Motu Note 53

Distinctive neighbourhood housing patterns in Aotearoa New Zealand

economic & public policy research

Hannah Kotula & David C Maré

August 2024

Document information

Author contact details
Hannah Kotula & David C Maré
Motu Economic and Public Policy Research
PO Box 24390, Wellington
hannah.kotula@motu.org.nz; dave.mare@motu.org.nz

Acknowledgements

The research was funded through the MBIE Endeavour Fund project *WERO: Working to End Racial Oppression* (Contract UOWX2002) `https://wero.ac.nz`

Disclaimer
Access to the data used in this study was provided by Stats NZ under conditions designed to give effect to the security and confidentiality provisions of the Data and Statistics Act 2022. The results presented in this study are the work of the authors, not Stats NZ or individual data suppliers.

These results are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI) which is carefully managed by Stats NZ. For more information about the IDI please visit `https://www.stats.govt.nz/integrated-data/`.

Motu Economic and Public Policy Research
PO Box 24390 info@motu.org.nz +64 4 9394250
Wellington www.motu.org.nz
New Zealand

 ISSN 2422-9059.

Abstract
This paper summarises distinct housing and demographic patterns across neighbourhoods in New Zealand's main urban areas, using data from the 2018 Census of Population and Dwellings. It uses exploratory factor analysis to classify neighbourhood types. It contributes background information for a broader research programme - *WERO: Working to End Racial Oppression*.

DOI
10.29310/WP.2024.53

JEL codes
J15, R31

Keywords
Housing markets; housing-market discrimination

1 Introduction

This research note summarises a set of distinctive housing and demographic patterns across neighbourhoods within urban areas of Aotearoa New Zealand. The research is motivated by a desire to identify locations where racism and discrimination within housing markets are potentially having an adverse effect on the socioeconomic outcomes of residents. The neighbourhood-level patterns described here cannot, by themselves, pin down the existence, extent, or nature of discrimination. Spatial variation in housing market outcomes reflects the complex interaction of many factors, including variation in the distribution of income and wealth, in the preferences of residents, and in the supply and demand of different forms of housing. Housing-related racism and discrimination act to limit the range of options available to people, or to raise the costs of some otherwise desirable options. Systemic racism, or racism affecting non-housing-related domains such as employment, education or social capital, can also contribute to housing inequality − even in the absence of housing-market discrimination.

The myriad of factors that can influence housing patterns presents a challenge to detecting the presence (or absence) of discrimination in housing outcomes. There are several approaches for identifying where discrimination may be occurring in housing markets. The most direct approach consists of audit studies. This involves multiple applicants who differ in only one respect, generally ethnicity, applying for the same rental property or house purchase. Audit studies have been used to detect discrimination against minority renters and buyers (Ondrich et al., 2003; Yinger, 1986; Christensen and Timmins, 2022).

Surveys can also provide information about discrimination by capturing people's reported experiences and perceptions of discrimination (*i.e.,* subjective responses) (e.g., Yeung and Crothers, 2016). Critical discourse analysis can be used to detect discriminatory attitudes and behaviours by examining the language people use in spoken or written discourse. In New Zealand, Lewis et al. (2020) and Norris and Nandedkar (2022) used critical discourse analysis to examine how ethnicity is framed in housing-related political discourse and peer-reviewed articles, respectively.

Indirect approaches based on identifying patterns of outcomes that might be expected to arise in the presence of housing-market discrimination can also be used to detect discrimination. These approaches include examining measures which are based on where people live, and examining other housing-market outcomes such as purchase prices (e.g., Cutler et al., 1999; King and Mieszkowski, 1973).

In this paper we take an indirect approach to detecting discrimination by focusing on housing patterns. We apply statistical methods (specifically, exploratory factor analysis) to identify distinctive neighbourhood housing and demographic patterns, using data from the 2018 New Zealand Census of Population and Dwellings. These patterns capture co-occurring neighbourhood-level characteristics. Examining such patterns cannot identify the extent of housing discrimination, but is useful in focusing attention on patterns that can be further investigated by in-depth research. Section 2 documents the data and variables that we use. Section 3 documents the methods that we use, followed by results in Section 4, and a concluding discussion in Section 5.

The current study is being prepared as an input for further research into the *geographies of racism*. This workstream aims to identify ethnicity-based residential patterns in different urban areas of Aotearoa, and the institutional drivers and discriminatory practices that underpin geographically uneven impacts of housing inequalities. This study forms part of a broader research programme *Working to End Racial Oppression (WERO)*[1] that takes up the challenge of confronting institutional and interpersonal racism in Aotearoa. The contribution of the current paper is to provide contextual information for other components of the *geographies of racism* stream: a survey-based study of residential choices and constraints; a qualitative study of rental housing market practices, and a multi-method study focused on health outcomes of deprivation related to segregation.

2 Data description

We use individual-, household-, and dwelling- level data from the 2018 census, focusing on variation within and between the 26 main urban areas. We restrict attention to the usually resident population of New Zealand (*i.e.,*

[1] For further details, see `https://wero.ac.nz`.

we exclude visitors, and individuals who were overseas or absent at the time of the 2018 census). The specific variables that we use are described below in Section 2.2. We weight household and dwelling variables by the number of usual residents within a household or dwelling, respectively. These variables thus represent the proportion of individuals who, for example, live in a crowded home rather than the proportion of homes that are crowded.

2.1 Defining neighbourhoods

Data are analysed for all *neighbourhoods* within major urban areas. We base our concept of neighbourhood on Statistics New Zealand meshblocks. Meshblocks typically contain 30–60 dwellings, though some meshblocks have a population of zero (*e.g.,* meshblocks covering water).[2] For each meshblock, we summarise the characteristics of an area surrounding (and including) the meshblock. Population proportions are calculated as the spatially weighted sum of people with a particular characteristic (or in a particular type of household or dwelling), divided by the spatially weighted sum of the whole population. We use spatial weights based on an Epanechnikov kernel with a bandwidth of $500m$, and use straight-line distance between meshblock centroids to determine the edge of the bandwidth. This approach gives spatially-smoothed measures, which are by construction correlated across nearby neighbourhoods. These spatially-smoothed measures are less susceptible to the modifiable areal unit problem.[3]

2.2 Variables

Individual-level variables

- Ethnicity
  - In census data, individuals can identify multiple ethnic affiliations. We classify ethnicity based on the 10 most prevalent combinations of responses within the usually resident population of New Zealand, along with one residual category capturing all other responses. The 10 main categories are: New Zealand European, a combined category of single ethnicity individuals not classified elsewhere, Chinese, Indian, Māori, New Zealand European *and* Māori dual ethnicity, Samoan, Filipino, Tongan, and English.[4]
- Country of birth
  - We identify the 10 most common countries of birth within the usually resident population of New Zealand. These are: New Zealand, China, England, India, the Philippines, Fiji, Samoa, South Africa, Australia, and Korea.[5] We also include one residual category capturing all other responses.
- Migrant status
  - We define *recent migrants* as individuals who arrived in New Zealand less than five years prior to the census.
- Individual income
  - Income is measured as total annual individual income from all sources, for individuals aged 18 and over. Income is reported in bands and we assign responses to the mid-point of each band. We then classify incomes by quartile. We do this by determining the cut-off values of individual income that divide the usually resident New Zealand population into four equal groups. These values are: $lowerquartile = \$17502$, $median = \$32502$, $upper\ quartile = \$55002$.[6]
  - We focus on individuals in the 1^{st} quartile (*i.e.,* individuals with an annual income lower than $\$17502$). We consider only those individuals aged 18 and over for this variable.
- Highest qualification

[2] Statistics New Zealand (2016). Statistical standard for meshblock. Available from `www.stats.govt.nz`. A meshblock is the smallest geographic unit for which statistical data is reported by StatsNZ.
[3] Analysing spatial data using different boundary definitions can lead to inconsistent results. This problem is known as the modifiable areal unit problem.
[4] Ethnic categories are ordered by prevalence based on smoothed counts.
[5] Country of birth is ordered by prevalence based on smoothed counts.
[6] Note that excluding individuals under 15 (or under 18) does not change these cut-off values.

- ◦ We classify highest qualification into two levels (*low* and *high*). Individuals who have completed secondary school or hold no qualification are assigned a qualification level of *low*, while those who have completed a post-school qualification or a Bachelor's degree (or above) are assigned a qualification level of *high*. We consider only those individuals aged 15 and over for this variable.

Household-level variables

- Household income
  - ◦ Household income is calculated as the sum of individual incomes of all adults (aged 18 and over) in the household.
  - ◦ We use equivalised household income, based on the Jensen equivalisation scale. We determine the cut-off values of equivalised income that would divide the usually resident population of New Zealand into four equal groups. We focus on individuals living in households in the first quartile of the national income distribution (*i.e.,* those in the lowest 25%).
- Canadian crowding index
  - ◦ We focus on the proportion of individuals living in a crowded home.[7]

Dwelling-level variables

We consider only occupied private permanent dwellings.

- Number of bedrooms
  - ◦ We use the proportion of individuals living in a dwelling with three or fewer bedrooms.
- Dwelling tenure
  - ◦ We use the proportion of individuals living in a dwelling they do not own nor partly own nor hold in a family trust.
- Sector of landlord
  - ◦ This includes the categories listed below. Note that we consider this variable only for those renting (and paying rent).
    - ∗ Private person, trust, or business
    - ∗ Local authority or city council
    - ∗ Housing New Zealand (now Kāinga Ora)
    - ∗ Iwi, hapū, or Māori land trust
    - ∗ Other community housing provider
    - ∗ Other state or government owned housing

3 Methods

We provide an overview of our methods before providing a full description.

Overview

We used factor analysis to identify neighbourhood-level patterns in housing and socio-demographic characteristics. We applied factor analysis to the variables listed in Section 2.2. Each of these variables was included as a spatially-smoothed population proportion within a neighbourhood. Factor analysis groups variables that tend to occur together into factors (*i.e.,* patterns). This allows us to identify co-occurring neighbourhood-level characteristics. We then calculated a factor score for each neighbourhood (*i.e.,* meshblock) that reflects the similarity of its composition to that of each of the commonly-occurring patterns. Recall that each pattern represents a set of variables that tend to occur together, and that each pattern captures co-occurring characteristics at

[7] For more information see: `https://datainfoplus.stats.govt.nz/Item/nz.govt.stats/f5494ff5-c7f7-45aa-b4cc-e14a8c3544e0`

the neighbourhood level (rather than the individual level). We focus on the first pattern because the remaining patterns do not capture distinctive housing characteristics. To visualise the distribution of the first pattern, we provide maps of the main urban areas (with darker shades of grey representing a high similarity with the given pattern). The variables included in each pattern are as described in the Section 4. We provide a series of boxplots to visualise the variability of factor scores within and between urban areas (Figures 1, 3 and 4), and a table with the share and incidence of individuals residing in *high* factor areas (Table 2).

Full description

We use exploratory factor analysis to identify commonly occurring patterns of neighbourhood housing and socio-demographic characteristics.[8] By analysing correlations between meshblock-level measures, factor analysis groups variables together thereby capturing sets of characteristics that tend to occur together. Each neighbourhood is assigned a set of factor scores that reflect how similar its composition is to that of each of the commonly-occurring patterns. Note that these patterns are at the neighbourhood level (rather than the individual level). We use the Kaiser–Meyer–Olkin statistic to test the strength of correlation between the neighbourhood characteristics. For our data, the statistic has a value of 0.79, indicating that factor analysis is appropriate for our data.

We apply principal-component factor analysis to the variables listed in Section 2.2.[9] We omit one category from each set of effects to avoid collinearity. Other estimated effects are then measured relative to the omitted category. The omitted category for each variable are shown in brackets in the following list: ethnicity (New Zealand European), birthplace (New Zealand), recent migrant (missing), individual income (first quartile), qualification level (low), household income (fourth quartile), crowding (not crowded), number of bedrooms (fewer than three), tenure (dwelling owned or held in a family trust), sector of landlord (private). For each variable, the omitted group is the largest group. The choice of omitted group does not affect the results of the analysis.

Each variable included in the factor analysis is captured as a spatially-smoothed population proportion within a neighbourhood (see section 2.1 above for a description of how we construct these spatially-smoothed measures). As mentioned above, we consider the proportion of individuals who, for example, live in a crowded home rather than the proportion of homes that are crowded. Factors are formed as linear combinations of meshblock-level variables that explain common variance between variables, and thus represent a summary measure of multiple variables. We use the Kaiser criterion to determine the number of factors to retain. There were 9 factors that met this criterion (*i.e.,* had an eigenvalue greater than one).

Factor analysis calculates a loading of each variable on each factor, allowing us to determine which variables are most strongly captured by each factor. In other words, factor loadings can be used to associate variables with factors. We focus on variables with an absolute loading ≥ 0.75, noting that a negative factor loading corresponds to a negative correlation between a given variable and factor. Variables that load strongly on a given factor are highly correlated with that factor (*i.e.,* the factor explains a high proportion of the variance in the variable), whereas variables that load weakly on a factor are poorly explained by the factor. Factor assignment can be ambiguous if a variable has a similar absolute loading on more than one factor. None of the variables in our dataset load strongly onto more than one factor. We rotate the factors to increase interpretability. We include a table with the rotated factor loadings in the Section 4 (Table 1).

After rotating the factors (using an orthogonal *varimax* rotation), we calculate factor scores using the default regression method. The regression method z-standardizes each factor to have zero mean and unit standard deviation. This means we can interpret an observation's score in terms of units of standard deviations from the mean of the factor. For example, an observation's factor score of 0.79 implies that this observation is 0.79 standard deviations above the average with regard to the corresponding factor.

For the 26 main urban areas in New Zealand, we map the scores of each factor to show its spatial distribution using Google Maps as a base layer. Maps showing the spatial distribution of factor 1 are included below in

[8] Note that all analyses were conducted in the Statistics New Zealand Datalab. The maps and boxplots included in this report are based directly on data released from the Datalab.

[9] We use the pcf option for the factor command in Stata.

Section 4.1. Because each factor captures a set of variables that are correlated with one another (*i.e.,* a set of variables with a similar spatial distribution), these maps represent a summary of many variables. As mentioned above, each factor captures co-occurring characteristics at the neighbourhood level (rather than co-occurring characteristics at the individual level).

Our focus is on understanding how discrimination may shape housing patterns. Thus we map only factor 1 as the remaining factors do not capture distinctive housing characteristics. However, we include written descriptions of factor 2 and 3 (in Section 4) as these factors explain a significant amount of variance (13% and 12%, respectively).

We use a consistent colour scheme so that colours have the same meaning across urban areas. This means maps of factor 1 (which we focus on, as mentioned above) can be compared across urban areas. To do this, we calculate a population weighted national lower quartile, median and upper quartile for the factor.[10] We use these values as the cut-offs for determining the colour of each meshblock (in our case four shades from light to dark grey) across all urban areas. Meshblocks are coloured the lightest shade of grey if the factor score for the meshblock lies between 0 and the lower quartile (*i.e.,* meshblocks in the lowest 25% for factor 1 are coloured light grey). Note that the first and last value in the legend (*e.g.,* -1.22 and 0.51 for factor 1 in Whangarei) correspond to the minimum and maximum factor score within the given urban area, respectively, and these values may vary across urban areas. Meshblocks with a population of zero (*e.g.,* those covering water) or those with a suppressed value are shown by a cross-hatch pattern.[11]

To visualise the variability of factor scores within and between urban areas, we show a series of boxplots for the first three factors (Figures 1, 3 and 4). We include boxplots for factors 2 and 3 because, as mentioned above, these factors capture a significant amount of variance. Each boxplot shows the population weighted lower quartile, median, and upper quartile values for a factor for a given urban area.[12] In these figures, urban areas are ordered by median value. Within urban area variability is shown by the width of the boxplot – wider boxes indicate greater variability. Between urban area variability is shown by the alignment of boxes – greater overlap indicates greater similarity.

We also present a table with the share and incidence of individuals residing in a *high* factor area for factors 1, 2, and 3 (Table 2). We define an area as *high* if it has a factor score in the top 25% of the national distribution for the given factor. *Incidence* refers to the proportion of individuals within a given area that live in a high factor area. *Share* allows us to answer the question: considering only individuals who live in high factor areas, what proportion of these individuals live in each urban area?

4 Results

We use factor analysis to identify distinctive housing and demographic patterns across neighbourhoods within urban areas of Aotearoa New Zealand. These patterns describe the co-occurrence of neighbourhood-level characteristics (i.e., these patterns cannot be used to determine the co-occurrence of characteristics at the individual level). Our results highlight three main neighbourhood-level patterns. [13] The first pattern (*i.e.,* factor 1) captures areas with a high proportion of individuals who live in a crowded home and/or live in a home owned by Housing New Zealand (now Kāinga Ora). These areas also correspond to areas that have a high proportion of individuals who identify as Samoan or Tongan. The factor loadings of these variables are shown in Table 1. Note that this pattern does not inform us whether individuals living in a crowded home also live in a home

[10] We use the 26 main urban areas in New Zealand to calculate these national values. We weight by meshblock-specific (*i.e.,* not spatially weighted) total population counts.

[11] In line with StatsNZ's confidentiality rules, we suppressed factor score values if the population of the meshblock was 20 or less. Note that all other values released from the datalab fall under the category of highly derived and thus did not require any confidentiality rules to be applied to them.

[12] Note that we weighted by unsmoothed meshblock-level population data.

[13] As explained in the method, we focus on the first three factors because they capture a substantial proportion of variance (19%, 13%, and 12%, respectively). The eigenvalues of the first three factors are: 10.5, 7.9, and 3.7, respectively.

owned by Kāinga Ora or identify as Samoan or Tongan – rather this pattern informs us that at a neighbourhood level these characteristics tend to occur together.

The second and third patterns (*i.e.,* factors 2 and 3) capture income and ethnic sorting, respectively. Neither of these patterns correspond strongly to any of the housing characteristics included in our dataset.[14] In other words, none of the housing characteristics included in our dataset load strongly on either factor 2 or 3. Pattern 2 captures areas with a high proportion of individuals who have a low income. These areas have little overlap with areas containing a high proportion of people who hold a high qualification. Pattern 3 indicates that individuals who were born in Korea or China, or who identify as Chinese, tend to live in similar areas.

As mentioned in the methods section above, we focus on the first three factors because these factors collectively explain 44% of the total variance. Pattern 1 accounts for 19% of this explained variance, while pattern 2 and 3 account for 13% and 12%, respectively.

The strength of housing patterns can vary both within and across urban areas. The median factor 1 score is similar across urban areas indicating that, overall, the prevalence of high-factor 1 areas is similar across urban areas. However, the amount of variability within an urban area varies across Aotearoa. Specifically, there was much greater diversity within Porirua and South Auckland (Figure 1). In each of these urban areas, there are places with high factor 1 scores and places with low factor 1 scores, *i.e.,* these urban areas are heterogeneous with distinct neighbourhood patterns (Figure 2iii,xvii). The strength of pattern 1 varies little within the remaining urban areas.

The proportion of individuals living in a high factor 1 area (*i.e.,* the incidence of factor 1) is highest for Porirua, South Auckland, and West Auckland – 60%, 69%, 72%, respectively, of individuals in these areas live in a high factor 1 area (Table 2). Moreover, South and Central Auckland have the highest share of individuals who live in a high factor 1 area – together these two areas account for 54% of all individuals who live in a high factor 1 area. In other words, 54% of individuals who live in a high factor 1 area, live in South or Central Auckland. Recall that we define *high* factor areas as those in the top 25% of the national distribution.

Pattern 2 captures income sorting, and this pattern varies little within and across urban areas (Figure 3). The share of individuals living in a high factor 2 area is reasonably evenly distributed across urban areas, though Christchurch and Hamilton have slightly higher shares (13% and 8%, respectively).The incidence of factor 2 is also similar across urban areas, with Wanganui and Gisborne having a slightly higher incidence of individuals living in a high factor 2 area.

Pattern 3 captures ethnic sorting, and the strength of this pattern varies significantly within South and North Auckland (Figure 4). In South Auckland, there are distinct areas with high factor 3 values. The median value for factor 3 is similar across urban areas, with the exception of Auckland which has a higher median value than the rest of the country. The incidence of individuals living in a high factor 3 area is highest for Central and North Auckland (59% and 74%, respectively). The share of individuals living in a high factor 3 area is also highest for Central and North Auckland (23% and 28%, respectively).

[14] Housing characteristics refer to the following variables: sector of landlord, number of bedrooms, crowding, and tenure. Note that housing characteristics are different from household-level variables (as described in Section 2.2).

Table 1: Rotated factor loadings and unique variances

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9	Uniqueness
Crowding - yes	0.834	0.2485	0.0288	0.269	0.2559	0.1774	0.0662	0.024	0.0322	0.0666
Sector of landlord - Housing New Zealand	0.8069	0.267	-0.087	0.1781	-0.1417	-0.0049	0.0376	0.0523	-0.0867	0.2065
Dwelling tenure - dwelling not owned nor held in a trust	0.3358	0.1853	0.0237	0.6623	0.1819	0.4118	-0.1324	0.1153	0.1042	0.1693
Number of bedrooms - more than three	-0.0314	-0.3089	0.4008	-0.4657	-0.0762	-0.4716	0.07	0.0176	-0.1409	0.2727
Sector of landlord - local authority	-0.0981	0.1476	-0.1387	0.1979	0.026	0.0199	-0.5836	0.1324	0.2401	0.4933
Sector of landlord - other state	0.0603	0.0267	-0.0136	0.0591	-0.0305	0.03	-0.0272	0.7724	-0.0829	0.386
Sector of landlord - other community	0.1747	0.2929	-0.0991	0.2221	0.0226	0.1072	-0.1506	0.5324	-0.0796	0.5001
Sector of landlord - Iwi	-0.0283	-0.0787	-0.0362	-0.0031	0.0043	-0.008	0.2811	0.4883	0.252	0.6106
Sector of landlord - missing	-0.0357	-0.017	-0.0391	-0.0011	0.1031	-0.016	0.0876	0.1099	-0.7333	0.4286
Crowding - missing	0.1277	0.1059	-0.1326	0.2977	0.1975	0.0035	0.1783	0.1627	0.4428	0.5729
Ethnicity - Samoan	0.9458	0.1147	-0.0837	-0.0047	0.0739	-0.0346	-0.0554	-0.0349	0.0202	0.0739
Birthplace - Samoa	0.9403	0.1315	-0.0897	-0.0027	0.0721	-0.0309	-0.0501	-0.0327	0.0249	0.0802
Ethnicity - Tongan	0.8576	0.0672	-0.0264	0.0361	0.054	-0.0276	-0.0888	0.0923	-0.0665	0.2334
Ethnicity - other	0.8497	-0.0515	0.0017	0.1348	0.1845	0.0992	0.0365	0.0177	0.0564	0.2084
Qualification level - missing	0.7179	0.3375	-0.0946	-0.2587	0.0261	-0.0818	0.324	0.0623	0.0286	0.1777
Household income - missing	0.7175	-0.085	0.2644	0.2188	0.1725	-0.0845	0.1902	0.0718	0.0684	0.2773
Birthplace - missing	0.5108	0.0797	-0.011	0.4977	0.2918	0.108	-0.0045	0.0722	0.1316	0.3655
Individual income - fourth quartile	-0.318	-0.8615	0.0693	-0.1707	-0.1351	-0.1504	0.0373	-0.0257	-0.0303	0.0789
Household income - first quartile	0.1312	0.8483	-0.0898	0.365	-0.0705	0.0539	-0.0859	0.0517	-0.0159	0.1037
Qualification level - high	-0.4575	-0.7922	0.1811	0.0833	0.0372	0.053	-0.0421	-0.0431	-0.0025	0.1156
Individual income - second quartile	-0.0863	0.7683	-0.3856	-0.1139	-0.0754	-0.0515	-0.0833	0.0134	0.0333	0.224
Household income - second quartile	-0.0769	0.7575	-0.2072	-0.4277	0.0343	0.1398	0.0355	-0.0359	0.024	0.1706
Ethnicity - English	-0.4388	-0.5969	-0.1172	-0.008	-0.2823	-0.069	-0.2755	0.0423	0.1408	0.2554
Ethnicity - Maori	0.2344	0.5497	-0.3006	0.1469	-0.1182	-0.0913	0.4916	0.0917	0.1392	0.2391
Birthplace - England	-0.4957	-0.5299	-0.0619	-0.0872	-0.3361	-0.2019	-0.3019	0.0443	0.0753	0.2096
Ethnicity - New Zealand European Maori	-0.1891	0.5239	-0.5153	-0.017	-0.2618	-0.0933	0.3699	0.0555	0.1009	0.1965
Birthplace - Australia	-0.4132	-0.5056	-0.3454	0.2358	-0.2951	-0.1322	0.0002	0.0007	0.0297	0.2932
Birthplace - China	-0.1277	-0.0938	0.9215	0.0467	0.1253	0.1279	0.0452	-0.0293	-0.0395	0.087
Ethnicity - Chinese	-0.1264	-0.1278	0.9189	0.0111	0.1323	0.0748	0.0453	-0.0274	-0.0498	0.0948
Birthplace - Korea	-0.1574	-0.0414	0.7516	0.1731	-0.0552	0.2947	0.0695	-0.0499	0.0681	0.2767
Recent migrant - yes	0.3507	-0.2641	0.7063	-0.218	0.3766	0.0396	-0.1488	0.0361	-0.0023	0.0941
Ethnicity - remainder	0.3422	-0.2021	0.6245	0.3057	0.1667	0.366	-0.1245	0.0023	0.0973	0.1717
Birthplace - South Africa	-0.192	-0.1291	0.5913	-0.3949	-0.1966	-0.1023	-0.1251	0.1066	0.1548	0.3408
Household income - third quartile	-0.2082	-0.0515	-0.0187	-0.8597	0.0907	0.1757	-0.0592	-0.0534	0.0138	0.169
Recent migrant - no	-0.132	-0.1859	0.4444	0.5113	0.3128	0.4905	-0.0731	-0.0414	0.0731	0.1383
Individual income - third quartile	0.2916	0.2727	-0.2408	-0.5106	0.278	0.3963	0.1276	-0.0283	0.156	0.2462
Ethnicity - Indian	0.1878	0.0155	0.1749	0.0038	0.939	0.117	-0.011	-0.0054	-0.018	0.0381
Birthplace - India	0.0075	0.0053	0.1985	0.1596	0.8719	0.2329	-0.0133	-0.0259	-0.0283	0.119
Birthplace - Fiji	0.4796	0.0812	0.0071	-0.1915	0.7212	-0.0334	-0.0068	0.0342	0.0226	0.2037
Birthplace - Philippines	-0.0076	0.0821	0.2152	-0.0245	0.129	0.8921	-0.0031	0.0223	-0.0203	0.133
Ethnicity - Filipino	0.0034	0.0788	0.2235	-0.0535	0.1276	0.8844	-0.0018	0.026	-0.0198	0.1415

Note: Values ≥ 0.75 or ≤ -0.75 are highlighted in green (for housing characteristics) or blue (for non-housing characteristics). We also highlight the maximum value in each row (using a lighter shade if the value is $\leq |0.75|$).

Table 2: Incidence and share of individuals living in high factor areas within urban areas

	Incidence			Share			
Urban area	**Factor1**	**Factor2**	**Factor3**	**Factor1**	**Factor2**	**Factor3**	**Proportion of population**
West Auckland	0.72	0.02	0.42	0.17	0.01	0.10	0.07
South Auckland	0.69	0.08	0.39	0.33	0.04	0.19	0.14
Wanganui	0.01	0.83	0.00	0.00	0.04	0.00	0.01
Gisborne	0.14	0.79	0.00	0.01	0.04	0.00	0.01
North Auckland	0.16	0.07	0.74	0.05	0.03	0.23	0.09
Central Auckland	0.44	0.03	0.59	0.21	0.02	0.28	0.13
Christchurch	0.08	0.27	0.18	0.03	0.13	0.07	0.11
Hamilton	0.14	0.34	0.26	0.03	0.08	0.05	0.06
Porirua	0.60	0.15	0.07	0.04	0.01	0.00	0.02
Lower Hutt	0.44	0.27	0.07	0.05	0.04	0.01	0.03
Wellington	0.18	0.01	0.08	0.04	0.00	0.02	0.06
Whangarei	0.05	0.74	0.01	0.00	0.05	0.00	0.02
Invercargill	0.00	0.74	0.03	0.00	0.05	0.00	0.01
Hastings	0.27	0.67	0.00	0.02	0.06	0.00	0.02
Napier	0.10	0.67	0.00	0.01	0.05	0.00	0.02
Rotorua	0.02	0.66	0.01	0.00	0.05	0.00	0.02
Te Awamutu	0.01	0.61	0.00	0.00	0.01	0.00	0.01
Blenheim	0.02	0.59	0.00	0.00	0.02	0.00	0.01
Palmerston North	0.08	0.55	0.12	0.01	0.06	0.01	0.02
Nelson	0.03	0.45	0.00	0.00	0.04	0.00	0.02
Dunedin	0.02	0.41	0.13	0.00	0.06	0.02	0.03
New Plymouth	0.00	0.39	0.02	0.00	0.03	0.00	0.02
Tauranga	0.02	0.30	0.06	0.00	0.05	0.01	0.04
Kapiti	0.00	0.25	0.00	0.00	0.01	0.00	0.01
Cambridge	0.00	0.16	0.08	0.00	0.00	0.00	0.01
Upper Hutt	0.10	0.14	0.03	0.00	0.01	0.00	0.01

Note: The top two values for each column are shaded in grey.

4.1 Figures & maps

Figure 1: Box plots of factor 1 scores by urban area. Urban areas are ordered by median value.

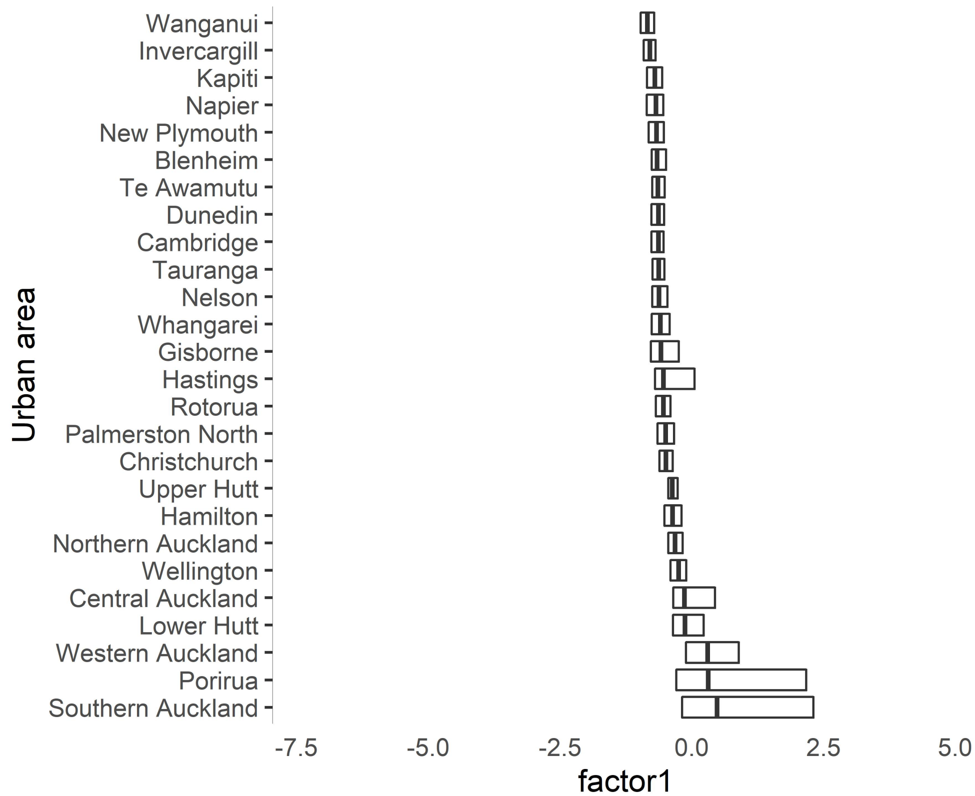
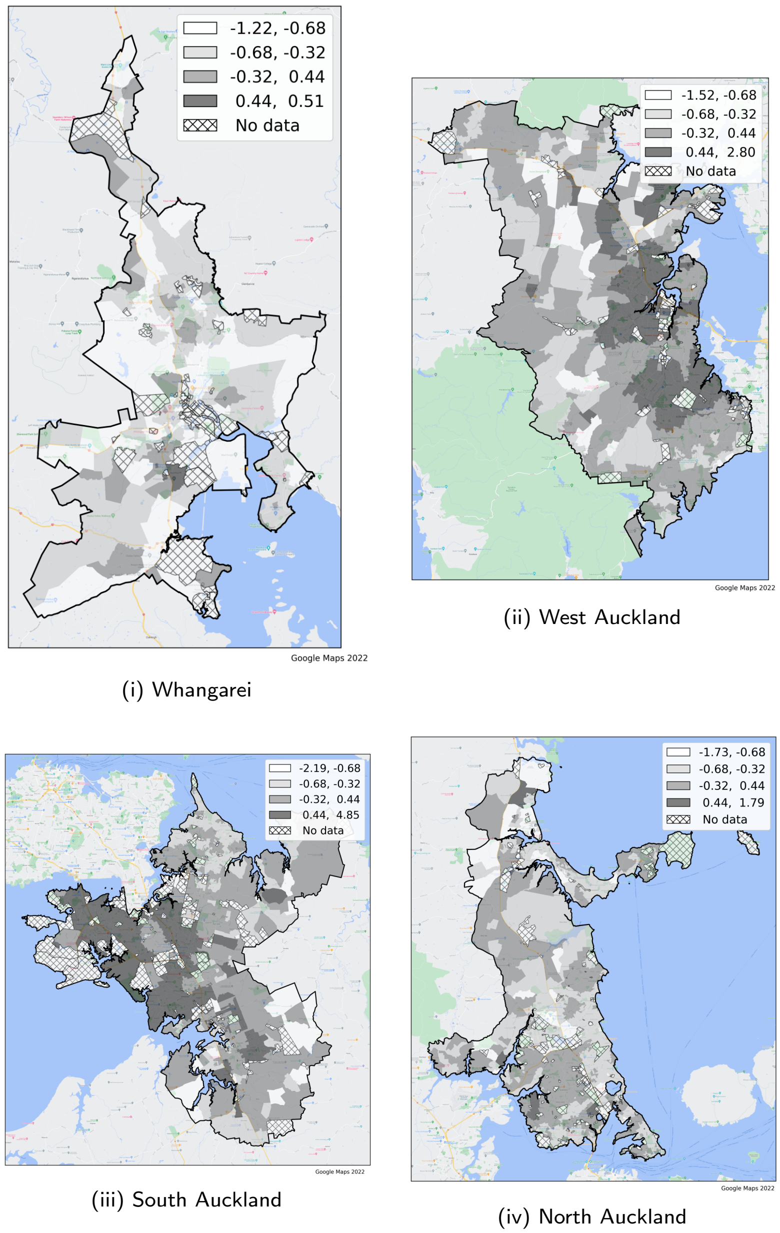


Figure 2: Maps of factor 1

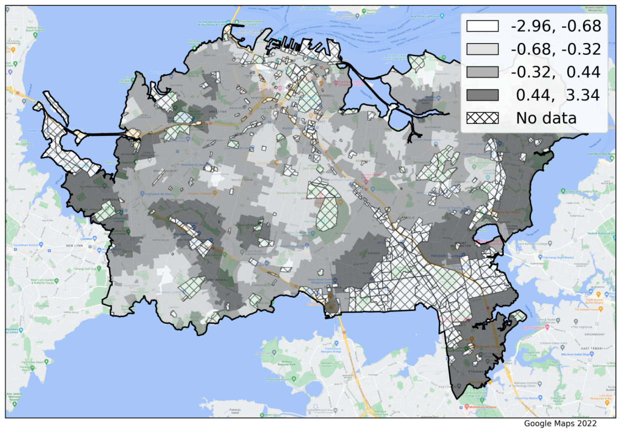


(i) Whangarei

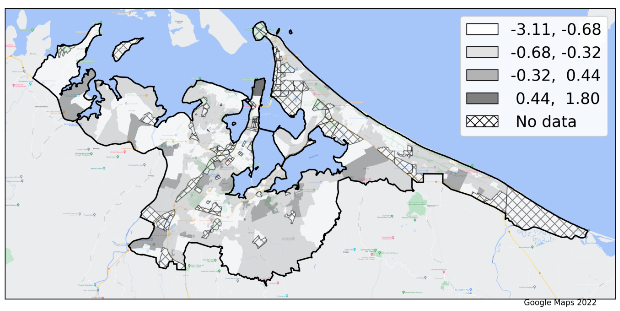
(ii) West Auckland

(iii) South Auckland

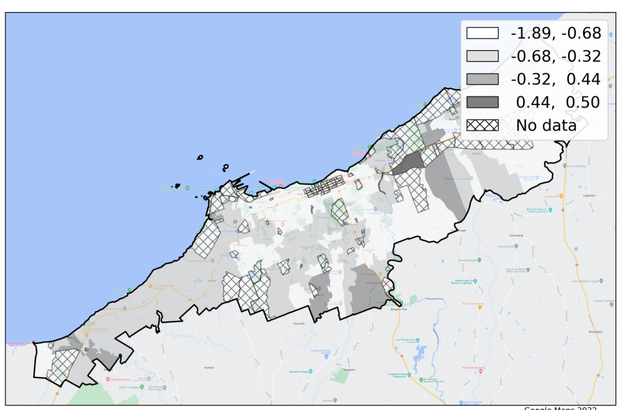
(iv) North Auckland



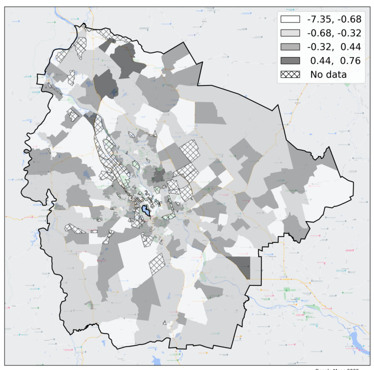
(v) Central Auckland



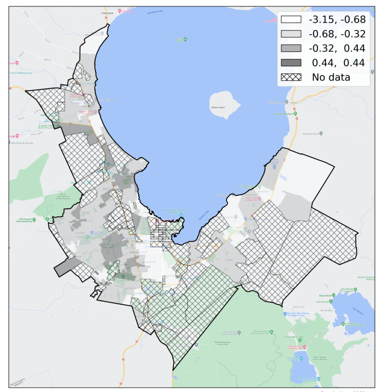
(vi) Tauranga



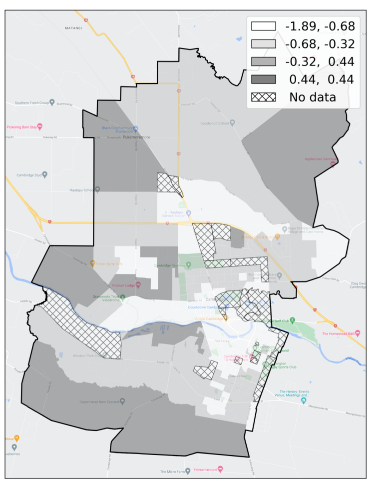
(vii) New Plymouth



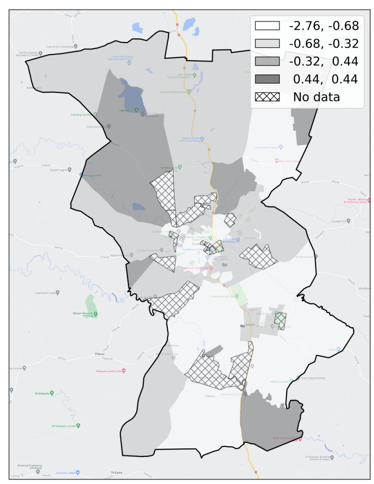
(viii) Hamilton



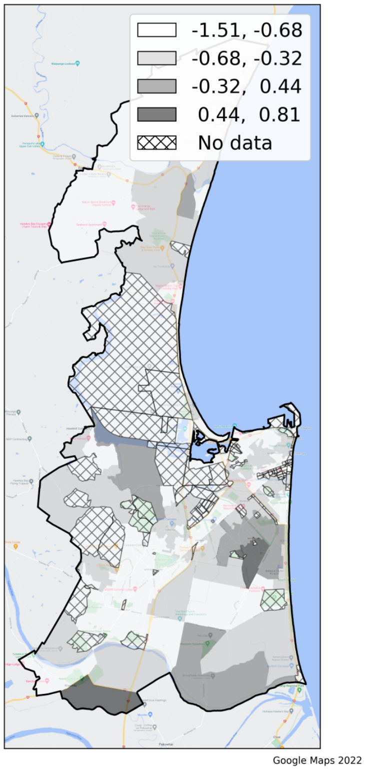
(ix) Rotorua



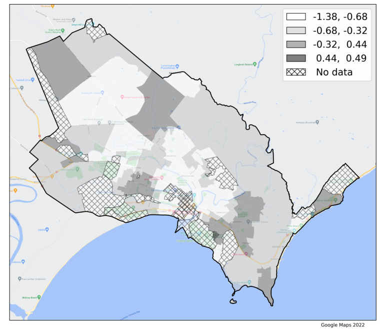
(x) Cambridge



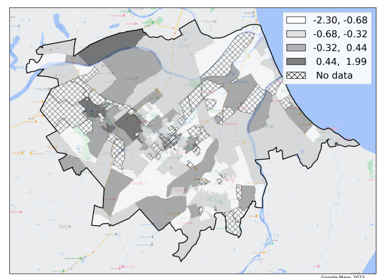
(xi) Te Awamutu



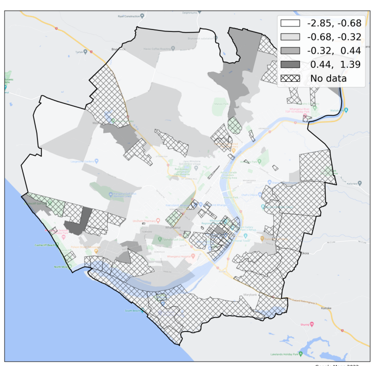
(xii) Napier



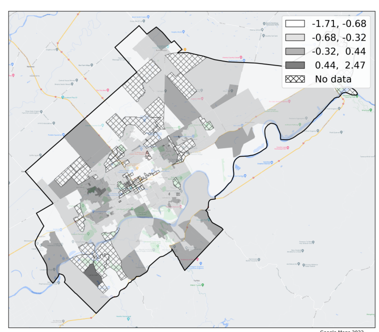
(xiii) Gisborne



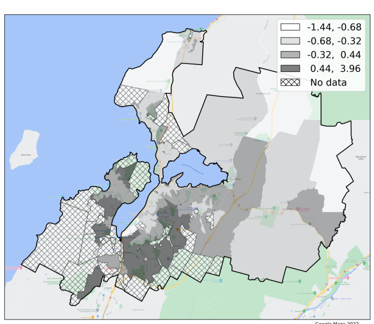
(xiv) Hastings



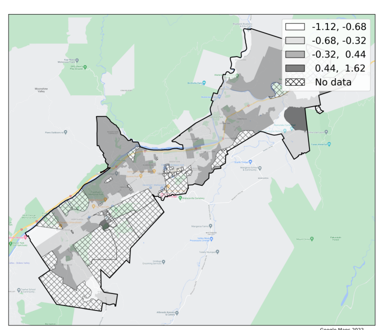
(xv) Wanganui



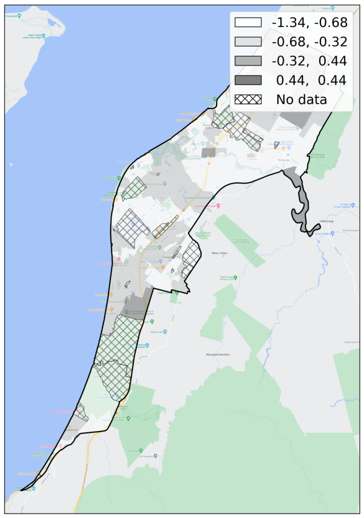
(xvi) Palmerston North



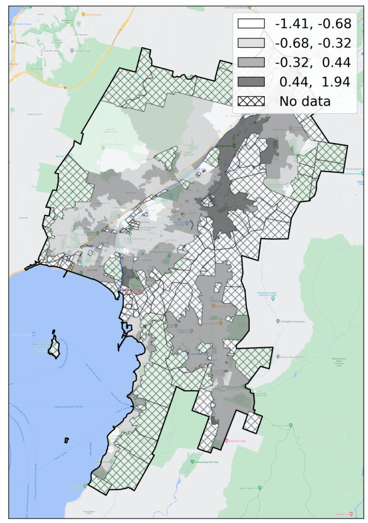
(xvii) Porirua



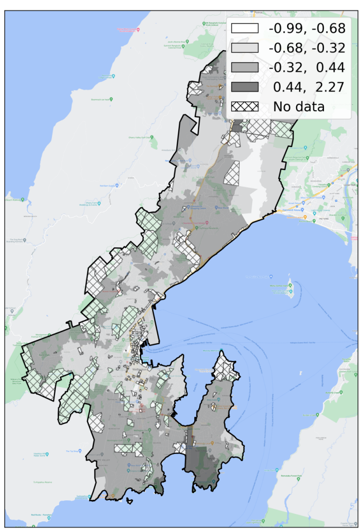
(xviii) Upper Hutt



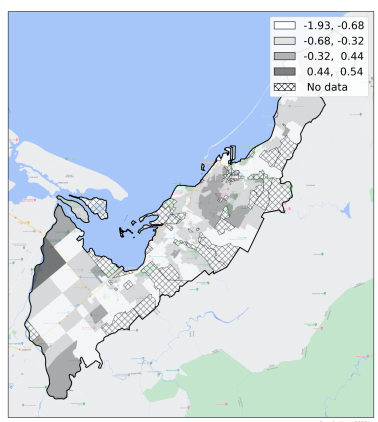
(xix) Kapiti



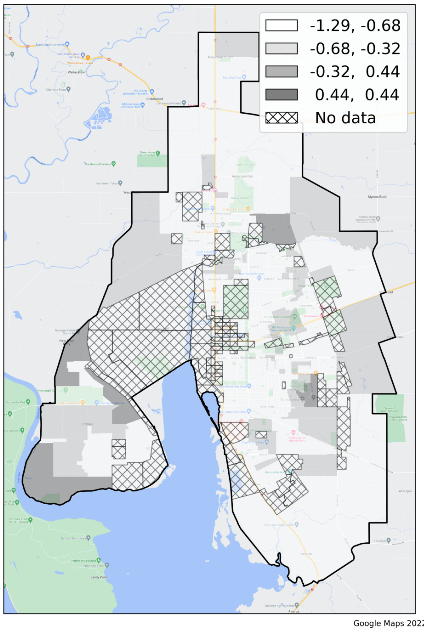
(xx) Lower Hutt



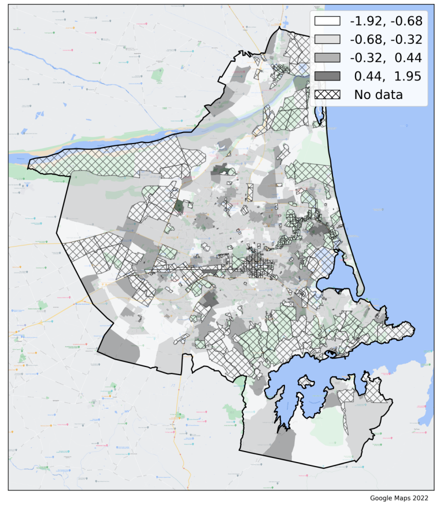
(xxi) Wellington



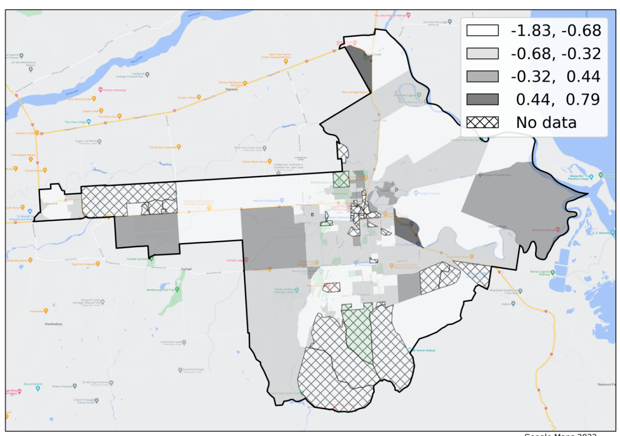
(xxii) Nelson



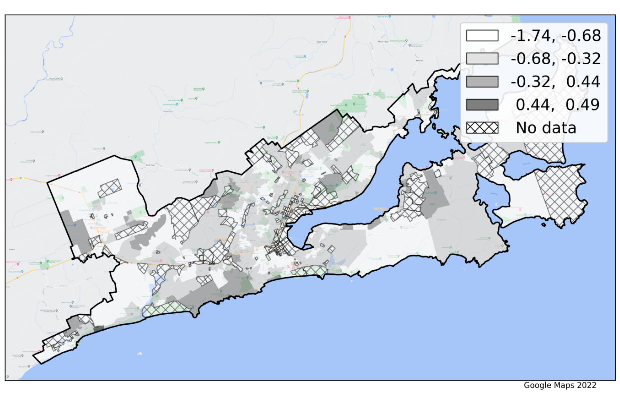
(xxiii) Invercargill



(xxiv) Christchurch



(xxv) Blenheim



(xxvi) Dunedin

Figure 3: Box plots of factor 2 scores by urban area. Urban areas are ordered by median value.

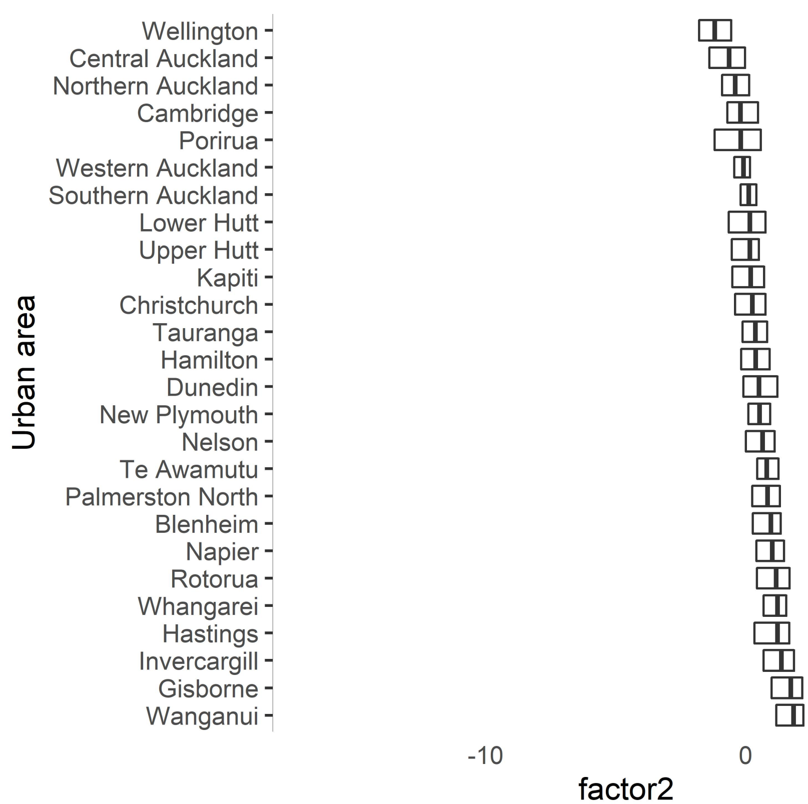
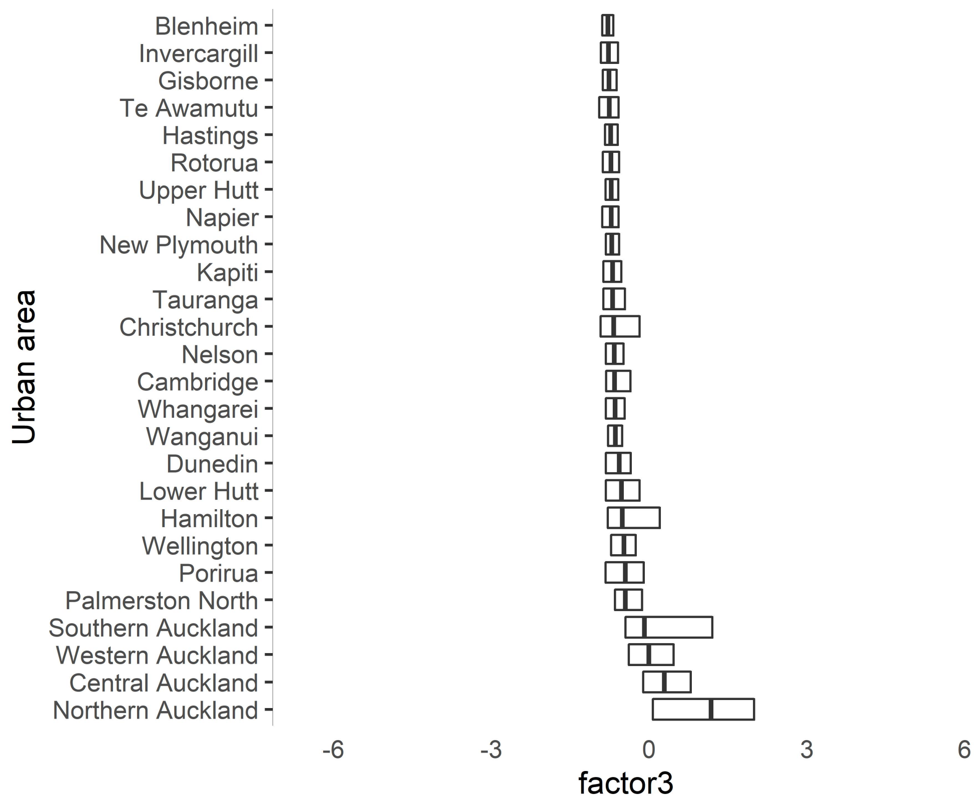


Figure 4: Box plots of factor 3 scores by urban area. Urban areas are ordered by median value.



5 Concluding discussion

The approach taken in this paper allows us to summarise patterns in housing; however, on their own, these patterns cannot pin down the existence, extent, or nature of discrimination. Our results identify three main patterns; we discuss these below along with key limitations of using housing patterns to infer the presence of housing discrimination.

The first pattern indicates that areas with a high proportion of individuals living in a crowded home and/or in a home owned by Housing New Zealand (now Kāinga Ora) correspond to areas that also have a high proportion of individuals who identify as Samoan or Tongan (Table 1). As mentioned above, this is a neigbourhood-level pattern and cannot be used to determine the co-occurrence of individual-level characteristics (*i.e.,* this pattern does not inform us whether individuals living in a crowded home also live in a home owned by Kāinga Ora or identify as Samoan or Tongan – rather this pattern informs us that at a neighbourhood level these characteristics tend to co-occur). Qualitative research is required to more fully understand this pattern. It is also important to note that our measure of household crowding is based on the Canadian Crowding Index, and may not capture cultural differences in living arrangements (*e.g.,* multi-generational households).

The second and third patterns capture income and ethnic sorting, respectively. These patterns do not correspond strongly to any of the housing characteristics in our dataset. Individual and household income can constrain housing choices leading to differences in housing outcomes across income levels. Although we find evidence of ethnic sorting (specifically, the tendency of individuals who were born in Korea or China, or who identify as Chinese to live in similar areas), we are unable to determine the pathways leading to this pattern. This pattern could reflect a preference to live near people belonging to the same ethnic group, or similar preferences for amenities (leading to people of a given ethnic group to live near certain amenities). Alternatively, ethnic sorting could reflect discriminatory practices that similarly influence members of a given ethnic group.

Housing patterns are shaped by a myriad of factors; inequalities in housing outcomes can arise via many pathways. Detecting the presence of discrimination in housing markets based on inequalities in housing outcomes (such as crowding or housing quality) requires disentangling discriminatory factors from non-discriminatory factors. Moreover, factors that shape housing decisions, such as income and education, may themselves be influenced by discrimination (*e.g.,* discrimination in employment could lead to inequalities in income which in turn could constrain housing choices). This advocates for a holistic approach capturing both direct and indirect pathways through which discrimination could be occurring.

The *WERO* research programme on the geographies of racism will provide more focused evidence on the role that racism has played and continues to play, either directly or indirectly, in the creation and perpetuation of unequal housing outcomes for different ethnic groups. It also aims to identify and highlight what types of actions or interventions would be most effective in combatting racism.

References

Christensen, Peter and Christopher Timmins, "Sorting or Steering: The Effects of Housing Discrimination on Neighborhood Choice," *Journal of Political Economy*, August 2022, *130* (8), 2110–2163. Publisher: The University of Chicago Press.

Cutler, David M, Edward L Glaeser, and Jacob L Vigdor, "The rise and decline of the American ghetto," *Journal of political economy*, 1999, *107* (3), 455–506.

King, A Thomas and Peter Mieszkowski, "Racial discrimination, segregation, and the price of housing," *Journal of political economy*, 1973, *81* (3), 590–606.

Lewis, Cassandra, Adele N Norris, Waimirirangi Heta-Cooper, and Juan Tauri, "Stigmatising gang narratives, housing, and the social policing of Māori women," *Neo-colonial injustice and the mass imprisonment of indigenous women*, 2020, pp. 13–33.

Norris, Adele N and Gauri Nandedkar, "Ethnicity, racism and housing: discourse analysis of New Zealand housing research," *Housing Studies*, 2022, *37* (8), 1331–1349.

Ondrich, Jan, Stephen Ross, and John Yinger, "Now you see it, now you don't: why do real estate agents withhold available houses from black customers?," *Review of Economics and Statistics*, 2003, *85* (4), 854–873.

Yeung, Stanley and Charles Crothers, "Patterns of perceived discrimination in New Zealand and their social contexts," *New Zealand Sociology*, 2016, *31* (7), 205–226.

Yinger, John, "Measuring racial discrimination with fair housing audits: Caught in the act," *The American Economic Review*, 1986, pp. 881–893.