buildings

Article

Disentangling the Modifiable Areal Unit Problem in Housing Density and Price Associations

Ka-Shing Cheung [1,*], Chiu-Wing Sham [2] and Chung-Yim Yiu [1]

[1] University of Auckland Business School, The University of Auckland, Auckland 1142, New Zealand; edward.yiu@auckland.ac.nz
[2] School of Computer Science, The University of Auckland, Auckland 1142, New Zealand; b.sham@auckland.ac.nz
* Correspondence: william.cheung@auckland.ac.nz

Abstract: Urban planning education must address the Modifiable Areal Unit Problem (MAUP) to comprehend the critical impact of urban density on sustainable city development. Quantitative studies using administrative area units face indefinite aggregate level biases. This paper introduces an efficient block-searching method to calculate property densities around residences of various boundary scales and empirically examines their relationship with housing prices in Auckland, New Zealand. Results reveal negative associations between housing prices and densities within neighbourhoods, emphasising the limitations of administrative boundaries. These findings underscore the necessity for planning education to navigate MAUP's complexities in shaping urban development policies.

Keywords: urban planning; housing density; public housing; Airbnb; neighbourhood; hedonic price model; Modifiable Areal Unit Problem (MAUP)

check for updates

Citation: Cheung, K.-S.; Sham, C.-W.; Yiu, C.-Y. Disentangling the Modifiable Areal Unit Problem in Housing Density and Price Associations. *Buildings* **2024**, *14*, 1840. https://doi.org/10.3390/buildings14061840

Academic Editor: Chyi Lin Lee

Received: 10 May 2024
Revised: 11 June 2024
Accepted: 13 June 2024
Published: 17 June 2024

1. Introduction

Urban economics explores the interconnection between the layout of cities, economic dynamics, and personal preferences within an urban setting. Central to this field is the concept of agglomeration effects. Individuals agglomerate to denser urban environments for better access to amenities, job opportunities, and cultural experiences. This choice is motivated by the desire to minimise transportation costs and to enhance productivity, innovation, and knowledge exchange. The trend towards such urban agglomeration is fuelled by the strong demand for land in central locations, driven by the convenience and advantages of proximity to essential services, workplaces, and social networks [1]. On a global scale, many countries are implementing policies to encourage the development of densely populated, compact urban areas. The compact city model, characterised by high urban density and mixed land use, has become a key strategy in modern urban planning for sustainable city development [2]. By promoting greater urban density, cities can achieve several benefits, including improved public transportation use, shorter commute times for residents, and the creation of vibrant mixed-use neighbourhoods that enhance overall liveability.

Since then, the scholarly literature has begun to illuminate the complex relationship between urban density and its various outcomes. Although adjustments in urban density can lead to positive effects in some areas, they may also result in negative impacts elsewhere. Ahlfeldt and Pietrostefani [3] conducted a comprehensive review identifying 15 different urban outcome categories linked to density effects and quantified each in monetary terms. Their research indicates that although significant benefits and costs are associated with increases in urban density, the benefits tend to surpass the costs in most large, developed cities. In a similar vein, Melia et al. [4] highlighted the trade-off between these urban outcomes, referring to it as the 'paradox of intensification' in the context of transport policy. This paradox suggests that while higher population densities reduce per capita car use and benefit the built environment, they can also lead to higher levels of motor traffic, which may adversely affect the local area. Furthermore, the paradox of intensification associated with urban density often arises from a too simplistic understanding of this complex concept. Although density is a useful measure for planning compact cities, its simplicity may mask the intricacies of urban form [5]. Therefore, it is crucial to dissect the factors contributing to the effects of density in compact urban environments [3].

Yet, the task of analysing and contrasting the effects of varying urban densities becomes much more complicated when employing spatially aggregated data. Compilations of data at disparate spatial scale levels or within different zoning systems for identical regions often fail to provide consistent analytical outcomes. This challenge is known as the Modifiable Areal Unit Problem (MAUP), initially highlighted by [6] and subsequently elaborated by Openshaw and Taylor [7]. In essence, the MAUP represents an ecological fallacy in spatial analysis, occurring when data, which are originally specific to particular locales, are aggregated into broader spatial categories such as regions or districts. This aggregation process, relevant to various metrics, including population density or illness rates, substantially affects the derived measures—such as totals, rates, proportions, and densities—depending on the size and shape of the spatial units selected for analysis. The continued practice of inputting census data (from census tracts) into regression models for urban policy development highlights the persistence of MAUP as a significant concern [8]. This is because the use of these areal units at various scales of aggregation has been demonstrated to lead to aggregation biases [9,10], which can result in inaccurate conclusions and policies [11].

The Modifiable Areal Unit Problem (MAUP) in the Housing Market

Imagine a city being analysed to understand the impact of ethnic density on housing prices, aiming to identify areas with high and low prevalences of ethnic minority groups. The city is divided into various spatial units for this study, such as districts or neighbourhoods. In one scenario, the city is divided into broad districts, while in another, it is segmented into finer neighbourhoods. Analysing larger districts might reveal a seemingly uniform impact of ethnic density on housing prices, potentially masking true variations within those districts. For example, within a single district, an affluent area may have low ethnic density, while an adjoining less affluent area may exhibit higher ethnic density. Averaging the data across the entire district could inaccurately suggest a uniformly moderate influence of ethnic density on housing prices.

For example, the visual below (Figure 1) clearly demonstrates how the Modifiable Areal Unit Problem (MAUP) can impact the interpretation of housing economic data, paralleling the discussion on ethnic density. In the upper diagrams, "Region Set A" and "Region Set B" depict residential data points within varying regional boundaries. Each square symbolises a residence, with the number inside indicating the ethnic density (1 for high ethnic density, 0 for low ethnic density), signifying areas with higher or lower ethnic diversity. In "Region Set A", residences are encompassed within a single boundary, leading to an aggregated indication of a 50% level of ethnic density—suggesting an even distribution of high and low ethnic density across the residences. In contrast, "Region Set B" divides the same residences across two distinct boundaries. Post-aggregation, one region displays a 0% level (indicating low ethnic density across all homes), while the other records a 100% level (indicating high ethnic density at every residence). This outcome variance arises not from genuine differences in ethnic density but from how the regional boundaries are defined. As a result, this arbitrary division can significantly distort the perceived distribution of ethnic diversity across the city, potentially leading to inaccurate analyses of how ethnic density influences housing economics, thereby, urban planning and policy inappropriately.

This variability can skew interpretations of variable relationships, impacting decision-making in urban planning, public health, and environmental management. MAUP remains a critical and intricate issue within spatial analysis, bearing considerable consequences for understanding spatial relationships and deriving valid conclusions from geographical data [12]. In urban planning, neighbourhoods are frequently delineated by census block groups such as census tracts in the US or statistical areas in Australia and New Zealand. These units of geographical location, varying in spatial scale, can introduce MAUP into regression analyses.

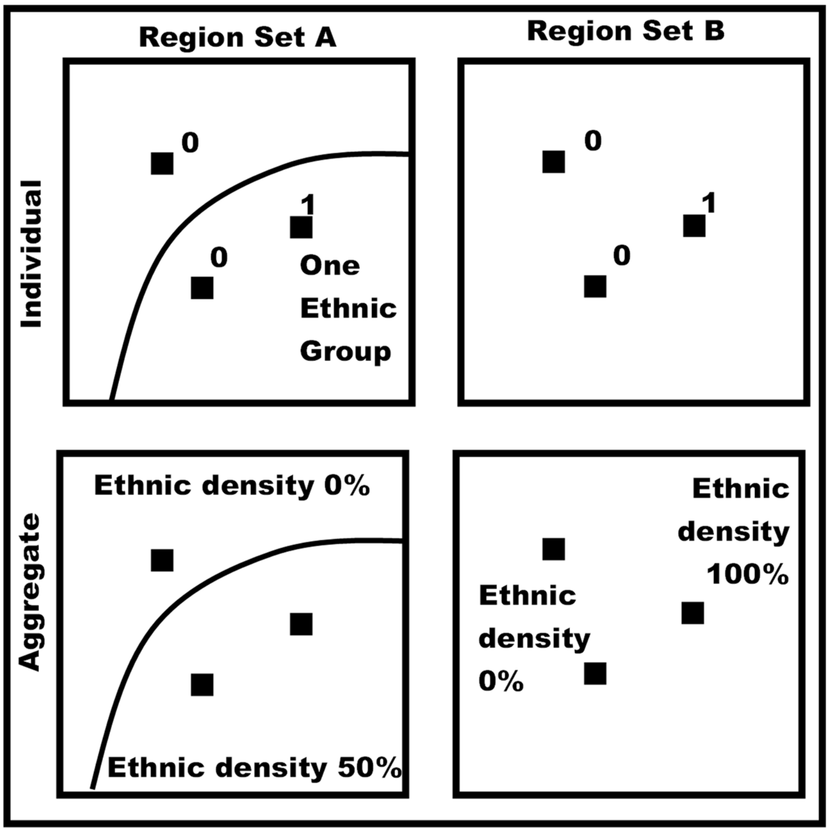


Figure 1. An example of the modifiable areal unit problem and the distortion of ethnic density.

One approach to mitigate the Modifiable Areal Unit Problem (MAUP) involves utilising smaller and more consistent areal units. Consistency in an estimator means it converges in probability to the true value of the parameter as the sample size increases indefinitely. Sensitivity analysis is another effective method, assessing how outcomes fluctuate with alterations in the spatial unit of analysis and offering insights into the robustness of the results. Additionally, adopting alternative spatial divisions that align more accurately with the natural or social boundaries of the data, as opposed to arbitrary administrative divisions, can provide a truer representation of spatial phenomena. Su et al. [13] proposed the creation of homogeneous boundaries within each district, such as block-block group, while Kwan [14] advocated for identifying the optimal scales to address MAUP. Parenteau and Sawada [15] and Chen et al. [16] employed sensitivity analysis to counteract MAUP's effects. Consequently, an effective analytical approach requires the exploration of various consistent boundaries with sensitivity tests rather than depending on predetermined third-party boundaries for data aggregation.

This study introduces a block-searching model that aims to create customised, consistent areal units to minimise the Modifiable Areal Unit Problem (MAUP) impact. To illustrate its application, it employs an empirical analysis of neighbourhood effects on property densities and housing prices. The research provides a detailed exploration of the relationship between housing density and property prices, factoring in internal and external densities and the specific characteristics of different housing types within neighbourhoods. Internal housing density refers to site-specific measures such as floor area ratio and the utilisation of floor space within individual properties. On the other hand, external housing density encompasses the concentration of private residences in the vicinity of a property and the prevalence of non-residential property types, such as industrial facilities, public housing, or Airbnb listings within neighbourhoods. The objective is to shed light on how various aspects of density affect the quality of urban neighbourhoods. This includes assessing the detrimental impacts of industrial property density, the effects of high residential density, and the presence of public housing or Airbnb units. The study posits that these density-related factors can significantly influence the attractiveness and liveability of urban areas, potentially leading to either price premiums or discounts on housing, all other factors being constant.

To the best of our knowledge, the existing literature in urban studies lacks a holistic analysis and comparison of different measures of urban density within a single context, limiting our understanding of their effects and the spectrum of positive and negative outcomes they may yield. The impacts of housing density can vary, encompassing both positive and negative aspects [5]. Higher urban densities offer sustainability advantages, such as efficient land use, social equity, and diversity. Conversely, drawbacks include social and psychological stresses resulting from limited space, safety concerns (crime), and inadequate shared amenities, traditionally considered private goods. Consequently, numerous studies have found a negative correlation between higher densities and housing prices, indicating a preference for lower-density neighbourhoods among home buyers [17].

The study will be structured as follows: In Section 2, we review the literature on density effects and neighbourhood effects. In Section 3, we discuss the research design and the block-searching method, and in Section 4, we present and discuss the empirical results. In Section 5, we summarise the findings and implications of the study. In the Appendix A, we report three robustness checks on our density models.

2. Literature Review on Density Effects

The influence of density on health, particularly mental health, has been extensively explored in prior research. Spacious internal and external environments have been identified as beneficial for health and well-being, especially for families with children [18–20]. Numerous studies have investigated the effects of internal density on housing, focusing on factors such as interior space and the number of residents per dwelling unit [18,21,22]. However, there is a noticeable gap in research on the external effects of housing density [23]. Gomez-Jacinto and Hombrados-Mendieta [24] argued that external density, also known as "community density", has a more significant impact on human stress compared to internal density. Urban planning acknowledges the significance of externalities and strives to regulate development intensity by considering the carrying capacity of urban areas [20]. Furthermore, Bramley and Power [25] identified various negative consequences of high urban density on quality of life, including reduced community cohesion, inadequate urban amenities, and a lack of open spaces. Moreover, previous studies examining urban land density analysis have lacked a solid foundation and tend to be arbitrary [26]. Research on residents' perceptions of housing density in their neighbourhoods has mainly relied on property agents or qualitative methods such as surveys and interviews [27–29].

Conversely, a significant body of literature in urban economics has highlighted the benefits of urban density, including reduced commuting times [30] and increased access to amenities [31]. The positive association between higher housing density and property values in a neighbourhood has been extensively researched and documented [3,32]. Yet, there are also costs associated with urban density, such as crowding and crime. The negative relationship between density and property values has been observed in some studies, particularly in high-density cities [33,34]. More recently, Duranton and Puga [1] provided a comprehensive literature review on both the costs and benefits of urban density. Samsudin et al. [35] reviewed the studies on both the positive and negative impacts of high-density environment on social capital. Wang et al. [36] examine the spatial variation characteristics of housing conditions in China, using 2846 counties as the basic research unit. Their study highlights significant spatial clustering of housing conditions, with better conditions southeast of the "Hu Line" and worse conditions northwest. Key findings include the importance of elevator configuration and the superior housing conditions in urban areas compared to non-urban areas. The study also reveals that urbanisation significantly impacts housing conditions, particularly in urban settings.

In recent discussions, there has been growing recognition of the potential associations of public housing density and Airbnb in neighbourhoods, particularly concerning inclusivity. Thackway et al. [37] investigate the spatial variability of Airbnb's impact on housing prices in Sydney. Using a hedonic property valuation model and geographically weighted regression (GWR), the study finds that Airbnb generally increases property prices, with a 1% rise in Airbnb density leading to a 2% increase in property sales prices. However, the effect is geographically uneven, with significant value uplifts in Sydney's northern beaches and parts of western Sydney, while traditional tourist areas around the CBD and eastern suburbs see insignificant or negative impacts. This research highlights the need for tailored Airbnb regulations based on local housing market conditions. However, empirical studies on this subject are notably scarce, and existing studies primarily rely on stated preference approaches, which may not provide robust large-scale evidence. This paper aims to fill this research gap by empirically examining the associations of internal, external and property type-specific housing densities on housing prices using a revealed preference approach. By analysing actual transacted housing prices across neighbourhoods with varying levels of housing density, public housing density, and Airbnb density, etc., this study seeks to identify and quantify these associations.

In fact, analysing the associations of density with housing prices, both over time and across different areas, can yield conflicting findings. Temporal analyses often found a positive price effect of upzoning [19,38,39]. However, Murray and Limb [40] found no relationship of zoned capacity with housing prices in their studies of upzoned densification areas in Brisbane, Australia. Yet they considered median home price data of the administrative area units (Statistical Areas SA2) instead of the upzoned areas. The discrepancy between the boundaries of the SA2 and the upzoned areas can lead to questionable results. It is therefore crucial to identify variables, such as house density and median housing price, from the same community boundary.

Furthermore, these studies primarily examine the temporal increase in development intensity and fail to consider the negative externality of higher neighbourhood densities in a cross-sectional context. Regarding cross-sectional comparisons of housing prices across neighbourhoods, varying housing densities have generally yielded negative results [17]. Nevertheless, it is worth noting that most cross-sectional studies on density effects rely on administrative area units, such as state population densities [41], localised densities of individual development sites [33], or densities of census tracts [17]. These administrative units vary in shape and size but have similar populations or numbers of housing units, making it difficult to draw meaningful comparisons of density effects on a like-for-like basis. For instance, Fesselmeyer, Seah and Kwok [33] compared density effects from development projects ranging from 17.86 to 645.08 units/acre on land areas spanning 0.04 to 18.52 acres. Despite having the same average density of 84.75 units/acre, the impact on housing prices can differ significantly between two plots of land that vary in size by a factor of 400. Some studies have employed a fixed-area method to analyse density effects, but with certain limitations. For example, Sequeira and Filippova [42] measured concentrations of social housing within a fixed 500 m radius but did not include the external housing density variable. They categorised densities into three dummy variables (High-Medium-Low) and incorporated other neighbourhood attributes, such as a deprivation index based on administrative area units. Essentially, they used two different definitions of neighbourhoods within the same model.

Dong [43] provided insights into the development and characteristics of multifamily homes in the Portland metropolitan area and evaluated the influence of density and density-related factors on the pricing of such homes. The findings of this study revealed that medium-density multifamily homes tend to have lower selling prices than other housing types. Surprisingly, the cost-saving effect typically associated with higher-density development appears to be weak or even negative in the case of multifamily homes. Instead, dwelling size emerges as the primary determinant of prices in this housing category. Similarly, Cheung and Yiu [44] measured the number of houses of different cohorts within a fixed-radius distance from homes but conducted a panel data analysis involving spatial and temporal comparisons. This approach may have mixed the effects of cross-sectional negative externalities and longitudinal positive upzoning effects, leading to a complex interpretation of density effects.

In a similar vein, the positive and negative effects of public housing on housing prices, which indicate positive and negative externalities, have been studied piecemeal. Positive externalities such as accommodation subsidies are typically applied at the administrative area unit level [45]. On the other hand, negative externalities such as the stigma of public housing [46] or higher crime rates [47,48] can be related to the density of public housing in proximity. In some situations, measuring these neighbourhood impacts by means of a buffer with a certain radius is more relevant. Moreover, the emergence of Airbnb has also been shown to impact housing prices. While there can be negative externalities for residents in the neighbourhoods [49,50], previous studies have generally found a positive effect of Airbnb on housing prices due to increased rental income [24] and reduced housing supply for local residents [51–53] based on administrative area units.

The choice between an administrative area unit or a radial buffer matters in the density effect research. The analysis depends on the specific attributes being studied. While administrative area units may be appropriate for studying the price effect of school zones, they may not accurately represent the neighbourhood effect of housing density due to variations in shape and size. This can introduce boundary bias and misrepresent the attributes of households near the administrative area unit boundaries. Demographic variables such as average household income, crime rates, and job opportunities are unlikely to align with administrative area units. Government statistics departments typically provide aggregate demographic data in administrative area units. However, accessing raw data from government data laboratories and efficiently processing density-related studies can be challenging. Counting metrics from raw data can be a time-consuming process, mainly when dealing with large numbers of entities. Moreover, while calculating housing density for each house may seem straightforward, it requires substantial computational resources, especially in cities with a large number of houses. Bangura and Lee [54] highlight the complexity of defining housing submarkets by demonstrating that the determinants of homeownership affordability vary significantly across different regions of Greater Sydney. This variability underscores the importance of considering submarket differences when analysing housing affordability and density effects, aligning with the need to address the Modifiable Areal Unit Problem (MAUP) in housing studies.

So, the fundamental yet significant research inquiry: Can we formulate a comprehensive framework within the urban housing market that enables planners to disentangle various measures of urban density, thereby aiding in the identification and attainment of a balance between the positive and negative consequences of higher density development in practical contexts? Based on the literature review concerning the effects of density, we have reformulated the framework of various density impacts, the details of which are provided in Table 1. This reformulated framework also guides our statistical analysis in the empirical tests that follow.

Table 1 compares three different research approaches investigating the impact of density on housing prices, yielding conflicting empirical results. The first approach explores the internal density effects of development intensity within a land parcel or within a housing unit on housing prices. This approach has been commonly applied in architectural research. The second approach examines the price implications of temporal changes in density. For instance, cities undergoing upzoning or densification often experience positive effects on housing prices. However, these studies typically focus solely on the increase in development intensity, overlooking potential negative externalities associated with higher housing density. The third approach involves cross-sectional comparisons of various property-type densities across different districts or neighbourhoods to evaluate their price effects. This approach, exploring the external density effects, forms the basis of this study. So far, empirical studies investigating the impact of neighbourhood density on housing prices remain scarce, likely due to challenges in accurately estimating representative density metrics. Overall, these three research lines have contributed to our understanding of density effects on housing prices. However, the existing literature exhibits conflicting findings and knowledge gaps, necessitating further exploration to develop a comprehensive understanding of the complexities surrounding housing density and its implications for property values.

Table 1. Multifaceted associations of urban density on housing prices.

Internal Housing Density Effect	Temporal Changes in Housing Density	External Density Effect—Various Property-Type Densities in a Neighbourhood
Higher development intensity within a land parcel may lead to a negative association with housing prices due to overcrowding and inadequate amenities.	Upzoning initiatives can result in a positive association with housing prices by potentially increasing the sellable floor area.	In neighbourhoods, higher housing density can have a negative impact on prices due to increased housing supply, while elevated public housing density can also contribute to price depreciation.
Increased floor space occupancy within individual housing units may lead to decreased housing prices due to overcrowding concerns.		The densities of non-residential property types can exert either positive or negative effects on prices, depending on the complementary or conflicting relationships between these property types and residential housing.

3. Research Design

Many urban studies necessitate counting the number of properties and/or amenities within a neighbourhood, such as measuring residential density [55,56] or analysing the number of points of interest [57]. Identifying nearest neighbours typically involves estimating their distances from the subject point of interest (POI). This computational challenge is commonly addressed using a range-query-oriented spatial index like KD-Tree, R-Tree, or Ball-Tree, which may be efficient for full-distance matrix computation [58]. However, a predetermined boundary is often set in urban planning issues such as housing density estimation, a radial buffer or a census tract/statistical area unit are examples of the boundary used. In other words, the maximum number of properties in a cluster is known and constant. Our block-searching method is a bespoke algorithm designed for a known-size dataset to enhance estimation efficiency.

3.1. Block Searching Method (BSM)

We have devised a block-searching technique to expedite the counting process. The proposed method significantly reduces computational time and memory requirements, making it applicable to various counting processes for counting other neighbourhood points of interest. Identifying whether two houses are neighbours within a boundary necessitates estimating the distance between all houses from the subject house. For instance, this study encompasses approximately 400,000 houses in Auckland, New Zealand. Simply counting neighbouring houses requires computing the distance between two houses 400,000 $\times$ 400,000 times. Using conventional methods, this computation could take days to complete.

3.1.1. Divide the Study Area into Blocks

We define a neighbourhood by a fixed dimension of a buffer radius, such as 0.5 km or 1.0 km from a house, calculated using the Haversine formula. The BSM divides the map into several grid blocks based on longitude and latitude. For instance, if the dimension of a neighbourhood is defined as a d km radius from a house, then the study area is initially divided into grid blocks with dimensions of at least $d \times d$ (Figure 2a).

(**a**) (**b**)

Figure 2. (**a**) An illustration of dividing the study area into $d \times d$ blocks; (**b**) An illustration of searching the neighbouring houses in nine blocks. This example shows a 3 × 3 blocks and the numbering sequence from 0 (the subject property location) and the surrounding eight blocks.

3.1.2. Searching and Counting Method

The nine grid blocks around a house set the boundary for searching and counting neighbouring properties of each house, and properties beyond the nine grid blocks are not considered neighbouring properties. For example, as shown in Figure 2b, if a house located at grid block 0 needs to count the number of neighbouring houses, the search area can be limited to the nine grid blocks {0, 1, . . ., 7, 8} surrounding the subject house.

3.1.3. Data Structure to Save Memory

To conserve computational memory, we utilise a unique data structure called Property Node Indexing (PNI) [59]. Housing attributes, not utilised in the searching and counting processes, are stored in a "Property" node, as shown as Pi in Figure 3, where i = 1, 2, 3, . . . As the study area is divided into blocks, each house is designated with a grid block number. We store only the pointer of each "Property" node in a block to save memory. This data structure enables efficient calculation of distances between properties by sequentially going through the index of the "Block" node.

The "Block" node includes an index of the "Property" node data and the count of corresponding properties. This data structure enables us to calculate the distance between properties by sequentially accessing the index of the "Block" node. We store only pointers because the size of the "Property" node data is large. By containing only pointers, the index reduces memory usage, potentially increasing search speed. Additionally, no extra storage is necessary during the searching and counting processes.

Furthermore, this data structure of entity nodes enhances the versatility of the searching and counting program. It allows for the integration of datasets from various entity nodes for searching and counting within the same area. For instance, in this study, the "Property" node data structure is applied to all types of property in the Auckland Region, including housing units, public housing units, Airbnb, commercial property, and industrial property. This method can also be extended to search and count other community or institutional buildings, as well as special points of interest. Overall, this approach makes the program universally applicable for most spatial searching and counting purposes.

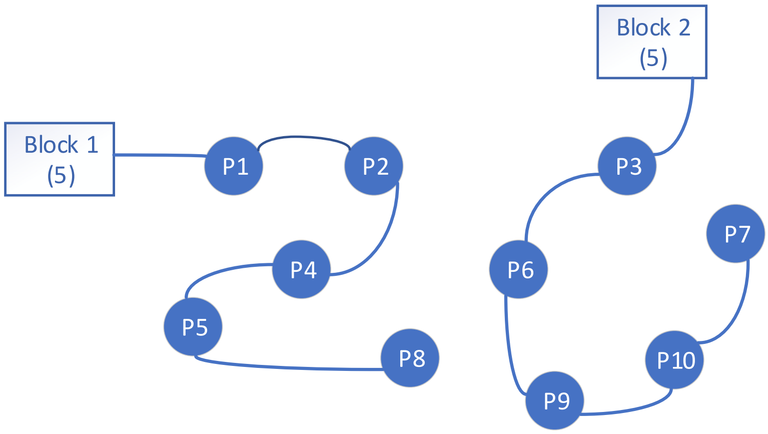


Figure 3. Data structure of the region node. Notes: as an illustration, the figure shows a situation where there are two block nodes, block 1 and block 2. Both have five neighbouring houses. The "Block" node connects the corresponding neighbouring houses to an index list without consuming too much in storage or memory resources.

3.1.4. Enhancement of Computational Efficiency

To demonstrate the computational efficiency enhancement of the BSM through block counting, we conducted experiments to evaluate the computational time required for various block sizes. The results as shown in Figure 4 indicate a significant reduction in execution time compared to traditional methods. For instance, when the buffer radius of neighbourhoods is increased from 0.5 km to 14.5 km, the computation time for searching and counting the number of houses and public houses is sharply reduced. The BSM enhances computational efficiency by 98%.

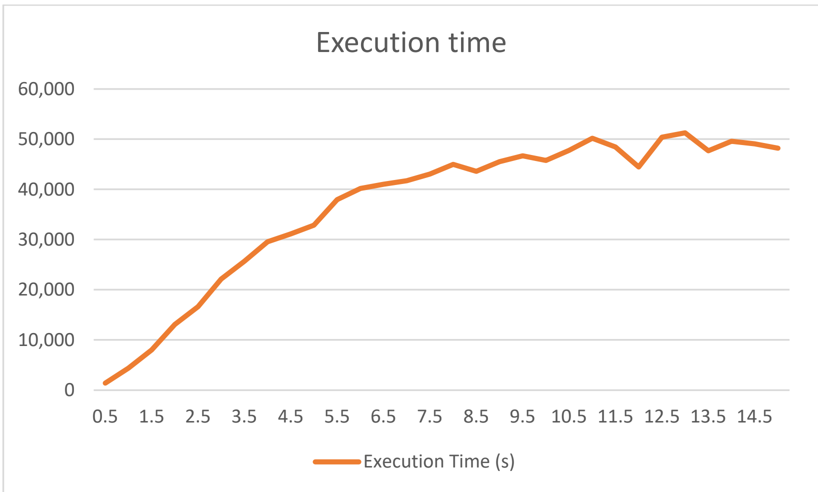


Figure 4. Execution time in seconds of searching and counting houses within a radial buffer of each house from 0.5 km to 14.5 km by the BSM. The vertical axis refers to the execution time in seconds, while the horizontal axis refers to the radial buffer size in km.

The figure depicts a linear increase in execution time from 0.5 km to 6.5 km, followed by a levelling off. This trend suggests that shorter search distances yield similar housing densities within neighbourhoods, whereas longer search distances encompass rural areas with lower housing densities.

3.2. Empirical Models

This study counts the number of housing units, public housing units, and Airbnb units to estimate both the external housing density and the property type-specific housing densities of various building types within three walking distances: a 0.5 km, 1.0 km, and 1.5 km radius. For comparison, data on private and public housing densities will also be collected based on the administratively defined boundary—Statistical Area 1 (SA1) unit (Public housing in New Zealand is generally provided by Housing New Zealand (Kāinga Ora) and Community Housing Providers (CHPs). The former is a state organisation operating under the Crown Entities Act 2004, the latter are NGOs. Kāinga Ora is the largest public housing provider providing 63,589 housing units, whereas CHPs provides 7730 units in June 2020 [60]. Auckland, being the highest populated city in New Zealand, possesses most of the public housing units provided).

3.2.1. Baseline Model

Equation (1) shows the baseline hedonic price model in the semi-log specification, including the structural attributes, location dummy attributes in suburbs, and time dummy variables, etc.

$$ln(P_{its}) = \alpha_0 + \sum_{k=1}^{K} \gamma_k X_{ki} + \sum_{v=1910}^{V} \beta_v C_{iv} + \sum_{t=1}^{T} \alpha_t D_{it} + \sum_{s=1}^{S} \theta_s L_{is} + \varepsilon_{ist} \ldots \quad (1)$$

where P_{its} denotes the transaction price of property i at time t in suburb s (i = 1, . . ., n; t = 1, . . ., T; s = 1, . . ., S), γ_k denotes the implicit price for the k^{th} property characteristic X_{jk} (k = 1, . . ., K); C_{iv} denotes the cohort dummy, which is set to 1 if the i^{th} house was built in decade v and to 0 otherwise (v = 1910s, 1920s, . . ., 2010s, with the 1900s cohort as the omitted base case); D_{it} denotes the month-of-sale dummy, which is set to 1 if the i^{th} house sold at time t, and otherwise to 0; L_{is} denotes the suburb location dummy, which is set to 1 if the i^{th} house sold is located in suburb s, and otherwise to 0; and ε_{ist} denotes the error term with the mean zero and the variance σ^2. The coefficients γ_k, β_v, α_t, and θ_s can be estimated by the ordinary least squares method.

3.2.2. Housing Density Model

Equation (2) presents the housing density hedonic price model in the semi-log specification estimated using the ordinary least squares (OLS) method, which accounts for the number of houses within a fixed walking distance from the transacted house. This study presents results for three different radial buffers within walking distance: a 0.5 km radius (Model 3), a 1.0 km radius (Model 4), and a 1.5 km radius (Model 5). The density considered here encompasses not only external housing density, determined by the number of housing units, but also property type-specific housing densities, including public housing density and Airbnb density. It achieves this by relating the proportion of public housing units or Airbnb units to the total number of housing units in the neighbourhood. Mathematically,

$$\begin{aligned} ln(P_{its}) = \alpha_0 + \alpha_r NH_{ir} + \delta_r^p \tfrac{nPUB_{ir}}{NH_{ir}} + \delta_r^a \tfrac{nAIR_{ir}}{NH_{ir}} + \sum_{k=1}^{K} \gamma_k X_{ki} + \sum_{v=1920}^{V} \beta_v C_{iv} \\ + \sum_{t=1}^{T} \alpha_t D_{it} + \sum_{s=1}^{S} \theta_s L_{is} + \varepsilon_{ist} \ldots \end{aligned} \quad (2)$$

where NH_{ir} denotes the total number of housing units in the r-radius buffer neighbourhood from the i^{th} house (Equation (2)) and α_r measures the housing density effect; δ_r^p and δ_r^a denote the implicit prices for the proportion of public housing and Airbnb to the total number of housing units in the r-radius buffer neighbourhood, r = 0.5, 1.0, 1.5, as defined in Equations (3) and (4), respectively:

$$NP\_PUB_{ir} = \frac{nPUB_{ir}}{NH_{ir}} \ldots \quad (3)$$

$$NP\_AIR_{ir} = \frac{nAIR_{ir}}{NH_{ir}} \ldots \quad (4)$$

where $nPUB_{ir}$ and $nAIR_{ir}$ denote the numbers of public housing and Airbnb units in the r-radius buffer neighbourhood from the i^{th} house, respectively.

4. Data and Empirical Results

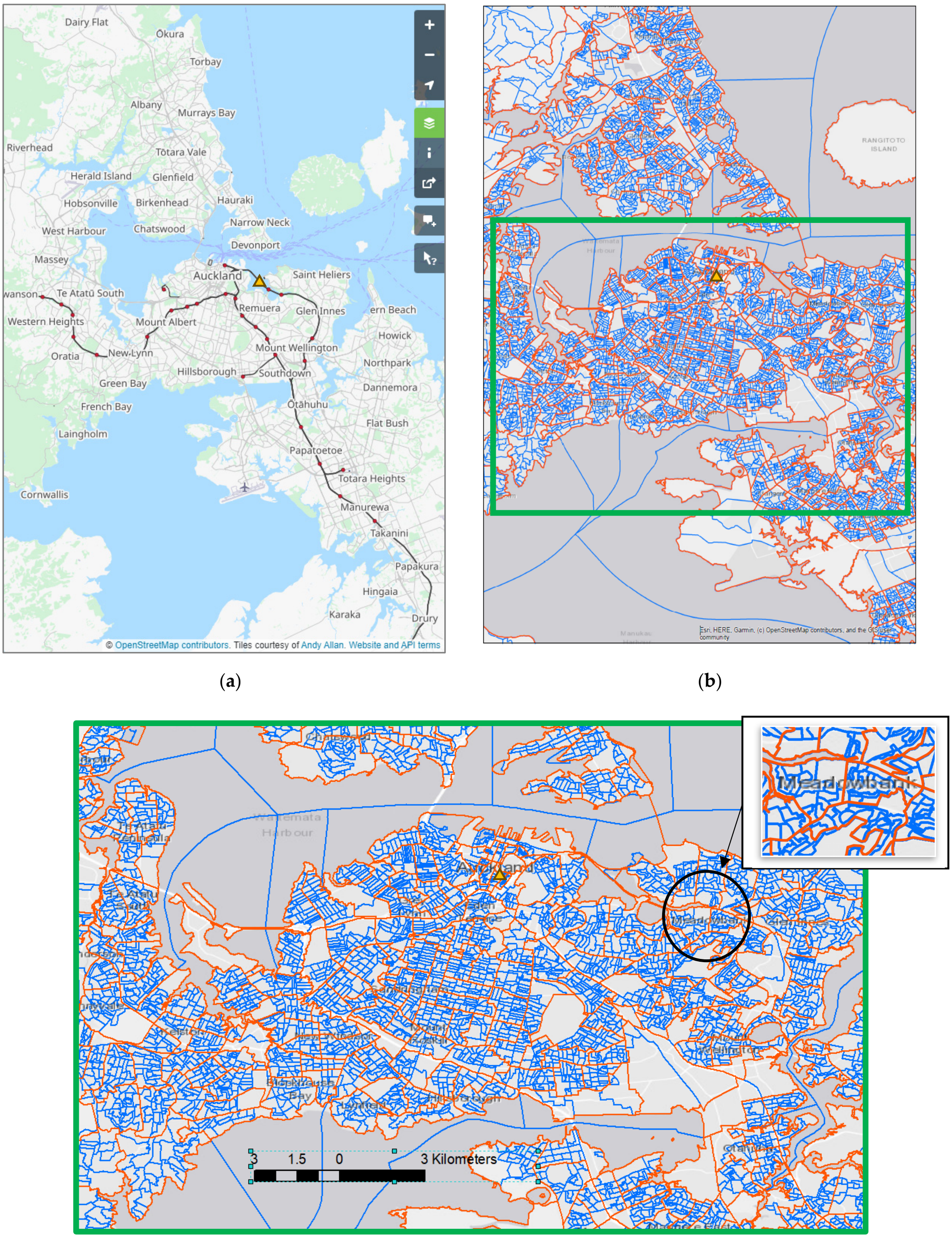
4.1. Data

The models are empirically tested using housing transaction data recorded in Auckland Central, New Zealand, spanning from January 2020 to December 2020 (12 months). Figure 5a shows an overview map of Auckland, New Zealand, highlighting major roads and locations. Figure 5b is a detailed map showing Statistical Area 1 (SA1) in blue and Statistical Area 2 (SA2) in orange. Figure 5c at the bottom is a zoomed-in map with a scale illustrating the layout of SA1 (blue) and SA2 (orange) within a 1-km radius. The map provides a detailed view of the statistical areas and their boundaries. As the COVID-19 pandemic commenced in 2020, the impacts of density on housing prices should be more pronounced, as the spread of disease depends on density. The dataset utilised in this study is sourced from CoreLogic. To maintain consistency in housing types, the hedonic price analysis excludes all non-house dwelling types, such as townhouses and apartments, from the transaction data. However, when counting the number of houses in the neighbourhood to test the density effect, all types of dwellings, including apartments and townhouses, are considered.

After excluding outliers, approximately twenty thousand valid records of housing transactions were identified for the period under study. Housing data were from Auckland Council's (2022) [61] Rating Information Database (RID). The RID records all property information in Auckland and is used for setting and assessing property rates. Additionally, the dataset provides a comprehensive list of variables concerning property and neighbourhood characteristics. Airbnb data in Auckland, New Zealand, were obtained from the AirDNA (2021) [62] subscription.

Table 2 shows the summary statistics of the housing and public housing densities of the neighbourhoods and other variables.

SA1 and SA2 refer to statistical area units of varying dimensions designated by the New Zealand Government to maintain a consistent population range within each unit. For instance, SA1s in New Zealand typically have a population of 100–200 residents, with a maximum of 500. SA2s in city council areas typically accommodate a population of 2000–4000 residents, whereas SA2s in district council areas usually encompass a population of 1000–3000 residents [63]. However, because this measurement does not account for the measured land area and may introduce boundary bias, it is generally not considered a reliable metric for assessing density associations with housing prices.



(**a**) (**b**)

(**c**)

Figure 5. (**a**) (top left) An overview map of Auckland, New Zealand, highlighting major roads and key locations. (**b**) (top right) A detailed map of Statistical Area 1 (SA1) in blue and Statistical Area 2 (SA2) in red, offering a closer look at the statistical boundaries and their distribution within Auckland. (**c**) (bottom) A zoomed-in map with a black radial circle illustrating how a radial circle covers the SA1 (blue) and SA2 (red) administrative boundaries. The yellow triangle indicates the CBD of the city.

Table 2. Descriptive Statistics of Variables of the Hedonic Price Model.

Variable	Description	Mean	SD	Min	Max
$ln(P_{its})$	Natural logarithm of the sale price of the i^{th} house at time t, suburb s	13.49	12.86	10.02	14.76
D_{it}	Month-of-sale dummy variable	12 months (January–December of 2020)			
L_{is}	Suburbs where the i^{th} house is located, 1736 suburbs of the Auckland Region				
	Demographic characteristics of neighbourhoods in various levels				
NH_{iSA1}	No. of housing units in the SA1	21.82	59.95	0.00	776.00
NP_PUB_{iSA1}	Proportion of the no. of public housing to total no. of housing units in the SA1	0.03	0.08	0.00	0.92
NH_{iSA2}	No. of housing units in the SA2	1021.99	425.21	12.00	2887
NP_PUB_{iSA2}	Proportion of the no. of public housing to total no. of housing units in the SA2	0.04	0.06	0.00	0.51
NH_{ir500}	No. of housing units in the 0.5 km radius	812.76	959.72	1.00	8211.00
NP_PUB_{ir500}	Proportion of the no. of public housing to total no. of housing units in the 0.5 km radius	0.04	0.06	0.00	0.62
NP_AIR_{ir500}	Proportion of the no. of Airbnb to total no. of housing units in the 500 m radius	0.001	0.007	0.00	1.00
NH_{ir1000}	No. of housing units in the 1 km radius	2614.21	2558.68	1.00	19,424
NP_PUB_{ir1000}	Proportion of the no. of public housing to total no. of housing units in the 1 km radius	0.04	0.05	0.00	0.43
NP_AIR_{ir1000}	Proportion of the no. of Airbnb to total no. of housing units in the 1000 m radius	0.001	0.007	0.00	1.00
NH_{ir1500}	No. of housing units in the 1.5 km radius	5090.00	4314.53	1.00	26,044
NP_PUB_{ir1500}	Proportion of the no. of public housing to total no. of housing units in the 1.5 km radius	0.04	0.05	0.00	0.38
NP_AIR_{ir1500}	Proportion of the no. of Airbnb to total no. of housing units in the 1.5 km radius	0.001	0.004	0.00	0.50
	X_{ik} Housing characteristics k of the i^{th} house, including:				
BD	Number of bedrooms	3.08	0.85	1.00	5.00
BATH	Number of bathrooms	1.50	0.66	1.00	4.00
AREA	Building floor area in square metres	142.72	61.41	24.00	410.00
COHORT	Dummy variables of the decade in which the house was built. 15 cohorts (1880, 1890, . . ., 2020)				
TENURE	Two types of tenure (freehold, leasehold)				

Notes: The 0.5% outliers are excluded; non-house type housing transactions are excluded. The attributes of the number of bedrooms, bathrooms, and floor area are continuous variables, whereas the attributes of year built (cohorts in decades) are dummy variables. The sample contains 24,800 transactions.

4.2. Empirical Results

Table 3 presents the results regarding density associations with housing prices. Column (1) displays the outcome for Model 1, which encompasses external housing density and public housing density based on SA1 units. Nearly all variables are statistically significant at the 1% level, and the model exhibits a reasonably high explanatory power (adjusted R-squared = 76%). Housing characteristic variables control various effects, including cohort, housing type, size, location, and time effects.

Table 3. The results of Models 1–5.

	Administrative Area Unit		*Radius Distancer from Each House*			*Radius Distancer from Each House (With CBD_DIST$_i$)*		
	Model 1 (SA1)	***Model 2 (SA2)***	***Model 3a (r = 0.5 km)***	***Model 4a (r = 1.0 km)***	***Model 5a (r = 1.5 km)***	***Model 3b (r = 0.5 km)***	***Model 4b (r = 1.0 km)***	***Model 5b (r = 1.5 km)***
Variables	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***
NH_{ir}	0.00005	−0.00003	−0.0001	−0.00002	−0.00001	−0.0001	−0.00002	−0.00001
	(2.38) **	(−5.49) ***	(−8.46) ***	(−6.59) ***	(−6.53) ***	(−6.12) ***	(−5.47) ***	(−4.59) ***
$nPUB_{ir}/NH_{ir}$	−0.107	−0.259	−0.363	−0.484	−0.490	−0.313	−0.367	−0.370
	(−6.48) ***	(−7.82) ***	(−11.42) ***	(−9.62) ***	(−7.52) ***	(−10.78) ***	(−8.10) ***	(−6.31) ***
$nAIR_{ir}/NH_{ir}$	-	-	0.224	0.172	0.226	0.599	−0.004	−8.14
			(2.43) **	(3.83) ***	(2.06) **	(0.77)	(−0.17)	(−3.28) ***
CBD_DIST_i	−0.011	−0.013	-	-	-	−0.003	−0.004	−0.006
	(−5.43) ***	(−6.31) ***				(−1.78) *	(−2.16) **	(−2.46) **
FIXED EFFECTS	Structure attributes, Month dummies, Suburb dummies, Cohort dummies, and Tenure dummy							
No. of Obs.	17,773	17,773	17,775	17,775	17,775	10,560	10,560	10,560
Adj. R-Squared	0.76	0.76	0.76	0.76	0.76	0.60	0.59	0.59

Notes: The dependent variable $ln(P_{its})$ is the logarithm of the net transacted housing prices in New Zealand dollars, and *, **, *** mean that the coefficient is significant at the 10%, 5%, and 1% levels, respectively. The figures in parentheses are the t-statistics. Outliers are excluded. The standard errors are Huber–White–Hinkley (HC1) heteroskedasticity consistent. Structure attributes (number of bedrooms, number of bathrooms, floor area, cohort dummies, tenure dummy) fixed effects, location (suburb dummies) fixed effects, and time (monthly dummies) fixed effects are controlled in all models.

When considering neighbouring houses at the SA1 level, the external housing density effect is found to be positive and significant at the 5% level. However, at the SA2 level (Model 2), this effect becomes negative and significant. These contradictory results highlight the challenges of measuring density associations using administrative area units. Conversely, the associations between external housing density and housing prices are consistently negative and significant across the 0.5 km, 1.0 km, and 1.5 km neighbourhood levels. Furthermore, the negative effect of housing density is weaker in larger buffers. These results agree with Fotheringham and Wong [8] and Ye and Rogerson's [64] (p. 53) conclusion that "when individual observations are aggregated differently, the sign and magnitude of aggregate-level bias are indefinite".

The effect of public housing density is negative and significant at both the SA1 and SA2 levels, as well as across the three radial neighbourhood buffer levels. However, the negative magnitudes estimated by the SA models (−0.107 at SA1 and −0.259 at SA2) are much weaker than those estimated by the radial neighbourhood buffer models (−0.363 at 0.5 km to −0.490 at 1.5 km). These results suggest that certain positive externalities arise from housing and social welfare policies favourable to public housing residents, which are better measured using administrative area unit models. In contrast, radial buffer models mostly capture the stigmatised negative externalities of public housing.

Models 3a, 4a, and 5a reveal a positive association between Airbnb density and housing prices in neighbourhoods, consistent with previous studies indicating asset value appreciation by Airbnb landlords [24]. However, the strong positive associations between Airbnb density and housing prices may be influenced by location, as these associations become insignificant or even negative when including the distance to the CBD variable, CBD_DIST_i, while the signs and significance of other coefficients remain largely unchanged, as depicted in Models 3b, 4b, and 5b. These results offer a clearer understanding of the associations between Airbnb and housing prices in neighbourhoods, as they measure Airbnb density within a certain walking distance instead of relying on the commonly used numbers of Airbnb listings in zip codes [53].

Three robustness tests on property density associations with housing prices are conducted, and the results are provided in the Appendix A. The first test examines the associations of different property types on housing prices, including residential, commercial, industrial, and government, institutional, and community (GIC) properties. The second test involves counting building floor areas instead of property numbers. The third test incorporates additional neighbourhood attributes, such as household income and ethnicity ratios, in the regression models to detect any potential variable omission biases.

5. Conclusions

The modifiable areal unit problem (MAUP) poses significant challenges in planning education, particularly in the analysis and interpretation of spatial data. This study found empirically that the association between property density in a neighbourhood and housing prices can vary depending on the aggregation of geographical area units. This variability highlights the sensitivity of analysis outcomes to the chosen spatial scale, which can lead to misleading conclusions if not carefully addressed. In planning education, understanding the MAUP underscores the importance of spatial analysis techniques that account for scale effects and encourage critical thinking about the implications of spatial aggregation. It emphasises the need for planners to be aware of the limitations and potential biases introduced by different spatial scales when making decisions or drawing conclusions based on spatial data analysis.

The paper makes a dual contribution. Firstly, it introduces a BSM to dramatically enhance computational efficiency in defining various boundaries to search and count neighbouring building structures, reducing computation time by 98% (from 14 h to 0.3 h). Secondly, employing hedonic price analysis initiates an examination of two distinct associations of housing density with housing prices: external housing density and property type-specific density, encompassing public housing density and Airbnb density. A sensitivity analysis of the areal unit scale is also conducted to demonstrate the impacts of the MAUP on the effects of property densities on housing prices. A conflicting sign of the density effects is found when administrative area units are used. More reasonable and significant results are obtained by using circular buffers of 0.5 km, 1.0 km and 1.5 km. The tests reveal negative associations between external housing density and public housing density, but there is a positive association between Airbnb density and housing prices in neighbourhoods. Specifically, external housing density incurs a discount of approximately −15.9% per one per cent increase in residential units within a 1.0 km radius neighbourhood. However, the effect varies by housing type, with detached/semi-detached houses yielding a positive premium of about 12.5%, while public housing density results in a discount of approximately −36.0%. Conversely, Airbnb density entails a premium of about 18.1% for a one per cent increase in Airbnb units within the neighbourhood.

Additionally, considering floor area proportion instead of the number of housing units, the density effect of residential properties in the 1.0 km radius neighbourhood is −4.8%. Detached/semi-detached houses, apartments, and public housing exhibit density effects of 5.4%, −2.2%, and −56.8%, respectively. The consistently negative and significant density associations of public housing are noteworthy. Unlike previous studies relying on surveys or interviews, i.e., by stated preference [25], this study leverages actual housing transaction data to provide empirical evidence on external and property type-specific densities. This revealed preference approach aids discussions on inclusive and sustainable neighbourhood planning, assisting urban planners in design decisions.

The study does encounter limitations, such as the assumption that density is uniformly associated with housing prices [65], and the risk of endogeneity bias when analysing cross-sectional data [66]. While the results confirm negative associations of density, particularly public housing density, with housing values, their practical implications may be subject to controversy. First, the dissatisfaction of homeowners may simply be a kind of statistical discrimination, i.e., discriminatory practices motivated by attempts to increase economic profits [67,68]. Second, many international organisations, such as the EU, OECD, and UN-Habitat, consider compact cities to be a sustainable urban development idea [69]. Densification or intensification is also often used to improve housing affordability in cities. Thus, the density issue can become part of a class struggle between homeowners and renters. Third, the desirability of compact developments is not universally true but is found to be "shaped by socioeconomic and cultural contexts surrounding the development" [70] (p. 4). For example, "in developing economies, urban densification often seems to exacerbate rather than mitigate urban challenges such as inequitable access to urban amenities, crowdedness, and urban poverty". Despite these constraints, the results provide valuable perspectives for conducting cost–benefit analyses in inclusive planning and assessing the implications of upzoning for urban planners. Also, while our block-searching technique provides a robust method for analysing housing density and price associations, it is not without limitations. One potential bias is the assumption of uniformity within each block, which may not perfectly capture the heterogeneity of urban environments. Additionally, the choice of border scales can influence the results, highlighting the need for sensitivity analyses. As a future direction, the application of remote sensing technologies could be explored to address the MAUP in urban planning studies. Remote sensing offers high-resolution, consistent data over large areas, which can help reduce the dependence on predefined administrative boundaries. This technology allows for the collection of detailed spatial and temporal information, facilitating the analysis of dynamic urban phenomena with greater accuracy. Integrating remote sensing data with existing spatial analysis methods can enhance the robustness of findings and provide more reliable insights for urban planning and policy-making.

Author Contributions: Conceptualisation, K.-S.C., C.-W.S. and C.-Y.Y.; methodology, K.-S.C., C.-W.S. and C.-Y.Y.; software, C.-W.S.; validation, C.-W.S. and C.-Y.Y.; formal analysis, C.-W.S. and C.-Y.Y.; investigation, K.-S.C., C.-W.S. and C.-Y.Y.; resources, K.-S.C. and C.-Y.Y.; data curation, K.-S.C. and C.-W.S.; writing—original draft preparation, K.-S.C. and C.-Y.Y.; writing—review and editing, K.-S.C., C.-W.S. and C.-Y.Y.; visualisation, K.-S.C., C.-W.S. and C.-Y.Y.; project administration, K.-S.C.; funding acquisition, K.-S.C. and C.-Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: The project is financially supported by the University of Auckland Faculty Research Development Fund (Ref no.: 3729664), and the University of Auckland Early Career Research Excellence Award (Ref no.: 3726886).

Data Availability Statement: Due to their proprietary nature, supporting data cannot be made openly available. Further information about the data and conditions for access are available at https://auckland.primo.exlibrisgroup.com/permalink/64UAUCK_INST/1lk16jl/alma99217168414002091 (accessed on 4 May 2024). The GitHub link of the block searching method. https://github.com/cwsham/Spatial-clustering-method (accessed on 10 May 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A Robustness Tests

We further conduct two robustness tests on the neighbourhood density associations with housing prices. In the first one, we count the numbers of all types of buildings in the neighbourhood of the house and analyse them by classifying them into four land use categories, viz. (a) residential, (b) commercial, (c) industrial, and (d) government, institutional and community (GIC). In the second test, we count the building floor areas of all types of buildings in the house's neighbourhood. We conduct each of the robustness tests at two levels of analysis. The first level is based on the four land use categories, and the second level is based on the sub-categories of each land use. For residential land use, in particular, we consider the neighbourhood density associations of the residential sub-categories with housing prices: houses, apartments, and public housing, and the commercial sub-categories: retail and office.

Appendix A.1 Neighbourhood Density Associations of All Building Types in Numbers

Table A1 shows the results of the level 1 models at the designated buffer radius from the subject house. $NB_{ir}, nRES_{ir}, nCOM_{ir}, nIND_{ir}$ represent the number of buildings, residential properties, commercial properties and industrial properties, respectively. The results confirm the less favourable neighbourhood density association of all types of properties in comparison with the omitted one, i.e., GIC properties. The magnitudes of the negative density association of residential properties are reduced when the neighbourhood radius is larger. However, the small magnitude of the negative associations of industrial building density with housing prices probably reflects that counting the number of buildings may not be appropriate when the property size is excessively large. We therefore further conduct another robustness test counting the building floor areas of properties.

Table A1. The Results of Models 6–8.

	Model 6 (r = 0.5 km)	***Model 7 (r = 1.0 km)***	***Model 8 (r = 1.5 km)***
Variables	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***
NB_{ir}	−0.0001	−0.00002	−0.00001
	(−6.21) ***	(−5.56) ***	(−4.17) ***
$nRES_{ir}/NB_{ir}$	−0.208	−0.159	−0.042
	(−4.05) ***	(−2.16) **	(−0.43)
$nCOM_{ir}/NB_{ir}$	−0.370	−0.343	−0.367
	(−12.06) ***	(−6.47) ***	(−5.48) ***
$nIND_{ir}/NB_{ir}$	−0.191	−0.130	0.062
	(−3.06) ***	(−1.42)	(0.51)
$DIST_i$	−0.003	−0.005	−0.006
	(−1.85) *	(−2.54) **	(−2.83) ***
FIXED EFFECTS	Structure attributes, Month dummies, Suburb dummies, Cohort dummies, Tenure dummy		
No. of Obs.	11,329	11,329	11,329
Adj. R-Squared	0.61	0.60	0.60

Notes: The dependent variable $ln(P_{its})$ is the logarithm of the net transacted housing prices in New Zealand dollars, and *, **, *** mean that the coefficient is significant at the 10%, 5%, and 1% levels, respectively. The figures in parentheses are the *t*-statistics. Outliers are excluded. Non-house-type housing is also excluded. The standard errors are Huber–White–Hinkley (HC1) heteroskedasticity consistent. Structure attributes fixed effects, location fixed effects and time fixed effects are controlled in all models.

Table A2 shows the results of the level 2 models at the designated buffer radius from the subject house. $NB_{ir}, nHSE_{ir}, nAPT_{ir}, nPUB_{ir}, nAIR_{ir}, nRET_{ir}, nOFF_{ir}, nIND_{ir}$ represent the numbers of buildings, houses, apartments, public housing, Airbnb, retail, office and industrial properties. The results confirm the less favourable neighbourhood density of all types of properties in comparison with the GIC properties. The results confirm the negative and positive neighbourhood density associations of public housing and Airbnb, even considering the total number of buildings. However, the impact of many other building types is either insignificant or inconsistent, which supports our second robustness test using floor areas.

Table A2. The Results of Models 9–11.

Variables	Model 9 (r = 0.5 km)	Model 10 (r = 1.0 km)	Model 11 (r = 1.5 km)
	Coefficient (t-Stats)	Coefficient (t-Stats)	Coefficient (t-Stats)
NB_{ir}	−0.0001	−0.00002	−0.00001
	(−6.45) ***	(−5.58) ***	(−3.74) ***
$nHSE_{ir}/NB_{ir}$	−0.009	0.125	0.114
	(−0.47)	(4.74) ***	(2.82) ***
$nAPT_{ir}/NB_{ir}$	−0.001	0.095	0.180
	(−0.01)	(1.51)	(1.76) *
$nPUB_{ir}/NB_{ir}$	−0.336	−0.360	−0.390
	(−11.40) ***	(−7.55) ***	(−5.58) ***
$nAIR_{ir}/NB_{ir}$	0.118	0.181	0.113
	(1.73) *	(2.04) **	(0.91)
$nRET_{ir}/NB_{ir}$	−0.204	0.158	−0.476
	(−2.15) **	(0.86)	(−1.86) *
$nOFF_{ir}/NB_{ir}$	−0.049	0.527	0.626
	(−0.24)	(1.61)	(1.58)
$nIND_{ir}/NB_{ir}$	−0.002	0.096	0.120
	(−0.06)	(2.15) **	(1.93) *
$DIST_i$	−0.005	−0.005	−0.005
	(−2.57) ***	(−2.42) **	(−2.50) ***
FIXED EFFECTS	Structure attributes, Month dummies, Suburb dummies, Cohort dummies, Tenure dummy		
No. of Obs.	11,329	11,329	11,329
Adj. R-Squared	0.61	0.60	0.60

Notes: The dependent variable $ln(P_{its})$ is the logarithm of the net transacted housing prices in New Zealand dollars, and *, **, *** mean that the coefficient is significant at the 10%, 5%, and 1% levels, respectively. The figures in parentheses are the t-statistics. Outliers are excluded. Non-house-type housing is also excluded. The standard errors are Huber–White–Hinkley (HC1) heteroskedasticity consistent. Structure attributes fixed effects, location fixed effects and time fixed effects are controlled in all models.

Appendix A.2 Neighbourhood Density Associations of All Building Types in Floor Areas

Table A3 shows the results of the level 1 models at the designated buffer radius from the subject house. $AB_{ir}, aRES_{ir}, aCOM_{ir}, aIND_{ir}$ represent the total floor areas of all buildings, residential properties, commercial properties and industrial properties, respectively, in the isochrone. The results confirm the less favourable neighbourhood density of all types of properties in comparison with the omitted one, i.e., GIC properties. The adverse associations of industrial building density are significantly more substantial and consistent, showing that counting the floor area of buildings is a more appropriate approach to assessing neighbourhood density associations.

Table A4 shows the results of the level 2 models at the designated buffer radius from the subject house. $AB_{ir}, aHSE_{ir}, aAPT_{ir}, aPUB_{ir}, aRET_{ir}, aOFF_{ir}, aIND_{ir}$ represent the total floor areas of all buildings, houses, apartments, public housing, retail, office and industrial properties, respectively, in the isochrone. Airbnb is not included, as the data of Airbnb's floor area are not available. First, the results confirm a reducing strength of the negative neighbourhood density association of the total floor areas of properties with housing prices when the isochrone radius increases. Second, the results also confirm an increasing strength of the negative neighbourhood density association of public housing floor areas. Third, the neighbourhood density associations of houses and apartments are consistently positive and negative, respectively, but some estimates are statistically insignificant.

Table A3. The Results of Models 12–14.

	Model 12 (r = 0.5 km)	***Model 13 (r = 1.0 km)***	***Model 14 (r = 1.5 km)***
Variables	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***
AB_{ir}	-5.41×10^{-8}	-5.69×10^{-8}	-2.11×10^{-8}
	(−1.09)	(−3.04) ***	(−1.73) *
$aRES_{ir}/AB_{ir}$	−0.039	−0.048	−0.097
	(−1.92) *	(−1.88) *	(−2.68) ***
$aCOM_{ir}/AB_{ir}$	−0.050	0.018	−0.047
	(−2.64) ***	(0.73)	(−1.60)
$aIND_{ir}/AB_{ir}$	−0.049	−0.070	−0.092
	(−2.17) **	(−2.71) ***	(−2.72) ***
$DIST_i$	−0.001	−0.002	−0.002
	(−0.62)	(−0.78)	(−1.05)
FIXED EFFECTS	Structure attributes, Month dummies, Suburb dummies, Cohort dummies, Tenure dummy		
No. of Obs.	11,329	11,329	11,329
Adj. R-Squared	0.61	0.60	0.60

Notes: The dependent variable $ln(P_{its})$ is the logarithm of the net transacted housing prices in New Zealand dollars, and *, **, *** mean that the coefficient is significant at the 10%, 5%, and 1% levels, respectively. The figures in parentheses are the t-statistics. Outliers are excluded. Non-house type housing is also excluded. The standard errors are Huber–White–Hinkley (HC1) heteroskedasticity consistent. Structure attributes fixed effects, location fixed effects and time-fixed effects are controlled in all models.

Table A4. The Results of Models 15–17.

	Model 15 (r = 0.5 km)	***Model 16 (r = 1.0 km)***	***Model 17 (r = 1.5 km)***
Variables	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***	***Coefficient (t-Stats)***
AB_{ir}	-1.78×10^{-7}	-7.92×10^{-8}	-2.33×10^{-8}
	(−3.50) ***	(−4.06) ***	(−1.77) *
$aHSE_{ir}/AB_{ir}$	0.010	0.054	0.038
	(0.66)	(2.48) **	(1.18)
$aAPT_{ir}/AB_{ir}$	−0.085	−0.022	−0.276
	(−1.30)	(−0.22)	(−1.47)
$aPUB_{ir}/AB_{ir}$	−0.458	−0.568	−0.721
	(−14.02) ***	(−10.80) ***	(−9.18) ***
$aRET_{ir}/AB_{ir}$	−0.062	0.091	−0.007
	(−1.65) *	(1.70) *	(−0.11)
$aOFF_{ir}/AB_{ir}$	0.083	0.155	−0.044
	(0.87)	(1.06)	(−0.21)
$aIND_{ir}/AB_{ir}$	−0.031	0.003	−0.015
	(−1.70) *	(0.14)	(−0.47)
$DIST_i$	−0.004	−0.004	−0.004
	(−1.91) *	(−2.01) **	(−1.81) *

Table A4. *Cont.*

	Model 15 (r = 0.5 km)	*Model 16 (r = 1.0 km)*	*Model 17 (r = 1.5 km)*
Variables	*Coefficient (t-Stats)*	*Coefficient (t-Stats)*	*Coefficient (t-Stats)*
FIXED EFFECTS	Structure attributes, Month dummies, Suburb dummies, Cohort dummies, Tenure dummy		
No. of Obs.	11,329	11,329	11,329
Adj. R-Squared	0.61	0.60	0.60

Notes: The dependent variable $ln(P_{its})$ is the logarithm of the net transacted housing prices in New Zealand dollars, and *, **, *** mean that the coefficient is significant at the 10%, 5%, 1% levels, respectively. Figures in parentheses are the t-statistics. Outliers are excluded. Non-house type housing is also excluded. The standard errors are Huber–White–Hinkley (HC1) heteroskedasticity consistent. Structure attributes fixed effects, location fixed effects and time fixed effects are controlled in all models.

Appendix A.3 Controlling More Neighbourhood Attributes

Table A5. The results of Models 3B, 4B, 5B with neighbourhood variables.

	Radius Distancer from Each House (With CBD_DIST and Neighbourhood Variables)		
	Model 3B (r = 0.5 km)	*Model 4B (r = 1.0 km)*	*Model 5B (r = 1.5 km)*
Variables	*Coefficient (t-Stats)*	*Coefficient (t-Stats)*	*Coefficient (t-Stats)*
NH_{ir}	−0.00004 (−4.90) ***	−0.00002 (−4.33) ***	−0.00001 (−3.72) ***
$nPUB_{ir}/NH_{ir}$	−0.222 (−6.70) ***	−0.204 (−3.62) ***	−0.144 (−2.10) **
$nAIR_{ir}/NH_{ir}$	0.830 (1.05)	−0.255 (−0.16)	−8.081 (−3.27) ***
$DIST_i$	−0.004 (−1.91) *	−0.005 (−2.24) **	−0.006 (−2.60) ***
$\ln(Income_i)$	−0.030 (−2.58) ***	−0.028 (−2.42) **	−0.025 (−2.13) **
$\ln(crime_i)$	−0.012 (−4.45) ***	−0.014 (−4.81) ***	−0.013 (−4.56) ***
$Asian_i$	−0.027 (−0.53)	−0.032 (−0.63)	−0.043 (−0.84)
$Maori_i$	−0.257 (−4.42) ***	−0.255 (−4.41) ***	−0.271 (−4.66) ***
$MELAA_i$	−0.632 (−2.01) **	−0.597 (−1.86) *	−0.629 (−1.92) *
$Pacific_i$	−0.126 (−2.71) ***	−0.157 (−3.29) ***	−0.197 (−4.14) ***
$Others_i$	0.414 (1.15)	0.398 (1.11)	0.265 (0.73)
FIXED EFFECTS	Structure attributes, Month dummies, Suburb dummies, Cohort dummies, and Tenure dummy		
No. of Obs.	10,418	10,418	10,418
Adj. R-Squared	0.60	0.60	0.60

Notes: Compared with the model specifications of Models 3b, 4b and 5b, seven neighbourhood variables are added in these Models 3B, 4B, 5B. They are (1) the natural logarithm of the median personal income per each neighbourhood statistical area unit where house *i* is located, $\ln(Income_i)$, (2) the natural logarithm of the number of crimes per each neighbourhood statistical area unit where house *i* is located, $\ln(crime_i)$, (3) the proportion of Asian people per each neighbourhood statistical area unit where house *i* is located, $Asian_i$, (4) the proportion of Maori people per each neighbourhood statistical area unit where house *i* is located, $Maori_i$, (5) the proportion of Middle Eastern, Latin American and African people per each neighbourhood statistical area unit where house *i* is located, $MELAA_i$, (6) the proportion of Pacific Islanders per each neighbourhood statistical area unit where house *i* is located, $Pacific_i$, (7) the proportion of other ethnic group people per each neighbourhood statistical area unit where house *i* is located, $Others_i$. The dominant ethnic group European people is omitted to allow estimations without multicollinearity. The dependent variable $ln(P_{its})$ is the logarithm of the net transacted housing prices in New Zealand dollars, and *, **, *** mean that the coefficient is significant at the 10%, 5%, 1% levels, respectively. The figures in parentheses are the t-statistics. Outliers are excluded. The standard errors are Huber–White–Hinkley (HC1) heteroskedasticity consistent. Structure attributes (number of bedrooms, number of bathrooms, floor area, cohort dummies, tenure dummy) fixed effects, location (suburb dummies) fixed effects and time (monthly dummies) fixed effects are controlled in all models.

References

1. Duranton, G.; Puga, D. The Economics of Urban Density. *J. Econ. Perspect.* **2020**, *34*, 3–26. [CrossRef]
2. Holman, N.; Mace, A.; Paccoud, A.; Sundaresan, J. Coordinating density; working through conviction, suspicion and pragmatism. *Prog. Plan.* **2015**, *101*, 1–38. [CrossRef]

3. Ahlfeldt, G.M.; Pietrostefani, E. The economic effects of density: A synthesis. *J. Urban Econ.* **2019**, *111*, 93–107. [CrossRef]
4. Melia, S.; Parkhurst, G.; Barton, H. The paradox of intensification. *Transp. Policy* **2011**, *18*, 46–52. [CrossRef]
5. Boyko, C.T.; Cooper, R. Clarifying and re-conceptualising density. *Prog. Plan.* **2011**, *76*, 1–61. [CrossRef]
6. Gehlke, C.E.; Biehl, K. Certain Effects of Grouping Upon the Size of the Correlation Coefficient in Census Tract Material. *J. Am. Stat. Assoc. Suppl.* **1934**, *29*, 169–170.
7. Openshaw, S.; Taylor, P.J. A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. In *Statistical Applications in the Spatial Sciences*; Wrigley, N., Bennett, R.J., Eds.; Pion: London, UK, 1979; Volume 21, pp. 127–144.
8. Fotheringham, A.S.; Wong, D.W.-S. The modifiable areal unit problem in multivariate statistical analysis. *Environ. Plan. A Econ. Space* **1991**, *23*, 1025–1044. [CrossRef]
9. Flowerdew, R. How serious is the modifiable areal unit problem for analysis of English census data? *Popul. Trends* **2011**, *145*, 102–114. [CrossRef] [PubMed]
10. Li, L.; Ban, H.; Wechsler, S.P.; Xu, B. 1.22—Spatial Data Uncertainty. In *Comprehensive Geographic Information Systems*; Huang, B., Ed.; Elsevier: Amsterdam, The Netherlands, 2018; pp. 313–340. [CrossRef]
11. Prouse, V.; Ramos, H.; Grant, J.L.; Radice, M. How and when Scale Matters: The Modifiable Areal Unit Problem and Income Inequality in Halifax. *Can. J. Urban Res.* **2014**, *23*, 61–82.
12. Wong, D.W.-S. The modifiable areal unit problem (MAUP). In *The SAGE Handbook of Spatial Analysis*; Fotheringham, A.S., Rogerson, P.A., Eds.; Sage Publications Ltd.: London, UK, 2009; pp. 105–123.
13. Su, M.D.; Lin, M.-C.; Wen, T.H. Spatial Mapping and Environmental Risk Identification. In *Encyclopaedia of Environmental Health*; Nriagu, J.O., Ed.; Elsevier: Amsterdam, The Netherlands, 2011; pp. 228–235. [CrossRef]
14. Kwan, M.-P. The Limits of the Neighbourhood Effect: Contextual Uncertainties in Geographic, Environmental Health, and Social Science Research. *Ann. Am. Assoc. Geogr.* **2018**, *108*, 1482–1490.
15. Parenteau, M.-P.; Sawada, M. The modifiable areal unit problem (MAUP) in the relationship between exposure to NO2 and respiratory health. *Int. J. Health Geogr.* **2011**, *10*, 58. [CrossRef] [PubMed]
16. Chen, X.; Ye, X.; Widener, M.J.; Delmelle, E.; Kwan, M.P.; Shannon, J.; Jia, P. A systematic review of the modifiable areal unit problem (MAUP) in community food environmental research. *Urban Inform.* **2022**, *1*, 22. [CrossRef]
17. Bramley, G.; Dunmore, K.; Dunse, N.; Gilbert, C.; Thanos, S.; Watkins, D.; The Implications of Housing Type/Size Mix and Density for the Affordability and Viability of New Housing Supply. National Housing and Planning Advice Unit, Fareham. 2010. Available online: https://www.thinkhouse.org.uk/site/assets/files/1618/nhpau_20size.pdf (accessed on 10 May 2024).
18. Lepore, S.J. Crowding: Effects on health and behaviour. In *Encyclopedia of Human Behavior*; Ramachandran, V.S., Ed.; Academic Press: San Diego, CA, USA, 1994; Volume 2, pp. 43–51.
19. Miles, D. Population Density, House Prices and Mortgage Design. *Scott. J. Political Econ.* **2012**, *59*, 444–466. [CrossRef]
20. Oh, K.; Jeong, Y.; Lee, D.; Lee, W.; Choi, J. Determining development density using the urban carrying capacity assessment system. *Landsc. Urban Plan.* **2005**, *73*, 1–15. [CrossRef]
21. Ng, A.W.L. Mental Health Impacts on People Living in Subdivided Flats in Hong Kong. *Prehospital Disaster Med.* **2019**, *34*, S150. [CrossRef]
22. Leung, K.M.; Yiu, C.Y. Rent determinants of sub-divided units in Hong Kong. *J. Hous. Built Environ.* **2019**, *34*, 133–151. [CrossRef]
23. Tang, B.S.; Yiu, C.Y. Space and scale: A study of development intensity and housing price in Hong Kong. *Landsc. Urban Plan.* **2010**, *96*, 172–182. [CrossRef]
24. Cheung, K.S.; Yiu, C.Y. Touristification, Airbnb and the tourism-led rent gap: Evidence from a revealed preference approach. *Tour. Manag.* **2022**, *92*, 104567. [CrossRef]
25. Bramley, G.; Power, S. Urban form and social sustainability: The role of density and housing type. *Environ. Plan. B Plan. Des.* **2009**, *36*, 30–48. [CrossRef]
26. Jiao, L. Urban land density function: A new method to characterise urban expansion. *Landsc. Urban Plan.* **2015**, *139*, 26–39. [CrossRef]
27. Myers, D.; Gearin, E. Current preferences and future demand for denser residential environments. *Hous. Policy Debate* **2001**, *12*, 633–659. [CrossRef]
28. Roe, B.; Irwin, E.G.; Morrow-Jones, H.A. Changes in homeowner preferences for housing density following 11 September 2001. *Appl. Econ. Lett.* **2005**, *12*, 73–78. [CrossRef]
29. Xiong, C.; Cheung, K.S. Understanding sellers' agents in the residential property market. *Int. J. Strateg. Prop. Manag.* **2021**, *25*, 179–189. [CrossRef]
30. Ewing, R.; Cervero, R. Travel and the Built Environment: A Synethesis. *J. Am. Plan. Assoc.* **2010**, *76*, 265–294. [CrossRef]
31. Daly, J.; Gronow, S.; Jenkins, D.; Plimmer, F. Consumer behaviour in the valuation of residential property. *Prop. Manag.* **2003**, *21*, 295–314. [CrossRef]
32. Acolin, A.; Colburn, G.; Walter, R.J. How do single-family homeowners value residential and commercial density? It depends. *Land Use Policy* **2022**, *113*, 105898. [CrossRef]
33. Fesselmeyer, E.; Seah, K.Y.S.; Kwok, J.C.Y. The effect of localised density on housing prices in Singapore. *Reg. Sci. Urban Econ.* **2018**, *68*, 304–315. [CrossRef]
34. Lee, J.S. Measuring the value of apartment density? The effect of residential density on housing prices in Seoul. *Int. J. Hous. Mark. Anal.* **2016**, *9*, 483–501. [CrossRef]

35. Samsudin, R.; Yok, T.P.; Chua, V. Social capital formation in high density urban environments: Perceived attributes of neighbourhood green space shape social capital more directly than physical ones. *Landsc. Urban Plan.* **2022**, *227*, 104527. [CrossRef]
36. Wang, Y.; You, Y.; Yue, X.; Wu, Y.; Zhou, Y.; Zhang, H.O. Spatial variation characteristics of housing conditions in China. *Buildings* **2023**, *13*, 2028. [CrossRef]
37. Thackway, W.T.; Ng, M.K.M.; Lee, C.L.; Shi, V.; Pettit, C.J. Spatial variability of the 'Airbnb effect': A spatially explicit analysis of Airbnb's impact on housing prices in Sydney. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 65. [CrossRef]
38. Cheung, K.S.; Monkkonen, P.; Yiu, C.Y. The Heterogeneous Impacts of Widespread Upzoning: Lessons from Auckland, New Zealand. *Urban Stud.* **2024**, *61*, 943–967. [CrossRef]
39. Greenaway-McGrevy, R.; Pacheco, G.; Sorensen, K. The effect of upzoning on house prices and redevelopment premiums in Auckland, New Zealand. *Urban Stud.* **2020**, *58*, 959–976. [CrossRef]
40. Murray, C.; Limb, M. We Zoned for Density and Got Higher House Prices: Supply Price Effects of Upzoning over 20 Years. *Urban Policy Res.* **2022**, *41*, 129–147. [CrossRef]
41. Souders, A. The Effect of Population Density on Housing Prices: A Cross State. AnalysisPh.D. Thesis, The University of Akron, Akron, OH, USA, 2016.
42. Sequeira, V.; Filippova, O. Does concentration of social housing influence house prices? Evidence from New Zealand. *Int. J. Hous. Mark. Anal.* **2020**, *14*, 720–744. [CrossRef]
43. Dong, H. Higher density development for lower cost housing? Understanding the multifamily housing market and the role of density in multifamily home prices. *J. Plan. Educ. Res.* **2020**, *43*, 0739456X20912829. [CrossRef]
44. Cheung, K.S.; Yiu, C.Y. The economics of architectural aesthetics: Identifying price effect of urban ambiences by different house cohorts. *Environ. Plan. B Urban Anal. City Sci.* **2022**, *49*, 1741–1756. [CrossRef]
45. Hyslop, D.R.; Rea, D. Do housing allowances increase rents? Evidence from a discrete policy change. *J. Hous. Econ.* **2019**, *46*, 101657. [CrossRef]
46. Palmer, C.; Ziersch, A.; Arthurson, K.; Baum, F. Challenging the Stigma of Public Housing: Preliminary Findings from a Qualitative Study in South Australia. *Urban Policy Res.* **2004**, *22*, 411–426. [CrossRef]
47. Woo, A.; Joh, K.; van Zandt, S. Unpacking the impacts of the low-income housing tax credit program on nearby property values. *Urban Stud.* **2016**, *53*, 2488–2510. [CrossRef]
48. Sandler, D.H. Externalities of public housing: The effect of public housing demolitions on local crime. *Reg. Sci. Urban Econ.* **2017**, *62*, 24–35. [CrossRef]
49. Guttentag, D. Airbnb: Disruptive innovation and the rise of an informal tourism accommodation sector. *Curr. Issues Tour.* **2015**, *18*, 1192–1217. [CrossRef]
50. Lambea Llop, N. A policy approach to the impact of tourist dwellings in condominiums and neighbourhoods in Barcelona. *Urban Res. Pract.* **2017**, *10*, 120–129. [CrossRef]
51. Horn, K.; Merante, M. Is home-sharing driving up rents? Evidence from Airbnb in Boston. *J. Hous. Econ.* **2017**, *38*, 14–24. [CrossRef]
52. Franco, S.F.; Santos, C.D.; Longo, R. *The Impact of Airbnb on Residential Property Values and Rents: Evidence from Portugal (FEUNL Working Paper Series wp630)*; Universidade Nova de Lisboa, Faculdade de Economia: Carcavelos, Portugal, 2019. [CrossRef]
53. Barron, K.; Kung, E.; Proserpio, D. The effect of home-sharing on house prices and rents: Evidence from Airbnb. *Mark. Sci.* **2020**, *40*, 23–47. [CrossRef]
54. Bangura, M.; Lee, C.L. The determinants of homeownership affordability in Greater Sydney: Evidence from a submarket analysis. *Hous. Stud.* **2023**, *38*, 206–232. [CrossRef]
55. Braitman, D.; Koomen, E. Residential density change: Densification and urban expansion. *Comput. Environ. Urban Syst.* **2015**, *54*, 32–46. [CrossRef]
56. Hargreaves, A.J. Representing the dwelling stock as 3D generic tiles estimated from average residential density. *Comput. Environ. Urban Syst.* **2015**, *54*, 280–300. [CrossRef]
57. Milias, V.; Psyllidis, A. Assessing the influence of point-of-interest features on the classification of place categories. *Comput. Environ. Urban Syst.* **2021**, *86*, 101597. [CrossRef]
58. Sun, K.; Hu, Y.; Ma, Y.; Zhou, R.Z.; Zhu, Y. Conflating point of interest (POI) data: A systematic review of matching methods. *Comput. Environ. Urban Syst.* **2023**, *103*, 101977. [CrossRef]
59. Sham, C.W.; Cheung, K.S.; Yiu, C.Y. Cluster based Indexing for Spatial Analysis on Read-only Database. In Proceedings of the 9th International Conference on Applied System Innovation (ICASI), Chiba, Japan, 21–25 April 2023. [CrossRef]
60. MHUD. *Public Housing Quarterly Report June 2020*; Ministry of Housing and Urban Development (MHUD), New Zealand Government: Wellington, New Zealand, 2020.
61. Auckland Council. Rating Information Database (RID). Auckland Council, New Zealand Government. 2022. Available online: https://www.aucklandcouncil.govt.nz/property-rates-valuations/change-your-address-or-property-details/Pages/change-property-information-object-rid.aspx (accessed on 10 May 2024).
62. AirDNA. *Auckland City, Airbnb Data*; AirDNA, LLC: Denver, CO, USA, 2021; Available online: https://www.airdna.co/vacation-rentaldata (accessed on 10 May 2024).
63. StatsNZ. Statistical Area 2020. 2020. Available online: https://datafinder.stats.govt.nz/layer/104271-statistical-area-2-2020-generalised/ (accessed on 10 May 2024).

64. Ye, X.; Rogerson, P.A. The impacts of the modifiable areal unit problem (MAUP) on omission error. *Geogr. Anal.* **2022**, *54*, 32–57. [CrossRef]
65. Redfearn, C.L. How informative are average effects? Hedonic regression and amenity capitalisation in complex urban housing markets. *Reg. Sci. Urban Econ.* **2009**, *39*, 297–306. [CrossRef]
66. Lynch, S.M.; Brown, J.S. Stratification and Inequality Over the Life Course. In *Handbook of Aging and the Social Sciences*, 7th ed.; Binstock, R.H., George, L.K., Eds.; Academic Press: Cambridge, MA, USA, 2011; pp. 105–117. [CrossRef]
67. Phelps, E.S. The statistical theory of racism and sexism. *Am. Econ. Rev.* **1972**, *62*, 659–661.
68. Cheung, K.S.; Yiu, C.Y.; Lin, B.P. Love thy neighbours or do ethnic neighbourhood qualities matter? Ethnic price differentials in a multi-ethnic housing market. *N. Z. Econ. Pap.* **2024**, *58*, 20–39. [CrossRef]
69. European Commission. Green Paper on the Urban Environment: Communication from the Commission to the Council and Parliament. 1990. Available online: https://publications.europa.eu/en/publication-detail/-/publication/0e4b169c-91b8-4de0-9fed-ead286a4efb7/language-en (accessed on 10 May 2024).
70. Tan, P.Y.; Rinaldi, B.M. Landscapes for compact cities. *J. Landsc. Archit.* **2019**, *14*, 4–7. [CrossRef]