(0.00295)
lag_listing_sr_cumu_dummy	-0.0263***	-0.0247***	-0.0251***	-0.0265***	-0.0181***	-0.0123***	-0.0177***	-0.0114***
	(0.00196)	(0.00187)	(0.00178)	(0.00207)	(0.00284)	(0.00307)	(0.00269)	(0.00335)
lag_vicinity_listing_radius_pct	-0.0117***	-0.00449	-0.00780**	-0.00942***	-0.00261	-0.0126**	-0.00308	-0.0122**
	(0.00370)	(0.00346)	(0.00385)	(0.00335)	(0.00564)	(0.00535)	(0.00589)	(0.00516)
lag_log_crimewhole_cumu_norr	0.0512***	0.171***	0.0427***	0.170***	-0.0496***	-0.0561***	-0.0478***	-0.0625**
	(0.0111)	(0.0137)	(0.00950)	(0.0168)	(0.0179)	(0.0213)	(0.0150)	(0.0265)
lag_log_review_utd	0.00425***	0.00434***	0.00346***	0.00534***	0.0137***	0.00967***	0.0135***	0.00909***
	(0.000587)	(0.000588)	(0.000544)	(0.000642)	(0.000924)	(0.000996)	(0.000870)	(0.00108)
log_listing_size_zip	-0.0273***	-0.0141***	-0.0212***	-0.0142***	0.0242***	0.00709*	0.0229***	0.00697
	(0.00286)	(0.00266)	(0.00281)	(0.00269)	(0.00421)	(0.00423)	(0.00409)	(0.00437)
log_rating_overall	0.0246***	0.0272***	0.0239***	0.0283***	-0.00297	-0.00151	-0.00500**	0.00196
	(0.00180)	(0.00181)	(0.00158)	(0.00217)	(0.00270)	(0.00298)	(0.00254)	(0.00325)
superhost_dummy	0.0175***	0.0171***	0.0176***	0.0172***	0.00595***	0.0106***	0.00673***	0.0102***
	(0.000817)	(0.000840)	(0.000773)	(0.000896)	(0.00115)	(0.00124)	(0.00112)	(0.00128)
cross_listing	0.0356***	0.0243***	0.0307***	0.0310***	-0.00764	-0.00254	-0.00801*	-0.000418
	(0.00369)	(0.00420)	(0.00343)	(0.00476)	(0.00468)	(0.00651)	(0.00447)	(0.00736)
strict_cp	0.000885	0.000465	0.00132	-0.000385	0.0130***	0.0116***	0.0129***	0.0115***
	(0.00114)	(0.00113)	(0.00107)	(0.00121)	(0.00169)	(0.00188)	(0.00161)	(0.00203)
Constant	0.456***	0.311***	0.428***	0.328***	4.857***	4.645***	4.863***	4.580***
	(0.0179)	(0.0176)	(0.0179)	(0.0175)	(0.0262)	(0.0281)	(0.0260)	(0.0284)
Observations	1,484,474	1,381,764	1,716,774	1,149,464	1,484,474	1,381,764	1,716,774	1,149,464
R-squared	0.552	0.569	0.551	0.573	0.921	0.924	0.919	0.925
Time*City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PropertyID FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CSEs	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID
Sample	H	L	W	M	H	L	W	M

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Listing level: Regression with all Airbnb listings (4 areas)

VARIABLES	(1) log occupancy_rate	(2) log occupancy_rate	(3) log occupancy_rate	(4) log occupancy_rate	(5) log adr	(6) log adr	(7) log adr	(8) log adr
lag_vicinity_sr_cumu_dummy	-0.0161***	-0.0210***	-0.0301***	-0.0372**	-0.0131***	-0.0171***	-0.0235*	-0.0164
	(0.00171)	(0.00249)	(0.00915)	(0.0181)	(0.00275)	(0.00376)	(0.0130)	(0.0408)
lag_listing_sr_cumu_dummy	-0.0236***	-0.0308***	-0.0289***	-0.0621**	-0.0201***	-0.00809**	-0.0102	0.0269
	(0.00158)	(0.00265)	(0.0108)	(0.0285)	(0.00248)	(0.00406)	(0.0168)	(0.0410)
lag_vicinity_listing_radius_pct	-0.0107***	-0.00419	-0.0282**	0.0154	-0.00485	-0.0115**	-0.0154	0.0310
	(0.00338)	(0.00396)	(0.0130)	(0.143)	(0.00540)	(0.00571)	(0.0277)	(0.205)
lag_log_crimewhole_cumu_norr	0.0439***	0.120***	0.169***	-0.411**	-0.0694***	-0.00732	-0.302**	0.969***
	(0.00996)	(0.0153)	(0.0617)	(0.167)	(0.0156)	(0.0233)	(0.136)	(0.320)
lag_log_review_utd	0.00480***	0.00551***	-0.0201***	0.0181**	0.0145***	0.00754***	-0.00161	0.0336**
	(0.000518)	(0.000725)	(0.00255)	(0.00794)	(0.000851)	(0.00114)	(0.00486)	(0.0139)
log_listing_size_zip	-0.0157***	-0.0276***	-0.0179*	0.0701**	0.0190***	0.0217***	-0.0251	-0.0591
	(0.00252)	(0.00290)	(0.00934)	(0.0324)	(0.00383)	(0.00465)	(0.0185)	(0.0576)
log_rating_overall	0.0251***	0.0262***	0.0328***	-0.0176	-3.99e-05	-0.00583*	-0.0117	-0.0425
	(0.00166)	(0.00211)	(0.00692)	(0.0521)	(0.00250)	(0.00331)	(0.0151)	(0.0355)
superhost_dummy	0.0145***	0.0209***	0.0367***	0.0396**	0.0101***	0.00444***	0.0114	-0.0218
	(0.000718)	(0.00102)	(0.00491)	(0.0154)	(0.00106)	(0.00138)	(0.00913)	(0.0262)
cross_listing	0.0327***				-0.00909**			
	(0.00277)				(0.00386)			
strict_cp	-0.00111	0.00179	0.0146***	-0.0182	0.0146***	0.00912***	0.0132	-0.0366*
	(0.000985)	(0.00143)	(0.00566)	(0.0122)	(0.00154)	(0.00225)	(0.00981)	(0.0216)
Constant	0.394***	0.432***	0.301***	0.209	5.068***	4.250***	4.000***	4.499***
	(0.0160)	(0.0185)	(0.0601)	(0.259)	(0.0239)	(0.0305)	(0.124)	(0.420)
Observations	1,745,432	1,016,553	94,722	9,531	1,745,432	1,016,553	94,722	9,531
R-squared	0.540	0.581	0.593	0.621	0.887	0.858	0.894	0.918
Time*City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PropertyID FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CSEs	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID	PropertyID
Sample	EH	PR	SR	HR	EH	PR	SR	HR

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Listing Level: Regression with the all Airbnb listings (4 listing types)

	Panel A: VS users				Panel B: Normal users			
VARIABLES	mean	p50	sd	N	mean	p50	sd	N
reservation_pre	2.76	2.00	1.51	2,252	2.72	2.00	1.43	4,504
log_ave_crime_cumu_norm_pre	0.93	0.28	1.95	2,252	0.81	0.27	1.44	4,504
ave_vsr_cumu_pre	0.63	0.50	0.44	2,252	0.64	0.50	0.43	4,504
ave_vs_listing_zip_pct_pre	0.06	0.05	0.04	2,252	0.06	0.05	0.04	4,504
ave_vs_listing_radius_pct_pre	0.09	0.07	0.07	2,252	0.08	0.07	0.06	4,504
log_ave_r_wct_cumu_pre	4.37	4.39	0.64	2,252	4.36	4.39	0.63	4,504
propensity_score	0.74	0.72	0.15	2,252	0.73	0.71	0.15	4,504

Table 8: Data summary statistics – users characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	reservation_monthly	vicinity_sr_cumu	vicinity_sr_cumu_dummy	crimewhole_cumu_norm	vicinity_listing_zip_pct	vicinity_listing_radius_pct
treat	0.430***	-0.00402	-0.000754	0.0991***	0.00143***	-6.14e-05
	(0.0309)	(0.00371)	(0.0319)	(0.0185)	(0.000228)	(0.000985)
post	-1.396***	-0.578***	-1.221***	1.078***	0.0271***	0.0290***
	(0.0462)	(0.0657)	(0.0584)	(0.0850)	(0.00187)	(0.00299)
interaction	-0.918***	-0.697***	-0.490***	-0.927***	-0.0250***	-0.0247***
	(0.0601)	(0.135)	(0.113)	(0.112)	(0.00267)	(0.00505)
Constant				0.712***	0.0576***	0.0830***
				(0.0145)	(0.000298)	(0.000592)
Observations	254,056	22,265	22,237	22,415	22,415	22,415
Number of pair_index	2,252	2,230	2,225			
PairID FE	Yes	Yes	Yes	Yes	Yes	Yes
CSEs	PairID	PairID	PairID	PairID	PairID	PairID
Sample	Monthly Reservation	Reserved Property	Reserved Property	Reserved Property	Reserved Property	Reserved Property
Log pseudolikelihood:	-80903	-15474	-9995			
R-squared				0.364	0.335	0.248

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9: User level: DID for all users

	(1)	(2)	(3)	(4)
VARIABLES	h_zip	h_zip	w_zip	w_zip
treat	1.340***	-0.0857**	1.205***	-0.656***
	(0.0693)	(0.0423)	(0.0587)	(0.0453)
post	-0.269***	0.264***	-0.0248	0.0866
	(0.0904)	(0.0592)	(0.0749)	(0.0628)
interaction	-0.351**	0.316***	-0.628***	0.682***
	(0.160)	(0.0990)	(0.135)	(0.105)
Observations	6,205	14,830	8,880	12,815
Number of pair_index	604	1,446	878	1,266
PairID FE	Yes	Yes	Yes	Yes
CSEs	PairID	PairID	PairID	PairID
Sample	1st_vsr_h_zip	1st_vsr_l_zip	1st_vsr_w_zip	1st_vsr_m_zip
Log pseudolikelihood:	-2793	-6782	-3999	-6120
R-squared				

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10: User level: DID for 4 groups of users based on their first safety review booking's zip code

Area	# of Listings	Avg. Occupancy (Days)	Avg. days per Reservation	# of Reservations	Avg. ADR ($)	Avg. % of VSLs in 0.3-mi-r	EST. GBV changes if no VSRs are available (%)	EST. GBV changes if Safety-Score is available (%)	EST. GBV changes in High Alert (%)
Panel A: Vicinity Safety Listings									
H	36,312	18.48	3.27	204,954	161.61	4.73%	6.32%	-0.14%	-4.08%
L	90,556	18.59	3.20	525,909	123.14	12.63%	6.20%	-0.24%	-4.49%
W	56,041	18.58	3.15	330,484	160.28	9.09%	6.70%	-0.11%	-3.81%
M	70,827	18.54	3.28	400,379	113.48	11.39%	6.29%	-0.32%	-4.84%
Panel B: Normal Listings									
H	1,448,162	13.74	3.82	5,211,927	191.10	5.07%	0.17%	-0.09%	-3.75%
L	1,291,208	14.19	3.76	4,874,547	138.06	8.45%	0.21%	-0.17%	-4.34%
W	1,660,733	13.83	3.81	6,036,619	189.28	5.14%	0.13%	-0.06%	-3.47%
M	1,078,637	14.14	3.76	4,049,855	130.41	9.00%	0.35%	-0.24%	-4.87%
Panel C: All Listings									
All	2,866,238	14.16	3.75	10,817,337	164.69	6.83%	0.52%	-0.13%	-4.06%

Table 11: GBV of the platform using back of envelope calculation

VARIABLES	(1) log_adr	(2) utility	(3) utility	(4) utility
lag_log_eh_radius_ave_review_utd	-0.00558***			
	(0.00171)			
log_eh_radius_ave_rating_overall	0.00307			
	(0.0112)			
eh_radius_ave_superhost_dummy	0.00479			
	(0.00370)			
eh_radius_ave_cross_listing	-0.00115			
	(0.00953)			
eh_radius_ave_strict_cp	0.00306			
	(0.00334)			
log_adr			-1.100***	
			(0.00903)	
log_adr_iv				-6.735***
				(1.609)
lag_vicinity_sr_cumu_dummy	-0.00985***	-0.0805***	-0.0914***	-0.147***
	(0.00263)	(0.0123)	(0.0121)	(0.0199)
lag_listing_sr_cumu_dummy	-0.0184***	-0.0806***	-0.101***	-0.204***
	(0.00247)	(0.0108)	(0.0107)	(0.0315)
lag_vicinity_ehlisting_radius_pc	-0.0129	-0.107*	-0.129**	-0.240***
	(0.00944)	(0.0550)	(0.0549)	(0.0634)
lag_log_crimewhole_cumu_norm	-0.0284	0.278***	0.249***	0.102
	(0.0215)	(0.0935)	(0.0932)	(0.105)
lag_log_review_utd	0.0197***	0.000443	0.0220***	0.132***
	(0.000874)	(0.00351)	(0.00347)	(0.0317)
log_rating_overall	0.0123*	0.291***	0.304***	0.373***
	(0.00676)	(0.0278)	(0.0279)	(0.0339)
superhost_dummy	0.0166***	0.0467***	0.0651***	0.159***
	(0.00105)	(0.00431)	(0.00424)	(0.0272)
cross_listing	-0.00634*	0.0640***	0.0570***	0.0211
	(0.00381)	(0.0147)	(0.0144)	(0.0180)
strict_cp	0.00443***	-0.0455***	-0.0405***	-0.0151*
	(0.00136)	(0.00542)	(0.00534)	(0.00905)
Constant	5.127***	-3.971***	1.655***	30.48***
	(0.0337)	(0.0858)	(0.0978)	(8.234)
Observations	1,014,301	1,014,301	1,014,301	1,014,301
R-squared	0.913	0.789	0.800	0.789
Time*City FE	Yes	Yes	Yes	Yes
PropertyID FE	Yes	Yes	Yes	Yes
CSEs	PropertyID	PropertyID	PropertyID	PropertyID
Sample	EH	EH	EH	EH
F statistic:	288.5			

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 12: Utility estimation

	Airbnb VS Listing	Airbnb Normal Listing	VRBO Listing	All Listing
CS (Zero-DSC w/o P change) - CS (Status quo)	77,374.58	-62,822.50	-17,936.54	-3,384.46
%	12.249%	-0.699%	-1.471%	-3.120%
CS (Zero-DSC w/ P change) - CS (Status quo)	40,951.36	-29,707.59	-14,053.40	-2,809.64
%	6.483%	-0.330%	-1.152%	-2.590%
CS (Safety-Score w/o P change) - CS (Status quo)	-4,283.25	-9,329.88	14,731.78	1,118.66
%	-0.678%	-0.104%	1.208%	1.030%
CS (Safety-Score w/ P change) - CS (Status quo)	30,388.23	-40,648.58	10,865.19	604.84
%	4.811%	-0.452%	0.891%	0.560%
CS (High-Alert w/o P change) - CS (Status quo)	-532,013.67	487,538.75	64,603.71	20,128.79
%	-84.223%	5.422%	5.298%	18.560%
CS (High-Alert w/ P change) - CS (Status quo)	-525,669.86	482,025.29	63,678.51	20,033.95
%	-83.219%	5.361%	5.222%	18.480%
Market share (Status quo)	5.984%	83.069%	10.947%	100.000%
Market share (Zero-DSC w/o P change)	6.723%	82.491%	10.786%	100.000%
Market share (Zero-DSC w/ P change)	6.383%	82.796%	10.821%	100.000%
Market share (Safety-Score w/o P change)	5.943%	82.978%	11.079%	100.000%
Market share (Safety-Score w/ P change)	6.266%	82.690%	11.045%	100.000%
Market share (High-Alert w/o P change)	0.935%	87.539%	11.527%	100.000%
Market share (High-Alert w/ P change)	0.993%	87.488%	11.519%	100.000%

Table 13: Consumer Surplus estimation

A Mathematics Proof

A.1 Proof on how to do reduced-form back of envelope calculation for the "high-alert" counterfactual

Denote:

X^H_{vsl} = # bookings of vicinity safety listings in high-income neighborhoods, under the status quo

X^L_{vsl} = # bookings of vicinity safety listings in low-income neighborhoods, under the status quo

X^H_{nl} = # bookings of normal listings in high-income neighborhoods, under the status quo

X^L_{nl} = # bookings of normal listings in low-income neighborhoods, under the status quo

Y^H_{vsl} = # bookings of vicinity safety listings in high-income neighborhoods, under "high-alert"

Y^L_{vsl} = # bookings of vicinity safety listings in low-income neighborhoods, under "high-alert"

Y^H_{nl} = # bookings of normal listings in high-income neighborhoods, under "high-alert"

Y^L_{nl} = # bookings of normal listings in low-income neighborhoods, under "high-alert"

We observe $\{X^H_{vsl}, X^L_{vsl}, X^H_{nl}, X^L_{nl}\}$ and reduced-form impacts of a user submitting a safety review out of self-experience. Our goal is to solve for $\{Y^H_{vsl}, Y^L_{vsl}, Y^H_{nl}, Y^L_{nl}\}$.

Self-experience of vicinity safety makes a guest β_1 (percent) more likely to make any bookings on Airbnb according to the Column 1 of Table 9. The coefficient on $safetyuser \times post = -0.918$ in a Poisson regression, implying that safety review reduces monthly reservations by $\beta_1 = exp(-0.918) - 1 = -0.6007$. In our notation, we have function 4:

$$\begin{aligned} & (Y^H_{vsl} + Y^L_{vsl} + Y^H_{nl} + Y^L_{nl}) \\ = \; & \underbrace{(X^H_{vsl} + X^L_{vsl}) \cdot (1 + \beta_1) + X^H_{nl} + X^L_{nl}}_{\text{item A}} \end{aligned} \tag{4}$$

Self-experience of vicinity safety makes a guest β_2 (percent) more likely to make a booking in an H neighborhood, conditional on she makes any booking on Airbnb. Note that the safety experience, if it occurs, has a 30% chance to occur in an H neighborhood and a 70% chance to occur in an L neighborhood.

According to Table 10 column 1, the coefficient of $safetyuser \times post = -0.351$ for the probit of booking in H if VS in H, which implies that having a safety review in H will change the probability of the booking in H by $\beta_{2|VS\ in\ H} = -0.0722$. According to Table 10 Column 2, the coefficient of $safetyuser \times post = +0.316$ for the probit of booking in H if VS in L, which implies that having a safety review in L will change the probability of the booking in L by $\beta_{2|VS\ in\ L} = +0.0781$. And denote the probability of vicinity safety listings in H(L) area as $Pr_{V|H} = 0.0245$ ($Pr_{V|L} = 0.0655$) In our notation, this means:

$$\begin{aligned}
& prob(booking\ in\ H|anybooking) \\
= & \tfrac{Y^H_{vsl}+Y^H_{nl}}{Y^H_{vsl}+Y^L_{vsl}+Y^H_{nl}+Y^L_{nl}} \\
= & \underbrace{\frac{X^H_{vsl}+X^H_{nl}}{X^H_{vsl}+X^L_{vsl}+X^H_{nl}+X^L_{nl}}}_{\text{item B}} \cdot (0.3 \cdot (1 + Pr_{V|H} \cdot \beta_{2|VS\ in\ H}) + 0.7 \cdot (1 + Pr_{V|L} \cdot \beta_{2|VS\ in\ L}))
\end{aligned} \tag{5}$$

Self-experience of vicinity safety makes a guest β_2 (percent) more likely to make a booking in a VSL, conditional on she makes any booking on Airbnb. According to Table 9 column 6, the coefficient of $safetyuser \times post = -0.490$ for the probit of booking in any VSL, which implies that having a safety review will change the probability of the booking in any VSL by $\beta_3 = -0.0747$ and the probability of vicinity safety listings is $Pr_V = 0.0443$.

In our notation, this means:

$$\begin{aligned}
& prob(booking\ in\ VSL|anybooking) \\
= & \tfrac{Y^H_{vsl}+Y^L_{vsl}}{Y^H_{vsl}+Y^L_{vsl}+Y^H_{nl}+Y^L_{nl}} \\
= & \underbrace{\frac{X^H_{vsl}+X^L_{vsl}}{X^H_{vsl}+X^L_{vsl}+X^H_{nl}+X^L_{nl}}}_{\text{item C}} \cdot (1 + Pr_V \cdot \beta_3)
\end{aligned} \tag{6}$$

So far, we have four unknowns and three equations, so we need an extra equation, which implies that the fraction of VSL in each neighborhood does not change. Denote this fraction as α^H_{vsl}, this amounts to:

$$\frac{Y^H_{vsl}}{Y^H_{vsl}+Y^H_{nl}} = \frac{X^H_{vsl}}{X^H_{vsl}+X^H_{nl}} \cdot (1+Pr_V \cdot \beta_3) = \alpha^H_{vsl} \cdot (1+Pr_V \cdot \beta_3) \tag{7}$$

Combining Equations 4, 5, 6 and 7 and solving for $\{Y^H_{vsl}, Y^L_{vsl}, Y^H_{nl}, Y^L_{nl}\}$, we have:

$$\frac{Y^H_{vsl}+Y^H_{nl}}{Y^H_{vsl}+Y^L_{vsl}+Y^H_{nl}+Y^L_{nl}} = \frac{Y^H_{vsl}+Y^H_{nl}}{A} = B \cdot (0.3 \cdot (1+Pr_{V|H} \cdot \beta_{2|VS\ in\ H}) + 0.7 \cdot (1+Pr_{V|L} \cdot \beta_{2|VS\ in\ L})) \tag{8}$$

$$Y^H_{vsl}+Y^H_{nl} = A \cdot B \cdot (0.3 \cdot (1+Pr_{V|H} \cdot \beta_{2|VS\ in\ H}) + 0.7 \cdot (1+Pr_{V|L} \cdot \beta_{2|VS\ in\ L})) \tag{9}$$

$$\frac{Y^H_{vsl}+Y^L_{vsl}}{Y^H_{vsl}+Y^L_{vsl}+Y^H_{nl}+Y^L_{nl}} = \frac{Y^H_{vsl}+Y^L_{vsl}}{A} = C \cdot (1+Pr_V \cdot \beta_3) \tag{10}$$

$$Y^H_{vsl}+Y^L_{vsl} = A \cdot C \cdot (1+Pr_V \cdot \beta_3) \tag{11}$$

$$\frac{Y^H_{vsl}}{Y^H_{vsl}+Y^H_{nl}} = \frac{Y^H_{vsl}}{A \cdot B \cdot (0.3 \cdot (1+Pr_{V|H} \cdot \beta_{2|VS\ in\ H}) + 0.7 \cdot (1+Pr_{V|L} \cdot \beta_{2|VS\ in\ L}))} = \alpha^H_{vsl} \cdot (1+Pr_V \cdot \beta_3) \tag{12}$$

So the solutions are:

$$\begin{aligned}
Y^H_{vsl} &= \alpha^H_{vsl} \cdot (1+Pr_V \cdot \beta_3) \cdot A \cdot B \cdot (0.3 \cdot (1+Pr_{V|H} \cdot \beta_{2|VS\ in\ H}) + 0.7 \cdot (1+Pr_{V|L} \cdot \beta_{2|VS\ in\ L})) \\
Y^H_{nl} &= A \cdot B \cdot (0.3 \cdot (1+Pr_{V|H} \cdot \beta_{2|VS\ in\ H}) + 0.7 \cdot (1+Pr_{V|L} \cdot \beta_{2|VS\ in\ L})) - Y^H_{vsl} \\
Y^L_{vsl} &= A \cdot C \cdot (1+Pr_V \cdot \beta_3) - Y^H_{vsl} \\
Y^L_{nl} &= A - Y^H_{vsl} - Y^H_{nl} - Y^L_{vsl}
\end{aligned} \tag{13}$$

Similar calculations are performed for W/M areas.

A.2 Extend the DID results of VS users to the consumer surplus counterfactual based on utility

The coefficient on $safetyreview \times post = -0.918$ in a Poisson regression according to the Table 9 Column 1, which is a 60.07% decrease. Given the average number of reservations per month for a single VS user in our sample is 0.1092 and review rate is 44.56%, the reservation that a VS user book in Airbnb is 0.066/0.4456 less than a normal user after she has reported a VS issue in her first VSR. A VS user is less likely to book Airbnb reservations than a normal user after she has reported a VS issue in her first VSR is described by:

$$\begin{aligned} &[\#Airbnbbooking_{VS\ user,aft} - \#Airbnbbooking_{VS\ user,bef}] \\ &-[\#Airbnbbooking_{NM\ user,aft} - \#Airbnbbooking_{NM\ user,bef}] = -0.147 \end{aligned} \tag{14}$$

Assuming VS and normal users have the same tendency to book short-term rental (i.e. # of total short-term rentals are the same), the above equation can be rewritten as user i's market share for all Airbnb choices $\sum_{j \in Airbnb} s_{ij}$:

$$\left(\frac{\partial \sum_{j \in Airbnb} s_{ij}}{\partial 1_{VSR}}\right)_{i=VS\ user} - \left(\frac{\partial \sum_{j \in Airbnb} s_{ij}}{\partial 1_{VSR}}\right)_{i=NM\ user} = -0.147 \tag{15}$$

Assume utility function is:

$$U_{ij} = \beta X_j + \gamma_{NM} + \Delta\gamma \cdot 1_{VSR,j} + \varepsilon_{ij} \tag{16}$$

Where γ_{NM} indicates normal users' sensitivity to observing any VSR in a listing, $\gamma_{NM} + \Delta\gamma$ indicates VS users' sensitivity to VSR. The market share of all Airbnb reservations is:

$$\sum_{j \in Airbnb} s_{ij} = 1 - s_{i,VRBO} = 1 - \frac{1}{1 + \sum_{j \in Airbnb} exp(U_{ij})} \tag{17}$$

Then:

$$\left(\frac{\partial \sum_{j\in Airbnb} s_{ij}}{\partial 1_{VSR}}\right)_{i=NM\ user} = +\gamma_{NM} \cdot s_{NM\ user,VRBO} \cdot \sum_{j\in Airbnb\ \&\ 1_{VSR}} s_{NM\ user,j} \tag{18}$$

$$\left(\frac{\partial \sum_{j\in Airbnb} s_{ij}}{\partial 1_{VSR}}\right)_{i=VS\ user} = +(\gamma_{NM}+\Delta\gamma) \cdot s_{VS\ user,VRBO} \cdot \sum_{j\in Airbnb\ \&\ 1_{VSR}} s_{VS\ user,j} \tag{19}$$

Denote a user's total probability of choosing any Airbnb listing with $VSR > 0$ as:

$$s_{NM\ user,Airbnb\ \&\ 1_{VSR}} = \sum_{j\in Airbnb\ \&\ 1_{VSR}} s_{NM\ user,j} \tag{20}$$

$$s_{VS\ user,Airbnb\ \&\ 1_{VSR}} = \sum_{j\in Airbnb\ \&\ 1_{VSR}} s_{VS\ user,j} \tag{21}$$

The DID results can be written as:

$$\begin{aligned} &+(\gamma_{NM}+\Delta\gamma) \cdot s_{VS\ user,VRBO} \cdot s_{VS\ user,Airbnb\ \&\ 1_{VSR}} \\ &-\gamma_{NM} \cdot s_{NM\ user,VRBO} \cdot s_{NM\ user,Airbnb\ \&\ 1_{VSR}} = -0.147 \end{aligned} \tag{22}$$

Note that we observe normal users' market shares in the data because almost all users are normal users, but we do not observe VS users' market shares because we cannot track VS users in all Airbnb and $VRBO$ bookings. However, the utility framework spells out how these two types of users differ. More specifically, the model implies:

$$\begin{aligned} \frac{s_{NM\ user,VRBO}}{s_{VS\ user,VRBO}} &= \frac{(1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR}))^{-1}}{(1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR}+\Delta\gamma\cdot 1_{VSR}))^{-1}} \\ &= \frac{1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR}+\Delta\gamma\cdot 1_{VSR})}{1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR})} \\ &= \frac{1+\sum_{j\in Airbnb} exp(\beta X)+exp(\Delta\gamma)\cdot\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM})}{1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR})} \\ &= s_{NM\ user,VRBO} + s_{NM\ user,Airbnb\ \&\ VSR=0} + exp(\Delta\gamma)\cdot s_{NM\ user,Airbnb\ \&\ 1_{VSR}} \end{aligned} \tag{23}$$

This implies:

$$s_{VS\ user,VRBO} = \frac{s_{NM\ user,VRBO}}{s_{NM\ user,VRBO} + s_{NM\ user,Airbnb\ \&\ VSR=0} + exp(\Delta\gamma) \cdot s_{NM\ user,Airbnb\ \&\ 1_{VSR}}} \tag{24}$$

Similarly:

$$\begin{aligned}
\frac{s_{NM\ user,Airbnb\ \&\ 1_{VSR}}}{s_{VS\ user,Airbnb\ \&\ 1_{VSR}}} &= \frac{\frac{\sum_{j\in Airbnb\ \&\ 1_{VSR}} exp(\beta X+\gamma_{NM})}{1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR})}}{\frac{\sum_{j\in Airbnb\ \&\ 1_{VSR}} exp(\beta X+\gamma_{NM}+\Delta\gamma)}{1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR}+\Delta\gamma\cdot 1_{VSR})}} \\
&= \frac{\frac{\sum_{j\in Airbnb\ \&\ 1_{VSR}} exp(\beta X+\gamma_{NM})}{1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR})}}{\frac{exp(\Delta\gamma)\sum_{j\in Airbnb\ \&\ 1_{VSR}} exp(\beta X+\gamma_{NM})}{1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR}+\Delta\gamma\cdot 1_{VSR})}} \\
&= exp(\Delta\gamma)\frac{1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR}+\Delta\gamma\cdot 1_{VSR})}{1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR})} \\
&= exp(\Delta\gamma)\frac{1+\sum_{j\in Airbnb\ \&\ VSR=0} exp(\beta X)+exp(\Delta\gamma)\sum_{j\in Airbnb\ \&\ 1_{VSR}} exp(\beta X+\gamma_{NM})}{1+\sum_{j\in Airbnb} exp(\beta X+\gamma_{NM}\cdot 1_{VSR})} \\
&= exp(\Delta\gamma)\cdot\big(s_{NM\ user,VRBO} + s_{NM\ user,Airbnb\ \&\ VSR=0} + exp(\Delta\gamma) \\
&\quad \cdot s_{NM\ user,Airbnb\ \&\ 1_{VSR}}\big)
\end{aligned} \tag{25}$$

This implies:

$$\begin{aligned}
s_{VS\ user,Airbnb\ \&\ 1_{VSR}} &= \frac{1}{exp(\Delta\gamma)} \\
&\cdot\frac{s_{NM\ user,Airbnb\ \&\ 1_{VSR}}}{s_{NM\ user,VRBO}+s_{NM\ user,Airbnb\ \&\ VSR=0}+exp(\Delta\gamma)\cdot s_{NM\ user,Airbnb\ \&\ 1_{VSR}}}
\end{aligned} \tag{26}$$

Plug these into the DID results:

$$\begin{aligned}
&(\gamma_{NM}+\Delta\gamma)\cdot s_{VS\ user,VRBO}\cdot s_{VS\ user,Airbnb\ \&\ 1_{VSR}} \\
&-\gamma_{NM}\cdot s_{VS\ user,VRBO}\cdot s_{VS\ user,Airbnb\ \&\ 1_{VSR}} = -0.147
\end{aligned} \tag{27}$$

$$\begin{aligned}
&\frac{\gamma_{NM}+\Delta\gamma}{exp(\Delta\gamma)}\cdot\frac{s_{NM\ user,VRBO}\cdot s_{NM\ user,Airbnb\ \&\ 1_{VSR}}}{(s_{NM\ user,VRBO}+s_{NM\ user,Airbnb\ \&\ 1_{VSR=0}}+exp(\Delta\gamma)\cdot s_{NM\ user,Airbnb\ \&\ 1_{VSR}})^2} \\
&= -0.147+\gamma_{NM}\cdot s_{NM\ user,VRBO}\cdot s_{NM\ user,Airbnb\ \&\ 1_{VSR}}
\end{aligned} \tag{28}$$

Because almost all users are normal users, the data gives us $s_{NM\ user,VRBO}$ (market share of $VRBO$), $s_{NM\ user,Airbnb\ \&\ VSR=0}$ (total market share of all normal Airbnb listings), and $s_{NM\ user,Airbnb\ \&\ 1_{VSR}}$ (total market share of all Airbnb VS listings). We also know γ_{NM} from the utility regression. Thus, the only unknown in the above equation is $\Delta\gamma$. We can solve it easily and get $\Delta\gamma = -2.17$.