Blessing in Disguise? Environmental Shocks and Performance Enhancement

Sumit Agarwal†
Long Wang‡
Yang Yang §

This version: October 6, 2019

†Department of Finance and Department of Real Estate, National University of Singapore, 15 Kent Ridge Drive, Singapore 119245 (email: ushakri@yahoo.com).

‡School of Entrepreneurship and Management, ShanghaiTech University, 393 Middle Huaxia Road, Pudong, Shanghai, China 201210 (email: wanglong@shanghaitech.edu.cn)

§School of Hotel and Tourism Management, CUHK Business School, The Chinese University of Hong Kong, 12 Chak Cheung Street, N.T. Hong Kong SAR (email: zoeyang@cuhk.edu.hk).

Blessing in Disguise? Environmental Shocks and Performance Enhancement

Abstract

This study examines the consumer satisfaction and performance of firms in the service sector in response to reputation shocks caused by random and exogenous air pollution events. The study incorporates machine learning techniques, such as social media sentiment analysis and topic modelling, to analyze review texts and show that pollution shocks during haze episodes lead to significant decreases in consumer satisfaction, which can be explained by the changes in consumers' mood rather than service quality. Moreover, while the level of consumer satisfaction immediately reverts to and then exceeds previous levels after the haze dissipates, such improvement in service quality is not persistent in the long run. The underlying mechanism is that firms with managers closely monitoring customer reviews and actively responding to negative feedback show signicant improvements following a temporary reputation crisis due to a pollution shock. Our findings provide novel empirical evidence on how drivers of organizational reputation change in times of crisis and highlight the importance of realizing deciencies in operation even in the absence of negative shocks.

Keywords: Air Pollution, Service Quality, Customer Review, Natural Language Processing, Machine Learning

JEL Code: Q51, Q53, Z30, D83, D22

1 Introduction

Effective management in the presence of crises and negative shocks is crucial for firms to maintain a high level of performance and remain competitive in the market; however, there is scant empirical evidence attesting to the performance of firms in the service sector as capturing a complete, accurate and timely picture of customer experience and measuring service quality are difficult. In this study, we first develop a web crawler using Python to collect the real-time hotel online review data, and then apply the VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis tool and the LDA (Latent Dirichlet Allocation) topic modelling technique in combination with the econometric models to study the dynamic changes of reviewers' scores and firms' responses in the presence of a negative shock that severely damages the firms' online reputation.

With the increasing importance and share of the service sector (69%) in the global economies (World Bank, 2017), providing superior services and products has become a priority for firms to remain competitive in the market. The productivity of the service sector has been challenging to measure due to difficulties in quantifying the inputs and outputs (Grönroos and Ojasalo, 2004). Thanks to the rapid development of information technology and e-commerce (Chevalier and Mayzlin, 2006; Chen and Xie, 2008; Vermeulen and Seegers, 2009; Liu and Park, 2015), online reviews have become useful measures of productivity in the service sector, which help to provide valuable insight into consumers' perception of quality (Ba and Pavlou, 2002; Duan et al., 2008; Mudambi and Schuff, 2010)[1]. With review sites becoming more significant in shaping customer opinions and behaviors, analyzing online reviews can help firms to improve their service quality and manage their online reputation. However, traditional analyzing methods that were designed for well-structured and quantitative data cannot be directly applied to study unstructured, voluminous and textual data. The advent of the age of big data and the rapid development of machine learning algorithms pose opportunities for businesses to analyze massive unstructured review data and track changes in customers' emotions in a more accurate and ecient way.

This study investigates the impact of a random and exogenous air pollution shock on the customer experience and service quality using the hospitality industry in Singapore as a laboratory. The evidence in this study shows that customer satisfaction decreases during haze episodes because of the unpleasant mood caused by the air pollution rather than the service quality provided by the hoteliers. In an effort to restore their online reputation after a pollution shock, hoteliers respond more actively to negative reviews and improve their service quality to earn more positive feedback. This implies that, in the absence of a threat to their online reputation, hotels do not optimize their productivity.

The paper focuses on four areas of interest. First, we investigate how air pollution impacts the

[1]A great deal of literature has thoroughly documented the importance of online review or electronic word of mouth (eWOM) in relation to sales (Chevalier and Mayzlin, 2006; Duan et al., 2008), trust (Ba and Pavlou, 2002), and consumer decision-making (Mudambi and Schuff, 2010). The existing studies on hotels also show that online reviews, especially negative reviews, can inform hotel managers about their guests' satisfaction with the services provided during their stay, helping managers to identify areas that require improvement (Vermeulen and Seegers, 2009; Chaves et al., 2012). During haze episodes, travellers experiencing the ill effects of air pollution may post a negative review of the hotel even if the hotel maintains a high quality of services.

numerical review score. More specifically, we study the dynamic change of review score before, during and after the haze shock. Second, we determine whether the change in review score after the haze episode persists in the long run. Third, we identify the underlying mechanisms that drive the effect of air pollution on the numerical review score. Fourth, we estimate the benefit gains and losses for guests staying in hotels in Singapore during and after the haze episodes.

The haze episodes occurred in Singapore from 2012 to 2016, especially the first severe air pollution shock took place in June 2013, provide us an ideal opportunity to examine the dynamic responses of travellers and hoteliers to pollution shocks. In particular, the haze episode in Singapore is purely random and exogenous because the air pollutants originate from Indonesia and the possibility of haze episodes in Singapore largely depends on wind directions (Sheldon and Sankaran, 2017). Therefore, potential bias due to endogeneity and sorting in the existing pollution studies (Dominici et al., 2014) being unlikely to be concerns in quantifying the impact of air pollution on travellers' subjective well-being in our paper. Besides, as Singapore is a small island country spanning 709 square kilometers (274 square miles, e.g. less than half the size of Houston in the U.S), air pollution is homogeneously spread island-wide, which means everyone in Singapore is exposed to the air pollutants.

We collect 1,692,659 reviews for 480 hotels in Singapore and 647 hotels in Hong Kong between Jun 2012 and Dec 2016 from *Booking, TripAdvisor, Agoda,* and *Expedia*. We begin the analysis by identifying a causal relationship between pollution shocks and online review scores in Singapore. We find a temporary decrease of 0.409 points, which amounts to 5.2% of the average review score on a 0-to-10 grading scale, during the haze period, and an average increase of approximately 0.295 points for the review scores of hotels in Singapore after the haze shock. We also find that the average review score rises immediately after the haze shock by 0.502 points by the end of the second month; the sharp rise is short-lived and observed only in the first three months following the haze shock. We then examine the underlying mechanisms of review score changes during and after the haze shock.

The decrease of the review score during the haze period points to two possible explanations. First, the "service quality effect", which suggests that severe air pollution may lower productivity in the workplace (He et al., 2019), may explain the decline in the hotels' service quality and review score. Second, the "mood effect", which suggests that air pollution can negatively affect the mood of travellers (Evans et al., 1987; Kim et al., 2010; Zheng et al., 2019), may explain the lower review score. We examine the two explanations by first conducting analyses using the subcategory review scores. *TripAdvisor* and *Expedia* provide information on the subcategory ratings on six aspects of the accommodation experience, including *cleanliness, service, location, sleep, value* and *room* (the definitions of the six subcategories are discussed in Section 4.1). We first classify *cleanliness* and *service* into the improvable group (subjective), and *location* and *room* into the non-improvable (objective) group. We then investigate which group is responsive to the air pollution during the haze shock. This allows us to study whether the reactions of travellers are affected by the perception of experienced hotel service or by their mood. The results show that, while the scores of improvable aspects, *cleanliness* and *service*, do not change much during the haze episode, those of non-improvable aspects, *location* and *room*, decline significantly. This suggests that the "mood effect", rather than the "service quality effect", accounts for the decrease of the review score during the haze episode.

To further investigate the "mood effect" hypothesis, we implement two machine learning techniques: sentiment analysis and topic modelling. The results show that the sentiment intensity is 0.089 points lower in the haze period than in other periods, and proportion of negative lexicons (neutral proportion) is 9.9% (10.5%) higher (lower) in the haze period than in other periods. We then investigate the impact of sentiment intensity on the numerical review score and reveal that guests' sentiment largely (as of 61.1% to 83.4%) explains the decrease of review score during the haze period. The topic modeling analysis shows that the review score and compound sentiment score for the environment-related review, are 0.146 and 0.011 points lower than that for reviews of other topics. More interestingly, the the environment-related topic during the haze period gives rise to lower review scores and compound sentiment scores. The results from the machine learning analysis verify the hypothesis that the severe air pollution may depress the mood of the travellers, driving them to give a lower rating of their stay experience, especially the objective aspects.

The increase of the review score after the haze period also points to two possible explanations. First, the "service quality effect" may explain the improvement in hotel service quality after the haze shock[2]. Second, the "promotion effect" may explain the reduction in hotel room prices after the haze shock. The results of subcategory analysis show that, in the post-shock period, the scores of *cleanliness*, *service* and *value* are significantly positive, while the scores of *location*, *rooms* and *sleep* are not statistically significant. This indicates that the improvement in the service quality of the improvable factors, *cleanliness* and *service*, leads to the increase of the overall numerical review score. This is supportive of the "service quality effect" explanation.

To further study the "service quality effect" explanation and the underlying mechanisms, we utilize the information on manager responses provided by *TripAdvisor* and *Expedia*. More specifically, we study whether hoteliers respond more actively to negative reviews after a haze shock in order to improve their service quality and restore their online reputation. The results show that pollution shocks and decreased consumer satisfaction prompt managers to reflect on their operating efficiency and improve their services, suggesting that a sizable share of firms operate at a sub-optimal level while firms have the capacity to improve their services and reputation. In particular, hotels with managers closely monitoring online reviews and responding to negative reviews show significant improvements. Moreover, we rule out the "promotion effect" explanation for the increase of the review score after the haze shock by using a set of district-daily hotel performance data to show that the haze shock does not have a significant impact on the average daily occupation rates and room prices.

We then address the possible endogeneity issue through a twofold strategy. First, we include online review data of all the hotels in Hong Kong, which shares strong similarities in economy, culture and geography with Singapore and is not affected by haze shocks, as a control group

[2]Mizerski (1982) and Cheema and Papatla (2010) show that hotel managers may be prompted by negative online reviews to improve their service quality in order to restore the hotel's online reputation

and conduct a difference-in-differences (DID) analysis. The essential internal validity of the DID approach is well verified in this study. We find that the review scores of hotels in Singapore are 0.463 points lower on a 0-to-10 scale during the haze episode, compared to the review scores of hotels in Hong Kong. Moreover, online review scores for hotels in Singapore rise by 0.282 to 0.300 points on average in the post-shock periods, relative to those for hotels in Hong Kong. Second, we use satellite fire data in Indonesia to instrument for the haze pollution in Singapore, to address the concerns that PSI readings may capture pollution from local activities. The results are consistent with that in the baseline analysis.

The concern on the change of customer composition is addressed by using a small sample consisting of "frequent travellers", who visit Singapore and Hong Kong before and after a haze period. We also exclude hotels built after 2012 to address the concern that the composition and quality of hotel may change over time. The results are very robust. Another concern on the outliers is addressed by employing the quantile estimation (Koenker and Bassett Jr, 1982). We find that the results of the 25th, 50th and 75th quantiles are all consistent with the baseline results, suggesting that hotels in different quantiles respond to the haze shock.

The heterogeneity of responses show that business travellers and Europeans travellers are more sensitive to the haze shock than their counterparts. Managers of four-star hotels are more likely to respond to the reviews that cover for a haze-episode stay than managers of non-four-star hotels. Lastly, we conduct a back-of-the-envelope calculation and show that tourists in Singapore on average enjoy a service improvement estimated at S$11.16 per room per night, which could be translated into S$138.68 million in total for all rooms in the subsequent 12 months following a haze shock.

Our paper contributes to the existing literature along three dimensions. First, to the best of our knowledge, our study is the first that examines the impact of detrimental environmental shocks on the performance of firms in the service sector. The results highlight the firms' ability to achieve higher productivity and provide better service quality after the shocks. Our study also identifies the ways in which firms can maintain their highest level of performance. Moreover, the study examines the negative environmental externalities from a new angle. While previous studies focus on the detrimental effects of environmental shocks and their overall economic impact or health related consequences[3], this paper not only focuses on the negative effects of pollution shocks, but also highlights the constructive consequences after the pollution shocks. Adding to the existing literature on the relationship between psychological factors and pollution, the analysis on subcategory ratings provides further evidence on how pollution affects the mood of consumers and contributes to the decreases in consumer satisfaction.

[3]Studies have found that the effect of particulates on the lungs can be increased through the absorption of certain chemicals linked to pulmonary cancer and hyperactivity in children (Jayachandran, 2009). Research by Dominici et al. (2014) supports a moderate association between pollutant levels and respiratory discomfort and related illnesses, such as eye irritation. Moreover, health related research documents the causal relationship between exposure to air pollution and depression, anxiety, tension, and anger (Evans et al., 1987), which may adversely affect economic outcomes (Hirshleifer and Shumway, 2003; Levy and Yagil, 2011). Poor air conditions have not only increased health concerns over the past two decades, but also lead to behavioral responses, such as labor supply and labor productivity (Graff Zivin and Neidell, 2012; Hanna and Oliva, 2015; Chang et al., 2016), defensive investments (Deschênes et al., 2017), housing market dynamics (Chay and Greenstone, 2003), and avoidance behavior (Zivin and Neidell, 2009).

Second, our research is also related to the analysis of "big data" using user-generated online reviews. In an era where consumers experiences plays a vital role to achieve successful operations, firms can utilize useful information to identify current or potential problems. Using online review data is a creative way to measure service output, given that most existing studies employ traditional customer satisfaction surveys and questionnaires on service quality. Marketing literature has well documented the extent to which online review texts directly reflect reviewers' perception of the quality of the product or service (Hennig-Thurau et al., 2004; Bansal et al., 2005). The electronic word of mouth (E-WOM) has been found to affect sales (Chevalier and Mayzlin, 2006; Duan et al., 2008; Luca and Zervas, 2016), trust (Ba and Pavlou, 2002), and consumer decision-making (Vermeulen and Seegers, 2009). The existing studies justify the validity to use online review scores from hotel-booking platforms to proxy the reviewers' evaluations of various aspects of hotel services, allowing us to examine how the supply side and demand side of the service sector respond dynamically to environmental shocks as well as to identify the underlying mechanisms driving the changes. The bulk of the data contained in such reviews are textual in nature, and therefore voluminous, unstructured, and challenging to be directly analyzed using traditional methods that were designed for well-structured, quantitative data. To extract new types of features from the text of consumer reviews or to perform computations more efficiently, a great amount of literature has described the well-developed methodologies and presented representative studies in computer science liu2010sentiment, marketing science (Lee and Bradlow, 2011), finance (Loughran and McDonald, 2016), and economics (Varian, 2014).

Third, our paper adds to the understanding of consequences of negative shocks. On the one hand, researchers have revealed the negative consequences of interruption in workplace (Herrmann and Rockoff, 2012; Coviello et al., 2014; Cai et al., 2017). Recent work by Cai et al. (2017) estimate the impact of machine-breakdown on workers' productivity and find a 3.3% decrease in the worker's productivity after the interruption. On the other hand, some studies have provided eye-catching evidence to show how positive outcomes can be triggered by negative shocks. Larcom et al. (2017) show that people who live around the underground train lines that underwent a 48-hour strike now save 20 seconds per journey after the exogenous strike because they were forced by the strike to choose a new route that turned out to be better. Hornbeck and Keniston (2017) present that the Boston fire created an opportunity for people to realize the constraints on urban growth and the land value in the burned areas and nearby un-burned areas increased enormously after the fire. Aggarwal et al. (2012) show that negative posts increase the readership of blogs.

The remainder of this paper is structured as follows. We start by providing background information on the air pollution in Southeast Asia and its general impact on surrounding economics, as well as the prevalence of online review in hotel industry in Section 2. Section 3 introduces the two maching learning techniques. Section 4 describes our dataset and descriptive statistics. Section 5 presents the reduced form evidence, including the analysis of the causal relationship between the negative environmental shock on hotel online reputation, the reputation and performance relationship, the reasons for sub-optimal performance, and the mechanism underlying the main effects. Section 6 discusses several alternative explanations and Section 7 conducts additional heterogeneity tests. Section 8 provides a back-of-the-envelope analysis and Section 9 concludes.

2 Background

2.1 Wildfires in Southeast Asia and Haze Pollution in Singapore

In recent years, heat waves, droughts, and climate changes such as El Niño have led to big fires in several parts of the world, including the western United States, western Canada, the Amazon in South America, and Southeast Asia. Fires are consuming millions of hectares of forest around the world, costing billions of dollars to fight and causing deaths and extensive destruction of property as well as environment. Damage from fires has also contributed to the high rate of deforestation. Although many wildfires occur during periods of high temperatures and drought, human activity has also made fire events more frequent and more intense. In particular, intentional burning for forest cultivation and agriculture has increased fire incidences in tropical areas.

In the ASEAN region, nearly all of the fires and haze over the past two decades have been caused directly by human intervention rather than by natural events (Qadri, 2001). For example, farmers and owners of agricultural land in Southeast Asia have for many years used burning as a way to clear land for agriculture, even though it is illegal. This method has been a primary cause of huge wildfires in the Indonesian archipelago. Prevailing winds blow smoke, ash, toxic gases and other pollutants from this area to nearby countries, such as Singapore and Malaysia.

The 1997 heat haze event experienced by Singapore, when the 24-hour Pollution Standard Index (PSI) reached the "unhealthy" level of 138, was the first to receive international attention. During the event of 2013, a new record was set when the three-hour PSI reached 401, which is considerably higher than the threshold set for the "hazardous" level (Sheldon and Sankaran, 2017). However, the worst haze event in Singapore to date was that of Oct 2015, when the one-hour PSI reading reached 471. The haze causes irritation to eyes and, when inhaled for prolonged periods, can have harmful long-term effects on the lungs, heart, and respiratory system (Jayachandran, 2009). Since haze pollutant contains carbon dioxide and sulfur dioxide, along with aerosols and toxic particulates as well as a strong acrid and burning smell, it is easily detectable by public, and the air is clearly distinguishable from that of normal days without haze[4]. Associated economic impacts include disruption to transport and tourism (Quah, 2002; Lee et al., 2016).

An air pollution event in Singapore can be considered as a random and exogenous shock (Sheldon and Sankaran, 2017). First, local emissions are not the cause of the haze pollution experienced in Singapore. The local air pollution is trivial because of stringent industrial emission regulations

[4]Information on haze pollution is publicly available from the National Environmental Agency (NEA) of Singapore. To keep residents informed about the air quality, the agency regularly reports information on one-, two-, and three-hour measurements through radio, television, internet, mobile applications and other media. NEA also provides five different PSI descriptors to indicate the levels of pollution risks based on the PSI measures. PSI readings above 100 are considered threatening to health. PSI values from 201 to 300 are considered as 'very unhealthy', and PSI values above 300 are considered 'hazardous'. The NEA and the Ministry of Health (MOH) provide general advisories to local residents, especially concerning sensitive groups such as children, the elderly, pregnant women, and people with respiratory illness, so they will reduce their exposure to the pollution outdoors.

in Singapore; most of the oil refineries and petrochemicals plants are located in Jurong Island, a reclaimed island west of the main island. This is reflected by the monthly average PSI value at around 40 in the absence of the haze pollution brought to Singapore by the wind. Second, the haze events in Singapore are exogenous because the air pollutants originate from Indonesia and the severity of haze pollution in Singapore largely depend on wind factors. Besides, as Figure 3 shows, the four haze episodes took place in different months in different years, which indicates that the haze shock is less likely to be seasonable and predictable. Since Singapore offers an ideal environment for clearly identifying the causal relationship between air pollution and economic outcomes, the Jun 2013 and Sep-Oct 2015 haze episodes are particularly suitable for research.

2.2 Online Reviews and Consumer Satisfaction

Tourism has become one of the largest and fastest-growing economic sectors in the world. The number of international tourist arrivals grew from 0.89 billion in 2009 to 1.24 billion in 2015[5], and the travel expenditure rose by around 50% for the same period. In Singapore, the tourism receipts increased by 10% in 2017, reaching $12.7 billion (Statistics Singapore, 2018).

Online review websites and travel communities have become the most influential information source for travelers (Vermeulen and Seegers, 2009; Liu and Park, 2015). As reported by Blanke and Chiesa (2013), around 87% of international travelers have used the Internet for trip planning, and 43% have read the online travel reviews. The reviews enable tourists to share their experiences of places and products as well as to communicate with other travelers and with industry managers. In order to guarantee authenticity, the online process restricts reviews and comments to travelers who made corresponding online bookings. Online reviews provide comprehensive information about consumer satisfaction with various attributes, such as the room facilities, the value for money, the location, sleep quality, cleanliness of the room, and service quality.

The user-generated contents are an electronic form of traditional word-of-mouth (E-WOM) marketing, which has a signicant impact on consumer buying decisions and willingness to pay (Chevalier and Mayzlin, 2006; Vermeulen and Seegers, 2009; Mudambi and Schuff, 2010). Online reviews that can reach large numbers of potential consumers have substantial E-WOM impact, and can therefore influence business outcomes for hotels. Ye et al. (2011) find that a 10% increase in traveler review ratings can bring 5% or more additional reservations to hotels. Our paper argues that declining scores in online reviews touch a nerve of hotel managers, gain the attention of their teams, and generate incentives for hotels to improve services and efficiency.

Organizations have reasons to pay close attention to online reviews. As suggested in previous literature, E-WOM exerts great influence on trust, decision-making, and sales (Ba and Pavlou, 2002; Chevalier and Mayzlin, 2006; Duan et al., 2008; Vermeulen and Seegers, 2009). It can be considered as a way to reduce the information asymmetry for other consumers (Liu and Park, 2015). Thus, organizations and managers should be concerned about the content of reviews and review scores. Moreover, negative reviews are more important to managers than positive ones

[5] Data source: https://data.worldbank.org/indicator/ST.INT.ARVL

when it comes to first-time consumers (Mizerski, 1982). According to the economic theory and risk preference, first-time consumers tend to have higher risk-aversion, and negative reviews collide with their purchasing intention (Holt and Laury, 2002; Thompson, 2005). Thus, a negative review will collide with consumers' attitude and purchasing intention, implying that organizations should pay more attention to the negative comments than to the positive reviews (Cheema and Papatla, 2010). Also, organizations are aware that negative reviews can hurt their reputations, which subsequently causes customer churn and performance loss (Roberts and Dowling, 2002; Boyd et al., 2010). To repair the damage caused by negative reviews and address the potential service deficiency, managers should respond actively to negative reviews. Our analysis presents the links and mechanisms that help to explain why managers respond to negative reviews and why hotels improve their service quality.

3 Big Data Analytics Techniques

Using web crawling tools written in Python and capable of iteratively and automatically downloading web pages, we first collect online user-generated review contents and all the information on the review pages from the four largest travel websites with hotel reviews (i.e., *Booking, TripAdvisor, Agoda,* and *Expedia*). In order to analyze the features derived from the qualitative and corresponding quantitative components of a review, we combine machine learning and natural language processing techniques with various econometric approaches. Section 4 describes the datasets in detail and this section describes techniques concerning big data analytics.

3.1 Sentiment Analysis

Air pollution shocks affect online reputation through two possible channels: reduction in the quality of service provided by hotel staff and negative emotions in customers. We apply sentiment analysis to identify the channel through which air pollution can depress mood levels, which is the main channel we argue in the study. Based on the computational treatment of subjectivity in a text, the analysis examines customers' opinions, sentiments, and emotions related to their stay experiences.

Although firms monitor the quantitative online review scores on various online platforms, they take online review text less seriously due to the considerable challenges using qualitative review text for analysis. The challenges include not only the large amounts of qualitative data involved, but also the specific sentiment embedded in multiple languages, e.g., short forms, slang terms, memes and emojis. In this study, we apply VADER (Valence Aware Dictionary and sentiment Reasoner)[6], a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media to extract both the sentiment polarity (positive/negative) and the sentiment intensity (on a scale from -1 to 1) of a review (Hutto and Gilbert, 2014).

VADER has two main advantages over traditional methods of sentiment analysis. First, it performs well on the social media by interpreting the emojis, slang, and emoticons that form

[6] VADER is fully open-sourced under the MIT License. Source: `https://choosealicense.com/`. Please refer to the Github page `https://github.com/cjhutto/vaderSentiment` for details of VADER scoring method.

important components of the social media environment, and by analyzing sentiments based on such key points as punctuation, capitalization, degree modifiers, conjunctions, and preceding tri-gram. Second, unlike other machine learning and deep learning techniques, VADER does not require any training data. It uses a combination of qualitative and quantitative methods, and is constructed from a generalized, valence-based, and human-curated gold standard sentiment lexicon.

In Table 1, we illustrate the strength of VADER sentiment analysis using several examples. We analyze sentences with different features (i.e., the specific punctuation, capitalization, degree modifiers, conjunctions, emojis, slang and emoticons in sentences) to show how well VADER works on the social media type of text, which contains various lexical forms. *Positive*, *Negative* and *Neutral* scores represent the proportion of text that falls in the corresponding categories, and the *Compound Score* is a metric that calculates the sum of all the ratings of the different lexicons, which have been normalized between -1 (most extreme negative) and +1 (most extreme positive). We utilize the VADER method to calculate the sentiment intensity for each review, and then analyze whether the guest sentiment affects the numerical rating of the stay. The details of the sentiment outcomes extracted from the online review texts are presented in Table 2.

[Table 1 inserted here]

[Table 2 inserted here]

Our study also deals with the difficulty of executing sentiment analysis across multiple languages (Han et al., 2016; Schuckert et al., 2015) by using *Googletrans*, a free and unlimited Python library that implemented Google Translate API[7]. This allows us to translate all the non-English reviews into English before applying VADER, because VADER does not work directly with other languages.

3.2 Topic Modeling

Since travellers tend to make specific comments based on their personal experiences and preferences in hotels, review text may reflect only the few dimensions (topics) that concern them the most. Our interest is to examine whether different dimensions in customer perceptions affect the corresponding online rating. In particular, we want to investigate the extent to which the reviews related to pollution or environment play a vital role in shaping customer satisfaction.

We utilize topic modeling, a frequently used text-mining tool for discovery of hidden semantic structures in a text body, to classify the massive review text into different topics. A topic model is a type of algorithm that scans a set of documents (known in the fields of machine learning and natural language processing as a corpus), examines how words and phrases co-occur in them, and automatically “learns” groups or clusters of words that best characterize those documents. The sets of words often appear to represent a coherent theme or topic (Huang et al., 2017). Commonly, a variety of topic modeling algorithms are used — including non-negative matrix factorization, Latent Dirichlet Allocation (LDA), and Structural Topic Models. LDA[8] is a popular topic modeling

[7]For details on Google Translate API, please refer to the API Documentation. Source: `https://py-googletrans.readthedocs.io/en/latest/`

[8]LDA was developed by Blei et al. (2003). This is an extension of Probabilistic Latent Semantic Analysis (PLSA) developed in 1999 by Thomas Hoffman with a minute difference in treatment of per-document distribution. For the

techniques used in the areas of computer science, marketing and economics (Guo et al., 2017; Puranam et al., 2017). It builds a topic-per-document model and a words-per-topic model, modeled as Dirichlet distributions. LDA can complete many steps of the textual analysis with little human intervention, even labeling dimensions, and is more suitable for dealing with large and unstructured online reviews, thus creating meanings that are more realistic.

Two general assumptions underlie LDA topic modeling: 1), documents that contain similar words usually share the same topic, since documents are probability distributions over latent topics; and 2), documents containing groups of words that frequently occur together usually share the same topic, since topics are probability distributions over words. For instance, environment-related topics would contain such words as "environment", "pollution", "dirty", etc., and if these words occur together frequently in various documents, those documents may belong to the same category. We also evaluate the accuracy of the topic model by calculating the perplexity metric, which measures the probability of new, unseen data given the model that was learned earlier; the lower the perplexity, the better the model. Using LDA topic modeling, the hidden semantic structure of an online review is converted for use in a supervised classification problem for the later regression analysis. Section 4.1 presents the output of topic modeling..

4 Data and Descriptive Statistics

The data is obtained from multiple sources, and grouped into two broad categories: online reviews and measures of ambient conditions. Table 2 presents a summary description of the data set, with online review data of all hotels in Singapore reported in Panel A, and monthly ambient conditions reported in Panel B. Appendix A provides detailed definitions of all variables used in the analysis. Appendix Table B1 summarizes the statistics of the Hong Kong dataset, which is used to address an alternative hypothesis. Appendix Table B2 uses a star rating to provide a breakdown of hotels. The demographic characteristics of travelers in our sample, including country of origin and traveller type, are listed in Appendix Table B3.

[Table 2 inserted here]

4.1 Online Reviews

We obtain the review data of all hotels in Singapore between June 2012 and December 2016 from four widely used platforms: *Booking, TripAdvisor, Agoda,* and *Expedia.* Sample reviews from various websites are presented in Appendix C. Figure 1 plots the geographic distribution of 413 hotels and 67 hostels in Singapore[9]. Luxury hotels (four- and five-star hotels) are located mainly in the central business districts in Singapore. The hotel prices at the neighborhood level are reflected

details of methodology and source code for LDA program, please refer to Blei et al. (2003) and Github page. Source: `https://github.com/susanli2016/NLP-with-Python/blob/master/LDA_news_headlines.ipynb`.

[9] 413 licensed hotels were registered in Singapore Tourism Board as of Dec 2016, and our data includes online reviews from another 67 hostels where travellers can make reservations through the online platforms. See `https://www.stb.gov.sg/content/stb/en/industries/hotels.html` for more details.

in a thematic map and are classified by standard deviation (i.e., the darkest color is assigned to values above $305 and other breaks are introduced in 1 standard deviation intervals).

The review data contains a rich set of information regarding the hotels and reviewers, including hotel name, address, star rating, average room price, as well as a traveller's review score, reviewed date, stayed month, account name, guest type (such as business, couple, family, friends, group, solo traveller, and other), and guest country of origin (there are 216 countries in the sample), as well as the contents of each review. The key variables we use in this study are the numerical review score and the full review text posted by individuals for each of their stays. We utilize the natural language processing and machine learning techniques introduced in Sections 3.2 and 3.3 to construct several measures on the sentiment intensity of each review text and to categorize the topic of each review. Note that in each review, all four platforms report only the stayed year-month, rather than the specific stayed date. For instance, Appendix Figure C1 records an online review at the year-month level. We therefore match the online review data to monthly weather and pollution data.

As shown in Table 2, among the 869,115 online reviews of hotels of Singapore, 102,276 (11.77%) are from *Expedia*, 173,382 (19.95%) are obtained from *TripAdvisor*, 195,841 (22.53%) are provided by *Booking* and 397,616 (45.75%) come from *Agoda*. The review scores on *TripAdvisor* and *Expedia* are on a 0-to-5 rating scale, while the review scores on *Agoda* and *Booking* are on a 0-to-10 rating scale. For consistency and ease of interpretation, we re-scale the review score in *TripAdvisor* and *Expedia* from 0 to 10, with a higher score corresponding to a higher level of customer satisfaction. The average review score is at 7.83, with a standard deviation of 1.8. The mean value of compound sentiment score extracted from the review texts in Singapore is estimated at 0.44. In a review text, 71% of the lexicons are neutral, 25% positive, and 4% negative.

To implement the LDA topic model and to set the parameter value, we follow Mankad et al. (2016). In determining the best model, we evaluate the "perplexity" goodness-of-fit measure by varying the number of topics. Since a lower perplexity value represents a better model fit, we set the number of topics as 10 because it bears the lowest "perplexity" value (at -7.610) compared to the choices of 5, 10, 15, ..., 30 topics. Figure 2 presents the 10 most frequently used key words in each of 10 topics. It shows that Topic 7 is the most related to environment or pollution due to its frequent use of keywords: pollution, dirty, and environment. We also show the interactive plot[10], in Appendix Figure B2, where each bubble represents a topic. The larger the bubble, the more prevalent that topic is. The bar plot on the right-hand side of Appendix Figure B2 shows the frequency of the terms in the topic, relative to the total term frequency in the documents. To determine how the environment-related review (topic 7) affects the numerical review score and the sentiment score, we create a dummy (indicated as *Topic7*) that equals 1 if a review is assigned to topic 7, and 0 otherwise. *Topic7* averages at 0.08, which suggests that 8% of the reviews in Singapore belong to the environment-related topic.

[Figure 1 inserted here]

[Figure 2 inserted here]

Moreover, the data from *TripAdvisor* and *Expedia* provides additional information on whether

[10]We also output the interactive plots for other topics, and upload the data to the online appendix.

hotel managers respond to guests' reviews. Specifically, the data includes the response contents and response date. For example, among 275,658 reviews from *TripAdvisor* and *Expedia* in Singapore, 134,399 (48.76%) reviews are responded by the hotel managers.

TripAdvisor also presents the subcategory ratings of six different attributes related to a stay: *cleanliness, service, location, room, value* and *sleep*. More specifically, *cleanliness* assesses the hygiene and cleanliness standards of hotel, and the efficiency of housekeeping; *service* evaluates the friendliness and communication skills of staff, as well as their efficiency in solving problems; *location* represents the accessibility and proximity to attractions, city center, and airport/railway stations; *room* refers to evaluation of the amenities in the room/bathroom, and size and layout of the room; *value* refers to the utility derived from the money spent on room, food and beverage, and other costs; *sleep* measures factors related to the overall sleep experience, such as quality of bed and pillow, noise from air conditioners and hallway, and the indoor environment.

We classify subcategory reviews into two groups, based on whether the evaluation in a subcategory can be improved by effort from hotel staff through the daily operations. Specifically, *cleanliness* and *service* are assigned to the improvable group, while *location* and *room* are assigned to the non-improvable group. We will discuss this in greater detail in Section 5.2.

4.2 Air Pollution and Weather Measures

For the air quality information in Singapore, we collect the 24-hour Pollution Standard Index (PSI) readings from the National Environment Agency (NEA)[11] for the period between Jun 1, 2012 and Dec 31, 2016[12]. The readings, which range from 0 to 500, are reported by the NEA through mass media, such as television, radio, the Internet and mobile applications, to inform residents about air quality. The 24-hour PSI reading provides an hourly indication of the air quality by averaging the data collected over the past 24 hours. The 24-hour PSI reading is a composite measure of the concentrations of multiple pollutants, which includes particulate matter (PM_{10}), fine particulate matter ($PM_{2.5}$), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), Ozone (O_3), and carbon monoxide (CO). We calculate the monthly average and monthly maximum of PSI readings based on the hourly PSI readings.

The trends of monthly mean PSI (dashed line) and monthly maximum PSI (dotted line), as well as the monthly average online review score (solid line) from Jun 2012 to Dec 2016 in Singapore, are plotted in Figure 3. Figure 3 presents an unconditional negative relationship between review score

[11] Data Source: `https://www.haze.gov.sg/`

[12] For pollution information in Hong Kong, we turn to the Air Quality Index (AQI) provided by the Environmental Protection Department (EPD) of Hong Kong. Source: `http://www.epd.gov.hk/epd/english/environmentinhk/air/air_quality/air_quality.html`. The AQI contains the raw daily records of PM_{10}, $PM_{2.5}$, SO_2, NO_2, O_3, and CO from 13 ambient air quality monitors throughout Hong Kong. Since PSI measures are not readily available in Hong Kong, we construct the PSI measure following the method used by NEA Singapore. Source: `https://www.haze.gov.sg/docs/default-source/faq/computation-of-the-pollutant-standards-index-(psi).pdf`. More specifically, we compute a sub-index value for each pollutant based on the ambient air concentration of the pollutant, and take the highest sub-index value as the PSI value. Compared to AQI, PSI is determined by the pollutant with the most significant concentration. For analysis use, we calculate the monthly average and monthly maximum of PSI readings.

and PSI: the review score drops significantly during the periods with larger PSI readings (marked by the shaded bars). More interestingly, we find that the review score reverts to, and even exceeds the magnitude before the haze episode. Other than relying on the continuous measure of PSI, we identify two severe and two mild haze shocks in Singapore based on the severity and duration of PSI readings. The dark shaded areas and large spikes highlight two severe haze shocks in Jun 2013 and Sep-Oct 2015, while the light shaded areas represent two mild haze episodes in Oct 2014 and Aug 2016. The width of shaded areas indicates the duration of air pollution events.

[**Figure 3 inserted here**]

In relation to travellers' experiences, weather conditions could confound the effects of air pollution. To avoid the possible contamination, we collect other weather information, such as temperature, rainfall, and wind speed from *Wunderground*[13]. Wunderground also includes 18 weather keywords to indicate the weather status of each hour. Based on these keywords, we create the "Days of Haze (DoH)" variable, which aggregates the number of days per month that contain the keywords *"light haze," "haze," or "heavy haze"*. *DoH* is considered as an alternative measure of air quality in our study.

5 Reduced Form Evidence

We study the impact of the random and exogenous pollution shocks on consumer satisfaction and service quality of firms following a fourfold empirical strategy. First, using an event study method, we examine whether exogenous variations in air quality affect consumer satisfaction. In particular, we study the dynamic responses of review scores before, during, and after the haze shock. Second, to understand the underlying mechanisms that account for changes to online reputation during the haze episode, we study the heterogeneity in response across subcategory review scores, and then apply two machine learning techniques: sentiment analysis and topic modeling. Third, to understand the underlying mechanisms through which online review scores change after the haze dissipates, we utilize the data on manager responses. Fourth, in performing the DID test, we include the reviews of all the hotels in Hong Kong during the sample period as a control group, to address the possible endogeneity; We further strengthen the causal relationship between haze pollution and online reviews by using a two stage least square approach and employ the satellite fire data in Indonesia to instrument for the air pollution in Singapore.

5.1 The Causal Relationship between Air Pollution and Consumer Satisfaction

First, we employ a pollution-shock approach to investigate the reduced-form relationships between various haze measurements and consumer satisfaction. The reduced-form pollution-shock approach makes few identification assumptions and allows strong causal interpretation (Dell et al., 2014). A necessary condition for the event study is the exogeneity of the pollution shock. Since the haze episodes in Singapore are random and exogenous, they suit the study particularly well.

[13]Source: `https://www.wunderground.com/`

To analyze the responses of online review scores to the changes of air quality in Singapore, we use the following OLS estimation:

$$Score_{i,j,k,t} = \alpha + \beta \cdot Haze_t + \gamma \cdot X_t + \mu_g + \delta_o + \zeta_j + \eta_k + \theta_{year} + \theta_{month} + \epsilon_{i,j,k,t} \tag{1}$$

where i, j, k, t, g, and o index reviewers, hotels, websites, stayed year-month, guest type, and guest origin country, respectively. The dependent variable $Score_{i,j,k,t}$ is the online review score of reviewer i rates for his/her stay in year-month t at hotel j on website k. It should be noted that t is the year-month in which the reviewers stayed at the hotel, not the time they posted their review. $Haze_t$ represents one of the haze measures in the stayed year-month t. We use six different measures to identify haze episodes from Jun 2012 to Dec 2016. First, we create two binary variables to identify the haze shocks, $Shock^a$ and $Shock^b$. As shown in Figure 3, two severe haze shocks (labelled by the dark shaded areas in Jun 2013 and Sep-Oct 2015) occurred during the period, as identified by $Shock^a$. $Shock^b$ represents two mild haze shocks in Sep-Nov 2014 and Aug 2016. Notably, we use $Shock^a$ as the main measure of air pollution simply to differentiate the haze episode from the non-haze episode because Singapore is free of air pollution most of the time. We also use the monthly average (PSI^{mean}) and monthly maximum (PSI^{max}) 24-hour PSI readings, which are the most direct way of measuring the haze intensity in the study period. In addition, we use DoH to measure the number of days per month with haze status; and we create categorical variables to identify three levels of haze by classifying the monthly average PSI^{max} into three categories: 0-50 (PSI_{0-50}), 51 to 100 (PSI_{50-100}), and above 100 (PSI_{100+}), which allows us to examine non-linear response to the haze level. To control for weather conditions that could possibly affect guests' lodging experiences, we add a vector of other time-varying observations, X_t, which includes the logarithmic terms of temperature, rainfall, and wind speed in year-month t.

In addition, we control for a rich set of fixed effects, which isolates our estimations on the haze effects from other unobservable confounding factors. μ and δ stand for the guest type fixed effect and the guest's country of origin fixed effect, respectively, which capture factors that possibly impact online review scores at the guest level. ζ is hotel fixed effect, which absorbs the fixed spatial unobservable characteristics across hotels. η is a website fixed effect that eliminates the differences of constructing the review score across the four websites. θ_{year} and θ_{month} represent the stayed year fixed effect and stayed month fixed effect, respectively, which absorb the time variations of online review scores and neutralizes the seasonality. The reason we control for year fixed effect and month fixed effects rather than year-month fixed effect is that including year-month fixed effect would absorb any month-to-month variation in economic and weather conditions, which would therefore eliminate our key variable $Haze_t$. Given that online reviewers may be influenced by the existing reviews of the hotel they are rating, review scores for a hotel may be correlated. Thus, all standard errors are robust and clustered at the hotel level, which allows an arbitrary variance-covariance matrix to capture the potential serial correlations in the residual error terms.

Notably, we restrict the sample period of the main analysis from Jun 2012 to Aug 2014 and conduct the event study on the first severe haze shock, which took place in Jun 2013, for two reasons. First, the test on the first severe haze shock is cleaner because it is immune from the expectation issue, as well as the previous haze effect. For instance, tests on the Sep-Oct 2015 haze would be less precise as they might suffer from contamination of the Jun 2013 haze shock. Second, focusing on a relatively short period mitigates the possibility of our estimation being contaminated by other events or weather shocks, e.g., the Singapore general selection.

Table 3 presents the regression results of estimating Equation (1) using online review data of hotels in Singapore from Jun 2012 to Aug 2014. As shown in Column (1), the average online review score is around 0.409 points lower during the first severe haze episode in Jun 2013 than that in other months during the sample period. Given that the average review score of hotels in Singapore is approximately 7.83, this amounts to a 5.2% decrease in average score during the haze episode. Columns (2) and (3) show that as the PSI readings double, the review scores decrease on average by 0.297 to 0.496[14]. The results implied that during a severe haze shock, when the 24-hour PSI readings increase five times from 50 to 300 and a linear relationship is assumed between the PSI readings and online review scores, the estimated drop of online review scores can be as large as 2.48 on a 0-to-10 scale. Column (4) indicates that one additional hazy day decreases the monthly average online review score by 0.040. Following Chang et al. (2016), we model the monthly PSI reading with a series of indicator variables to allow for a nonlinear effect of PSI. As shown in Column (5), we continue to find consistent evidence that online reviews respond negatively to PSI. We find that on average, PSI levels from 51 to 100 lower the review scores by 0.182 points, and that when PSI levels jump over 100, the magnitude of negative effects increases to 0.484 points.

[Table 3 inserted here]

In addition, we conduct the same analyses using the full sample period from Jun 2012 to Dec 2016 and report the results in Appendix Table B4. The estimated coefficients of haze measures are negative and statistically significant, which are highly consistent with our results in Table 3. A new result in Appendix Table B4 is the estimation of $Shock^b$, which differentiate the mild haze episode from other periods. $Shock^b$ bears the negative sign and statistical significance and its magnitude is smaller than $Shock^a$.

We then study the responses of online review scores after haze shocks in the full sample period from Jun 2012 to Dec 2016 and give the specification as following:

$$Score_{i,j,k,t} = \alpha + \beta \cdot Post + \gamma \cdot X_t + \mu_g + \delta_o + \zeta_j + \eta_k + \theta_{year} + \theta_{month} + \epsilon_{i,j,k,t} \tag{2}$$

where $Post$ is a binary variable equal to 1 for the periods after the haze shocks, and 0 for the periods before the haze shocks. The indicators and fixed effects are the same as those in Equation (1). Four haze episodes, including two severe and two mild ones, are excluded when estimating Equation (2). To precisely examine the change of the review score before and after the strong haze shocks, we divide the sample period into four phases (see the phase labels in the text-boxes of Figure 3): Jun 2012 to May 2013 (phase 1, the pre-shock period of the Jun 2013 haze), Jul 2013 to

[14]In the standardized interpretation, a 1 standard deviation increase in mean (max) PSI value results in a 0.095 (0.064) standard deviation increase in the review score.

Aug 2014 (phase 2, the post-shock period of the Jun 2013 haze), Dec 2014 to Aug 2015 (phase 3, the pre-shock period of the Sep-Oct 2015 haze), and Nov 2015 to Dec 2016 (phase 4, the post-shock period of the Sep-Oct 2015 haze). It is worth noting that phase 2 is the post-shock period of the Jun 2013 haze shock; and phase 3, the following period, is the pre-shock period of the Sep-Oct 2015 haze episode. Figure 3 shows that a mild haze shock took place in 2014, which might have affected the online review scores in the pre-shock period of the Sep-Oct 2015 haze. Since online reviews in phase 1 would not be affected by previous air pollution events, the comparison between phases 1 and 2 is therefore a cleaner test.

Table 4 provides the results comparing the responses among different phases. The coefficients of *Post* in Columns (1) and (3) capture the differences between the online review scores before and after the two severe haze shocks. Columns (2) and (4) conduct additional robustness checks by comparing the review scores to alternative pre- or post-shock periods. More specifically, the coefficient of *Post* in Column (1) indicates that the average online review score increases by 0.295 points for the post-shock phase 2 after the Jun 2013 haze shock, compared to the average online review score in the pre-shock phase 1. This amounts to 3.77% of the average review score in Singapore. Column (3) presents the estimation for the Sep-Oct 2015 haze shock and shows that the average online review score in the post-shock phase 4 is 0.251 points higher than that in the pre-shock phase 3. In Column (2), we compare phase 3 to phase 1, and the coefficient on *Post* is positive but statistically insignificant, suggesting that the increment in online review scores after the Jun 2013 haze shock may not be persistently large. In Column (4), we compare phase 4 to phase 1, which is the pre-shock period free of any haze effect. The coefficient on *Post* is positive and statistically significant at 0.144.

[Table 4 inserted here]

In summary, the results in this section show that, as reflected in the review scores, online reputation decreases significantly during the air pollution shocks, but then immediately reverts to, or even exceeds, its previous level when the haze episode ends. Understanding not only the significance and magnitude of the shock, but also the underlying mechanisms driving such results, is critical: how and why do review scores change during and after air pollution shocks?

5.2 Why do Online Review Scores Decrease during the Haze Episode?

Damage to online reputation during exogenous air pollution episodes can be explained either by a decreased level of hotel service provided by hotel staff ("service quality effect"), or by the negative mood triggered in hotel customers ("mood effect"), or by a combination of both. To determine the underlying cause of plummeted review scores, we look at subcategory online review scores that provide information on which aspects of the stayed experience are affected by the general service quality, and which are affected by the customer mood.

In particular, *TripAdvisor* provides the numerical review scores of six subcategories: *cleanliness*, *service*, *value*, *location*, *rooms*, and *sleep*. Based on the definitions of the six subcategories, *cleanliness* and *service* reflect the general service quality of housekeeping staff and hotel management, and thus can be considered as improvable aspects since the ratings on these two categories would increase given greater effort by the hoteliers. *location* and *rooms* describe the objective aspects of the hotel, such as the proximity to attractions, room size and hotel facilities, and thus are considered non-improvable because the location, room size, and the hotel facilities cannot be changed, at least not in a short time or at a low cost. *sleep*[15] measures the overall sleeping experience of the guest and therefore. It is difficult to categorize *sleep* into the either improvable or non-improvable groups given its definition. *value* refers to the general utility derived from the money spent on the stayed experience, which broadly covers *cleanliness*, *service*, *location*, *rooms*, and *sleep*. To this end, we do not assign *value* and *sleep* into either the improvable or non-improvable groups.

Bad air quality exerts great effects on people's mood (Evans et al., 1987; Kim et al., 2010; Zheng et al., 2019). If air pollution, mediated by mood, lead to a collective changes in the level of customer satisfaction, then we would expect they give lower review scores for all categories. On the other hand, if haze makes it more difficult for the hotel workers to exert effort, then decreases in service quality should bring down ratings for the improvable areas such as *cleanliness* and *service*.

We then examine guests' responses in each subcategory of the online reviews during and after the first severe haze shock following Equations (1) and (2). Table 5 presents some interesting results, where Panel A and Panel B correspond to the responses during and after the Jun 2013 haze shock, respectively. In Panel A, we find that the scores in all six categories decrease during the June 2013 haze episode, but that only three (*value*, *location*, and *room*) are statistically significant. Insignificant changes in ratings of the subjective aspects (*cleanliness* and *service*) during the haze episode suggest that service quality provided by the hotel staff remains the same; however, the significant decrease of ratings on the objective aspects (*location* and *rooms*) of the hotel points to the possibility that the air pollution shock affects the mood of the guests, which has a substantial negative effect on their ratings of the objective aspects of the hotel. For instance, the rating of *location* declines significantly because the utility of a guest's stay is greatly reduced if they can barely see anything from their window during the haze episode. This suggests that the "mood effect", rather than "service quality effect", accounts for the lower review score during haze episodes.

[Table 5 inserted here]

To establish a direct link between customer mood and air pollution for the "mood effect" argument, we utilize sentiment analysis and topic model in analyzing the unstructured review contents. More specifically, we investigate whether guest mood, which can be extracted from the review text, changes during and after the haze episode. Using the four sentiment outcomes (*compound score, positive proportion, negative proportion*, and *neutral proportion*) generated by the VADER program as the dependent variables, we re-estimate Equations (1) and (2) and report the results in Table 6.

The coefficient of $Shock^a$ for compound score in Column (1) of Panel A is statistically negative

[15] The sleep quality is largely a subjective experience, such as how tired guests feel upon waking, sleepiness, restfulness, and difficulty falling asleep and maintaining sleep. On the other hand, hotels can provide better sleep experience by improving bedding quality, using soundproof walls and blackout curtains, and upgrading climate control appliances. Therefore, it is relatively difficult to assign *sleep* into either improvable or non-improvable groups.

at -0.089, suggesting that the sentiment intensity is 0.089 points lower in the haze period than that in normal periods. However, the coefficient of *Post* for compound score in Panel B turns out to be statistically indifferent from zero, which means that the sentiment intensity bounces back to its original level after the haze episode ends. The results are consistent with existing research that indicates a strong association between air pollution and depressed mood, and therefore supports our "mood effect" argument that the negative mood causes travellers to rate their stay experience lower, and it is evident for the decreased rating on the objective aspects.

[Table 6 inserted here]

The pollution effects on other three sentiment outcomes are presented in Columns (2) to (4). We find that the effect of pollution shock on the proportion of negative sentiment is statistically significant at 0.099, and that the effect of pollution shock on the proportion of neutral sentiment is statistically significant at -0.105; the impact of $Shock^a$ on positive proportion is statistically different from zero. This suggests that the proportion of negative (neutral) sentiment lexicons in review texts is 9.9% (10.5%) higher (lower) in the haze period than that in other periods. More importantly, the increment of negative sentiment lexicons comes mainly from the decrease of neutral lexicons. The coefficients of *Post* for the three shares in Panel B are all statistically indifferent from zero, implying that customer mood recovers to its previous level after the pollution shock.

We then investigate the impact of sentiment intensity on review scores. In particular, we add *compound score* as the explanatory variable to Equation (1) and rerun the regressions using the review score as the dependent variable. The results are reported in Table 7 and can be directly compared with the results in Table 3. The coefficient on the compound score is statistically significant at around 1.206 in all columns, implying that the customers' mood directly affects the numerical review score. Compared to the results in Table 3, the magnitudes of *Haze* coefficients in Table 7 decrease by 61.1% (Column 2) to 83.4% (Column 1). In particular, the inclusion of compound score reduces the impact of $Shock^a$ on review score from -0.409 (as shown in Column 1 of Table 3) to -0.068, which suggests that the consumer mood can largely explain the decrease of review score during the haze period. This strongly supports our "mood effect" argument that the negative mood caused by the air pollution gives rise to the lower review score during the haze episode.

[Table 7 inserted here]

To examine whether the review text that reflects the environment or pollution related topic (Topic 7) is associated with lower review scores and more negative sentiment, we use a topic modeling technique, namely Latent Dirichlet Allocation (LDA). The results in Table 8 show that both the review score and the compound sentiment score for environment-related reviews are 0.146 and 0.011 points lower, respectively, than that for other topics. Interestingly, the interactions of $Topic_7 * Shock^a$ are statistically negative, implying that the environment-related reviews result in lower review scores and compound sentiment scores in the haze period.

[Table 8 inserted here]

In summary, the results from the subcategory analysis, sentiment analysis, and topic modeling analysis support the "mood effect" that the severe air pollution may depress the sentiment of the travellers and therefore drive them to rate their stay experience lower, especially in the objective aspects.

5.3 Why do Online Review Scores Increase after the Haze Episode?

The sharp increase of review scores after the haze episode can also be explained either by hotels improving service quality to regain online reputation ("service quality effect"), or by promotions or discounts introduced after the haze shock ("promotion effect"), or a combination of both.

The results in Panel B of Table 5 show that the coefficients of *Post* on *cleanliness*, *service* and *value* scores are significantly positive, implying that the ratings in these three subcategories increase substantially in the post-shock period. However, we do not observe the significant increment on *location*, *rooms* and *sleep* after the haze episode. The increase in ratings in all improvable categories after shock supports the "service quality effect" explanation that negative online reviews may prompt hotel managers to improve their service quality in order to raise their hotels' online reputation.

To understand whether managers' responses play a role in consumer satisfaction after shock, we propose the following testable hypotheses:

H1. Managers are more likely to respond to reviews of lower scores and to reviews of a haze-episode stay.

H2. Managers respond more positively to reviews of lower scores and to reviews of a haze-episode stay.

H3. Managers' responses positively influence the consumer satisfaction in the next period.

H1 and **H2** concern the behaviors of managers or hoteliers facing the negative reviews[16]. **H3** relates to the results of manager response[17] and points to the intuition of why hoteliers care about active responses to the negative reviews, providing suggestive evidence for the "service quality effect" explanation.

We examine **H1** and **H2** by estimating the following models respectively:

$$\begin{aligned} logit(Response_{i,j,k,t}) = &\alpha + \beta \cdot Haze_t + \phi \cdot Score_{i,j,k,t} + \gamma \cdot X_t + \upsilon \cdot Sentence_{i,j,k,t} + \\ &\mu_g + \delta_o + \eta_k + \zeta_j + \theta_{year} + \theta_{month} + \epsilon_{i,j,k,t} \end{aligned} \tag{3}$$

[16] Managers' responses to online reviews have become an important part of customer relationship management (Gu and Ye, 2014). As suggested by the reciprocation theory (Jones, 1966), managers can show that they listen to and appreciate their customers by responding to positive online reviews. Moreover, according to the service recovery theory (Wallin Andreassen, 2000), managers' responses to negative online reviews can help the management team to address service issues and improve customer satisfaction (Xie et al., 2014).

[17] Research indicates that managers' responses to online reviews impact business outcomes (Gu and Ye, 2014). In an analysis of the online reviews on *Ctrip.com*, a review and e-commerce website for travel goods in China, Gu and Ye (2014) find that consumer satisfaction increases when hotel managers respond to consumer complaints. In a study of the online reviews and manager responses on *TripAdvisor*, Xie et al. (2014) find that managers' responses to online reviews about the location of the hotel have a positive impact on the hotel's performance, whereas responses to reviews about the cleanliness of the hotel have a negative effect. While managers' responses to online reviews play an important role in customer satisfaction, managers do not respond to every online review and may choose to pay particular attention to specific types of reviews (Park and Allen, 2013).

$$ResponseScore_{i,j,k,t} = \alpha + \beta \cdot Haze_t + \phi \cdot Score_{i,j,k,t} + \gamma \cdot X_t + \upsilon \cdot Sentence_{i,j,k,t} + \mu_g + \delta_o + \eta_k + \zeta_j + \theta_{year} + \theta_{month} + \epsilon_{i,j,k,t} \tag{4}$$

where the dependent variable $Response_{i,j,k,t}$ in Equation (3), is a binary variable equal to 1 if a manager responds to a review rated by individual i for the stay in year-month t at hotel j on website k, and 0 otherwise. The dependent variable $ResponseScore_{i,j,k,t}$ in Equation (4) is the compound sentiment score extracted from managers' responses using the VADER sentiment analysis. Equation (4) is a subsample analysis of Equation (3) because only responded review contains the response text. Following Mudambi and Schuff (2010) and Liu and Park (2015), we include the count of sentences in a review to proxy the review depth in above Equations. Since only *TripAdvisor* and *Expedia* provide information on manager response, the samples used in Equations (3) and (4) are smaller than that used in the baseline analysis. The indicators and fixed effects are the same as in Equation (1).

Column (1) of Table 9 presents the results of estimating the logit model in Equation (3). The coefficient on $Shock^a$ is positive and statistically significant at 0.110, indicating that managers are more likely to respond to reviews covering a haze-episode stay. The coefficient on *Review Score* is statistically significant and negative, indicating that hotel managers are more likely to respond to reviews with lower scores. The results support **H1** and are consistent the results in Proserpio and Zervas (2017), which indicate that hotels tend to respond more to negative shocks. Column (2) reports the results of estimating Equation (4). Similar to the results in Column (1), $Shock^a$ carries a significantly positive coefficient, and *Review Score* bears a significantly negative coefficient. This suggests that managers' responses contain more positive words for reviews with lower scores and reviews during the pollution shock, which is supportive of **H2**. In general, the results in the first two columns provide evidence of crisis management to recover reputation after environmental shocks.

[Table 9 inserted here]

Next, we test **H3** to see whether managers' responses affect hotels' review scores in the next period using the following specification:

$$Score_{j,k,t} = \alpha + \beta \cdot Response_{j,k,t-1} + \phi \cdot Haze_t + \gamma \cdot X_{t,l} + \eta_k + \zeta_j + \theta_{year} + \theta_{month} + \epsilon_{j,k,t} \tag{5}$$

where the dependent variable is the average online review score for hotel j on website k in year-month t. $Response_{j,k,t-1}$ refers to the total number of responses or the response rate (the percentage of comments been responded) for hotel j on website k in the previous year-month $t-1$. The fixed effects of website, hotel, year, and month are controlled.

Columns (3) to (4) of Table 9 reports the results, with $Response_{j,k,t-1}$ in Column (3) being the logarithmic total number of responses and $Response_{j,k,t-1}$ in Column (4) being the response rate (in percentage). We find $Response_{j,k,t-1}$ is significantly positive in both columns, implying that responses to online reviews in the previous period lead to an increase in the online review score in the current period. The results indicate that managers' responses acknowledge the existing service issues, which allows management teams to address problems and increase customer satisfaction. Hoteliers largely improve their operations and services following a temporary reputation crisis caused by exogenous pollution events, with the effect being stronger in hotels whose managers closely monitor their online reviews.

Although the results in this section show that the hoteliers' active responses lead to the improvement of service quality and a resulting increase in review score, at this stage we are not able to quantify the effort devoted by the hoteliers due to the lack of data regarding hoteliers' operation cost. Nevertheless, the results in Tables 5 and 9 suggest that space exist for hoteliers to improve service quality.

Another possible explanation ("promotion effect") for the post-shock increases is that hotels offer discounted room prices following a haze shock in order to boost consumer satisfaction, since guests pay lower prices for the same service levels. In particular, lowering prices would influence the composition of guests and therefore the guests before and after a haze shock could be different. For instance, guests who previously stayed at three-star/economic hotels could choose to move up to four-star/luxury hotels at the same price level. Moreover, improving service level could be easier for hotels if occupation rates were affected during the haze episodes.

We employ a set of hotel performance data at district-daily level[18] to measure the impact of haze on hotel room prices and occupation rates. As shown in Appendix Table B5, we find that the haze shock did not have a significant impact on the average daily occupation rates and room prices, which contradicts the promotion explanation. In addition, the official statistics provided by the Singapore government[19] show that the monthly average room prices of all hotels and the number of visitor arrivals do not abruptly change during and after the haze period (as shown in Appendix Figures B4 and B5).

Thus, our results suggest that the improved online reputation is caused by better service quality rather than discounts or promotions after the air pollution shock.

6 Analysis of Competing Hypothesis

6.1 Exogeneity of the Pollution Shock: A Difference-in-differences Estimation

Although the pollution-shock approach allows for a strong causal interpretation of the relationship between air pollution and consumer satisfaction, we do not observe a counterfactual status as all hotels in Singapore have been exposed to the haze episodes. To measure the causal effects of

[18] The daily hotel performance indices are obtained from Smith Travel Research (STR), a data analytics company that tracks hotel performance in Singapore. We collect the daily hotel performance data, which include hotel room rate and occupancy rate, of a sample of 33,472 hotel rooms in 2015. The hotel sample accounts for about 76.8% of the hotel rooms in Singapore as of Jun 30, 2016. The hotels are categorized by price into upper mid-scale, mid-scale, and economy classes, and by the geographic area into four regions: Marina Bay, Sentosa, Orchard, and River Valley. 73% of the sample hotel rooms are located in the Central Region of Singapore. Hotels located outside of the Central region are excluded in the sample due to unavailability of data.

[19] Singapore Tourism Board (STB) provides monthly Gazetted Hotel Statistics, including the total room revenue, average room rate, average occupancy rate, gross occupied hotel rooms per month, etc., see `https://data.gov.sg/dataset/monthly-gazetted-hotel-statistics-summary` for more detail.

air pollution on online review scores, it is essential to simultaneously observe reviews on hotels affected by haze episodes and reviews on hotels unaffected by haze episodes. This drives us to generate proxies of the counterfactual outcomes using online review data of all hotels in Hong Kong as a control group to perform a difference-in-differences analysis. Three assumptions must hold in order to estimate the causal effect: first, the pollution shock is not caused by online review scores at baseline; second, hotels in Singapore (treatment) and hotels in Hong Kong (control) are comparable and have parallel trends in online reviews before the haze shock; third, composition of comparison groups is stable and no spillover effects. We collect hotel reviews, weather and pollution measures in Hong Kong and present the summary statistics in Appendix Table B1.

The first assumption is valid because the pollution shock in Singapore is caused by wildfires in Indonesian island rather than the changes in online reviews of hotels. Unlike Singapore, Hong Kong is free of severe haze, with a relatively stable monthly PSI mean value that is lower than 50 during the sample period[20]. For the second assumption, we show that Singapore and Hong Kong share great similarities in geological, economic, and cultural aspects[21]. Moreover, Panel A of Appendix Table B2 indicates that hotels in Singapore and Hong Kong are comparable in star rating; and Appendix Table B3 illustrates that the guests' characteristics, including country of origin and guest types, in both regions are similar[22]. In Appendix Figure B5, we also present the parallel trends of tourism statistics[23] such as number of reviews and number of international visitors arrival in two regions. Given the similarity in customer pool, hotel prices, and hotel rating distribution, as well as the business environment, we believe hotels in Singapore and Hong Kong are comparable. The parallel trends assumption is also by verified by plotting the unconditional trends of review scores in both regions, as shown in Figure 4, and formal tests on this assumption are discussed in Table 11. We address the competing hypotheses that threat the third assumption in Section 6.3.

[**Figure 4 inserted here**]

Our difference-in-differences specifications are as follows:

$$Score_{i,j,k,t} = \alpha + \beta \cdot Haze_t + \phi \cdot Treatment \cdot Haze_t + \gamma \cdot X_t + \mu_g + \delta_o + \zeta_j + \eta_k + \theta_t + \epsilon_{i,j,k,t} \tag{6}$$

[20]As shown in Table 2 and Table B1, both the monthly average and monthly maximum PSI readings in Singapore are around 1.5 times large as that in Hong Kong. The peak of the max PSI in Singapore reached 297, while the highest max PSI in Hong Kong just touched 87. For other weather data, the average temperature was 5.22 centigrade higher in Singapore than in Hong Kong, while Hong Kong experienced more rainfall and wind.

[21]Table B2 in the Appendix compares Hong Kong and Singapore with regional statistics. Both are Southeast Asian islands with a similar population size and population density. Hong Kong and Singapore are regional hubs for transport by air and sea, as well as being regional financial centers. Both are advanced economies with over 40,000 GDP per capita (PPP) in US dollars. Historically, Hong Kong and Singapore were colonies of the United Kingdom; and the economies in both regions started to rise rapidly in the 1970s to 1980s. The similarity between Singapore and Hong Kong relieves the possible treatment bias in the difference-in-differences estimations in the following sections.

[22]19 of the top 20 country of origin for the hotel guests in Singapore and Hong Kong are the same. The compositions of the guest types in Singapore and Hong are quite similar.

[23]Data source of Singapore: `https://data.gov.sg/dataset/total-visitor-international-arrivals-to-singapore`; Data source of Hong Kong: `https://www.discoverhongkong.com/ca/about-hktb/news/visitor-arrival.jsp`

$$Score_{i,j,k,t} = \alpha + \beta_{pre} \cdot Treatment \cdot Pre + \beta_{post} \cdot Treatment \cdot Post + \gamma \cdot X_t + \mu_g + \delta_o + \zeta_j + \eta_k + \theta_t + \epsilon_{i,j,k,t} \tag{7}$$

where Equation (6) studies the response of the online review scores during the haze shocks; and Equation (7) compares the changes of the online review scores before and after the haze shocks, excluding the period with haze shocks. Still, the sample period for estimating Equations (6) and (7) is from Jun 2012 to Aug 2014. More specifically, $Haze_t$ in Equation (6) represents different haze measures in year-month t, and $Treatment$ is a binary variable equal to 1 for hotels in Singapore, and is equal to 0 for hotels in Hong Kong. ϕ captures the causal effects of air pollution to scores. It is worth noting that we control for year-month fixed effect θ_t in Equations (6) and (7) instead of the year fixed effect and month fixed effect included in Equations (1) and (2), because monthly weather conditions are different in Singapore and Hong Kong. Other fixed effects are the same as those in Equation (1). The standard error is clustered at the hotel level.

Pre in Equation (7) is a binary variable equal to 1 for the six months (i.e., Dec 2012-May 2013) before the Jun 2013 haze shock, and $Post$ is a binary variable equal to 1 for twelve months after the haze shock. β_{post} captures the average post-shock responses of scores for hotels in Singapore (compared to the benchmark period, i.e., from Jun 2012 to Nov 2012), relative to the post-shock changes on online review scores of hotels in Hong Kong. On the other hand, β_{pre} measures the differences of scores between the treatment group and the control group during the six pre-shock months (compared to the benchmark period). For a robustness check, we also use two alternative Pre, Sep 2012-May 2013 and Feb 2013-May 2013. Validity of the difference-in-differences design assumes a parallel trend in online review scores of hotels in Singapore and Hong Kong before the shock, and requires β_{pre} to be statistically indistinguishable from 0.

Table 10 reports the results of estimating Equation (6) using various haze measurements, and the haze effects are captured by the interaction term $Treatment * Haze_t$. In all specifications, the interaction coefficient β is consistently estimated to be significantly negative. Column (1) suggests that review scores of hotels in Singapore are 0.418 points lower on a 0-to-10 scale during the haze episode, compared to review scores of hotels in Hong Kong, which were not affected by air pollution shocks. Columns (2) and (3) use a logarithm of monthly mean PSI and maximum PSI to measure the intensity of air pollutants, and show that the online review scores of hotels in Singapore decrease by 0.083 to 0.128 points relative to that in Hong Kong if PSI values double. In addition, as shown in Column (4), each additional hazy day decreases the review scores of hotels in Singapore by 0.061 points, on average, relative to that of Hong Kong. We also re-estimate Equation (6) for the sentiment compound score and report the results in Appendix Table B6. Consistent with the results in Table 10, the sentiment intensity extracted from the review text decreases more for travellers in Singapore than for those in Hong Kong during the haze period.

[Table 10 inserted here]

Table 11 reports the results of estimating Equation (7). Columns (1) to (3) present results using a different pre-trend duration and a different benchmark period; and the coefficients of the pre-treatment period variable $Treatment$*Pre are statistically insignificant with small magnitudes, suggesting that the parallel trend assumption is verified. The coefficients on $Treatment^{*}Post$ capture the online review scores responses after the haze shocks compared to the benchmark period. Overall, the ex-post responses on online review scores are statistically significant, implying that online review scores for hotels in Singapore rise by 0.228 to 0.254 points, on average, in the post-shock periods, relative to that in Hong Kong. We also report the effects on compound sentiment score in Appendix Table B7. We find both the coefficient of $Treatment^{*}Pre$ and the coefficient of $Treatment^{*}Post$ are not statistically significant, suggesting that consumers' mood recovers to the previous level when the haze episode ends.

[Table 11 inserted here]

We also plot the dynamic effect of the air pollution shock on review scores for a time window covering 6 months before and 12 months after the shock by repeating estimating the following specification:

$$\begin{aligned} &Score_{i,j,k,t} = \alpha + \beta_s \cdot Treatment \cdot Post_s + \gamma \cdot X_t + \mu_g + \delta_o + \zeta_j + \eta_k + \theta_t + \epsilon_{i,j,k,t} \\ &s = -6, -5, -4, ..., 0, ..., 10, 11, 12 \end{aligned} \tag{8}$$

where $Post_s$ is a binary variable equal to 1 for month s (-6, -5, -4,..., 0,..., 10, 11, 12) before/after June 2013. The coefficient β_s measures the difference in the response of review score between the treatment and control groups in each month, compared with that in the benchmark period (2012:06 - 2012:12). More specifically, the coefficient β_0 measures the immediate response in review score during the haze episode month. The coefficients $\beta_1, \ldots, \beta_{12}$ measure the responses in the first, $\ldots$, and twelfth month after the haze episode, respectively. Similarly, coefficients $\beta_{-1}, \ldots, \beta_{-6}$ capture the difference indifference of response in review scores between the treatment and control groups in each of the six pre-shock months.

Figure 5 presents the entire path of dynamic coefficients β_s (indicated by the solid line), where s=-6, -5,..., 11, 12. The dotted lines depict the corresponding 95% confidence interval. As previously noted, the patterns of review scores between hotels in Singapore and Hong Kong during the six-months pre-shock period are statistically insignificant. Review scores of hotels in Singapore drop substantially during the haze shock (in Jun 2013), but the drop is temporary; as Figure 5 shows, the average review score rises immediately after the haze shock by 0.502 points by the end of the second month after the Jun 2013 shock, compared to the average score six months before the shock[24]. It is important to note that the sharp rise in the average review score is short-lived and observed only in the first three months following the haze shock.

[Figure 5 inserted here]

[24] We conjecture that the reason the review score did not rise sharply in the first month ex post the haze shock is twofold: first, the guests did not write the review during the stay; second, the manager should first see the review and then supervise the staff to improve service. Both are reasons for the delay in the rise in review scores after the haze shock.

6.2 Causal Relationship between PSI and Online Reviews: An Instrumental Variable Approach

To address the concerns that PSI readings may capture pollution from local activities rather than the forest fires in Indonesia, we follow Sheldon and Sankaran (2017) and Agarwal et al. (2017) to estimate the causal relationship between air pollution and online review scores using satellite fire data in Indonesia to instrument for the transboundary haze pollution. The fire radiative power (FRP)[25] in megawatts (MW) from all Indonesian latitudes and longitudes between January 1, 2012 to Dec 31, 2015. We estimate the following two-stage least square equations:

$$PSI_t = \beta_1 \cdot FRP_t + X_t + \mu_g + \delta_o + \zeta_j + \eta_k + \theta_{year} + \theta_{month} + \epsilon_t \tag{9}$$

$$Score_{i,j,k,t} = \alpha + \beta_2 \cdot \widehat{PSI_t} + \gamma \cdot X_t + \mu_g + \delta_o + \zeta_j + \eta_k + \theta_{year} + \theta_{month} + \epsilon_{i,j,k,t} \tag{10}$$

where PSI_t is adjusted monthly 24-hour PSI reading for air pollution at month t. FRP_t is the monthly fire radiative power in Indonesia. $\widehat{PSI_t}$ is the predicted value of the PSI from the estimation of Equation (9). Other indicators and fixed effects are the same as those in Equation (1).

The results are reported in Appendix Table B8. Column (1) shows the first stage results of IV and indicates that Indonesian's FRP is a statistically significant determinant of PSI readings in Singapore. The estimation of Equation (10) shows that a 100% increase in the predicted PSI in Singapore causes the online review scores to decrease by 0.191 points, and reduces the compound sentiment score by 0.038 points. The results are consistent with that in the baseline analysis.

6.3 Selection Bias

To address the possibility the changes of customer composition may affect the review score, we construct a panel data of "frequent travellers", who visit Singapore (and Hong Kong) before and after a haze period. We identify the same customer using the information on reviewer ID, which is provided by *TripAdvisor*. We repeat the DID estimation on Equation (7) using both the review score and compound sentiment score as the outcome variables and present the results in Appendix Table B9. The results indicate that the review score increases by 0.320 points for Singapore hotels than that for Hong Kong hotels after the haze shock, while the treatment effect is insignificant on estimating the compound sentiment score. The results in Appendix Table B9 are in great consistency with the DID results (Table 11 and Appendix Table B7) using the full sample. Therefore, we conclude that the composition of hotel guests is unlikely to threaten our findings.

The second concern relates to the composition and quality of hotel may change over time. In particular, if guest prefer newly-built hotels and those were built after the haze shock, we would observe lower review scores during the haze shocks due to the absence of new hotels in the

[25] The data was obtained from the Fire Information for Resource Management System at National Aeronautics and Space Administration (NASA). We further aggregate the FRP data into monthly level.

sample. To address the potential sample selection bias, our sample only contains hotels built before 2012. Newly-built hotels are excluded in the analysis to avoid selection bias. Another alternative explanation is that only very few hotels show significant response to haze shocks, and the standard linear regression only estimates the average effect and hides effects at different quantiles of the distribution.

To provide more robust results against outliers in the responses measurement, we employ the quantile estimation (Koenker and Bassett Jr, 1982) to examine the Jun 2013 haze shock. As shown in Appendix Table B10, we find that the coefficient estimates of 25th, 50th, and 75th quantiles are all positive and statistically significant, suggesting hotels in different quantiles respond to the haze shock. The results of quantile regressions provide us with a more comprehensive picture of the change of review score after the haze episode and show that our conclusions are not driven by the outliers or abnormal distributions of the hotel responses.

Besides, as Luca and Zervas (2016) document, the credibility of online reviews is undermined as businesses have incentives to commit review fraud facing increased competition. We could not eliminate the possibility that hotels have incentives to fake good reviews or delete bad reviews to improve their reputation, which could threat the reliability of the review sample. Since we still observe significant drop on review scores during the haze periods, faking or deleting reviews is less likely to be an issue in our study.

7 Heterogeneity Tests

In this section, we present the heterogeneity tests across traveller types, travellers' continent of origin, and hotel star rating. Specifically, Table 12 shows the heterogeneity in the responses of online review scores to the haze shock across traveller types and travellers' continent of origin. As shown in Column (1), the interaction term, $Shock^{a}$*Business, is significantly negative, suggesting that business travellers are more sensitive to the haze shock relative to non-business ones. The severe haze shock led to a 0.470 points decrease in business travellers' review score, compared to a 0.396 points decrease in non-business travellers' review score. However, as shown in Column (2), the increment of review scores of business travellers is 0.074 points higher than that of non-business travellers after the air pollution shock.

In the heterogeneity tests by continent of origin (Columns 3 and 4), the Asian travellers serve as the reference group. Columns (3) and (4) shows that travellers from Europe are the most environmentally sensitive travellers, which is reflected by the largest decrease (-0.500) of review scores during the haze episode and the largest increase (0.423) of review score after the haze episode.

[Table 12 inserted here]

Managers' response to online reviews varies dramatically among hotels, regardless whether the hotels have similar guest ratings or whether the reviews are positive or negative (Levy et al., 2013; Park and Allen, 2013). Table 13 presents the heterogeneity in the responses of online review scores to the haze shock across hotel star ratings (Column 1), the heterogeneity in the responses of online review scores ex-post the haze shock across hotel star ratings (Column 2), and the heterogeneity in the manager response across hotel star ratings (Column 3). Column (1) shows that drops in the review scores of three-star hotels are greater than those of the reference group, one-star and two-star hotels, and the luxury four- and five-star hotels during the haze period. In Column (2), we see that the increment of review scores of four-star hotels after the haze shock appears to be the largest. Column (3) indicates that managers of four- and five-star hotels are more likely to respond to the reviews that cover for a haze-episode stay than managers of other hotels. This suggests luxury hotels are more concerned about loss of reputation due to haze shocks.

[**Table 13 inserted here**]

8 The Back-of-the-Envelope Calculation

An exogenous air pollution event in Singapore provides us with an opportunity to explore how consumers react to negative pollution shocks, and how the service sector might react to a temporary reputation crisis by improving service quality. Specifically, the types of responses in the review input from consumers suggest that their mood during the period of the visit was negatively affected by the pollution event rather than by concerns related to the quality of service they received. The substantial damage to the online reputations then triggers hotels to improve their service quality after the event, which subsequently raises customer satisfaction level, creating a virtuous circle that generates substantial economic gains.

Existing research has strived to utilize economic analysis to estimate the dollar benefits of air quality improvements. One approach is to proxy willingness to pay by measuring the additional cost to society from diminished air quality (Deschênes et al., 2017). The results in this paper allow us to conduct a simple back-of-the-envelope calculation on the change of subjective utility in terms of the price paid for the stay of a customer during and after the haze period, which can be measured through the changes of online review scores.

The cost-benefit analysis calculates the welfare losses during air pollution period and the welfare gains from the improved services that triggered by the negative shock using the following parameters: the online review scores dropped substantially by around 0.409 points on average during the Jun 2013 haze episode (Tables 3), and then increase by around 0.295 points after the haze shock (Tables 4). In addition, the aggregate statistics from Singapore Tourism Board and regression results using micro-level data (Appendix Table B5) show that the room price barely changes during the sample period.

We define the welfare gain (or loss) from the increase (or decrease) in the guests' subjective well-being as follows:

$$Average\ Gain = \frac{\sum_{t=0}^{n} Revenue_t \times \frac{\Delta Score_t}{Score_{sg}}}{\sum_{t=0}^{n} Room_t} \tag{11}$$

where t indexes the year-month. $Revenue_t$ stands for the total room revenue[26] of the hospitality

[26] Singapore Tourism Board (STB) provides monthly Gazetted Hotel Statistics, including the total room revenue,

industry in Singapore in year-month t. $\Delta Score_t$ refers to the coefficients obtained from Tables 3 and 4. $Score_{sg}$ stands for a constant number 7.67, which is the average review score of hotels in Singapore before the June 2013 haze shock (from Jun 2012 to May 2013). Therefore, $\frac{\Delta Score_t}{Score_{sg}}$ stands for the percentage change of the online review scores in year-month t relative to the score during the pre-shock period. $\sum_{t=0}^{n} Revenue_t \times \frac{\Delta Score_t}{Score_{sg}}$ estimates the total welfare change in the form of total room revenue between time t and time n. Dividing the total welfare change by the number of gross occupied rooms between time t and time n, $\sum_{t=0}^{n} Room_t$, we get the average welfare change per room per night.

To illustrate, in Jun 2013, the relative change of score (in percentage) is $\frac{-0.409}{7.67} = -5.33\%$, and the total revenue[27] is S\$246.5 million with an average room price of S\$259. Ideally, in the absence of the air pollution shock, a guest should enjoy an average subjective well-being of 7.67 by paying S\$259. Due to the haze shock, the travellers' subjective well-being or relative utility decreases by 5.33% compared to the counter-factual state, but they are still paying for S\$259 per stay. Therefore, the average welfare loss for one occupied room per night during the haze episode in Jun 2013 is $5.33\% * 259 = S\$13.80$. And the total welfare losses in Jun 2013 are calculated as: $13.80 * 951,709 = S\$13.13\, million$.

Similarly, we compute the welfare gains reflected by the guests' increased subjective sense of well-being for the 12 post-shock months (from Jul 2013 to Aug 2014). For instance, thanks to the improvements in service levels that took place after the haze shock, an guest enjoys $(0.295/7.67) = 3.85\%$ more of the benchmark subjective well-being of 7.67 by paying S\$290 per night[28]. The welfare gain per stay in the 12 post-shock months is $3.85\% * 290 = S\$11.16$ and the total welfare gain for the 12 months is $11.16 * 12,427,085 = S\$138.68\, million$

In summary, one guest (assuming a guest occupies one room) experiences a S\$13.80 welfare loss per night in Jun 2013 as a result of the haze shock, but then receives a S\$11.16 welfare gain per night during the 12 months that follow.

9 Conclusion

In this study, we examine the impacts of exogenous air pollution shocks on the performance of firms in the service sector and the subjective well-being of customers. Using data obtained from four prominent hotel-booking websites, the study shows that consumer satisfaction decreased significantly during haze periods, and then immediately reverted to previous levels in the following month. More interestingly, the review score continued to rise sharply in the second month after the haze episode, before starting to decline and finally reach a plateau with a score a little higher

average room rate, average occupancy rate, gross occupied hotel rooms per month, etc., see `https://data.gov.sg/dataset/monthly-gazetted-hotel-statistics-summary` for more detail.

[27] The statistics provided by the Singapore Tourism Board show that the number of room occupied in June 2013 is 951,709. Source: `https://data.gov.sg/dataset/monthly-gazetted-hotel-statistics-summary`

[28] The statistics provided by the Singapore Tourism Board show that the total room revenue and number of room occupied between Jul 2013 to Aug 2014 are S\$3602.9 million and 12,427,085, respectively. Source: `https://data.gov.sg/dataset/monthly-gazetted-hotel-statistics-summary`

than the original level. Exogenous environmental shocks, such as serious haze episodes, significantly reduce the satisfaction level of tourists and negatively impact the online reputation of hotels during guests' stays. The lower consumer satisfaction can be attributed to the negative moods of guests, rather than to any lowering in the quality of hotel service during the air pollution shocks. We verify this explanation by exploiting two machine learning techniques, sentiment analysis and topic modelling. Then, we provide solid evidence to show that the increase of review score in the post-haze period is explained by the improved service quality other than by the decreased room price. Furthermore, this response appears to be stronger in the cases of hoteliers closely monitoring their online reviews during the haze episodes.

We also address the possible endonegeity concern in a twofold strategy. First, we include the reviews of all hotels in Hong Kong, which is free of haze shocks in the sample period, as a control group and conduct the difference-in-differences analysis. Second, we use an instrumental variable approach to address that our PSI measure may capture the pollution at local level. We then address the guest selection bias by creating a subsample of frequent travellers and address the hotel selection bias by excluding hotels built before 2012. To provide more robust results against outliers or abnormal distributions of the hotel responses, we employ the quantile estimation (Koenker and Bassett Jr, 1982) and find that the coefficient estimates of 25th, 50th, and 75th quantiles are all positive and statistically significant, suggesting hotels in different quantiles respond to the haze shock.

In addition, we demonstrate that managers are more likely to respond to reviews with lower scores and to reviews that cover a stay during a haze period. Our analysis indicates that responses to reviews contribute to a rise in review scores during the next period. The results, together with the subcategory results, suggest a channel used by hotels to improve their service level following a haze episode.

A growing body of literature has proved the negative impacts of various types of pollution on economic growth (Jorgenson and Wilcoxen, 1990), productivity (Chang et al., 2016), health (Deschênes et al., 2017), housing price (Chay and Greenstone, 2005), and many other areas. Our investigation focuses on the haze effects to the service sector and highlights another direction for pollution studies. We provide empirical evidence to show that detrimental environmental shocks could trigger positive outcomes by spurring hotels to improve service quality. This paper has policy implications related to deficiencies in management for firms as well as regulators. The estimation of the welfare gains indicates that travellers in Singapore on average enjoy a service improvement estimated at S$11.16 per room per night in the subsequent 12 month following a haze shock.

References

Agarwal, S., T. F. Sing, and Y. Yang (2017). Impact of transboundary haze pollution on household utilities consumption. *Georgetown McDonough School of Business Research Paper* (2942096).

Aggarwal, R., R. Gopal, R. Sankaranarayanan, and P. V. Singh (2012). Blog, blogger, and the firm: Can negative employee posts lead to positive outcomes? *Information Systems Research 23*(2), 306–322.

Ba, S. and P. A. Pavlou (2002). Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly*, 243–268.

Bansal, H. S., S. F. Taylor, and Y. St. James (2005). "migrating" to new service providers: Toward a unifying framework of consumers' switching behaviors. *Journal of the Academy of Marketing Science 33*(1), 96–115.

Blanke, J. and T. Chiesa (2013). The travel & tourism competitiveness report 2013. In *The World Economic Forum*.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent dirichlet allocation. *Journal of machine Learning research 3*(Jan), 993–1022.

Boyd, B. K., D. D. Bergh, and D. J. Ketchen Jr (2010). Reconsidering the reputation—performance relationship: A resource-based view. *Journal of Management 36*(3), 588–609.

Cai, X., J. Gong, Y. Lu, and S. Zhong (2017). Recover overnight? work interruption and worker productivity. *Management Science 1*(2), 28–57.

Chang, T., J. Graff Zivin, T. Gross, and M. Neidell (2016). Particulate pollution and the productivity of pear packers. *American Economic Journal: Economic Policy 8*(3), 141–69.

Chaves, M. S., R. Gomes, and C. Pedron (2012). Analysing reviews in the web 2.0: Small and medium hotels in portugal. *Tourism Management 33*(5), 1286–1287.

Chay, K. Y. and M. Greenstone (2003). The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. *Quarterly Journal of Economics 118*(3), 1121–1167.

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

Cheema, A. and P. Papatla (2010). Relative importance of online versus offline information for internet purchases: Product category and internet experience effects. *Journal of Business Research 63*(9-10), 979–985.

Chen, Y. and J. Xie (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Science 54*(3), 477–491.

Chevalier, J. A. and D. Mayzlin (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research 43*(3), 345–354.

Coviello, D., A. Ichino, and N. Persico (2014). Time allocation and task juggling. *American Economic Review 104*(2), 609–23.

Dell, M., B. F. Jones, and B. A. Olken (2014). What do we learn from the weather? the new climate-economy literature. *Journal of Economic Literature 52*(3), 740–98.

Deschênes, O., M. Greenstone, and J. S. Shapiro (2017). Defensive investments and the demand for air quality: Evidence from the nox budget program. *American Economic Review 107*(10), 2958–89.

Dominici, F., M. Greenstone, and C. R. Sunstein (2014). Particulate matter matters. *Science 344*(6181), 257–259.

Duan, W., B. Gu, and A. B. Whinston (2008). The dynamics of online word-of-mouth and product sales—an empirical investigation of the movie industry. *Journal of Retailing 84*(2), 233–242.

Evans, G. W., S. V. Jacobs, D. Dooley, and R. Catalano (1987). The interaction of stressful life events and chronic strains on community mental health. *American Journal of Community Psychology 15*(1), 23–34.

Graff Zivin, J. and M. Neidell (2012). The impact of pollution on worker productivity. *American Economic Review 102*(7), 3652–73.

Grönroos, C. and K. Ojasalo (2004). Service productivity: Towards a conceptualization of the transformation of inputs into economic results in services. *Journal of Business Research 57*(4), 414–423.

Gu, B. and Q. Ye (2014). First step in social media: Measuring the influence of online management responses on customer satisfaction. *Production and Operations Management 23*(4), 570–582.

Guo, Y., S. J. Barnes, and Q. Jia (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management 59*, 467–483.

Han, H. J., S. Mankad, N. Gavirneni, R. Verma, et al. (2016). What guests really think of your hotel: Text analytics of online customer reviews.

Hanna, R. and P. Oliva (2015). The effect of pollution on labor supply: Evidence from a natural experiment in mexico city. *Journal of Public Economics 122*, 68–79.

He, J., H. Liu, and A. Salvo (2019). Severe air pollution and labor productivity: Evidence from industrial towns in china. *American Economic Journal: Applied Economics 11*(1), 173–201.

Hennig-Thurau, T., K. P. Gwinner, G. Walsh, and D. D. Gremler (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing 18*(1), 38–52.

Herrmann, M. A. and J. E. Rockoff (2012). Worker absence and productivity: Evidence from teaching. *Journal of Labor Economics 30*(4), 749–782.

Hirshleifer, D. and T. Shumway (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance 58*(3), 1009–1032.

Holt, C. A. and S. K. Laury (2002). Risk aversion and incentive effects. *American Economic Review 92*(5), 1644–1655.

Hornbeck, R. and D. Keniston (2017). Creative destruction: Barriers to urban growth and the great boston fire of 1872. *American Economic Review 107*(6), 1365–98.

Huang, A. H., R. Lehavy, A. Y. Zang, and R. Zheng (2017). Analyst information discovery and interpretation roles: A topic modeling approach. *Management Science 64*(6), 2833–2855.

Hutto, C. J. and E. Gilbert (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*.

Jayachandran, S. (2009). Air quality and early-life mortality evidence from indonesia's wildfires. *Journal of Human Resources 44*(4), 916–954.

Jones, S. C. (1966). Some determinants of interpersonal evaluating behavior. *Journal of Personality and Social Psychology 3*(4), 397.

Jorgenson, D. W. and P. J. Wilcoxen (1990). Environmental regulation and us economic growth. *The Rand Journal of Economics*, 314–340.

Kim, C., S. H. Jung, D. R. Kang, H. C. Kim, K. T. Moon, N. W. Hur, D. C. Shin, and I. Suh (2010). Ambient particulate matter as a risk factor for suicide. *American journal of psychiatry 167*(9), 1100–1107.

Koenker, R. and G. Bassett Jr (1982). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica*, 43–61.

Larcom, S., F. Rauch, and T. Willems (2017). The benefits of forced experimentation: striking evidence from the london underground network. *Quarterly Journal of Economics 132*(4), 2019–2055.

Lee, J. S. H., Z. Jaafar, A. K. J. Tan, L. R. Carrasco, J. J. Ewing, D. P. Bickford, E. L. Webb, and L. P. Koh (2016). Toward clearer skies: challenges in regulating transboundary haze in southeast asia. *Environmental Science & Policy 55*, 87–95.

Lee, T. Y. and E. T. Bradlow (2011). Automated marketing research using online customer reviews. *Journal of Marketing Research 48*(5), 881–894.

Levy, S. E., W. Duan, and S. Boo (2013). An analysis of one-star online reviews and responses in the washington, dc, lodging market. *Cornell Hospitality Quarterly 54*(1), 49–63.

Levy, T. and J. Yagil (2011). Air pollution and stock returns in the us. *Journal of Economic Psychology 32*(3), 374–383.

Liu, Z. and S. Park (2015). What makes a useful online review? implication for travel product websites. *Tourism Management 47*, 140–151.

Loughran, T. and B. McDonald (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research 54*(4), 1187–1230.

Luca, M. and G. Zervas (2016). Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science 62*(12), 3412–3427.

Mankad, S., H. an, J. Goh, and S. Gavirneni (2016). Understanding online hotel reviews through automated text analysis. *Service Science 8*(2), 124–138.

Mizerski, R. W. (1982). An attribution explanation of the disproportionate influence of unfavorable information. *Journal of Consumer Research 9*(3), 301–310.

Mudambi, S. M. and D. Schuff (2010). Research note: What makes a helpful online review? a study of customer reviews on amazon. com. *MIS Quarterly*, 185–200.

Park, S.-Y. and J. P. Allen (2013). Responding to online reviews: Problem solving and engagement in hotels. *Cornell Hospitality Quarterly 54*(1), 64–73.

Proserpio, D. and G. Zervas (2017). Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science 36*(5), 645–665.

Puranam, D., V. Narayan, and V. Kadiyali (2017). The effect of calorie posting regulation on consumer opinion: A flexible latent dirichlet allocation model with informative priors. *Marketing Science 36*(5), 726–746.

Qadri, S. T. (2001). *Fire, smoke, and haze: The ASEAN response strategy.* Asian Development Bank.

Quah, E. (2002). Transboundary pollution in southeast asia: the indonesian fires. *World Development 30*(3), 429–441.

Roberts, P. W. and G. R. Dowling (2002). Corporate reputation and sustained superior financial performance. *Strategic Management Journal 23*(12), 1077–1093.

Schuckert, M., X. Liu, and R. Law (2015). Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing 32*(5), 608–621.

Sheldon, T. L. and C. Sankaran (2017). The impact of indonesian forest fires on singaporean pollution and health. *American Economic Review 107*(5), 526–29.

Statistics Singapore, S. T. B. (2018). Singapore annual report on tourism statistics 2017. *Singapore: Singapore Tourism Board*.

Thompson, C. J. (2005). Consumer risk perceptions in a community of reflexive doubt. *Journal of Consumer Research 32*(2), 235–248.

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives 28*(2), 3–28.

Vermeulen, I. E. and D. Seegers (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management 30*(1), 123–127.

Wallin Andreassen, T. (2000). Antecedents to satisfaction with service recovery. *European Journal of Marketing 34*(1/2), 156–175.

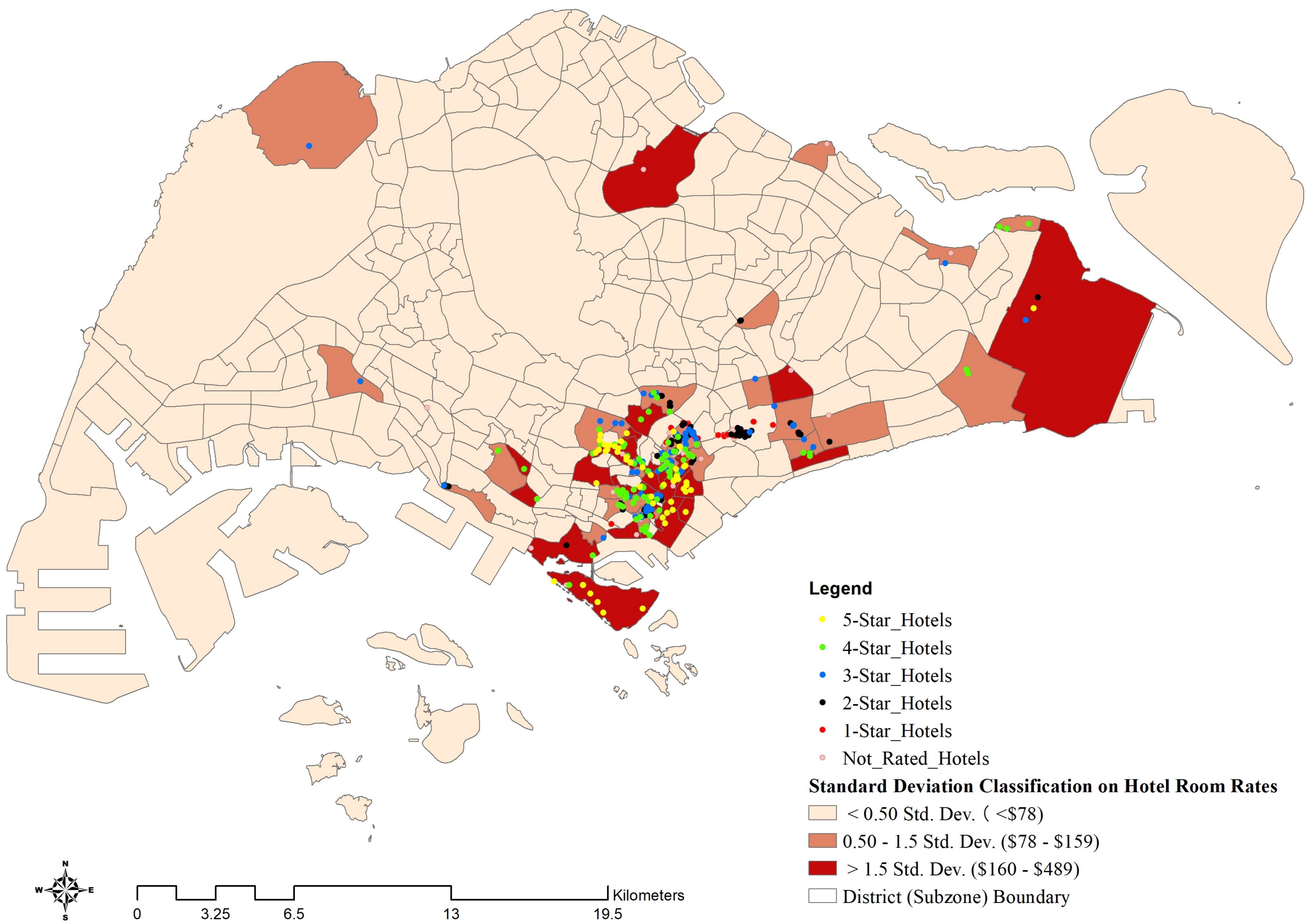
Xie, K. L., Z. Zhang, and Z. Zhang (2014). The business value of online consumer reviews and management response to hotel performance. *International Journal of Hospitality Management 43*, 1–12.

Ye, Q., R. Law, B. Gu, and W. Chen (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior 27*(2), 634–639.

Zheng, S., J. Wang, C. Sun, X. Zhang, and M. E. Kahn (2019). Air pollution lowers chinese urbanites' expressed happiness on social media. *Nature Human Behaviour 3*(3), 237.

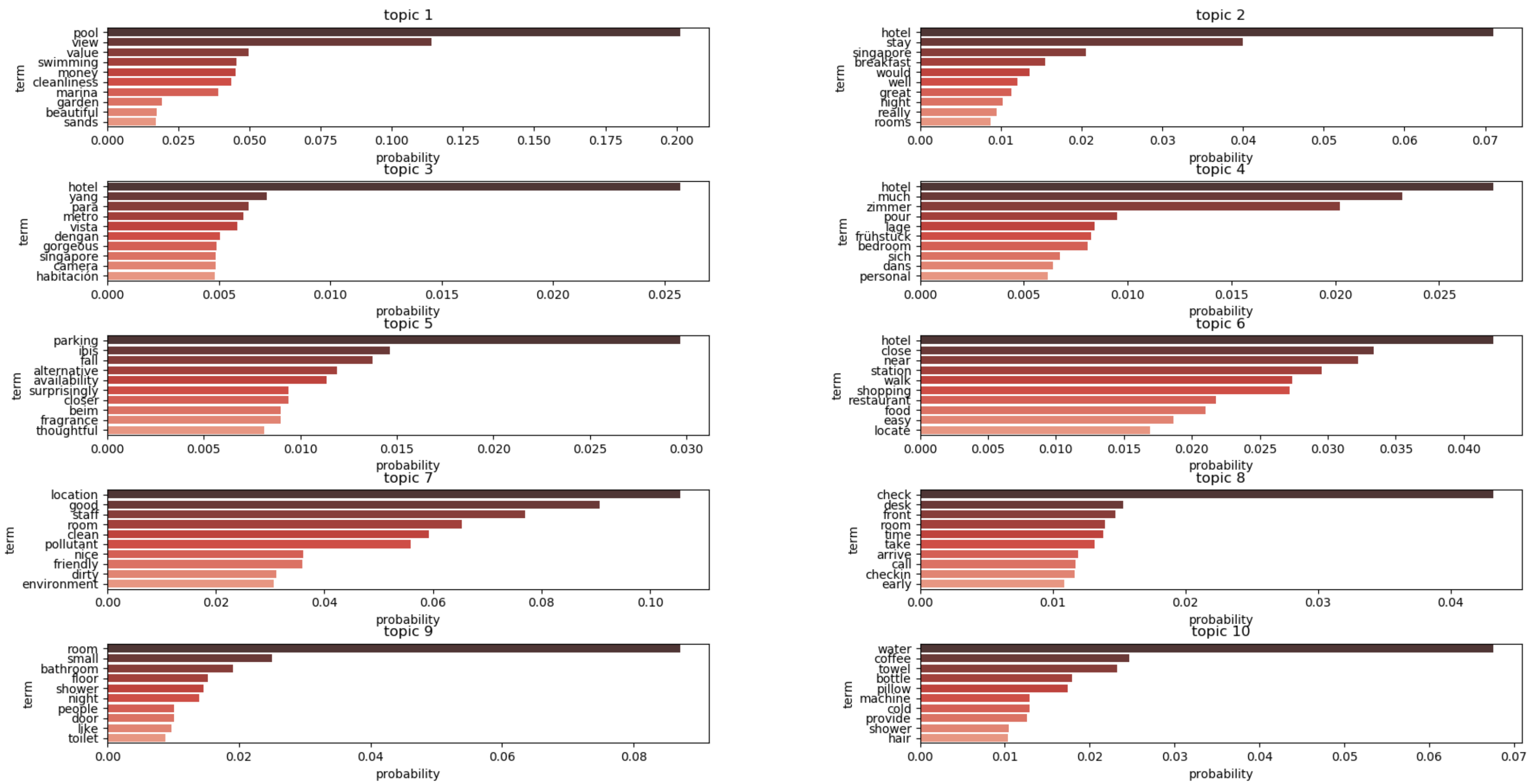
Zivin, J. G. and M. Neidell (2009). Days of haze: Environmental information disclosure and intertemporal avoidance behavior. *Journal of Environmental Economics and Management 58*(2), 119–128.

Figure 1: Geographic Distributions of Hotels by Star Rating in Singapore



Notes: This figure plots the geographic distributions of hotels and hostels in Singapore by star rating. We plot the quantile distribution by hotel daily room prices at the neighborhood (sector) level.

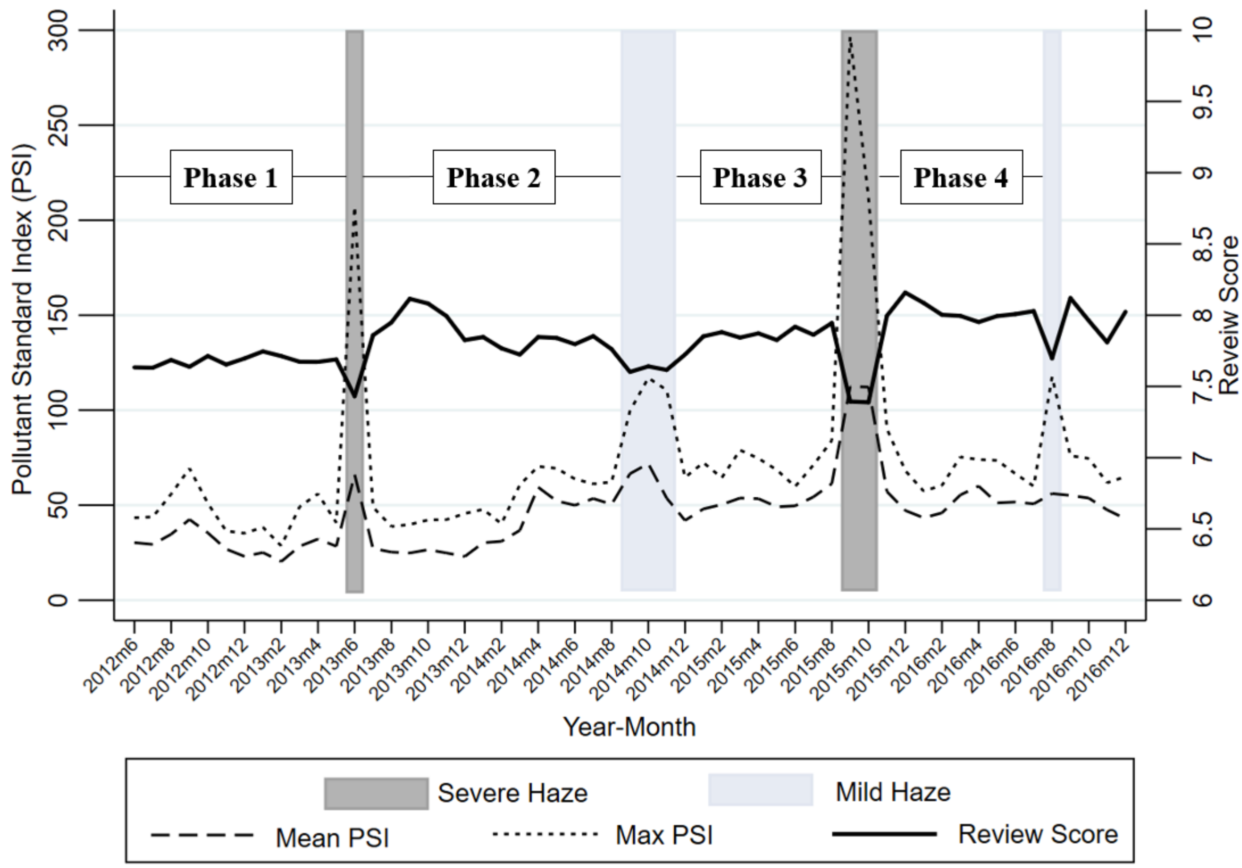
Figure 2: The 10 Most Likely Words in Each Topic in Singapore Reviews



Notes: This figure presents the 10 most likely used words in each of the 10 topics extracted from the review texts using the LDA program. Tocp 7 is considered environment-related due to its frequent use of keywords: pollution, dirty, and environment.

Figure 3: Monthly Trends of Air Pollution Measures and Online Review Scores

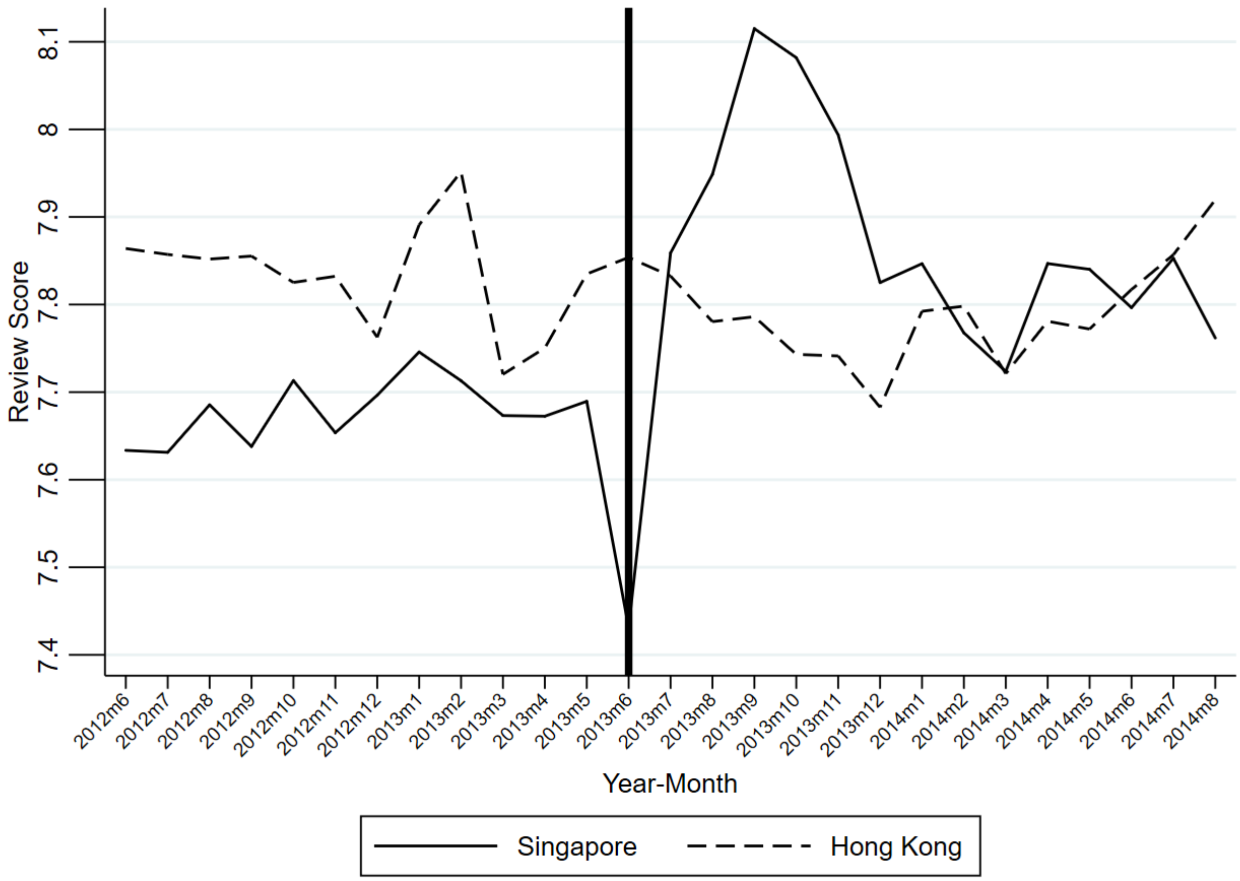
(Singapore, Jun 2012 - Dec 2016)



Notes: This figure plots the monthly trends of mean PSI (dashed line) and maximum PSI (dotted line), as well as the average online review score (solid line) of hotels from Jun 2012 to Dec 2016 in Singapore. The dark shaded areas and large spikes highlight two strong haze shocks in Jun 2013 and Sep-Oct 2015, while the light shaded areas represent two mild haze episodes in Oct 2014 and Aug 2016. The width of shaded areas indicates the duration of air pollution events. We divide the sample period into four phases: Jun 2012 to May 2013 (phase 1, the pre-shock period of the Jun 2013 haze), Jul 2013 to Aug 2014 (phase 2, the post-shock period of the Jun 2013 haze), Dec 2014 to Aug 2015 (phase 3, the pre-shock period of the Sep-Oct 2015 haze), and Nov 2015 to Dec 2016 (phase 4, the post-shock period of the Sep-Oct 2015 haze).

Figure 4: Unconditional Trends of Review Scores

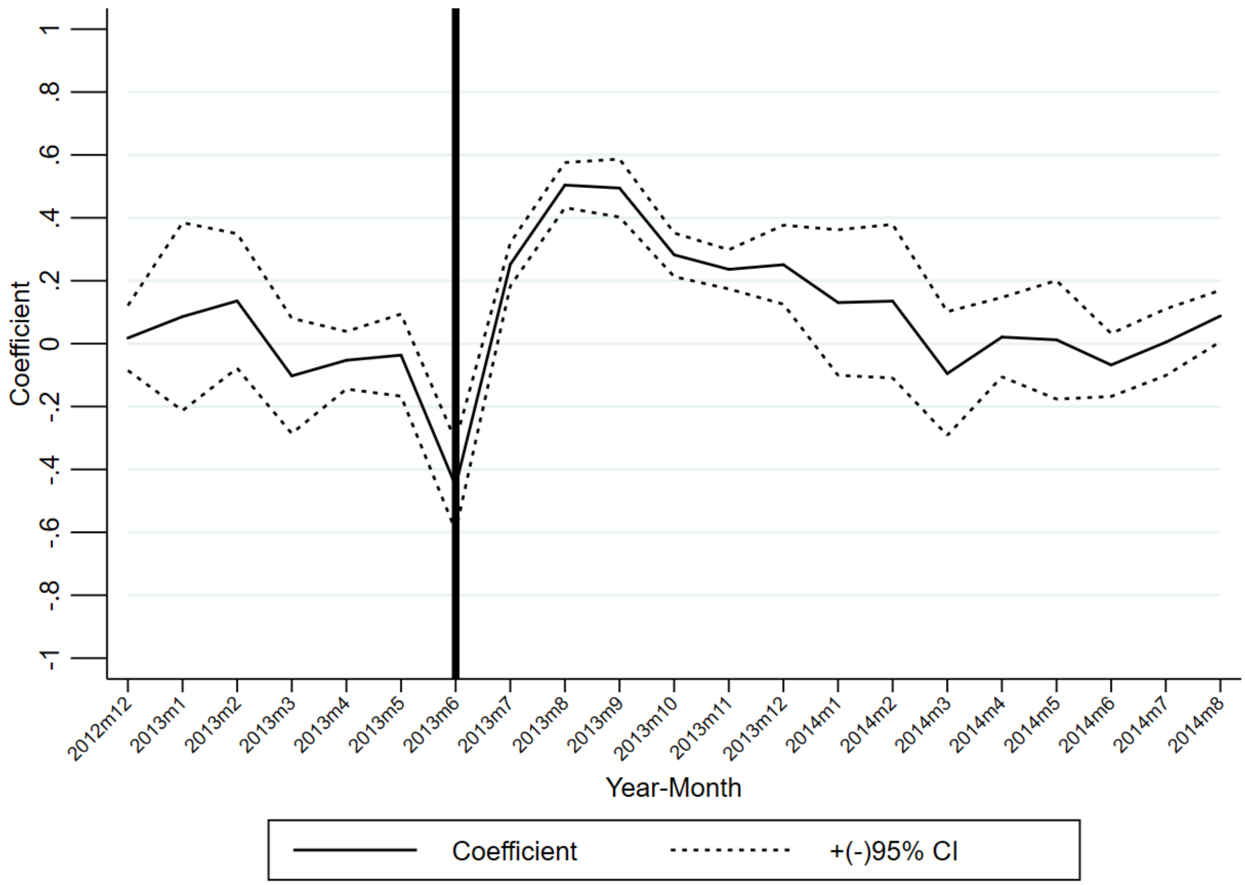
(Singapore and Hong Kong, Jun 2012 - Agu 2014)



Notes: This figure plots the unconditional trends of review scores in Singapore and Hong Kong. The vertical line indicates the sever haze shock took place in June 2013.

Figure 5: Estimated Response Dynamics of Review Scores

(Singapore and Hong Kong, Dec 2012 - Aug 2014)



Notes: This figure plots the entire path of dynamic response of the online review scores following Equation (8), along with its corresponding 95 percent confidence intervals. The x-axis denotes the year-month and y-axis shows the estimated coefficient of β_s. The vertical line indicates the sever haze shock took place in June 2013.

Table 1: The Example of Sentiment Analysis

review text	Negative	Neutral	Positive	Compound Score
Weather is bad, but the hotel is great.	0.151	0.470	0.379	0.6597
Weather is bad, but the hotel is really great.	0.138	0.490	0.373	0.7040
Weather is bad, but the hotel is really great !	0.135	0.481	0.384	0.7296
Weather is bad, but the hotel is really GREAT !	0.127	0.451	0.422	0.8037
Weather is bad, but the hotel is really GREAT :)	0.103	0.366	0.531	0.9073
Weather is bad, but the hotel is really GREAT, ☺	0.096	0.342	0.562	0.9313
Weather kinda sux, but the hotel is really GREAT, lol.	0.077	0.382	0.541	0.9136

Notes: This table reports the sentiment outcomes of seven sample sentences with different features (i.e., the specific punctuation, capitalization, degree modifiers, conjunctions, emojis, slang and emoticons in sentences).

Table 2: Descriptive Staitstics

Panel A: Online Reviews and Manager Responses for Hotels in Singapore

Variables	Obs.	Mean	Median	S.D	Min	Max
Review Score	869,115	7.83	8.00	1.80	1.30	10.00
Compound Score	869,115	0.44	0.52	0.46	-1.00	1.00
Positive Proportion	869,115	0.25	0.20	0.23	0.00	1.00
Negative Proportion	869,115	0.04	0.00	0.08	0.00	1.00
Neutral Proportion	869,115	0.71	0.75	0.23	0.00	1.00
Topic7 (Environment)	869,115	0.08	0.00	0.28	0	1
# of Sentences	869,115	5.11	2.00	8.29	1	610
Manager Response	275,658	0.49	0.00	0.50	0	1

Panel B: Monthly Weather and Pollution Measures in Singapore

PSI_{max}	55	72.43	63.80	45.73	28.60	296.60
PSI_{mean}	55	46.08	48.02	18.57	20.33	112.44
Days of Haze (DOH)	55	1.49	0.00	3.94	0.00	20.00
Temperature	55	31.62	31.70	0.80	29.60	33.40
Rainfall	55	156.61	126.60	92.43	0.20	395.20
WindSpeed	55	10.49	10.22	3.25	4.18	18.84
FRP	43	26.78	26.95	5.92	16.66	41.40

Notes: This table presents a summary description of the data, with online review data at the individual level for hotels in Singapore reported in Panel A, and ambient conditions in Singapore on a monthly basis reported in Panel B. The sample is from Jun 2012 to Dec 2016 in Singapore. Please refer to Appendix A for the detailed definition of variables.

Table 3: Responses of Review Score to the June 2013 Haze Shock

(Singapore, Jun 2012 - Aug 2014)

Dep. Variable	Review Score				
Haze Measure	$Shock^a$	$ln(PSI^{mean})$	$ln(PSI^{max})$	DoH	PSI Category
Model	(1)	(2)	(3)	(4)	(5)
Haze	-0.409***	-0.496***	-0.297***	-0.040***	
	(0.026)	(0.030)	(0.017)	(0.002)	
PSI_{51-100}					-0.182***
					(0.016)
PSI_{100+}					-0.484***
					(0.027)
ln(Temperature)	0.035**	0.053***	-0.021	-0.015	0.043**
	(0.015)	(0.018)	(0.015)	(0.015)	(0.017)
ln(Rainfall)	-0.015***	-0.012**	-0.014***	-0.012***	-0.021***
	(0.004)	(0.005)	(0.004)	(0.004)	(0.006)
ln(WindSpeed)	0.200***	0.138***	0.271***	0.211***	0.232***
	(0.027)	(0.027)	(0.028)	(0.027)	(0.027)
Constant	7.672***	6.616***	8.187***	7.072***	5.072***
	(0.477)	(0.521)	(0.466)	(0.472)	(0.538)
Observations	363,976	363,976	363,976	363,976	363,976
R-squared	0.204	0.204	0.204	0.204	0.204
Year FE and Month FE	Yes	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes	Yes	Yes
Website FE	Yes	Yes	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the regression results of estimating Equation (1) using hotel online review data in Singapore from Jun 2012 to Aug 2014. The dependent variable is review score. The headers in the second row indicate the measures of air pollution. $Shock^a$ represents the severe shocks in Jun 2013; PSI^{mean} and PSI^{max} represent the monthly average PSI readings and monthly maximum PSI readings, respectively; DoH measures the number of days with haze status per month; and $PSICategory$ classifies the monthly average PSI^{max} into three categories: 0-50, 51 to 100, and above 100. Please refer to Appendix A for the detailed definition of variables. Fixed effects of year, month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table 4: The Changes on Review Score before and after the Severe Haze Shocks

(Singapore, Jun 2012 - Dec 2016)

Dep. Variable	Review Score			
Event	**Jun 2013 Haze**		**Sep-Oct 2015 Haze**	
Post vs. Pre Period	phase 2 vs. phase 1	phase 3 vs. phase 1	phase 4 vs. phase 3	phase 4 vs. phase 1
Model	(1)	(2)	(3)	(4)
Post	0.295***	0.050	0.251***	0.144***
	(0.032)	(0.037)	(0.023)	(0.038)
ln(Temperature)	0.058***	-0.007	0.026*	0.044***
	(0.019)	(0.027)	(0.014)	(0.014)
ln(Rainfall)	-0.025***	0.010	0.006	-0.008
	(0.006)	(0.015)	(0.008)	(0.009)
ln(WindSpeed)	-0.057	0.090**	0.229***	0.064**
	(0.037)	(0.045)	(0.067)	(0.028)
Constant	5.463***	6.949***	7.965***	6.275***
	(0.532)	(0.867)	(0.596)	(0.669)
Observations	348,661	274,869	431,631	454,842
R-squared	0.203	0.198	0.192	0.193
Year FE and Month FE	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes	Yes
Website FE	Yes	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes	Yes

Notes: This table provides the results of estimating Equation (2) using different combinations of periods. The coefficients on *Post* show the differences between the online review scores before and after two severe haze shocks. The dependent variable is review score. We divide the sample period into four phases: Jun 2012 to May 2013 (phase 1, the pre-shock period of the Jun 2013 haze), Jul 2013 to Aug 2014 (phase 2, the post-shock period of the Jun 2013 haze), Dec 2014 to Aug 2015 (phase 3, the pre-shock period of the Sep-Oct 2015 haze), and Nov 2015 to Dec 2016 (phase 4, the post-shock period of the Sep-Oct 2015 haze). The headers in the third row indicate the combinations of periods used for regression. Please refer to Appendix A for the detailed definition of variables. Fixed effects of year, month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table 5: Changes on Subcategory Review Scores

(Singapore, Jun 2012 - Aug 2014)

Panel A: Relationship of Subcategory Review Scores and Air Pollution

Dep. Variable	Cleanliness	Service	Value	Location	Rooms	Sleep
Model	(1)	(2)	(3)	(4)	(5)	(6)
Shock[a]	-0.058	-0.065	-0.390***	-0.253***	-0.184***	-0.087
	(0.050)	(0.060)	(0.070)	(0.057)	(0.068)	(0.072)
Observations	50,876	55,559	50,673	50,790	50,629	50,491
R-squared	0.193	0.159	0.118	0.237	0.248	0.178

Panel B: The Ex-post Responses of Subcategory Review Scores to the Haze Shocks

Post	0.176***	0.307***	0.388***	0.014	0.054	0.034
	(0.058)	(0.064)	(0.058)	(0.056)	(0.058)	(0.062)
Observations	48,989	53,670	48,785	48,907	48,754	48,614
R-squared	0.193	0.159	0.118	0.238	0.247	0.178
Weather Control	Temperature, Wind, Rainfall					
Fixed Effects	Year, Month, Country of Origin, Website, Guest Type, Hotel FE					

Notes: This table presents the results estimating the impact of Jun 2013 haze shock in Singapore on subcategory review scores during the sample period between Jun 2012 and Aug 2014. The headers in the first row indicate the dependent variable used for each estimation. Panel A examines the changes of subcategory review scores during the Jun 2013 haze episode, and Panel B compares the subcategory review scores between the post-shock and pre-shock periods. Please refer to Appendix A for the detailed definition of variables. The three weather measures and fixed effects of year, month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table 6: Changes on Sentiment Outcomes

(Singapore, Jun 2012 - Aug 2014)

Panel A: Relationship of Sentiment Outcomes and Air Pollution Measures

Dep. Variable	Compound Score	Positive Proportion	Negative Proportion	Neutral Proportion
Model	(1)	(2)	(3)	(4)
Shock[a]	-0.089***	0.006	0.099***	-0.105***
	(0.005)	(0.007)	(0.002)	(0.003)
Observations	363,976	363,976	363,976	363,976
R-squared	0.143	0.089	0.141	0.071
Panel B: The Ex-post Responses of Sentiment Outcomes to the Haze Shocks				
Post	0.016	0.008	-0.003	-0.006
	(0.019)	(0.008)	(0.003)	(0.005)
Observations	348,661	348,661	348,661	348,661
R-squared	0.144	0.088	0.084	0.062
Weather Control	Temperature, Wind, Rainfall			
Fixed Effects	Year, Month, Country of Origin, Website, Guest Type, Hotel FE			

Notes: This table presents the results estimating the impact of Jun 2013 haze shock in Singapore on sentiment outcomes during the sample period between Jun 2012 and Aug 2014. The headers in the first row indicate the dependent variable used for each estimation. Panel A examines the changes of sentiment outcomes during the Jun 2013 haze episode, and Panel B compares the sentiment outcomes between the post-shock and pre-shock periods. Please refer to Appendix A for the detailed definition of variables. The three weather measures and fixed effects of year, month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table 7: Relation between Review Score and Compound Score

(Singapore, Jun 2012 - Aug 2014)

Dep. Variable	Review Score				
Haze Measure	$Shock^a$	$ln(PSI^{mean})$	$ln(PSI^{max})$	DoH	PSI Category
Model	(1)	(2)	(3)	(4)	(5)
Compound Score	1.206***	1.207***	1.208***	1.208***	1.206***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.028)
Haze	-0.068**	-0.193***	-0.110***	-0.015***	
	(0.031)	(0.033)	(0.019)	(0.003)	
PSI_{51-100}					-0.182***
					(0.015)
PSI_{100+}					-0.143***
					(0.032)
ln(Temperature)	-0.022	0.029	-0.002	0.001	0.056***
	(0.015)	(0.019)	(0.015)	(0.015)	(0.016)
ln(Rainfall)	-0.004	0.009*	-0.001	-0.001	0.031***
	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)
ln(WindSpeed)	0.210***	0.193***	0.243***	0.221***	0.242***
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)
Constant	6.640***	5.679***	6.334***	5.892***	4.044***
	(0.463)	(0.502)	(0.448)	(0.455)	(0.512)
Observations	363,976	363,976	363,976	363,976	363,976
R-squared	0.309	0.309	0.309	0.309	0.310
Year FE and Month FE	Yes	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes	Yes	Yes
Website FE	Yes	Yes	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the results of estimating Equation (1) with the compound score included as an explanatory variable. The dependent variable is review score. Please refer to Appendix A for the detailed definition of variables. Fixed effects of year, month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table 8: Topic Modelling

(Singapore, Jun 2012 - Aug 2014)

Dep. Variable Model	Review Score (1)	Compound Score (2)
Topic7	-0.146***	-0.011***
	(0.010)	(0.002)
Shock[a]	-0.358***	-0.083***
	(0.027)	(0.005)
Topic7**Shock*[a]	-0.311***	-0.041***
	(0.044)	(0.008)
ln(Temperature)	-0.035**	-0.008***
	(0.015)	(0.003)
ln(Rainfall)	-0.015***	-0.008***
	(0.004)	(0.001)
ln(WindSpeed)	0.199***	-0.011**
	(0.027)	(0.005)
Constant	7.704***	0.891***
	(0.475)	(0.091)
Observations	363,976	363,976
R-squared	0.205	0.144
Year FE and Month FE	Yes	Yes
Country of Origin FE	Yes	Yes
Website FE	Yes	Yes
Guest Type FE	Yes	Yes
Hotel FE	Yes	Yes

Notes: This table presents the results of estimating the impact of environment-related topic 7 on the review score (Column 1) and compound score (Column 2). Please refer to Appendix A for the detailed definition of variables. Fixed effects of year, month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table 9: Negative Comments, Managers' Responses, and Changes on Review Scores

(Singapore, Jun 2012 - Dec 2016)

Dep. Variable	Manager Response	Response Score	Review Score	
Equation	Equation (3)	Equation (4)	Equation (5)	
Model	(1)	(2)	(3)	(4)
$Shock^a$	0.110***	0.094***	-0.561***	-0.558***
	(0.009)	(0.005)	(0.048)	(0.048)
Review Score	-0.026***	-0.033***		
	(0.002)	(0.001)		
ln(# of Sentences)	0.001	-0.001		
	(0.003)	(0.002)		
$ln(\# of Responses)_{t-1}$			0.048***	
			(0.017)	
$ResponseRate_{t-1}$				0.118**
				(0.048)
Constant	0.440*	0.552**	5.172***	5.167***
	(0.243)	(0.234)	(0.813)	(0.812)
Observations	264,236	132,492	13,637	13,637
R-squared	0.464	0.075	0.530	0.530
Year FE and Month FE	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	No	No
Website FE	Yes	Yes	Yes	Yes
Guest Type FE	Yes	Yes	No	No
Hotel FE	Yes	Yes	Yes	Yes

Notes: This table reports the results of estimating Equations (3), (4), and (5). The headers in the first row indicate the dependent variable used for each estimation. The dependent variable, *Response Score*, in Column (2) is the sentiment compound score extracted from the manager response text. $\# of Responses_{t-1}$ is the number of responses in year-month $t-1$ and $ResponseRate_{t-1}$ is the response rate (the percentage of comments been responded) in year-month $t-1$. Please refer to Appendix A for the detailed definition of variables. Fixed effects of year, month, guests' country of origin, website, guest type, and hotel are included in regressions (1) an (2). Fixed effects of year, month, website, and hotel are included in regressions (3) an (4). Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table 10: DID Estimations: Changes on Online Review Scores during the June 2013 Haze Shock

(Singapore and Hong Kong, Jun 2012 - Aug 2014)

Dep. Variable	Review Score			
Haze Measure	$Shock^a$	$ln(PSI^{mean})$	$ln(PSI^{max})$	DoH
Model	(1)	(2)	(3)	(4)
$Treatment$*Haze	-0.418***	-0.128***	-0.083***	-0.061***
	(0.025)	(0.023)	(0.023)	(0.005)
Haze	0.044*	-0.087***	-0.225***	0.021***
	(0.024)	(0.017)	(0.023)	(0.004)
ln(Temperature)	-0.004***	-0.008***	-0.016***	0.002
	(0.001)	(0.001)	(0.002)	(0.002)
ln(Rainfall)	0.004	-0.011***	-0.022***	0.007**
	(0.003)	(0.003)	(0.003)	(0.003)
ln(WindSpeed)	0.059***	0.091***	0.092***	0.076***
	(0.014)	(0.013)	(0.014)	(0.018)
Constant	7.399***	8.056***	8.663***	7.165***
	(0.164)	(0.182)	(0.192)	(0.170)
Observations	675,272	675,272	675,272	675,272
R-squared	0.200	0.199	0.200	0.200
Year-Month FE	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes	Yes
Website FE	Yes	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes	Yes

Notes: This table reports the results of estimating Equation (6). The treatment sample consists of all the online reviews on hotels in Singapore and the control sample consists of online reviews of hotels in Hong Kong. The dependent variable is review score. $Treatment$ is a binary variable equal to 1 for online reviews of hotels in Singapore, and is equal to 0 for online reviews of hotels in Hong Kong. The headers in the second row indicate the measures of air pollution. Please refer to Appendix A for the detailed definition of variables. Fixed effects of year-month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table 11: DID Estimations: Changes on Online Review Scores after the June 2013 Haze Shock

(Singapore and Hong Kong, Jun 2012 - Aug 2014)

Dep. Variable	Review Score		
Pre-shock period	-9 Months	-6 Months	-3 Months
Post-shock period	+12 Months	+12 Months	+12 Months
Model	(1)	(2)	(3)
$Treatment * Pre$	-0.038	-0.033	0.001
	(0.026)	(0.022)	(0.025)
$Treatment * Post$	0.228***	0.235***	0.254***
	(0.033)	(0.029)	(0.025)
ln(Temperature)	-0.002**	-0.002*	-0.003***
	(0.001)	(0.001)	(0.001)
ln(Rainfall)	0.022***	0.022***	0.021***
	(0.003)	(0.003)	(0.003)
ln(WindSpeed)	-0.103***	-0.100***	-0.098***
	(0.014)	(0.015)	(0.016)
Constant	7.626***	7.610***	7.646***
	(0.167)	(0.169)	(0.167)
Observations	646,304	646,304	646,304
R-squared	0.200	0.200	0.200
Year-Month FE	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes
Website FE	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes

Notes: This table reports the results of estimating Equation (7). The treatment sample consists of all the online reviews on hotels in Singapore and the control sample consists of online reviews of hotels in Hong Kong. The dependent variable is review score. $Treatment$ is a binary variable equal to 1 for online reviews of hotels in Singapore, and is equal to 0 for online reviews of hotels in Hong Kong. Please refer to Appendix A for the detailed definition of variables. Fixed effects of year-month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table 12: Heterogeneity Test across Traveller Types and Travellers' Continent of Origin

(Singapore, Jun 2012 - Aug 2014)

Dep. Variable	Review Score			
Group	By Traveller Types		By Travellers' Continent of Origin	
Model	(1)	(2)	(3)	(4)
Shock^a	-0.396***		-0.395***	
	(0.027)		(0.026)	
Shock^a*Business	-0.074**			
	(0.036)			
Post		0.283***		0.250***
		(0.032)		(0.032)
Post*Business		0.074***		
		(0.019)		
Shock^a*Africa			0.363	
			(0.224)	
Shock^a*Europe			-0.105**	
			(0.053)	
Shock^a*NorthAmerica			0.036	
			(0.078)	
Shock^a*Oceania			-0.080*	
			(0.044)	
Shock^a*SouthAmerica			-0.165	
			(0.309)	
Post*Africa				0.061
				(0.097)
Post*Europe				0.173***
				(0.023)
Post*NorthAmerica				0.080**
				(0.035)
Post*Oceania				0.130***
				(0.022)
Post*SouthAmerica				0.100
				(0.093)
Constant	7.675***	5.459***	7.664***	5.471***
	(0.477)	(0.533)	(0.496)	(0.555)
Observations	363,976	348,661	342,888	328,207
R-squared	0.204	0.203	0.205	0.205
Fixed Effects	Year, Month, Country of Origin, Website, Guest Type, Hotel FE			

Notes: This table results of heterogeneity tests across traveller types and travellers' continent of origin. The dependent variable is the review score. Columns (1) and (2) study the difference in response between business travellers and non-business travellers. Columns (3) and (4) examines the difference in response of travellers by continent of origin. Please refer to Appendix A for the detailed definition of variables. Weather control variables, Year, month, guests' country of origin, website, guest type, and hotel fixed effect are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table 13: Heterogeneity Test by Hotel Star

(Singapore, Jun 2012 - Aug 2014)

Dep. Variable Model	Review Score (1)	Review Score (2)	Manager Response (3)
$Shock^a$	-0.371***		0.020
	(0.041)		(0.021)
Post		0.219***	
		(0.043)	
$Shock^a$*Star3	-0.148**		-0.023
	(0.062)		(0.028)
$Shock^a$*Star4	0.061**		0.032**
	(0.028)		(0.014)
$Shock^a$*Star5	0.097***		0.051***
	(0.014)		(0.011)
Post*Star3		0.054	
		(0.049)	
Post*Star4		0.147***	
		(0.043)	
Post*Star5		0.057	
		(0.037)	
Review Score			-0.003
			(0.002)
ln(# of Sentences)			0.004
			(0.003)
Constant	7.673***	5.422***	0.566***
	(0.477)	(0.537)	(0.127)
Observations	363,944	348,630	96,362
R-squared	0.204	0.203	0.562
Year FE and Month FE	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes
Website FE	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes

Notes: This table results of heterogeneity tests across hotel stars. The headers in the first row indicate the dependent variable used in each estimation. Please refer to Appendix A for the detailed definition of variables. Weather control variables, Year, month, guests' country of origin, website, guest type, and hotel fixed effect are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Appendices

Appendix A. The Definitions of Variables

Review Score: the review score made by a guest on the hotel-booking website after his/her stay at the hotel, which is used to grade the guest's experience in the hotel. The review scores on *TripAdvisor* and *Expedia* are on a 0-to-5 rating scale, while the review scores on *Agoda* and *Booking* are on a 0-to-10 rating scale. For consistency and ease of interpretation, we re-scale the review score in *TripAdvisor* and *Expedia* from 0 to 10, with a higher score corresponding to a higher level of customer satisfaction.

Compound Score: is a sentiment score extracted from the review text by using the VADER program. It is the sum of all the ratings of the different lexicons, which have been normalized between -1 (most extreme negative) and +1 (most extreme positive).

$Positive\ Proportion$: the proportion of positive words in a review text.

$Negative\ Proportion$: the proportion of negative words in a review text.

$Neutral\ Proportion$: the proportion of neutral words in a review text.

$Topic7$: a dummy that equals 1 if a review is assigned to environment-related topic 7, and 0 otherwise.

$\#ofSentences$: the number of sentences in a review.

$Manager\ Response$: is a binary that equals 1 if a review is responded by the hotel manager, and 0 otherwise. Only *TripAdvisor* and *Expedia* provide the information on Response.

Response Score: is a sentiment score extracted from the response text by using the VADER program. It is the sum of all the ratings of the different lexicons, which have been normalized between -1 (most extreme negative) and +1 (most extreme positive).

$Cleanliness$: the subcategory review score made by a guest on the *TripAdvisor* website after his/her stay at the hotel, which is used to grade the cleanliness of the hotel.

Service: the subcategory review score made by a guest on the *TripAdvisor* website after his/her stay at the hotel, which is used to grade the service quality of the hotel.

Location: the subcategory review score made by a guest on the *TripAdvisor* website after his/her stay at the hotel, which is used to grade the location of the hotel.

Sleep: the subcategory review score made by a guest on the *TripAdvisor* website after his/her stay at the hotel, which is used to grade the guest's sleep quality of the hotel.

$Value$: the subcategory review score made by a guest on the *TripAdvisor* website after his/her stay at the hotel, which is used to grade the value of the hotel.

Room: the subcategory review score made by a guest on the *TripAdvisor* website after his/her stay at the hotel, which is used to grade the room of the hotel.

PSI^{max}: is the maximum value of the daily PSI in a month.

PSI^{mean}: is the mean value of daily PSI in a month.

DoH: is the number of days that experience haze in a month.

$Temperature$: is the average value of daily temperature in a month.

$Rainfall$: is the total milliliter of rainfall in a month.

$WindSpeed$: is the mean value of the daily speed of wind in a month.

$Shock^a$: is a binary variable equals to 1 if the guest/reviewer stayed at the hotel during the two severe haze episodes in Jun 2013 and Sep-Oct 2015 (as shown in Figure 3).

$Shock^b$: is a binary variable equals to 1 if the guest/reviewer stayed at the hotel during the two mild haze episodes in Oct 2014 and Aug 2016 (as shown in Figure 3).

PSI_{0-50}: is a binary variable that equals 1 if the Max PSI ranges from 0-50.

PSI_{51-100}: is a binary variable that equals 1 if the Max PSI ranges from 51-100.

PSI_{100+}: is a binary variable that equals 1 if the Max PSI is higher than 100.

FRP: the monthly fire radiative power in Indonesia.

$\# of Responses_{j,k,t-1}$: the number of responses for hotel j on website k in year-month $t-1$

$ResponseRate_{j,k,t-1}$: the response rate (the percentage of comments been responded) for hotel j on website k in year-month $t-1$

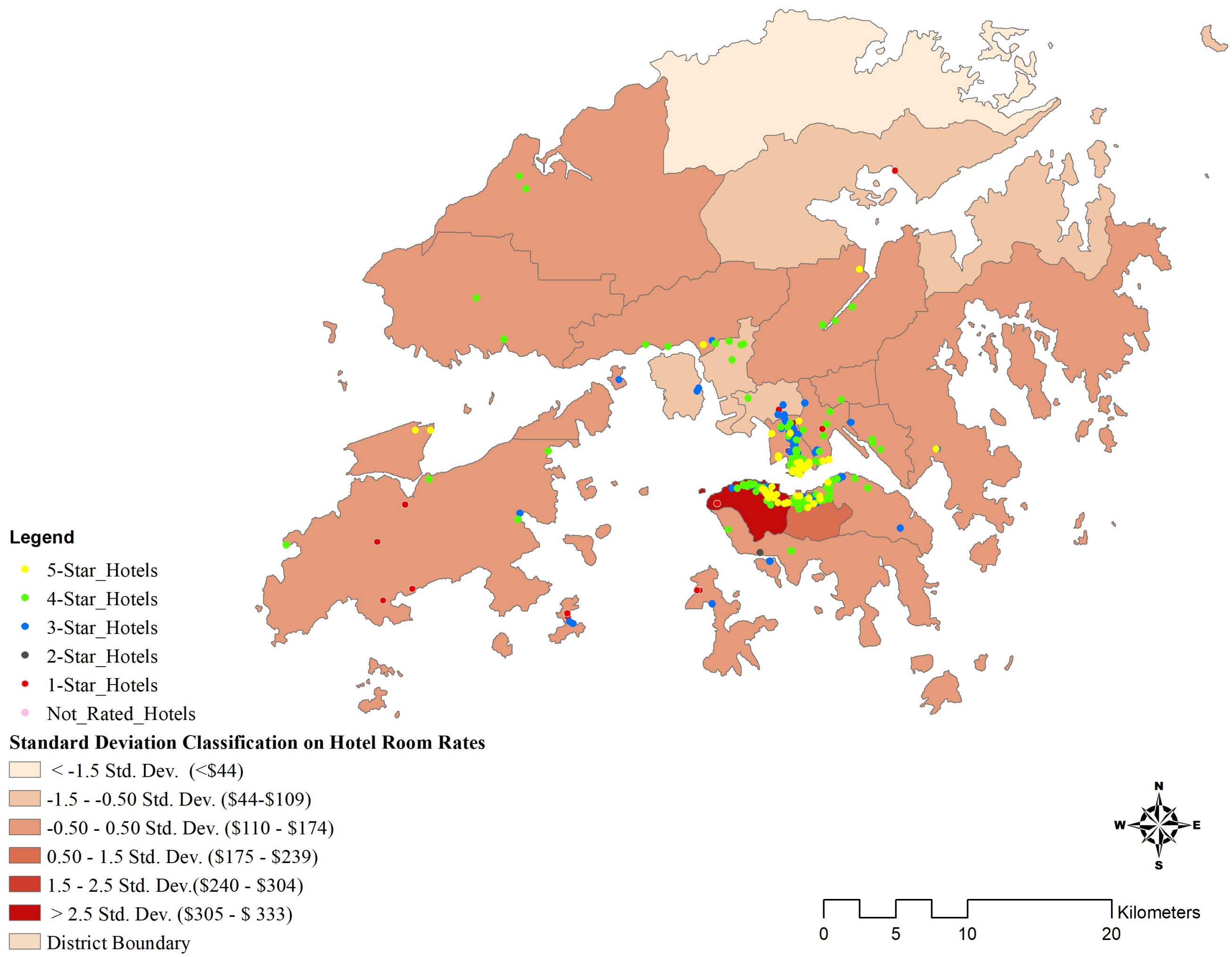
$Treatment$: is a binary variable that equals 1 if the review score is on the hotels that are located in Singapore, and 0 if the review score is on the hotels that are located in Hong Kong.

Pre: is a binary variable that equals 1 if the review is made by a guest for his/her stay before the haze shock in June 2013.

$Post$: is a binary variable that equals 1 if the review is made by a guest for his/her stay after the haze shock in June 2013.

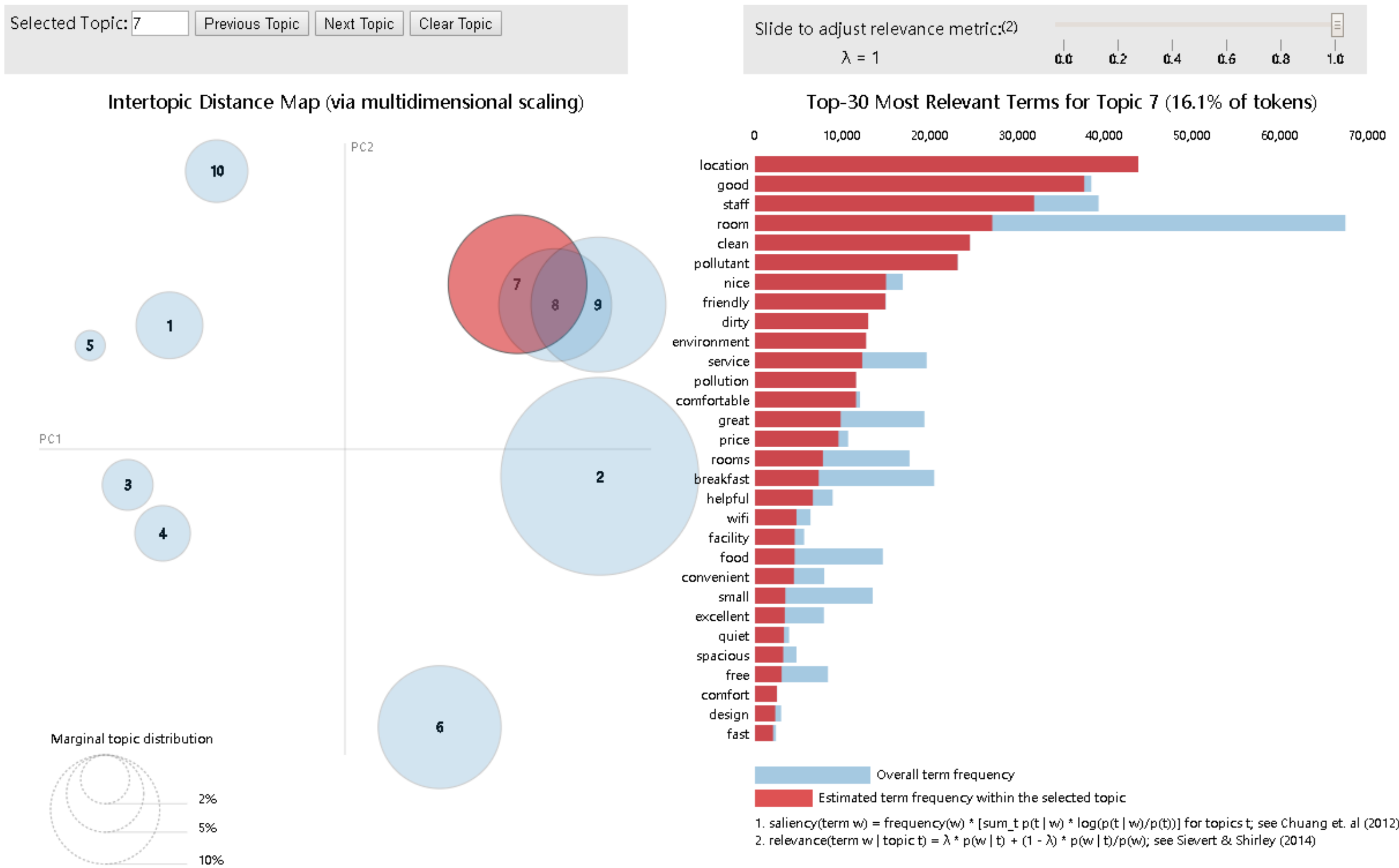
Appendix B. Figures and Tables

Figure B1: Geographic Distributions of Hotels by Star Rating in Hong Kong



Notes: This figure plots the geographic distribution of hotels and hostels in Hong Kong by star rating. We plot the quantile distribution by hotel daily room prices at the neighborhood (sector) level.

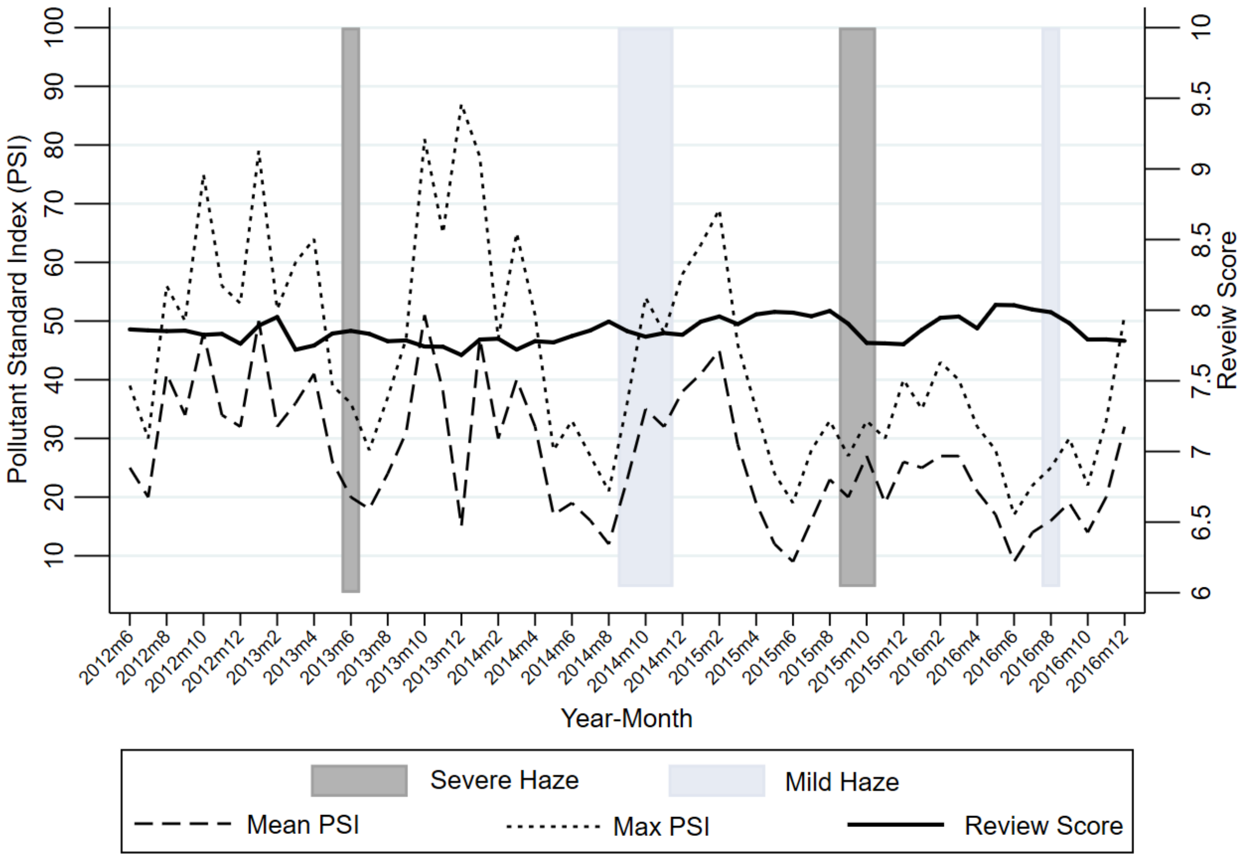
Figure B2: The 30 Most Likely Used Words in the Pollution Related Topic 7



Notes: This figure presents the interactive plots for the Topic 7. Each bubble represents a topic. The larger the bubble, the more prevalent that topic is. The bar plot on the right-hand side of the figure shows the frequency of the terms in the topic, relative to the total term frequency in the documents.

Figure B3: Monthly Trends of Air Pollution Measures and Online Review Scores

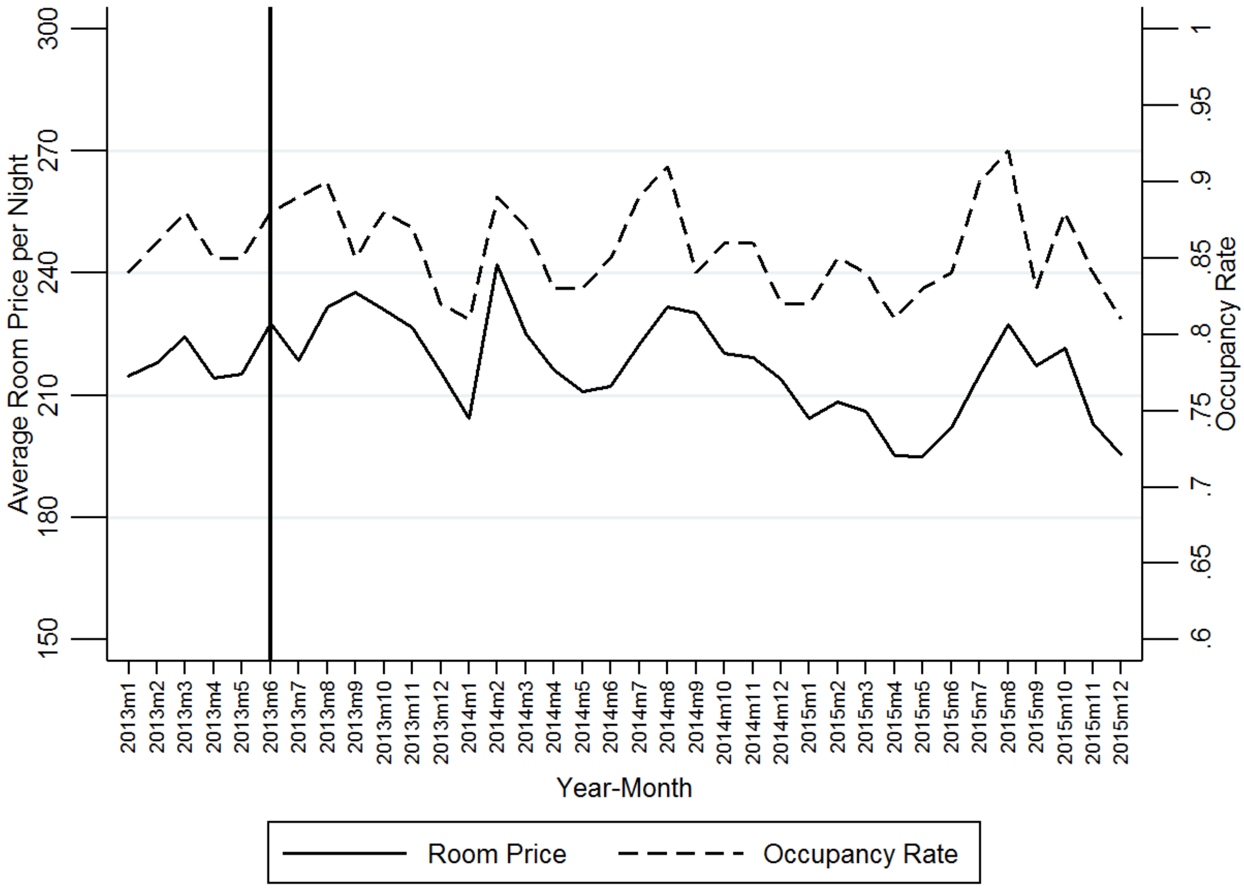
(Hong Kong, Jun 2012 - Dec 2016)



Notes: This figure plots the monthly trends of mean PSI (dashed line), maximum PSI (dotted line), wind speed, as well as the average online review score (solid line) of hotels in Hong Kong from Jun 2012 to Dec 2016.

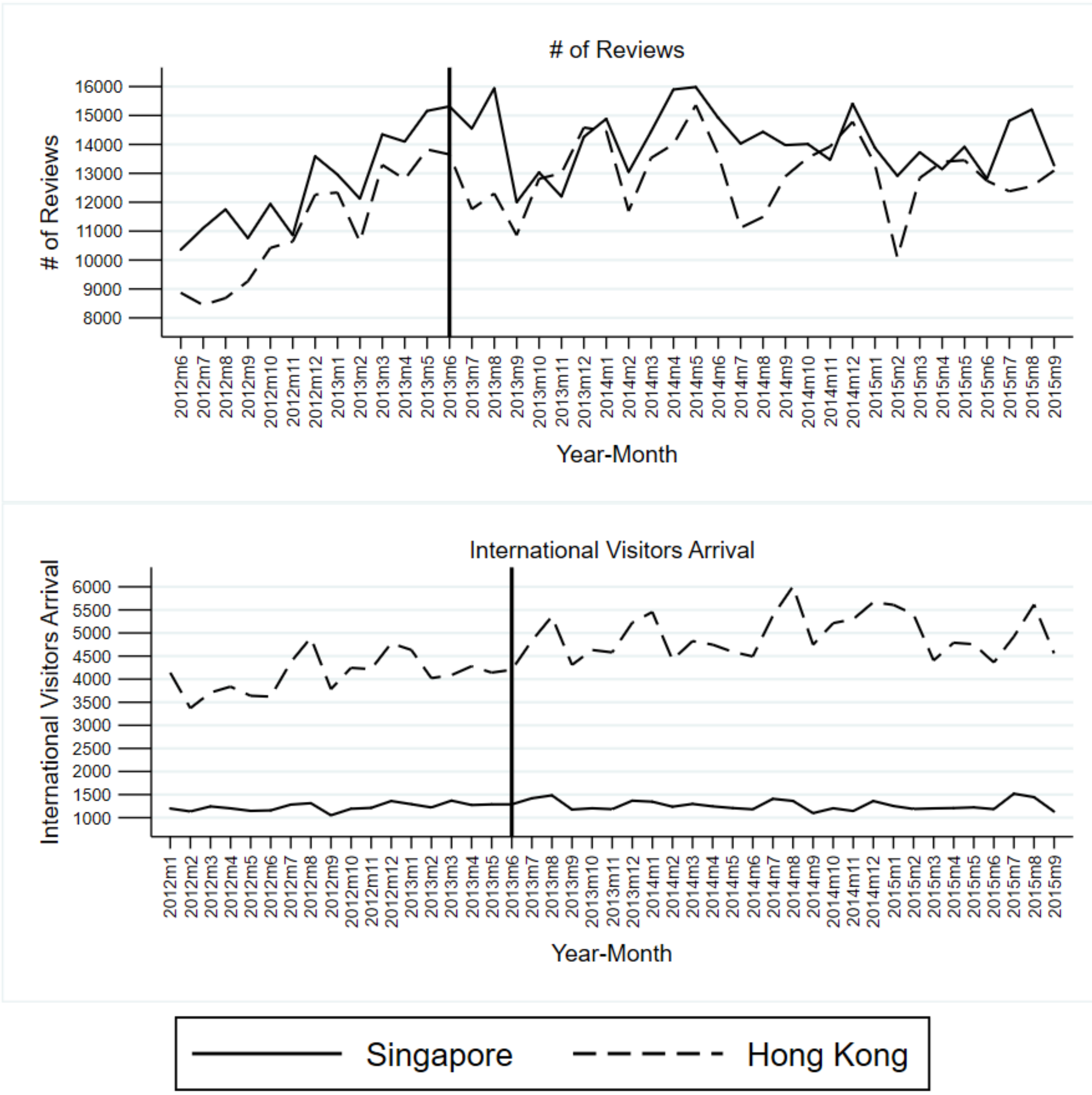
Figure B4: Average Room Price and Occupancy Rate

(Singapore, Jan 2013 - Dec 2015)



Notes: This figure plots the monthly trends of average room prices (solid line) and average room occupancy rate in Singapore.
Source of Singapore: `https://data.gov.sg/dataset/monthly-gazetted-hotel-statistics-summary`

Figure B5: Trends of Tourism Statistics in Singapore and Hong Kong



Notes: This figure plots the monthly trends of review counts (top graph) and number of international visitors arrival (bottom graph) in Singapore (solid line) and Hong Kong (dashed line). average room prices (solid line) and average room occupancy rate in Singapore.
Source of Singapore: `https://data.gov.sg/dataset/total-visitor-international-arrivals-to-singapore`;
Source of Hong Kong: `https://www.discoverhongkong.com/ca/about-hktb/news/visitor-arrival.jsp`

Table B1. Descriptive Staitstics (Hong Kong Sample)

Variables	Obs.	Mean	Median	S.D	Min	Max
Panel A: Online Reviews and Manager Responses for Hotels in Hong Kong						
Review Score	823,544	7.86	8.00	1.73	0.00	10.00
Compound Score	823,544	0.43	0.51	0.46	-1.00	1.00
Positive Share	823,544	0.25	0.19	0.23	0.00	1.00
Negative Share	823,544	0.04	0.00	0.09	0.00	1.00
Neutral Share	823,544	0.71	0.75	0.23	0.00	1.00
# of Sentences	823,544	4.77	2.00	7.96	1.00	816.00
Manager Response	248,225	0.34	0.00	0.47	0.00	1.00
Panel B: Monthly Weather and Pollution Measures in Hong Kong						
PSI_{maxn}	55	43.73	39.00	17.66	17.00	87.00
PSI_{mean}	55	26.98	26.00	10.91	9.00	51.00
DoH	55	0.89	0.00	1.65	0.00	8.00
Temperature	55	26.40	28.30	5.05	17.80	32.60
Rainfall	55	212.91	148.70	186.35	1.50	687.30
WindSpeed	55	16.79	16.89	1.57	12.35	19.22

Notes: This table presents a summary description of the data, with online review data at the individual level for hotels in Hong Kong reported in Panel A, and ambient conditions in Hong Kong on a monthly basis reported in Panel B. The sample is from Jun 2012 to Dec 2016 in Hong Kong. Please refer to Appendix A for the definition of key variables.

Table B2. A Comparison of Hong Kong and Singapore

Panel A. Hotels by Star Rating Number of Hotels (Average Room Rate)	Singapore	Hong Kong
5-star	62 (S$306.7)	42 (S$400.6)
4-star	112 (S$159.1)	131 (S$210.7)
3-star	116 (S$116.3)	146 (S$132.0)
2-star	129 (S$82.1)	148 (S$59.8)
1-star	61 (S$55.1)	180 (S$53.0)
Total	480	647

Panel B. Statistics	Singapore	Hong Kong
Annual GDP (2016)	309,754 million$	320,881 million$
GDP per capita (2016)	55,241$	43,497$
Service sector as GDP (2016)	75%	93%
Unemployment rate (2017Q1)	3.0%	3.2%
Top tax rate (2016)	22.0%	15.0%
Surface Area (2017)	709 km^2	1,050 km^2
Population (2017)	5,607,000	7,410,000
% of Chinese (2016)	77%	94%
Birth Rate (2015)	9.70	8.20
Fertility Rate (2015)	1.24	1.20
Corruption Index (2017)	84	77
CO_2 Tons per capita (2015)	8.66	6.27
Period under British Rule	1819-1959	1841-1997
Period under Japanese Occupation	1942-1945	1942-1945

Notes: This Table compares and contrasts Hong Kong and Singapore with known quantitative statistics. Panel A shows the number of hotels and average prices (in bracket) of hotels breakdown by star rating and Panel B compares socioeconomic statistics of two regions.

Table B3. Top 20 Countries of Travellers and Summary of Guest Types in Singapore and Hong Kong

Singapore				Hong Kong			
Guest Country	Obs.	Guest Types	Obs.	Guest Country	Obs.	Guest Types	Obs.
China	94,295	Couple	274,150 (31.54%)	China	79,070	Couple	227,490 (27.62%)
Indonesia	74,039	Family	188,733 (21.72%)	Singapore	52,312	Family	150,846 (18.32%)
Australia	71,758	Business	107,959 (12.24%)	Taiwan	47,322	Business	101,501 (12.32%)
Malaysia	60,415	Solo	121,355 (13.96%)	United States	36,434	Solo	144,441 (17.54%)
Japan	33,186	Group	56359,732 (6.48)	Australia	36,338	Group	63,340 (7.69%)
United Kingdom	33,170	Friends	18,128 (2.09%)	Malaysia	36,022	Friends	16,095 (1.95%)
Thailand	30,500	Other	102,431 (11.79%)	Japan	32,578	Other	119,831 (14.55%)
United States	26,068			Thailand	32,182		
Philippines	25,570			United Kingdom	27,519		
Hong Kong	24,171			Philippines	26,592		
Taiwan	19,347			South Korea	25,775		
South Korea	17,129			Indonesia	17,198		
India	17,129			Canada	13,115		
Vietnam	10,966			Macau	10,579		
Germany	9,841			India	8,465		
New Zealand	7,597			Russia	7,648		
Canada	6,999			Germany	6,367		
France	6,924			France	5,839		
Russia	5,476			New Zealand	4,821		
Italy	5,240			Italy	3,613		

Notes: This table lists the top 20 countries of origin of travellers who visited Singapore and Hong Kong from Jun 2012 to Dec 2016. It also reports the sample distributions by guest type.

Table B4. Responses of Online Review Scores to the Air Pollution Shocks

(Singapore, Jun 2012 - Dec 2016)

Dep. Variable	Review Score					
Haze Measure	$Shock^a$	$Shock^b$	$ln(PSI^{mean})$	$ln(PSI^{max})$	DoH	PSI Category
Model	(1)	(2)	(3)	(4)	(5)	(6)
Haze	-0.522***	-0.441***	-0.624***	-0.420***	-0.034***	
	(0.017)	(0.013)	(0.020)	(0.012)	(0.001)	
PSI_{51-100}						-0.103***
						(0.010)
PSI_{100+}						-0.505***
						(0.016)
ln(Temperature)	0.042***	0.114***	0.100***	0.112***	0.035***	0.110***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
ln(Rainfall)	-0.002	-0.019***	-0.024***	-0.018***	-0.001	-0.027***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
ln(WindSpeed)	0.139***	0.074***	0.152***	0.166***	0.070***	0.102***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)	(0.017)
Constant	5.824***	3.604***	5.900***	5.028***	6.197***	3.689***
	(0.297)	(0.308)	(0.287)	(0.298)	(0.294)	(0.309)
Observations	869,115	869,115	869,115	869,115	869,115	869,115
R-squared	0.195	0.196	0.195	0.196	0.195	0.195
Year FE and Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes	Yes	Yes	Yes
Website FE	Yes	Yes	Yes	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the regression results of estimating Equation (1) using hotel online review data in Singapore from Jun 2012 to Dec 2016. The dependent variable is review score. The headers in the second row indicate the measures of air pollution. $Shock^a$ represents the two severe shocks and $Shock^b$ represents the two mild shocks; PSI^{mean} and PSI^{max} represent the monthly average PSI readings and monthly maximum PSI readings, respectively; DoH measures the number of days with haze status per month; and $PSICategory$ classifies the monthly average PSI^{max} into three categories: 0-50, 51 to 100, and above 100. Fixed effects of year, month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table B5. The Impact of a on Hotel Room Prices and Occupancy Rate

(Singapore, Jan 2013 - Jun 2015)

Dep. Variable	Occupancy Rate			ln(Room Price)		
Haze Measure	$Shock^a$	$ln(PSI)$	$ln(PSI_{lag})$	$Shock^a$	$ln(PSI)$	$ln(PSI_{lag})$
Model	(1)	(2)	(3)	(4)	(5)	(6)
Haze	-0.001	-0.001	-0.009	-0.029	-0.008	-0.017
	(0.013)	(0.005)	(0.009)	(0.033)	(0.014)	(0.014)
ln(Temperature)	-0.008	-0.008	-0.006	-0.025	-0.019	-0.019
	(0.043)	(0.043)	(0.043)	(0.107)	(0.107)	(0.107)
ln(Rainfall)	0.000	0.000	0.001	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)
ln(WindSpeed)	0.008	0.008	0.009*	-0.008	-0.008	-0.007
	(0.005)	(0.005)	(0.005)	(0.013)	(0.013)	(0.013)
Constant	0.766***	0.767***	0.789***	5.787***	5.795***	5.823***
	(0.143)	(0.143)	(0.143)	(0.359)	(0.360)	(0.361)
Observations	4,240	4,240	4,240	4,240	4,240	4,240
R-squared	0.190	0.190	0.191	0.025	0.025	0.025
Year FE and Month FE	Yes	Yes	Yes	Yes	Yes	Yes
DoW FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table examines the impact of air pollution on hotel occupancy rate (Columns 1 to 3) and hotel room prices (Columns 4 to 6). The headers in the second row indicate the measures of air pollution. $Shock^a$ represents the two severe shocks. PSI is the daily PSI reading. PSI_{lag} is the one month lag daily PSI reading. Fixed effects of year, month, day of week and district are included in all regressions. Robust standard errors are clustered at the district level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table B6. DID Estimations: Changes on Sentiment Compound Scores during the June 2013 Haze Shock

(Singapore and Hong Kong, Jun 2012 - Aug 2014)

Dep. Variable	Compound Score			
Haze Measure	$Shock^a$	$ln(PSI^{mean})$	$ln(PSI^{max})$	DoH
Model	(1)	(2)	(3)	(4)
$Treatment$*Haze	-0.128***	-0.033***	-0.038***	-0.009***
	(0.006)	(0.005)	(0.005)	(0.001)
Haze	0.077***	-0.005	-0.021***	-0.001
	(0.005)	(0.003)	(0.005)	(0.001)
ln(Temperature)	0.001***	0.001	0.001*	0.001***
	(0.000)	(0.001)	(0.000)	(0.000)
ln(Rainfall)	0.000	-0.001**	-0.004***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
ln(WindSpeed)	0.003	0.005*	0.007**	-0.007*
	(0.003)	(0.003)	(0.003)	(0.004)
Constant	0.617***	0.703***	0.817***	0.635***
	(0.030)	(0.032)	(0.035)	(0.032)
Observations	675,272	675,272	675,272	675,272
R-squared	0.144	0.143	0.144	0.144
Year-Month FE	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes	Yes
Website FE	Yes	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes	Yes

Notes: This table reports the results of estimating Equation (6). The treatment sample consists of all the online reviews on hotels in Singapore and the control sample consists of online reviews of hotels in Hong Kong. The dependent variable is sentiment compound score. $Treatment$ is a binary variable equal to 1 for online reviews of hotels in Singapore, and is equal to 0 for online reviews of hotels in Hong Kong. The headers in the second row indicate the measures of air pollution. Fixed effects of year-month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table B7. DID Estimations: Changes on Sentiment Compound Scores after the June 2013 Haze Shock

(Singapore and Hong Kong, Jun 2012 - Aug 2014)

Dep. Variable	Compound Score		
Pre-shock period	-9 Months	-6 Months	-3 Months
Post-shock period	+12 Months	+12 Months	+12 Months
Model	(1)	(2)	(3)
$Treatment * Pre$	-0.012	-0.012	-0.004
	(0.014)	(0.015)	(0.016)
$Treatment * Post$	-0.005	-0.003	0.002
	(0.006)	(0.006)	(0.005)
ln(Temperature)	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
ln(Rainfall)	0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)
ln(WindSpeed)	-0.001	0.000	-0.001
	(0.003)	(0.003)	(0.003)
Constant	0.615***	0.609***	0.624***
	(0.031)	(0.031)	(0.031)
Observations	646,304	646,304	646,304
R-squared	0.145	0.145	0.145
Year-Month FE	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes
Website FE	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes

Notes: This table reports the results of estimating Equation (7). The treatment sample consists of all the online reviews on hotels in Singapore and the control sample consists of online reviews of hotels in Hong Kong. The dependent variable is sentiment compound score. *Treatment* is a binary variable equal to 1 for online reviews of hotels in Singapore, and is equal to 0 for online reviews of hotels in Hong Kong. Fixed effects of year-month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table B8. IV Estimations

(Singapore, Jun 2012 - Dec 2015)

Dep. Variable	$ln(PSI^{max})$	Review Score	Compound Score
Stage	First-Stage	Second-Stage	Second-Stage
Model	(1)	(2)	(3)
$ln(FRP)$	0.601***		
	(0.003)		
$ln(PSI^{max})$		-0.191***	-0.038***
		(0.029)	(0.007)
ln(Temperature)	0.292***	-0.012	-0.014***
	(0.001)	(0.013)	(0.003)
ln(Rainfall)	0.082***	-0.003	-0.007***
	(0.001)	(0.003)	(0.001)
ln(WindSpeed)	-0.262***	0.270***	0.023***
	(0.002)	(0.016)	(0.004)
Constant	-6.927***	6.365***	0.950***
	(0.051)	(0.370)	(0.095)
Observations	605,142	605,142	605,142
R-squared	0.671	0.199	0.150
Year FE and Month FE	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes
Website FE	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes

Notes: This table reports the results of estimating the two-stage least square Equations (9) and (10). The headers in the first row indicate the dependent variable used in each estimation. Column (1) reports the first stage results. Column (2) and Column (3) report the second stage results on review score and sentiment compound score, respectively. Fixed effects of year, month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table B9. DID Estimations using Frequent Travellers Only

(Singapore and Hong Kong, Jun 2012 - Aug 2014)

Dep. Variable	Review Score	Compound Score
Model	(1)	(2)
$Treatment * Pre$	-0.133	0.003
	(0.129)	(0.034)
$Treatment * Post$	0.320***	-0.024
	(0.102)	(0.028)
ln(Temperature)	0.006	0.001
	(0.008)	(0.002)
ln(Rainfall)	0.020	-0.012**
	(0.019)	(0.005)
ln(WindSpeed)	-0.127	0.035
	(0.121)	(0.030)
Constant	6.308***	0.381***
	(0.400)	(0.125)
Observations	16,231	16,231
R-squared	0.580	0.555
Year-Month FE	Yes	Yes
Website FE	Yes	Yes
Hotel FE	Yes	Yes
Guest FE	Yes	Yes

Notes: This table reports the results of estimating Equation (7) using frequent travellers only. The treatment sample consists of all the online reviews on hotels in Singapore and the control sample consists of online reviews of hotels in Hong Kong. The headers in the first row indicate the dependent variable used in each estimation. $Treatment$ is a binary variable equal to 1 for online reviews of hotels in Singapore, and is equal to 0 for online reviews of hotels in Hong Kong. Fixed effects of year-month, website, guest, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Table B10. Quantile Regressions (QR) for Ex-ante versus Ex-Post Responses of Online Review Scores

(Singapore, Jun 2012 - Aug 2014)

Dep. Variable	Review Score			
Method	OLS	QR (0.25)	QR (0.5)	QR (0.75)
Model	(1)	(2)	(3)	(4)
Post	0.295***	0.258***	0.300***	0.190***
	(0.032)	(0.010)	(0.010)	(0.024)
ln(Temperature)	0.058***	-0.103***	0.001	0.104***
	(0.019)	(0.006)	(0.006)	(0.015)
ln(Rainfall)	0.025***	0.014***	0.001	0.039***
	(0.006)	(0.003)	(0.003)	(0.008)
ln(WindSpeed)	-0.057	-0.152***	0.001	0.223***
	(0.037)	(0.016)	(0.016)	(0.038)
Constant	5.463***	10.149***	7.700***	4.976***
	(0.532)	(0.227)	(0.226)	(0.529)
Observations	348,661	348,661	348,661	348,661
R-squared	0.203	0.008	0.008	0.008
Year FE and Month FE	Yes	Yes	Yes	Yes
Country of Origin FE	Yes	Yes	Yes	Yes
Website FE	Yes	Yes	Yes	Yes
Guest Type FE	Yes	Yes	Yes	Yes
Hotel FE	Yes	Yes	Yes	Yes

Notes: This table provides the quatile estimations for the Equation (2). The dependent variable is review score. Fixed effects of year, month, guests' country of origin, website, guest type, and hotel are included in all regressions. Robust standard errors are clustered at the hotel level and significance levels are indicated by *** p<0.01, ** p<0.05,* p<0.1.

Appendix C. Sample Online Review

Figure C1. Sample Online Review on **Booking.com**

Katelyn
United Kingdom

10

Reviewed: 16 September 2019

Exceptional

· Apart from what i didnt like, pretty good of everything

· Check in and out system took so long. They must have more staff all the time as there are so many tourists and it's the most well-known hotel in Singapore.
Also reception desk should prepare first-aid medical supplies. I asked plaster when I check out, but they said i should have asked in advance, which is weird compare to their hotel scale. It's a shame and a bit disappointed of their service.

Stayed in September 2019

Figure C2. Sample Online Review on **TripAdvisor.com**

AHall27 wrote a review Oct 2010
Perth, Australia • **1** contribution • **1** helpful vote

Great architecture, hotel, location, facilities and staff

"Great piece of architecture. Casino and shops on site. Pool is out of this world. Clean, friendly and great atmosphere. Slightly more pricey on the beverage front than most of Singapore, but well worth it for the views from the 57th floor overlooking the Marina Bay. Well organised, rooms are large with terrific views. Pricey but a great experience. Can't wait to go back."

Read less

Date of stay: September 2010

Value	Rooms
Location	Cleanliness
Service	Sleep Quality

Figure C3. Sample Online Review on **Expedia.com**

5/5 Excellent

Rich, chula Vista, USA
Traveled with partner
Apr 6, 2019

If you have only one night in Singapore, spend it here! The view from the rooftop pool is worth every penny of the price and the whole area around the hotel is full of activities. The awe factor sticks around on the second day but if you are staying longer and are not "crazy rich" there are less expensive options located closer to other activities. As awesome as this place us, it is only a part of what Singapore has to offer.

Stayed 1 night in Mar 2019

Figure C4. Sample Online Review on **Agoda.com**

10.0 Exceptional

Charles from United States
Couple
Deluxe King Garden View
Stayed 1 night in April 2019

"You definitely get your money's worth here"

Nice spacious room overlooking Gardens by the Bay. I loved the infinity pool, especially during the laser light shows every night. The exclusive observation deck was great too, with lots of Instagram-worthy photo ops. Breakfast buffet was one of the largest I've ever experienced. I'll definitely stay there again!

Reviewed May 01, 2019