Public Transport, Noise Complaints, and Housing: Evidence from Sentiment Analysis in Singapore

Yi Fan,[*] Ho Pin Teo,[†] Wayne Xinwei Wan[‡]

December 18, 2019

Abstract

This paper investigates the effect of a new bus route on subjective noise complaints of residents and the influence of noise on housing price. To overcome the challenge of mapping noise data with subjective emotion, we use a novel data source—text-based noise complaint records from residents in a town in Singapore—and apply natural language processing (NLP) tools to conduct sentiment analysis. To address the endogeneity concern regarding the bus route, we use hypothetical least cost path as an instrument for the existing bus route. We find that living closer to the bus route for every 100 meters increases noise complaints by around 10 percentage points, and the effect is more severe on medium floor levels (5^{th}- 8^{th} floors) and near bus stops (within 100 meters). We further link noise with housing price and discover a price reduction of 3% with a 1-scale-point increase in noise complaints. This implies that bus noise offsets 18.8% of the benefit from convenience, which sheds light on the importance of noise insulation policy and design.

Key words: noise, public transport, sentiment analysis, housing
JEL Codes: R41, Q53, R30

[*]Department of Real Estate, National University of Singapore. Email: rstfyi@nus.edu.sg. Yi Fan acknowledges financial support from NUS Startup Grant No. R-297-000-134-133.

[†]Department of Building, National University of Singapore. Email: bdgthp@nus.edu.sg

[‡]Department of Land Economy, The University of Cambridge. Email: xw357@cam.ac.uk

1 Introduction

Public transport is an important service for improving accessibility and addressing congestion in cities (Anderson, 2014). It also constitutes a significant proportion of road transportation, which is one of the major sources of noise pollution in urban environments (Chui et al., 2004). With rapid urbanization, the problem of urban noise pollution has been attracting increasingly more attention from governors, scholars, and the public. The World Health Organization (WHO) reports that noise pollution has significant impacts on physical and mental health (WHO, 2016). Excessive noise also leads to social issues such as violence and generates economic losses (De Borger and Proost, 2013; Jamir et al., 2014). However, a few studies have differentiated the transportation accessibility benefits and negative externalities from harmful emissions for housing price (Chasco and Le Gallo, 2015; Higgins et al., 2019), especially from the noise of public transport. In this paper, we use a natural experiment with a newly opened bus route across a dense residential area in Singapore to examine the impact of public transport on noise complaints and its influence on housing price. As one of the world's most densely populated city-states, Singapore has always faced a severe issue of urban noise pollution (Lam et al., 2013). More than 80% of its citizens live in public housing, with short distances between buildings. Around 70,000 complaints of excessive noise are made to government agencies every year (Wan, 2016).

It has been widely acknowledged in the literature that better accessibility to public transport results in higher housing price and the intensification of its service increases the social welfare of residents (Monchambert and De Palma, 2014; Holmgren, 2014; Chalak et al., 2016). At the individual level, better accessibility enables participation in social activities and is associated with positive health outcomes (De Vos et al., 2013; Lucas, 2012). However, the elevation of service frequency also introduces traffic noise, possibly aggravates environmental pollution (Bilger and Carrieri, 2013; Nega et al., 2013), and imposes negative externalities on housing price (Ossokina and Verweij, 2015). Chasco and Le Gallo (2015) estimate the households' willingness to pay for properties with less noise, using the households' subjective perceptions of noise from the census data. Diao et al. (2016) document that the removal of train noise externalities increases housing prices in the affected area by 13.7%. Higgins et al. (2019) find the accessibility benefits of the new highways are offset by the environmental costs of air pollution, specifically for the housing units with both high accessibility and high exposure to pollution. However, there still lacks empirical studies about the accessibility benefit and noise cost of public transport. Our study aims to bridge this literature gap by directly identifying the negative impact of launching public bus services on housing prices through the influence of noise, which will potentially offset the benefit brought by its convenience.

The real-time measurement of noise can be very costly at individual building levels in the dense urban environments (Segura-Garcia et al., 2014), which has been a major challenge for past studies. Although several cities around the world are providing the city-level noise maps,[1] only sparse measurements of noise samples at the district or regional level at are taken, and noise maps are estimated using propagation models (Mircea et al., 2008). To overcome this challenge, one branch of research focuses on improving noise-measurement instruments or building empirical mathematical models to simulate real-time sound environments and noise distribution (Alam et al., 2010; Mak et al., 2010; Rana et al., 2010). However, these studies usually cover a limited number of buildings. The results are thus easily biased by unobserved building factors. Other studies propose to use in-house surveys to construct residents' noise perception index as a proxy for actual noise levels (Brown and Lam, 1987; Jakovljevic et al., 2009; Park et al., 2016). Nevertheless, the survey method suffers from a number of drawbacks, such as memory error and retrospective bias in responses (Taylor et al., 2013). Furthermore, most previous studies in this stream focus on the frequency of troublesome cases, possibly because it is difficult to address subjectivity when measuring the perceived severity of noise incidents (Weinhold, 2013). However, incidence and intensity are two distinct dimensions, and are expected to have different patterns (Figures 1a-1b).

In this study, we propose a novel sentiment analysis method to study residents' perception of noise pollution—specifically, noise from the intensification of public bus services—from residents'

[1]Examples of the city-level noise maps include the U.S. National Transportation Noise Map (`https://www.transportation.gov/highlights/national-transportation-noise-map`), and the London Road Traffic Noise Map (`http://www.londonnoisemap.com/`).

noise complaints in Singapore's public housing. This sentiment analysis method originates in the field of computer natural language processing (NLP), in which it is used to learn customers' attitudes toward products based on their text comments. In Singapore, residents' noise complaints in public housing are recorded by the local government in words, which facilitates the generation of sentiment scores based on these words as a measure of residents' emotion. Our study covers $2,032$ noise complaint records from all 142 public housing blocks in three planning subzones of the Bukit Panjang area in Northwest Singapore from 2010 to 2018. SentimentR, which is based on a sentiment dictionary containing approximately 6,800 ranked positive and negative sentiment words, is used in this study (Hu and Liu, 2004; Lam, 2016).[2]

Our sentiment analysis method has two advantages: First, government agencies in many major cities worldwide encourage the residents to report noise incidents, so the database of noise complaint records naturally exists, such as the Noise Complaints Open Data in New York City. In other words, this method is not only applicable in Singapore, but also externally valid. Second, our sentiment analysis method captures the subjective perceptions of noise pollution, which is found to better explain the impact of noise on housing price than the objective measurements (Boyle and Kiel, 2001; Chasco and Le Gallo, 2013, 2015). Different from past studies using surveys (Weinhold, 2013), this method captures a real-time measurement of the noise sentiment intensity and is powerful for baselining the subjectivity of individual responses.

Another major challenge in past studies on public transport and urban noise pollution is the problem of endogeneity (Cropper and Gordon, 1991; Higgins et al., 2018). Unobserved factors may be simultaneously associated with public transport and the perception of urban noise, which undermines the reliability of empirical results. For instance, the intensification of public transport services may be an ex post action by the government to address the growing population in the area, while the higher resident density also induces more noise. Residents' unobserved personal attributes may also threaten the estimation, as residents who are sensitive to noise are likely to choose quieter locations and are less tolerant of a sudden increase in noise. The main empirical

[2]Details of the toolkit are discussed in Appendix A.

approach to address this challenge has been to develop instrumental variables (IVs) that are highly correlated with the actual route of the transportation service, but are intended to affect noise output only through this correlation (Redding and Turner, 2015). In the literature, commonly applied instruments for the actual transportation route include the initially planned route (Baum-Snow, 2007; Michaels et al., 2012; Donaldson, 2018), the historical route (Duranton and Turner, 2011; Hsu and Zhang, 2014; Martincus et al., 2017), and other related hypothetical routes (Banerjee et al., 2012; Faber, 2014; Jedwab et al., 2017).

Singapore's public housing provides an ideal context to address the endogeneity problem. The buildings in our study area were almost all constructed during same period in which a new satellite town was initially planned, and few further developments have been carried out after that. Public housing blocks, which accommodate more than 80% of Singapore residents, have uniform building plans, room layouts, and construction materials. Demographics such as nationality and ethnicity are controlled for to be evenly distributed based on the nation's "Ethnic Integration Policy and Permanent Resident Quota" system. During our study period of 2010-2018, a new regional bus service, Route No. 972, was launched in November 2013, and no other bus routes were introduced in this region during the same period. This allows us to apply a difference-in-difference strategy to examine the noise impact from the launch of the bus service. Further, we modify the hypothetical least cost route IV strategy of Faber (2014) and Jedwab et al. (2017) to an urban scale and use it to consistently estimate the causal effect of the bus service improvement on residential noise. The main IV in the estimation is the Euclidean cost path (ECP), which is the shortest linear distance connecting all of the bus stops along the bus service route. The other IV used in the robustness check is the least cost path (LCP), which is the shortest driving route connecting the points at which the bus enters and exits the study area.

Our empirical results reveal that at an individual level, if the distance between the housing unit and the bus route decreases by 100 meters, the launch of the new bus service has worsened the sentiment of noise complaints by 9.5 percentage points. If using a binary variable to classify whether the housing unit is near the new bus route, we find the sentiment from residents living near the bus route (within 100 meters) increases by 10.9 percentage points compared to those living between 100 and 200 meters. These results remain robust if we include controls for the noise sentiment in the previous year, use LCP as the alternative IV, or use aggregated sentiment scores at building level. The adverse effect also exhibits heterogeneity across floor height and distance to bus stops. For units on the 1st to 4th floors and on 9th floor or above, this adverse effect is not statistically significant, probably due to noise insulation infrastructure on the ground such as shade trees and attenuation of noise on higher floors. On medium floors (5th-8th levels), however, this adverse effect has doubled in comparison with the average: A decrease of distance by 100 meters increases the sentiment by 24.3 percentage points, and living near the bus route (within 100 meters) increases the sentiment by 21.4 percentage points. The noise effect from the bus service has also shown larger magnitude on buildings closer to bus stops, which implies that the introduction of visitors is a major source of noise annoyance.

Finally, we conduct cost and benefit analysis on the impact of public transport on housing price. By explicitly controlling for the change in accessibility by comparing distance to the nearest bus stops before and after the launch of a new bus route, we find that an increase of 1 scale point in noise sentiment is associated with a 3.26% decrease in housing price. This implies that for properties closer to the bus route by 100 meters, noise generated by buses leads to an implicit 0.31% decrease in property value. Using the same transaction data set we also find that, after the new bus service launches, the prices of properties closer to the bus route by 100 meters has increased by 1.34%. Therefore, our empirical results estimate that over 18.79% of the benefit from improved accessibility brought by the new bus routes is offset by the negative externality from its noise pollution.

Our paper contributes to the literature from both conceptual and methodological perspectives. First, we isolate the negative impact of noise pollution on housing price apart from the accessibility convenience by launching a new public bus route. It is closely related to the studies of investment into public transport to address the last-mile connectivity issue, which is trending not only in Singapore but in many other congested cities as well (Xie et al., 2010). Prior literature intensively examines the overall impact of public transport on housing price (Baum-Snow and Kahn, 2000; McMillen and McDonald, 2004; Xu et al., 2015), but a few studies present empirical evidences on cost-benefit analysis (Chasco and Le Gallo, 2015; Higgins et al., 2019), possibly because of the costly real-time measures of noise intensity and the complexity of the public transport system. Taking advantage of the clean setting of the public bus system in Singapore's context, we advance current understanding by weighing the advantages and disadvantages of public transport for housing price and provide empirical evidence for policy makers.

Second, as Segura-Garcia et al. (2014) document, current noise map data in the major cities worldwide are estimations based on sparse measurements and mathematical propagation models, while measuring noise at the individual building level in the dense urban environments is costly. We propose a novel methodology to measure the noise sentiment at finer detailed level using the noise complaints data, which naturally exists with the government agencies in many major cities, such as Singapore, New York and London. In addition, different from measuring noise incidents by counting frequency through surveys (Tamura et al., 2017; Weinhold, 2013), our methodology of using real-time noise complaints to measure noise sentiment also contributes to the literature on understanding residential noise pollution based on subjective perception, which is proved to show a pattern that is complementary to that in previous literature (Boyle and Kiel, 2001; Chasco and Le Gallo, 2013; Dzhambov and Dimitrova, 2014).

The rest of the paper is structured as follows. Section 2 introduces the institutional background. Section 3 describes the data and the sentiment analysis tool, while Section 4 presents the empirical specifications. Results are summarized in Section 5, followed by cost benefit analysis in Section 6. Section 7 concludes.

2 Institutional Background

2.1 Public Housing and Urban Residential Noise in Singapore

Noise pollution—the excessive propagation of noise—is widely known to be harmful for human life and activities. According to the WHO, noise exposure is responsible for a wide range of negative public health effects, such as ischemic heart disease, cognitive impairment in children, and stress-related mental health risks. Exposure to residential road traffic noise is also associated with a higher risk of diabetes and cardiovascular disease (Sørensen et al., 2013; Münzel et al., 2018). From a social perspective, community noise pollution is shown to increase violent behavior and crime rates and to lower the birth rate and newborn weights (Jamir et al., 2014; Nieuwenhuijsen et al., 2017). Total socioeconomic loss from road noise pollution in the UK is estimated to be similar to the loss from road accidents, and it exceeds the loss from climate change (DEFRA, 2013).

The issue of residential noise pollution is more serious in densely populated modern cities such as Singapore. According to The Straits Times, "over 7,000 residents are living in each square kilometer of land in Singapore," and "more than 85 percent of its residents live in the nation's public housing flats with very close distance to roads, constructions or other building blocks." Around 70,000 complaints are made to various government agencies each year about excessive noise in Singapore, which means that around 5% of local households make a noise complaint every year (Wan, 2016). To manage noise pollution in public housing, the local public housing authority—the Housing Development Board (HDB)—has promoted the Neighborliness Campaign among residents, which encourages them to respect the neighborhood by avoiding producing excessive noise. The National Environment Agency (NEA) has also established regulations to control for noise origins, such as setting maximum permissible noise levels for factories and construction work, a no-work rule during certain periods of the day, and maximum noise emission limits for air conditioning and mechanical ventilation systems in buildings. In case of excessive noise, residents can also call or email the hotlines located in individual town centers.

Public housing in Singapore provides an ideal setting to study urban residential noise pollution. On the one hand, residential noise pollution in public housing is much more serious than in private estates, which leads to policy incentives for social equality. Due to concerns about land and construction cost, most public housing is densely constructed using economical materials, which provide limited performance in noise insulation. Since residents living in public housing have stronger demand for public transport, most public housing is located closer to transportation hubs or major roads. This leads to more exposure to traffic noise. On the other hand, different building and urban attributes, such as building typology and floor level (Lam et al., 2013; Mak et al., 2010), will have significant noisescape influence. This makes it challenging to control for these noise-related variables in a complex urban context. In Singapore, however, public housing has undergone around four waves of morphological changes, and the buildings constructed in each era have uniform morphology (Pow, 2009). This makes it feasible to exclude the impact from variance morphology in this study.

2.2 Study Area and the New Bus Service

Our study covers 142 HDB blocks in the subzones of Fajar, Jelebu, and Bungkit of Bukit Panjang District in Northwest Singapore. The district of Bukit Panjang, which was previously known as Zhenghua, is one of the oldest low-rise residential suburban towns and is situated on a low-lying elongated hill. To meet demand stemming from the population surge after the nation's independence, the development of public housing in the town and advanced earthworks began in the early 1980s. Most of the HDB development was completed by the mid-1990s, and only a few new residential projects or redevelopments have been completed since then. As a result, the morphology of HDB blocks in the study area mostly follows the two standard building prototypes for HDB buildings in that generation: 132 Slab Blocks and 8 Point Blocks (Appendix Figure B1). Slab Blocks are between 12 and 14 stories, and Point Blocks are all 25 stories. Only two blocks have been redeveloped in recent years, for which a new building typology was adopted. HDB buildings from the same generation are also constructed using standard materials, and layouts of the units and the number of rooms are almost the same. Like any other public housing in Singapore, regular repainting and upgrading programs are also conducted in this district every 5 to 7 years. Therefore, the maintenance of these buildings is also kept at a similar level.

Moreover, sales of HDB blocks are restricted to Singapore citizens or permanent residents, and there is a quota for noncitizen owners in each building. The nation's Ethnic Integration Policy requires that the proportion of owners' ethnicity in each individual building strictly follow the national average proportion of Singapore's three major races (Chinese, Malay, and Indian). If the number of owners from one race hits the threshold, owners from this race can only sell the unit to buyers of the same race. Therefore, on the aggregate block level, the demographics are relatively uniform.

The road network in the area follows a typical town planning hierarchy in Singapore. The north and east sides of the site are enclosed by the two highest standard expressways in Singapore, and on the other side of the expressways are reserved forest land without urban development. One major road (Bukit Panjang Road) cuts through the site to connect with the expressways and a secondary ring road forms a loop to direct traffic to major roads. All other minor roads in the community are connected to the ring road loop. Since 1999, there has been a light rail train (LRT) line looping in the district, which connects to the nearby Choa Chu Kang Mass Rapid Transit (MRT) station on the North-East Line. The Bukit Panjang MTR station—the terminal station of the Downtown Line—also connects to this LRT line in the study district; it started operations on December 27, 2015.

Like many other major cities worldwide, the Singapore government aims to address the trending last-mile connectivity issue in the city by improving the service coverage and frequency of its public buses (Xie et al., 2010). Public buses are the most frequently used transportation mode (41.3%, according to Singapore's General Household Survey 2015) by the population traveling to work in Singapore. Singapore's Ministry of Transport is aiming to increase the peak-hour public transport mode share to 75% by 2030, and it launched the Bus Service Enhancement Programme (BSEP) to expand the bus fleet by 35% before 2017. As part of the program, a new regional bus service, Route No. 972, was launched in November 2013 in our study area (Figure 2a). Our study site has an area of 1.2 square kilometers (km), and the segment of the bus route in our site is over 2.5 km long with 6 bus stops. The bus stops are sequentially allocated along the road at intervals of about 400 meters. Unlike buses intended to connect the community with other districts in the city, this bus is mainly designed to improve last-mile connectivity within the community. Residents take this bus to the nearest MRT station, which is more than 1 km away. Therefore, instead of driving on the fastest path along major roads, this bus zigzags along minor community roads to ensure that most of the blocks are within 200 meters of its service cover. Faber (2014) and Jedwab et al. (2017) demonstrate that on a regional scale, the theoretical shortest path between cities is an effective instrument for an actual inter-city expressway network. Similarly, on an urban scale, we can construct the theoretical shortest path for this bus service as the instrument for empirical analysis. Apart from Route No. 972, no other bus routes were introduced along a similar route during our study period from 2010 to 2018.

3 Data

3.1 Noise Complaints and Sentiment Analysis Tool

In Singapore, residents in public housing report their noise complaints to the local government through hotlines, email, or other written notices. These records are then centralized by local Community Development Councils (CDC) for further action. The data in our study cover all noise complaints made by HDB residents in the subzones of Fajar, Jelebu, and Bungkit of Bukit Panjang District from March 2010 to February 2018.

We collected 2,032 records of noise complaints in all the 142 buildings. For each complaint, the incident date, time, and complaint content are recorded by the agency that receives the complaint. This information is believed to be more accurate on the sentiment intensity of noise than any subsequent survey of residents after the noise incident. A cleansing process is applied before calculation of their sentiment score. This is because, by nature of the lexicon-based sentiment analysis tools, sentences must be transformed into a hierarchical structure of meaningful segments based on their meaning and grammar. As a result, scoring is sensitive to grammatical errors and typos, and these must be manually corrected in the preprocessing. In addition, when complaint data were paraphrased and recorded, the registrar used a number of abbreviations, such as "plz" for "please." These abbreviations are also manually revised.

The concept of sentiment analysis was introduced in the early 2000s, when computer scientists tried to extract polarized opinions (either positive or negative) from customer reviews of commercial goods or movies (Pang et al., 2002; Turney, 2002). It has since been applied in various empirical studies involving human perceptions (Cambria et al., 2013). Tetlock (2007) finds that the frequency of negative words in *Wall Street Journal* articles predicts stock returns, and Garcia (2013) finds that this predictive power is stronger during recessions. Using search results from Google, Zheng et al. (2016) conclude that investor confidence is a determinant of China's housing price. Nevertheless, the search engine method, which uses a limited number of predefined vocabulary as key words (e.g., "good" or "bad"), may not be able to capture all variations in sentiment, especially when data entries only contain short sentences.

We employed a lexicon-based method using the SentimentR toolkit, which is a more mature alternative tool.[3] This is a supervised machine learning process and is dependent on the scoring lexicon—i.e., dictionary—the algorithm is associated with (Gao et al., 2015). After constructing a reliable sentiment scoring dictionary and transforming sentences into hierarchies, the algorithm assigns corresponding scores to these patterns and derives the sentiment of sentences (Pang and Lee, 2008; Kim et al., 2011). These tools are able to present a continuous scoring range for sentiment, instead of simply identifying the polarity of the text (Thelwall et al., 2010). In our study, the sentiment score of each individual noise report is calculated by running the SentimentR tool on the cleaned complaint contents. Results are winsorized to the top and bottom 1% of the distribution to remove the impact of extreme cases. This is followed by a normalization of scores to the range from 0 to 1, which represents the most low-key and the most severe emotions, respectively.

[3] Details on the toolkit selection process and algorithm of SentimentR are presented in Appendix A.

A subset of 593 records in the database contain the complainant's full unit number and floor.[4] There are 328 complainants within 100 to 200 meters of the bus route, and 265 complainants living within the 100-meter boundary. Distributions on demographics and the physical features of buildings are consistent with the standard for Singapore's public housing. When the complainant calls the hotline, the registrar will usually record the complainant's title (Mr., Ms., etc.), so that the information on their gender can be obtained (527 records). We use these 527 records as our main sample in the baseline regressions. There are 194 cases before the launch of the bus and 253 cases after. Geo-referenced information on the site's administrative boundary is obtained from the Singapore government's online public data portal, while information on building blocks, road networks, MRT and LRT lines, and bus stops are downloaded from OpenStreetMap. We also code the morphological classification of the buildings and the locations of Residential Committee (RC) Centers. The list of shopping centers in the study area is from the online OneMap of Singapore's Urban Redevelopment Authority. These are geo-referenced to OpenStreetMap's shapefile of building blocks by matching the postcode of each building. The age of each building block is from the official HDB website. The distance from each building block to roads, bus stations, or other public facilities is calculated using ArcGIS. The LRT line in our study area operates on a viaduct, which causes noise exposure, so the distance from buildings to the viaduct is also calculated. Table 1 presents summary statistics for these complaint cases.

Within each building block, the sentiment of complaints in the 12 months before or after launching the bus service is averaged separately to construct the aggregate building level sentiment index. In this way, records without full floor, unit, or demographic information can then be included in the analysis. In 47 buildings, complaints were made during both the 12 months before and after launching the bus service. We use the subset of these 47 buildings as the sample in our robustness check at the building level. Appendix Table C1 summarizes the changes in this aggregate level sentiment and the physical features of these building blocks. Except for distance to nearest bus stop and the bus service route, there are no statistically significant differences be-

[4] We also conduct the analysis at building level to address the issue of missing unit information for individual cases.

tween the physical features of these buildings, which aligns with our previous understanding that the physical features of Singapore's public housing are highly uniform.

Data on housing transactions, including transaction price, date, floor, and size of the unit, are also obtained from the government's online public data portal. From 2000 to 2017, there are $1,450$ transactions for 131 buildings in our study area. All transactions are second-hand resales of public housing units between Singaporean citizens or permanent residents. The information is also summarized in Table 1. The averaged transacted housing price is 375,207 SGD, which is equivalent to approximately 278,365 USD. The dwelling size is 104 square meters on average, with a mean floor level of 6.7. Our sample is thus representative of public housing transactions in Singapore. Of these transactions, 893 can be matched with noise complaints in the same building 12 months before the transaction date.

3.2 Least Cost Path

The major objective of this study is to examine the distance effect from public traffic noise in a high-density urban environment. Similar empirical analyses in the literature usually face a challenge from the strong ex ante assumption of random shocks. Nonetheless, this assumption may not hold in this context, because new public bus service is usually launched to meet rising demand from residents, such as increasing numbers of employed residents for daily commuting or in-transit visitors. Specifically in our study, this newly launched bus line is mainly designed for local community service, which means that its route will pass through more crowded sub-town centers to serve more residents. Also, unobserved personal traits may both impact the selection of residential location and the noise sentiment. These unobserved factors will bias OLS estimates of the noise effect from bus service.

To address this concern, we construct a hypothetical least cost bus route as the instrumental variable for the actual bus route, following the IV construction strategies from Faber (2014) and Jedwab et al. (2017). Specifically, a theoretical shortest route for Bus 972 to pass through the study area exists, and it still connects the original entrance and exit points. The actual bus route deviates from the theoretical shortest route to cover more residential areas, but is not likely to deviate too far from the theoretical path. This is because bus service must also compete with alternative travel modes, such as trains and private driving, and must take passengers' total commuting time into consideration. Thus, the actual bus route is expected to be highly correlated with the theoretical shortest route. The theoretical path, however, is believed to be uncorrelated with noise annoyance beyond its correlation with the actual bus service.

We define two such theoretical shortest paths. The first one, following the strategy of Faber (2014), assumes that the original bus stops along the service route are the fixed nodes and uses Euclidean straight lines to connect these nodes as the shortest path. This path is denoted as the Euclidean cost path (ECP) and is presented in Figure 2a. In the second construction, instead of using geological elevation and land use to represent the cost as Faber (2014) and Jedwab et al. (2017) do, we consider actual travel time as the cost indicator. By fixing only the bus's entrance and exit points in the site, we use Google Maps to calculate the shortest-time commuting path connecting these two points. This path is denoted as the least cost path (LCP) and is drawn in Figure 2b. Distances from buildings to these paths are calculated using ArcGIS.

4 Empirical Strategy

A standard difference-in-difference (DID) strategy estimates the net impact of closeness to new public transport services on housing prices (Diao et al., 2017). However, it is not able to differentiate the benefit of accessibility and the cost of environmental externalities, because both the benefit and the cost correlate with the closeness to the public transport. Therefore, we apply a two-step strategy to specifically estimate the negative impact of noise from public transport on housing price. First, we apply a DID strategy to estimate the impact of closeness to the new bus route on the noise sentiment. Second, we estimate the impact of lagged noise sentiment on subsequent housing price, by explicitly controlling for the change in accessibility. Combining the results from the two steps, we interpolate the negative impact of closeness to the bus route on housing price due to the noise. Finally, we compare the net impact of public transport on housing price and its negative impact of noise, and we estimate how much benefit of accessibility is offset by the noise externalities. The following part of this section explains our empirical specifications in details.

4.1 Public Bus and Noise Complaints: Baseline Estimation

The baseline estimation uses 527 complaint records during March 2010 to February 2018 with full floor and unit information. The following DID specification is applied:

$$SI^t_{ij} = \beta_1 Distance_i + \beta_2 Launch_t + \beta_3 Launch_t * Distance_i + X_i^{'}\theta \ + U_j^{'}\mu \ + \varphi_t + \omega_i + \epsilon_{ijt}, \quad (1)$$

$$SI^t_{ij} = \beta_1 Near_i + \beta_2 Launch_t + \beta_3 Launch_t * Near_i + X_i^{'}\theta \ + U_j^{'}\mu \ + \varphi_t + \omega_i + \epsilon_{ijt}, \quad (2)$$

where SI^t_{ij} is the sentiment score from a complaint made at time t by a resident living in building i and unit j. $Launch_t$ is a dummy variable, and it is 1 if complaint time t is later than the launch of bus service. Otherwise, it equals to 0. In the first specification, we include $Distance_i$, a variable denoting the distance of building i to the new bus route. Therefore, the coefficient of the interaction between $Launch_t$ and $Distance_i$ is the estimate of the causal impact of closeness to the new bus route on noise complaints. In addition, we specify $Near_i$, a dummy variable indicating whether block i is within 100 meters of the actual bus route, as an alternative to $Distance_i$, in Equation (2). X_i is a vector controlling for building i's physical properties, which include the morphology of the building (slab block, point block, or new HDB block), age of the block, and its distance to MRT/LRT stations, the LRT viaduct line, bus stops, expressways, major roads, and RC Centers. These factors are common sources of residential noise. U_j is the vector controlling for unit-specific properties, including the floor level and its squared form, and the gender of the complainant. φ_t is the year times month fixed effect, and ω_i is the block fixed effect. ϵ_{ijt} is the error term. Standard errors are clustered by building blocks. The parallel trend between treatment group and comparison group in the DID model is also verified (Appendix Figure B2).

OLS estimates from Equations (1) and (2) are likely to be biased if the design of the bus service route is not random, which is highly possible in common urban management practice. The two instrumental variables described in the previous section are applied to address this concern. We use the Euclidean cost path (ECP) as the main IV and include the least cost path (LCP) in the robustness check. To facilitate the latter cost-benefit analysis in this paper on the impact of bus service on housing price, we also replicate Equations (1) and (2), with the outcome variable of housing price. Details will be discussed in Section 6.

4.2 Noise Complaints and Housing Price

Further, we examine the impact of residents' noise sentiment on housing prices. We test this effect from the yearly aggregated noise sentiment in each block on housing price.[5] The empirical specification is as follow:

$$log(Price_{ijt}) = \beta SI_{i,t-1} + X_i^{'}\theta \ + U_j^{'}\mu \ + \lambda M_t + \varphi_t + \omega_i + \epsilon_{ijt}, \tag{3}$$

where $log(Price_{ijt})$ is the log form of the total transaction price for unit j in block i sold at time t. $SI_{i,t-1}$ is the average noise sentiment in block i during the 12 months before transaction time t. The coefficient β is thus the estimate of the effect of the noise sentiment on housing price. X_i is the same vector controlling for characteristics of building block i as in Equations (1) and (2). We also include an explicit control for building i's change in accessibility due to the new bus service. U_j controls the properties of the housing unit, including its size and floor level. M_t represents the macroeconomic index; specifically, the prime lending rate for bank mortgages at time t. φ_t is year and quarter fixed effects, while ω_i is the block fixed effect. ϵ_{ijt} denotes the error term. Standard errors are clustered by building blocks.

[5] Noise compliant records are too scarce to map onto individual housing transactions.

5 Effects of Public Bus Routes on Noise Complaints

5.1 Baseline Estimation

Table 2 reports first-stage IV regression results for the effects of bus routes on noise complaints at the individual level. Columns (1) and (2) examine the numerical distance to the bus route, as specified in Equation (1), using ECP and the combination of ECP and LCP as the instrument(s), respectively. Columns (3) and (4) include the binary indicator of closeness to bus route instead, as specified in Equation (2), and also apply ECP and the combination of ECP and LCP as the instrument(s), respectively. First-stage results reveal that both ECP and the combination of ECP and LCP are strongly and statistically significantly correlated with the actual bus service route, controlling for the physical features, gender of the complainants, and the fixed effect from time and building. The F-statistics in all specifications are around 20-30, mitigating concerns about weak instruments.

Table 3 presents OLS and second-stage IV estimation results. Column (1) presents OLS estimates from Equation (1), while Columns (2) and (3) display IV estimates using ECP and the combination of ECP and LCP, respectively. Columns (4)-(6) show corresponding estimates from Equation (2). The OLS estimate is negative, though with no statistical significance, in Column (1). As discussed in Section 3.2, it is likely contaminated by unobserved factors. For example, residents who are more sensitive to noise will choose to live in units farther from roads, and will also make more noise complaints. Therefore, the OLS estimate of the distance effect will be upward biased. Using ECP as an instrument, being close to the bus service route by 100 meters results in higher negative a sentiment score of 0.095 (Column (2)), and 0.097 (Column (3)) using both ECP and LCP as IVs. Both of the two IV estimates are statistically significant at the 5% level. Since the sentiment score is normalized to the range of 0 to 1, this indicates that by living closer to the new bus route by 100 meters, the negative noise sentiment will increase by around 10 percentage points on average. The IV estimates have larger magnitudes than the OLS estimate, which also aligns with our expected direction of the bias. The binary closeness effect is estimated to be 0.109

(Column (5)) using ECP as the IV, or 0.112 (Column (6)) using the combined IVs. Both estimates are statistically significant at the 5% level. Consistent with previous findings using the continuous distance, the binary estimates reveal that intensification of public transport services worsens the noise sentiment in surrounding public housing by approximately 10 percentage points.

One assumption in the interpretations above is that the intensification of the bus service is the only source of changing noise levels along the bus route. A possible debate about the validity of this assumption is that with the newly open bus route, increased bus traffic will change the dynamics with other vehicles on the route. Commuters using private vehicles other than public buses may choose other routes to avoid congestion. However, under this scenario, the traffic volume from other vehicles is expected to decrease, and thus our estimate possibly provides a lower bound of the true estimate. In other words, the real impact of the bus service route on noise complaints is likely higher than our estimates.

5.2 Robustness Checks

One possible concern about the baseline estimation result is that an individual's sentiment about a new noise incident is dependent on previous exposure to noise. If the surrounding environment has been noisy for years, residents may not notice additional noise incidents, while residents living in quiet housing units may be less tolerant of a sudden increase in the noise level. To alleviate this concern, the average sentiment score of noise complaints made in the same building with a 1-year lag is included as an additional control. Results are reported in Appendix Table C2. The estimates remain robust in magnitude and level of significance.

Moreover, to address the potential selection issue of excluding individual cases without unit-level information, we further conduct the analysis at an aggregated building level. Specifically, using a 12-month window before and after the opening of the new bus route, we investigate the impact of bus service on average noise sentiment within buildings. This specification follows the empirical strategies applied by Faber (2014):

$$SI_i^{After} - SI_i^{Before} = \beta Distance_i + X_i^{'}\theta \ + \epsilon_i, \quad (4)$$

$$SI_i^{After} - SI_i^{Before} = \beta Near_i + X_i^{'}\theta \ + \epsilon_i, \quad (5)$$

where SI_i^{Before} and SI_i^{After} are the average sentiment scores of noise complaints in block i recorded 12 months before and after launch of the new bus service. $Distance_i$ is the distance from block i to the actual bus route, and $Near_i$ is the dummy variable indicating whether block i is within 100 meters of the actual bus route. X_i is the same vector controlling for building i's physical properties, as in the individual case-level estimation. ϵ_i is the error term. Like the case-level estimation, we use the ECP as IV for $Distance_i$ and $Near_i$ in the main regressions, and include LCP in the robustness check.

Table 4 presents the OLS and IV estimation results of Equations (4) and (5).[6] On the continuous scale, buildings closer to the bus route by every 100 meters present worsened noise annoyance by 13.2 percentage points (Column (2)) using ECP as the IV, and 14.8 percentage points (Column (3)) using the combined IVs. Living in buildings within 100 meters of the bus route has an incremental noise sentiment score of 0.204 (Column (5)) if using ECP as the IV, while the effect is 0.220 (Column (6)) if using the combined IVs. All of the IV estimates are statistically significant at conventional levels of significance. As public housing unit allocations and resales in Singapore follow a strict ethnicity and citizenship quota scheme, aggregated sentiments at building level are not likely biased by demographic distributions. Such evidence at an aggregate level is consistent with the individual-level pattern, which consolidates the negative impact of bus route on noise complaints.

[6] First-stage results are presented in Appendix Table C3.

5.3 Heterogeneous Effects across Floor Levels and Distance to Bus Stops

Transmission of sound differs across height and sound insulation at various levels. To explore the heterogenous noise effects of the bus service on noise complaints from different floor levels, we classify the samples into several categories and include the subsamples in separate regressions. Since the majority of buildings in our study areas are 12 stories, we naturally divide the samples into three categories: low floor ($1^{st} - 4^{th}$ floor), medium floor ($5^{th} - 8^{th}$ floor) and high floor (above 8^{th} floor).[7] Results are presented in Panel A of Table 5 using ECP as the IV. The empirical specification follows Equations (1) and (2). R-squared values in the estimations have improved significantly from the baseline specification, indicating that the heterogeneous floor effect is strong. For the low floor range and high floor range, the noise effect from the bus service is statistically insignificant. However, the noise effect on the medium floor range has doubled in comparison with the baseline estimation. The continuous distance effect for 100 meters is estimated as -0.243 (Column (2)), and the binary closeness negative effect is about 0.214 (Column (5)).

There are several plausible explanations for the strong noise impact on the medium floors. First, both the real noise data and the simulation results in literature support that residents living on medium floors may suffer the most from traffic noise exposure, due to the reflection of noise from the road surface and the natural attenuation of noise on higher floors (Barclay et al., 2012; Chew, 1991; Mak et al., 2010; Walerian et al., 2001). Specifically, Walerian et al. (2001) use simulation noise data and document that the noise of road transport peaks at the 5^{th} and 6^{th} floors of the surround parallel buildings with a width of approximately 40 meters in between, which is similar to the urban form in our study area. Second, it may be explained by the existing occupants at lower floors, because elderly people who are more like to be hearing impaired may also prefer to live on the lower floors. Third, it is also probably due to the sound insulation infrastructure in the local context. Using landscaping and noise barriers along the major roads (expressways) to reduce transport noise is common in many global cities including Singapore (DEFRA, 2019), but there lacks policy regulation for noise insulation along community roads and across different

[7] Our result that the noise impact is stronger on medium floors remains robust if we further divide the samples into four categories (i.e. $1^{st} - 3^{rd}$ floor, $4^{th} - 6^{th}$ floor, $7^{th} - 9^{th}$ floor, and above 9^{th} floor).

floors. Specifically, in the public housing of Singapore, there is normally no soundproofing design across different floors. Noise insulation infrastructures are only installed along rail tracks and expressways, but not along the community roads in our study area (LTA, 2015). As a result, only the shade trees along the roads and some noise tolerant buildings (i.e. multiple-floor car parks) may serve as the natural noise barriers in the study site (Bin et al., 2019), and based on our field study, all these natural barriers are around three to four stories in height. In summary, our result implies the need for further investigation to improve urban noise prevention for medium floor units in Singapore's public housing as well as in the urban context of other cities beyond, especially considering the random assignment of new units to public housing applicants.

In addition, we examine how distance to a bus stop impacts noise complaints, as bus stops are special nodes on the entire bus route. Samples within 100 meters of a bus stop are separated from those outside the 100-meter radius in estimations, using ECP as IV. Results are presented in Panel B of Table 5. Columns (1) and (2) show estimates using distance variations, while Columns (3) and (4) present estimates using the dummy variable measuring closeness. The effects at buildings within 100 meters of a bus stop show larger magnitude (-0.288 for the continuous distance variable and 0.511 for the binary closeness variable), with statistical significance at the 1% level. This indicates that the noise effect from bus service is more severe around bus drop-off points. A possible explanation is the influx of visitors around bus stops. In addition, it is possible that buses generate more noise when they stop at or leave from the bus stops. Also, with frequent visual exposure to noise sources at bus stops, subjective feelings of annoyance by surrounding residents may be intensified.

6 Cost-Benefit Analysis: Public Transport and Housing Price

Public transport offers convenience to surrounding residents, and thus raises housing price in the neighborhood. Nonetheless, it also generates noise pollution, which exerts a negative impact on the housing price. To implement a cost benefit analysis, we start with an estimation on the overall impact of bus routes on housing price. Specifically, we follow the same difference-in-differences specification in Equation (1), but replace the dependent variable of noise sentiment with housing price (Diao et al., 2017):

$$log(Price_{ijt}) = \beta_1 Distance_i + \beta_2 Launch_t + \beta_3 Launch_t * Distance_i + X_i^{'}\theta \ + U_j^{'}\mu \ + \varphi_t + \omega_i + \epsilon_{ijt}. \quad (6)$$

The definitions for the other variables are the same as in Equation (1). Appendix Table C4 displays the corresponding results using all of the resale transaction data in our study area from 2000 to 2017. Before the bus starts operation, the prices for housing units closer to the bus route by 100 meters are 1.04% lower. This is probably because the route is designed along the community road, generating neighborhood noise but offering limited convenience before the new bus service begins. After the introduction of bus service, the prices for units closer to the bus route by 100 meters increase by 1.34%. The estimates are statistically significant at a high 1% level. A similar conclusion can be drawn if we use a binary indicator of closeness to estimate the benefit from the convenience of public transport on housing prices within 100 meters and within 100 to 200 meters of the bus route.

Nevertheless, this explicit benefit is likely to be offset by an implicit decrease in housing price due to the noise generated by vehicles on the new bus route. We thus examine the effect of noise sentiment on housing price using the same set of resale transaction records. The OLS estimate of Equation (3) may serve as an upper bound of the true effect of noise pollution on housing price, since it is not able to differentiate the positive impact of improving accessibility on housing price and the negative impact of noise pollution on housing price. In order to estimate the exact economic costs of noise pollution on housing price, we include an explicit control for accessibility before and after the launch of bus No.972. Specifically, we use distance to the nearest bus stop connecting toa the Central Business District (CBD) of Singapore as a proxy for change in accessibility. As illustrated in Appendix Figure B3a, the new bus route No.972, which zigzags within the residential community, aims to solve last-mile connectivity to CBD. Before the launch of bus No.972, the only bus connecting the site and the CBD was route No.190, which operates along a major road

(Appendix Figure B3b). Therefore, before the launch of the new bus route, the nearest bus stop connecting to the CBD lies only on the route of No.190. After the treatment, it will be on either bus route No.190 or bus route No.972.[8]

Table 6 presents the corresponding OLS regression results. In Column (1), we include the same 47 building blocks as sampled in the building-level estimation in Section 5.2. This reveals that a 1-scale-point increase in noise sentiment over the past 12 months is correlated with a 3.26% decrease in housing price. This estimate is statistically significant at the 5% level. In Column (2), all the buildings with transaction records in our study period are included. The point estimate on the exact impact of noise sentiment on housing price is -0.03, with statistical significance at the 5% level.

Finally, we analyze the cost and benefit for intensifying public transport on housing price. Since living closer to the new bus service route by 100 meters has caused the noise sentiment score to increase by around 9.53% (Table 3), the negative effect of its noise on housing price can be interpolated as approximately 0.31% per 100-meter distance to the bus route. Taking the explicit 1.34% increase in housing price due to the new bus route, we estimate that the implicit effect of traffic noise on housing value has offset around 18.79% of the economic benefit brought by the improvement of public transport.

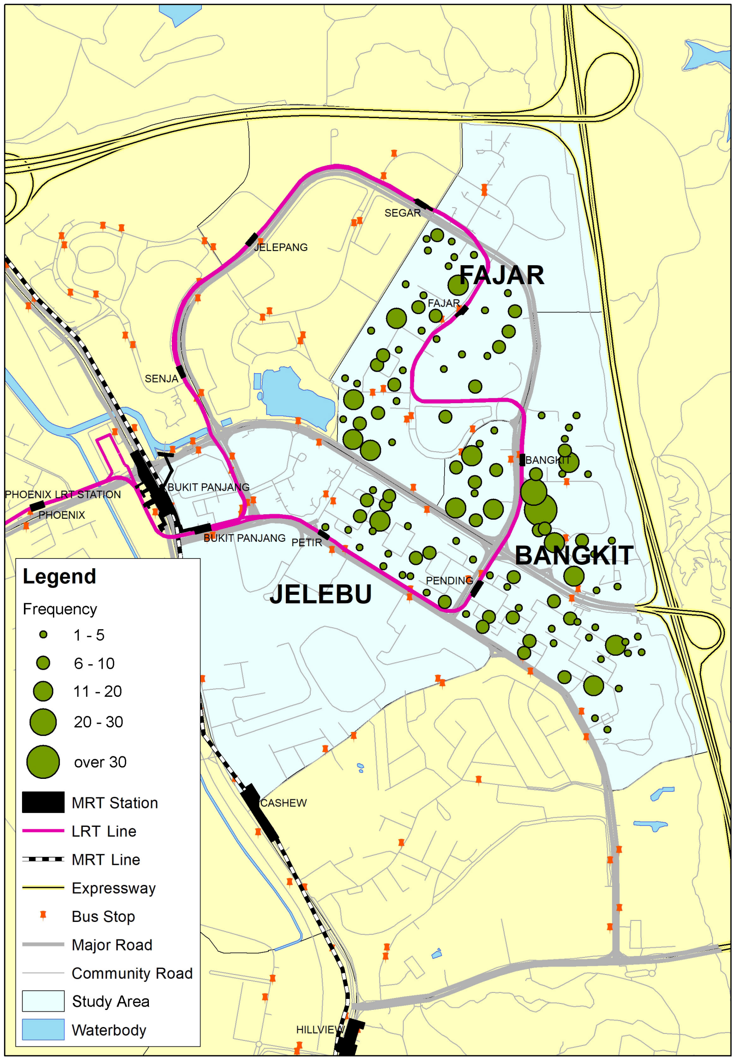
7 Conclusion

Public transport is supposed to raise housing price in the neighborhood because of its convenience. The byproduct of excessive noise pollution from public transport, however, may offset its benefit to the region, and such effect is considered more severe in urban areas with high density. Using the opening of a new bus route zigzagging through a public housing neighborhood in Singapore, we contribute to the literature on public transport and housing price by presenting new empirical evidence in the form of a cost-benefit analysis (Baum-Snow and Kahn, 2000; McMillen and McDon-

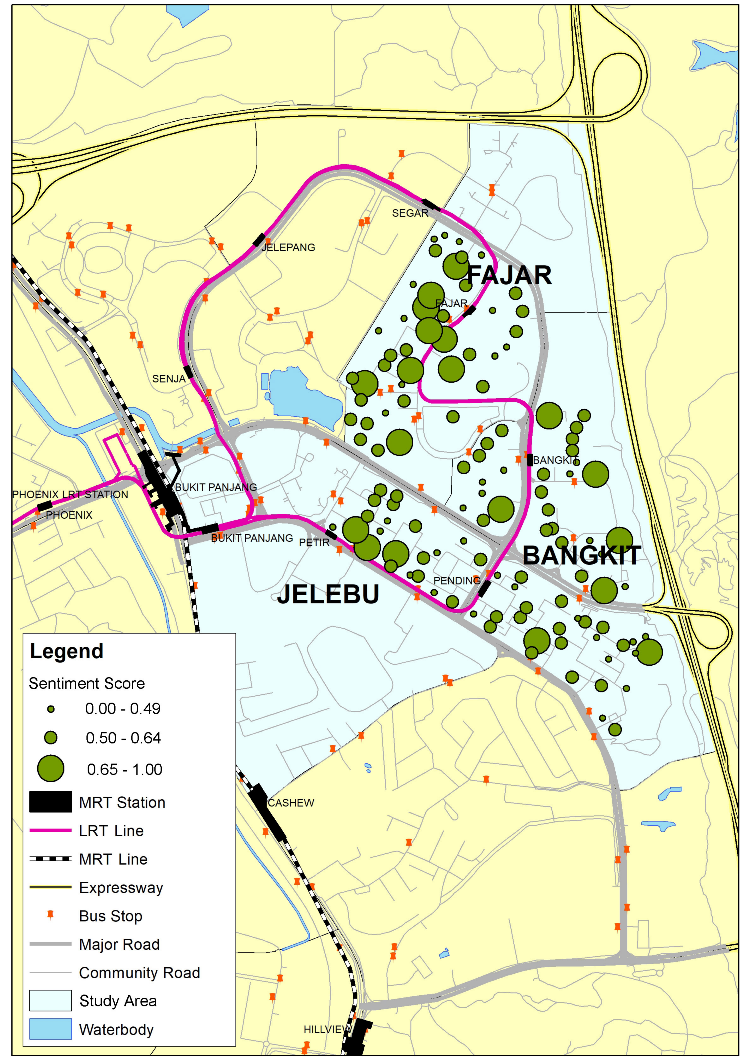
[8]After the treatment, 63 out of the 142 blocks in our study area have the nearest bus stop connecting to the CBD on the new route No. 972.

ald, 2004; Xu et al., 2015). We also use the natural language processing tools to generate real-time noise sentiments, which exhibits patterns that are different from and complementary to those in the literature that counts the frequency of noise incidents in retrospective surveys (Dzhambov and Dimitrova, 2014; Weinhold, 2013). Applying an IV estimation, we show that the intensification of public transport service significantly increases residents' noise complaints. Individuals living within 100 meters of a the bus route are about 11 percentage points more likely to make noise complaints to government agencies than those living further away. This adverse effect is more serious at median floor levels and locations near bus stops. We further link noise complaints to housing transactions. With the change in accessibility explicitly controlled for, it is estimated that traffic noise offsets 21.8% of the economic benefits brought by the improvement of public transport.

Our results shed light on urban policies to address the negative impact of noise pollution on economic development. Many global cities are undergoing intensification in rapid urbanization (Searle and Filion, 2011). While the congestion caused by private transport has always been an issue, economies of scale from intensified population encourages governments to intensify public transport as a solution to congestion (Goetzke, 2008). Services like public buses, which aim to ease last-mile mobility from trains or subways, better connect with residential buildings. However, as demonstrated in this paper, they may introduce adverse noise effects as well. The results of this paper provide an empirical basis for policy design, such as a differentiated residential noise insulation infrastructure, for urban environments in high-density cities like Singapore.

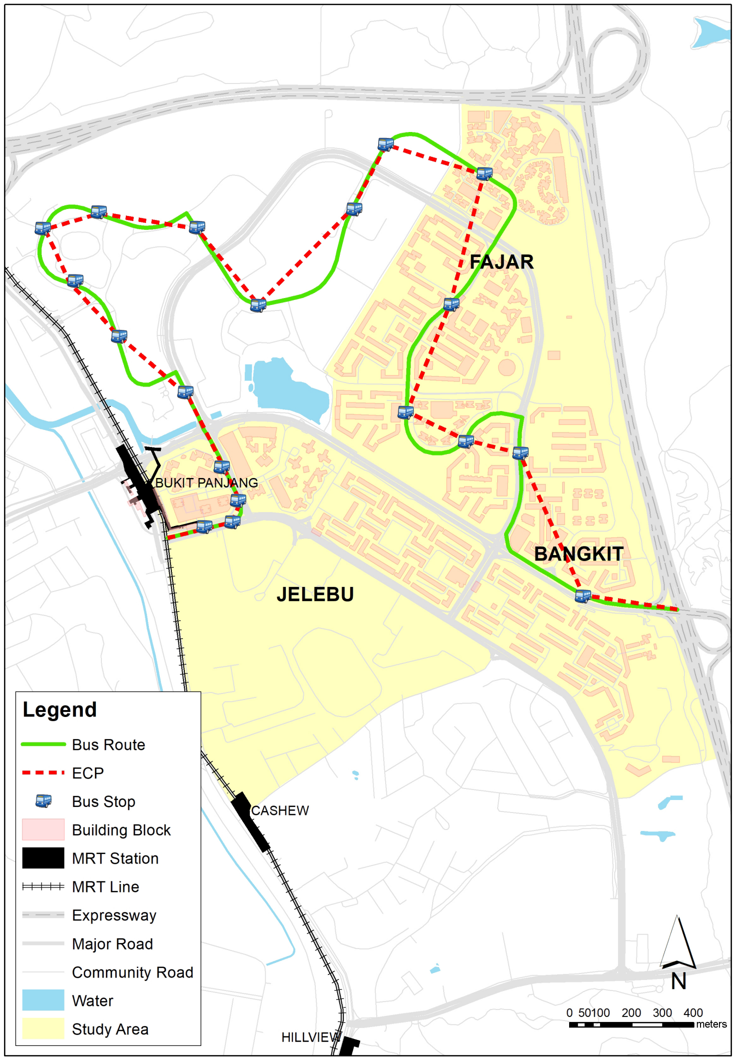


(a) Complaint Frequency

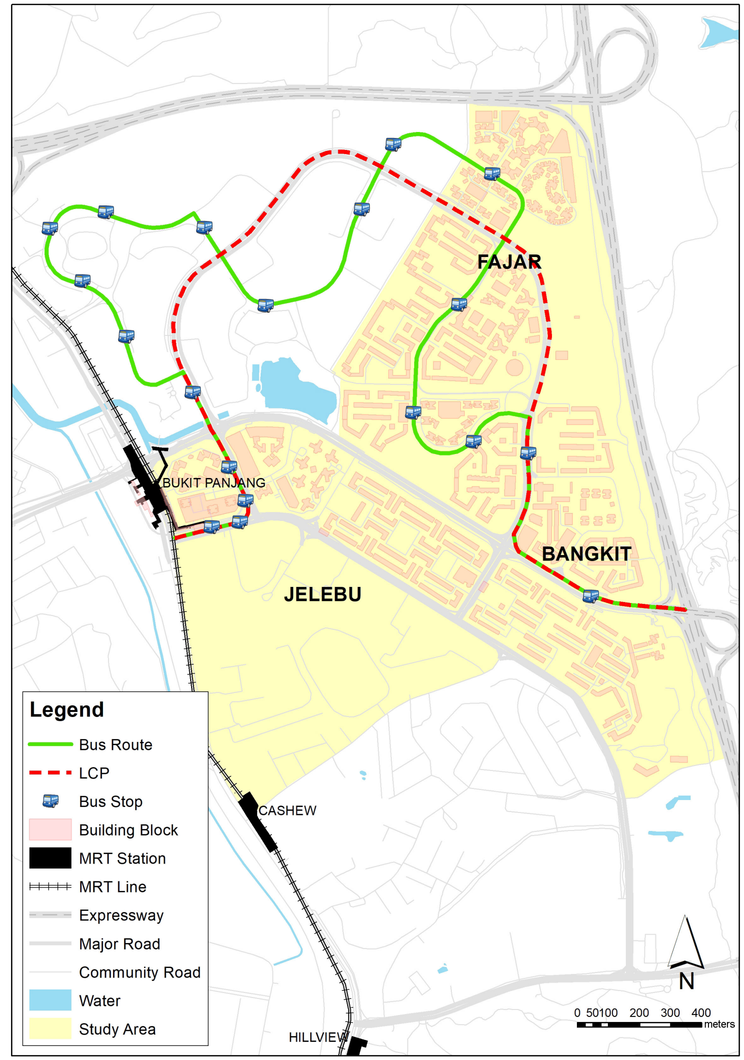


(b) Complaint Severity

Figure 1: Map of Noise Complaint Frequency (a) and Sentiment Severity (a) in Singapore's Bukit Panjang Area during 2010-2018



(a) ECP and Actual Bus Route



(b) LCP and Actual Bus Route

Figure 2: Actual Service Route of Bus No.972 and the Constructed Theoretical Least Cost Path of (a) ECP and (b) LCP.

Table 1: Summary Statistics of Noise Complaints and Housing Transactions

		Total			Near = 0 (100m - 200m)			Near = 1 (within 100m)		Diff((5)-(8))
	N	mean	SD	N	mean	SD	N	mean	SD	t-test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sentiment	593	0.409	0.165	328	0.416	0.170	265	0.399	0.159	0.017
Distance (100m)	593	1.111	0.493	328	1.483	0.317	265	0.65	0.187	0.833***
LCP (100m)	593	1.737	1.323	328	2.136	1.349	265	1.243	1.108	0.893***
ECP (100m)	593	1.068	0.794	328	1.516	0.649	265	0.514	0.578	1.001***
Launch (after = 1)	593	0.575	0.495	328	0.512	0.501	265	0.653	0.477	-0.141***
Near (within 100m = 1)	593	0.447	0.498							
RC Center (with = 1)	593	0.069	0.254	328	0.0671	0.251	265	0.072	0.258	-0.005
Building Age	593	29.02	2.267	328	28.9	2.584	265	29.17	1.794	-0.267
Floor Level	593	6.408	3.960	328	6.479	3.976	265	6.321	3.946	0.158
Morphology	593	2.926	0.293	328	2.878	0.371	265	2.985	0.122	-0.107***
Gender (Male = 1)	527	0.467	0.499	300	0.503	0.501	227	0.419	0.494	0.085*
To train station (100m)	593	2.125	1.155	328	2.435	1.224	265	1.741	0.933	0.693***
To LRT line (100m)	593	1.531	1.154	328	1.884	1.155	265	1.093	0.994	0.790***
To bus stop (100m)	593	1.202	0.501	328	1.335	0.548	265	1.038	0.377	0.233***
To expressway (100m)	593	4.281	2.030	328	4.553	2.357	265	3.945	1.468	0.608***
To major road (100m)	593	1.217	0.816	328	1.359	0.758	265	1.041	0.851	0.318***
To shopping center (100m)	593	5.302	2.863	328	5.181	3.145	265	5.452	2.467	-0.271
Price (SGD)	1,450	375,207	73,564	1,003	364,924	69,114	447	398,282	77,968	-33,359***
Floor	1,450	6.710	3.630	1,003	6.671	3.588	447	6.799	3.724	-.1277
Area (sq.m.)	1,450	103.595	20.399	1,003	99.984	19.56	447	111.696	19.928	-11.7117***

Notes: Morphology equals to 1 if the block follows the new generation building layout, 2 if it follows the point block layout, or 3 if it follows the slab block layout.
*** p<0.01, ** p<0.05, * p<0.1

Table 2: First-stage Results of the Effect of Bus Route on Noise Complaints at Individual Level

	(1) Launch*Distance	(2) Launch*Distance	(3) Launch*Near	(4) Launch*Near
Launch*ECP	0.4184***	0.4025***	-0.3653***	-0.3306***
	(0.0520)	(0.0602)	(0.0546)	(0.0615)
ECP	-0.0070	-0.0491	0.0190	0.0367
	(0.0364)	(0.0395)	(0.0419)	(0.0451)
Launch*LCP		0.0491		-0.0787**
		(0.0352)		(0.0360)
LCP		0.1558**		-0.0790
		(0.0644)		(0.0651)
Launch	0.8278***	0.7932***	0.7781***	0.8410***
	(0.2051)	(0.2057)	(0.2049)	(0.2067)
Building Age	-0.0117	-0.0158	-0.0040	-0.0023
	(0.0104)	(0.0102)	(0.0119)	(0.0119)
Floor	0.0029	0.0033	0.0007	-0.0000
	(0.0102)	(0.0105)	(0.0123)	(0.0126)
Floor Squared	-0.0008	-0.0008	0.0005	0.0005
	(0.0006)	(0.0006)	(0.0007)	(0.0007)
Point Block	0.3776**	0.3339**	0.0401	0.0605
	(0.1789)	(0.1692)	(0.1951)	(0.1908)
Slab Block	0.1492	0.0789	0.0943	0.1315
	(0.1397)	(0.1330)	(0.1448)	(0.1457)
Male	0.0156	0.0149	-0.0402	-0.0405
	(0.0239)	(0.0228)	(0.0299)	(0.0286)
RC Center	-0.0135	-0.0083	-0.0094	-0.0105
	(0.0684)	(0.0692)	(0.0796)	(0.0803)
First Satge F-Stats	32.67	38.07	22.69	26.81
Observations	527	527	527	527

Notes: Columns (1) and (2) report results for the continuous distance variable. Columns (3) and (4) report results for the binary closeness indicator. Columns (1) and (3) use the ECP, the Euclidean straight lines connecting bus stops, as the instrument. Columns (2) and (4) use both ECP and LCP, the least travel time path, as the instrument. Unreported control variables include the distance to train stations, LRT viaduct line, bus stops, expressways, major roads, and shopping centers. Standard errors are clustered by building blocks. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: OLS and IV Estimates on Effects of the Bus Route on Noise Complaints at Individual Level

	(1)	(2)	(3)	(4)	(5)	(6)
			Noise Sentiment			
	OLS	ECP IV	Both IVs	OLS	ECP IV	Both IVs
Launch*Distance	-0.0405	-0.0953**	-0.0965**			
	(0.0338)	(0.0430)	(0.0422)			
Distance	0.0074	0.0125	0.0301			
	(0.0328)	(0.0462)	(0.0423)			
Launch*Near				0.0496	0.1090**	0.1118**
				(0.0324)	(0.0459)	(0.0437)
Near				-0.0136	-0.0105	-0.0300
				(0.0303)	(0.0529)	(0.0476)
Launch	-0.0400	0.0319	0.0333	-0.1123	-0.1330	-0.1365
	(0.1167)	(0.1140)	(0.1114)	(0.1217)	(0.1069)	(0.1054)
Building Age	-0.0019	-0.0024	-0.0022	-0.0016	-0.0011	-0.0015
	(0.0048)	(0.0044)	(0.0043)	(0.0048)	(0.0044)	(0.0044)
Floor	-0.0128*	-0.0122*	-0.0125**	-0.0130*	-0.0126**	-0.0128**
	(0.0071)	(0.0063)	(0.0063)	(0.0070)	(0.0063)	(0.0063)
Floor Squared	0.0008*	0.0007*	0.0007*	0.0008*	0.0007*	0.0007*
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Point Block	0.0171	0.0166	0.0156	0.0181	0.0195	0.0186
	(0.0199)	(0.0187)	(0.0186)	(0.0201)	(0.0188)	(0.0186)
Slab Block	-0.0423	-0.0457	-0.0415	-0.0487	-0.0764	-0.0593
	(0.0792)	(0.0739)	(0.0728)	(0.0811)	(0.0786)	(0.0786)
Male	-0.0409	-0.0484	-0.0393	-0.0429	-0.0675	-0.0505
	(0.0508)	(0.0475)	(0.0480)	(0.0543)	(0.0549)	(0.0554)
RC Center	0.0453*	0.0420*	0.0463*	0.0484*	0.0448*	0.0501**
	(0.0260)	(0.0236)	(0.0239)	(0.0269)	(0.0251)	(0.0251)
Block & Time Fixed Effect	Y	Y	Y	Y	Y	Y
Observations	527	527	527	527	527	527
R-squared	0.213	0.206	0.208	0.214	0.203	0.207

Notes: Columns (1)-(3) report results for the continuous distance variable and columns (4)-(6) report results for the binary closeness indicator. Columns (1) and (4) are OLS estimation results. Columns (2) and (5) use the ECP, the Euclidean straight lines connecting bus stops, as the instrument. Columns (3) and (6) use both ECP and LCP, the least travel time path, as the instrument. Unreported control variables include the distance to train stations, LRT viaduct line, bus stops, expressways, major roads, and shopping centers. Standard errors are clustered by building blocks. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: OLS and IV Estimates on Effects of the Bus Route on Noise Complaints at Building Level

	(1)	(2)	(3)	(4)	(5)	(6)
			Noise Sentiment			
	OLS	ECP IV	Both IVs	OLS	ECP IV	Both IVs
Distance	-0.1095*	-0.1324*	-0.1484**			
	(0.0605)	(0.0741)	(0.0700)			
Near				0.0765	0.2037*	0.2201**
				(0.0660)	(0.1175)	(0.1056)
Building Age	0.0114	0.0112	0.0111	0.0127	0.0136	0.0137
	(0.0148)	(0.0129)	(0.0129)	(0.0147)	(0.0131)	(0.0131)
Slab Block	-0.0686	-0.0707	-0.0723	-0.0545	-0.0483	-0.0475
	(0.0780)	(0.0708)	(0.0721)	(0.0751)	(0.0785)	(0.0809)
RC Center	0.0687	0.0640	0.0606	0.0726	0.0411	0.0370
	(0.0579)	(0.0538)	(0.0538)	(0.0631)	(0.0761)	(0.0759)
First-stage F-stats		25.62	17.66		8.94	6.04
Observations	47	47	47	47	47	47
R-squared	0.245	0.242	0.237	0.215	0.119	0.092

Notes: Columns (1)-(3) report results for the continuous distance variable and columns (4)-(6) report results for the binary closeness indicator. Columns (1) and (4) are OLS estimation results. Columns (2) and (5) use the ECP, the Euclidean straight lines connecting bus stops, as the instrument. Columns (3) and (6) use both ECP and LCP, the least travel time path, as the instrument. Unreported control variables include the distance to train stations, LRT viaduct line, bus stops, expressways, major roads, and shopping centers. Standard errors are clustered by building blocks. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Heterogenous Effects of Bus Route on Noise Complaints

Panel A. Noise Sentiment Across Floor Levels

	(1) 1-4 Storey	(2) 5-8 Storey	(3) Above 8 Storey	(4) 1-4 Storey	(5) 5-8 Storey	(6) Above 8 Storey
Launch*Distance	-0.0552	-0.2428***	0.0282			
	(0.0527)	(0.0820)	(0.0608)			
Launch*Near				0.0738	0.2141***	-0.0290
				(0.0674)	(0.0799)	(0.0802)
Block & Time Fixed Effect	Y	Y	Y	Y	Y	Y
First-stage F-stats	32.74	20.66	11.65	14.83	18.11	7.85
Observations	221	144	162	221	144	162
R-squared	0.479	0.598	0.564	0.477	0.605	0.544

Panel B. Noise Sentiment with Different Distance to Bus Stops

	(1) Within 100m	(2) Out of 100m	(3) Within 100m	(4) Out of 100m
Launch*Distance	-0.2875***	-0.1330**		
	(0.0853)	(0.0607)		
Launch*Near			0.5114***	0.1315*
			(0.1924)	(0.0739)
Block & Time Fixed Effect	Y	Y	Y	Y
First-stage F-stats	38.59	26.27	17.24	21.35
Observations	179	348	179	348
R-squared	0.590	0.314	0.447	0.274

Notes: Columns (1)-(3) of Panel A report results for the continuous distance variable. Columns (4)-(6) of Panel A report results for the binary closeness indicator. Columns (1) and (2) of Panel B report results for the continuous distance variable. Columns (3) and (4) of Panel B report results for the binary closeness indicator. All estimations use ECP, the Euclidean straight lines connecting bus stops, as the instrument. Unreported control variables include building age, floor and floor squared, gender, morphology, RC center, and the distance to train stations, LRT viaduct line, bus stops, expressways, major roads, and shopping centers. Standard errors are clustered by building blocks. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Effects of Noise Complaints on Housing Price

	(1)	(2)
	log(price)	
	sampled blocks	all blocks
Sentiment	-0.0326**	-0.0300**
	(0.0160)	(0.0149)
Distance to Nearest Bus Stop Connecting CBD	0.0005	-0.0027*
	(0.0030)	(0.0016)
Prime Lending Rate	-2.4724***	-3.5344***
	(0.3513)	(0.4423)
Floor	0.0067***	0.0064***
	(0.0007)	(0.0005)
Area	0.0075***	0.0075***
	(0.0004)	(0.0002)
Building Age	-0.0018	-0.0035***
	(0.0012)	(0.0009)
Block & Time Fixed Effect	Y	Y
Observations	488	893
R-squared	0.927	0.919

Notes: Column (1) includes the same 47 building blocks used to estimate the effect of bus route on block level. Column (2) includes all the building blocks in the study area. Unreported control variables include the distance to train stations, LRT viaduct line, bus stops, expressways, major roads, and shopping centers. Standard errors are clustered by building blocks. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix A: Selection of Sentiment Analysis Toolkit and the Algorithm of SentimentR

Given the requirement to more accurately capture sentiment variations, a lexicon-based method is applied in this work. It is then essential to select the most suitable toolkit for the context of our study. Over 50 types of sentiment scoring tools have been developed so far[9] and are designed for specific context, such as evaluating the sentiment of comments on movies or the posts on Twitter. By comparing the functional specification (Abbasi et al., 2014) and classification accuracy of these tools[10], five tools—SentimentR (Rinker, 2016), SentiWordNet (Esuli and Sebastiani, 2007), SentiWordNet3.0 (Baccianella et al., 2010), VaderSentiment(Gilbert, 2014) and IBM Watsons[11]—selected for the pre-test.

SentiWordNet is a Python-based open-source tool that is widely acknowledged to be effective by including neutral emotion as well as polarized emotions, and SentiWordNet 3.0 is the updated version of it with training on a larger database. Although the literature has shown that including a neutral class of vocabulary improves classification accuracy for general content (Koppel and Schler, 2006; Hamed et al., 2016), our contents are already negatively polarized and it is therefore less effective for capturing the variance than SentimentR. VaderSentiment is another Python-based tool, but with special focus on social media text instead of using customer reviews, as SentimentR does. It turns out that VaderSentiment also underperforms in capturing variance in our context. IBM Watson, the famous java-based commercial toolkit, focuses more on its ability to capture the extremes of emotions. Although it performs well in identifying the most intense negative emotions, it also ends up with a large proportion of zero scores (neutral) in the results such that we are not able to capture the continuous distribution of the sentiment in the middle range.

[9]The Natural Language Toolkit (NLTK), a suite of libraries developed by Steven Bird and Edward Loper in the Department of Computer and Information Science at the University of Pennsylvania, summarizes up-to-date tools for natural language processing in English. The full set of tools and their demonstration can be retrieved from `https://www.nltk.org/`.

[10]SemantAPI(http://semantapi.com/) offers a free, open-source toolkit that allows easy comparison of the most popular NLP and sentiment analysis solutions.

[11]IBM Watsons Tone Analyzer: https://www.ibm.com/watson/services/tone-analyzer/

SentimentR is a R language-based tool developed by Rinker (2016) and is based on a sentiment lexicon dictionary containing a list of approximately 6,800 positive and negative sentiment words generated from customer reviews (Hu and Liu, 2004). This tool has been cited as effective for capturing emotion about social and economic issues with fast speed at the document level (Lam, 2016). It is believed to be the best fit for our context. Although this tool was originally developed for the opinion mining at the document level, it has also shown strong power in classification at the sentence level. Using different tools normally results in the same sign of estimates, but the statistical significance may vary due to their different focuses in capturing sentiment. These differences lead to context-specific weighting toward polarization, and therefore result in different standard errors when evaluating the significance.

The detailed algorithm for SentimentR is presented as follows.

For a certain segment of text under examination, it can be decomposed into individual element sentences s_i, and W_i refers to the list of individual words in the sentence such that $W_i = \{w_{i1}, w_{i2}, \ldots, w_{ij}, \ldots, w_{iJ}\}$, where j is the order of the word in the sentence and J is the total length of the sentence. Pause punctuation, including commas, colons, and semicolons, are also decoded as cw and the other punctuation is eliminated. SentimentR tool uses the lexicon developed by Hu and Liu (2004), and the words are tagged with a score of either +1 or -1 to form polarized words (pw), which represent positive and negative emotions, respectively. Further, scoring is modified based on the surrounding clustering of vocabularies, which are termed as "contextual valence shifters". The clustering of context is defined as two words before and four words after polarized words, denoted as $c_{i,j} = \{pw_{i,j-2}, \ldots, pw_{i,j}, \ldots, pw_{i,j+4}\}$. Valence shifts, or the clustering of vocabularies, further modify the sentiment of the original polarization score in four ways: "neutral"; "negators" (e.g., adding "not" in front of an adjective); "amplification" (e.g., adding "very" or "extremely"); or "deamplifications" (e.g., using "hardly"). They are represented as neutral ($w^0_{i,j}$), negator ($w^n_{i,j}$), amplifier ($w^a_{i,j}$) or deamplifier ($w^d_{i,j}$), respectively. Weightings are proposed by Rinker for these clusters and logical calculations are applied to determine the sign of the effect if multiple numbers of valence shifts exist. Finally, it is possible to calculate the sentiment score of the sentence δ_i

by summing these weighted valence shifts and averaging the sum by the square root of the word count.

The above methodology proposed to calculate the sentiment score on the sentence level can be simplified as follows:

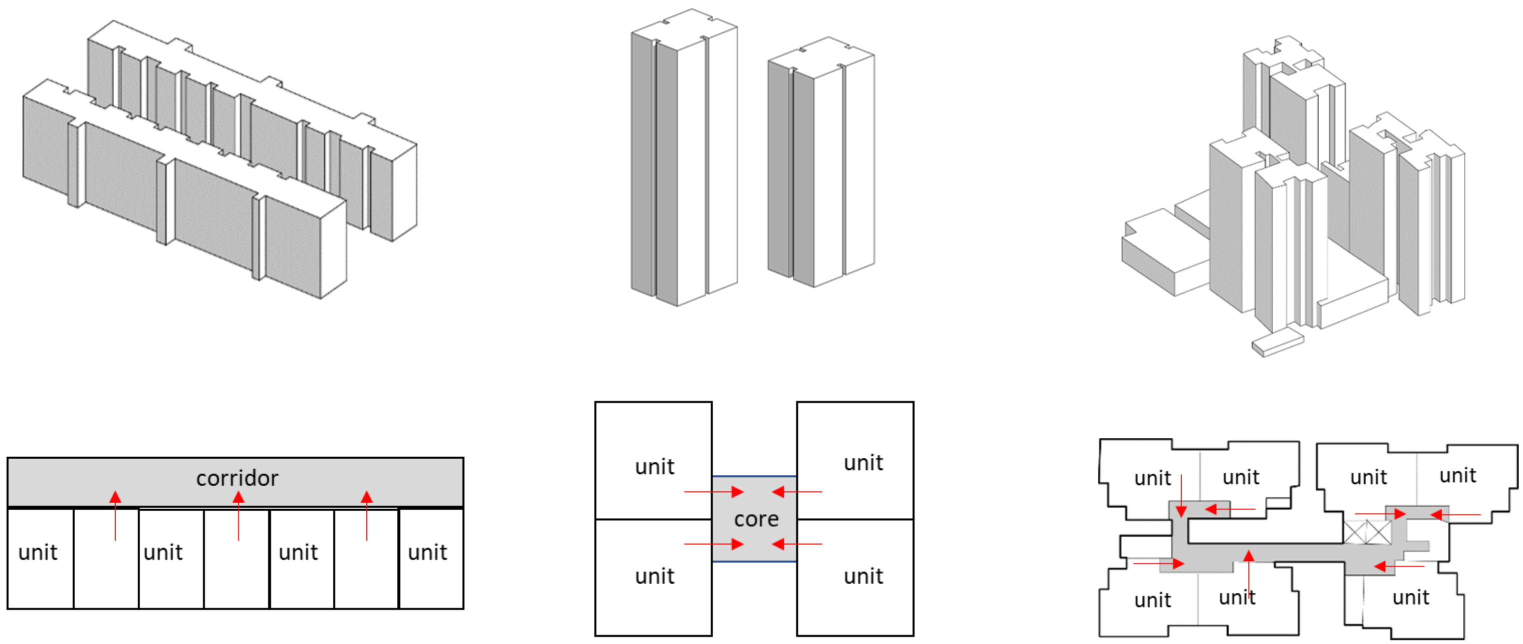
$$\delta_i = \sum c'_{ij} / \sqrt{J} \tag{7}$$

where:

$$c'_{ij} = \sum((1 + w_{amp} + w_{deamp}) \cdot w_{i,j}(-1)^{2+w_{neg}}) \tag{8}$$

If the measurement is conducted for a document, the average of sentences is further conducted for paragraphs and for the whole article sequentially.

Appendix B: Supplementary Figures



(a) Slab Block **(b)** Point Block **(c)** New Block

Figure B1: Morphological Prototypes of Public Housing Blocks in the Study Area

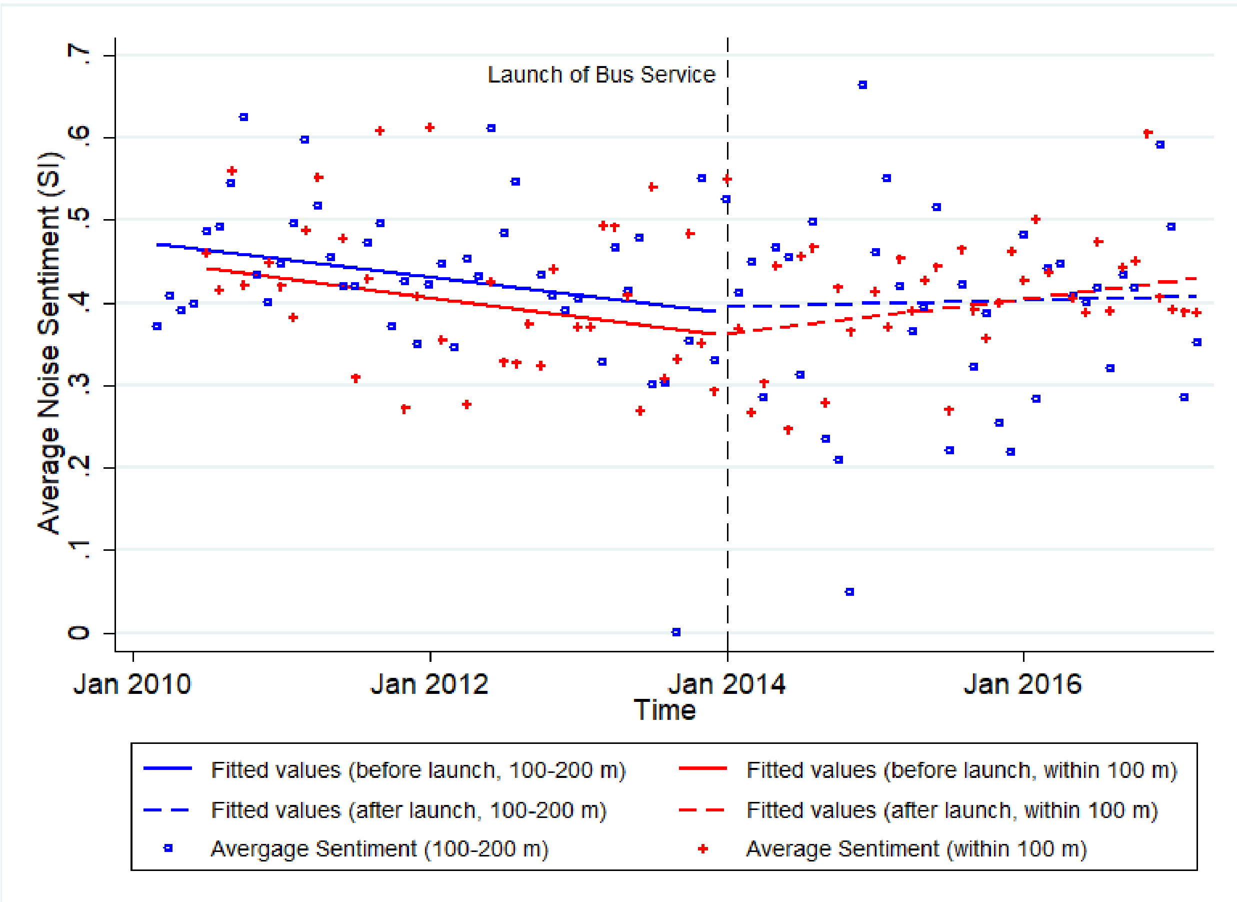
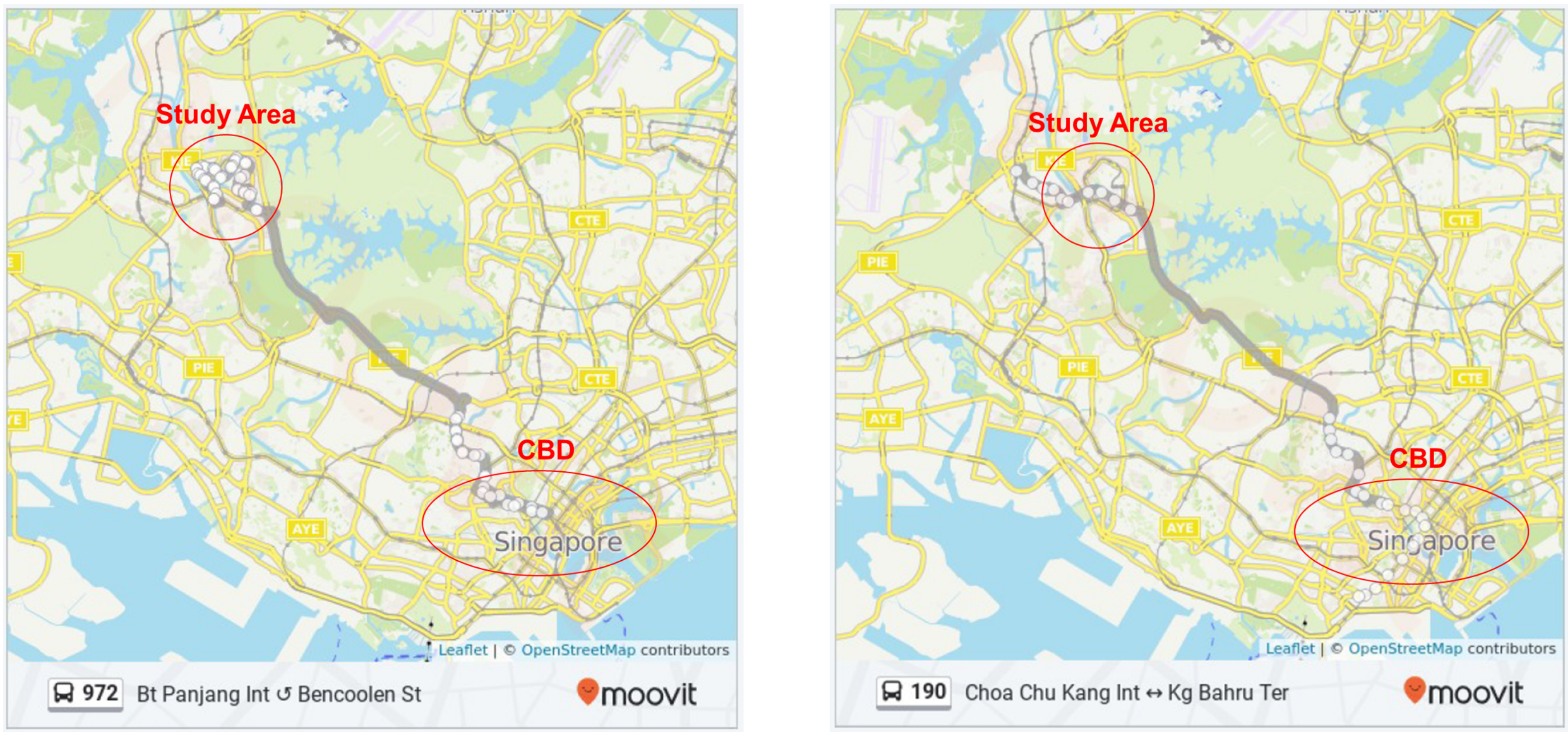


Figure B2: By-month Average of Noise Sentiment for Units within 100 Meters and within 100-200 Meters of a Bus Route



(a) Bus 972 (New Bus Line)

(b) Bus 190 (Existing Bus Line)

Figure B3: Service Route of Bus 972 and Bus 190

Source: Moovit Singapore

Appendix C: Supplementary Tables

Table C1: Summary Statistics of Noise Complaints at Building Level

	Total			Near = 0 (out of 100m)			Near = 1 (within 100m)			Diff ((5)-(8))
	N	mean	sd	N	mean	sd	N	mean	sd	t-test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Diff_sentiment	47	-0.030	0.182	28	-0.054	0.181	19	0.004	0.183	-0.057
Sentiment_before	47	0.436	0.155	28	0.433	0.152	19	0.441	0.163	-0.007
Sentiment_after	47	0.406	0.120	28	0.380	0.126	19	0.445	0.102	-0.065*
Distance (100m)	47	1.146	0.508	28	1.501	0.300	19	0.624	0.200	0.877***
LCP (100m)	47	2.000	1.417	28	2.254	1.320	19	1.625	1.506	0.629*
ECP (100m)	47	1.307	0.731	28	1.610	0.695	19	0.861	0.539	0.748***
RC Center (with = 1)	47	0.064	0.247	28	0.036	0.189	19	0.105	0.315	-0.070
Building Age	47	28.810	2.787	28	28.790	2.820	19	28.840	2.814	-0.056
Morphology	47	1.957	0.204	28	1.929	0.262	19	2.000	0.000	-0.071
Near (within 100m = 1)	47	0.404	0.496							
To train station (100m)	47	2.310	1.186	28	2.460	1.250	19	2.090	1.080	0.371
To LRT line (100m)	47	1.708	1.225	28	1.919	1.192	19	1.397	1.238	0.523
To bus stop (100m)	47	1.228	0.529	28	1.349	0.560	19	1.049	0.434	0.300*
To expressway (100m)	47	4.526	2.269	28	4.760	2.468	19	4.182	1.953	0.578
To major road (100m)	47	1.356	0.983	28	1.383	0.827	19	1.317	1.199	0.824
To shopping center (100m)	47	5.200	3.123	28	5.156	3.162	19	5.266	3.150	0.908

Notes: Morphology equals to 1 if the block follows the new generation building layout, 2 if it follows the point block layout, or 3 if it follows the slab block layout. Floor and Area refer to the corresponding feature of the sold units. *** p<0.01, ** p<0.05, * p<0.1

Table C2: Effects of Bus Route on Noise Complaints at Individual Level: With the Control for Previous Noise Level

	(1)	(2)	(3)	(4)	(5)	(6)
			Noise Sentiment			
	OLS	ECP IV	Both IVs	OLS	ECP IV	Both IVs
Launch*Distance	-0.0367	-0.1511**	-0.1500**			
	(0.0430)	(0.0726)	(0.0698)			
Distance	-0.0141	0.0182	0.0477			
	(0.0431)	(0.0748)	(0.0672)			
Launch*Near				0.0396	0.1767**	0.1602**
				(0.0353)	(0.0859)	(0.0743)
Near				-0.0096	-0.0150	-0.0404
				(0.0346)	(0.0782)	(0.0669)
Launch	0.0579	0.1968	0.2075	0.0151	-0.0256	-0.0095
	(0.1594)	(0.1747)	(0.1739)	(0.1425)	(0.1321)	(0.1299)
Sentiment in Previous Year	-0.0845	-0.0565	-0.0625	-0.0846	-0.0318	-0.0434
	(0.1161)	(0.1074)	(0.1089)	(0.1161)	(0.1074)	(0.1075)
Building Age	-0.0027	0.0000	0.0009	-0.0020	0.0013	0.0004
	(0.0058)	(0.0060)	(0.0057)	(0.0055)	(0.0056)	(0.0053)
Floor	-0.0085	-0.0096	-0.0097	-0.0084	-0.0092	-0.0091
	(0.0108)	(0.0093)	(0.0093)	(0.0108)	(0.0095)	(0.0094)
Floor Squared	0.0008	0.0008	0.0008	0.0008	0.0007	0.0008
	(0.0007)	(0.0006)	(0.0006)	(0.0007)	(0.0006)	(0.0006)
Male	0.0212	0.0178	0.0176	0.0229	0.0277	0.0256
	(0.0273)	(0.0226)	(0.0223)	(0.0270)	(0.0244)	(0.0233)
Point Block	0.0199	-0.0113	-0.0031	0.0279	-0.0326	0.0034
	(0.1427)	(0.1166)	(0.1132)	(0.1422)	(0.1232)	(0.1179)
Slab Block	-0.0028	-0.0174	-0.0039	0.0140	-0.0103	0.0189
	(0.1041)	(0.0796)	(0.0767)	(0.1051)	(0.0876)	(0.0841)
RC Center	0.0138	0.0112	0.0229	0.0256	0.0195	0.0301
	(0.0375)	(0.0389)	(0.0383)	(0.0399)	(0.0407)	(0.0399)
Block & Time Fixed Effect	Y	Y	Y	Y	Y	Y
First-stage F-Stats		19.00	32.69		7.26	12.41
Observations	311	311	311	311	311	311
R-squared	0.273	0.255	0.261	0.270	0.225	0.246

Notes: Columns (1)-(3) report results the continuous distance variable and columns (4)-(6) report results for for the binary closeness indicator. Columns (1) and (4) are OLS estimation results. Columns (2) and (5) use the ECP, the Euclidean straight lines connecting bus stops, as the instrument. Columns (3) and (6) use both ECP and LCP, the least travel time path, as the instrument. Unreported control variables include the distance to train stations, LRT viaduct line, bus stops, expressways, major roads, and shopping centers. Standard errors are clustered by building blocks. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table C3: First-stage Results of the IV Estimation at Building Level

	(1) Distance	(2) Distance	(3) Near	(4) Near
ECP	0.5058***	0.4300***	-0.3288***	-0.2602**
	(0.0999)	(0.0971)	(0.1100)	(0.1029)
LCP		0.3015**		-0.2732**
		(0.1236)		(0.1261)
Building Age	0.0014	-0.0070	-0.0126	-0.0049
	(0.0257)	(0.0260)	(0.0339)	(0.0346)
Slab Block	-0.2475	-0.3409*	0.0506	0.1352
	(0.1667)	(0.2010)	(0.2313)	(0.2511)
RC Center	-0.0804	-0.0620	0.1647	0.1480
	(0.1833)	(0.2015)	(0.1987)	(0.2116)
First-stage F-stats	25.62	17.66	8.94	6.04
Observations	47	47	47	47

Notes: Columns (1) and (2) report results for the continuous distance variable. Columns (3) and (4) report results for the binary closeness indicator. Column (1) and (3) use the ECP, the Euclidean straight lines connecting bus stops, as the instrument. Column (2) and (4) use both ECP and LCP, the least travel time path, as the instrument. Unreported control variables include the distance to train stations, LRT viaduct line, bus stops, expressways, major roads, and shopping centers. Standard errors are clustered by building blocks. Robust standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table C4: Effect of Launching Public Bus Service on Housing Price

	(1) log (price)	(2) log (price)
Launch*Distance	-0.0134***	
	(0.0050)	
Distance	0.0104***	
	(0.0037)	
Launch*Near		0.0129**
		(0.0054)
Near		-0.0097**
		(0.0038)
Launch	0.0299***	0.0101
	(0.0094)	(0.0077)
Prime Lending Rate	-2.8815***	-2.8766***
	(0.3563)	(0.3558)
Floor	0.0070***	0.0070***
	(0.0003)	(0.0003)
Area	0.0030***	0.0030***
	(0.0008)	(0.0008)
Building Age	-0.0009	-0.0009
	(0.0006)	(0.0006)
Distance to Train Station	-0.0036**	-0.0035**
	(0.0015)	(0.0015)
Distance to Bus Stop	-0.0046	-0.0042
	(0.0032)	(0.0031)
Distance to Expressway	0.0024***	0.0025***
	(0.0008)	(0.0008)
Distance to Major Road	-0.0010	-0.0014
	(0.0017)	(0.0016)
Distance to Shopping Center	0.0037***	0.0037***
	(0.0009)	(0.0009)
Block & Time Fixed Effect	Y	Y
Observations	1,450	1,450
R-squared	0.941	0.941

Notes: Column (1) reports results for the continuous distance variable and column (2) reports results for the binary closeness indicator. Unreported control variables include unit layout, building morphology and RC Center. Standard errors clustered by building blocks. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

References

Abbasi, A., Hassan, A., and Dhar, M. (2014). Benchmarking twitter sentiment analysis tools. In *Proceedings of the 9th Language Resources and Evaluation Conference*, volume 14, pages 26–31.

Alam, S. M., Eang, L. S., Tan, A., and Tiong, T. S. (2010). An investigation of community noise in high-rise residential environments. *The Journal of the Acoustical Society of America*, 127(6):3511–3518.

Anderson, M. L. (2014). Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review*, 104(9):2763–96.

Baccianella, S., Esuli, A., and Sebastiani, F. (2010). Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of the 7th Language Resources and Evaluation Conference*, volume 10, pages 2200–2204.

Banerjee, A., Duflo, E., and Qian, N. (2012). On the road: Access to transportation infrastructure and economic growth in china. Technical report, National Bureau of Economic Research.

Barclay, M., Kang, J., and Sharples, S. (2012). Combining noise mapping and ventilation performance for non-domestic buildings in an urban area. *Building and Environment*, 52:68–76.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics*, 122(2):775–805.

Baum-Snow, N. and Kahn, M. E. (2000). The effects of new public projects to expand urban rail transit. *Journal of Public Economics*, 77(2):241–263.

Bilger, M. and Carrieri, V. (2013). Health in the cities: When the neighborhood matters more than income. *Journal of Health Economics*, 32(1):1–11.

Bin, C. K., Lee, C. C. K., Teh, S. K., Chan, H. Z., and Sim, M. T. (2019). Land traffic noise management at recipient's end in singapore. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 259, pages 1465–1473. Institute of Noise Control Engineering.

Boyle, M. and Kiel, K. (2001). A survey of house price hedonic studies of the impact of environmental externalities. *Journal of Real Estate Literature*, 9(2):117–144.

Brown, A. and Lam, K. (1987). Urban noise surveys. *Applied Acoustics*, 20(1):23–39.

Cambria, E., Schuller, B., Xia, Y., and Havasi, C. (2013). New avenues in opinion mining and sentiment analysis. *IEEE Intelligent Systems*, 28(2):15–21.

Chalak, A., Al-Naghi, H., Irani, A., and Abou-Zeid, M. (2016). Commuters' behavior towards upgraded bus services in Greater Beirut: Implications for greenhouse gas emissions, social welfare and transport policy. *Transportation Research Part A: Policy and Practice*, 88:265–285.

Chasco, C. and Le Gallo, J. (2013). The impact of objective and subjective measures of air quality and noise on house prices: a multilevel approach for downtown madrid. *Economic Geography*, 89(2):127–148.

Chasco, C. and Le Gallo, J. (2015). Heterogeneity in perceptions of noise and air pollution: a spatial quantile approach on the city of madrid. *Spatial Economic Analysis*, 10(3):317–343.

Chew, C. (1991). Prediction of traffic noise from expressways—part ii: Buildings flanking both sides of expressway. *Applied Acoustics*, 32(1):61–72.

Chui, H., Heng, R. B., and Ng, K. (2004). Study of traffic noise levels in singapore. *Proceedings of Acoustics 2004*, pages 513–517.

Cropper, M. L. and Gordon, P. L. (1991). Wasteful commuting: a re-examination. *Journal of Urban Economics*, 29(1):2–13.

De Borger, B. and Proost, S. (2013). Traffic externalities in cities: the economics of speed bumps, low emission zones and city bypasses. *Journal of Urban Economics*, 76:53–70.

De Vos, J., Schwanen, T., Van Acker, V., and Witlox, F. (2013). Travel and subjective well-being: A focus on findings, methods and future research needs. *Transport Reviews*, 33(4):421–442.

DEFRA (2013). Noise pollution: economic analysis. *Department for Environment, Food & Rural Affairs, UK Government* `https://www.gov.uk/guidance/noise-pollution-economic-analysis`. Accessed: 2018-08-16.

DEFRA (2019). Noise action plan: Roads. *Department for Environment, Food & Rural Affairs, UK Government* `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/813666/noise-action-plan-2019-roads.pdf`. Accessed: 2019-12-04.

Diao, M., Leonard, D., and Sing, T. F. (2017). Spatial-difference-in-differences models for impact of new mass rapid transit line on private housing values. *Regional Science and Urban Economics*, 67:64–77.

Diao, M., Qin, Y., and Sing, T. F. (2016). Negative externalities of rail noise and housing values: Evidence from the cessation of railway operations in singapore. *Real Estate Economics*, 44(4):878–917.

Donaldson, D. (2018). Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934.

Duranton, G. and Turner, M. A. (2011). The fundamental law of road congestion: Evidence from us cities. *American Economic Review*, 101(6):2616–52.

Dzhambov, A. and Dimitrova, D. (2014). Neighborhood noise pollution as a determinant of displaced aggression: A pilot study. *Noise and Health*, 16(69):95–101.

Esuli, A. and Sebastiani, F. (2007). Sentiwordnet: a high-coverage lexical resource for opinion mining. *Evaluation*, 17:1–26.

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from china's national trunk highway system. *Review of Economic Studies*, 81(3):1046–1070.

Gao, J., He, X., and Deng, L. (2015). Deep learning for web search and natural language processing. In *Proceedings of the 8th ACM International Conference on Web Search and Data Mining*. WSDM 2015.

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, 68(3):1267–1300.

Gilbert, C. H. E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International Conference on Weblogs and Social Media*.

Goetzke, F. (2008). Network effects in public transit use: evidence from a spatially autoregressive mode choice model for new york. *Urban Studies*, 45(2):407–417.

Hamed, A.-R., Qiu, R., and Li, D. (2016). The importance of neutral class in sentiment analysis of arabic tweets. *International Journal of Computer Science and Information Technology*, 8:17–31.

Higgins, C. D., Adams, M. D., Réquia, W. J., and Mohamed, M. (2019). Accessibility, air pollution, and congestion: Capturing spatial trade-offs from agglomeration in the property market. *Land Use Policy*, 84:177–191.

Higgins, C. D., Sweet, M. N., and Kanaroglou, P. S. (2018). All minutes are not equal: travel time and the effects of congestion on commute satisfaction in Canadian cities. *Transportation*, 45(5):1249–1268.

Holmgren, J. (2014). A strategy for increased public transport usage–the effects of implementing a welfare maximizing policy. *Research in Transportation Economics*, 48:221–226.

Hsu, W.-T. and Zhang, H. (2014). The fundamental law of highway congestion revisited: Evidence from national expressways in Japan. *Journal of Urban Economics*, 81:65–76.

Hu, M. and Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 168–177. ACM.

Jakovljevic, B., Paunovic, K., and Belojevic, G. (2009). Road-traffic noise and factors influencing noise annoyance in an urban population. *Environment International*, 35(3):552–556.

Jamir, L., Nongkynrih, B., and Gupta, S. K. (2014). Community noise pollution in urban india: Need for public health action. *Indian Journal of Community Medicine*, 39(1):8.

Jedwab, R., Kerby, E., and Moradi, A. (2017). History, path dependence and development: Evidence from colonial railways, settlers and cities in kenya. *The Economic Journal*, 127(603):1467–1494.

Kim, H. D., Ganesan, K., Sondhi, P., and Zhai, C. (2011). Comprehensive review of opinion summarization.

Koppel, M. and Schler, J. (2006). The importance of neutral examples for learning sentiment. *Computational Intelligence*, 22(2):100–109.

Lam, K. C., Ma, W., Chan, P. K., Hui, W. C., Chung, K. L., Chung, Y.-t. T., Wong, C. Y., and Lin, H. (2013). Relationship between road traffic noisescape and urban form in hong kong. *Environmental Monitoring and Assessment*, 185(12):9683–9695.

Lam, Y. T. (2016). Economic sentiments in singapore. *Economic Survey of Singapore*, 2.

LTA (2015). Road noise barrier trial extended and low-noise road surfacing to be tested. *Factsheet: Mitigating Road Noise*. Accessed: 2019-12-04.

Lucas, K. (2012). Transport and social exclusion: Where are we now? *Transport policy*, 20:105–113.

Mak, C., Leung, W., and Jiang, G. (2010). Measurement and prediction of road traffic noise at different building floor levels in hong kong. *Building Services Engineering Research and Technology*, 31(2):131–139.

Martincus, C. V., Carballo, J., and Cusolito, A. (2017). Roads, exports and employment: Evidence from a developing country. *Journal of Development Economics*, 125:21–39.

McMillen, D. P. and McDonald, J. (2004). Reaction of house prices to a new rapid transit line: Chicago's midway line, 1983–1999. *Real Estate Economics*, 32(3):463–486.

Michaels, G., Rauch, F., and Redding, S. J. (2012). Urbanization and structural transformation. *The Quarterly Journal of Economics*, 127(2):535–586.

Mircea, M., Kovacs, I., Stoian, I., Marichescu, A., and Tepes-Bobescu, A. (2008). Strategic mapping of the ambient noise produced by road traffic, accordingly to european regulations. In *2008 IEEE International Conference on Automation, Quality and Testing, Robotics*, volume 3, pages 321–326. IEEE.

Monchambert, G. and De Palma, A. (2014). Public transport reliability and commuter strategy. *Journal of Urban Economics*, 81:14–29.

Münzel, T., Schmidt, F. P., Steven, S., Herzog, J., Daiber, A., and Sørensen, M. (2018). Environmental noise and the cardiovascular system. *Journal of the American College of Cardiology*, 71(6):688–697.

Nega, T. H., Chihara, L., Smith, K., and Jayaraman, M. (2013). Traffic noise and inequality in the twin cities, minnesota. *Human and Ecological Risk Assessment: An International Journal*, 19(3):601–619.

Nieuwenhuijsen, M. J., Ristovska, G., and Dadvand, P. (2017). Who environmental noise guidelines for the european region: A systematic review on environmental noise and adverse birth outcomes. *International Journal of Environmental Research and Public Health*, 14(10):1252.

Ossokina, I. V. and Verweij, G. (2015). Urban traffic externalities: Quasi-experimental evidence from housing prices. *Regional Science and Urban Economics*, 55:1–13.

Pang, B. and Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2):1–135.

Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing*, pages 79–86. Association for Computational Linguistics.

Park, S. H., Lee, P. J., Yang, K. S., and Kim, K. W. (2016). Relationships between non-acoustic factors and subjective reactions to floor impact noise in apartment buildings. *The Journal of the Acoustical Society of America*, 139(3):1158–1167.

Pow, C.-P. (2009). Public intervention, private aspiration: Gated communities and the condominisation of housing landscapes in singapore. *Asia Pacific Viewpoint*, 50(2):215–227.

Rana, R. K., Chou, C. T., Kanhere, S. S., Bulusu, N., and Hu, W. (2010). Ear-phone: an end-to-end participatory urban noise mapping system. In *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*, pages 105–116. ACM.

Redding, S. J. and Turner, M. A. (2015). Transportation costs and the spatial organization of economic activity. In *Handbook of Regional and Urban Economics*, volume 5, pages 1339–1398. Elsevier.

Rinker, T. W. (2016). *sentimentr: Calculate Text Polarity Sentiment*, volume 3. University at Buffalo/SUNY, Buffalo, New York, 0.5 edition.

Searle, G. and Filion, P. (2011). Planning context and urban intensification outcomes: Sydney versus toronto. *Urban Studies*, 48(7):1419–1438.

Segura-Garcia, J., Felici-Castell, S., Perez-Solano, J. J., Cobos, M., and Navarro, J. M. (2014).

Low-cost alternatives for urban noise nuisance monitoring using wireless sensor networks. *IEEE Sensors Journal*, 15(2):836–844.

Sørensen, M., Andersen, Z. J., Nordsborg, R. B., Becker, T., Tjønneland, A., Overvad, K., and Raaschou-Nielsen, O. (2013). Long-term exposure to road traffic noise and incident diabetes: a cohort study. *Environmental Health Perspectives*, 121(2):217.

Tamura, K., Elbel, B., Chaix, B., Regan, S., Al-Ajlouni, Y., Athens, J., Meline, J., and Duncan, D. (2017). Residential and GPS-defined activity space neighborhood noise complaints, body mass index and blood pressure among low-income housing residents in New York City. *Journal of Community Health*, 42(5):974–982.

Taylor, J., Eastwick, C., Lawrence, C., and Wilson, R. (2013). Noise levels and noise perception from small and micro wind turbines. *Renewable Energy*, 55:120–127.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., and Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558.

Turney, P. D. (2002). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 417–424. Association for Computational Linguistics.

Walerian, E., Janczur, R., and Czechowicz, M. (2001). Sound level forecasting for city-centers. part 1: sound level due to a road within an urban canyon. *Applied Acoustics*, 62(4):359–380.

Wan, W. (2016). Of noisy neighbours and silent nights. *The Straits Times* `https://www.straitstimes.com/opinion/of-noisy-neighbours-and-silent-nights`. Accessed: 2018-08-16.

Weinhold, D. (2013). The happiness-reducing costs of noise pollution. *Journal of Regional Science*, 53(2):292–303.

WHO (2016). Noise. `http://www.who.int/sustainable-development/transport/health-risks/noise/en/`. Accessed: 2018-08-16.

Xie, C., Gong, H., and Wang, F. (2010). A solution for the last mile problem of the Beijing rapid transit network: Local shuttle bus system. In *2010 18th International Conference on Geoinformatics*, pages 1–6. IEEE.

Xu, Y., Zhang, Q., and Zheng, S. (2015). The rising demand for subway after private driving restriction: Evidence from beijing's housing market. *Regional Science and Urban Economics*, 54:28–37.

Zheng, S., Sun, W., and Kahn, M. E. (2016). Investor confidence as a determinant of china's urban housing market dynamics. *Real Estate Economics*, 44(4):814–845.