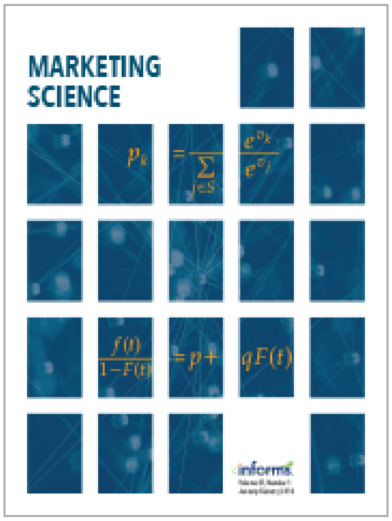


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Safety Reviews on Airbnb: An Information Tale

Aron Culotta, Ginger Zhe Jin, Yidan Sun, Liad Wagman

informs
https://pubsonline.informs.org/journal/mksc

MARKETING SCIENCE
Articles in Advance, pp. 1–38
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Safety Reviews on Airbnb: An Information Tale

Aron Culotta,[a] Ginger Zhe Jin,[b,c,*] Yidan Sun,[d] Liad Wagman[e]

[a] Computer Science, Tulane University, New Orleans, Louisiana 70118; [b] Department of Economics, University of Maryland, College Park, Maryland 20742; [c] National Bureau of Economic Research, Cambridge, Massachusetts 02138; [d] Leadership and Organization Science, School of Management, Binghamton University, Binghamton, New York 13902; [e] Lally School of Management, Rensselaer Polytechnic Institute, Troy, New York 12180

*Corresponding author

Contact: aculotta@tulane.edu, https://orcid.org/0000-0003-2660-7575 (AC); ginger@umd.edu, https://orcid.org/0000-0001-7912-3780 (GZJ); ysun20@binghamton.edu, https://orcid.org/0009-0009-2220-7655 (YS); wagman@rpi.edu, https://orcid.org/0000-0002-1570-4858 (LW)

Received: November 14, 2023
Revised: November 2, 2024; March 15, 2025; May 20, 2025
Accepted: May 26, 2025
Published Online in Articles in Advance: September 3, 2025

https://doi.org/10.1287/mksc.2023.0552

Abstract. Consumer reviews, especially those expressing concerns of product quality, are crucial for the credibility of online platforms. However, reviews that criticize a product or service may also dissuade buyers from using the platform, creating a potential incentive to blur the visibility of critical reviews. Using Airbnb and official crime data in five major U.S. cities, we find that both reviews and personal experiences concerning the safety of a listing's vicinity decrease guest bookings on the platform. Counterfactual simulations suggest that a complete removal of vicinity safety reviews (VSRs) could hurt guests if they do not adjust their beliefs accordingly, and such removal can increase revenues from reservations on Airbnb, with positive sorting toward listings formerly with VSRs. Conversely, highlighting VSRs would generate opposite effects. However, the incentive to suppress VSRs can be mitigated if guests have a rational expectation of average vicinity risk after all VSRs are removed or if guests can learn from their own vicinity safety experience for a long-enough time. Because VSRs are more closely correlated with official crime statistics in low-income and minority neighborhoods, our findings suggest that suppressing or highlighting VSRs would have different effects on different neighborhoods.

History: This paper has been accepted for the *Marketing Science* Special Issue on Digital Platforms in Marketing Science. Tat Chan served as the senior editor.

Funding: Dr. Culotta is supported in part by the Harold L. and Heather E. Jurist Center of Excellence for Artificial Intelligence at Tulane University, the Tulane Center for Community-Engaged Artificial Intelligence, and NSF awards [IIS-III-2107505, IIS-HCI-2333537, and SCC-IRG-2427237].

Supplemental Material: The online appendix and data files are available at https://doi.org/10.1287/mksc.2023.0552.

Keywords: Airbnb • vicinity safety • online review • information design • digital platform

1. Introduction

Addressing negative information about product quality is a classic problem facing business managers. For example, tobacco manufacturers were reluctant to reveal the health risks associated with cigarettes, pharmaceutical manufacturers may hesitate to acknowledge side effects found in clinical trials, and Sport Utility Vehicle (SUV) producers did not publish detailed data on SUV rollover risks until the government threatened regulation (Jüni et al. 2004, Fung et al. 2007). Behind these examples is the concern that negative news about product quality may reduce demand for the focal product or category, and this market-reducing effect may dominate any market-stealing effects that one may obtain by being less negative than competitors.

Digital platforms are better positioned to address this thorny problem because they are open to sellers of *all* types of product quality that meet their standard. Because platforms can earn commission from any sales on the platform and consumers are willing to pay more for better quality, platforms have incentives to help consumers discern high-quality products from low-quality ones. This explains why nearly all digital platforms gather consumer feedback in a standardized format, make it available globally, and aggregate it in a way that is salient and easy to digest and search if they so choose (see reviews by Einav et al. 2016, Tadelis 2016, and Luca 2017). This, in turn, can attract high-quality sellers to join the platform and encourages on-platform sellers to maintain high quality, forming a virtuous circle.

However, is there a limit to this market-driven solution? Is it possible that digital platforms do not always have the incentive to fully reveal and highlight critical feedback of product quality?[1] For example, suppose all consumers expect some minimum quality from every product listed on the platform, but some unlucky consumers have experienced below-minimum quality from a small number of listings. In this scenario, the platform may choose from a spectrum of information policies. At one extreme, it may disallow any critical feedback about the substandard experience in its online review system (while finding nonpublic ways to compensate the unlucky consumers or punish the substandard sellers); if the negative experience is rare enough, other buyers may not find out by themselves for a long time. At the other extreme, the platform may encourage and broadcast the critical feedback, and it may alert every future consumer of the substandard risk. Between the two extremes, the platform may allow critical feedback but make it hard to find or filter the content of the feedback before posting.

From a platform's perspective, the key economic trade-off is how surprising the negative experience is and how quickly that experience—if it is reflected in an authentic review—can find its way to influence the platform's future business. Intuitively, the bigger the negative surprise is and the more that future readers of the review may extend that negative surprise to other listings on the platform, the more harmful the review could be for the platform. For example, a buyer who gets burned by paying thousands of dollars for a counterfeit product may infer that all sellers that share a certain attribute with the cheating seller also sell counterfeits. If this buyer—and everyone else equally alerted by her experience—choose to switch away from the platform rather than switch toward other on-platform sellers that do not share this problematic attribute, the platform could lose significant business in the future.

Conversely, online review systems often suffer from information frictions. The probability of experiencing a negative event may be small for any individual buyer. The degree of the shock may depend on the subjective opinion of the buyer. Some burned buyers may be reluctant to leave a negative review (even if they choose to exit the platform), some negative reviews may not be read by all future buyers, and some readers may have difficulty deciphering the real content of a review as they believe some reviews are fake or misleading but cannot tell which is which (Gandhi et al. 2025). When these frictions add up to mute the negative surprise from most future buyers, a profit-maximizing platform may prefer to keep these frictions or even add more obfuscation into the system as long as it can still maintain sufficient credibility with future buyers.

In short, whether to encourage or discourage critical feedback on a digital platform depends on how much negative information spillover the feedback may generate for the platform—both concurrently and in the future—after taking into account the information frictions in its online review system and the potential of consumers learning from both their own experiences and other channels beyond online reviews.

In this paper, we use safety reviews on Airbnb as an example to understand why and when critical feedback about product quality can create the aforementioned trade-off for the platform. In particular, we use all Airbnb listings in five major U.S. cities (Atlanta, Chicago, Los Angeles, New Orleans, and New York City) from July 2015 to December 2019 and a lexicon approach to identify safety reviews posted by Airbnb guests. We find that 0.51% of the 4.8 million guest reviews express concerns about safety, among which 48.08% are about safety issues near but outside the focal property (such as local crime, which is referred to as vicinity safety reviews (VSRs)) rather than safety issues inside the property (such as a slippery tub or compromised lock, which is referred to as listing safety reviews (LSRs)). Both VSRs and LSRs are significantly more negative in sentiment than an average review, which is not surprising as guests that have *chosen* to stay at a dwelling owned or managed by an anonymous host usually assume that the neighborhood and property are reasonably safe.[2] A comparison with official crime statistics further suggests that the VSRs, although noisy and subjective, do reflect real safety risks in the related zip codes to some degree.

In general, critical consumer feedback may generate at least two information spillovers on a digital platform. First, buyer A's critical feedback on product listing X may deter herself and other buyers from buying X in the future. This "within-listing-crossbuyer" effect is typical in a reputation system and is well studied.[3] Second, a poor experience with listing X may motivate buyer A to give critical feedback to X and reassess other buyers' similar critical feedback toward other listings or even the whole feedback system. This "crosslisting-within-buyer" effect is often omitted because Bayesian updating assumes that learning from others' experience is the same as learning from self-experience if the information has the same accuracy. However, in practice, self-experience can be much more salient to an individual. Few researchers have quantified the second spillovers explicitly; one exception is Nosko and Tadelis (2015), who show that buyers who have bought from a more (less) reputable seller on eBay are more (less) likely to return to the platform to transact with *any* sellers, above and beyond the likelihood to transact with the same seller that created that good (bad) experience.

Although both VSRs and LSRs are likely to be feedback that criticizes an Airbnb listing, we highlight their differences in a few ways. By definition, LSRs are about safety issues inside the listed property, which is under the control of the host and can be addressed by changing the structure or amenities inside the property. It is hard to imagine that guests would blame the host of listing Y for the LSRs of listing X (assuming that the hosts of the two listings are unrelated). However, the host cannot do much about safety in listing X's vicinity. The VSRs associated with X may inform guests of the vicinity safety (VS) risk of nearby listings, which is a built-in spillover because of geographic proximity. In comparison, the "crosslisting-within-buyer" effect may occur regardless of geographic distance. Specifically, buyer A's self-experience of vicinity safety issues associated with listing X may lead A to recognize that similar negative shocks may be behind all VSRs written by other guests on other listings. Arguably, a similar logic could apply to LSRs as well, but the host's ability to address LSRs can mitigate the negative spillover of LSRs. Over time, guests may recognize that past LSRs on a listing are no longer relevant if the host fixed the issues and subsequent reviews were positive.

Empirical evidence supports the presence of both "within-listing-crossbuyer" and "crosslisting-within-buyer" spillovers. In particular, when we follow the same listings before and after they receive any VSRs or LSRs, there is a significant drop in the listings' monthly occupancy rates as well as average paid prices per night. The effect on occupancy is stronger for LSRs (−2.41%) than for VSRs (−1.45%), but the effect on price is comparable (−1.47% for VSRs and −1.46% for LSRs). Robustness checks that compare similar listings with and without safety reviews confirm that these effects are likely driven by the random arrival time of the VSR or LSR rather than omitted local demand or supply shocks. These findings suggest that prospective guests are concerned about both listing and vicinity safety and that they seem more sensitive to LSRs than to VSRs.

In addition to this classical "within-listing-crossbuyer" effect in listing reputation, we also find significant "crosslisting-within-buyer" effects for VSR and LSR. In particular, we compare the guests that wrote VSRs on Airbnb (referred to as vicinity safety guests) with the non-VS guests that booked similar listings (in terms of crime and VSRs) with similar frequency but never wrote any VSRs in our data set. A difference-in-differences (DID) analysis finds that VS guests are 60.07% less likely than non-VS guests to book future stays on Airbnb after posting the VSR, and when they do book on Airbnb, they tend to book in areas with fewer official crimes, fewer overall VSRs, and a lower percentage of listings with any VSR. The learning is weaker if the focal listing that triggered the VS guest's VSR had previously received any VSRs from other guests, but even in this case, the VS guests are still 51.62% less likely to book future stays on Airbnb after posting their own VSR. This suggests that self-experience is much more salient than reading other guests' VSRs; thus, the online review system is not fully effective as far as conveying all of the information embedded in VSRs. When we conduct a parallel exercise for guests who have written LSRs (as compared with similar guests who have not written LSRs), we find effects in the same direction but of a lower magnitude, suggesting that both LSRs and VSRs have a "crosslisting-within-buyer" effect but that the negative spillover of VSR is greater. The finding that VSRs have a greater "crosslisting-within-buyer" effect but a lower "within-listing-crossbuyer" effect than LSRs suggests that VSRs generate a greater negative shock in self-experience than LSRs.

Given these results, there is a possibility that the second type of information spillover, namely VS guests' stronger reactions to their own vicinity safety experiences (the crosslisting-within-buyer effect), may undermine a platform's incentives to post and highlight VSRs as critical feedback. This could occur because the platform's information policy may affect how a VS user's negative self-experience may change other guests' belief about the VSRs that they read on the platform without self-experience. Interestingly, in a recent policy change that took effect on 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.[4] This policy change, despite no evidence of strict enforcement, suggests that Airbnb is willing to consider a separate information policy for VSRs, apart from the traditional collection and posting policy for LSRs and other listing attributes under the host's control.

This consideration along with the differential information spillovers that we have documented for LSRs and VSRs motivate us to examine what would happen for guests, hosts, and the platform should Airbnb implement one of four counterfactual information policies for VSRs: (i) eliminating all VSRs while assuming no belief update among guests ("no disclosure no belief update"), (ii) eliminating all VSRs but allowing guests to form rational belief of average VSR risk conditional on observable listing attributes ("no disclosure but with rational belief"), (iii) alerting all guests to the existing VSRs and making them as informed as those who have written VSRs themselves ("high alert"), and (iv) keeping the information system as is but removing listings with 1+ or 2+ VSRs ("listing removal").

To conduct the counterfactuals, we incorporate competition between Airbnb and other short-term lodging options as within- and crossplatform sorting would have different implications for platform revenue. To account for such competition, we use a discrete choice model to estimate consumer utility from each Airbnb entire-home listing while treating VRBO listings and hotel stays in the same city-month as the outside good. We then use the structural estimates to quantify consumer surplus and Airbnb gross booking value (GBV) under the status quo of our sample (i.e., VSRs are largely permitted) versus the four counterfactual regimes.

Because VS guests are rare and we cannot track these guests in the data over time until they have continued to book on Airbnb and leave another review (with these actions being endogenous), the discrete choice model cannot identify how the self-experience of VSRs affects future booking by VS guests. To address this problem, we use our DID estimate of the "crosslisting-within-buyer" effect of VSRs to calibrate the coefficient of VSRs in the utility function, which measures how much bigger the shock of VSRs in self-experience must be—relative to reading VSRs written by other guests—to justify the future booking behavior of VS guests as observed in the raw data. This calibration enables us to distinguish between the actual utility that a guest may obtain from a listing with VSRs and the utility that the guest perceives at the time of booking.

Compared with the status quo, we find that not disclosing VSRs and no belief updates upon VSR removal would decrease consumer surplus in the market by 1.183% and increase Airbnb's GBV from the sample cities by 0.327%. This occurs because the no-disclosure policy generates a positive sorting toward listings formerly with VSRs and away from listings without VSRs and listings off Airbnb. Interestingly, the perverse incentive to suppress VSRs can be mitigated if we allow guests to form a rational belief of the average VSR risk conditional on observable listing attributes. In that case, the decline in consumer surplus is less (−0.993%) because VSR removal reminds guests of average VSR risk, which generates a negative information shock to listings without VSRs and motivates guests to shift demand away from Airbnb, although the positive information shock brings more bookings to listings with VSRs. In sum, the two countervailing forces reduce Airbnb's overall GBV by 0.047% and thus, discourage the platform from adopting a no-disclosure policy. In both no-disclosure regimes (with or without guests' belief updates), the effects can be softened if we allow listings to change their price up to 1%, depending on whether the counterfactual policy brings a negative or positive information shock to specific Airbnb listings.

Conversely, if Airbnb highlights VSRs and makes all guests as informed as those who have written VSRs themselves, the high alert would increase consumer surplus in the market by 9.599%–10.340% and decrease Airbnb's GBV by 2.726%–6.026%, depending on whether we allow listing price to change by 1% in response and whether we assume that the high alert on vicinity safety also applies to the VSRs for nearby listings. In comparison, removing listings with 1+ or 2+ VSRs would reduce consumer surplus by 1.187%–5.008% and depress Airbnb's GBV by 1.523%–2.883%. Both consumers and Airbnb suffer from listing removal because it reduces consumers' choice set.

In a dynamic simulation, we also consider a situation where Airbnb keeps the online review system as is (i.e., neither suppresses nor highlights VSRs) but consumers who experienced VSRs become high alert organically, even if everyone else with no such self-experience continues to hold their perception of VSRs as observed in our data. Our simulation suggests a slow process that decays Airbnb GBV but enhances consumer surplus, and its convergence toward platform-wide high alert depends on how much VSR experience is underreported in our data and how likely it is that consumers staying in VS listings end up with self-experiences that are reported as VSRs.

In short, we find that the interests of consumers and the platform do not always align, especially with respect to two extreme information policies. At one extreme, where consumers are not aware of the platform's suppression of VSR and do not update their beliefs of vicinity safety accordingly, misalignment could occur because removing VSRs encourages more guests to book on Airbnb and facilitates within-Airbnb sorting toward VS listings, although these changes end up hurting some consumers. Fortunately, a few market mechanisms—including consumers learning from self-experience and from updating their beliefs upon VSR suppression—help to realign the incentives and discourage the platform from suppressing VSR.

At the other extreme, where Airbnb highlights VSRs in a way that makes every potential host as alert as guests who have written VSRs themselves, misalignment could occur because such high alert drives consumers away from VS listings, and the sorting toward hotel and non-Airbnb listings may exceed the sorting toward non-VS listings on Airbnb, hurting the overall GBV of Airbnb. Although this suggests that Airbnb may lack incentives to adopt a high-alert policy right away, we show that consumer self-experience alone would push the market toward high alert over time.

Although the overall welfare effects are moderate (because VSRs are rare in the data), they mask large distributional effects; more VSR transparency benefits Airbnb listings without VSRs as well as the outside good at the cost of Airbnb listings with VSRs. Because listings with VSRs are more likely to be located in low-income or minority neighborhoods, consumer sorting upon VSR transparency would generate sizable revenue shifts across hosts in different neighborhoods. These effects highlight a potential trade-off as far as generating greater revenues and attracting hosts in low-income and minority areas on the one hand, which can enhance the economic impact of the platform in a city's underserved neighborhoods, and possibly providing additional value to guests on the other hand.

As detailed below, we contribute to the rising literature on the information design of online platforms and the empirical literature of online feedback and seller reputation. As information intermediaries, digital platforms have more incentives than traditional sellers to alleviate information asymmetries between buyers and sellers. But, they are still inherently different from a social planner because they may place 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 other competing platforms and outside options. Our empirical findings highlight these differences and quantify the extent to which consumers' self-experience and belief update upon review suppression can help to realign the incentives of the platform and consumers. We also document how the impact of a platform's information policy may vary for neighborhoods of different incomes or with different minority representation as being inclusive could be important for the platform and/or the social planner. These findings can help facilitate ongoing discussions of 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 on Airbnb's review system. Section 4 describes the data set, defines VSRs and LSRs, and provides summary statistics. Section 5 reports reduced-form evidence for the "within-listing-crossbuyer" and "crosslisting-within-buyer" effects of safety reviews. Section 6 incorporates these effects into a structural demand model and predicts how listings' GBV and consumer surplus would change under four counterfactual regimes and a dynamic simulation of the status quo. Section 7 discusses the implications of our findings and concludes with future research directions.

2. Related Literature

Our work is related to three strands of literature. First and foremost, we contribute to the growing literature on information design in online platforms.[5] Because consumer feedback is underprovided and there is a selection against critical feedback, researchers have studied the design of feedback systems as far as who is allowed to provide feedback (Mayzlin et al. 2014, Klein et al. 2016, Zervas et al. 2021), how to improve the authenticity of feedback (Wagman and Conitzer 2008, Conitzer et al. 2010, Conitzer and Wagman 2014, Gandhi et al. 2025), what kind of feedback is shown to the public, when to reveal the feedback to the public (Bolton et al. 2013, Fradkin et al. 2021), and how to aggregate historical feedback (Staats et al. 2017, Dai et al. 2018).

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

In our setting, we offer a counterexample where a platform's review policy has the potential to discourage buyers from providing a specific type of critical feedback. The discouragement can occur when a platform hides, obfuscates, or deletes critical 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 (Mayzlin et al. 2014, Luca and Zervas 2016, Gandhi et al. 2025).

At the same time, prior theoretical work has shown that platforms may be strategically motivated to omit certain information, including critical feedback. For instance, Kovbasyuk and Spagnolo (2024) explain why platforms may sometimes seek to erase certain historical bad records of sellers in order to increase matching rates. Romanyuk and Smolin (2019) show that platforms such as Uber and Lyft may seek to hide some buyer information (say, destination) prior to completing a buyer-seller match because doing so may 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, whereas 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.

In a different setting (online advertising auctions), Decarolis et al. (2023) use q-learning simulations to show that search engine platforms (such as Google and Bing) can increase their auction revenue by withholding bidding information from advertisers that bid repeatedly via artificial intelligence algorithms. Using similar simulations, Banchio and Skrzypacz (2022) show that the platform's gain from withholding bidding information occurs in second-price auctions but not in first-price auctions. Empirically, Blake et al. (2021) show that an online platform that matches buyers and sellers of the secondary-market sales of event tickets can increase the volume and quality of tickets sold by obfuscating the full purchase price to buyers until the final checkout step.

Our paper presents an empirical example of highlighting or withholding product *quality* information instead of *price* information. As shown in our counterfactual analysis, the platform may have economic incentives to downplay VSRs in some situations because VSRs may generate negative information spillovers to the rest of the platform. The bigger the negative shock is that VSRs can generate in self-experience, the more likely the other guests are to be as alerted about vicinity safety as VS guests, and the lower the matching rate is for the broader platform. In theory, such negative effects could be dominated by a sorting effect if posting or highlighting VSRs would direct buyers toward safer listings on the same platform and motivate safer listings to increase their prices sufficiently to compensate for the platform's loss from a lower matching rate. Conversely, the negative effects of highlighting VSRs may overwhelm the within-platform sorting effect as shown in our counterfactual analysis. In addition, the low probability of self-experiencing VSRs and the information frictions present in the current feedback system (such as buyer reluctance to post any critical feedback) could serve as natural information barriers to limit the negative spillovers of VSRs and therefore, encourage a platform to maintain the status quo rather than remove these information barriers for the benefits of consumers. On the positive side, we also find that a few market mechanisms—including consumers learning from self-experience and belief updating upon VSR suppression—help to realign the incentives of the platform and consumers. These market mechanisms counter the platform's incentive to suppress or downplay VSR, especially if the platform values its business in the long run.

The second literature to which we contribute is about online feedback and seller reputation. Our findings on the "within-listing-crossbuyer" effect of VSRs and LSRs confirm the classical literature of online seller reputation (see reviews by Bajari and Hortacsu 2004, Einav et al. 2016, and Tadelis 2016) and consumer response to critical feedback in particular (Chakravarty et al. 2010).

In addition, to our knowledge, we are among the few who have attempted to quantify crosslisting spillover effects of critical feedback. By definition, VSRs may generate spillovers among listings in nearby geographies should guests infer the overall safety of the vicinity from multiple nearby listings. Although this spillover is specific to the nature of vicinity safety (and difficult to distinguish from common shocks to listings in the same area), the crosslisting-within-buyer effect of VSRs and LSRs is more generalizable to other online platforms. As shown by Nosko and Tadelis (2015), buyers who had a good (bad) experience with a reputable seller on eBay are more (less) likely to return to eBay for sales with *any* sellers. Similarly, we show that having a negative safety experience tends to motivate a guest to subsequently avoid booking *any* listings on Airbnb in our sample cities and if they book again at all, to avoid both the listings and the areas that have any safety reviews. Compared with Nosko and Tadelis (2015), we show that the crosslisting-within-buyer spillover is not only limited to the extensive margin (whether to return to the platform for future transactions), but it also motivates experienced buyers to be more discerning and adjust how they interpret the presence of safety reviews in other listings.

The difference between VSRs and LSRs also allows us to separately identify the "crosslisting-within-buyer" effects of VSRs and LSRs. Their relative magnitudes suggest that VSRs may generate a larger negative shock than LSRs in self-experience, although the classical within-listing-crossbuyer effect of VSRs is smaller than that of LSRs. This difference highlights the importance of paying attention to the information spillovers of consumer feedback that tend to be missing in the classical seller reputation literature.

The crosslisting-within-buyer effect of consumer feedback could apply to many other platforms beyond eBay and Airbnb. For example, buyers of processed food may worry about contamination in food preparation, parents may worry about unsafe toys from countries with poor quality control standards, consumers of moving services may worry about road delays, and restaurant patrons may worry about the difficulty of finding parking. Some of these risks may be avoidable by the seller if she has full information and expertise to screen the supply chain, but often, individual sellers cannot change the production environment of their country of origin, cannot easily change the location of their business, and have little control over road conditions. Yet, consumers have legitimate concerns in these risky dimensions, although the risk is usually not observable until the small probability of negative outcomes manifests in practice. Once the negative outcome occurs in self-experience or is made equally known to consumers, consumers may quickly attribute the risk to sellers who receive similar critical feedback and intentionally avoid them. In some cases, wary consumers may even begin to watch out for the risk among all sellers on the platform. These potential negative effects present a dilemma to the platform; should the platform highlight such negative information at the risk of losing buyers and sellers, or should the platform withhold action and then act to minimize the impact of the negative outcomes when they occur? As previously indicated, this dilemma is not dissimilar to the dilemma facing tobacco, pharmaceutics, and SUV manufacturers, but the extent of the problem and the market-driven incentives to address it depend on the nature and impact of negative information for the whole platform as well as changes in consumer information through self-experience and belief updating.

Of course, the crosslisting-within-buyer spillovers are not necessarily limited to specific seller attributes. In our analysis, we assume that the presence of LSRs or VSRs is the only inference linkage between listings. In practice, a buyer who experiences a listing safety (LS) or vicinity safety issue with listing X may infer that other listings that are located in another neighborhood with similar demographics as the focal listing carry a similar LSR or VSR risk, even if these listings and their nearby listings do not have LSR or VSR at all. Because it is impossible to list all of the potential inference linkages that an affected buyer may use to expand their safety experience to the safety risk of other listings, we restrict our estimate to the inference linkage based on the presence of LSRs or VSRs in different listings. In doing so, we provide a conservative estimate for the impact of hiding or highlighting safety reviews because the more linkages that a buyer uses, the bigger the crosslisting-within-buyer spillovers there should be.

Another contribution that we make to the literature of online seller reputation is highlighting some long-run consequences of rare critical feedback, especially on product quality that is out of the control of sellers (vicinity safety). Because vicinity safety is a small probability event and buyers may be reluctant to submit critical feedback, VSRs on any Airbnb listing accumulate slowly over time, which could affect their overall informativeness. As we later show, between 2015 and 2019, we observe a growing rank correlation between a zip code's normalized cumulative VSR count and the zip code's normalized official crime statistics in low-income and minority areas. This suggests that cumulative VSRs do contain useful information regarding a zip code's actual safety status, and its informativeness may increase over time because of the law of large numbers. Furthermore, the rarity of VSRs highlights the importance of the platform's information policy because it affects the dissemination of the crosslisting-within-buyer effect from rare self-experience and hence, the informativeness of the gradually accumulated VSRs. In comparison, a few studies argue that online feedback systems may become less informative over time because of feedback bias (Klein et al. 2009, Barach et al. 2020, 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 and highlight that self-experience of vicinity safety issues can be much more salient than reading VSRs written by other guests.

Finally, we are not the first to study safety issues regarding online short-term rental (STR) platforms. Suess et al. (2020) find that nonhosting residents with 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, Kim et al. 2017, Nieuwland and Van Melik 2020). Zhang et al. (2022) show 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 rates in a city, whereas noncommercial 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, Han et al. 2020, Filieri et al. 2021, Roth 2021, Maldonado-Guzmán 2022). 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 2022) and to skirting local regulations (Jia and Wagman 2020). A recent study of Chicago's short-term rental regulations found that the incidence of burglaries has declined near buildings that prohibit STR listings (Jin et al. 2024).

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 as well as clarify how the interests of guests, hosts, and the platform may diverge with respect to the disclosure of VSRs. As shown in our counterfactuals, disclosing and highlighting VSRs can 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 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.[6] 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 of crimes, such as theft or burglary, at short-term rental lodgings reportedly increased by 555% from 2017 to 2019.[7] 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 that they may experience in a listing's vicinity.[8]

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,[9] and perhaps most readily linked to each listing, the listing's reviews from prior guests.[10] Airbnb enables guests and hosts to blindly review each other after a guest's stay.[11] 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 on 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.[12] 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 because such safety aspects are outside the control of the host. 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 that they suffer from "irrelevant" reviews about the vicinity of their listings,[13] and this policy change could be a way to address these complaints. Another reason might be concerns over review accuracy; arguably, vicinity safety is a subjective feeling, which is subject to the reviewer's priors and interpretation, and it is often difficult to prove correct or wrong. However, similar accuracy concerns could apply to other review content, although the degree of objectiveness may vary. A third reason may have something to do with the aspiration of being inclusive. Airbnb has advocated for "building a more inclusive travel community" and provided "education and inclusion resources for hosts."[14] The same aspiration may have motivated Airbnb to adopt an antidiscrimination policy, establish a permanent antidiscrimination team, and encourage designs and services friendly to users with disabilities. To the extent that VSRs are more present in low-income or minority neighborhoods, the new review policy could be another effort to make the platform friendlier to hosts in economically disadvantaged neighborhoods.

The frequency of VSRs in our raw data from mid-2015 to December 2020 presents no evidence indicating that Airbnb has enforced this policy post-December 2019 as far as vicinity safety is concerned. However, anecdotes suggest that some reviews that touched on neighborhood safety had been removed.[15] Our work does not depend on whether and how Airbnb enforces this policy as our analysis sample ends in December 2019 (to avoid potential market shifts because of the coronavirus disease 2019 (COVID-19) pandemic). Nevertheless, this new policy suggests that Airbnb is willing to consider different feedback policies depending on whether the focal issue is under the control of the host or not. This motivates us to distinguish between LSRs and VSRs and explore why these two types of buyer feedback may introduce different incentives for the platform's information design.

To be clear, Airbnb has adopted other methods to address neighborhood safety directly. For example, Airbnb introduced a neighborhood support hotline in December 2019,[16] around the same time that 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, because our main analysis sample ends in December 2019 and we do not know how many guests who left VSRs in our sample would have used the hotline should the hotline have existed at the time of the review, we cannot predict how the hotline may 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 one- to five-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."[17] Hence, the 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.

When potential guests search on Airbnb, the platform may not provide the precise address of each listing and depicts nearby listings on the same map. This setting makes it simple to identify nearby listings; thus, a guest observing VSRs on Airbnb listing X can extend the vicinity safety concern to all nearby Airbnb listings on the same map. However, the lack of an exact address makes it more difficult to (i) combine the listing information on Airbnb with external information sources, such as local news and crime statistics, and (ii) extend the same concern to listings on VRBO or other short-term rental platforms. Guests may not be familiar with streets and neighborhoods in the destination city, which further exacerbates the challenges with drawing connections among listings on different platforms, especially given that platforms may not provide precise addresses. Guests also may not always be able to tell whether two listings on Airbnb and VRBO are in fact the same listing. These information frictions imply that the potential spillover from one listing's VSRs to nearby listings is more salient for nearby listings on Airbnb than for potentially nearby listings on VRBO.

4. Data

We use several data sources to track short-term rental listings, official crime statistics, and some fundamentals of the short-term lodging market in each sample city. We describe each data source separately.

Data on Short-Term Rental Listings. The main data set that we use has information on short-term rental listings that had been advertised on Airbnb from July 2015 to December 2019 and on VRBO from June 2017 to December 2019 in five U.S. cities (Atlanta, Chicago, Los Angeles, New Orleans, and New York). The data were acquired from AirDNA, a company that specializes in collecting Airbnb and VRBO data. For Airbnb listings, this data set includes the textual contents of all Airbnb listing reviews in those cities. We have no access to reviews on VRBO. The original data from AirDNA extend to December 2020, but demand for short-term rentals subsequently changed dramatically because of the COVID-19 pandemic; therefore, our main analysis uses data up to December 2019.

Each listing is identified by a unique property identification and comes with time-invariant characteristics, such as the listing zip code and the listing property type (entire home, private room, shared room, or hotel room) as well as the host's unique identifier. Listings also have time-variant characteristics, including average daily rate (ADR),[18] the number of reservations, days that are reserved by guests, occupancy rate,[19] number of reviews, overall rating scores,[20] the listing's superhost status,[21] the listing's guest-facing cancellation policy,[22] the average number of words in the listing's reviews, the number of listings in the same zip code, and whether the listing is crosslisted on VRBO.[23] Although Airbnb and VRBO only provide proxy longitude and latitude for each listing, we are able to compare the proxy and actual locations in a few example listings based on our own or our friends' real Airbnb bookings. We find that the proxy location is usually within 150 meters of the actual location; thus, we treat the zip code corresponding to a listing's proxy longitude and latitude as its actual zip code, and we use proxy locations to define whether two listings are within each other's 0.3-mile radius.

Our unit of observation is listing-month. We focus on "active listings" (listings whose calendars are not indicated as "blocked" in the data set for an entire month) and exclude observations with an ADR of over $1,000 as some hosts may set their rates prohibitively high in lieu of blocking their calendars. We use regular monthly scrapes between July 2015 and December 2019 on Airbnb (from June 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.

Definition of Safety Reviews on Airbnb. Because we only observe guest reviews on Airbnb, we can only define LSRs and VSRs on Airbnb. LSRs 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"). VSRs 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"). Although 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 because of its simplicity and transparency. Lexicons are also found to have high levels of precision as compared with machine learning approaches (Hutto and Gilbert 2014, Zhang et al. 2014) and have been used extensively in the literature (Dhaoui et al. 2017, Monroe 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 male and female as well as different ethnicities) labeled 1,400 reviews that were generated from the lexicon approach algorithm with the initial keyword set for both LSRs and VSRs. 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) three specific 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 critical reviews regarding vicinity safety (listing safety) in the

1,300 sample with such a human-selected keyword is greater than 0% (10%).

In the second iteration of labeling, two research assistants (male and female) with different ethnicities labeled 3,100 reviews that were generated from the lexicon approach algorithm with the updated keyword set for both LSRs and VSRs such that five reviews associated with each keyword were randomly selected. In this iteration, reviewers labeled whether each review pertains to negative sentiment about vicinity 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 A1 in Online Appendix A.[24]

The keyword lists developed above are not the only inputs that we use to define vicinity or listing safety reviews. As far as VSRs, 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 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," and "outside") in Table A1 in Online Appendix A. 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.[25] Next, we selected 13 "negative" keywords and filtered out double-negative reviews where the keyword occurs within 5 words of a safety keyword. The whole procedure of our VSR and LSR definitions is illustrated by Figure A1 in Online Appendix A.

Overall, our approach resulted in 11,800 matched VSRs and 12,800 matched LSRs across the five sample cities. In total, they account for 0.25% and 0.27% of all of the observed Airbnb reviews, respectively. From July 2015 to December 2019, only 4.43% of listings ever had any VSR, 4.95% ever had any LSR, and 8.49% ever had any safety review (VSR or LSR). Conditional on having any VSRs by December 2019, 81.04% of listings have one VSR, 11.96% have two VSRs, and the remaining 7% have 3+ VSRs. Conditional on having any LSR by December 2019, 86.46% of listings have one LSR, 10.71% have two LSRs, and the remaining 2.83% have 3+ LSRs.

As shown in Figures A2 and A3 in Online Appendix A, 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.[26] 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"). Although 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.

To check whether the safety reviews defined above are indeed critical feedback, as we intended to identify, we employ a pretrained Natural Language Processing model from Hugging Face to determine the sentiment score of all reviews.[27] According to the analysis, the overall average sentiment score across all available reviews is 0.79. Specifically, VSRs show a relatively neutral average sentiment score of 0.06, whereas sentences containing VSR safety keywords tend to have a negative average sentiment score of −0.31. In comparison, LSRs demonstrate a lower average sentiment score of −0.41, and sentences with LSR safety keywords have the most negative average sentiment score of −0.76. By contrast, the non-VSRs or non-LSRs have an average sentiment score matching the overall average of 0.79. These patterns suggest that our lexicon approach has successfully captured the negative sentiment when guests comment on listing or vicinity safety issues during their stay.

Sensitivity Test on Safety Review Definitions. Because the sets of safety keywords are selected based on manual labeling, we conduct a sensitivity check. In particular, at the end of the first-round iteration, we refined our keyword selection by focusing on the keywords for which the percentage of critical reviews regarding vicinity safety (listing safety) in the 1,300-keyword sample is greater than 50% (50%) rather than 0% (10%). This means that we included only those with higher relevance and more critical sentiment for the second-round iteration. As a result, the alternative definition identified 32 vicinity safety keywords (e.g., "homeless" and "drugs") and 47 listing safety keywords (e.g., "mold" and "stained") as shown in Figures A6 and A7 in Online Appendix A. This refined set of keywords resulted in 5,272 VSRs and 12,150 LSRs, which are roughly 55% and 5% less than what we find in the main definition, respectively. Consequently, 1.82% of listings had any VSRs and 4.71% had any LSRs as compared with 4.43% and 4.95% in the main definition, respectively. Despite these differences, we find similar results in the listing-level regressions (defined in Section 5.1). In particular, the coefficients reflecting the effects of a listing's VSRs and LSRs on its own price and occupancy become stronger in magnitude (in the same direction as using the main definition of VSRs and LSRs), likely because the VSRs and LSRs under the alternative definition have a higher probability of capturing actual and severe safety issues.[28]

Official Crime and Demographic Statistics. A second data set that we collected covers official crime records from databases tracking crimes in Chicago,[29] New Orleans,[30] New York City,[31] Atlanta,[32] and Los Angeles.[33] 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 U.S. Census Bureau.[34] 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 in which it is located. Similarly, we refer to a zip code as minority (M) or white (W) as a function of whether the percentage of its population that is identified as minority is below or above the city median.

Hotel Lodging Data, Air Travel Data, and Zillow Home Value Index. To capture potential competition with Airbnb and VRBO in the short-term lodging market, we use two supplemental data sets to define market size. First, we obtain data from Smith Travel Research on total hotel booking volume and revenue by zip code and month from 2015 to 2019 in our sample cities. These data do not contain hotel-specific information, so we cannot distinguish among different types of hotels within the same zip code. It turns out that only 24.6% of zip codes in our data have any hotel data because hotels tend to concentrate in the commercial areas of a city, whereas Airbnb and VRBO listings can be spread out in all kinds of neighborhoods throughout the city. About 40% of the Airbnb listings that we observe in the AirDNA data are located inside these hotel-present zip codes. Second, we use the U.S. Department of Transportation's T100 (top 100 metropolitan areas) data to calculate total incoming air travelers (domestic and international) per city-month.

If we define the short-term lodging market by city-month, we can measure the market size by (a) the total amount of occupancy in hotels, Airbnb, and VRBO or (b) the total count of incoming air travelers. The latter is five to nine times bigger than the former on average because many incoming air travelers may live in the city or leave the city on the same day. Nevertheless, the two measures are highly correlated, with a correlation coefficient ranging from 0.5 to 0.9 across the five sample cities.

An alternative way of defining the short-term lodging market is by zip code and month. This detailed definition may better capture the head-to-head competition within each zip code, but given the fact that most zip codes do not have any hotels, VRBO would be the only outside option competing with Airbnb in these markets. This is imperfect because VSRs of an Airbnb listing may remind guests of the potential vicinity safety risk of nearby listings on VRBO, although the lack of precise addresses may make it difficult to pin down exactly what VRBO listings are close to the focal listing on Airbnb. In Section 6, we check how sensitive our structural estimation results are to the market definition (city-month or zip code-month) and to the definition of market size (Airbnb + VRBO, hotel + Airbnb + VRBO, and incoming air travelers).

As detailed in Section 6, we use Zillow's Home Value Index (ZHVI; by zip code and month) to construct instruments for listing price. Zillow defines ZHVI as a measure of the typical home value and market changes across a given region and housing type. It reflects the typical value for homes in the 35th–65th percentile range. Although ZHVI is an imperfect measure of the cost of running Airbnb listings in a particular zip code-month, it embodies property tax, property maintenance costs, and the opportunity costs of using the property for alternative uses. We download the seasonally adjusted version of ZHVI[35] and merge it with other data by zip code and month.

Table A2 in Online Appendix A defines the key variables that we use, including listing attributes (such as price, occupancy rate, safety reviews, and ratings) and neighborhood attributes (such as income, population, and crime statistics by zip code).

Summary of VSRs and LSRs on Airbnb. Table 1 summarizes the data at the listing-month level, where vicinity safety Airbnb listings are defined as observations that have a positive number of vicinity safety reviews before the reporting month, whereas non-VS Airbnb listings do not have any VSRs before the reporting month. As Table 1 indicates, about 4% of the total observations are VS listings. On average, VS listings have a higher occupancy rate, a higher number of reservations, a higher fraction of superhosts, and a higher number of reviews than non-VS listings. In contrast, the nightly rates and overall ratings of VS listings are lower on average than non-VS 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 is 0.06. Figures A4 and A5 in Online Appendix A demonstrate the distribution of VS keywords for four groups of zip codes (high income, low income, white, and minority).

Table 1. Summary Statistics of Airbnb Listings (July 2015 to December 2019)

	All listings (N = 2,866,238)		VS listings (N = 126,868)		Normal listings (N = 2,739,370)	
Variables	Mean	P50	Mean	P50	Mean	P50
Occupancy rate (0–1)	0.56	0.64	0.68	0.78	0.56	0.64
1 if any occupancy in the month	0.85	1.00	0.95	1.00	0.85	1.00
Price (average daily rate $)	164.69	125.51	134.15	106.31	166.10	126.67
# of reservations in the month	3.77	3.00	5.76	5.00	3.68	3.00
# of reservation-days in the month	14.16	14.00	18.56	21.00	13.95	14.00
1 if any VSRs since 2015/7 to last month	0.04	0.00	1.00	1.00	0.00	0.00
1 if any LSRs since 2015/7 to last month	0.05	0.00	0.20	0.00	0.04	0.00
# of VSRs since 2015/7 to last month	0.06	0.00	1.34	1.00	0.00	0.00
# of LSRs since 2015/7 to last month	0.06	0.00	0.26	0.00	0.05	0.00
% of any VSRs within 0.3-mile radius	0.07	0.04	0.10	0.07	0.07	0.03
Overall ratings (1–10)	9.18	9.60	9.09	9.20	9.18	9.60
# of reviews	33.71	15.00	93.02	70.00	30.96	14.00
# of listing within zip code	540.67	449.00	554.66	481.00	540.02	447.00
1 if crosslisting on VRBO	0.02	0.00	0.03	0.00	0.02	0.00
1 if superhost	0.23	0.00	0.26	0.00	0.23	0.00
1 if strict cancellation policy	0.50	0.00	0.58	1.00	0.49	0.00
Avg word count in a review since 2015/7	53.83	50.43	57.49	53.91	53.66	50.20
Median income in zip code ($)	57,187	50,943	42,645	34,432	57,861	51,427
Population in zip code	48,158	45,747	42,514	36,654	48,419	46,025
% white in zip code	0.53	0.59	0.41	0.38	0.53	0.60
1 if zip code is high income	0.52	1.00	0.29	0.00	0.53	1.00
1 if zip code is white	0.60	1.00	0.44	0.00	0.61	1.00
Normalized crime reports in zip code since 2015/7	0.86	0.21	1.69	0.33	0.83	0.20

Notes. This table summarizes Airbnb listings from July 2015 to December 2019 in the five sample cities. The variable for crime reports is reported by zip code-year-month and normalized by the population of the zip code. Unit of observation = listing-month. P, percentile.

Comparing the high-income and low-income (or white and minority) groups, it appears that the low-income (minority) group dominates the volume of VSRs.

How Do VSRs Correlate with Official Crime Statistics? We normalize the total number of reported crime cases in a zip code-month by population size in that zip code. The Pearson correlation between this normalized crime flow and the flow of all VSRs reported in a zip code-month is low (0.04). If we count VSRs cumulatively from July 2015 to the focal month[36] and correlate them with the flow of official crime counts, the correlation increases to 0.08. If we use cumulative counts in both, the correlation is 0.14.

Although the Pearson correlation between VSRs and total crime counts is fairly low at the zip code-month level, the ordinal order of vicinity safety across zip codes in the same city might be more informative than the absolute magnitude of either statistics. This motivates us to compute the *rank* correlation between the two. In particular, for crime counts, we rank the normalized flow crime data per zip code within each city-month and determine the percentile crime rank of the zip code for that month. For VSRs, we use the percentile rank of the number of flow VSRs in the zip code in the reporting month within each city. The correlation between these two ranks is 0.32. If we compute the percentile rank of VSRs by the number of *cumulative* VSRs in the zip code up to the reporting month within each city, its correlation with the percentile rank of flow crime data is 0.58, and its correlation with the percentile rank of cumulative crime data is 0.59. These numbers suggest that it is more important to capture VSR in cumulative counts because VSRs are rare, whereas whether to use flow or cumulative measures for normalized crime data is less crucial. In our reduced-form and structural analyses, we always use the raw data of VSRs (at the listing-month level) and crime reports (at the zip code-month level), not their percentile ranks, and therefore, we do not have a collinearity problem given their low correlation in the raw data.

To explore how VSRs and crime statistics correlate differently for different types of demographic areas, we compute the percentile rank correlation index between the zip code-level VSR count (cumulative) and crime count (flow) data in each month for the whole sample and the four groups of zip codes (high income, low income, white, and minority) separately. Figure 1 indicates that the percentile rank correlation exhibits an increasing trend, especially in low-income and minority groups, suggesting that the percentile rank of cumulative VSRs in a zip code has increasingly more power, reflecting the actual flow of crime reports over time in these areas.

Heterogeneity by Type and Area of Listings. Table A3 in Online Appendix A provides summary statistics based on the type of an Airbnb listing. The majority of listing-months in our whole sample are entire-home listings (60.9%), which tend to charge a much higher daily average price ($212.81) than private-room ($91.67) and shared-space ($58.23) listings. Some hotels are listed on Airbnb as well; their daily price ($197.16) is similar to that of entire-home listings, but hotel listings only account for 0.3% of all listing-month observations in our sample. Hotel and entire-homing listings are more likely to have any VSRs (cumulative since July 2015) than private-room and shared-space listings, but the likelihood of having any LSRs (cumulative since July 2015) is the highest among entire-home listings followed by private-room listings and hotel listings and is the least in shared-room listings. For nonhotel listings, the averages of the cumulative numbers of VSRs and LSRs are similar to the dummy of having any VSRs or LSRs because most listings with any VSRs or LSRs tend to have one rather than multiple such safety reviews. The average cumulative number of VSRs for hotel listings is higher than the average of having any VSRs, likely because each hotel listing may correspond to multiple hotel rooms.

Table A3 in Online Appendix A provides summary statistics based on whether a listing is located in a high-income or low-income zip code and in a white or minority zip code. The number of listing-months is comparable between H and L zip codes but higher in W zip codes than in M zip codes. Listings in L and M areas are much more likely to have any VSRs and any VSRs nearby than those in H and W areas. These differences are typically between 0.06 and 0.07 in L and M zip codes versus between 0.02 and 0.03 in H and W zip codes. However, the likelihood of having any LSRs is comparable across L, M, H, and W zip codes (all around 0.05). The cumulative crime counts, normalized by zip code population, are of a completely different scale, with an average of 0.56 in H zip codes and 1.19 in L zip codes. Although the average normalized crime counts are higher in W zip codes than in M zip codes (1.10 versus 0.51, respectively), the median is higher in M zip codes than in W zip codes (0.23 versus 0.19, respectively). This suggests that the normalized crime count in W zip codes is more skewed than that in M zip codes.

5. Reduced-Form Effects of Safety Reviews

We first present reduced-form evidence from listing-level and guest-level analyses. The listing-level analysis documents the within-listing-crossbuyer effects of VSRs and LSRs. It also explores the possibility that VSRs of nearby listings could affect the focal listing's price and occupancy. The guest-level analysis aims to capture the crosslisting-within-buyer effects of VSRs and LSRs.

5.1. Listing-Level Analysis

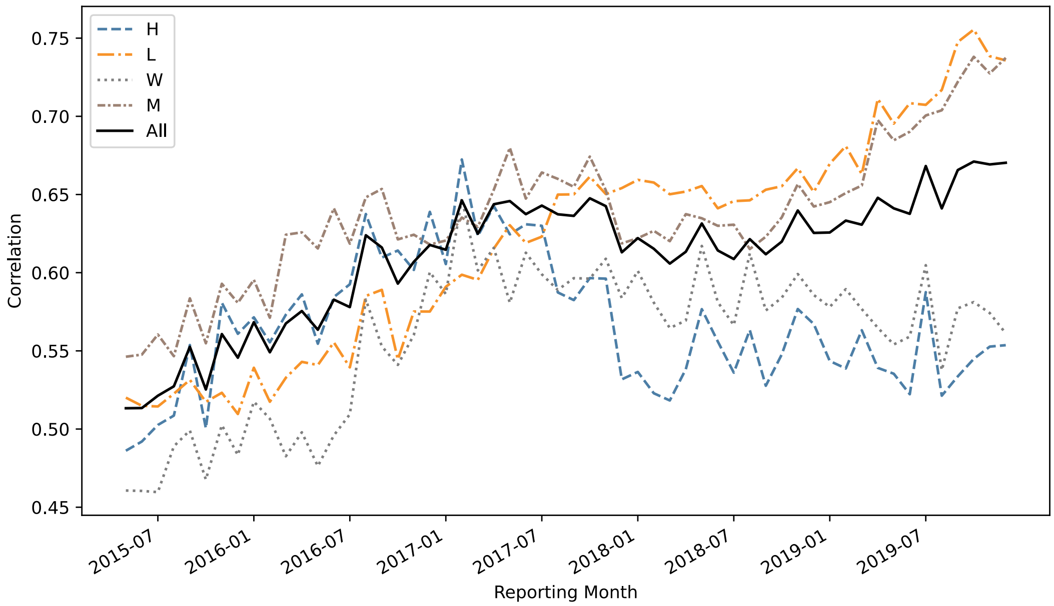
Baseline Results. We begin by assessing the effects of VSRs and LSRs by listing-month. Our hypothesis is that if potential guests view VSRs and LSRs as a proxy for safety around or within a listing, such reviews would reduce the guests' willingness to book the listing. Our base specification is given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

$$y_{j,t} = \alpha_j + \alpha_{k,t} + \delta X_{j,t} + [\beta_1 Crime_{j,t-1}] + \beta_2 LSR_{j,t-1} + \beta_3 VSR_{j,t-1} + \beta_4 VSRADIUS_{j,t-1} + \epsilon_{j,t}, \tag{1}$$

where j denotes a listing j-month t observation and $Crime_{j,t-1}$ 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 j is located. $LSR_{j,t-1}$ and $VSR_{j,t-1}$ are two dummy variables that equal one if the listing has at least one LSR or one VSR, respectively, before month t. The variable $VSRADIUS_{j,t-1}$ is the percentage of listings that have at least one VSR within a 0.3-mile radius of listing j prior to month t; $X_{j,t}$ 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, crosslisting status (i.e., whether the listing is also listed on VRBO), and whether the listing is hosted by a superhost. The dependent variable $y_{j,t}$ is either the log of listing j's average daily rate (price) in month t or the log of listing j's monthly occupancy rate (calculated as log of one plus the occupancy rate).[37] Coefficient α_j denotes listing fixed effects, and $\alpha_{k,t}$ denotes city-year-month or zip code-year-month fixed effects as we experiment with various controls for local shocks. Standard errors are clustered by Airbnb property identification. The primary assumption is that within a listing, the presence and timing of safety reviews are correlated with the true safety conditions 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.

Panel A of Table 2 presents three versions of the Ordinary Least Squares (OLS) results. Columns (1)–(4) in panel A of Table 2 control for city-year-month fixed effects, with and without $Crime_{j,t-1}$ on the right-hand side; Columns (5) and (6) in panel A of Table 2 control for zip code-year-month fixed effects, which automatically absorb $Crime_{j,t-1}$. We prefer columns (5) and (6) in panel A of Table 2 because they control for arbitrary local demand or supply shocks at the zip code level and address the concern that official crime statistics may include safety issues related to past Airbnb activities

Figure 1. (Color online) Percentile Rank Correlation Between Normalized Crime Flow and Cumulative VSR per Zip Code (Ranks Are Computed Within Each City-Month)



and therefore, that they may be endogenous and confound the interpretation of other coefficients.

Across all columns in panel A of Table 2, we observe that having any VSRs or LSRs on the listing is associated with a significant decrease in a listing's price and occupancy. Specifically, according to columns (5) and (6) in panel A of Table 2, for an average Airbnb listing in our sample, having any VSRs before the study month is associated with a 1.45% reduction in the listing's monthly occupancy rate and a 1.47% reduction in its average price per reserved night; having any LSRs is associated with a 2.41% drop in occupancy and 1.46% reduction in price. LSRs thus have a larger effect on occupancy than VSRs. The coefficient on *VSRADIUS* is negative and significant in columns (1) and (2) in panel A of Table 2 but becomes less significant after we control for *Crime* in columns (3) and (4) in panel A of Table 2 and statistically nondistinguishable from zero after we control for zip code-year-month fixed effects in columns (5) and (6) in panel A of Table 2. These results suggest that although nearby listings' VSRs could have a negative spillover on the focal listing, it is difficult to distinguish this effect from zip code-year-month shocks that apply to focal and nearby listings at the same time.

Because Equation (1) includes listing fixed effects and defines *VSR* and *LSR* cumulatively since May 2015, their coefficients capture the within-listing changes of occupancy and price before and after the listing receives its *first* VSR or LSR. We choose this definition because most listings that have any VSRs (LSRs) have only one VSR (LSR); hence, this margin is the most salient variation in our data. Results are similar if we exclude listings with 2+ VSRs or 2+ LSRs from the sample.

Still, a curious question is when the effects of VSRs and LSRs kick in and persist over time. To answer it, we redefine *VSR* and *LSR* as having any VSRs/LSRs within the past 12 months, more than 12 months ago, within the past 6 months, or more than 6 months ago. As reported in columns (1) and (2) in panel B of Table 2, when we only define *VSR* and *LSR* as having any VSRs/LSRs within the past 12 months (while controlling for zip code-year-month fixed effects and thus, absorbing *Crime*), the coefficients of *VSR* and *LSR* have the same sign and significance as what we obtain by using cumulative measures, but the magnitudes are smaller, especially when the dependent variable is occupancy rate. In columns (3) and (4) in panel B of Table 2, we further control for having any VSRs or LSRs more than 12 months ago, and the coefficients of VSR and LSR variables are much more similar in magnitude to what we obtain by using cumulative measures. In particular, the VSR or LSR coefficients on price are stable, but the coefficients on occupancy suggest that the negative impacts of VSR and LSR on occupancy are strengthened over time within a listing. The same patterns occur when we redefine *VSR*/*LSR* as having any VSRs/LSRs in last 6 months and more than 6 months ago. In an unreported table, we have tried to rerun the regressions in panel B of Table 2 excluding listings with 2+ VSRs or 2+ LSRs. The same pattern remains, suggesting that the strengthened effect of having any VSRs or LSRs is not driven by listings accumulating more VSRs/LSRs over time.

Table 2. Baseline Results of Reduced-Form Listing-Level Analysis of Airbnb Listings

Dependent variable	(1) Log *occupancy rate*	(2) log(*price*)	(3) Log *occupancy rate*	(4) log(*price*)	(5) Log *occupancy rate*	(6) log(*price*)
Panel A: Cumulative *VSR*, *LSR*, *VSRADIUS*, and *Crime*						
Any VSR since 2015/7 to last month	−0.0171***	−0.0156***	−0.0160***	−0.0154***	−0.0145***	−0.0147***
	(0.00140)	(0.00219)	(0.00140)	(0.00219)	(0.00136)	(0.00209)
Any LSR since 2015/7 to last month	−0.0253***	−0.0156***	−0.0249***	−0.0155***	−0.0241***	−0.0146***
	(0.00135)	(0.00210)	(0.00135)	(0.00210)	(0.00130)	(0.00200)
% of Any VSR within 0.3-mile radius	−0.00593**	−0.0107***	−0.00323	−0.0103***	−0.00228	−0.00224
	(0.00253)	(0.00393)	(0.00252)	(0.00390)	(0.00239)	(0.00377)
log(*crimes in zip code since 2015/7 to last month*)			−0.0720***	−0.0107	Absorbed	Absorbed
			(0.00950)	(0.0152)		
Property identification FE	Yes	Yes	Yes	Yes	Yes	Yes
City-year-month FE	Yes	Yes	Yes	Yes	No	No
Zip code-year-month FE	No	No	No	No	Yes	Yes
Observations	2,866,238	2,866,238	2,866,238	2,866,238	2,866,238	2,866,238
R^2	0.559	0.928	0.559	0.928	0.566	0.929
Panel B: More detailed lags of VSR and LSR						
Any VSR in last 12m	−0.00371***	−0.00905***	−0.00918***	−0.0135***		
	(0.00111)	(0.00166)	(0.00122)	(0.00185)		
Any VSR more than 12m ago			−0.0205***	−0.0163***		
			(0.00195)	(0.00329)		
Any LSR in last 12m	−0.0101***	−0.0111***	−0.0187***	−0.0147***		
	(0.00108)	(0.00152)	(0.00120)	(0.00175)		
Any LSR more than 12m ago			−0.0336***	−0.0145***		
			(0.00202)	(0.00330)		
Any VSR in last 6m					−0.00393***	−0.0129***
					(0.00118)	(0.00171)
Any VSR more than 6m ago					−0.0204***	−0.0141***
					(0.00165)	(0.00264)
Any LSR in last 6m					−0.0136***	−0.0160***
					(0.00117)	(0.00164)
Any LSR more than 6m ago					−0.0317***	−0.0127***
					(0.00165)	(0.00266)
log(*lagged crimes in zip code*)	Absorbed	Absorbed	Absorbed	Absorbed	Absorbed	Absorbed
Property identification FE	Yes	Yes	Yes	Yes	Yes	Yes
Zip code-year-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,866,238	2,866,238	2,866,238	2,866,238	2,866,238	2,866,238
R^2	0.566	0.929	0.567	0.929	0.567	0.929

Notes. This table reports the baseline results following Equation (1). The sample consists of all Airbnb listings from July 2015 to December 2019 in the five sample cities. All regressions control for property identification fixed effects (FEs), and listing attributes including the number of reviews, star ratings, whether the listing is a superhost, whether the listing is crosslisted on Airbnb and VRBO, whether the listing offers a strict cancellation policy, and the number of Airbnb listings in the same zip code. Columns (1)–(4) in panel A control for city-year-month fixed effects, columns (5) and (6) in panel A and columns (1)–(6) in panel B control for zip code-year-month fixed effects. Standard errors are in parentheses.

***$p < 0.01$.

The growing impact of VSRs/LSRs on a listing is somewhat surprising; by default, Airbnb presents consumer reviews by recency, and thus, a review posted months ago may become less visible to prospective guests if the listing accumulates a large number of reviews over time. However, Airbnb expands rapidly during our sample period, and media reports on safety and community concerns of Airbnb listings have grown over time. It is possible that newer guests are more wary about safety issues and pay more attention to safety reviews. Past critical feedback, like VSRs or LSRs, may act as an "anchor" for interpreting subsequent reviews, even if newer reviews do not mention safety issues directly. It is also possible that consumer reviews (including safety reviews) play some role in Airbnb's sorting and recommendation algorithms, and thus, listings with VSRs/LSRs are less discoverable by guests over time, which hurts the listings' occupancy.

Alternative Specification (DID + Matching). Admittedly, the baseline specification assumes that *when* safety reviews (LSRs or VSRs) appear in the online review record of an Airbnb listing is random and independent of time-varying demand shocks to that listing once we control for listing fixed effects, zip code-year-month fixed effects, and observable listing attributes. However, listings vary in many ways; their different experience and history on Airbnb could affect the occurrence of safety review(s) as well as today's occupancy rate and price regardless of LSRs or VSRs.

To address this concern, we identify 1,566 listings that have any VSRs in our data (hereafter, "VS listings"); for each of them, we use propensity score matching to find another two non-VS listings that never receive any VSRs but look most similar to the treated listing up to the month before the VS listing received its first VSR. The variables that we use to match VS and control listings include listing type, number of bedrooms, log of average number of reviews, log of rating score, superhost status, cancellation policy, crosslisting status, average zip code category (high-income and white majority), and log of average number of listings in the zip code. Because different VS listings may receive their first VSR at different times, we organize VS listings into cohorts by the month of their first VSR and perform the aforementioned matching for each cohort separately. To measure the matching quality between VS and control listings, Figure A8 in Online Appendix A shows that the propensity score distribution is well overlapped between these two groups, and Table A4 in Online Appendix A shows that the two groups are well balanced in listing attributes.

Pooling the observed months for the 1,566 VS listings and the corresponding 3,132 matched control listings, we run a difference-in-differences specification:

$$y_{jt} = \alpha_t + \alpha_j + \beta_1 \cdot VS\_listing_i + \beta_2 \cdot post\_1st\_VSR_{p,t} + \beta_3 \cdot VS\_listing_j \times post\_1st\_VSR_{p,t} + \epsilon_{j,t}, \quad (2)$$

where j denotes listing, p denotes the treatment-control pair, $VS_{listing}$ is a dummy of whether the listing is VSR treated, $post\_1st\_VSR$ is a dummy indicating that t is after the first VSR of the treated listing (or the matched treated listing if j is a control listing), and the DID coefficient of the interaction captures how the listing's occupancy or price changes after it receives the first VSR as compared with similar control listings. We control for listing fixed effects and city-year-month fixed effects. Standard errors (in parentheses) are clustered by treatment-control pair.

As shown in columns (1) and (2) in panel A of Table 3, the estimated DID coefficients are negative and significant with 99% confidence, confirming that listings receiving any VSRs do suffer from a decrease in occupancy and price.

We repeat the exercise by identifying 1,759 listings that have received any LSR (hereafter, LS listings) and matching each of them with two control listings that have no LSR but are most similar to the LS listing in observable attributes. The matching quality between LS listings and their corresponding controls is presented in Figure A9 and Table A5 in Online Appendix A. Pooling the observed months of 1,759 LS listings and 3,518 corresponding control listings, we estimate a parallel DID specification:

$$y_{jt} = \alpha_t + \alpha_j + \beta_1 \cdot LS\_listing_j + \beta_2 \cdot post\_1st\_LSR_{p,t} + \beta_3 \cdot LS\_listing_j \times post\_1st\_LSR_{p,t} + \epsilon_{j,t}, \quad (3)$$

where *LS_listing* is a dummy indicating whether j is an LS listing, p denotes the treatment-control pair, $post\_1st\_LSR$ is a dummy indicating whether t is after j (or the LS listing paired with j if j is a control listing) has received its first LSR, and the DID coefficient of the interaction term captures the average impact of LSRs on the performance of LS listings. We control for listing fixed effects and city-year-month fixed effects. Standard errors are clustered by treatment-control pair.

The estimated DID coefficients are reported in columns (1) and (2) in panel B of Table 3. Again, both of them are negative and significant with 99% confidence, confirming the OLS finding that a listing tends to suffer in price and occupancy after it starts to receive any LSRs.

Note that for occupancy, the DID coefficients based on the matched samples (Table 3) are of greater magnitudes than the coefficients of VSRs and LSRs in the baseline OLS regressions (columns (5) and (6) in panel A of Table 2). We can think of two reasons. First, the DID samples compare VSR and LSR listings with a selected group of control listings that look most similar to them in observable attributes; hence, the DID + matching design is more immune to potential confounding factors in the whole-sample OLS regression. In this sense, the DID coefficients should be closer to the true effect of VSRs or LSRs. Second, the treated and control definitions in the DID samples are based on a single binary indicator, and the DID coefficient can only identify the effect of this single binary variable switching from zero to one. In practice, multiple "treatments" may occur simultaneously or sequentially; a listing can have both VSRs and LSRs, and a listing with VSRs may also have other nearby listings with VSRs. To the extent that these "treatments" are positively correlated, the DID coefficients may end up capturing the sum of all of them.

For both DID analyses, columns (3) and (4) in Table 3 explore how the DID coefficients change 1–3, 4–6, 7–12, and 13+ months after the listing receives its first VSR or LSR. Consistent with the baseline OLS results (panel B of Table 2), we observe that the effects of VSRs and LSRs occur soon after their presence in a listing; this effect is stable for price but somewhat strengthened

Table 3. DID Results of Matched Airbnb Listings with or Without VSRs or LSRs

Dependent variable	(1) log(*occupancy rate*)	(2) log(*price*)	(3) log(*occupancy rate*)	(4) log(*price*)
		Panel A: Matched sample by ever VSR		
VS listing × post 1st VSR	−0.0273***	−0.00697***		
	(0.00177)	(0.00201)		
VS listing × 1–3m post 1st VSR			−0.0195***	−0.00927***
			(0.00300)	(0.00340)
VS listing × 4–6m post 1st VSR			−0.0321***	−0.00942***
			(0.00305)	(0.00347)
VS listing × 7–12m post 1st VSR			−0.0278***	−0.0101***
			(0.00264)	(0.00299)
VS listing × 13+m post 1st VSR			−0.0366***	0.00252
			(0.00245)	(0.00278)
No. of observations	147,576	147,576	147,576	147,576
R^2	0.446	0.927	0.447	0.927
		Panel B: Matched sample by ever LSR		
LS listing × post 1st LSR	−0.0363***	−0.0107***		
	(0.00178)	(0.00202)		
LS listing × 1–3m post 1st LSR			−0.0237***	−0.0164***
			(0.00298)	(0.00339)
LS listing × 4–6m post 1st LSR			−0.0443***	−0.0127***
			(0.00305)	(0.00347)
LS listing × 7–12m post 1st LSR			−0.0337***	−0.0168***
			(0.00263)	(0.00300)
LS listing × 13+m post 1st LSR			−0.0503***	0.00603***
			(0.00254)	(0.00289)
No. of observations	161,427	161,427	161,427	161,427
R^2	0.454	0.925	0.455	0.925

Notes. This table reports the DID results at the listing level. The sample in panel A consists of Airbnb listings that ever have VSRs and the control listings that are similar to them in observable attributes and Airbnb history before the first VSR occurs. The sample in panel B consists of Airbnb listings that ever have LSRs and the control listings that are similar to them in observable attributes and Airbnb history before the first LSR occurs. All regressions control for listing fixed effects, city-year-month fixed effects, and the post dummy itself. Standard errors (in parentheses) are clustered by treatment-control pair. Standard errors are in parentheses.
***$p < 0.01$.

over time for occupancy. The event-study plots for these DID analyses (Figures A10 and A11 in Online Appendix A) confirm this pattern.

In short, the DID + matching results give us confidence with regard to the negative impact of VSRs and LSRs on the performance of Airbnb listings, but the OLS estimates in Table 2 provide us with a more comprehensive picture of how *VSR*, *LSR*, and *VSRADIUS* relate to listing performance. They further allow us to distinguish the within-listing-crossbuyer effects of VSRs and LSRs from a possibility that VSRs of nearby listings may raise a vicinity safety concern regarding the focal listing. Thus, in the remainder of Section 5.1, we use OLS estimates with zip code-year-month fixed effects and cumulative measures of VSRs and LSRs (columns (5) and (6) in panel A of Table 2) as the baseline results to explore mechanisms and heterogeneous effects.

Mechanisms. To explore whether the baseline effects are driven by the extensive or intensive margins, column (1) in Table 4 considers as the dependent variable a dummy that equals one when a listing's occupancy rate is positive and zero otherwise. Column (2) in Table 4 reruns the OLS baseline specification (Equation (1)), conditional on a listing's occupancy rate being positive. The coefficients of *VSR* and *LSR* are always negative and significant in columns (1) and (2) in Table 4 as in the baseline results. This robustness suggests that these two variables are negatively correlated with listing performance on both the extensive and intensive margins. As in the baseline results, once we control for zip code-year-month fixed effects, the coefficient of *VSRADIUS* is statistically insignificant from zero, and *Crime* is absorbed by the fixed effects.

Another mechanism that we explore is whether the baseline results are driven by the visibility of VSRs or LSRs to potential guests. To do so, we split the sample by whether a listing-month has more than 13 reviews, where 13 is close to the median in the sample, recognizing that prospective guests are more likely to notice

Table 4. Mechanisms for Reduced-Form Listing-Level Analysis of Airbnb Listings

Sample	(1) Whole	(2) Occupancy > 0	(3) # Reviews ≤ 13	(4) # Reviews > 13
Dependent variable	*Occupancy rate dummy*	log *occupancy rate*	log *occupancy rate*	log *occupancy rate*
Any VSR since 2015/7 to last month	−0.00883***	−0.0107***	−0.0164***	−0.00764***
	(0.00153)	(0.00115)	(0.00498)	(0.00144)
Any LSR since 2015/7 to last month	−0.0118***	−0.0201***	−0.0350***	−0.0154***
	(0.00148)	(0.00109)	(0.00419)	(0.00137)
% of Any nearby VSR w/in 0.3-mile radius	−0.00254	−0.00237	−0.00150	0.000663
	(0.00400)	(0.00228)	(0.00350)	(0.00335)
R^2	0.429	0.509	0.576	0.535
Dependent variable		log(*price*)	log(*price*)	log(*price*)
Any VSR since 2015/7 to last month		−0.0110***	−0.00413	−0.0110***
		(0.00189)	(0.00659)	(0.00221)
Any LSR since 2015/7 to last month		−0.0105***	−0.00223	−0.0116***
		(0.00179)	(0.00617)	(0.00210)
% of Any nearby VSR w/in 0.3-mile radius		−0.00176	0.00306	−0.00786*
		(0.00325)	(0.00563)	(0.00453)
R^2		0.945	0.933	0.940
log(*lagged crimes*)	Absorbed	Absorbed	Absorbed	Absorbed
Property identification FE	Yes	Yes	Yes	Yes
Zip code-year-month FE	Yes	Yes	Yes	Yes
Observations	2,866,238	2,441,566	1,370,655	1,495,583

Notes. This table explores mechanisms behind the baseline results in columns (5) and (6) in panel A of Table 2. The whole sample in column (1) consists of all Airbnb listings from January 2015 to December 2019 in the five sample cities. Columns (2), (3), and (4) use subsamples. The occupancy and price regressions in the same column use the same subsample. All regressions control zip code-year-month fixed effects (FEs); property identification FEs; and listing attributes, including number of reviews, star ratings, whether the listing is a superhost, whether the listing is crosslisted on Airbnb and VRBO, whether the listing offers a strict cancellation policy, and the number of Airbnb listings in the same zip code. Standard errors (in parentheses) are clustered by property identification. Standard errors are in parentheses.

***$p < 0.01$.

safety reviews (both VSRs and LSRs) when listings have a lower number of reviews.

Indeed, columns (3) and (4) in Table 4 report that in the subsample of listings with 13 or fewer reviews, the negative effects of having any VSRs and LSRs on occupancy rate (1.64% for VSRs and 3.5% for LSRs) are higher than the corresponding negative effects for listings with more than 13 reviews (0.764% for VSRs and 1.54% for LSRs). When the dependent variable is listings' log price, the coefficients of *VSR* and *LSR* remain negative in both subsamples, but they are of a larger magnitude and more significant for listings with more than 13 reviews, possibly because hosts of newer listings may still be in the process of identifying their pricing for those listings. The results are similar if we add additional controls for the average word count or average sentiment of a listing's review.[38]

Heterogeneous Effects. Table 5 reports the baseline results for high-income, low-income, white, and minority (M) zip codes separately. Although having any VSRs or LSRs has negative effects on occupancy rate and price across all four subsamples, this negative effect tends to have a slightly higher magnitude in H and W zip codes than in L and M zip codes. One potential explanation is that guests may have different prior beliefs and different sensitivities to safety issues, perhaps more so if their search targets a specific area that they believe is relatively safe. The coefficient of *VSRADIUS* is statistically zero except for high-income zip codes in the occupancy regression.

Table A8 in Online Appendix A consider subsamples comprising different listing types (entire home, private room, shared space, and hotel room). Additional heterogeneous effects may arise here because for instance, for guests who seek nonentire dwellings (private room or shared space) within an accommodation, safety issues may be more salient. Results in Table A8 in Online Appendix A confirm this prior; the magnitudes of the negative effects from having any VSRs and LSRs on occupancy are larger for private rooms and shared spaces (1.58% and 2.28% for VSRs and 2.83% and 2.35% for LSRs, respectively) in comparison with entire-home listings (1.34% for VSRs and 2.19% for LSRs).

5.2. Guest-Level Analysis

The baseline results demonstrate a robust negative within-listing-crossbuyer effect of VSRs and LSRs on listing price and occupancy but do not capture the

Table 5. Reduced-Form Listing-Level Analysis of Airbnb Listings by Four Area Types

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	H	L	W	M	H	L	W	M
Dependent variable	log *occupancy rate*	log *occupancy rate*	log *occupancy rate*	log *occupancy rate*	log(*price*)	log(*price*)	log(*price*)	log(*price*)
Any VSR since 2015/7 to last month	−0.0163***	−0.0137***	−0.0150***	−0.0142***	−0.0165***	−0.0136***	−0.0159***	−0.0135***
	(0.00249)	(0.00162)	(0.00209)	(0.00178)	(0.00368)	(0.00254)	(0.00313)	(0.00281)
Any LSR since 2015/7 to last month	−0.0250***	−0.0231***	−0.0233***	−0.0251***	−0.0173***	−0.0120***	−0.0174***	−0.0105***
	(0.00189)	(0.00180)	(0.00172)	(0.00199)	(0.00271)	(0.00292)	(0.00258)	(0.00315)
% Any nearby VSR w/in 0.3-mile radius	−0.00705**	−0.00123	−0.00471	−0.00069	0.00197	−0.00537	0.00486	−0.00681
	(0.00346)	(0.00328)	(0.00366)	(0.00315)	(0.00556)	(0.00510)	(0.00562)	(0.00503)
log(*lagged crimes*)	Absorbed	Absorbed	Absorbed	Absorbed	Absorbed	Absorbed	Absorbed	Absorbed
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^2	0.558	0.575	0.557	0.581	0.923	0.926	0.921	0.927

Notes. This table reports heterogeneous effects behind the baseline results in columns (5) and (6) in panel A of Table 2. The whole sample consists of all Airbnb listings from July 2015 to December 2019 in the five sample cities. All regressions control zip code-year-month fixed effects; property identification fixed effects; and listing attributes, including number of reviews, star ratings, whether the listing is a superhost, whether the listing is crosslisted on Airbnb and VRBO, whether the listing offers a strict cancellation policy, and the number of Airbnb listings in the same zip code. Standard errors (in parentheses) are clustered by property identification.

**$p < 0.05$; ***$p < 0.01$.

crosslisting-within-buyer effects of safety reviews because the baseline regressions track listings but not guests. To capture the within-buyer effects, we need to track individual guests over time. In particular, we need guest-level analysis to test whether guests who leave any VSRs (henceforth, VS guests) or any LSRs (henceforth, LS guests) act differently before and after they post their first VSR or LSR in comparison with otherwise similar guests who did not leave any VSRs or LSRs. Because VSRs and LSRs are rare, we conduct the analysis for them separately.

Guest-Level VSR Analysis. 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 that we consider, with reviews beginning in May 2015) is the first VSR that this guest posted. Guests who never posted any VSRs, referred to as "non-VS" users, are the potential control group for VS users. To ensure that we can match VS and non-VS users in their Airbnb experience prior to leaving any VSR, we focus on the subset of VS users who left at least two reviews in the five sample cities before leaving their first VSR.

In order to match VS users with non-VS users, we use propensity score matching with *K*-nearest neighbor (KNN) to select the two most similar non-VS users for each VS-user. Note that our setting is different from the typical propensity score matching scenario for two reasons; (a) the treatment (when a VS user wrote her first VSR) is staggered at different calendar times, and (b) the starting time of each guest in our data (when a user started to write her first review on Airbnb) can differ greatly. The recent econometrics literature reviewed by Roth et al. (2023) has provided a few new ways to address (a), but they usually require a balanced sample in which researchers can observe both treated and control units from the same beginning and end periods, whereas treated units may get treated at different times. In our case, a balanced sample is difficult to achieve because of (b).

One potential way to address this challenge is only matching VS and non-VS guests up to a common calendar time t_0 (e.g., prior to any VSRs appearing in our sample of VS guests). This leads to a compromise in matching quality because a VS guest who wrote her first VSR at t_i may end up matching with a non-VS user who differs significantly from her between t_i and t_0, although they look identical up to t_0. Another potential solution is lining up every VS guest's treatment time (time of writing their first VSR) as zero and randomly assigning time 0 for every non-VS guest. This way, we may have a good match for the VS guest's historical experience up to time 0, but the matched nonguest could have a seemingly similar experience from a very different calendar time. Given that Airbnb expanded quickly throughout the United States in our sample period, the public's general expectations of price and quality from Airbnb listings may change drastically over time; thus, the mismatch in calendar time is not ideal either.

In light of these challenges, we conduct our propensity score matching for each cohort of VS guests separately. In particular, we group VS guests who wrote their first VSR in the same year-month as the same cohort. For all cohort-*k* VS guests who wrote their first VSR in month t_k^0, we can compute their average attributes up to t_k^0; for all non-VS guests, we also compute their average attributes up to t_k^0. This gives us a snapshot of VS guests and non-VS guests as of t_k^0. For this snapshot, we regress the dummy of a user being a VS guest on a list of user attributes up to t_k^0. The pretreatment user attributes that we use include the number of reservations that the user made on Airbnb, 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. This procedure gives us two nearest non-VS guests for each VS guest in cohort k. Repeating this process for all cohorts of VS guests,[39] we identify 2,252 VS users and 4,504 matched non-VS users. Table A6 in Online Appendix A reports that the VS users and their matched non-VS 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 A12 in Online Appendix A.

Following each VS user and their matched non-VS users over time (by the reviews that they wrote on Airbnb), our panel data include which users are paired, their user identifications, and the time and attributes of the listings that they book on Airbnb. To 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 run the following DID regression:

$$\begin{aligned} y_{it} = \alpha_t + \alpha_p + \beta_1 \cdot VS\_user_i + \beta_2 \times post\_1st\_VSR_{pt} \\ + \beta_3 \cdot VS\_user_i \times post\_1st\_VSR_{pt} + \epsilon_{i,t}, \end{aligned} \quad (4)$$

where the subscript p denotes the treatment-control pair identified in the sample construction and the dummy $post\_1st\_VSR$ indicates whether t is after the time of the first VSR of the VS user herself or the VS user with whom the non-VS user is matched.

We construct several measures for the dependent variable y_{it}. 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 and thus, captures both the extensive and intensive margins. Because it is a count variable, we use a Poisson regression instead of ordinary least squares. The other measures of y_{it} include the normalized cumulative count of officially reported crimes in the zip codes of user i reserved listings in month t, the number of VSRs in i reserved listings, the percentage of VS listings in the zip codes as well as in the 0.3-radius area of the i reserved listings, and whether the reserved listings have any VSRs. These variables capture the types of listings that i books on Airbnb conditional on her booking at all (the intensive margin). The dummy VS_user_i equals one for VS users and zero otherwise, and the dummy $post_t$ equals one if t is after the time of the first VSR of VS user i. The key variable is the interaction between VS_user_i and $post\_1st\_VSR_{pt}$ in year-month fixed effects α_t. Treatment-control pairs fixed effects are denoted by α_p. Standard errors are robust and clustered by treatment-control pairs.

Column (1) in panel A of Table 6 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.[40] Columns (2)–(6) in panel A of Table 6 assess whether VS users are more sensitive to safety information when booking subsequent Airbnb listings after posting their first VSR. Results suggest that the subsequent listings chosen by VS users tend to locate in zip codes that have fewer normalized crime reports, are less likely to have VSRs, and are less likely to locate 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 note that our DID specification assumes that every matched non-VS user has no VS experience in their Airbnb stays. This assumption may not hold if some guests experienced VS issues but chose not to mention it in consumer feedback. Given the fact that 44.56% of Airbnb stays in our sample result in a consumer review and buyers tend to underreport critical feedback on the internet, the DID effect reported in Table 6 may understate the true effect. In particular, if the fraction of those encountering a VS issue and writing about it in VSRs is x and the probability of writing any review after a stay (regardless of the nature of experience) is δ, then the fraction of having a VS experience (regardless of writing about it or not) would be x/δ. This implies that in the "control" group of non-VS users, only a fraction of $(1 - x/\delta)/(1 - x)$ are true non-VS users. If the true treatment effect of having a VS experience is β, then our DID estimate $\hat{\beta}$ in Table 6 captures the difference between β and $\beta \times (1 - ((1 - x/\delta)/(1 - x)))$; hence, the true effect $\beta = \hat{\beta} \times (1 - x)/(1 - x/\delta)$. In our data, VS listings account for 8% of Airbnb bookings, and the average probability of writing any review after a stay is 44.56%, implying $x = 8\%$, $\delta = 0.4456$, and thus, $\beta = \hat{\beta} \times 1.1213$. In other words, the DID estimates may underestimate the true effect by roughly 12%.

One may argue that the extent of learning through self-experience would depend on a guest's prior about vicinity safety. Unfortunately, we have no data on each guest's hometown and therefore, cannot approximate their prior with the type of vicinity in which they normally live. Nevertheless, some VS users may have seen

Table 6. Reduced-Form Guest-Level Analysis: DID for VS Users (Treated) and Non-VS Users (Control)

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Monthly reservation	Reserved property	Reserved property	Reserved property	Reserved property	Reserved property
Model	Poisson	Poisson	Logit	OLS	OLS	OLS
Dependent variable	*# reservations in a month*	*# VSR in booked listing*	*1 if any VSR in booked listing*	*Crime in booked zip*	*% VS listing in booked zip*	*% VS listing in 0.3m radius*
Panel A: Full sample						
VS_user × post	−0.918***	−0.697***	−0.490***	−0.927***	−0.0250***	−0.0247***
	(0.0601)	(0.135)	(0.113)	(0.112)	(0.00267)	(0.00505)
Observations	254,056	22,265	22,237	22,415	22,415	22,415
Panel B: Subsample = VS user's 1st VSR is the 1st VSR of the listing						
VS_user × post	−0.961***	−0.793***	−0.696***	−0.961***	−0.0280***	−0.0275***
	(0.0667)	(0.146)	(0.129)	(0.127)	(0.00271)	(0.00551)
Observations	202,262	17,743	17,726	17,893	17,893	17,893
Panel C: Subsample = VS user's 1st VSR is not the 1st VSR of the listing						
VS_user × post	−0.726***	−0.372	0.256	−0.710***	−0.00872	−0.00854
	(0.139)	(0.298)	(0.239)	(0.228)	(0.00838)	(0.0129)
Observations	51,794	4,522	4,511	4,522	4,522	4,522

Notes. This table presents the DID results of VS users and non-VS users who are similar to the VS users in user attributes and Airbnb history before the VS user posts her first VSR. Standard errors are in parentheses. All regressions control treatment-control pair identification fixed effects and the post dummy. Standard errors are clustered by pair identification.
***$p < 0.01$.

some VSRs left by a listing's prior guests, and that listing eventually triggered their own VSR; therefore, they would not respond as vigorously to their own vicinity safety experience as other VS users. To test this, we create a dummy (First-Is-First) indicating whether a VS user's own VSR was the first VSR on the focal listing. About 79.6% of VS users have First-Is-First = 1. We then rerun the DID analysis for the subsamples of First-Is-First = 1 and First-Is-First = 0, respectively. Each subsample includes the VS users with the specific value of First-Is-First and their matched normal users. Regression results are reported in panels B and C of Table 6. If the above prediction is correct, the VS users with First-Is-First = 1 should demonstrate greater changes after their own VS experience as compared with those with First-Is-First = 0.

Indeed, the coefficients reported in panel B of Table 6 (for First-Is-First = 1) are of a larger magnitude than those in panel C of Table 6 (for First-Is-First = 0). The estimates in panel C of Table 6 are noisier and sometimes insignificant, in part because only 20.4% of VS users may have seen prior VSRs on the focal listing before posting their own VSR. That being said, even these VS users demonstrate a strong decline of Airbnb bookings after their own VS experience (−51.62% in column (1) in Table 6) as compared with −61.75% for VS users with First-Is-First = 1 and −60.07% for all VS users. These numbers are not statistically different from each other, suggesting that the VSRs left on the focal listings before our VS users' own VS experience have a limited influence on their prior of vicinity safety before booking the focal listing and that one's own VS experience is a still a salient shock ex post. This points to a significant crosslisting-within-buyer effect of VSRs.

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.

From the interaction term in panel A of Table 7, it is apparent that VS users exhibit a positive effect on subsequent reservations in opposite types of zip codes (columns (2) and (4) in panel A of Table 7) and a negative effect in the same type of zip codes (columns (1) and (3) in panel A of Table 7). One explanation is that VS users expect a certain level of safety in the area of their booking, and when they encounter a negative shock, they prefer to avoid that type of area in subsequent stays.

One may argue that the tendency to avoid the same type of areas is driven by mean reversion rather than active learning. To address this, we repeat the exercise for the subsamples with First-Is-First = 1 and First-Is-First = 0 separately. Results are reported in panels B and C of Table 7. Three of the four columns (columns (2)–(4) in panels B and C of Table 7) are consistent with the argument that learning through self-experience is stronger when the VS user did not see any other VSRs on the focal listing before her own VSR. The only exception is when the self VSR is in a high-income zip code (column (1) in panels B and C of Table 7). In that case,

Table 7. Reduced-Form Guest-Level Analysis: DID for VS Users by the Zip Code of Their VSR Bookings

	(1)	(2)	(3)	(4)
Sample	1st_vsr_h_zip	1st_vsr_l_zip	1st_vsr_w_zip	1st_vsr_m_zip
Model	Logit	Logit	Logit	Logit
Dependent variable	*1 if book in any H zip*	*1 if book in any H zip*	*1 if book in any W zip*	*1 if book in any W zip*
Panel A: Full sample				
VS_user × post	−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
Panel B: Subsample = VS user's 1st VSR is the 1st VSR of the listing				
VS_user × post	−0.287*	0.370***	−0.646***	0.729***
	(0.169)	(0.111)	(0.149)	(0.117)
Observations	5,539	11,377	7,181	10,113
Panel C: Subsample = VS user's 1st VSR is not the 1st VSR of the listing				
VS_user × post	−0.887*	0.143	−0.545*	0.494**
	(0.496)	(0.218)	(0.327)	(0.247)
Observations	666	3,453	1,699	2,702

Notes. This table presents the DID results of VS users vs. non-VS users who are similar to the VS users in user attributes and Airbnb history before the VS user posts her first VSR. Standard errors are in parentheses. All regressions control treatment-control pair identification fixed effects and the post dummy, with standard errors clustered by pair identification. Columns (1)–(4) use the subsample corresponding to the VS users whose first VSR is posted on a property listing in an H, L, W, or M area, respectively.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

both VS users of First-Is-First equal to one or zero decrease their likelihood of booking future Airbnb stays in high-income zip codes (which is consistent with learning), but the coefficient on the DID interaction term is of a larger magnitude for those with First-Is-First = 0, although the difference is not statistically significant. Compared with the other columns, column (1) in panels B and C of Table 7 has less statistical power because VSRs are rarer in high-income zip codes. Overall, we conclude that the tendency to avoid the type of zip code that triggered VS users' own VSR is partially driven by learning from one's own VS experience.

To push it further, we reorganize our DID sample into another eight subsamples depending on whether a VS user previously had Airbnb stays in the same type of area that triggered her own VS experience. For example, if her own VS experience was in a low-income area, she may or may not have had Airbnb stays in low-income areas before. This gives us the subsamples of HL, LL, LH, and HH, where the second letter indicates the income type of the area that triggered the VS user's own VSR and the first letter represents the income-type area of her prior experience. Similarly, we can create the subsamples of WM, MM, MW, and WW depending on whether the area is primarily white or minority. All matched normal users belong to the same subsample as the VS users with whom they are paired.

Results are reported in Tables A9 and A10 in Online Appendix A. In the raw data, we know that VSRs are more likely to occur in low-income and minority areas, but listings in these areas also account for over 60% of all Airbnb bookings; thus, the sample sizes of LH and LL are larger than those of HL and HH, and the sample sizes of MW and MM are larger than those of WM and WW. If we focus on column (1) in Table A9 in Online Appendix A, the VS users in LH are the most "surprised" (in terms of reducing future reservations on Airbnb) among the four L/H groups, and the VS users in MW are the most surprised among the four M/W groups. This is intuitive because the VS users with at least one L or M stay before their own VS experience in H or W may have high-vicinity-safety expectations in H or W and are consequently most disappointed when vicinity safety issues arise in those areas. In contrast, the VS users in HL or WM are not as surprised (column (1) in Table A9 in Online Appendix A), likely because they had a lower prior for vicinity safety in the L or M areas. Nevertheless, conditional on booking on Airbnb, they tend to book listings with fewer VSRs following their own VSR. These patterns confirm the crosslisting-within-buyer effect of self-experience with vicinity safety.

Guest-Level Analysis for LS Guests. For comparison, we repeat the same DID + matching procedure for guests who ever wrote LSRs on Airbnb (LS users) and those who never wrote LSRs (non-LS users), and we rerun a similar DID regression on the panel data that track the Airbnb activities of LS users and their matched non-LS users:

$$y_{it} = \alpha_t + \alpha_p + \beta_1 \cdot LS\_user_i + \beta_2 \cdot post\_1st\_LSR_{pt} + \beta_3 \cdot LS\_user_i \times post\_1st\_LSR_{pt} + \epsilon_{i,t}, \quad (5)$$

where the subscript p denotes the treatment-control pair identified in the sample construction and the dummy $post\_1st\_LSR$ indicates whether t is after the time of the first LSR of the LS user herself or the LS user with whom the non-LS user is matched. We construct the dependent variables the same way as in Equation (4).

Results in Table 8 suggest that LSRs trigger some crosslisting-within-buyer effect as well; having experienced and written about LSRs in an Airbnb listing makes the LS user 48.21% less likely to book on Airbnb afterward,[41] and conditional on having future bookings, the LS user is less likely to book listings with any LSRs (columns (2) and (3) in Table 8), listings in a zip code that has a higher percentage of LS listings, or listings that are in a 0.3-mile radius of any listings with LSRs.

It is worth noting that these effects are stronger for VS users than for LS users on both the extensive margin of not making any subsequent Airbnb booking after self-experience (−60.07% for VS users versus −48.21% for LS users) and the intensive margin of shying away from the listings that have the same type of safety reviews that the treated user has written about herself (−38.74% for VS users versus −28.39% for LS users).[42] The extensive margin results are further confirmed in the event study plot (Figure A14 in Online Appendix A). Combined with the fact that LSRs tend to have a greater within-listing-crossbuyer effect than VSRs as shown in Table 2, the larger crosslisting-within-buyer effects of VSRs relative to LSRs imply that guests may receive a bigger surprise from a vicinity safety experience than from a listing safety experience and therefore, react more strongly to the negative shock. It is also possible that guests believe that LSRs can be addressed by hosts, and therefore, they can find non-LSR listings on Airbnb, but VSRs describe a problem out of the host's control and harder to avoid on Airbnb.

6. Structural Estimation and Counterfactual Analysis

So far, the reduced-form evidence supports (i) the classical within-listing-crossbuyer effect of VSRs and LSRs as listing performance worsens after a listing receives its first VSR or LSR and (ii) the crosslisting-within-buyer effect of VSRs or LSRs as a guest who wrote VSRs or LSRs tends to avoid other listings/vicinities with any VSRs or LSRs in their future bookings or avoid booking on Airbnb altogether. Interestingly, the former is stronger for LSRs than for VSRs, but the latter is stronger for VSRs than for LSRs, suggesting that self-experience in VSR generates a greater negative surprise to guests. In comparison, the spillover from a listing's VSR to nearby listings is weak and hard to identify from other local shocks at the zip code level.

It is difficult to use these reduced-form estimates to understand the implications of the externalities for hosts and platforms because they do not address listing competition. In particular, listings with and without

Table 8. Reduced-Form Guest-Level Analysis: Compare DID Results for VS Users and LS Users

Sample Model	(1) Monthly reservation Poission	(2) Reserved property Poission	(3) Reserved property Logit	(4) Reserved property OLS	(5) Reserved property OLS	(6) Reserved property OLS
Panel A: Full sample results for VS user DID						
Dependent variable	*# reservations in a month*	*# VSR in booked listing*	*1 if any VSR in booked listing*	*Crime in booked zip*	*% VS listing in booked zip*	*% VS listing in 0.3m radius*
VS user × post	−0.918***	−0.697***	−0.490***	−0.927***	−0.0250***	−0.0247***
	(0.0601)	(0.135)	(0.113)	(0.112)	(0.00267)	(0.00505)
Observations	254,056	22,265	22,237	22,415	22,415	22,415
Panel B: Full sample results for LS user DID						
Dependent variable	*# reservations in a month*	*# LSR in booked listing*	*1 if any LSR in booked listing*	*Crime in booked zip*	*% LS listing in booked zip*	*% LS listing in 0.3m radius*
LS user × post	−0.658***	−0.269***	−0.334***	−0.671***	−0.0117***	−0.0238***
	(0.0516)	(0.103)	(0.109)	(0.112)	(0.000983)	(0.00277)
Observations	288,072	21,113	25,981	26,915	22,629	22,629

Notes. This table compares the DID results for VS users and LS users. The panel A sample consists of VS users and non-VS users who are similar to the VS users in user attributes and Airbnb history before the VS user posts her first VSR. The panel B sample consists of LS users and non-LS users who are similar to the LS users in user attributes and Airbnb history before the LS user posts her first LSR. Standard errors are in parentheses. All regressions control treatment-control pair identification fixed effects and the post dummy, with standard errors clustered by pair identification.

***$p < 0.01$.

VSRs/LSRs may compete against each other on Airbnb, and all Airbnb listings compete with the outside options (including listings on competing short-term-rental platforms, hotels, bed and breakfasts, a friend's or relative's couch in the destination city, or no travel at all). To address this shortcoming, we resort to a structural model that describes how guests choose among competing short-term lodging options.

6.1. Demand Model and Estimation

Possible Market Definitions. We have explored several ways of defining the short-term lodging market. Option (a) is limited to Airbnb and VRBO bookings in a zip code-month. Option (b) includes all hotels, Airbnb, and VRBO stays in a zip code-month. Option (c) includes hotels, Airbnb, and VRBO stays across all zip codes in a city-month. Option (d) includes all incoming air travelers in a city-month.

Among the four options, (a) is the narrowest because it assumes that a guest has a specific zip code in mind when she searches for short-term lodging and that there is no competition between hotels and STR listings. This can be problematic, not only because hotels compete with STR listings but also, because guests who are concerned about vicinity safety with regard to Airbnb listings may have similar concerns for nearby VRBO listings (if they can overcome the information friction across the two platforms to figure out what Airbnb listings and VRBO listings are geographically close). However, adding hotels to the market at the zip code-month level is also problematic because most zip codes do not have any hotels. Expanding the market from zip code-month to city-month can get around this problem, but it assumes that any potential guests would consider all zip codes in a city. This consideration set might be larger than what most guests actually consider, calling for model parameters that address different substitution patterns within and across zip codes in the same city.

As we explore the above four options of market definition, we focus on entire-home listings on STR platforms because only entire-home listings are available on VRBO, and the few hotel listings on Airbnb suggest that entire-home listings are much more similar to hotel listings than private-room or shared-space listings (Table A3 in Online Appendix A). Because our VRBO data period is from June 2017 to December 2019, our analysis in this subsection is limited to June 2017 to December 2019 only.

Guest Utility. Under any of the four market definitions, following Berry (1994) and Mansley et al. (2019), we assume that each prospective guest chooses an Airbnb entire-home listing or the outside good so as to maximize her utility from the listing, where the utility associated with listing j in zip code z of city k and month t can be written as

$$U_{j,t} = EU_{j,t} + \epsilon_{j,t} = \alpha_j + \alpha_{k,t} + \delta \cdot X_{j,t} + \beta_0 \cdot \log(P_{j,t}) + \beta_1 \cdot Crime_{z,t-1} + \beta_2 \cdot LSR_{j,t-1} + \beta_3 \cdot VSR_{j,t-1} + \beta_4 \cdot VSRADIUS_{j,t-1} + \zeta_j^{city} + (1-\sigma_{city})\zeta_j^{zip} + (1-\sigma_{zip})(1-\sigma_{city})\epsilon_{j,t}, \tag{6}$$

where $X_{j,t}$ stands for listing attributes,[43] $\alpha_{k,t}$ represents some area-time fixed effects,[44] and $\epsilon_{j,t}$ conforms to independent and identically distributed extreme value distribution. $\zeta_j^{city}, \zeta_j^{zip}$ follow the unique distributions such that $[\zeta_j^{city} + (1-\sigma_{city})\zeta_j^{zip} + (1-\sigma_{zip})\ (1-\sigma_{city})\epsilon_{j,t}]$ describes a two-level nested logit error. As shown in Figure 2, if the market definition is city-month, a guest first chooses between Airbnb and the outside good and then, within Airbnb, chooses among different zip codes before selecting an Airbnb listing in a specific zip code.

The nesting parameter $0 < \sigma_{city} < 1$ describes how Airbnb listings in different zip codes are closer substitutes to each other than the substitution between Airbnb and the outside good, and the nesting parameter $0 < \sigma_{zip} < 1$ describes how Airbnb listings in the same zip codes are closer substitutes to each other than the substitution between Airbnb listings across zip codes. When $\sigma_{zip} = \sigma_{city}$, the two-level nesting structure collapses to one nest (Airbnb versus the outside good); when $\sigma_{city} = 0$, it further collapses to a simple logit where the outside good is equivalent to another single listing available in the market. If the market definition is zip code-month rather than city-month, we can only have a one-nest structure.

Under this nesting structure, we can express the market share of listing j at time t as[45]

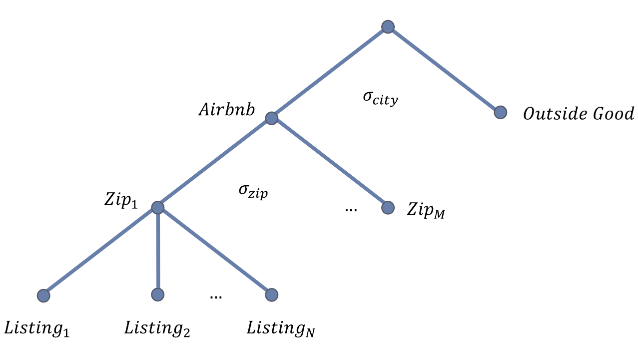
$$s_{j,t} = \bar{s}_{j,t|zip_z} \cdot \bar{s}_{zip_z|Airbnb} \cdot \bar{s}_{Airbnb}.$$

Thus,

$$\ln(s_{j,t}) - \ln(s_{0,t}) = EU_{j,t} + \sigma_{zip} \cdot \log(\bar{s}_{j,t|zip_z}) + \sigma_{city} \cdot \log(\bar{s}_{zip_z|city}). \tag{7}$$

This is equivalent to regressing the difference of log market shares between listing j and the outside good

Figure 2. (Color online) Two-Level Nested Demand for Short-Term Lodging



$(ln(s_{j,t}) - ln(s_{0,t}))$ on the attributes of listing j in month t plus its within-zip-code market share and the within-city market share of j's zip code. The right-hand side of Equation (7) is similar to Equation (1), except for three 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 daily price. Third, we include a listing's within-zip code market share and the zip code's within-Airbnb market share; these two within-market shares are endogenous, and we use the number of Airbnb listings within the corresponding zip code-month interacting with average listing attributes in that zip code-month as an instrument for $\log(\bar{s}_{j,t|zip_z})$ and the number of zip codes in a city interacting with the average Airbnb listing attributes in the city as instruments for $\log(\bar{s}_{zip_z|city})$.

As $\log(P)$ might be endogenous, we consider three instruments. First, following Berry et al. (1995), the average attributes of entire-home listings within a 0.3-mile radius of the focal listing in the same zip code-month can be instruments as they 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; we refer to them as BLP (Berry–Levinsohn–Pakes) instruments hereafter. Second, the average price of private-room listings within a 0.3-mile radius of the focal listing in the same zip code-month can be potential instruments under the assumption that guests of private-room listings are fundamentally different from guests of entire-home listings and hotels, but both types of listings are subject to similar cost shocks in the same location. We refer to them as private room instruments. Third, Zillow's Home Value Index by zip code-month can capture local property taxes, house maintenance costs, and nonrental usage of the property. Usually, short-term rental units only account for a tiny fraction of the housing stock in a city,[46] so the potential impact of short-term rental activities on ZHVI should be minimal. Because ZHVI is zip code specific but not property specific, we interact it with a listing's basic attributes (number of bedrooms and star ratings) to construct property specific instruments. We refer to them as $ZHVI \cdot x$ instruments.

Table 9 reports the results for 12 combinations of different market definitions and different instruments for listing price. In panel A of Table 9, we adopt a narrow market definition of zip code-month and use BLP instruments for listing price, but we explore whether to include hotels in the market and whether to allow one-level nesting (Airbnb versus the outside good). When VRBO is the only outside good, the nesting parameter is found to be 1.15, out of the regular range of 0–1. But, when the outside good consists of hotels and VRBO listings, we find a more reasonable nesting parameter (0.25), which suggests that substitution within Airbnb listings is closer than the substitution between Airbnb and hotel + VRBO in the same zip code-month. This is not surprising because the volume of hotel stays is much larger than Airbnb and VRBO listings if they are available in the zip code.

However, most zip codes do not have hotels; thus, in panel B of Table 9, we expand the market to city-month. We try two market size definitions (hotel + Airbnb + VRBO and the total count of incoming air travelers) while still using BLP instruments for listing price. The two market size definitions are highly correlated, but the count of incoming air travelers is five to nine times larger than the total count of hotel, Airbnb, and VRBO stays in a city-month. For each of them, we try one-level and two-level nesting models for comparison. Results between these two market size definitions are mostly similar except that we have difficulty identifying significant nesting parameters if we require the model to have two nesting parameters when the market is defined as all incoming air travelers. When the market is defined as hotel + Airbnb + VRBO, the zip code nesting parameter (σ_{zip}) is slightly smaller (0.219) than the city nest parameter (σ_{city}, 0.238). These estimates suggest that Airbnb listings are closer substitutes to each other than hotels and other short-term lodging options, and listings within the same zip codes are closer substitutes than across zip codes. Note that although both panels A and B of Table 9 use BLP instruments for listing price, the coefficient on price drops dramatically from somewhere between -5 and -9 in panel A of Table 9, where the market is defined as zip code-month, to between -1 and -2 in panel B of Table 9, where the market is defined as city-month. This suggests that BLP instruments may be good at capturing guest sensitivity to price within a zip code but not good at capturing it citywide.

In panel C of Table 9, we keep the market definition at the city-month level with market size defined as hotel + Airbnb + VRBO[47] but try ZHVI and private room instruments for the listing price. As in panel B of Table 9, we report results for one-level and two-level nesting for comparison. The nesting parameters are similar to what we find in panel B of Table 9, but the price coefficient differs. When we use ZHVI instruments, the price coefficient (-2.582) is more negative than using the BLP instruments in panel B of Table 9, suggesting that the ZHVI instruments can capture more price sensitivity. When we use private room instruments, the price coefficient is much smaller (around -1.4) and similar to that of panel B of Table 9.

In all 12 specifications, we find a consistent pattern for the coefficients of *VSR*, *LSR*, and *VSRADIUS*. Similar to our reduced-form results in Table 2, coefficients of *VSR* and *LSR* are negative and significant with 99% confidence, and the coefficient of *LSR* is consistently larger in magnitude than that of *VSR*, suggesting that guests perceive worse utility from a listing after it receives a safety review, especially if the safety review is about the property itself. At the same time, we continue to observe an insignificant

Table 9. Specification Test for the Structural Model

	(1)	(2)	(3)	(4)
Panel A: Market = zip code-month, BLP IV for listing price				
Market size	Airbnb + VRBO		Hotel + Airbnb + VRBO	
Model	Logit	Nlogit	Logit	Nlogit
log(*price*)	−7.762***	−5.744***	−9.163***	−8.724***
	(1.185)	(1.174)	(1.315)	(1.316)
nesting parameter (zip)		1.150***		0.250***
		(0.0354)		(0.0389)
Any VSR	−0.142***	−0.0705***	−0.152***	−0.137***
	(0.0159)	(0.0159)	(0.0186)	(0.0187)
Any LSR	−0.210***	−0.0913***	−0.230***	−0.204***
	(0.0226)	(0.0226)	(0.0256)	(0.0259)
% Any nearby VSR w/in 0.3m radius	−0.190***	−0.229***	−0.306***	−0.315***
	(0.0601)	(0.0598)	(0.0812)	(0.0812)
1st stage F-stat for P	293.8	293.8	293.9	293.9
1st stage F-stat for nesting par. (zip)		131.8		131.8
Property identification FE	Yes	Yes	Yes	Yes
City-year-month FE	Yes	Yes	Yes	Yes
Observations	921,092	921,092	921,092	921,092
R^2	0.755	0.756	0.937	0.937
Panel B: Market = city-month, BLP IV for listing price				
Market size	Hotel + Airbnb + VRBO		Incoming air travelers	
Model	N1Logit	N2logit	N1Logit	N2logit
IV for price	*BLP*	*BLP*	*BLP*	*BLP*
log(*price*)	−1.833***	−1.492***	−1.538**	−1.378**
	(0.550)	(0.551)	(0.549)	(0.548)
nesting parameter (city)	0.371***	0.238***	0.146***	−0.0527
	(0.0194)	(0.0613)	(0.0193)	(0.0611)
nesting parameter (zip)		0.219***		0.0310
		(0.0230)		(0.0229)
Any VSR	−0.0457***	−0.0506***	−0.0464***	−0.0497***
	(0.00908)	(0.00915)	(0.00907)	(0.00912)
Any LSR	−0.0831***	−0.0894***	−0.0931***	−0.0995***
	(0.0120)	(0.0121)	(0.0120)	(0.0121)
% Any nearby VSR w/in 0.3m radius	−0.0122	−0.00225	0.0426	0.0489
	(0.0363)	(0.0363)	(0.0361)	(0.0361)
1st stage F-stat for p	490.8	490.8	490.8	490.8
1st stage F-stat for nesting par. (city)	250.5	164.3	250.5	164.3
1st stage F-stat for nesting par. (zip)		216.5		216.5
Property identification FE	Yes	Yes	Yes	Yes
City-month (1–12) FE + year-month FE	Yes	Yes	Yes	Yes
Observations	921,092	921,092	921,092	921,092
R^2	0.681	0.680	0.646	0.646
Panel C: Market = city-month, other IV for listing price				
Market size	Hotel + Airbnb + VRBO		Hotel + Airbnb + VRBO	
Model	N1Logit	N2Logit	N1Logit	N2Logit
IV for price	$ZHVI \cdot x$	$ZHVI \cdot x$	*PrivRoom P*	*PrivRoom P*
log(*price*)	−8.911***	−2.582***	−1.439***	−1.429***
	(0.933)	(0.894)	(0.221)	(0.224)
nesting parameter (city)	0.444***	0.231***	0.360***	0.197***
	(0.0202)	(0.0614)	(0.0194)	(0.0615)
nesting parameter (zip)		0.225***		0.194***
		(0.0231)		(0.0233)

Table 9. (Continued)

Panel C: Market = city-month, other IV for listing price				
Market size	Hotel + Airbnb + VRBO		Hotel + Airbnb + VRBO	
Model	N1Logit	N2Logit	N1Logit	N2Logit
Any VSR	−0.106***	−0.0599***	−0.0427***	−0.0511***
	(0.0111)	(0.0110)	(0.00784)	(0.00791)
Any LSR	−0.199***	−0.108***	−0.0772***	−0.0903***
	(0.0172)	(0.0170)	(0.00828)	(0.00843)
% Any nearby VSR w/in 0.3m radius	−0.124***	−0.0193	−0.00530	−0.00014
	(0.0381)	(0.0378)	(0.0356)	(0.0356)
1st stage F-stat for p	502	502	516.2	516.2
1st stage F-stat for nesting par. (city)	250.5	164.3	250.5	164.3
1st state F-stat for nesting par. (zip)		216.5		216.5
Property identification FE	Yes	Yes	Yes	Yes
City-month (1–12) FE + year-month FE	Yes	Yes	Yes	Yes
Observations	921,092	921,092	920,815	920,815
R^2	0.681	0.680	0.681	0.681

Notes. This table reports the structural estimates following Equation (7). Standard errors are in parentheses. All regressions control for property identification fixed effects (FEs), listing attributes, and log of cumulative crime reports since July 2015 to last month in the zip code of the listing. Within-city and within-zip market shares (for the nesting parameters) are instrumented by the number of listings within inside goods or the number of zip codes in a city × average listing attributes. BLP, Berry–Levinsohn–Pakes; IV, instrumental variables; Nlogit, one-level nested logit model; N2logit, Two-Level Nested Logit Model.

$^{**}p < 0.05$; $^{***}p < 0.01$.

coefficient of *VSRADIUS*, echoing the mixed effect of VSRs on nearby listings in the reduced-form analysis at the listing level.

Given the robustness, from now on, we will use column (2) in panel C of Table 9—where market definition is city-month, market size is hotel + Airbnb + VRBO, and $ZHIV \cdot x$ are instruments for listing price—as our preferred structural model estimation for counterfactual analysis.

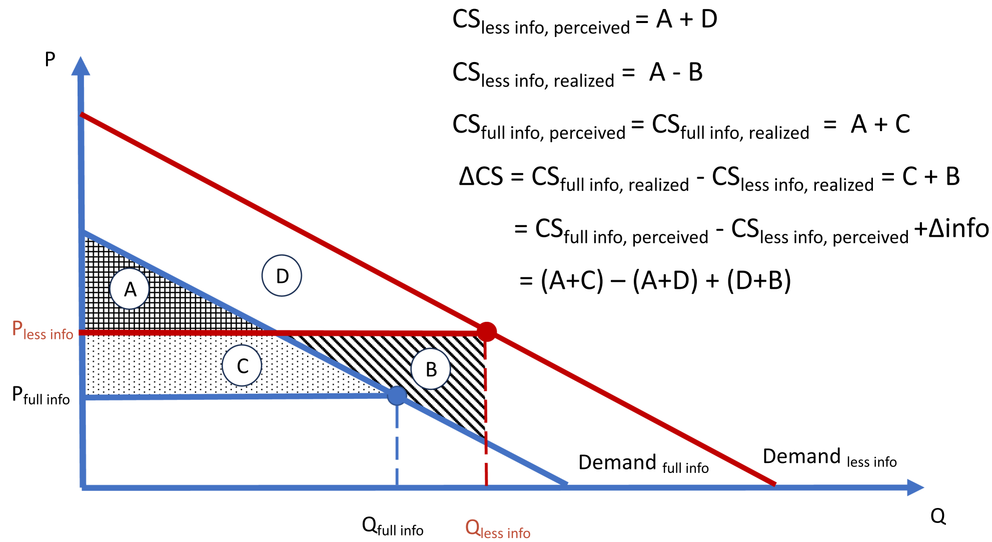
Note that the coefficient of *VSR* captures how an *average* prospective guest in our sample *perceives* the vicinity safety of listing i at the time of choice. Because VSRs only account for 0.25% of all guest reviews, the vast majority of the guests may have not experienced any vicinity safety issues on Airbnb (or have experienced but never reported them in a user review) before t. Indeed, if we rerun Equation (7) excluding the VS users identified in our reduced-form analysis, the estimated coefficients barely change. This means that Equation (7) can yield reliable estimates for the within-listing-crossbuyer effects but not the crosslisting-within-buyer effect driven by VS users learning from their own VS experience.

We have also considered including an interaction of the dummy of VS users with *VSR*. Aside from the aforementioned sample size issue, this interaction would add endogeneity to the specification because we do not observe whether a listing is booked by a VS user until the user has booked and left reviews for that listing on Airbnb. Because we cannot observe the situations where a VS user considers Airbnb listings but decides to not book on Airbnb, including this interaction will not tell us how self-experience of VSRs has changed the VS user's expected utility from Airbnb listings.

Fortunately, such learning from self-experience has already been captured in our reduced-form guest-level analysis; thus, a key question is how to incorporate the DID estimate into the structural framework. This is important not only because this extra crosslisting-within-buyer effect is in addition to the within-listing-crossbuyer effect that we can identify directly from the vast majority of Airbnb guests but also, because self-experience sheds light on the guest's *realized* utility when she stays in a listing with vicinity safety issues. Although such realized utility, as indicated by guest reviews, only occurred in a tiny fraction of Airbnb stays, a fully informed guest should expect the realized utility when she subsequently chooses where to stay. As documented by Jin and Sorensen (2006), Allcott (2011), Train (2015), and Reimers and Waldfogel (2021), the difference between realized utility and perceived utility is essential for evaluating how consumer surplus changes under different information regimes.

In particular, Figure 3 illustrates the role of perceived and realized utility in consumer surplus. Consider two demand curves; the inner one represents demand for Airbnb listings under a high-alert regime of VSRs, whereas the outer one represents demand under a regime with less information about VSRs. When prospective guests perceive the listings to be safer than they actually are, the market clears at a higher price and with more bookings than under high alert (i.e., $P_{\text{less info}} > P_{\text{full info}}$ and $Q_{\text{less info}} > Q_{\text{full info}}$). Those who book under less information comprise two guest types. Some have a high willingness to pay and would have booked on Airbnb even if they knew the high alert ex ante, with their realized

Figure 3. (Color online) Consumer Surplus Under High Alert of VSR (Realized) or Less Information of VSR (Perceived)



consumer surplus being area A; others have a relatively low willingness to pay and would not book on Airbnb had they known that the listings are actually less safe than they appear, with their realized consumer surplus being negative (area B). Hence, the total realized consumer surplus is A − B under the less-information regime. In comparison, under the high-alert regime, the realized consumer surplus is A + C, where C represents the extra consumer surplus that fully informed guests could obtain via a lower equilibrium price.

There is another way to arrive at the same conclusion. Let us denote the white trapezoid between the two demand curves in Figure 3 as area D. Under the less-information regime, the perceived consumer surplus is A + D, whereas the realized consumer surplus is A − B; under a full-information regime, both the perceived consumer surplus and realized consumer surplus are A + C. Thus, the difference between the realized consumer surplus under full- and less-information regimes, (A + C) − (A − B) = C + B, can also be written as the difference between their perceived consumer surplus plus an adjustment that reflects the shift of the demand curve for all consumers who would purchase under less information, namely (A + C) − (A + D) + (D + B) = C + B. We will use this alternative expression to compute consumer surplus changes in the counterfactual analysis.

As shown below, our counterfactual analysis assumes that the listing choices made by VS users *after* they wrote VSRs on Airbnb reflect their updated, subsequent interpretation of VSRs on all potential listings. Because this updated interpretation incorporates their true experience in the stay that triggered their VSR, we assume that it captures the realized utility from VSRs. This means that VS users would have a new β_3 in the utility function upon their own VS experience. Changes in their β_3 would capture the difference between perceived and realized utility driven by VSRs.

We calibrate a new β_3 that would explain why the number of Airbnb bookings of VS users dropped 60.07% after their own VS experiences as compared with similar non-VS users according to our guest-level DID analysis (column (1) in Table 6). Following the procedure described in Online Appendix B, our calibration suggests that the VS users must have changed their β_3 by −2.195, which is more than 35 times the estimated β_3 of the whole sample (−0.0599). If we consider the possibility that non-VS users in the control group of the DID analysis may include some true VS users that chose not to write any VSRs, the DID estimate (60.07%) would be underestimated. According to our calculation, the degree of underestimation depends on the degree of underreporting. Because the probability of leaving a review after any Airbnb stay is 44.56% in our sample, the calculation suggests that the true effect is 60.07% · 1.1213 = 67.36%, and the calibrated $\Delta\beta_3$ would be even bigger (−2.274). Overall, the calibration suggests that the crosslisting-within-buyer effect of VSRs—based on a guest's own VS experience—is strong and would have an impact on booking decisions should all non-VS users interpret VSRs in the same way as VS users.

Arguably, the same process may apply to the self-experience of LSRs as well. Because the reduced-form evidence suggests that the self-experience of VSRs is a greater negative shock to the guest than self-experience of LSRs, the counterfactual analysis below focuses on VSRs only for the ease of illustration.

6.2. Counterfactual Analysis

We consider four counterfactual regimes as compared with the status quo. The first is "no disclosure no belief update," where all VSRs are removed from the data and no guests update their belief of VSR risk despite the removals. Conceptually, this is equivalent to an extreme interpretation and implementation of Airbnb's

December 2019 policy on VSRs, where all VSRs are removed and all guests view all listings as if they have never received any VSRs.

One may argue that guests may change their belief about a listing's VSR risk if they know that no VSRs would ever be disclosed. To accommodate this possibility, we explore a second no-disclosure regime where Airbnb removes all VSRs but all guests form a rational expectation of a listing's VSR risk conditional on the listing's observable attributes. This regime of "no disclosure but with rational belief" can occur if the announcement of the no-disclosure policy is salient and if all prospective guests fully understand the statistical correlation between VSR risk and other observable listing attributes in the raw data.

The two no-disclosure regimes differ in information treatment. Under "no disclosure no belief update," VSR information on VS listings changes from some VSRs to zero VSRs, whereas non-VS listings remain clean of VSRs. Under "no disclosure but with belief update," VS listings look less risky than in the status quo, but all normal listings now look as risky as VS listings with similar attributes (rather than of zero VSR risk). This amounts to a positive information shock to VS listings *and* a negative information shock to non-VS listings. The positive information shock to VS listings is less in "no disclosure but with rational belief" than in "no disclosure no belief update" because by definition, rational guests should have predicted some probability of having VSRs for VS listings.

A priori, it is unclear which of the two no-disclosure regimes is closer to reality. "No disclosure no belief update" could occur if the platform quietly removes all VSRs without the notice of most customers (even if the platform makes a public announcement of the policy change). In that case, most guests may interpret that the list did not receive any VSRs in the past rather than the platform did not report any VSRs. By contrast, "no disclosure but with rational belief" could occur if the platform's public announcement of the no-disclosure policy is widely disseminated and if guests are fully aware of the statistical relationship between historical VSR occurrences and listing attributes. Given the facts that the platform does not have full incentives to broadcast a no-disclosure policy, not all potential guests pay close attention to every platform announcement, and it is rare for an average guest to have the same access to the comprehensive listing-month data as in this paper, we believe that the reality is likely somewhere between the two no-disclosure regimes. We report both to help readers understand their differences.

An extreme regime in contrast to no disclosure is "high alert," where we assume that all users react to VSRs as much as VS users react to their own reported VSRs. Comparing with the above three regimes, the fourth counterfactual regime keeps the information policy as is but removes listings with 1+ or 2+ VSRs. This is different from no disclosure because it removes VS listings from guests' choice set, whereas VS listings are kept alive and appear similar to non-VS listings (in the lack of VSRs) in the no-disclosure regimes. This "listing removal" counterfactual aims to mimic a change in the platform's listing screening policy rather than information policy.

We now elaborate on how we calculate consumer welfare under each counterfactual regime. For the status quo, we use the results in column (2) in panel C of Table 9 to calculate $EU_{j,t}$ for each Airbnb listing-month, and then, we use the price coefficient to normalize it to U.S. dollars. By definition, this is the guest's perceived utility. Following Small and Rosen (1981) and Mcfadden (2001), in a nested logit model such as ours, a consumer's expected utility from her utility-maximizing choice depends on the inclusive value of the choice set, namely $I = \log(1 + \exp(I_{Airbnb}))$, where $I_{Airbnb} = (1 - \sigma_{city}) \log\sum_{z\in city}\exp((I_{zip_z})/(1 - \sigma_{city}))$ is the Airbnb-specific inclusive value and $I_{zip_z} = (1 - \sigma_{zip}) \cdot \log\sum_{j\in zip_z}\exp((EU_{j,t})/(1 - \sigma_{zip}))$ is the zip code-specific inclusive value. As depicted in Figure 3, a consumer's perceived utility may guide her choice of listing ex ante, but her realized utility may deviate from her perceived utility. To measure the realized utility, we use the calibrated change of β_3 (-2.274 as described above) to update β_3 in the utility function (while taking everything else unchanged) and recompute the utility.

For the counterfactual of "no disclosure no belief update," we set all VSRs as zero in the (perceived) utility function, recompute $EU_{i,t}$ for each Airbnb entire-home listing, and simulate its market share. This assumes that everything else remains the same when the platform removes all VSRs, which could be violated if listings adjust prices after the regime shift. Because the vast majority of our data precede Airbnb's new review policy and Airbnb seems far from fully implementing the policy, we cannot observe such price adjustments directly. The reduced-form regressions in Table 2 associate the presence of VSRs in VS listings with a 1.47% difference in price. Hence, in an alternative calculation, we assume that the no-disclosure regime may enable a 1% price increase in VS listings, whereas the price of normal listings remains unchanged. This gives us a comparison between no disclosure with price changes versus no disclosure without price changes.

To implement the "no disclosure but with rational belief" counterfactual, we use a logit specification to regress the dummy of any VSRs observed in our data (at the listing-month level) on observable listing attributes up to the month before, including whether the listing has any LSRs, local crime statistics (cumulative by zip code), the listing's total number of reviews on Airbnb, the listing's average star ratings, whether the listing has a superhost status, whether the listing is crosslisted on VRBO, whether the listing has a strict cancellation policy, year-month fixed effects, and city-month fixed effects. We then use the estimated coefficients to predict the likelihood of having any VSRs per listing-month. Replacing the VSR dummy in the utility equation with this predicted VSR probability, we compute the subsequent market shares, consumer surplus, and Airbnb GBV.

As in the first no-disclosure regime, we run the second no-disclosure counterfactual with and without price changes. In particular, we assume that the price may be adjusted down for normal listings by predicted probability of VSRs × 1% because they look riskier in this counterfactual than in the status quo, and the price will be adjusted up for VS listings by (1 − predicted probability of VSRs) × 1% because they look less risky in this counterfactual than in the status quo.

The high-alert counterfactual is equivalent to assuming that guests have full information, and therefore, their perceived utility is the same as the realized utility calculated above. In other words, all guests use the calibrated β_3 (based on the DID results from VS users) in the utility of each listing for both perceived and realized utility. As for the listing removal regime, we use each listing's utility as in the status quo but remove listings with 1+ or 2+ VSRs from the guest's choice set. As in the no-disclosure regimes, in "high-alert" and "listing-removal" regimes, we first simulate market shares without price changes and then, introduce an ad hoc price change (−1% for VS listings) to illustrate how price changes may alleviate the impact of the counterfactual regime.

After we compute the perceived and realized utility under each regime, we can quantify changes in consumer surplus from the status quo to any counterfactual. Defining each city-month (k, t) as a unique market, our analysis includes 145 markets in total. We further define market size $M_{k,t}$ as the total reserved days in the market (hotel + Airbnb + VRBO). Following Reimers and Waldfogel (2021) and Figure 3, the consumer surplus changes in a single market from the status quo to the high-alert regime can be computed as

$$\Delta CS_{k,t} = \frac{M_{k,t}}{\beta_0} \cdot \Bigg[\ln(I \,|\, highalert) - \ln(I \,|\, perceived) + \sum_j ((U_{jt,perceived} - U_{jt,highalert}) \cdot s_j) \Bigg]. \quad (8)$$

Similar calculations are performed for other counterfactual regimes.

Table 10 reports the consumer surplus results under different counterfactuals for an average user with an average reservation day across all city-months. The first two rows in Table 10 refer to "no disclosure no belief update" with and without price changes. Rows 3 and 4 in Table 10 refer to "no disclosure but with rational belief" with and without price changes. Rows 5–8 in Table 10 refer to high alert with and without price changes and with and without a "radius effect," where the radius effect allows for the same updated distaste for VSRs to apply to the VSRs in nearby listings as well. To quantify the radius effect, we assume that the estimated coefficient of *VSRADIUS* (β_4) would increase in the same proportion as the calibrated coefficient of *VSR* (β_3) should guests extrapolate the high alert of vicinity safety concerns to nearby VSRs in the same way as a listing's own VSRs. This extreme regime is designed to illustrate the potential consequences in case prospective guests become sensitive to *any* VSRs under high alert. The last two rows in Table 10 refer to removing listings with 1+ or 2+ VSRs.

Table 10 indicates that under high alert without price changes and without a radius effect, overall consumer surplus under high alert (without a radius effect) increases by roughly 9.756%, which is slightly less if we incorporate the hypothetical 1% price drop of VS listings (9.599%) and slightly higher if we allow a radius effect in high alert (10.340% and 10.183%) because high alert helps guests reduce their stays in relatively unsafe listings.

Consumer surplus under no-disclosure regimes declines in comparison with the status quo. Under "no disclosure no belief update," consumer surplus may decline by 1.183% without price change (and 0.676% with price changes) because consumers cannot use VSRs to sort among Airbnb listings. Under "no disclosure but with rational belief," consumer surplus is still less than the status quo, but the decline is smaller (by 0.993% without price change and 0.571% with price changes). This is as expected because rational belief would associate all Airbnb listings with an average belief of VSR risk conditional on observable listing attributes. Consequently, the positive information shock on VS listings is less than in the regime without belief update, and the negative information shock on non-VS listings further alerts consumers of average VSR risk. In both no-disclosure regimes, the ad hoc 1% price adjustment can mitigate the loss in consumer surplus but is not enough to eliminate it. Interestingly, removing listings with 1+ or 2+ VSRs would generate a bigger decline of consumer surplus (−1.187% to −5.008%) than no-disclosure regimes because they narrow consumers' choice set. The second column in Table 10 reports bootstrapped standard errors for the consumer surplus changes. To compute standard errors, we redraw 99% of the data at the zip code-year-month level 100 times, rerun the counterfactual analysis for each redrawn sample, and report the standard deviation of counterfactual estimates.

These changes in consumer surplus are driven by changes in consumer beliefs, which in turn, shift market shares. Under "no disclosure no belief update," the lack of VSR information moves market share from hotels, VRBO, and normal Airbnb listings to VS listings. In

Table 10. Counterfactual Analysis: Simulated Changes in Consumer Surplus in the Market

Δ*Consumer surplus (vs. status quo)*	All listings estimate (%)	All listings standard error (%)
No disclosure no belief change w/o P change	−1.183	(0.043)
No disclosure no belief change w/ P change	−0.676	(0.012)
No disclosure w/ rational belief w/o P change	−0.993	(0.037)
No disclosure w/ rational belief w/ P change	−0.571	(0.010)
High alert w/o P change & w/o radius effect	9.756	(0.019)
High alert w/ P change & w/o radius effect	9.599	(0.023)
High alert w/o P change & w/ radius effect	10.340	(0.157)
High alert w/ P change & w/ radius effect	10.183	(0.148)
Remove listings with any VSR	−5.008	(0.033)
Remove listings with 2+ VSR	−1.187	(0.008)

Notes. Counterfactual simulations are based on (i) structural estimates from column (2) in panel C of Table 9, where market is city-year-month, market size is hotels + Airbnb + VRBO, and IV for price is $ZHVI \cdot x$, and (ii) a calibrated VSR coefficient based on the DID + matching of VS users (column (1) in panel A of Table 6). Bootstrapped standard errors are in parentheses. To compute standard errors, we redraw 99% of the data at the zip code-year-month level 100 times, rerun the counterfactual analysis for each redrawn sample, and report the standard deviation of counterfactual estimates. P, price.

comparison, "no disclosure but with rational belief" also introduces a negative information shock to non-VS listings, thus moving market share toward VS listings and away from non-VS listings. In comparison, the dramatic high-alert counterfactual would move almost all market shares away from VS listings. By definition, removing listings with one or more VSRs would eliminate VS listings' market share, whereas removing listings with two or more VSRs only reduces the market share of VS listings modestly because most VSR listings have only one VSR.

Table 11 reports GBV changes based on simulated market shares in each regime. "No disclosure no belief update" generates 0.327% more GBV for entire-home listings on Airbnb in our sample if no price changes or 0.285% more GBV if assuming that the price for VS listings increases by 1%. This suggests that the platform could have strategic incentives to hide VSRs if the no-disclosure regime can be implemented quietly without much consumer notice. However, "no disclosure but with rational belief" would *decrease* Airbnb's GBV by 0.047% if no price changes or increase the GBV by 0.013% if VS and non-VS listings may adjust their prices up to 1% according to changes in consumer belief. This suggests that consumers' rational belief based on observable listing attributes could mitigate the platform's incentive to hide VSRs. Price changes in response to the information changes can soften the effects on platform GBV or even overturn the direction of the GBV effects. We caution that the assumed 1% price adjustment is not necessarily the equilibrium change as we do not observe hosts' costs and do not model how hosts set their prices in reality. Rather, it points to the possibility that price changes can play an important role determining the platform's overall incentives in disclosing VSRs.

Compared with the no-disclosure regimes, high alert generates substantial GBV loss for the platform, ranging from −2.726% to −6.026% depending on whether we incorporate 1% price change and the radius effect. In short, under high alert and the "no disclosure no belief update," the interests of Airbnb and consumers are misaligned; consumers would prefer more transparency, but a GBV-centric Airbnb would prefer no disclosure without consumer belief update.

Interestingly, the interests of guests and the platform are aligned on listing removal; both would suffer from the removal of (all or some) listings with VSRs because it narrows consumers' choice set. The interests of guests and the platform are also partially aligned in the regime of "no disclosure but with rational belief," especially when price adjustment is small or nonexisting.

One caveat of all above counterfactual calculation is that we focus on consumers' static choice of short-term lodging but do not account for the fact that the status quo may decay over time without any change of platform policy because consumers burned by self-experience in VSRs would become high alert organically, even if everyone else with no such self-experience continues to hold her perception of VSRs as observed in our data. Because 0.8% of consumers would choose VS listings on Airbnb in the status quo, this means that every month, 0.8% of the not-yet-alerted consumers may become alerted by self-experience in VSR, and thus, the market at any time would reflect a mixture of the static status quo and the high alert as simulated in Tables 10 and 11.

The lower dashed line for CS and upper dashed line for GBV in the left panel of Figure 4 show this organic decay process in 25 years (300 months on the horizontal axis) under the status quo, assuming that consumers can only update their understanding of the real impact of VSRs based on self-experience. By the end of the third year, 25.7% of consumers would have booked any VS listings and get a self-experience of VSRs; this percentage increases to 38.2% by the 5th year and 61.9% by the 10th year. In the meantime, consumer surplus grows slowly by <6%, and Airbnb GBV drops slowly to

Table 11. Counterfactual Analysis: Simulated Market Shares and Changes in GBV

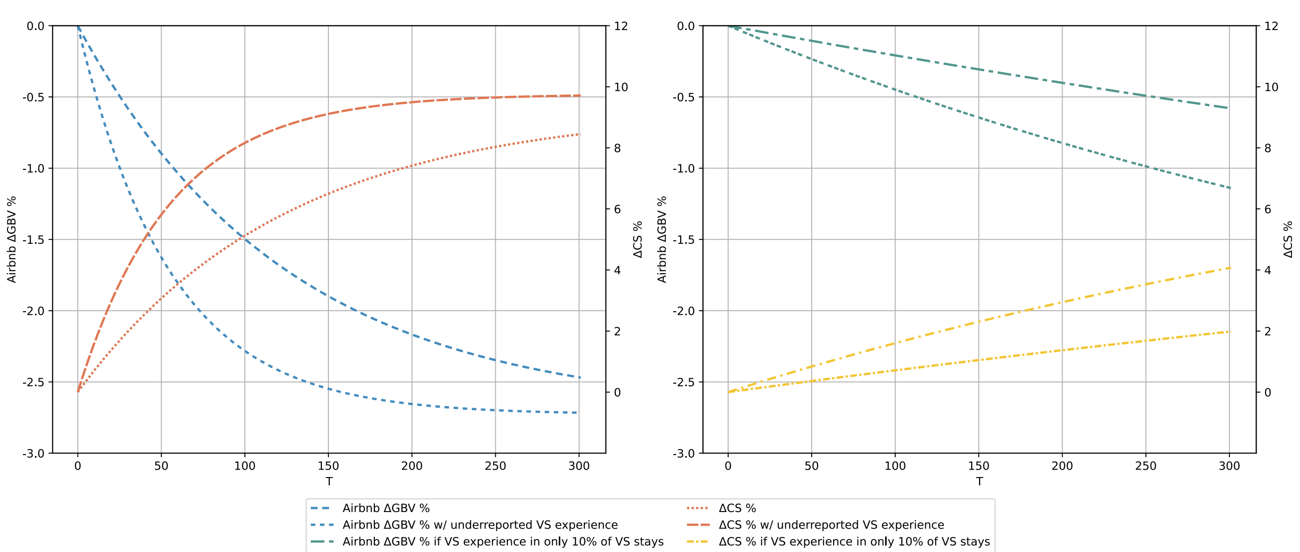
ΔGBV *(vs. status quo)*	Airbnb VS listings (%)	Airbnb non-VS listings (%)	Airbnb listings (%)	Hotel + VRBO listings (%)	All listings (%)
No disclosure no belief update w/o P change	8.052	−0.129	0.327	−0.058	−0.026
	(0.296)	(0.016)	(0.028)	(0.004)	(0.001)
No disclosure no belief update w/ P change	5.689	−0.034	0.285	−0.038	−0.011
	(0.081)	(0.021)	(0.022)	(0.002)	(0.001)
No disclosure w/ rational belief w/o P change	6.637	−0.442	−0.047	−0.014	−0.016
	(0.250)	(0.014)	(0.019)	(0.003)	(0.001)
No disclosure w/ rational belief w/ P change	4.715	−0.265	0.013	−0.012	−0.010
	(0.071)	(0.023)	(0.023)	(0.003)	(0.001)
High alert w/o P change w/o radius effect	−94.520	2.699	−2.728	0.563	0.291
	(0.091)	(0.037)	(0.030)	(0.003)	(0.001)
High alert w/ P change w/o radius effect	−94.390	2.693	−2.726	0.562	0.290
	(0.093)	(0.037)	(0.029)	(0.003)	(0.001)
High alert w/o P change w/ radius effect	−95.029	−0.764	−6.026	0.923	0.350
	(0.132)	(0.676)	(0.645)	(0.070)	(0.012)
High alert w/ P change w/ radius effect	−94.911	−0.769	−6.024	0.922	0.349
	(0.127)	(0.676)	(0.645)	(0.070)	(0.012)
Remove listings with any VSR	−100	2.859	−2.883	0.595	0.308
	(0.000)	(0.036)	(0.034)	(0.003)	(0.001)
Remove listings with 2+ VSR	−39.523	0.723	−1.523	0.150	0.012
	(0.066)	(0.010)	(0.010)	(0.001)	(0.000)

Notes. Counterfactual simulations are based on (i) structural estimates from column (2) in panel C of Table 9, where market is city-year-month, market size is hotels + Airbnb + VRBO, and IV for price is $ZHVI \cdot x$, and (ii) a calibrated VSR coefficient based on the DID + matching of VS users (column (1) in panel A of Table 6). Bootstrapped standard errors are in parentheses. To compute standard errors, we redraw 99% of the data at the zip code-year-month level 100 times, rerun the counterfactual analysis for each redrawn sample, and report the standard deviation of counterfactual estimates. P, price.

about 1.7% by the end of the 10th year. This is less and slower than the high-alert counterfactual (which generates >9% increase in consumer surplus and >5% drop in GBV). This contrast highlights the importance of the platform information policy when negative self-experience is rare but strong as we have seen in the case of VSRs.

To further illustrate how the organic evolution of the status quo depends on the extent of VSR experience, we add three alternative decay processes in Figure 4.

Figure 4. (Color online) Potential Evolution of the Status Quo Because of Self-Experience of VSR Only



Notes. This graph simulates how consumer surplus and Airbnb GBV change over time as guests learn about VSR via self-experiences over time. The simulation is based on results in Tables 10 and 11. CS, consumer surplus.

The first alternative—depicted by the upper dashed line for CS and lower dashed line for GBV in the left panel of Figure 4—incorporates the fact that only 44.56% of Airbnb guests write any review after a stay in our data. This implies that there may be more listings with VS experience than the number of VS listings that we can find in the data. Assuming that the actual number of VS listings shall be increased from x to $x/44.56\%$, the market share of actual VS listings would be $0.8\%/44.56\% = 1.8\%$ rather than 0.8%. As shown by the upper dashed line for CS and lower dashed line for GBV in the left panel of Figure 4, this change would make more consumers aware of VS and speed up the convergence.

The second alternative relaxes the assumption that every guest staying in a VS listing must encounter a VS experience. In our data, VS listings account for 8% of stays within Airbnb, but VSRs only account for 0.25% of all reviews. If review rate does not depend on the nature of experience, this implies that only $0.25\%/8\% = 3.1\%$ of stays in VS listings have resulted in a VS experience worth reporting in VSRs. From the literature, we know that negative experience may suffer from more underreporting than positive experience, but it is difficult to pin down the extent of this difference. Hence, for illustration purposes, we simulate an alternative process in the right panel of Figure 4, assuming that 10% of stays in VS listings generate a negative VS experience and that only such experiences would motivate the guest to update her coefficient of VSRs in the utility function. This amounts to only $0.8\% \times 10\% = 0.08\%$ of all not-yet-burnt short-term rental consumers who would get an update in each month. Consequently, the market evolution is much slower, with only 4.7% of consumers ever burnt by a VS experience by the end of the fifth year. As shown by the lower dashed line for CS and upper dashed line for GBV in the right panel of Figure 4, the changes in consumer surplus and platform GBV are much slower in this alternative process than what happens if we assume that VSR experience always occurs in any stays of VS listings (the left panel of Figure 4).

The third alternative includes the assumptions in both the first and second alternatives, namely VS experience is underreported by 44.56% but only 10% of VS stays trigger a VS experience. The results of this alternative process are displayed in the upper dashed line for CS and lower dashed line for GBV in the right panel of Figure 4. Again, incorporating underreporting in our organic evolution would speed up the information

Table 12. Counterfactual Analysis: Changes in GBV by Four Area Types

	Airbnb listings (%)			
ΔGBV (vs. status quo)	H	L	W	M
No disclosure w/o P change	0.158 (0.023)	0.593 (0.037)	0.241 (0.025)	0.535 (0.035)
No disclosure w/ P change	0.167 (0.021)	0.471 (0.022)	0.225 (0.021)	0.431 (0.022)
No disclosure w/ rational belief w/o P change	−0.193 (0.016)	0.183 (0.026)	−0.122 (0.018)	0.133 (0.024)
No disclosure w/ rational belief w/ P change	−0.086 (0.023)	0.167 (0.023)	−0.041 (0.023)	0.142 (0.023)
High alert w/o P change w/o radius effect	−0.715 (0.032)	−5.881 (0.033)	−1.697 (0.032)	−5.195 (0.028)
High alert w/ P change w/o radius effect	−0.715 (0.032)	−5.875 (0.032)	−1.696 (0.032)	−5.189 (0.028)
High alert w/o P change w/ radius effect	−4.074 (0.655)	−9.083 (0.627)	−5.026 (0.650)	−8.418 (0.631)
High alert w/ P change w/ radius effect	−4.074 (0.655)	−9.077 (0.627)	−5.025 (0.650)	−8.413 (0.630)
Remove listings with any VSR	−0.753 (0.035)	−6.219 (0.040)	−1.792 (0.035)	−5.493 (0.034)
Remove listings with 2+ VSR	−0.492 (0.009)	−3.138 (0.015)	−1.061 (0.011)	−2.628 (0.011)

Notes. Counterfactual simulations are based on (i) structural estimates from column (2) in panel C of Table 9, where market is city-year-month, market size is hotels + Airbnb + VRBO, and IV for price is $ZHVI \cdot x$, and (ii) a calibrated VSR coefficient based on the DID + matching of VS users (column (1) in panel A of Table 6). Bootstrapped standard errors are in parentheses. To compute standard errors, we redraw 99% of the data at the zip code-year-month level 100 times, rerun the counterfactual analysis for each redrawn sample, and report the standard deviation of counterfactual estimates. P, price.

update from self-experience and subsequent changes in consumer surplus and platform GBV, everything else being equal.

To explore the distributional effects of our counterfactual regimes, Table 12 breaks down the counterfactual GBV changes in Airbnb listings by the four types of zip codes. Because VS listings are more likely to locate in low-income and minority zip codes, both no-disclosure regimes benefit Airbnb listings in these zip codes. However, the regime of "no disclosure but with rational belief" can hurt listings in high-income and white zip codes, especially if price adjustment is not deep enough to counter the negative information shock to them. In both high-alert and listing removal regimes, low-income and minority areas would suffer from a bigger drop in GBV, either by consumer's informed choice or by removing some VS listings from consumer's choice set.

7. Conclusion

Taking safety reviews as an example of critical feedback on Airbnb, we show that vicinity safety reviews and listing safety reviews not only have the classical within-listing-crossbuyer effect of guiding future buyers toward listings without VSRs/LSRs, but they also motivate guests that have written VSRs/LSRs themselves to learn and update their understanding of the VSRs/LSRs of other listings. As a result, these guests are less likely to book future stays on Airbnb, and when they do book, they tend to book listings without VSRs/LSRs and in areas with fewer official crime reports and fewer VSRs/LSRs. More interestingly, such crosslisting-within-buyer effect is stronger for VSRs than for LSRs, although the classical within-listing-crossbuyer effect is greater for LSRs than for VSRs, suggesting that self-experience of VSRs is a greater negative shock for guests.

Using a structural approach to account for listing competition on and off Airbnb, we show that a revenue-centric platform may prefer to limit the disclosure of VSRs altogether, even though the aggregate surplus of guests appears to increase when the VSRs are instead emphasized to alert prospective guests. However, this strategic incentive to hide VSRs can be mitigated or even overturned if consumers can form rational beliefs about VSR risks (conditional on observable listing attributes) after the platform announces a no-disclosure policy. In that case, although no disclosure prevents consumers from distinguishing seemingly identical VS and non-VS listings, it generates a negative information shock on non-VS listings, which discourages consumers from booking non-VS listings and thus, could reduce the overall GBV of the platform. Put another way, consumers' rational beliefs under a no-disclosure regime help to align the interests of consumers and the platform. In comparison, removing listings with VSRs may hurt both consumer surplus and platform GBV because it narrows consumers' choice set.

Combined, our findings highlight the economic incentives and tensions behind a platform's information policy regarding critical feedback. For a rare but strong negative experience like VSR, allowing VSRs but not highlighting them on the platform may slowly decay guest trust via the organic within-buyer-crosslisting effect, resulting in a slow decline of the platform's general booking value and a slow increase of consumer surplus as guests learn from self-experience. The platform can hasten this process by adopting a more transparent information policy to warn consumers of the risks. Although doing so may lead to a significant GBV loss for the platform according to our calculations, it may be worthwhile for the platform if more transparency can boost user trust and attract sufficiently many new users to join the platform over the long run.

Another managerial implication from our work is the distributional effects of information policy. Under the high-alert regime, we show that listings in low-income and minority zip codes may stand to lose a disproportionate share of revenue relative to their counterparts in high-income and white zip codes, but consumer surplus under the high-alert regime is higher than under the status quo and the no-disclosure regimes. The platform thus faces a trade-off as far as generating greater 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, which may affect reviews that mention vicinity safety, our findings suggest that the policy, if fully implemented without rational belief updates on the consumer side, can have some unintended consequences for consumers and listings without VSR. How to balance the economic interests of all users is a challenge for platforms as well as for policymakers who 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 VSRs for each listing. Unfortunately, crime statistics (when available) may not fully capture all of the safety concerns that a guest may have in mind at the time of booking. Another alternative is to incorporate VSRs into the overall ratings of a listing, and how to adjust ratings in line with the platform's or a social planner's objectives certainly merits future research.

There are a number of limitations to our analyses. First, guest reviews in our data do not include potential responses from hosts. Second, in the guest-level analysis, we only observe guests' reservation provided that they have made any Airbnb reservations in the five major U.S. cities that we consider and posted a review on Airbnb. If VS users are more vocal and thus, more likely to post subsequent reviews after their first VSR, then our findings underestimate the negative effects on their subsequent booking activities; if, however, VS users are less likely to post subsequent reviews, then our findings overestimate the effects. In our analysis, we attempt to adjust for the potential underestimation by relying on the overall review rate observed in our data (44.56%), but this adjustment does not incorporate the possibility that underreporting might differ by the nature of guest experience. Third, our main 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, our counterfactual simulations are hypothetical.

These limitations suggest additional 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 introduction of its new review policy. In addition, 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 counterfactual analysis, a policy that encourages and highlights VSRs 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 economic parity across different neighborhoods.

Acknowledgments

The authors are grateful to AirDNA and Smith Travel Research for providing the data and to their home universities for financial support. Editor Tat Chan, an anonymous associate editor, and referees provided invaluable comments that helped the authors improve the paper significantly. Marshall Van Alstyne, Matthias Hunold, Xiang Hui, Meng Liu, Peter Coles, Francine Lafontaine, Ying Fan, Juan Pablo Atal, Juan Camilo Castillo, Sophie Calder-Wang, Yufeng Huang, Zhe Yuan, Jisu Cao, and Devesh Raval; seminar participants at the Federal Trade Commission, the Luohan Academy Webinar, Washington University at St. Louis, Boston University, the University of Michigan, and the University of Pennsylvania; and participants of the 2022 Mannheim Centre for Competition and Innovation (MaCCI) annual conference, the 2022 International Industrial Organization Conference (IIOC) annual conference, the 2022 INFORMS Marketing Science and virtual conference, and the 2023 Strategy and Economics of Digital Markets conference have provided constructive comments. Tejas Nazare, Nour Ben Ltaifa, Hunter Petrik, and Jingyi Xing 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 the authors have a financial relationship with Airbnb or competing short-term rental platforms. All errors are the authors' own.

Endnotes

[1] Recent examples of platform choice of which information to avail to users include YouTube, which has adopted a policy of hiding dislike counts on shared videos (see, e.g., https://rb.gy/xhhqnd), and Instagram, which has given users the option of hiding likes (see, e.g., https://rb.gy/tacuj5).

[2] Almost no hosts would volunteer to discuss safety in their listing descriptions because any mention (even the phrase "perfectly safe") may call guest attention to safety concerns.

[3] See reviews by Bajari and Hortacsu (2004), Einav et al. (2016), and Tadelis (2016).

[4] Airbnb's December 9, 2019 explanation about the review policy update effective December 11, 2019 can be found at https://web.archive.org/web/20200213215420/https://community.withairbnb.com/t5/Airbnb-Updates/Making-reviews-more-relevant-and-useful-for-our-community/td-p/1191576. The review policy update referred to in this explanation can be found at https://web.archive.org/web/20200615014646/https://www.airbnb.com/help/article/2673/airbnbs-review-policy.

[5] Bergemann and Morris (2019) offer a general review of information design, including but not limited to online platforms.

[6] 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.

[7] See, for example, https://rb.gy/1eohbw.

[8] See, for example, https://rb.gy/nwetrv and https://rb.gy/wrqvy4.

[9] See, for example, https://www.neighborhoodscout.com/.

[10] 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 critical 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).

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

[12] Airbnb's December 9, 2019 explanation about the review policy update effective December 11, 2019 can be found at https://web.archive.org/web/20200213215420/https://community.withairbnb.com/t5/Airbnb-Updates/Making-reviews-more-relevant-and-useful-for-our-community/td-p/1191576. The review policy update referred to in this explanation can be found at https://web.archive.org/web/20200615014646/https://www.airbnb.com/help/article/2673/airbnbs-review-policy.

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

[14] See Airbnb report "A Six-Year Update on Airbnb's Work to Fight Discrimination and Build Inclusion" (December 13, 2022) available at https://news.airbnb.com/wp-content/uploads/sites/4/2022/12/A-Six-Year-Update-on-Airbnbs-Work-to-Fight-Discrimination-and-Build-Inclusion-12122022.pdf.

[15] For example, on January 27, 2020, a tweet by user "PatrickR0820" stated: "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 [sic]) & 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).

16 See, for example, https://rb.gy/sykoim.

17 See, for example, https://web.archive.org/web/20220712202933/https://community.withairbnb.com/t5/Host-Voice/quot-Location-quot-As-A-Guest-Review-Point/idi-p/162137.

18 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.

19 Occupancy rate is calculated by dividing the number of booked nights by the sum of the available nights and booked nights.

20 Overall rating scores are normalized to the 0–10 range. Our data set also includes location star ratings. Adding them 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.

21 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 five-star reviews, and (iv) did not cancel guest reservations in the past 12 months.

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

23 Only listings of the entire home could be listed on both Airbnb and VRBO. The colisting indicator is a variable created by AirDNA; it is unclear to what extent an individual guest searching on Airbnb and VRBO can tell whether two listings are the same listing colisted on both platform because neither platform provides a precise address of a listing until the guest has booked and paid for the listing.

24 Most of the keywords appear relatively infrequently, and removing any one of them 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.

25 Although 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 one to two sentences around the keyword match.

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

27 The utilized model is a fine-tuned checkpoint of DistilBERT-base-uncased, which is accessible at https://huggingface.co/distilbert-base-uncased-finetuned-sst-2-english. It demonstrates a noteworthy accuracy of 91.3% on the development set. The sentiment scoring system ranges from −1 to 1, where a score of −1 indicates an extremely negative sentiment and a score of 1 indicates an extremely positive sentiment.

28 To save space, we omit the table of results for these alternative regressions; they are available upon request.

29 Official crime data for Chicago are at https://rb.gy/atjsss.

30 Official crime data for New Orleans are at https://rb.gy/4vue82.

31 Official crime data for New York City are at https://rb.gy/iwrwp2.

32 Official crime data for Atlanta are at https://rb.gy/96txbl.

33 Official crime data for Los Angeles are at https://rb.gy/tebnla.

34 See, for example, https://www.census.gov/data.html.

35 ZHVI data are available at https://www.zillow.com/research/data/.

36 We assume that VSRs begin with a clean slate (zero records) as of the beginning of our data set.

37 Some listing-month observations have an occupancy rate of zero and consequently, are missing an average reserved daily rate in the data set for those months, although the data set 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).

38 These results are available upon request.

39 We did the matching with replacement; thus, it is possible that the same non-VS guest is matched with two VS guests in two different cohorts. In that case, we include this non-VS guest twice in the DID sample, with a different pair identification and pseudotreatment time corresponding to the VS guest whom she matches.

40 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 $1 - e^{-.918}$).

41 Because the specification is Poisson, the marginal effect is $\exp(-0.658) - 1 = 0.4821$.

42 $\exp(-0.490) - 1 = -0.3874$, and $\exp(-0.334) - 1 = 0.2839$.

43 For VRBO listings in the outside good, we observe their X directly except for VSR and LSR because we observe no reviews on VRBO. We code their VSR and LSR as zero. For hotels in the outside good, we observe their average daily price and occupancy volume directly but not other listing attributes. Given the general difference between regular hotels and STR listings, we assume that all hotels have the highest ratings (in the Airbnb definition) and do not have a strict cancellation policy.

44 Note that the area-time fixed effects ($\alpha_{k,t}$) cannot be as detailed as the market definition as that way, the fixed effects would absorb the outside good market size and make the results independent of market definition. When we define the market as zip code-month, we use city-year-month fixed effects for $\alpha_{k,t}$. When we define the market as city-month, we use city-calendar month (1–12) and year-month fixed effects for $\alpha_{k,t}$ to control for common time effects and city-specific seasonality.

45 In particular, the within-zip-code market share is $\overline{s}_{j,t|zip_z} = \exp[(EU_{j,t}/(1-\sigma_{zip}))]/\exp[(I_{zip_z}/(1-\sigma_{zip}))]$, the zip code's within-Airbnb market share is $\overline{s}_{zip_z|Airbnb} = \exp[(I_{zip_z}/(1-\sigma_{city}))]/\exp[(I_{Airbnb}/1-\sigma_{city}))]$, Airbnb's overall market share in the city is $\overline{s}_{Airbnb} = \exp(I_{Airbnb})/\exp(I)$, the zip code-specific inclusive value is $I_{zip_z} = (1-\sigma_{zip}) \cdot \log \sum_{j \in zip_z} \exp(EU_{j,t}/(1-\sigma_{zip}))$, the Airbnb-specific inclusive value is $I_{Airbnb} = (1-\sigma_{city}) \log \sum_{z \in city} \exp(I_{zip_z}/(1-\sigma_{city}))$, and the overall inclusive value is $I = \log(1 + \exp(I_{Airbnb}))$.

46 For example, Chicago has 1.26 million housing units in total but only 5,499 Airbnb listings in an average month of our data (of which 3,420 are entire-home listings).

47 In an unreported table, we also tried a version that excludes VRBO from the outside good and use $ZHVI \cdot x$ as instruments for listing price. We find almost identical results as columns (1) and (2) in panel C of Table 9. This is reasonable because VRBO listings accounts for less than 1% of market share in a city-month.

References

Allcott H (2011) Consumers' perceptions and misperceptions of energy costs. *Amer. Econom. Rev.* 101:98–104.

Bajari P, Hortacsu A (2004) Economic insights from internet auctions. *J. Econom. Literature* 42:457–486.

Banchio M, Skrzypacz A (2022) Artificial intelligence and auction design. *Proc. 23rd ACM Conf. Econom. Comput. (EC'22)* (Association for Computing Machinery, New York), 30–31.

Barach MA, Golden JM, Horton JJ (2020) Steering in online markets: The role of platform incentives and credibility. *Management Sci.* 66:4047–4070.

Bergemann D, Morris S (2019) Information design: A unified perspective. *J. Econom. Literature* 57:44–95.

Berry S (1994) Estimating discrete choice models of product differentiation. *RAND J. Econom.* 25:242–262.

Berry S, Levinsohn J, Pakes A (1995) Automobile prices in market equilibrium. *Econometrica* 63:841–890.

Blake T, Moshary S, Sweeney K, Tadelis S (2021) Price salience and product choice. *Marketing Sci.* 40:619–636.

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

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

Conitzer V, Wagman L (2014) False-name-proof voting over two alternatives. *Internat. J. Game Theory* 43:599–618.

Conitzer V, Immorlica N, Letchford J, Munagala K, Wagman L (2010) False-name-proofness in social networks. Saberi A, ed. *Internet and Network Economics (WINE 2010)*, Lecture Notes in Computer Science, vol. 6484 (Springer, Berlin, Heidelberg), 209–221.

Dai WD, Jin G, Lee J, Luca M (2018) Aggregation of consumer ratings: An application to Yelp.com. *Quant. Marketing Econom.* 16:289–339.

Decarolis F, Rovigattie G, Rovigatti M, Shakhgildyan K (2023) Artificial intelligence & data obfuscation: Algorithmic competition in digital ad auctions. Preprint, submitted December 20, http://dx.doi.org/10.2139/ssrn.4660391.

Dhaoui C, Webster CM, Tan LP (2017) Social media sentiment analysis: Lexicon versus machine learning. *J. Consumer Marketing* 34(6):480–488.

Einav L, Farronato C, Levin J (2016) Peer-to-peer markets. *Annual Rev. Econom.* 8:615–635.

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

Fradkin A, Grewal E, Holtz D (2021) Reciprocity and unveiling in two-sided reputation systems: Evidence from an experiment on Airbnb. *Marketing Sci.* 40(6):1013–1029.

Fung A, Graham M, Weil D (2007) *Full Disclosure: The Perils and Promise of Transparency* (Cambridge University Press, Cambridge, UK).

Gandhi A, Hollenbeck B, Li Z (2025) Misinformation and mistrust: The equilibrium effects of fake reviews on Amazon.com. Working paper, UCLA Anderson School of Management, Los Angeles; Northwestern University, Evanston, IL.

Gurran N, Phibbs P (2017) When tourists move in: How should urban planners respond to Airbnb? *J. Amer. Planning Assoc.* 83:80–92.

Han W, Wang X (2019) Does home sharing impact crime rate? A tale of two cities. *Proc. 40th Internat. Conf. Inform. Systems (ICIS 2019)* (Association for Information Systems, Munich, Germany).

Han W, Wang X, Ahsen M, Wattal S (2020) Does home sharing impact crime rate? An empirical investigation. Preprint, submitted April 25, http://dx.doi.org/10.2139/ssrn.3520919.

Hui X, Klein TJ, Stahl K, (2021) When and why do buyers rate in online markets? Technical report, University of Bonn and University of Mannheim, Bonn, Germany.

Hutto CJ, Gilbert E (2014) VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proc. Internat. AAAI Conf. Weblogs Social Media (ICWSM 2014)* (AAAI Press, Ann Arbor, MI), 216–225.

Jia J, Wagman L (2020) Platform, anonymity, and illegal actors: Evidence of Whac-a-Mole enforcement from Airbnb. *J. Law Econom.* 63:729–761.

Jia J, Jin GZ, Wagman L (2021) Platform as a rule maker: Evidence from Airbnb's cancellation policies. NBER Working Paper No. 28878, National Bureau of Economic Research, Cambridge, MA.

Jin GZ, Sorensen A (2006) Information and consumer choice: The value of publicized health plan ratings. *J. Health Econom.* 25:248–275.

Jin GZ, Wagman L, Zhong M (2024) The effects of short-term rental regulation: Insight from Chicago. *Internat. J. Indust. Organ.* 96:203087.

Jüni P, Nartey L, Reichenbach S, Sterchi R, Dieppe PA, Egger M (2004) Risk of cardiovascular events and rofecoxib: Cumulative meta-analysis. *Lancet* 364:2021–2029.

Kim J-H, Leung TC, Wagman L (2017) Can restricting property use be value enhancing? Evidence from short-term rental regulation. *J. Law Econom.* 60:309–334.

Klein TJ, Lambertz C, Stahl KO (2016) Market transparency, adverse selection, and moral hazard. *J. Political Econom.* 124:1677–1713.

Klein TJ, Lambertz C, Spagnolo G, Stahl KO (2009) The actual structure of eBay's feedback mechanism and early evidence on the effects of recent changes. *Internat. J. Electronic Bus.* 7:301–320.

Kovbasyuk S, Spagnolo G (2024) Memory and markets. *Rev. Econom. Stud*. 91(3):1775–1806.

Luca M (2017) Designing online marketplaces: Trust and reputation mechanisms. *Innovation Policy Econom.* 17:77–93.

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

Maldonado-Guzmán DJ (2022) Airbnb and crime in Barcelona (Spain): Testing the relationship using a geographically weighted regression. *Ann. GIS* 28:147–160.

Mansley R, Miller N, Ryan C, Weinberg M (2019) Notes on the nested logit demand model. Working paper, Georgetown University, Washington, DC.

Mayzlin D, Dover Y, Chevalier J (2014) Promotional reviews: An empirical investigation of online review manipulation. *Amer. Econom. Rev.* 104:2421–2455.

Mcfadden D (2001) Economic choices. *Amer. Econom. Rev.* 91:351–378.

Monroe BL, Colaresi MP, Quinn KM (2017) Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Anal.* 16:372–403.

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

Nosko C, Tadelis S (2015) The limits of reputation in platform markets: An empirical analysis and field experiment. NBER Working Paper No. 20830, National Bureau of Economic Research, Cambridge, MA.

Reimers I, Waldfogel J (2021) Digitization and pre-purchase information: The causal and welfare impacts of reviews and crowd ratings. *Amer. Econom. Rev.* 111:1944–1971.

Romanyuk G, Smolin A (2019) Cream skimming and information design in matching markets. *Amer. Econom. J. Microeconomics* 11:250–276.

Roth JJ (2021) Home sharing and crime across neighborhoods: An analysis of Austin, Texas. *Criminal Justice Rev.* 46:40–52.

Roth J, Sant'anna PH, Bilinski A, Poe J (2023) What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *J. Econometrics* 235:2218–2244.

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

Small KA, Rosen HA (1981) Applied welfare economics with discrete choice models. *Econometrica* 49:105–130.

Staats BR, Dai H, Hofmann D, Milkman KL (2017) Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Sci.* 63:1563–1585.

Suess C, Woosnam KM, Erul E (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 Rev. Econom.* 8:321–340.

Train K (2015) Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples. *J. Choice Model.* 16:15–22.

Wagman L, Conitzer V (2008) Optimal false-name-proof voting rules with costly voting. *Proc. Natl. Conf. Artificial Intelligence (AAAI 2008)* (Association for the Advancement of Artificial Intelligence, Chicago), 190–195.

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

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

Zhang H, Gan W, Jiang B (2014) Machine learning and lexicon-based methods for sentiment classification: A survey. *Proc. 2014 11th Web Inform. System Appl. Conf. (WISA 2014)* (IEEE Computer Society, Tianjin), 262–265.

Zhang S, Lee D, Singh P, Mukhopadhyay T (2022) Demand interactions in sharing economies: Evidence from a natural experiment involving Airbnb and Uber/Lyft. *J. Marketing Res.* 59:374–391.