CReAM | Centre for Research & Analysis of Migration

Discussion Paper Series 03/2023

Towards a Dynamic Spatial Microsimulation Model for Projecting Auckland's Spatial Distribution of Ethnic Groups

Mohana Mondal, Michael P. Cameron & Jacques Poot

February 2023

Centre for Research and Analysis of Migration

Department of Economics, University College London
Drayton House, 30 Gordon Street, London WC1H 0AX

cream@ucl.ac.uk

www.cream-migration.org

Towards a Dynamic Spatial Microsimulation Model for Projecting Auckland's Spatial Distribution of Ethnic Groups

Mohana Mondal

Digital Engineering NZ, WSP

Wellington, New Zealand

Email: mohanamondal92@gmail.com

Michael P. Cameron

School of Accounting, Finance and Economics

University of Waikato, New Zealand

and

Te Ngira: Institute for Population Research

University of Waikato, New Zealand

Email: mcam@waikato.ac.nz

Jacques Poot

Te Ngira: Institute for Population Research

University of Waikato, New Zealand

and

School of Accounting, Finance and Economics

University of Waikato, New Zealand

Email: jacques.poot@waikato.ac.nz

Towards a dynamic spatial microsimulation model for projecting Auckland's spatial distribution of ethnic groups

Abstract

In this paper we describe the development, calibration and validation of a dynamic spatial microsimulation model for projecting small area (area unit) ethnic populations in Auckland, New Zealand's most culturally diverse city, in which about 40 percent of the population is foreign born. The key elements of the microsimulation model are a module that projects residential mobility within Auckland and migration between Auckland and the rest of the world, and a module that projects mobility in ethnic identity. The model is developed and calibrated using data on 1996-2001 linked populations in the New Zealand Longitudinal Census (NZLC) 1981-2006. We compare the microsimulation results with the actual 2006 population in each area unit. We find that in terms of indexes of overall residential sorting and ethnic diversity, our projected values are very close to the actual values. At a more disaggregated spatial scale, the model performs well in terms of the simulated normalised entropy measure of ethnic diversity in area units, but performs less well in terms of projecting residential sorting for each individual ethnic group.

Keywords: dynamic microsimulation model, ethnic identity, location transition, ethnic transition

JEL Classification: C53, F22, J10, R23

14 February 2023

Disclaimer: The results in this paper are not official statistics. They have been created for research purposes from Census unit record data in the Statistics New Zealand Datalab. The opinions, findings, recommendations, and conclusions expressed in this paper are those of the authors, not Statistics NZ. Access to the anonymised data used in this study was provided by Statistics NZ under the security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business, or organisation, and the results in this paper have been confidentialised to protect these groups from identification and to keep their data safe. Careful consideration has been given to the privacy, security, and confidentiality issues associated with using unit record census data.

Acknowledgements: This paper is a revised version of Chapter 5 of the first author's PhD thesis *Ethnic Mobility and the Spatial Distribution of Ethnicity in Auckland*. We thank Rhema Vaithianathan and Robert Tanton for constructive comments.

Section 1: Introduction

The preferences of individuals regarding their residential location constitute an important topic of study because residential location of households is one of the key components of urban dynamics. The literature on residential sorting suggests that people choose where to locate based on a variety of factors (e.g. Duncan and Duncan 1955; Uyeki 1964; Schelling 1971). Patterns of residential sorting have been observed to be influenced by ethnicity and race (e.g. Schelling 1971; Ho and Bedford 2006; Johnston et al. 2011; Mondal et al. 2021b), educational qualification (e.g. Farley 1977; Denton and Massey 1988; Domina 2006), occupational status (e.g. Duncan and Duncan, 1955; Simkus 1978), and income (e.g. Fischer 2003). Clearly, a better understanding of urban population dynamics is needed to provide insight into what the future spatial distribution of a population might look like and to enhance thereby the efficiency and efficacy of planning for future public services and housing demands (Cameron and Poot 2019).

Our understanding of residential sorting, and its causes and impacts, remains relatively limited (Bruch and Mare 2006). Better understanding of changing residential sorting patterns requires examination at different spatial levels, as different geographic scales portray different dimensions of residential sorting (Reardon et al. 2009). Urban households are likely to take current and anticipated spatial features that are apparent at different spatial scales into account when deciding on their residential location. Yet most of the research on the dynamics of individual transitions and residential sorting looks either backwards in time or focuses just on the present (Rees et al. 2017).

Ethnic diversity is an important contributor to residential sorting. Schelling (1971) noted that individuals prefer to stay in close contact with people with whom they share similar preferences, which may inter alia lead to people clustering together with others of the same ethnicity. Residential sorting may also occur in terms of other characteristics such as education, income, or occupation. However, in Auckland, New Zealand – the city this paper focuses on –residential sorting of the population is stronger in terms of the self-identified ethnicity of individuals than in terms of their economic characteristics (Mondal et al. 2021b). In this context, Auckland provides an important case study of residential sorting given that this city, with a population of 1.6 million (one third of the population of New Zealand), is one of the most culturally diverse cities in the world and also the most diverse city in New Zealand (Maré and Poot 2022; Mondal et al. 2021b).

Projections of ethnic diversity in a city require assessing the ethnic composition of the population at the neighbourhood level (O'Sullivan 2009). This makes the task of projecting ethnic populations more difficult. The data requirements for small-area projections are high, and the methods are currently under-developed (Cameron and Cochrane 2017). In this paper, we describe and evaluate a microsimulation model (MSM) of the population of the Auckland region that captures ethnic diversity at a fine spatial scale, namely that of census area units, and with the maximum feasible disaggregation of ethnic groups. The model is constructed with microdata from the 1981-2006 New Zealand Longitudinal Census (Didham et al. 2014), yielding 1996-2001 longitudinal data on ethnicity-specific populations along with their ethnic and spatial mobility. We test our model by comparing our simulated results to the actual 2006 census data.

This work represents the first attempt to develop a dynamic spatial MSM to project the future ethnic spatial distribution at a fine spatial scale in New Zealand. The model uses a greater level of disaggregation of ethnicity than was done in previous studies in New Zealand, but also in many other countries. This way we aim to capture better the heterogeneity that exists within the broad ethnic groups, in terms of preferences and choices (Mondal et al. 2021b). We develop and run our model in Stata, which is in itself a novel approach to dynamic spatial microsimulation modelling. The Stata statistical software is available inside the secured Statistics New Zealand Datalab. Hence we can run our model in the Datalab with the original microdata rather than first having to generate a sample of anonymized synthetic unit record data that can be taken out of the Datalab. Using the original microdata avoids any potential bias that might result from creating a synthetic base population. Moreover, our approach allows us to use the entire Auckland population that could be linked in the 1996 and 2001 censuses as our base population, rather than just a sample of the population.

The remainder of the paper is organised as follows. Section 2 reviews different types of MSMs and how they have been used in previous research. Sections 3 and 4 describe the data and the methods we employ respectively. Section 5 describes the results and the testing of the MSM model. Section 6 concludes.

Section 2: Literature Review

Microsimulation is a methodology to model outcomes at the micro level. The outcomes can be about people (e.g. Mot 1992), households (e.g. Rogers et al. 2014), or firms (e.g. Moeckel, 2009). Microsimulation has become increasingly popular in recent decades as ever-increasing computing power enables a growing range of applications developed by means of rich microdata (Li and O'Donoghue 2013). Among the many applications possible, a MSM can be used to simulate and project populations and their attributes. Simulation can be interpreted here as the process by which attributes are assigned to individual units (Lomax and Smith 2017), informed by unit record data. The base population of a MSM can come either from a survey or can be synthesised from various data sources (Zaidi and Rake 2001). MSMs have previously been used for tax-benefit analysis (Lambert et al. 1994; Spielauer 2011), projecting future socio-economic development trends under current (or forecast) policies (Favreault and Smith 2004; Harding 2007), modelling lifetime earnings distributions (Smith et al. 2007; Holmer et al. 2014), and in studies of wealth accumulation (Caldwell et al. 1998). MSMs have also been used to assess the future performance and sustainability of long-term public programmes such as pensions, healthcare, and educational financing (Goldman et al. 2009; Rowe and Wolfson 2000; Wolfson and Rowe 2013).

2.1 Types of MSMs

All MSMs require micro data (Wu et al. 2011), but differ in terms of the overall setup of the model (static or dynamic), the estimation of transition probabilities, exclusion or inclusion of behavioural responses of the micro-units (arithmetical or behavioural), treatment of time (discrete/continuous), and whether they are explicitly spatial or not.

Static MSMs usually take a cross-section of the population at a specific point in time, and measure the immediate effects of policy changes without modelling any of the specific processes that result in changes over time (Lambert et al. 1994; Spielauer 2011). This type of MSM has been mainly used to evaluate tax-benefit systems (Pechmen and Okner 1974) or to analyse the redistribution impacts of reforming existing tax systems (Paulus et al. 2009). For example, Immervoll et al. (2007) used a static MSM to estimate changes in marginal and participation tax rates in response to increasing traditional welfare and the introduction of in-work benefits in 15 countries of the European Union in 1998.[1] Eggink et al. (2016) used a static MSM to forecast the use of publicly funded long-term elderly care in the Netherlands from 2008 to 2030.

In contrast, *dynamic* MSMs are able to simulate changes over time for a population, by 'ageing' unit records based on the probabilities of numerous real-life events occurring. This type of model can therefore estimate the effects of policies separately for the long term and the short term (Lomax and Smith 2017). For example, Favreault and Smith (2004) designed DYNASIM3 (Dynamic Simulation of Income Model III) in order to analyse the long-term distributional consequences of retirement and ageing from 1992 to 2040 in the US. In the UK, PENSIM is a national dynamic microsimulation model designed to study the impact of policy changes on the income distribution of pensioners. This model follows 1935-1985 birth cohorts up to 2030 (Holmer et al. 2014; Hancock et al. 1992).

Dynamic MSMs can be probabilistically dynamic or implicitly dynamic. *Probabilistically dynamic* MSMs use event probabilities to project the characteristics of each unit record in the simulated database into the future. The event probabilities (or transition probabilities) are probabilities that govern the change in the variables studied from one time period to the next. For example, Ballas et al. (2005a) used a probabilistic model to project population change from 1991 until 1996 and between 1996 and 2002 at the District Electoral Division (DED) level in Ireland. Probabilistically dynamic MSMs require modellers to undertake the difficult task of determining the interdependencies between individual attributes and events, and so they require high quality suitable data, which are seldom available (Ballas et al. 2005b). In contrast, *implicitly dynamic* MSMs use independent small area projections and apply static simulation techniques to create small area microdata. For example, Ballas et al. (2005b) used data from the 1971, 1981 and 1991 British population censuses to estimate small area data for 2001, 2011 and 2021 in Wales. They then used these estimates, in combination with national survey data, to simulate future trends in car ownership, demography, and employment at the small area level.

Arithmetical MSMs are generally used to simulate distributional effects in response to changes in taxes, benefits and wages. This type of model takes as constant the individual's behavioural responses to the policy change being examined, i.e. the individual's behavioural responses to

[1] Participation tax rates are the difference between current household taxes and benefits and the household taxes and benefits when individual earnings are set to zero, divided by individual earnings (Immervoll et al. 2007).

the policies are not included in the model (Bourguignon and Spadaro 2006). Hence, any behavioural responses are considered exogenous, i.e. determined outside the model. Arithmetical models have been used to examine indirect taxes and tax reforms (Sahn and Younger 2003; Creedy 1999), to estimate incidence of public spending in health and education (Demery 2003), and also to compare fiscal policy effects (Atkinson et al. 1988, 2002; Callan and Sutherland 1997). For example, Atkinson et al. (1988) analysed the effect of replacing the French tax-benefit system with that of the British, for a given sample of French households.

In contrast, *behavioural* MSMs explicitly consider the changes in the behaviour of individuals in response to policy changes. These models are based on economic theory and may be policy-specific (Creedy and Duncan 2002). Behavioural MSMs have been used to evaluate the effects of direct tax reforms (Bonin et al. 2002; Das and van Soest 2001; Blundell et al. 2000) as well as indirect tax reforms (Kaplanoglou and Newbery 2003; Liberati 2001; Creedy 1999). The main advantages of behavioural MSMs are the ability to account for the heterogeneity within the population of interest, and the identification of both the mean and the distributional impact of a reform. However, these models require the estimation of a policy-specific behavioural model and they are often not generalizable to the evaluation of other policies (Zucchelli et al. 2010).

Dynamic MSMs can be represented in discrete or continuous time. In the case of *discrete-time dynamic* MSMs, each individual's characteristics are simulated at fixed time intervals. These models usually include a transition probability matrix for the simulations (Willekens 2006). In New Zealand, Milne et al. (2015) developed a discrete-time dynamic MSM that modelled child development from birth to age 13, focusing on factors that influence health service use, early literacy and conduct problems of children. They used 2006 New Zealand Census data and three New Zealand child cohort studies to build their model and transition probability estimates. [2]

Continuous-time dynamic MSMs treat time as continuous and are therefore able to estimate the time at which each event occurs. In these models, individuals are assigned characteristics that can change at any time. The continuous-time dynamic MSMs use survival functions to model the length of time that an individual will remain in his/her current state, and to simulate the timing of events (Willekens 2006). Although these models have theoretical advantages, they have higher data requirements than discrete time MSMs (Zaidi and Rake 2001). In Canada,

[2] These studies are the Christchurch Health and Development Study, the Dunedin Multidisciplinary Health and Development Study, and the Pacific Islands Families Study.

Rowe and Wolfson (2000) used a dynamic continuous-time MSM called 'LifePaths' to model health care treatment, student loans and public pensions. Their analysis started with the cohort born in 1892 and extended for two centuries. In Australia, DYNAMOD is a continuous-time dynamic MSM developed by the National Centre for Social and Economic Modelling (NATSEM), and was designed to project population characteristics and the implications of policy changes over a 50-year period (King et al. 1999).

A dynamic MSM can be classified as either open or closed, based on whether new individuals are introduced to the base population as the simulation progresses, or not. In an *open* MSM such as LifePaths in Canada, new individuals are generated if an individual in the initial population is selected to form a marital union. This differs from a *closed* MSM, such as DYNACAN in Canada, which generates a new unit only when a baby is born (Zaidi and Rake 2001), or not at all.

MSMs can also be non-spatial or spatial in nature. *Dynamic spatial* MSMs are used to project the *geographical* trends in socio-economic activities. For example, the SVERIGE model (Rephann 2004, Vencatasawmy et al. 1999) was the first national-level dynamic spatial MSM, and was developed from longitudinal socio-economic information on every resident in Sweden from 1985 until 1995. The model was used to study the spatial consequences of public policies at different geographical levels (national, regional and local). The model included specific events in a person's life, generated through deterministic models of behaviours that are functions of individual, household and regional socio-economic characteristics. Holm et al. (2002) studied population composition change in Sweden by simulating the development of all individuals in Sweden with respect to variations in demographic processes such as mortality, fertility and immigration using a dynamic spatial MSM. Their model was executed for 110 years (1990-2100).

Finally, MSMs differ in terms of how the base population is created. Some MSMs use census or survey data to form a base population. Census data do not always provide all of the variables necessary for analysis, so data may also be obtained from multiple alternative sources, generated for diverse purposes that are not always directly compatible. In these cases, a *synthetic population* that closely represents the actual population is created to be the base population in the MSM (Zaidi and Rake 2001). The synthetic unit records may be generated using existing datasets and a variety of techniques like iterative proportional fitting, linear programming, or complex combinatorial optimisation methods (Ballas 2001, Ballas and Clarke

2000; Williamson et al. 1998). For example, DYNACAN in Canada, DYNAMOD 2 in Australia, and PENSIM in the UK all use census or survey unit records as the base population, whereas NEDYMAS in The Netherlands and LifePaths in Canada uses a synthetic database of unit records created using the census and other data sources (Li and O'Donoghue 2013).

2.2 Previous MSMs Projecting Ethnic Population Change

Dynamic MSMs have been used previously to project the future ethnic composition of the population of several countries. For example, Demosim is a dynamic spatial MSM developed and maintained by Statistics Canada, which has been used to project the Canadian ethno-cultural population composition. Demosim produces dynamic population projections at various spatial levels that include provinces, territories, census metropolitan areas, and smaller geographical areas, based on individual demographic characteristics including age, sex, and place of birth (Statistics Canada 2018). Malenfant et al. (2015) used the Demosim model to provide insight into the ethno-cultural makeup of the Canadian population in 2031 at different spatial scales. Taking 20 percent of the 2006 Canadian census as the base population, they calculated transition probabilities for mortality, immigration, internal migration, emigration, and highest level of schooling. They found that there would be a significant increase in ethno-cultural diversity over time, both within the Canadian-born and the foreign-born populations, especially in certain metropolitan areas, such as Toronto and Vancouver.

Davis and Lay-Yee (2019) built a dynamic MSM (SociaLab) to simulate societal change in New Zealand from 1981 to 2038. They worked with linked microdata from the New Zealand Longitudinal Census that covers 1981 until 2006, to build, calibrate, and validate their model. They considered individual demographic characteristics like age, sex, place of birth, religion, and ethnicity as predictor variables. They used four broad ethnic groups (Māori, Pacific, Asian and NZ European/Other), considering them as time-invariant (i.e. each individual's ethnicity was assumed to remain constant throughout the simulation). The results from their model show that from 2006 to 2038, New Zealand will be ageing and becoming more ethnically diverse, which continues the observed trend over the past several decades.[3] Also, changing patterns in living arrangements, such as households shifting away from the nuclear family, were projected to continue.

[3] See also Mondal et al. 2021b, who show similar past trends for Auckland.

In the study most closely related to ours, Ardestani (2013) built a hybrid geosimulation model (a combination of an agent-based model and a microsimulation model) to investigate residential segregation in Auckland, New Zealand over the period 1991 to 2006. The author used New Zealand Census data to inform, calibrate and validate the model, and examined the changes in ethnic residential segregation for four major ethnic groups (New Zealand European, Māori, Pacific, and Asian). He took into account the link between micro level (individual preferences) and macro-level (number of groups, group size, and proportion) elements to model and predict (until 2021) the changing ethnic residential patterns within the Greater Auckland Urban Area at both meso (territorial authorities) [4] and macro levels (the entire Auckland urban area). He simulated several scenarios based on different assumptions about population growth, mobility rates of each ethnic group, housing vacancy rates, and freedom of movement (as a proxy for income). Ethnic population was projected to be consistently clustered over time in all of the area units in the Auckland urban area. Results also showed that the number of area units with a majority of Asian and Māori population will increase in the future in all of the territorial authorities they studied. In the Waitakere area, there would be several area units where the Pacific people were projected to be the largest group. It was also projected that in the Manukau area there would be an absolute decline in the New Zealand European population.

In a follow-up study, Ardestani et al. (2018) used a multi-scaled agent based model to simulate the relocation of residents in the five central territorial authorities (TAs) of the Auckland urban area. The aim was to study the dynamics of residential segregation. They focused again on the four major ethnic groups, and found that a high fertility and high migration scenario leads to lesser levels of residential segregation than a low fertility and low migration scenario. They also found that, in the low fertility and migration scenario, residential segregation observed across the whole Auckland urban area was less than the residential segregation observed separately in some of the TAs (e.g. Manukau). They also looked into the impact of housing vacancy rates on the dynamics of residential segregation, and found that a reduction in housing vacancy rates leads to higher degrees of residential sorting at both the territorial authority and metropolitan area scales.

As noted earlier, studies relating to the spatial ethnic distribution of future population at the local level have been rare, both globally and in New Zealand. With respect to New Zealand,

[4] The territorial authorities considered were Auckland City, Manukau, North Shore, Waitakere, and Papakura.

Ardestani (2013) and Ardestani et al. (2018) did not investigate the residential segregation patterns at the area unit level, and focused only on four broad ethnic groups. This overlooks the diversity *within* these ethnic groups (especially within the Asian and Pacific Peoples ethnic groups, Mondal et al. 2021a). Additionally, these studies did not consider inter-ethnic mobility (changes in ethnic affiliation over time), which plays an important role in social change and is an increasingly popular and important area of research both internationally and in New Zealand (Didham 2016; Carter et al. 2009). Our model extends this earlier work, and addresses these shortcomings to some extent.

Section 3: Data

The most recent population census in New Zealand was in 2018 and recorded a usually resident population of 4.7 million. Auckland is the most ethnically diverse metropolitan region in New Zealand and accounts for about one-third of the New Zealand population (Maré and Poot, 2022; Mondal et al. 2021b). The major ethnic groups present in Auckland in the 2018 Census were: European (53.5 percent), Asian (28.2 percent), Pacific Peoples (15.5 percent), Māori (11.5 percent), MELAA[5] (2.3 percent), and Other (1.1 percent) (Statistics New Zealand 2020a).[6] Because of its high ethnic diversity and relatively large population, we focus on Auckland for this microsimulation research.

We use data for the Auckland region from the 1996-2001 linked populations in the New Zealand Longitudinal Census 1981-2006 (NZLC) (Didham et al. 2014).[7] The longitudinal census links individual records across pairs of censuses in a deterministic way (for example, an individual with age a in census year t who declared to have not changed address during the intercensal period is the same person as the individual of age a–5 in census year t–5 at that address). Throughout this paper we use 'previous' to refer to data from the first census in each inter-censal period and 'current' for data from the following census. The link rate for individuals from the 1996 Census to the 2001 Census was 69.5 percent, and for the 2001 Census

[5] Middle Eastern/Latin American/African.

[6] Percentages do not sum to 100 percent, as people can report more than one ethnicity.

[7] The 2018 census has not yet been integrated into the NZLC dataset. Work has been undertaken to link data from the 2013 Census to the 2006 Census (Kang, 2017), but these data were unavailable at the time of writing.

to the 2006 Census was 70.3 percent (Didham et al. 2014).[8] The NZLC is the most comprehensive source of longitudinal socio-demographic information on individuals (e.g. sex, age, ethnicity, education, place of residence, etc.) in New Zealand. Our analysis is based on unit record data aggregated to the area unit level, using 2013 Auckland area unit boundaries.[9] In 2013, the Auckland region was comprised of 413 land-based area units, of which 409 had a non-zero usually resident population. We dropped area units with no usually resident population. The unit record data were accessed within Statistics New Zealand's secure data laboratory, to meet the confidentiality and security rules of the Statistics Act 1975.[10]

In New Zealand, ethnicity captures the ethnic group(s) that people feel a sense of belonging to. It is not a measure of race, ancestry, nationality or citizenship, but a measure of cultural affiliation. Ethnicity is self-recognised and declared. Individuals can identify with up to six ethnic groups in the census.[11] Individuals are able to choose one or more ethnicities in each census different from any they had chosen previously (Statistics New Zealand 2015).

The New Zealand Standard Classification of Ethnicity categorises ethnicity into four levels (Statistics New Zealand, 2013). The Level 1 classification of ethnicity has six categories and Level 2 has 21, which are shown in Table 1. The Level 1 ethnic groups are very broad and potentially mask heterogeneity in the characteristics of the ethnic groups, particularly for the Asian and the Pacific ethnic groups (Mondal et al. 2021a). Hence, we use Level 2 ethnic groups to better capture this heterogeneity. There are a non-negligible number of individuals among those who are European, Asian or Pacific Peoples, who were coded as belonging to the 'Not further defined' group or the 'Other' group. We combined these two groups for each of those three ethnicities. Hence we have 18 rather than 21 ethnic groups in the analysis. We do not use finer Level 3 ethnic groups as the group sizes are too small for some groups to develop a suitable model.

[Insert Table 1 here]

[8] The link rate for the censuses of 2006 and 2013 is unavailable. A census pair '*t*-5, *t*' refers to a pair of censuses where individual records in census *t* are linked to those of the previous census *t*-5. For example, if we are looking at linking records from the 1996 Census to those from the 1991 Census, we refer to this as the 1991–1996 census pair (Didham et al. 2014).
[9] Area units are non-administrative aggregations of adjacent meshblocks with common boundaries (Statistics New Zealand 2013). An area unit is approximately the size of a suburb in urban areas.
[10] As stated in the *Disclaimer* at the start of this paper.
[11] Individuals could choose up to three ethnic groups until 1996, which was increased to six in later censuses.

Two issues affect the use of ethnicity data. First, the format and wordings of the Census ethnicity question have been inconsistent between censuses. For instance, the ethnicity question in 2001 differed substantially from that in 1996.[12] These inconsistencies affect particularly the European ethnic groups (including New Zealand European) and the Māori ethnic group. In the 1996 data, the count for 'Other Europeans' was much higher than in the 2001 data. This was because the difference in format of the ethnicity question resulted in increased multiple responses, and a consequent reduction in single responses. This also resulted in some respondents answering the 1996 question on the basis of ancestry rather than ethnicity. The count for the New Zealand European category was much lower in 1996 than in 2001, which can be attributed to the fact that in 1996, people saw the additional 'Other European' category as being more suitable to describe their ethnicity than the 'New Zealand European' category (Statistics New Zealand 2017).

Second, there has also been inconsistency in the treatment of responses of 'New Zealander' to the Census ethnicity question. The standard for ethnicity statistics was developed in 2005. Previously, the 'New Zealander' response was included in the 'European' category, and was later moved to the 'Other ethnicity' category (Statistics New Zealand 2007a). New Zealand Europeans were the most likely group to be calling themselves 'New Zealander' in the census (Statistics New Zealand 2007b; Brown and Gray 2009). This resulted in an increase in the 'Other ethnicity' category, and a consequent reduction in the size and proportion of people reporting as being 'European' or 'New Zealand European'. 'New Zealander' was included explicitly as a new category in 2006, but not in 2001 or earlier. In 2001, individuals considering themselves to be a 'New Zealander' were likely to have been counted in the 'New Zealand European' ethnic category (Statistics New Zealand 2017).

Our model incorporates inter-censal migration flows. This requires that we observe the location of each individual in two successive censuses. That is problematic in the case of emigration (from Auckland to overseas), and deaths, as in both cases the individual is not observed in the second of each pair of linked censuses. To overcome this issue, we apportioned the number of

[12] In the 1996 Census, the ethnicity question had a different format compared to that used in 1991 and 2001. In 1996, there was an option to choose Other European with additional drop down answer boxes for English, Dutch, Australian, Scottish, Irish and Other. These options were absent in the 1991 and 2001 Censuses. Moreover, the first two answer boxes appeared in a different order in 1996 from that in 1991 and 2001. In 1996, 'NZ Māori' was listed first and 'NZ European or Pākehā' was listed second. The 1991 and 2001 questions only used the words 'New Zealand European' rather than 'NZ European or Pākehā' (Pākehā is the Māori word referring to a person of European descent). The 2001 question used the word 'Māori' rather than 'NZ Māori' (Statistics New Zealand, 2017).

emigrants from Auckland and the number of deaths in Auckland to each area unit according to area unit population.[13] For in-migration (from overseas or from elsewhere in New Zealand to Auckland) and births, we identified those individuals who were not present in the previous census in Auckland but present in Auckland in the current census. We use the census characteristics of these individuals. Thus, our model accounts for both population inflow into Auckland (due to births and inward migration) and population outflow (due to deaths and outward migration), but the inflows and outflows are not split into the contributions from migration and natural change.[14]

Section 4: Methodology

In this section, we describe the construction and calibration of a dynamic spatial MSM which can be used to project the future spatial patterns of ethnic diversity in Auckland, taking both ethnic and spatial mobility into consideration. Our model is a *discrete-time* (runs in five-year time steps) *probabilistic* (uses transitional probabilities to project forward) *dynamic* (includes time-varying parameters) and *spatial* (assigns an area unit of residence to each individual) MSM. Our model is also an *open* MSM as, in addition to people moving between area units within Auckland, it allows individuals to move out of Auckland (out-migration) as well as move into Auckland from other areas in New Zealand and from other countries (in-migration).

The MSM model we describe here is a *validation model*, which uses linked 1996-2001 data from the 1986-2006 NZLC to simulate and project the population in 2006, which is then validated against actual 2006 census data. This model can then be used to develop a *projection model* that will simulate and project the population in subsequent census years. However, projecting area unit populations after 2006 is beyond the scope of the present paper. The validation model is comprised of two modules: (1) an ethnic transition module; and (2) a locational transition module. For each of these two modules, we break the population into two age groups: (1) children/adolescents (0-17 years); and (2) adults (18 years and older).

The MSM captures individual ethnic transitions as well as spatial mobility i.e. individuals making choices regarding their ethnicity and location. Figures 1 and 2 outline the theoretical framework for the ethnic transition and locational transition modules respectively. In practice,

[13] Total emigration was calculated as a residual of 1996-2001 Auckland population change after accounting for recorded births, deaths and internal migration.
[14] Intercensal births can of course only affect the age group 0-4 in the current census.

the ethnic transition module runs first in each time step, followed by the locational transition module.

[Insert Figure 1 here]

[Insert Figure 2 here]

Table 2 summarises the variables used in the analysis. The ethnic transition module runs a separate logistic regression equation for each ethnicity. We take the individual's ethnic response, which is binary (1 = belongs to the ethnic group *i*, 0 = otherwise), in the current census as the dependent variable. This variable represents whether or not the individual identifies with that group, regardless of whether they also identify with one or more other groups. This substantially simplifies the analysis relative to a multinomial logit specification, which would require that every possible combination of ethnic affiliations be an option (Mondal et al. 2020).

An individual's ethnicity is in our model an 18x1 row vector of binary variables, with one binary variable for each of the 18 ethnic groups *i*. Our approach allows us to include multiple ethnic affiliation for individuals without requiring an order of priority for the determination of the ethnic choices, i.e. each individual's choice in regards to each ethnicity is given equal importance. From the logistic regression equations, we obtain the predicted probabilities of an individual belonging to ethnic group *i* in the current census. We then assign uniformly distributed random variables (over the interval 0 and 1) to each individual. Comparing the predicted probabilities with the random variables, the model determines whether the individual identifies with any of the possible ethnicities in the projected year or not.

The individual-level determinants of ethnicity in the ethnic transition module are the individual's ethnicity or ethnicities in the previous census, their age, sex and whether they were born in New Zealand. Neighbourhood level variables are the ethnic diversity and the percentage share of the different ethnic groups in the area unit they reside in. All independent variables in the logistic regressions were observed at the start of each inter-censal period.

The location transition module proceeds in two stages, following Willekens' (2016) migrant pool model for projecting migration. In the first stage, the number of out-migrants (i.e. people who change their usual residence) is projected. Specifically, we first use logistic regression equations (with separate coefficients for adults and children) to obtain predicted probabilities of moving for each individual in the current census. Similar to our ethnic transition model, we assign a uniformly distributed random variable to each individual. Then, comparing the values of the random variable and the predicted probabilities, the model determines whether the person is a mover or not in the current year.

In the second stage, the people who changed their location are then distributed over possible destinations using a distribution function that is solely dependent on the destination but not on the origin. In this step, movers are allocated to destination area units based on a column-standardised origin-destination matrix (with a zero diagonal) calculated using the intra-urban relocation data from the actual 1996-2001 linked census. A different origin-destination matrix is used for each ethnic group. For individuals with multiple ethnicities, one of their ethnicities is chosen at random, and the corresponding origin-destination matrix is used. The destination for each migrant is determined again using a uniformly distributed random variable, with the appropriate column of the origin-destination matrix used as a lookup table to determine the selected destination probabilistically. Those individuals where 'outside Auckland' (out-migration or death) is selected as the destination are removed from the dataset.

As the decision to move is affected by duration of stay (Poot 1987), we include the number of years the resident has lived in the origin area unit as an explanatory variable in the locational transitional equations along with all variables included in the ethnic transition equations.

[Insert Table 2 here]

4.1 Simulation Evaluation

We evaluate the performance of our model in two ways. First, we compare the proportion of people who changed their ethnicity, the proportion of people who changed their location, and the proportion of people who moved out of Auckland between 2001 and 2006 in our simulated data to those in the actual 2001-2006 linked census data. Second, we compare measures of residential sorting based on the simulated data for 2006 with those based on actual 2006 census data. In our comparisons, we use different forecast error measures to estimate forecast error and bias in the model.

Measures of residential sorting

There are many different measures that can be used as indicators of residential sorting (see e.g. Nijkamp and Poot 2015; Reardon and Firebaugh 2002; Massey and Denton 1988). We choose entropy-based measures, following the influential contribution by Theil and Finezza (1971). Entropy measures are conceptually and mathematically attractive and are the least biased by group size (Mondal et al. 2021a; Reardon and Firebaugh 2002). The measures used in our analysis are detailed in Table 3. In order to observe the extent to which ethnic groups are over- or under-represented in an area unit, we calculate the diversity (entropy) index (E_a) of the population in area unit *a* in terms of the given ethnic group classifications. Following Nijkamp and Poot (2015), we normalise the entropy diversity index to an evenness index I_a that varies between zero and one. The value of the diversity evenness index is zero (i.e. $E_a = 0$) when only one of the groups is present in area unit a and is one (i.e. $E_a = 1$) when all groups are equally represented in area unit *a* (Nijkamp and Poot 2015). We also use the Entropy Index of spatial sorting of group g (EIS_g), which measures the area-population weighted average of one minus the relative entropy of the areas $\left(\frac{E_{ga}}{\bar{E}_g}\right)$ with respect to group *g* (see Table 3). This index varies between zero (when the group is distributed proportionally to the total population in all area units) and one (when all areas in which group g is represented contain no other group). We also calculate an overall measure of residential sorting *(H*)*, by taking the group-population weighted average of the EIS_g values. This is an alternative way of calculating the *Theil's Multi-group Segregation Index H* (White 1986; Theil 1972; Theil and Finezza 1971). This calculation gives approximately the same value as *H* (for which the formula is not included in Table 3), but is easier to interpret. Finally, we also calculate the normalised diversity (entropy) index *I** of the whole Auckland population in terms of the given ethnic group classifications.[15]

[Insert Table 3 here]

Projection error measures

Following Cameron and Cochrane (2017) and Wilson (2015), we estimate multiple measures of projection error and bias. Projection error is defined as the difference between the index values based on the modelled / simulated population (M_t) and the actual population (A_t), standardised by the actual population size. Thus, the projection's Percentage Error at time *t* based on data at time *t*-5 ($PE_{t-5,t}$) is given as:

[15] Despite the entropy based diversity and sorting measures requiring us to take the natural logarithm of population shares when certain groups may be absent from certain areas, this does not cause a computational problem because $-\frac{P_{ga}}{P_a}\ln\frac{P_{ga}}{P_a} = 0$ when $P_{ga} = 0$, given that 0*ln(1/0)= $\lim_{q\to 0}[-q(\ln(q)] = 0$. See also the note at the bottom of Table 3.

$$PE_{t-5,t} = \frac{M_t - A_t}{A_t} x100\%$$

To report projection accuracy, we use the weighted mean absolute percentage error (WMAPE) as our primary measure. This is a weighted mean of the absolute Percentage Errors (PE_t), with weights equal to the actual group size proportions of the population in the year projected (Wilson 2012; Siegel 2002). WMAPE is preferable in cases where population sizes vary widely. In our study, population size of an area unit in Auckland varies from less than 9 to over 3000. WMAPE in projected year *t* is defined as:

$$WMAPE_{t-5,t} = \sum\nolimits_g \left(\left| PE_{t-5,t}^{g} \right| \frac{P_{gt}}{P_t} \right)$$

where g is the number of groups, P_{gt} is the population size of each group and P_t is size of the total Auckland population in year *t*.

The population projection error distribution is likely to be right-skewed due to small numbers of unusually high errors, resulting in the mean being a poor representation of the average error (Tayman and Swanson 1999). Thus, we also report the median absolute percentage error (MedAPE_t) and the median algebraic percentage error ($\mathrm{MedALPE}_t$), neither of which are affected by extreme outliers. MedAPE_t is the middle of the set of ranked absolute PE_t values. MedAPE_t is a measure of precision of a projection, because it is not influenced by the direction of the error. On the other hand, $\mathrm{MedALPE}_t$ measures the middle of a set of ranked non-absolute (i.e. algebraic) PE_t, values. This measure preserves the negative and the positive percentage error values.

4.2 Calibration Process

After performing the initial stages of model coding and running, we calibrated the model so that the simulated 2006 population using the 1996-2001 linked data in the NZLC would be as close as possible to the actual 2006 population. We expect that if the simulated proportion of people changing their location, the proportion of people in each ethnic group, and the proportion of each ethnic group changing their location are close to the actual proportions, then the model should be able to replicate the actual levels of ethnic diversity and residential sorting in the Auckland population in 2006. The calibration processes undertaken are described below.

Step 1: Calibrating the proportion of ’movers’

We observed that the percentage of people changing locations in our initial model was more than that observed in the actual data. We took the difference between the actual and the simulated proportion of people changing their location as our first calibration constant. We then added this calibration constant from the previously generated uniformly distributed random variable of staying at the current location, thereby ensuring that the model would decrease the number of ‘movers’. The model then uses this calibrated random variable to calculate the predicted probabilities to determine whether the person is a mover or not.

Step 2: Calibrating the proportion of people in each ethnic group

We calculated the difference between the proportion of people in each ethnic group between the simulated data and the actual data. We considered the difference for each ethnic group as a calibration constant for that ethnic group. For the cases where the model simulation generated too many members in an ethnic group, we added a calibration constant onto the uniformly distributed random variable. We subtracted the calibration constants from the random variable if the model simulation generated too few members of an ethnic group. This process was repeated several times, aiming to minimise the sum of the absolute differences between actual and simulated proportions.

Step 3: Calibrating the proportion of people in each ethnic group who are ‘movers’

We calculated the differences between the proportion of people changing location in the simulated data and the actual data for each ethnic group. We treated these differences for each ethnic group as ethnic-specific calibration constants. We then subtracted the calibration constant for ethnicity *i* from the predicted probability of moving for people who belong to ethnicity *i*. For people belonging to multiple ethnic groups we subtracted all of the ethnic-specific calibration constants that apply to them from the predicted probability of moving. Again, this process was repeated several times, aiming to minimise the sum of the absolute differences between actual and simulated proportions.

Section 5: Results

The ultimate aim of the dynamic spatial MSM model is to be a *projection model* that will project the population forward with errors that remain small enough for the results to be useful for informing local public policy and urban management. The outcome depends strongly on the extent to which we can accurately model transitions. To obtain the predicted probabilities for both ethnic transition and location transition, we ran logistic regression equations with clustered standard errors.[16]

There are too many coefficients to discuss the logistic regression results in detail. However, there are some general patterns that provide insight into the determinants of location and ethnicity transitions. Most generally, the coefficients often differ between adults and children (those aged less than 18).[17] The logistic regression of intra-urban mobility (in column (1) of Tables A1-A3) shows that NZ Europeans are more mobile than average while those with Pacific Island ethnicity are less mobile. As expected, residential mobility declines with age and with duration of residence. Females are less mobile. Ethnic diversity of area units and the various ethnic group shares do not appear to influence the rate of intra-urban mobility. However, New Zealand born children and adolescents are less mobile than the other 0-17 year olds.

With respect to ethnic mobility, there is – as expected – a lot of persistence: the most important predictor of ethnicity at time *t* is ethnicity at time *t*–5. There are also some interesting correlations between ethnic groups. For example, having declared to be 'Other European' at the previous census has a positive effect on declaring to be a 'New Zealand European' in the current census. Similarly, having declared to be Asian or from the Pacific in the previous census generally reduces the likelihood of declaring 'Other European' ethnicity in the current census. Ethnic mobility is lower at older ages and among the New Zealand born, i.e. the non-immigrants. High ethnic diversity of an area unit (i.e. a relatively large value of the entropy diversity index) leads to a greater likelihood of declaring 'Other European', 'Samoan' or 'Middle Eastern' ethnicity. A large 'own group' share of the area unit population does not always imply a stronger identification with that group. In fact the opposite is sometimes true. For example, in areas where the share of 'NZ European' or of 'Other European' is large, the likelihood of declaring these respective ethnicities is lower.

We validated the ability of the current model to replicate known 2006 census outcomes. Table 4 shows that 21 percent of people, who were in Auckland in 2001 and 2006, changed at least one of their ethnicities during the intercensal period, whereas for the simulated 2006 Auckland population, this proportion is very similar: 22 percent. The percentage of people reporting

---

[16] Appendix tables A1, A2 and A3 show the logistic regression results.
[17] However, no formal statistical tests of equality of coefficients were conducted.

moving from one area unit in 2001 to a different area unit in 2006 was 40 percent in the 2006 Census. The simulated percentage is 42 percent, i.e. very similar. The difference in the percentage of people moving out of Auckland between the actual and the simulated data is 3 percentage points, being 9 percent and 6 percent respectively.

[Insert Table 4 here]

Table 5 shows that in terms of overall ethnic residential sorting in Auckland, our simulated value for the Theil's multi-group spatial sorting index (H^*) is close to the actual value, the difference being -0.008 (or 9.7 percent). Table 5 also shows that the simulated ethnic diversity in Auckland (I^*) very closely matches the actual ethnic diversity observed in Auckland in 2006.

[Insert Table 5 here]

Table 6 summarises the three forecast error measures (WMAPE, MedAPE and MedALPE) for both the Entropy Index of Segregation measure for ethnic groups EIS_g and the Normalised Entropy Diversity measure for area units I_a. The WMAPE is smaller than the MedAPE for the simulated spatial sorting/segregation of the ethnic groups (19.34 and 28.53 respectively). The fact that the MedALPE has the same absolute value as the MedAPE indicates that the simulation underestimates group segregation for all groups.

The negative MedALPE value (-28.5 percent) reflects therefore that there is downward bias in the simulated values of the Entropy Index of Segregation measure, potentially resulting from the fact that not all determinants of ethnic mobility have been observed. The inconsistencies in the ethnic categorisations in the 1996 and 2001 Census data mentioned in Section 3, which were used to parameterise the initial model, contribute to the model performance. This is demonstrated by the fact that although the simulated and the actual measures of overall ethnic residential sorting in Auckland are very similar (Table 5), the model does not perform as well when we simulate the ethnic residential sorting for individual ethnic groups.

With respect to the diversity measure, the WMAPE is larger than the MedAPE, which is in turn larger than the MedALPE (4.07, 3.54 and 1.68 respectively). It is clear that the simulation performs better in projecting the diversity of areas than the spatial sorting of ethnic groups.

[Insert Table 6 here]

Section 6: Conclusion

The main aim of this paper is to describe the development and calibration of a microsimulation model that can be used for projecting the future spatial ethnic distribution in Auckland. The model described in this paper takes both ethnic and spatial mobility into consideration. Data from the 1986-2006 NZLC was used to simulate the spatial distribution of the Auckland population by ethnicity in 2006. The simulated results were then compared to the actual 2006 Census data.

We have demonstrated that census data can be used to inform, calibrate and validate our model. Our simulation is generally capable of reproducing the dynamics of residential sorting in Auckland, without detailed information on all the elements of an individual's residential decision-making process. Projection errors vary with population size of a region (Tayman et al. 1998; Smith and Shahidullah 1995). Smith and Shahidullah (1995) worked on projections of total population for all census tracts in three counties in Florida (Dade, Duval and Pinellas) and found that error measure values decline with increase in population size. Their reported Mean Absolute Percentage Errors (MAPEs) ranged from 17.3 per cent to 27.6 per cent. Tayman et al. (1998), in their work on census tracts projections in San Diego County reported that in the census tracts with population size between 1,000 and 1,500 the MAPE values were as high as 56.5 per cent and 46.2 per cent respectively. Keeping in mind that the area unit population composition in our work is around 1,500 on average, the results show that our model projects the spatial distribution of ethnicities in Auckland with a reasonable level of error.

Results from the locational transition module are fairly close to the actual data. However, our ethnic transition module appears to generate a lower degree of accuracy. We interpret this as caused by inconsistencies in the ethnic categorisation in the census data that were used in developing our model. We infer this from the fact that the way in which both the ethnic and locational transition modules work is similar.

This model is not without limitations. First, with a given set of predictor variables, logistic regression equations are used to predict the probability of a certain event occurring. Hence, only data from people who have been linked in the 1996-2001 NZLC could be used in estimating the logistic regression equation. However, the base population for the simulation is the whole Auckland population in the 2001 Census, whether linked in the 1996-2001 NZLC data or not. Thus, any extent to which unlinked and linked people differ in ways that are correlated with the transitions we estimate will generate bias in the results. However, some of this bias will be attenuated through the process of calibration.

Second, due to few people reporting as belonging to the ‘Not further defined (NFD)’ and ‘Other’ ethnic groups, we combined these into one broad ethnic group called ‘ONFD’. As the ‘NFD’ groups are a disaggregated Level 2 category in the ethnic classification under each broad Level 1 ethnic category, they are likely to behave more like the other sub-groups within their Level 1 broad ethnic group than they would to the ‘Other’ Level 1 ethnic group with which they have been merged. This problem could be eliminated by removing these ethnic groups from the model, but at a cost of deviating the model further from the underlying real-world data from the full census. Hence we preferred to retain these ethnic groups at this stage of model development. A future extension to this work could be to separate these ethnic groups or merge them into other Level 2 groups within the same Level 1 broad ethnic group, and observe the effect on the model results. These model extensions would become easier if the model were extended to consider the future ethnic diversity of the whole of New Zealand, wherein the problem of small cell counts for these groups would be reduced.

Third, an individual’s location decision and ethnic choices are dependent on a variety of factors other than the ones that are used in the model, one of these being their completed education level (which can also proxy for income). Although data on the completed education for adults is available in the Census, the same data for children transitioning to adulthood is not available. Including education within the model would require the addition of a module on educational attainment. We initially attempted to parameterise such a model, but it performed poorly.[18] Thus, we have not included education as a predictor variable in the model. As a future prospect for research, it would be interesting to see how adding an additional educational transition module to the model alters the results. Fourth, ethnic identity of the parents is important for the evolution of ethnic identity of adolescents (Mondal et al. 2020). However, the NZLC does not have this data. Thus, we could not include this variable in the model.

In spite of these limitations, this paper has described the development of a modelling approach to project urban ethnic diversity at a fine spatial scale and relatively narrowly defined ethnic groups. Our model was developed using Stata, which extends the number of resources previously used to build and run microsimulation models. Our future focus will be to use this

[18] Further details are available from the authors on request.

calibrated model, 2013-2018 NZLC data, and 2023 Census data when they become available, to project the future ethnic spatial distribution in Auckland forward to 2038.

References

Ardestani, B.M. (2013). Using a hybrid model for investigating residential segregation: an empirical and simulation-based study (Doctoral thesis, University of Auckland, New Zealand). Retrieved from https://researchspace.auckland.ac.nz/handle/2292/21618

Ardestani, B. M., O'Sullivan, D., & Davis, P. (2018). A multi-scaled agent-based model of residential segregation applied to a real metropolitan area. *Computers, Environment and Urban Systems*, 69, 1–16.

Atkinson A. B., Bourguignon, F., & Chiappori, P. A. (1988). What do we learn about tax reforms from international comparisons? France and Britain. *European Economic Review*, 32 (2-3): 343-52.

Atkinson, A. B., Bourguignon, F., O'Donoghue, C., Sutherland, H. & Utili, F. (2002). Microsimulation of social policy in the European Union: Case study of a European minimum pension. *Economica*, 69, 229-243.

Ballas, D. and Clarke, G.P. (2000). GIS and microsimulation for local labour market policy analysis. *Computers, Environment and Urban Systems*, 24, 305-330.

Ballas, D. (2001). A spatial microsimulation approach to local labour market policy analysis (unpublished PhD thesis). School of Geography, University of Leeds.

Ballas, D., Clarke, G., & Wiemers, E. (2005a). Building a dynamic spatial microsimulation model for Ireland. *Population, Space and Place*, 11(3), 157-172.

Ballas, D., Rossiter, D., Thomas, B., Clarke, G.P., and Dorling, D. (2005b). *Geography Matters: Simulating the Local Impacts of National Social Policies.* York, UK: York Publishing Services.

Blundell, R., Duncan, A., McCrae, J. & Meghir, C. (2000). The labour market impact of the working families' tax credit. *Fiscal Studies*, 21(1), 75–104.

Bonin, H., Kempe, W., & Schneider, H. (2002). Household Labour Supply Effects of Low-wages Subsidies in Germany. IZA Discussion Paper 637, Institute for the Study of Labor, Bonn, Germany.

Bourguignon, F., & Spadaro, A. (2006). Microsimulation as a tool for evaluating redistribution policies. *Journal of Economic Inequality*, 4, 77-106.

Brown, P. & A. Gray (2009). Inter-ethnic mobility between the 2001 and 2006 Censuses: The statistical impact of the New Zealander response. In Statistics New Zealand, *Final Report of a Review of the Official Ethnicity Standard 2009.* Wellington, New Zealand (pp 27-36).

Bruch, E. E., & Mare, R. D. (2006). Neighborhood choice and neighborhood change. *American Journal of Sociology*, 112(3), 667–709.

Caldwell, S., Clarke, G., & Keister, A. (1998). Modelling regional changes in US household income and wealth: A research agenda. *Environment and Planning C: Government and Policy*, 16(6), 707-722.

Callan, T., & Sutherland, H. (1997). The impact of comparable policies in European countries: Microsimulation approaches. *European Economic Review*, 41 (3-5), 327–333.

Cameron, M. P. & Cochrane, W. (2017). Using land-use modelling to statistically downscale population projections to small areas. *Australasian Journal of Regional Studies*, 23(2), 195-216.

Cameron, M. P. & Poot, J. (2019). Towards superdiverse Aotearoa: Dimensions of past and future ethnic diversity in New Zealand and its regions. *New Zealand Population Review*, 45, 18-45.

Carter, K. N., Hayward, M., Blakely, T., & Shaw, C. (2009). How much and for whom does self-identified ethnicity change over time in New Zealand? Results from a longitudinal study. *Social Policy Journal of New Zealand*, 36, 32-45.

Creedy, J. (1999). *Modelling Indirect Taxes and Tax Reform*. Northampton, UK, Edward Elgar.

Creedy, J. & Duncan, A. (2002). Behavioural microsimulation with labour supply responses. *Journal of Economic Surveys*, 16, 1-39.

Das, M. & van Soest, A. (2001). Family labour supply and proposed tax reforms in the Netherlands. *De Economist*, 149, 191–218.

Davis, P., & Lay-Yee, R. (2019). *Simulating Societal Change: Counterfactual Modelling for Social and Policy Inquiry*. Switzerland: Springer International Publishing.

Denton, N.A., & Massey, D.S. (1988). Residential segregation of blacks, Hispanics and Asians by socioeconomic status and generation. *Social Science Quarterly*, 69, 797–817.

Demery, L. (2003). Analysing the incidence of public spending. In F. Bourguignon & P. da Silva L.(eds.), *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools* (pp 41-68). World Bank Publications.

Didham, R., Nissen, K., & Dobson, W. (2014). Linking censuses: New Zealand longitudinal census 1981–2006. Available from www.stats.govt.nz.

Didham, R. (2016). Ethnic mobility in the New Zealand census, 1981-2013: A preliminary look. *New Zealand Population Review*, 42, 27–42.

Domina, T. (2006). Brain drain and brain gain: rising educational segregation in the United States, 1940–2000. *City Community*, 5, 387–407.

Duncan, O.D., & Duncan, B. (1955) Residential distribution and occupational stratification. *American Journal of Sociology,* 60, 493–503.

Eggink, E., Woittiez, I., & Ras, M. (2016). Forecasting the use of elderly care: a static micro-simulation model. *The European Journal of Health Economics: HEPAC: Health Economics in Prevention and Care*, 17(6), 681–691.

Farley, R. (1977) Residential segregation in urbanized areas of the United States in 1970: an analysis of social class and racial differences. *Demography,* 14, 497–518.

Favreault, M., & Smith, K.E. (2004). A primer on the dynamic simulation of income model (DYNASIM3). *The Urban Institute*.

Fischer, M. J. (2003). The relative importance of income and race in determining residential outcomes in U.S. urban areas, 1970-2000. *Urban Affairs Review*, 38, 669–696.

Goldman, D.P., Zheng, Y., Girosi, F., Michaud, P.C., Olshansky, S.J., Cutler, D., & Rowe, J.W. (2009). The benefits of risk factor prevention in Americans aged 51 years and older. *American Journal of Public Health*, 99(11), 2096-2101.

Hancock, R., Mallender, J., & Pudney, S. (1992). Constructing a computer model for simulating the future distribution of pensioners' incomes for Great Britain. In R. Hancock & H. Sutherland (eds). *Microsimulation Models for Public Policy Analysis: New Frontiers* (pp. 33-66). London UK: Suntory-Toyota International Centre for Economics and Related Disciplines.

Harding, A. (2007). APPSIM: The Australian dynamic population and policy microsimulation model. *National Centre for Social and Economic Modeling*, Canberra, Australia.

Ho, E., & Bedford, R. D., (2006). The Chinese in Auckland: changing profiles in a more diverse society. In: Li, W. (ed.) *From Urban Enclave to Ethnic Suburb.* University of Hawaii Press, Honolulu (pp 203-233).

Holm, E., Holme, K., Lindgren, U., & Makila, K. (2002). The SVERIGE spatial microsimulation model. Discussion paper 595, *Department of Social and Economic Geography*, Umea University.

Holmer, M., Janney, A., Cohen, B. (2014). *PENSIM overview*. Washington, DC: U.S. *Department of Labor*, *Office of Policy Research*.

Iceland, J., Weinberg, D.H., Steinmetz, E. (2002) Racial and ethnic residential segregation in the United States 1980–2000, Appendix B. US Census Bureau, Washington.

Immervoll, H., Kleven, H. J., Kreiner, C. T., & Saez, E. (2007). Welfare reform in European countries: a microsimulation analysis. *The Economic Journal*, *117*(516), 1–44.

Johnston, R., Poulsen, M., & Forrest, J. (2011). Evaluating changing residential segregation in Auckland, New Zealand, using spatial statistics. *Tijdschrift Economische en Sociale Geografie,* 102: 1–23.

Kang, I. (2017) New Zealand Longitudinal Census development. *International Journal of Population Data Science*, 1(1): 232 (Proceedings of the IPDLN Conference (August 2016)). doi: 10.23889/ijpds.v1i1.252.

Kaplanoglou, G. & Newbery, D.M. (2003). Indirect taxation in Greece: Evaluation and possible reform. *International Tax and Public Finance*, 10, 511–533.

King, A., Baekgaard, H., & Robinson, M. (1999). Dynamod2: An overview. *Technical Paper 19, NATSEM, University of Canberra*.

Lambert, S., Percival, R., Schofield, D., & Paul, S. (1994). An introduction to STINMOD: a static microsimulation model. (STINMOD technical paper; No. 1). Canberra: National Centre for Social and Economic Modelling (NATSEM).

Liberati, P. (2001). The distributional effects of indirect tax changes in Italy. *International Tax and Public Finance*, 8(1), 27–51.

Li, J., & O'Donoghue, C. (2013). A survey of dynamic microsimulation models: uses, model structure and methodology. *International Journal of Microsimulation*, 6(2), 3-55.

Lomax, N., & Smith, A. (2017). Microsimulation for demography. *Australian Population Studies*, 1(1), 73–85.

Massey, D.S., & Denton, N.A. (1988). The dimensions of racial segregation. *Social Forces*, 67, pp. 281-315.

Malenfant, É. C., Lebel, A., & Martel, L. (2015). Projections of the diversity of the Canadian population, 2006 to 2031. Statistics Canada. Retrieved from https://www150.statcan.gc.ca/n1/pub/91-551-x/91-551-x2010001-eng.htm

Maré, D.C. and Poot, J. (2022). Accounting for social difference when measuring cultural diversity. Motu Working Paper 22-04. Motu Economic and Public Policy Research. Wellington, New Zealand.

Milne, B., Lay Yee, R., McLay, J. M., Pearson, J., Von Randow, M., & Davis, P. (2015). Modelling the Early life-course (MELC): A microsimulation model of child development in New Zealand. 8(2), 28–60.

Moeckel, R. (2009). Simulation of firms as a planning support system to limit urban sprawl of jobs. *Environment and Planning B: Planning and Design*, 36: 883-905.

Mondal, M., Cameron, M.P., & Poot, J. (2020). Determinants of ethnic identity among adolescents: Evidence from New Zealand. (Working paper in Economics No.5/20). Hamilton, New Zealand: University of Waikato.

Mondal, M., Cameron, M.P., & Poot, J. (2021a). Group-size bias in the measurement of residential sorting. In S. Suzuki & R. Patuelli, (Eds.), *A Broad View of Regional Science: Essays in Honor of Peter Nijkamp.* Singapore: Springer Nature, Chapter 7, pp. 113-136.

Mondal, M., Cameron, M.P., & Poot, J. (2021b). Cultural and economic residential sorting of Auckland's population, 1991-2013: An entropy approach. *Journal of Geographical Systems*, *23*(2): 291-330.

Mot, E. S. (1992). *Survey of Microsimulation Models*. Social Security Research Committee, The Hague: VUGA.

Nijkamp, P., & Poot, J. (2015). Cultural diversity: a matter of measurement. In: Nijkamp, P., Poot, J., & Bakens, J. (eds.) *The Economics of Cultural Diversity*. Edward Elgar, Cheltenham, pp. 17-51.

O'Sullivan (2009). Changing neighbourhoods-neighbourhoods changing: A framework for spatially explicit agent-based models of social systems. *Sociological Methods & Research*, 37(4), 498-530.

Paulus, A., Čok,M., Figari, F., Hegedüs, P., Kump, N., Lelkes, O., Levy, H., Lietz, C., Lüpsik, S., Mantovani, D., Morawski, L., Sutherland, H., Szivos, P., & Võrk, A. (2009). The effects of taxes and benefits on income distribution in the enlarged EU. EUROMOD Working Paper (No. EM8/09).

Pechman, J., & Okner, B. (1974). *Who Bears the Tax Burden?* Washington, DC: The Brookings Institution.

Poot, J. (1987). Estimating duration-of-residence distributions: Age, sex and occupational differentials in New Zealand. *New Zealand Geographer*, *43*(1), 23–32.

Reardon, S.F., Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological Methods*, 32, 33–67.

Reardon, S.F., Farrell, C.R., Matthews, S.A., O'Sullivan, D., Bischoff, K., & Firebaugh G. (2009). Race and space in the 1990s: Changes in the geographic scale of racial residential segregation, 1990-2000. *Social Science Research*, 38(1), 55-70.

Rees P.H., Wohland P., Norman P., Lomax N., & Clark S.D. (2017). Population projections by ethnicity: Challenges and solutions for the United Kingdom. In: D. A. Swanson D (eds), *The Frontiers of Applied Demography* (Volume 9, pp.383-408). Springer.

Rephann, T. (2004). Economic-demographic effects of immigration: Results from a dynamic spatial microsimulation model. *International Regional Science Review*, 27(4), 379-410.

Rowe, G., & Wolfson, M., (2000). Public pensions-Canadian analysis based on the Lifepaths generational accounting framework. 6th Nordic Seminar on Microsimulation Models, Copenhagen, Denmark.

Rogers, S., Rineer, J., Scruggs, M., Wheaton, W., Cooley, P., Roberts, D., & Wagener, D. (2014). A geospatial dynamic microsimulation model for household population projections. *International Journal of Microsimulation*, *7*(2), 119-146.

Sahn, D., & Younger, S. (2003). Estimating the incidence of indirect taxes in developing countries. In F. Bourguignon & P. da Silva L., eds., *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools*. World Bank Publications (pp. 27-40).

Schelling, T.C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1, 143–186.

Siegel, J. S. (2002). *Applied Demography: Applications to Business, Government, Law and Public Policy.* Academic Press, San Diego, CA.

Simkus, A.A. (1978). Residential segregation by occupation and race in ten urbanized areas, 1950-1970. *American Sociological Review*, 43, 81–93.

Smith, S., & Shahidullah, M. (1995). An evaluation of population projection errors for census tracts. *Journal of the American Statistical Association*, 90(429), 64-71.

Smith, K.E., Favreault, M., Ratcliffe, C.E., Butrica, B.A., Toder, E.J., & Bakija, J. (2007). *Modeling Income in the Near Term 5*. http://dx.doi.org/10.2139/ssrn.2206397

Spielauer, M. (2011). What is social science microsimulation? *Social Science Computer Review*, 29(1), 9-20.

Statistics Canada. (2018). Demosim. Retrieved October 17, 2019, from https://www.statcan.gc.ca/eng/microsimulation/demosim/demosim

Statistics New Zealand. (2007a). *Profile of New Zealander Responses, Ethnicity Question: 2006 Census*. http://archive.stats.govt.nz/Census/about-2006-census/profile-of-nzer-responses-ethnicity-question-2006-census.aspx

Statistics New Zealand. (2007b). *Survey of dynamics and motivations for migration in New Zealand – information releases*. http://www.stats.govt.nz/

Statistics New Zealand. (2013). *2013 Census definitions and forms.* Available from www.stats.govt.nz.

Statistics New Zealand. (2015). *IDI data dictionary: 2013 census data*. Available from www.stats.govt.nz.

Statistics New Zealand. (2017). 2001 Census of Population and Dwellings: change in ethnicity question. Available from http://www.stats.govt.nz/Census/2001-census-data/change-in-ethnicity-question.aspx

Statistics New Zealand. (2020). *Ethnic group (detailed total response - level 3) by age and sex, for the census usually resident population count, 2006, 2013, and 2018 Censuses (RC, TA, SA2, DHB).* Available from http://nzdotstat.stats.govt.nz

Tayman, J., Schafer, E., & Carter, L. (1998). The role of population size in the determination and prediction of population forecast errors: An evaluation using confidence intervals for subcounty areas. *Population Research and Policy Review*, 17(1), 1–20.

Tayman, J. & Swanson, D.A. (1999). On the validity of MAPE as a measure of population forecasting accuracy. *Population Research and Policy Review*, 18, 523-528.

Theil, H. (1972). *Statistical Decomposition Analysis: With Applications in the Social and Administrative sciences*. North-Holland, Amsterdam.

Theil, H., & Finezza, A.J. (1971). A note on the measurement of racial integration of schools by means of informational concepts. *Journal of Mathematical Sociology*, 1, pp. 187-194.

Uyeki, E.S. (1964). Residential distribution and stratification, 1950-1960. *American Journal of Sociology*, 69, 491–498.

Vencatasawmy, C.P., Holme, K., Rephann, T., Esko, J., Swan, N., Öhman, M., Åström, M., Alfredsson,E., & Siikavaara, J. (1999). Building a spatial microsimulation model. Paper presented at the 11th European Colloquium on Quantitative and Theoretical Geography in Durham, England, on September 3-7, 1999. Retrieved from https://pdfs.semanticscholar.org/3f01/36d2633e1ca80c0a228aca492e25c1a9c9dc

Willekens, F. (2006) Description of the micro-simulation model (continuous-time micro-simulation). Deliverable D8 (first part), MicMac Bridging the micro-macro gap in population forecasting, NIDI, The Netherlands.

Willekens, F. (2016). Migration flows: measurement, analysis and modeling. In M.J. White (Ed.), *International Handbook of Migration and Population Distribution* (pp. 225-24). *Springer.* https ://DOI 10.1007/978-94-017-7282-2_11

Williamson, P., Birkin, M. & Rees, P. (1998). The estimation of population microdata by using data from small area statistics and samples of anonymised records. *Environment and Planning A*, 30, 785–816

Wilson, T. (2012). Forecast accuracy and uncertainty of Australian Bureau of Statistics state and territory population projections. *International Journal of Population Research,* 2012, 1-16. https://doi.org/10.1155/2012/419824

Wilson, T. (2015). Short-term forecast error of Australian local government area population projections. *Australasian Journal of Regional Studies*, 21, 253-275.

Wolfson, M, & Rowe, G. (2013). HealthPaths: Using health trajectories to estimate and simulate the relative importance of determinants of health-adjusted life expectancy (HALE): a statistical analysis. *The Lancet*, 381(2), Special issue, S148.

White, M.J. (1986). Segregation and diversity measures in population distribution. *Population Index*, 52, 198-221.

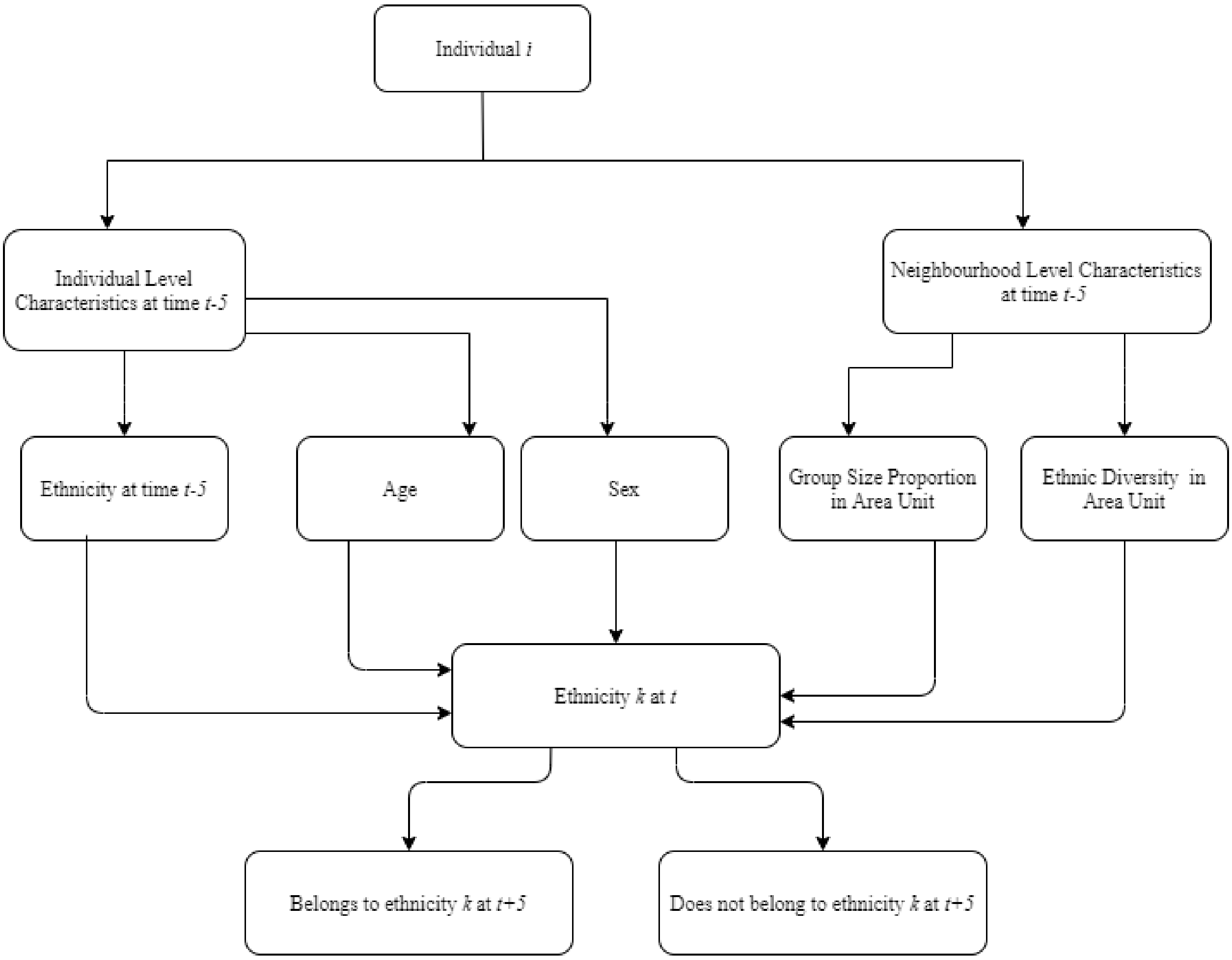
Wu, B. M., Birkin, M. H., & Rees, P. H. (2011). A dynamic MSM with agent elements for spatial demographic forecasting. *Social Science Computer Review*, 29(1), 145–160.

Zaidi, A., & Rake, K. (2001). Dynamic microsimulation models: a review and some lessons for SAGE. The London School of Economics, Simulating Social Policy in an Ageing Society (SAGE), Discussion Paper. 2.

Zucchelli, E., Jones, A.M., & Rice, N. (2010). The evaluation of health policies through microsimulation methods. (HEDG working paper number 10/03), Department of Economics, University of York.

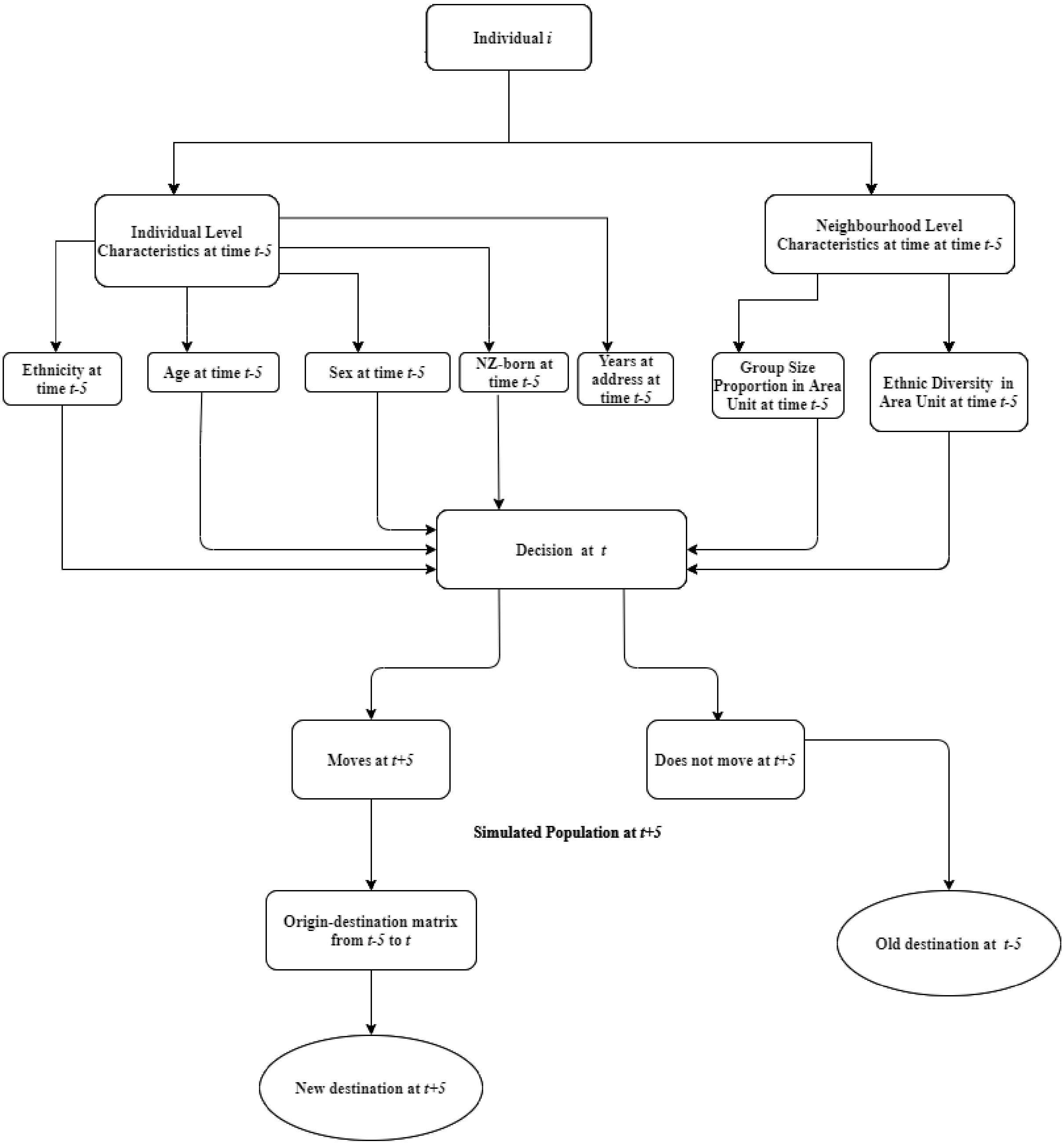
Figures

Figure 1: Theoretical framework – Ethnic Transition



Note: The figure does not include 'New Zealand born'/'Foreign born' status, which is an additional determinant of ethnicity at time *t*.

Figure 2: Theoretical framework – Location Transition



Tables

Table 1: Ethnic group classification in New Zealand

Ethnic group code (Level 1)	Ethnic group code description (Level 1)	Ethnic group code (Level 2)	Ethnic group code description (Level 2)	Ethnic group in simulation
01	European	10	European not further defined	2
		11	NZ European	1
		12	Other European	2
02	Māori	21	NZ Māori	3
03	Pacific Peoples	30	Pacific Island not further defined	10
		31	Samoan	4
		32	Cook Island Māori	5
		33	Tongan	6
		34	Niuean	7
		35	Tokelauan	8
		36	Fijian	9
		37	Other Pacific Island	10
04	Asian	40	Asian not further defined	14
		41	Southeast Asian	11
		42	Chinese	12
		43	Indian	13
		44	Other Asian	14
05	MELAA	51	Middle Eastern	15
		52	Latin American/Hispanic	16
		53	African	17
06	Other	61	Other ethnicity	18

Source: Statistics New Zealand (2013)

Table 2: Variables used in the analysis

Module	Predicted Variable	Level of variables	Predictor variables (all evaluated at the time of the previous census)
Ethnic Transition	Ethnic affiliation in current census (1=belongs to ethnic group *i*, 0=otherwise)	Individual	Ethnicity, age, sex, NZ-born
		Neighbourhood	Ethnic diversity in area unit, Ethnic group size proportions in area unit.
Location Transition	Moved[19] (1=moved, 0=otherwise)	Individual	Ethnicity, age, sex, NZ-born, years at address
		Neighbourhood	Ethnic diversity in area unit, Ethnic group size proportions in area unit.

Note: The coefficients of the predictor variables in these logit models are estimated separately for the population aged 0-17 ('Children and Adolescents') and the population aged 18 and over ('Adults')

[19] We created the binary variable 'moved' (1=if individual changed area unit during the intercensal period, 0=otherwise) from the census data on location of usual residence in the current census and the variable 'address five years ago' for the same individual.

This page has been left blank.

Table 3: Summary Measures of Residential Sorting

Entropy diversity (area unit)	$E_a = -\sum_{g=1}^{G} \frac{P_{ga}}{P_a} \ln \frac{P_{ga}}{P_a}$
Normalised Entropy diversity (area unit)	$I_a = -\frac{\sum_{g=1}^{G} \frac{P_{ga}}{P_a} \ln \frac{P_{ga}}{P_a}}{\ln(G)}$
Normalised Entropy diversity (city)	$I^* = -\frac{\sum_{g=1}^{G} \frac{P_g}{P} ln \frac{P_g}{P}}{ln(G)}$
Entropy Index of Segregation (group)	$EIS_g = \sum_{a=1}^{A} \frac{P_a}{P} \left(1 - \frac{E_{ga}}{\bar{E}_g}\right)$
	where: $E_{ga} = -\frac{P_{ga}}{P_a} \ln\left(\frac{P_{ga}}{P_a}\right) - (1 - \frac{P_{ga}}{P_a})\, ln\left(1 - \frac{P_{ga}}{P_a}\right)$
	$\bar{E}_g = -\frac{P_g}{P} ln\left(\frac{P_g}{P}\right) - (1 - \frac{P_g}{P})\, ln\left(1 - \frac{P_g}{P}\right)$
Theil's multi-group spatial sorting index (city)	$H^* = \sum_{g=1}^{G} \frac{P_g}{P} EIS_g$

Notes: P_{ga} refers to the population of group g (=1, 2,…G) in area a (= 1,2,….A). P_a is the total number of people in area unit a. P_g is the number of members of group g in Auckland and P is the total population of Auckland. Comparing group g with all other groups combined, we denote the entropy of area a as (E_{ga}) and of the whole Auckland city as $\bar{E}_g$. The calculation of *EIS* requires that we define 0*ln(1/0)= $\lim_{q\to 0}[-q \ln(q)] = 0$ to account for any cases in which group g is not represented in an area a. These summary measures of residential sorting are defined in Iceland et al. (2002).

Table 4: Comparison between simulated data and the actual 2006 Census data

Variable	Model	Actual	Difference (Model- Actual)
Ethnic change	22%	21%	1%
Location change	42%	40%	2%
Movement out of Auckland	6%	9%	-3%

Note: The table shows the difference in percentages of people, in the simulated 2006 data and the actual 2006 Census data.

Table 5: Actual and simulated spatial sorting in Auckland, 2006

Measures of Residential Sorting	Model	Actual	Absolute difference (Model-Actual) and percentage difference
Theil's multi-group index (*H**)	0.084	0.093	-0.008 (-9.7%)
Evenness Index (*I**)	0.654	0.656	-0.002 (-0.3%)

Note: The table shows the difference in the calculated sorting indexes based on the simulated 2006 Census data and the actual 2006 Census data.

Table 6: Model Performance

Error Measure	***EIS***	***I***
WMAPE (%)	19.34	4.07
MedAPE (%)	28.53	3.54
MedALPE (%)	-28.53	1.68

Notes: *EIS* refers to the Entropy Index of Segregation, calculated for ethnic group. *I* refers to Normalised Entropy diversity, calculated for each area unit. See also Table 3.

This page has been left blank.

Appendix

Table A1: Clustered Logistic Regression of Locational Change and Ethnic Identity

Variables	(1) Moved	(2) NZ European	(3) Other European	(4) NZ Māori	(5) Samoan	(6) Cook Island Māori	(7) Tongan	(8) Niuean	(9) Tokelauan	(10) Fijian	(11) Other PI	(12) SE Asian	(13) Chinese	(14) Indian	(15) Other Asian	(16) Middle Eastern	(17) Latin American	(18) African	(19) ONFD
									Adults (aged 18 and over)										
NZ European	0.163***	3.051***	-0.590***	-1.378***	-1.216***	-1.596***	-1.267***	-1.756***	-1.990***	-0.682***	-1.093***	-1.578***	-1.959***	-1.848***	-0.445	-0.483**	-0.923**	-0.612*	-0.314***
	(0.015)	(0.042)	(0.023)	(0.031)	(0.084)	(0.126)	(0.150)	(0.127)	(0.389)	(0.162)	(0.242)	(0.189)	(0.112)	(0.140)	(0.366)	(0.188)	(0.402)	(0.316)	(0.057)
Other European	0.029	1.137***	2.758***	1.684***	-1.145***	-1.151***	-1.308***	-1.149***	-1.270**	-0.386***	-0.441**	-2.393***	-2.214***	-2.285***	-1.150***	-1.130***	0.0251	0.225	1.447***
	(0.023)	(0.049)	(0.045)	(0.037)	(0.056)	(0.080)	(0.104)	(0.141)	(0.522)	(0.127)	(0.201)	(0.186)	(0.095)	(0.165)	(0.280)	(0.161)	(0.272)	(0.222)	(0.073)
NZ Māori	0.0615	-2.463***	0.629***	5.491***	-0.564***	-0.594***	-0.418***	-0.901***	-1.178**	0.334*	-0.197	-0.750***	-0.384***	-0.889***	0.375	-0.0392	0.305	-0.405	0.595***
	(0.0595)	(0.049)	(0.046)	(0.050)	(0.101)	(0.178)	(0.151)	(0.142)	(0.469)	(0.189)	(0.320)	(0.273)	(0.098)	(0.148)	(0.322)	(0.279)	(0.567)	(0.504)	(0.115)
Samoan	-0.418***	-2.420***	-0.360***	-1.493***	7.605***	-1.282***	-1.490***	-1.123***	0.231	-0.449**	0.289	-4.051***	-1.691***	-2.482***	-1.283**	-2.128***	-1.477*	-1.794**	-0.904***
	(0.095)	(0.049)	(0.078)	(0.078)	(0.060)	(0.159)	(0.177)	(0.193)	(0.326)	(0.201)	(0.316)	(0.503)	(0.140)	(0.257)	(0.588)	(0.493)	(0.778)	(0.768)	(0.167)
Cook Island Māori	-0.419***	-2.328***	-0.578***	-1.633***	-0.912***	7.705***	-1.129***	-1.250***	-0.964*	-0.530	1.402***	-3.397***	-1.881***	-2.357***	-0.247	-1.994**	-1.485	-0.663	-0.732***
	(0.089)	(0.067)	(0.084)	(0.106)	(0.174)	(0.075)	(0.207)	(0.218)	(0.525)	(0.457)	(0.347)	(0.607)	(0.268)	(0.339)	(0.455)	(0.793)	(1.112)	(0.805)	(0.204)
Tongan	-0.318***	-2.057***	-0.596***	-1.119***	-1.099***	-1.377***	7.669***	-1.762***	-1.227***	-0.367	0.113	-3.804***	-2.010***	-2.249***	-0.813	-0.970*	-2.063*	-0.244	-1.030***
	(0.090)	(0.076)	(0.102)	(0.104)	(0.157)	(0.245)	(0.086)	(0.203)	(0.429)	(0.256)	(0.432)	(0.600)	(0.301)	(0.318)	(0.587)	(0.562)	(1.224)	(0.510)	(0.251)
Niuean	-0.590***	-2.356***	-0.478***	-1.072***	-1.230***	-0.940***	-1.403***	8.293***	-0.683	-0.615*	-0.352	-2.913***	-1.326***	-2.249***	0.135	-1.196	-0.382	-0.399	-1.485***
	(0.106)	(0.085)	(0.122)	(0.095)	(0.234)	(0.256)	(0.208)	(0.099)	(0.681)	(0.332)	(0.614)	(0.536)	(0.271)	(0.542)	(0.557)	(0.770)	(0.591)	(0.716)	(0.388)
Tokelauan	-0.456***	-2.729***	-0.390	-1.207**	-1.064***	-1.731	-0.830	-0.144	9.050***	-2.062*	0.238		-2.032**		0.461				
	(0.127)	(0.383)	(0.417)	(0.533)	(0.408)	(1.698)	(0.827)	(0.775)	(0.261)	(1.128)	(0.909)		(0.860)		(1.649)				
Fijian	-0.097	-0.793***	-0.228*	-0.965***	-0.627*	-0.115	-0.353	-0.0513	1.693***	7.523***	1.933***	-3.170***	-0.696**	0.743***	0.394			0.311	1.993***
	(0.062)	(0.100)	(0.130)	(0.182)	(0.332)	(0.331)	(0.322)	(0.405)	(0.647)	(0.150)	(0.380)	(1.041)	(0.341)	(0.253)	(0.693)			(1.034)	(0.143)
Other PI	-0.004	-0.904***	0.863***	-1.161***	-0.532	1.187***	-0.920	0.218	-0.267	3.052***	7.094***	-0.648	-0.282	-1.172	2.053***			0.854	1.140***
	(0.088)	(0.173)	(0.163)	(0.341)	(0.428)	(0.449)	(0.639)	(0.360)	(0.876)	(0.392)	(0.187)	(0.684)	(0.548)	(0.772)	(0.596)			(1.496)	(0.307)
SE Asian	-0.050	-2.118***	-0.832***	-0.550**	-2.456***	-2.766***	-3.601***			-0.907	-0.159	5.708***	-0.183	-1.745***	0.210	-1.295**		0.236	1.205***
	(0.066)	(0.099)	(0.152)	(0.216)	(0.417)	(0.987)	(1.016)			(0.577)	(0.625)	(0.180)	(0.302)	(0.314)	(0.532)	(0.575)		(0.720)	(0.137)
Chinese	-0.210**	-2.756***	-1.014***	-0.649***	-0.784***	-1.436***	-2.178***	-1.965***	-1.696**	-0.489**	-0.845**	-1.226***	6.807***	-2.620***	-1.191*	-2.705***		-1.336**	-0.270*
	(0.087)	(0.106)	(0.099)	(0.080)	(0.085)	(0.191)	(0.299)	(0.314)	(0.847)	(0.248)	(0.377)	(0.171)	(0.088)	(0.255)	(0.638)	(0.764)		(0.625)	(0.153)
Indian	-0.179*	-3.029***	-1.064***	-1.131***	-2.172***	-2.041***	-3.175***	-3.374***	-1.349*	-1.559***	-0.0875	-2.518***	-2.396***	7.760***	-0.660	-1.024**		0.259	-0.620***
	(0.099)	(0.095)	(0.130)	(0.117)	(0.236)	(0.279)	(0.429)	(0.502)	(0.760)	(0.337)	(0.439)	(0.320)	(0.276)	(0.095)	(0.717)	(0.443)		(0.506)	(0.178)

Other Asian	0.051	-2.377***	-0.614***	-0.742	-1.769***	-1.675***	-1.209***	-1.993***	-0.109	-0.561		-2.513***	-2.125***	-0.274	9.325***				0.115
	(0.086)	(0.146)	(0.217)	(0.525)	(0.479)	(0.461)	(0.440)	(0.517)	(0.629)	(0.632)		(0.655)	(0.487)	(0.302)	(0.239)				(0.310)
Middle Eastern	0.066	-1.096***	-0.0886	-1.523***	-1.215**	-0.851		-1.184**	1.345	0.404	1.761**	-2.202**	-3.011***	-1.474**	0.436	7.402***	1.069	1.182	1.120***
	(0.107)	(0.150)	(0.170)	(0.337)	(0.599)	(0.627)		(0.529)	(1.189)	(1.017)	(0.798)	(0.917)	(0.767)	(0.737)	(0.853)	(0.162)	(1.005)	(1.076)	(0.220)
Latin American	0.142	-1.332***	1.286***	-0.082	-3.162***		-3.882***	-0.353		-0.599*		-1.409**	-0.841	-0.652			8.272***		1.332***
	(0.160)	(0.201)	(0.245)	(0.555)	(0.541)		(0.864)	(0.510)		(0.337)		(0.650)	(0.740)	(0.673)			(0.305)		(0.312)
African	-0.189	-0.317*	-0.276	-0.741***	-0.615	-0.155	-1.833**	-0.997*		0.703			-1.750*	0.0380			0.0239	7.596***	1.998***
	(0.129)	(0.168)	(0.239)	(0.253)	(0.646)	(0.817)	(0.803)	(0.606)		(1.034)			(0.930)	(0.896)			(0.634)	(0.221)	(0.223)
ONFD	-0.072	-0.389***	0.603***	-0.490*	0.114	-2.203*	-1.122	-1.353**		0.254	0.957	2.533***	-0.163	0.805*	1.611**	1.271**	2.294***	0.482	3.413***
	(0.084)	(0.128)	(0.149)	(0.295)	(0.359)	(1.133)	(0.774)	(0.688)		(0.379)	(0.725)	(0.282)	(0.478)	(0.435)	(0.630)	(0.603)	(0.677)	(0.707)	(0.132)
Children and Adolescents (aged 0-17)																			
NZ European	0.018	3.682***	-0.033	-0.650***	-1.008***	-0.978***	-1.124***	-0.788***	-0.920*	-0.757***	-0.468*	-0.793***	-1.443***	-1.692***	-0.762**	-0.016	-0.389	0.099	0.152
	(0.046)	(0.060)	(0.085)	(0.071)	(0.103)	(0.127)	(0.147)	(0.180)	(0.496)	(0.262)	(0.273)	(0.200)	(0.163)	(0.233)	(0.372)	(0.468)	(0.587)	(0.700)	(0.200)
Other European	0.020	1.578***	2.789***	1.793***	-0.394***	-0.479***	-0.302	-0.430*	0.651	-0.032	0.479	-0.376	-1.224***	-1.472***	0.578	-0.816	0.064	0.736	1.502***
	(0.040)	(0.065)	(0.063)	(0.056)	(0.117)	(0.137)	(0.202)	(0.259)	(0.696)	(0.347)	(0.390)	(0.304)	(0.221)	(0.356)	(0.425)	(0.530)	(0.639)	(0.499)	(0.134)
NZ Māori	0.155*	-0.755***	0.921***	5.040***	-0.343***	0.246**	-0.493***	0.115	0.493	0.272	1.078***	-0.917***	-0.201	-0.657***	0.990**	0.475	-0.235	-1.637*	0.389**
	(0.087)	(0.085)	(0.0950)	(0.070)	(0.126)	(0.119)	(0.147)	(0.179)	(0.509)	(0.249)	(0.277)	(0.315)	(0.195)	(0.252)	(0.403)	(0.662)	(0.847)	(0.947)	(0.180)
Samoan	-0.249**	-1.410***	0.470***	-1.054***	6.672***	-0.568***	-0.327	-0.430	1.877***	-0.471	1.007***	-1.329**	-0.876***	-1.330***	-1.364	-1.130	-1.560**	-0.804	0.100
	(0.101)	(0.060)	(0.099)	(0.074)	(0.087)	(0.192)	(0.208)	(0.265)	(0.501)	(0.321)	(0.375)	(0.550)	(0.225)	(0.371)	(0.965)	(1.071)	(0.721)	(0.588)	(0.295)
Cook Island Māori	-0.228***	-1.155***	0.693***	-1.073***	-0.377**	6.704***	-0.532**	-0.058	-0.104	-0.136	1.712***	-1.054	-1.057***	-1.634***	-0.270	0.591	0.483	0.796	0.843***
	(0.080)	(0.074)	(0.151)	(0.154)	(0.173)	(0.116)	(0.208)	(0.225)	(0.721)	(0.330)	(0.338)	(0.662)	(0.391)	(0.511)	(0.772)	(0.714)	(1.033)	(0.907)	(0.250)
Tongan	-0.083	-1.428***	0.477***	-1.136***	-0.789***	-0.657***	6.742***	-0.457	0.796	-0.015	0.024	-1.487***	-1.099***	-2.410***	0.101	0.981			0.187
	(0.097)	(0.112)	(0.181)	(0.124)	(0.204)	(0.247)	(0.126)	(0.419)	(0.633)	(0.524)	(0.544)	(0.555)	(0.397)	(0.620)	(0.711)	(0.738)			(0.415)
Niuean	-0.261**	-0.929***	0.479**	-0.430***	-0.484*	-0.211	-0.031	7.456***	0.062	0.689*	-0.152	-1.936**	-0.221	-2.176***		0.201	0.684		0.536
	(0.121)	(0.118)	(0.188)	(0.100)	(0.256)	(0.236)	(0.312)	(0.139)	(0.922)	(0.418)	(0.552)	(0.752)	(0.385)	(0.613)		(1.127)	(0.632)		(0.397)
Tokelauan	-0.129	-0.860*	0.564	-1.144***	-0.398		-1.937	0.798	8.803***		2.761***		0.641						
	(0.261)	(0.499)	(0.771)	(0.357)	(0.968)		(1.199)	(0.744)	(0.634)		(0.679)		(0.819)						
Fijian	-0.073	-0.595***	0.866***	-0.557***	-0.550	-0.115	0.431	0.231		7.149***	2.054***		0.470	0.972**	1.832*			0.307	2.130***
	(0.124)	(0.200)	(0.264)	(0.199)	(0.456)	(0.439)	(0.660)	(0.625)		(0.223)	(0.701)		(0.778)	(0.478)	(1.084)			(0.902)	(0.365)
Other PI	-0.703***	-0.105	1.128***	-1.166***	-0.896	-0.020	-0.606	-0.419	-0.538	2.604**	6.390***		0.978	-0.536	1.779**				1.939***
	(0.210)	(0.327)	(0.391)	(0.421)	(1.061)	(0.503)	(1.513)	(0.322)	(1.185)	(1.016)	(0.363)		(1.330)	(0.602)	(0.889)				(0.551)
SE Asian	-0.224*	-1.509***	0.035	-1.359***	-3.143***	-1.252*	-0.837*	-1.111				6.623***	0.746	-0.937**	0.240	1.054			2.863***
	(0.115)	(0.145)	(0.256)	(0.294)	(0.934)	(0.642)	(0.494)	(0.800)				(0.194)	(0.468)	(0.429)	(1.057)	(0.859)			(0.232)
Chinese	-0.0897	-1.741***	0.350**	-0.801***	-0.780***	-0.811***	-1.740***	-0.536	-0.139	-0.618	-0.0403	0.870*	6.399***	-1.830***	-0.385	1.105	-0.0796		1.445***

	(0.108)	(0.118)	(0.165)	(0.124)	(0.172)	(0.252)	(0.433)	(0.372)	(0.856)	(0.448)	(0.541)	(0.446)	(0.129)	(0.619)	(1.242)	(1.000)	(0.915)		(0.243)
Indian	-0.106	-2.091***	-0.099	-1.288***	-1.220***	-1.981***	-1.762***	-0.631*	0.443	-0.371	-0.410	-0.181	-1.679***	7.827***	0.335	-0.137	0.641	1.196	1.034***
	(0.105)	(0.122)	(0.239)	(0.139)	(0.337)	(0.410)	(0.462)	(0.358)	(1.180)	(0.651)	(0.810)	(0.491)	(0.604)	(0.170)	(0.843)	(1.078)	(1.343)	(1.031)	(0.380)
Other Asian	0.104	-1.563***	-0.067	-0.793**	-1.658	-0.841***	0.432	0.386					-1.765***	-2.185	8.719***				1.274***
	(0.198)	(0.266)	(0.576)	(0.361)	(1.627)	(0.326)	(0.661)	(1.081)					(0.457)	(1.959)	(0.373)				(0.494)
Middle Eastern	-0.170	-0.310	-0.112	-2.215***	0.994	-0.023	-0.023					-1.023	1.006			8.182***		2.912**	
	(0.185)	(0.272)	(0.571)	(0.617)	(1.011)	(1.041)	(1.016)					(0.915)	(1.101)			(0.580)		(1.234)	
Latin American	0.289	-0.147	2.126***	-2.323**	-2.250**	0.619		-0.298					-0.118	0.841			8.550***		2.093**
	(0.338)	(0.418)	(0.592)	(1.093)	(1.050)	(1.023)		(0.591)					(0.668)	(1.091)			(0.543)		(1.051)
African	-0.456	-0.734	1.071**	-1.097*	-0.667	1.057				1.300***								8.460***	1.478*
	(0.284)	(0.554)	(0.537)	(0.576)	(0.485)	(0.927)				(0.484)								(0.594)	(0.841)
ONFD	0.225	0.290	0.586	-0.369	1.081**	-1.591		1.114		1.011***	1.098	3.092***	0.502	-0.802	0.395	1.089		1.094	4.474***
	(0.184)	(0.266)	(0.410)	(0.458)	(0.443)	(1.001)		(0.975)		(0.357)	(0.751)	(0.469)	(0.948)	(2.145)	(0.737)	(2.231)		(0.690)	(0.284)

Notes: Each column reports one logistic regression. Tables A1, A2 and A3 are reporting results from the same regressions. For easy readability the results have been split across three different tables according to blocks of explanatory variables. Separate coefficients are estimated for adults and children.
The 'Other' and 'Not further defined' ethnic groups (among those who are European, Asian or Pacific Islanders) have been combined into one group 'ONFD'. Hence the analysis distinguishes 18 Level 2 ethnic groups instead of the 21 listed in Table 1.
The table reports logistic regression coefficients. Column (1) refers to the geographic mobility regression. Columns (2) to (19) refer to the ethnic mobility regressions (one for each of the 18 ethnic groups). Clustered standard errors are shown in parenthesis.
*p<0.1; **p<0.05; ***p<0.01
Blank cells refer to the cases in which variables have been omitted due to perfect collinearity, usually due to small cell sizes.
The number of observations and goodness of fit measures are given at the bottom of Table A3.

Table A2: Clustered Logistic Regression of Locational and Ethnic Transition-Effect of Individual-Level Variable

Variables	(1) Moved	(2) NZ European	(3) Other European	(4) NZ Māori	(5) Samoan	(6) Cook Island Māori	(7) Tongan	(8) Niuean	(9) Tokelauan	(10) Fijian	(11) Other PI	(12) SE Asian	(13) Chinese	(14) Indian	(15) Other Asian	(16) Middle Eastern	(17) Latin American	(18) African	(19) ONFD
Adults (aged 18 and over)																			
Sex	-0.044***	-0.027**	-0.039***	-0.115***	-0.018	-0.148**	0.036	0.031	0.082	-0.057	-0.390***	-0.211**	-0.033	0.151**	0.136	0.014	-0.171	0.121	-0.157***
	(0.007)	(0.014)	(0.014)	(0.021)	(0.050)	(0.074)	(0.071)	(0.082)	(0.219)	(0.081)	(0.105)	(0.083)	(0.049)	(0.066)	(0.158)	(0.099)	(0.205)	(0.156)	(0.033)
Age	-0.025***	0.013***	-0.012***	-0.023***	-0.036***	-0.035***	-0.039***	-0.029***	-0.039***	-0.024***	-0.027***	-0.035***	-0.025***	-0.034***	-0.033***	-0.008**	-0.039***	-0.028***	-0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.008)	(0.004)	(0.005)	(0.003)	(0.002)	(0.003)	(0.005)	(0.003)	(0.010)	(0.007)	(0.001)
NZ Born	-0.045	2.288***	-1.552***	3.256***	-1.181***	-0.627***	-1.501***	-0.501***	-0.270	-1.066***	-0.517***	-2.924***	-1.005***	-1.776***	-2.141***	-1.484***	-2.438***	-1.187***	-1.968***
	(0.038)	(0.038)	(0.035)	(0.052)	(0.077)	(0.119)	(0.132)	(0.122)	(0.290)	(0.136)	(0.180)	(0.142)	(0.087)	(0.131)	(0.332)	(0.184)	(0.371)	(0.280)	(0.076)
Years at address	-0.044***																		
	(0.002)																		
Children and Adolescents (aged 0-17)																			
Sex	-0.029*	-0.016	-0.121**	-0.064**	-0.024	-0.18**	-0.022	-0.030	-0.147	0.146	-0.326	-0.039	0.118	-0.259**	-0.019	-0.560**	-0.186	-0.247	0.111
	(0.016)	(0.029)	(0.050)	(0.030)	(0.069)	(0.089)	(0.086)	(0.117)	(0.256)	(0.180)	(0.213)	(0.133)	(0.089)	(0.130)	(0.246)	(0.236)	(0.423)	(0.392)	(0.115)
Age	-0.089***	-0.007	-0.018	-0.032***	-0.021	0.012	-0.062***	-0.013	0.014	0.002	-0.089*	-0.074**	-0.005	-0.028	-0.078	0.028	0.104	-0.043	-0.070***
	(0.006)	(0.007)	(0.0113)	(0.008)	(0.014)	(0.017)	(0.018)	(0.026)	(0.061)	(0.040)	(0.046)	(0.030)	(0.019)	(0.029)	(0.056)	(0.058)	(0.092)	(0.083)	(0.024)
NZ Born	-0.320***	1.380***	-1.078***	1.043***	-0.060	0.071	-0.282*	0.093	-0.115	-0.149	-1.058***	-0.984***	-0.527***	-0.599**	-2.001***	-0.926***	-1.089*	-0.308	-0.907***
	(0.046)	(0.068)	(0.072)	(0.085)	(0.136)	(0.175)	(0.154)	(0.232)	(0.664)	(0.273)	(0.285)	(0.237)	(0.150)	(0.237)	(0.372)	(0.307)	(0.569)	(0.671)	(0.144)
Years at address	-0.013***																		
	(0.002)																		

Notes: See the notes at the bottom of Table A.1.

Table A3: Clustered Logistic Regression of Locational and Ethnic Transition-Effect of Neighbourhood-Level Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
Variables	**Moved**	**NZ European**	**Other European**	**NZ Māori**	**Samoan**	**Cook Island Māori**	**Tongan**	**Niuean**	**Tokelauan**	**Fijian**	**Other PI**	**SE Asian**	**Chinese**	**Indian**	**Other Asian**	**Middle Eastern**	**Latin American**	**African**	**ONFD**
									Adults (aged 18 and over)										
Entropy	0.828	-0.486	0.974***	0.595*	3.348***	2.006	1.720	0.493	4.857	-0.443	-2.507**	1.158	1.319*	1.439	4.316*	4.983***	-0.512	-1.758	-0.033
	(1.404)	(0.319)	(0.293)	(0.362)	(1.015)	(1.541)	(1.107)	(1.260)	(4.075)	(0.977)	(1.263)	(1.320)	(0.748)	(1.234)	(2.255)	(1.889)	(2.421)	(3.184)	(0.703)
NZ European Gr	0.009	-0.009***	-0.030***	-0.058***	-0.042***	-0.041***	-0.039***	-0.046***	-0.092***	-0.068***	-0.057***	-0.035***	-0.028***	-0.020***	-0.061***	-0.060***	-0.043***	-0.084***	-0.041***
	(0.006)	(0.001)	(0.001)	(0.002)	(0.004)	(0.006)	(0.005)	(0.006)	(0.022)	(0.006)	(0.008)	(0.006)	(0.003)	(0.006)	(0.010)	(0.008)	(0.012)	(0.013)	(0.003)
Other European Gr	0.015	-0.003	-0.029***	-0.051***	-0.032***	-0.048***	-0.047***	-0.055***	-0.006	0.003	-0.009	-0.003	-0.034***	-0.068***	-0.065***	-0.076***	-0.101**	0.024	-0.022***
	(0.016)	(0.003)	(0.004)	(0.005)	(0.009)	(0.012)	(0.014)	(0.015)	(0.050)	(0.016)	(0.021)	(0.016)	(0.010)	(0.013)	(0.025)	(0.018)	(0.046)	(0.028)	(0.007)
NZ Māori Gr	0.006	-0.025***	-0.025***	0.007*	-0.019**	-0.017	-0.026**	-0.021	-0.091**	-0.034***	0.0002	-0.010	-0.044***	-0.021*	-0.062**	-0.113***	-0.069**	-0.031	-0.013*
	(0.017)	(0.003)	(0.003)	(0.004)	(0.009)	(0.014)	(0.012)	(0.013)	(0.038)	(0.011)	(0.016)	(0.013)	(0.010)	(0.011)	(0.025)	(0.019)	(0.029)	(0.033)	(0.007)
Samoan Gr	0.004	-0.013**	-0.030***	-0.045***	-0.012	-0.029*	-0.023*	-0.045***	-0.053	-0.048***	-0.045**	-0.033	-0.010	-0.018	-0.093***	-0.093*	-0.058	-0.0326	-0.029***
	(0.029)	(0.005)	(0.005)	(0.005)	(0.011)	(0.015)	(0.013)	(0.014)	(0.040)	(0.018)	(0.021)	(0.021)	(0.011)	(0.019)	(0.033)	(0.049)	(0.043)	(0.046)	(0.011)
Cook Island Māori Gr	0.001	-0.044***	-0.054***	-0.093***	-0.071***	-0.060**	-0.091***	-0.034	-0.090	-0.056*	-0.176***	-0.025	-0.097***	-0.081**	-0.199***	0.0189	0.0241	-0.078	-0.080***
	(0.052)	(0.009)	(0.011)	(0.012)	(0.016)	(0.024)	(0.025)	(0.030)	(0.070)	(0.029)	(0.047)	(0.041)	(0.023)	(0.034)	(0.072)	(0.056)	(0.058)	(0.069)	(0.020)
Tongan Gr	0.017	-0.020***	-0.047***	-0.049***	-0.076***	-0.006	0.006	-0.107***	-0.153*	-0.064**	0.038	-0.121***	-0.023	-0.040	-0.024	-0.204***	-0.224***	-0.010	-0.046**
	(0.048)	(0.007)	(0.009)	(0.010)	(0.021)	(0.026)	(0.032)	(0.040)	(0.085)	(0.031)	(0.047)	(0.037)	(0.025)	(0.037)	(0.054)	(0.058)	(0.087)	(0.065)	(0.019)
Niuean Gr	0.016	-0.005	0.005	0.002	-0.050	-0.043	0.019	0.092**	-0.023	-0.063	0.050	0.134**	-0.038	-0.025	0.002	0.010	0.091	-0.031	-0.019
	(0.077)	(0.012)	(0.013)	(0.019)	(0.032)	(0.039)	(0.034)	(0.042)	(0.132)	(0.054)	(0.069)	(0.059)	(0.040)	(0.060)	(0.105)	(0.087)	(0.140)	(0.101)	(0.030)
Tokelauan Gr	0.318	-0.017	0.033	0.026	0.380**	0.072	0.124	0.143	1.091	0.615*	0.756**	-0.352	0.129	0.073	0.207	-0.525	-0.891	0.211	0.150
	(0.461)	(0.064)	(0.078)	(0.094)	(0.160)	(0.198)	(0.193)	(0.293)	(0.773)	(0.318)	(0.368)	(0.394)	(0.218)	(0.344)	(0.692)	(0.544)	(0.841)	(0.616)	(0.166)
Fijian Gr	0.007	-0.022	0.140***	0.105**	0.205**	0.081	0.240**	0.184	-0.309	0.549***	0.029	0.039	0.042	0.176	0.210	0.118	1.009**	0.111	0.025
	(0.215)	(0.0295)	(0.0356)	(0.0466)	(0.0805)	(0.111)	(0.110)	(0.127)	(0.362)	(0.179)	(0.194)	(0.184)	(0.113)	(0.137)	(0.290)	(0.185)	(0.417)	(0.287)	(0.0805)
Other PI Gr	-0.110	-0.024	0.209***	0.169**	-0.014	0.182	0.197	0.127	-0.216	-0.085	0.826***	-0.322	-0.059	0.080	0.409	0.371	0.990	-0.478	0.0127
	(0.360)	(0.052)	(0.065)	(0.077)	(0.152)	(0.156)	(0.208)	(0.265)	(0.645)	(0.308)	(0.320)	(0.327)	(0.203)	(0.258)	(0.440)	(0.416)	(0.653)	(0.459)	(0.122)
SE Asian Gr	0.063	0.008	-0.007	0.034	-0.091	0.023	0.066	0.001	0.113	-0.033	0.154	0.295***	-0.045	-0.067	-0.011	-0.008	0.380**	-0.281	-0.106**
	(0.126)	(0.017)	(0.021)	(0.031)	(0.057)	(0.079)	(0.069)	(0.086)	(0.178)	(0.110)	(0.119)	(0.082)	(0.070)	(0.072)	(0.159)	(0.127)	(0.156)	(0.186)	(0.048)
Chinese Gr	0.015	-0.018***	-0.042***	-0.076***	-0.072***	-0.057**	-0.0388	-0.085***	-0.137	-0.029	-0.083**	0.0002	0.042***	-0.036	-0.047	-0.047	0.057	-0.062	-0.026**
	(0.035)	(0.005)	(0.006)	(0.010)	(0.022)	(0.027)	(0.029)	(0.027)	(0.099)	(0.030)	(0.037)	(0.031)	(0.015)	(0.024)	(0.047)	(0.033)	(0.076)	(0.061)	(0.013)
Indian Gr	0.024	-0.036***	-0.025**	-0.051***	-0.021	-0.055	-0.024	0.062**	-0.020	-0.0001	-0.051	-0.022	-0.050**	0.067**	-0.179***	-0.176***	-0.123	0.058	-0.007
	(0.044)	(0.007)	(0.010)	(0.010)	(0.020)	(0.036)	(0.031)	(0.029)	(0.090)	(0.032)	(0.050)	(0.038)	(0.024)	(0.029)	(0.060)	(0.057)	(0.089)	(0.066)	(0.018)

Other Asian Gr	0.172	-0.045	-0.015	-0.169***	0.276***	0.160	0.214	0.298*	-0.722	-0.031	0.059	-0.341*	0.011	0.090	0.232	0.110	-0.159	0.819***	0.017
	(0.236)	(0.034)	(0.038)	(0.061)	(0.099)	(0.148)	(0.131)	(0.179)	(0.516)	(0.180)	(0.267)	(0.204)	(0.109)	(0.131)	(0.266)	(0.242)	(0.398)	(0.302)	(0.077)
Middle Eastern Gr	0.243	-0.155***	0.230***	0.040	0.090	0.056	0.168	0.049	0.101	-0.390	0.101	0.107	-0.116	0.142	0.527***	0.916***	0.198	0.057	0.041
	(0.302)	(0.048)	(0.049)	(0.066)	(0.101)	(0.102)	(0.112)	(0.233)	(0.222)	(0.265)	(0.129)	(0.178)	(0.173)	(0.184)	(0.175)	(0.153)	(0.527)	(0.359)	(0.096)
Latin American Gr	-0.098	0.161*	-0.443***	-0.126	0.050	0.293	-0.130	-0.717	-0.966	-0.055	-0.623	-0.126	0.211	0.029	-0.738	0.872	1.827*	1.137	0.132
	(0.636)	(0.092)	(0.114)	(0.119)	(0.313)	(0.348)	(0.421)	(0.483)	(1.091)	(0.499)	(0.685)	(0.486)	(0.337)	(0.495)	(0.852)	(0.816)	(0.939)	(0.837)	(0.196)
African Gr	0.083	0.260***	-0.107	-0.026	0.002	0.544	0.441	0.558	2.148***	1.460***	-0.472	-0.245	0.005	-0.538	0.762	-0.355	-0.505	0.548	0.047
	(0.538)	(0.081)	(0.114)	(0.134)	(0.268)	(0.346)	(0.373)	(0.446)	(0.781)	(0.426)	(0.438)	(0.463)	(0.311)	(0.350)	(0.793)	(0.510)	(1.061)	(0.541)	(0.173)
ONFD Gr	0.180	0.063	-0.061	-0.044	0.016	-0.021	0.248	0.541*	0.897	0.169	0.391	-0.460	0.004	0.023	-0.149	-0.507	-0.737	0.699	0.212
	(0.400)	(0.060)	(0.074)	(0.089)	(0.179)	(0.234)	(0.229)	(0.315)	(0.597)	(0.397)	(0.386)	(0.325)	(0.191)	(0.239)	(0.650)	(0.405)	(1.070)	(0.590)	(0.130)
Observations	403,200	403,200	403,200	403,200	403,200	402,900	402,600	400,500	399,300	403,200	401,800	402,600	403,100	402,800	402,700	399,400	380,500	401,700	402,800
Pseudo R-squared	0.631	0.652	0.291	0.695	0.852	0.831	0.834	0.842	0.791	0.464	0.610	0.701	0.827	0.869	0.815	0.227	0.202	0.211	0.193

Children and Adolescents (aged 0-17)

Entropy	0.214	-2.225***	-1.102	0.899	1.949	3.464***	3.458*	1.444	-6.839***	-1.419	-1.260	1.346	0.677	7.174***	6.333	11.79**	6.622	-3.518	4.387**
	(1.420)	(0.711)	(0.707)	(0.555)	(1.413)	(1.283)	(1.990)	(1.954)	(1.732)	(2.569)	(2.854)	(2.501)	(1.212)	(2.354)	(4.524)	(5.423)	(5.322)	(10.54)	(1.891)
NZ European Gr	0.004	-0.005**	-0.030***	-0.047***	-0.053***	-0.0531***	-0.041***	-0.060***	-0.068***	-0.062***	-0.040**	-0.054***	-0.001	-0.054***	-0.041**	-0.072***	-0.073***	-0.100***	-0.059***
	(0.007)	(0.003)	(0.003)	(0.003)	(0.006)	(0.005)	(0.008)	(0.010)	(0.020)	(0.013)	(0.016)	(0.012)	(0.031)	(0.007)	(0.020)	(0.024)	(0.022)	(0.036)	(0.008)
Other European Gr	0.025	0.018***	-0.033***	-0.024***	-0.025	-0.075***	-0.078***	-0.073***	-0.010	-0.055	-0.029	-0.020	-0.0003	-0.069**	-0.179***	-0.163***	-0.151**	0.050	-0.033*
	(0.017)	(0.007)	(0.010)	(0.007)	(0.015)	(0.016)	(0.020)	(0.025)	(0.075)	(0.039)	(0.031)	(0.025)	(0.031)	(0.028)	(0.048)	(0.051)	(0.060)	(0.072)	(0.017)
NZ Māori Gr	0.023	-0.020***	-0.019**	0.029***	-0.005	-0.051***	-0.058***	-0.041**	0.061**	-0.021	-0.032	0.001	-0.007	-0.068***	-0.109**	-0.172***	-0.126	0.0185	-0.052***
	(0.018)	(0.005)	(0.009)	(0.006)	(0.012)	(0.012)	(0.018)	(0.020)	(0.030)	(0.022)	(0.032)	(0.026)	(0.024)	(0.025)	(0.044)	(0.044)	(0.079)	(0.086)	(0.019)
Samoan Gr	0.002	-0.005	-0.021	-0.029***	0.005	-0.042***	-0.033**	-0.048**	0.036	-0.054	-0.047	-0.041	-0.004	-0.031	0.011	-0.229**	0.117	0.039	-0.098***
	(0.031)	(0.010)	(0.015)	(0.008)	(0.011)	(0.015)	(0.016)	(0.022)	(0.052)	(0.047)	(0.055)	(0.027)	(0.032)	(0.035)	(0.071)	(0.108)	(0.088)	(0.080)	(0.031)
Cook Island Māori Gr	-0.007	-0.026	-0.061**	-0.112***	-0.132***	-0.061*	-0.072*	-0.097**	-0.275**	-0.164**	-0.108	-0.144***	-0.001	-0.082	-0.211**	-0.119	-0.445**	-0.293	-0.106*
	(0.054)	(0.017)	(0.024)	(0.018)	(0.024)	(0.034)	(0.041)	(0.042)	(0.131)	(0.081)	(0.079)	(0.054)	(0.055)	(0.056)	(0.092)	(0.138)	(0.191)	(0.181)	(0.057)
Tongan Gr	0.014	-0.017	-0.078**	-0.054***	-0.054*	-0.056*	0.039	-0.092**	0.143	-0.111	0.119**	0.057	0.017	-0.133*	-0.144	-0.131	-0.212	-0.454**	-0.075
	(0.049)	(0.015)	(0.031)	(0.016)	(0.028)	(0.034)	(0.033)	(0.042)	(0.099)	(0.071)	(0.056)	(0.061)	(0.047)	(0.079)	(0.110)	(0.128)	(0.147)	(0.199)	(0.050)
Niuean Gr	0.003	-0.009	0.079**	-0.006	-0.005	-0.036	-0.098**	0.121	-0.057	0.156	0.026	-0.035	-0.039	-0.121	-0.037	0.283	-0.074	0.368*	-0.044
	(0.081)	(0.022)	(0.038)	(0.022)	(0.044)	(0.041)	(0.044)	(0.075)	(0.135)	(0.156)	(0.109)	(0.090)	(0.057)	(0.098)	(0.133)	(0.193)	(0.207)	(0.204)	(0.093)
Tokelauan Gr	0.283	0.023	-0.071	-0.006	-0.058	0.122	-0.307	0.046	0.802	0.479	-0.456	0.201	-0.679*	-0.143	1.428	-1.315	-0.572	-0.001	0.315
	(0.487)	(0.124)	(0.198)	(0.135)	(0.257)	(0.311)	(0.319)	(0.403)	(1.449)	(0.557)	(0.647)	(0.580)	(0.387)	(0.494)	(0.920)	(1.307)	(1.402)	(1.169)	(0.308)
Fijian Gr	0.106	0.063	-0.031	0.029	-0.071	-0.044	0.399**	0.248	-0.431	0.545**	-0.302	0.167	-0.032	0.179	-0.571	-0.108	-0.521	0.195	-0.245
	(0.226)	(0.067)	(0.112)	(0.082)	(0.131)	(0.141)	(0.170)	(0.221)	(0.677)	(0.235)	(0.313)	(0.274)	(0.184)	(0.263)	(0.456)	(0.513)	(0.808)	(0.596)	(0.239)

Other PI Gr	-0.055	-0.125	0.093	0.255**	-0.137	0.193	0.597**	0.459	-0.750	-0.171	0.448	0.070	0.070	-0.613	0.115	-0.839	-0.313	-0.240	0.245
	(0.383)	(0.103)	(0.163)	(0.120)	(0.214)	(0.260)	(0.287)	(0.333)	(1.054)	(0.589)	(0.504)	(0.478)	(0.349)	(0.459)	(1.144)	(1.086)	(1.023)	(1.063)	(0.349)
SE Asian Gr	0.033	0.055	0.101*	-0.010	-0.067	-0.068	0.069	0.082	-0.012	0.081	-0.204	-0.130	0.028	-0.431***	-0.026	-0.127	0.009	-0.066	-0.242
	(0.133)	(0.034)	(0.052)	(0.034)	(0.067)	(0.079)	(0.094)	(0.104)	(0.319)	(0.174)	(0.200)	(0.128)	(0.090)	(0.141)	(0.238)	(0.265)	(0.659)	(0.432)	(0.150)
Chinese Gr	0.015	-0.011	-0.017	-0.065***	-0.042	-0.078**	-0.053*	-0.081*	-0.084	0.020	-0.111	0.013	0.105**	-0.060	-0.297***	-0.121*	-0.216	0.139	-0.077**
	(0.035)	(0.012)	(0.018)	(0.016)	(0.035)	(0.032)	(0.032)	(0.042)	(0.158)	(0.054)	(0.070)	(0.057)	(0.045)	(0.045)	(0.102)	(0.070)	(0.186)	(0.115)	(0.039)
Indian Gr	0.040	-0.025	-0.013	-0.015	-0.006	-0.033	-0.114**	-0.041	0.026	-0.126	0.008	-0.067	-0.009	-0.035	0.040	-0.251*	-0.344**	0.085	-0.168***
	(0.044)	(0.017)	(0.023)	(0.015)	(0.034)	(0.035)	(0.044)	(0.063)	(0.115)	(0.080)	(0.052)	(0.067)	(0.045)	(0.055)	(0.093)	(0.130)	(0.146)	(0.192)	(0.056)
Other Asian Gr	0.123	0.076	-0.046	-0.261**	0.068	-0.216	-0.012	0.374	-0.919	0.375	-0.660	-0.242	-0.124	-0.061	1.390***	-0.322	0.898	-0.612	0.241
	(0.240)	(0.073)	(0.125)	(0.112)	(0.172)	(0.190)	(0.198)	(0.232)	(1.320)	(0.338)	(0.531)	(0.301)	(0.168)	(0.234)	(0.414)	(0.748)	(0.758)	(0.844)	(0.203)
Middle Eastern Gr	0.036	-0.113	0.139	-0.107	-0.056	0.161	0.012	0.064	-0.482	0.224	0.472*	-0.206	0.186	0.280	0.445	0.559*	0.549	-0.186	0.153
	(0.275)	(0.104)	(0.127)	(0.121)	(0.178)	(0.129)	(0.126)	(0.283)	(0.497)	(0.372)	(0.280)	(0.292)	(0.167)	(0.242)	(0.271)	(0.327)	(0.586)	(1.000)	(0.199)
Latin American Gr	-0.110	-0.201	-0.507	-0.811***	-0.161	0.589	0.112	0.264	-1.651	-0.286	-0.035	0.149	-0.966*	-1.123	-1.894	-0.344	2.460**	-2.613	-1.247*
	(0.650)	(0.183)	(0.311)	(0.216)	(0.403)	(0.417)	(0.480)	(0.674)	(1.954)	(0.855)	(1.059)	(0.765)	(0.569)	(0.781)	(1.793)	(1.468)	(1.023)	(2.112)	(0.708)
African Gr	0.180	0.279	-0.286	0.255	0.562	0.909**	-0.117	0.817*	1.407	0.609	-0.436	0.0207	-0.910	-0.194	2.648***	-1.055	0.403	-0.0126	-0.371
	(0.573)	(0.182)	(0.274)	(0.207)	(0.472)	(0.409)	(0.527)	(0.461)	(1.537)	(1.142)	(0.939)	(0.792)	(0.612)	(0.633)	(0.977)	(0.895)	(1.612)	(1.280)	(0.519)
ONFD Gr	0.360	-0.172	-0.035	-0.030	0.086	0.160	-0.323	0.552	-0.630	0.591	-1.744**	-0.551	-0.009	0.448	0.262	0.758	-0.828	0.205	0.793**
	(0.427)	(0.122)	(0.194)	(0.154)	(0.259)	(0.290)	(0.321)	(0.438)	(0.859)	(0.826)	(0.870)	(0.474)	(0.359)	(0.424)	(1.032)	(0.992)	(2.066)	(1.090)	(0.368)
Observations	65,800	65,800	65,800	65,800	65,800	65,600	65,500	65,500	64,200	64,700	64,800	64,900	65,700	65,400	63,800	64,900	61,700	59,000	65,500
Pseudo R-squared	0.593	0.614	0.217	0.653	0.839	0.812	0.823	0.827	0.783	0.472	0.592	0.692	0.819	0.862	0.801	0.216	0.196	0.201	0.182

Notes: See the notes at the bottom of Table A.1.
Additionally. 'Gr' refers to group proportion. For example, 'Tongan Gr' refers to the share of those with Tongan ethnicity in the area unit the individual resides in.