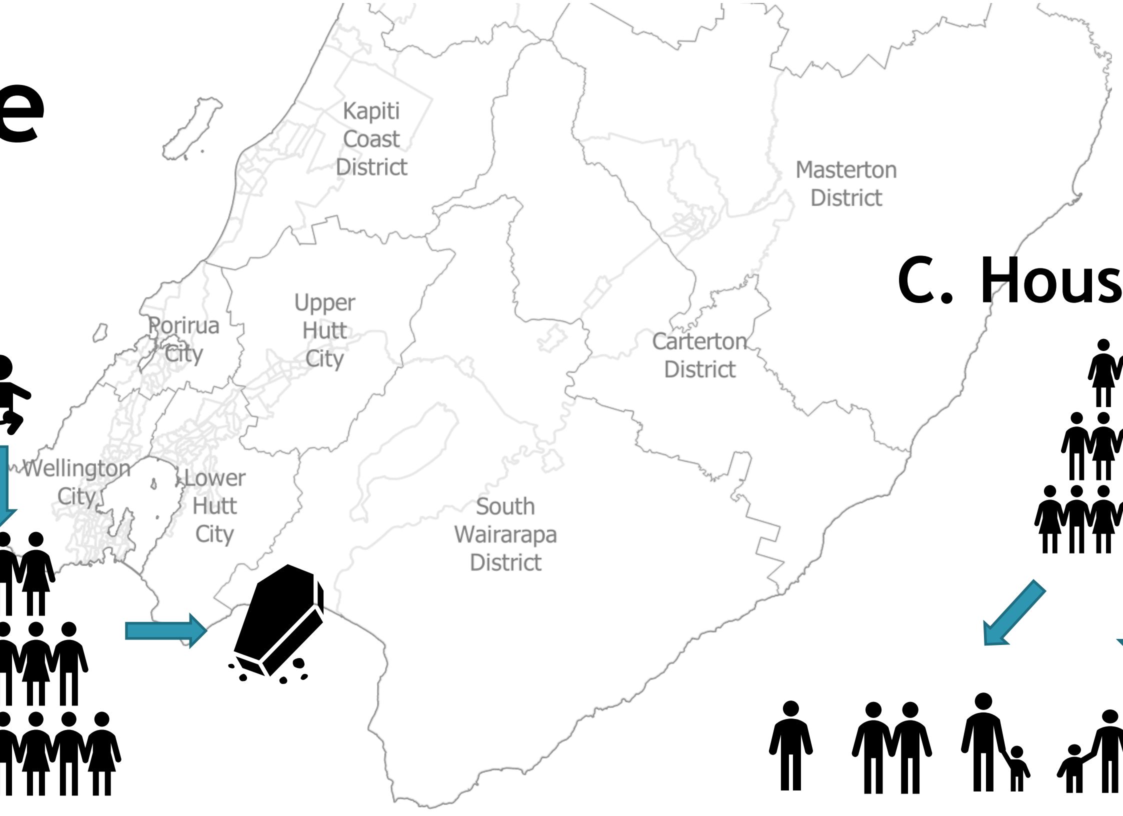
Wellington demographic and dwelling forecasts

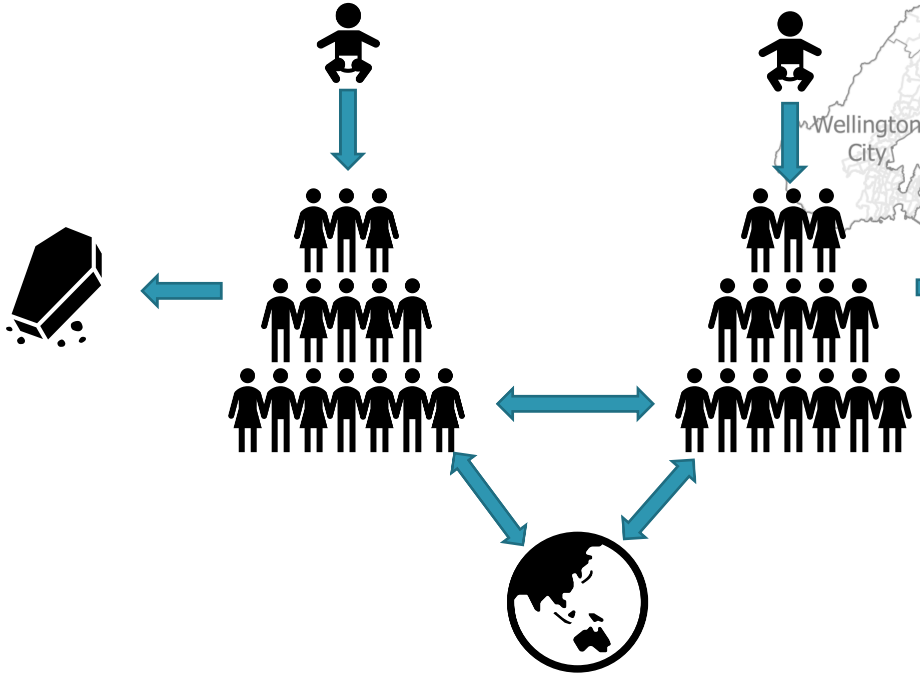
Methodology

SENSE PARTNERS
DATA LOGIC ACTION

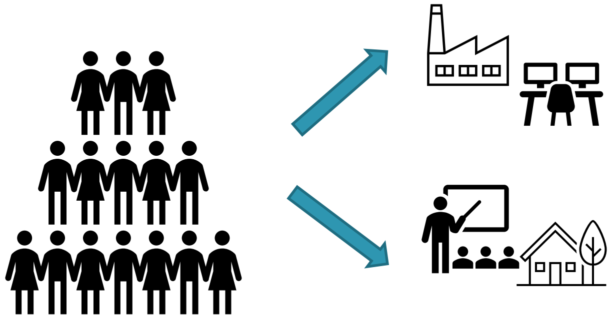
Summary scope



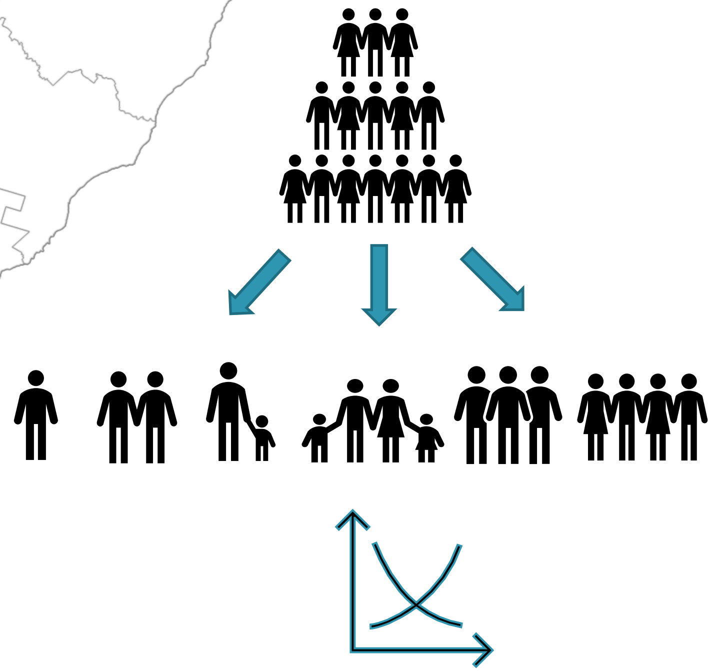
A. People



B. Work



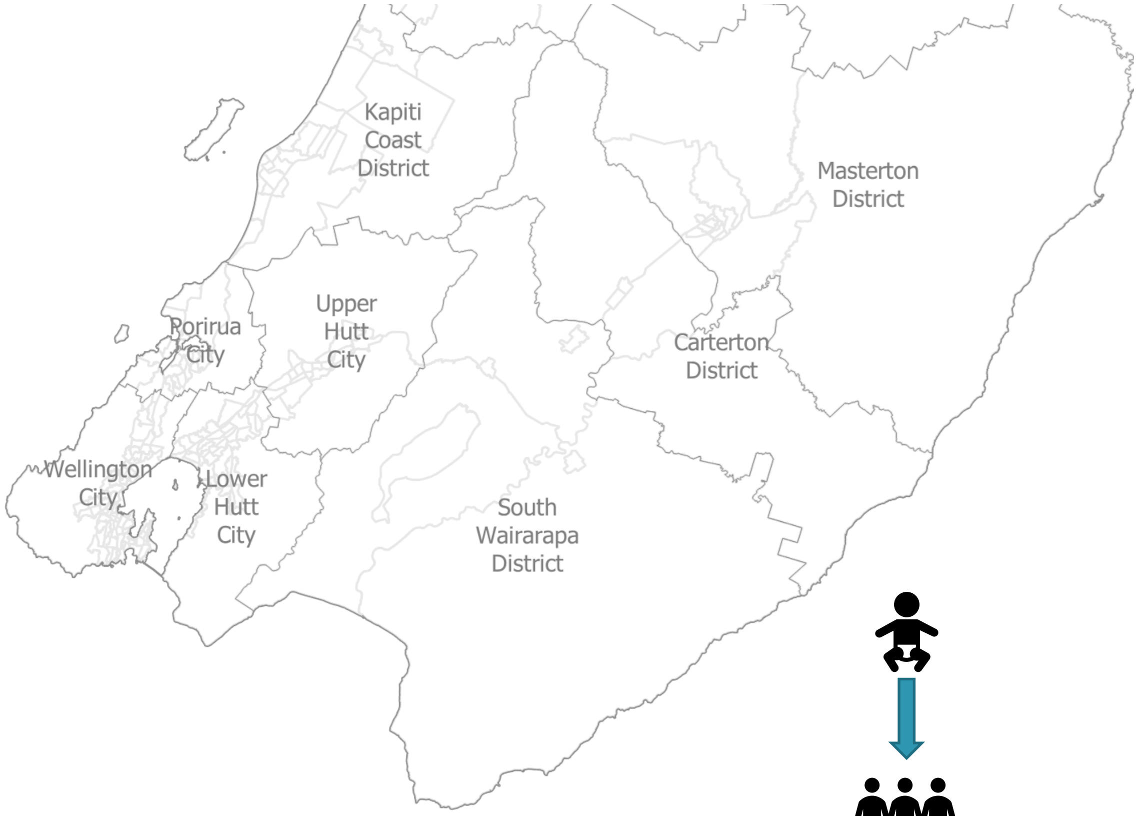
C. Households



D. Housing

Approach

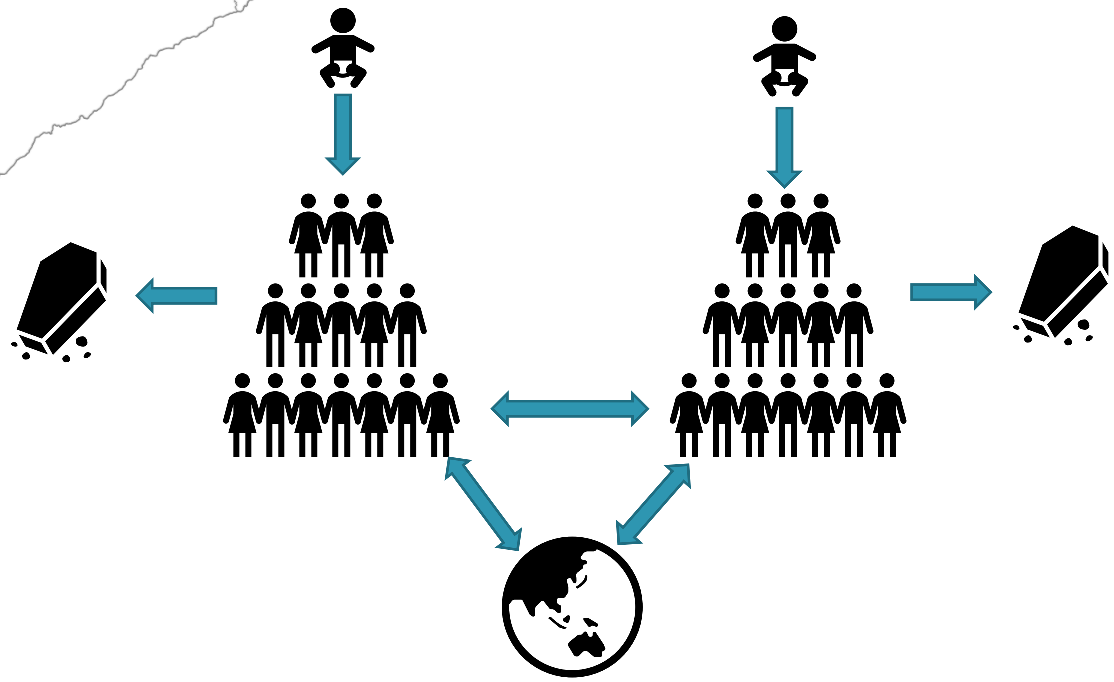
- The focus of our method is to use persistent and predictable structural and compositional characteristics of populations and economies to extrapolate future trends. The methods place a premium on respecting adding-up constraints (e.g. domestic migration must sum to zero) and consistency between forecasts. For this reason the model is a national model, with district details.
- The forecasts produced should be interpreted as potentials. There are a number of things that the forecasts do not take into account, such as national or local policy changes which can affect actual population and economic growth.
- To capture uncertainty around trends we conduct monte-carlo simulation, where inputs are varied randomly and repeatedly (500 times) to produce distributions over future values, rather than point estimates. This approach also helps to emphasise the considerable uncertainty that exists about the future and the extent to which this uncertainty grows the further out we look.
- The forecasts are based on 4 component models, as summarised in the earlier slide on scope. The modelling proceeds in a linear fashion through each of the models.



A. People

Demographic model

Fertility, mortality, migration



A. Model steps and components

- 500 sets of population forecasts are produced.
- In each set, population is forecast, sequentially, for each of the next 30 years following the steps set out below in each year.
- In each year external model input values are perturbed with random variation in values to simulate uncertainty*.

0

National population at the start of the current year ($N_{r,a,s,t}$)
by district (r) by age ($a \in \{0,1,2 \ldots 99\}$) by sex ($s \in \{1,2\}$) by year ($t$)

1

External model inputs (forecasts and parameter values, counts in upper case, rates in lower case)

- Forecast age-specific district fertility rates $f_{r,a,t}$
- Forecast age- and sex-specific district mortality rates $x_{r,a,s,t}$
- International age-specific immigration by district by age $M^{w}_{r,a,t}$
- International age-specific emigration rates by district $e^{w}_{r,a,t}$
- Rates of domestic migration from origin to destination by age $e^{i}_{o,d,a,t}$
- Forecast age- and sex-specific district labour force participation rates $l_{r,a,s,t}$

2

Internal model calculations of population change during the current year

- Births by district by sex $\mathrm{B}_{r,a=0,s,t} = \frac{B_s}{B_{\bar{s}}} \sum_a f_{r,a,t} \cdot N_{r,a,s=2,t}$
- Deaths by district by age $\mathrm{X}_{r,a,s,t} = x_{r,a,s,t} \cdot N_{r,a,s,t}$
- Inflows from the world by district by age and sex $M^{w}_{r,a,s,t} = \frac{M^{w}_{r,a,t}}{2}$
- Outflows to the world by district by age $E^{w}_{r,a,s,t} = N_{r,a,t} \cdot e^{w}_{r,a,t}$
- Internal/domestic inflows by district by age and sex $M^{i}_{r,a,s,t} = \sum_{o} N_{r=o,a,s,t} e^{i}_{o=r,d,a}$
- Internal/domestic outflows by district by age and sex $E^{i}_{r,a,s,t} = \sum_{d} N_{r,a,t} e^{i}_{o=r,a,d,t}$
- Labour force by district by age by sex $\mathrm{L}_{r,a,s,t} = l_{r,a,s,t} \cdot N_{r,a,s,t}$

3

Updated national population at the end of the current year and start of next year

$$N_{r,a,s,t+1} = N_{r,a,s,t} + B_{r,a,s,t} - X_{r,a,s,t} + M^{w}_{r,a,s,t} + M^{i}_{r,a,s,t} + E^{w}_{r,a,s,t} + E^{i}_{r,a,s,t}$$

*For simplicity of presentation, this overview excludes depiction of the random variations incorporated

A.1. Fertility

- Births (B) by district (r) by sex (s) are forecast with

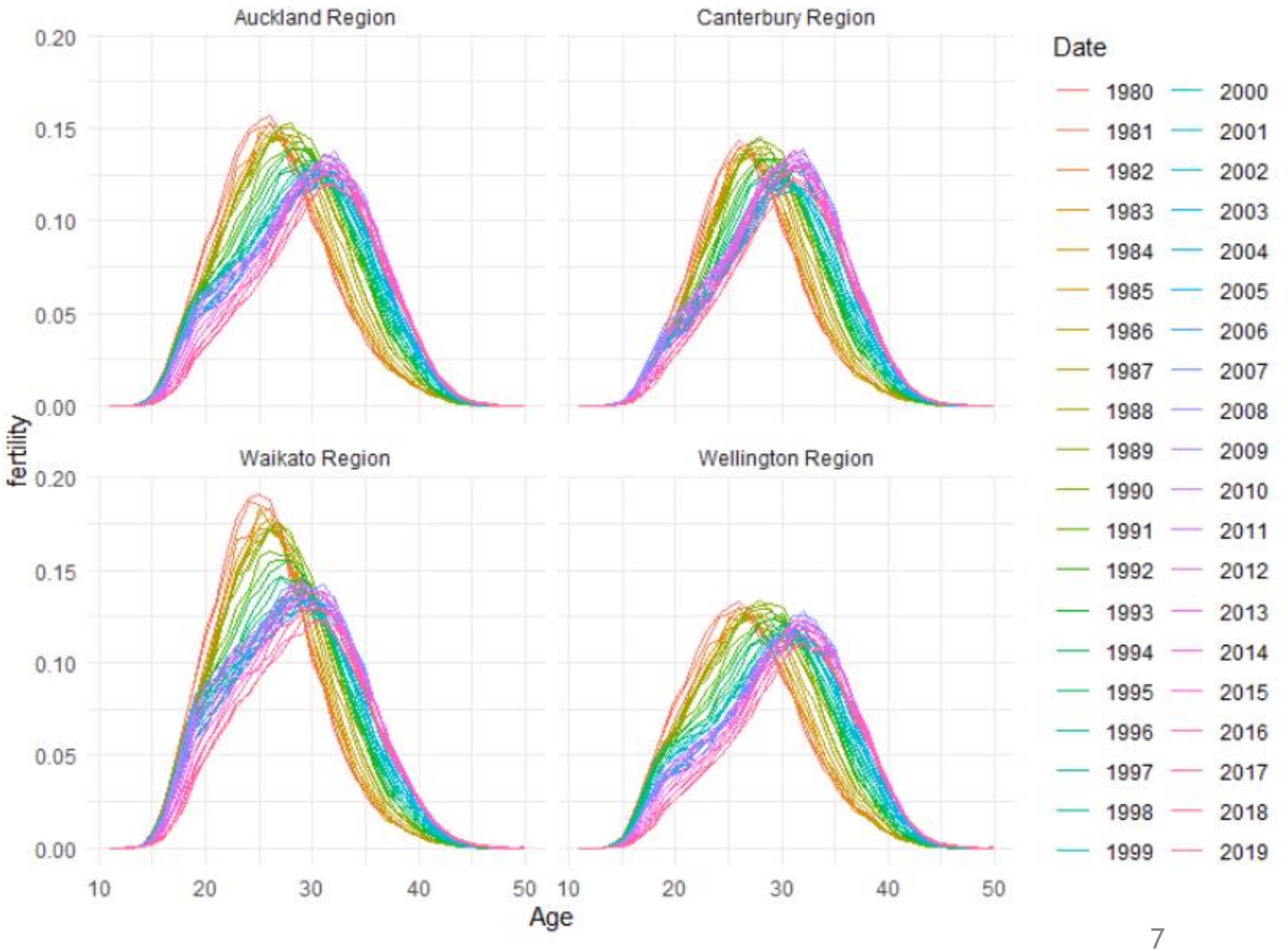
$$B_{r,a=0,s,t} = \frac{B_s}{\sum_s B_s} \sum_a f_{r,a,t} N_{r,a,s=2,t} + \epsilon_t$$

Where births (i.e. age (a) is 0) are a function of

- a fixed birth ratio $\frac{B_s}{\sum_s B_s}$ of 0.513 for males and 0.487 for females
- the size of the population by age ($N_{r,a,s=2,t}$) of females ($s = 2$) and forecast age-specific fertility rates ($f_{r,a,t}$)
- random variation in total births from year to year, using $\epsilon_t \sim N(0, \sigma)$ where the standard error (σ) is estimated from a model used to predict the district births using estimated age specific regional fertility rates.

A.1.1 Fertility data

- National age-specific fertility rates (Stats NZ, 1980-2019)
- Regional age-group-specific fertility rates (Stats NZ, 1996, 2001, 2006, 2013)
- Regional age-specific fertility rates are estimated using polynomials fitted to the ratio of regional observed age-group rates and national observed age-group rates (see Figure at right for sample of estimated age-specific regional rates)

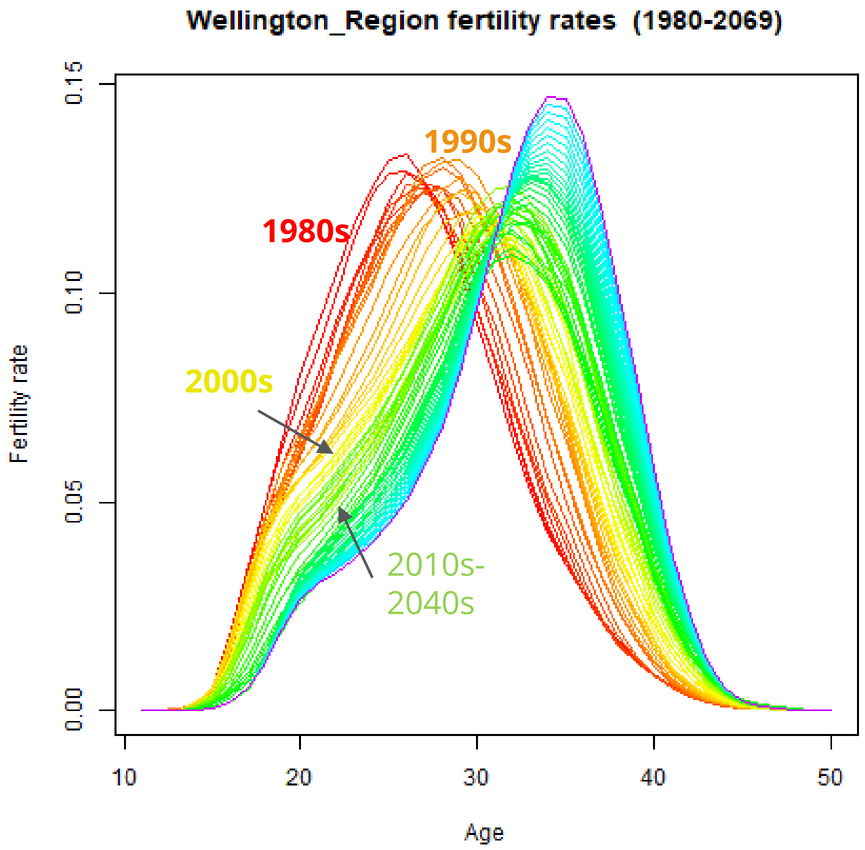


A.1.2 Fertility forecast method step 1

- Functional demographic model used for forecasting regional fertility rates, based on estimated region-specific rates[1]
- The functional demographic model is a generalisation of the standard and widely used 'Lee-Carter' model, which decomposes trends in age-specific demographic rates, such as fertility, into components e.g.

  Age-specific rate = average rate by age + time trend x interaction between age effects and time trends

- The interaction between age effects and time trends is there to account for e.g. displacement effects such as an increase in fertility rates at age 30 when fertility rates at age 29 decline.

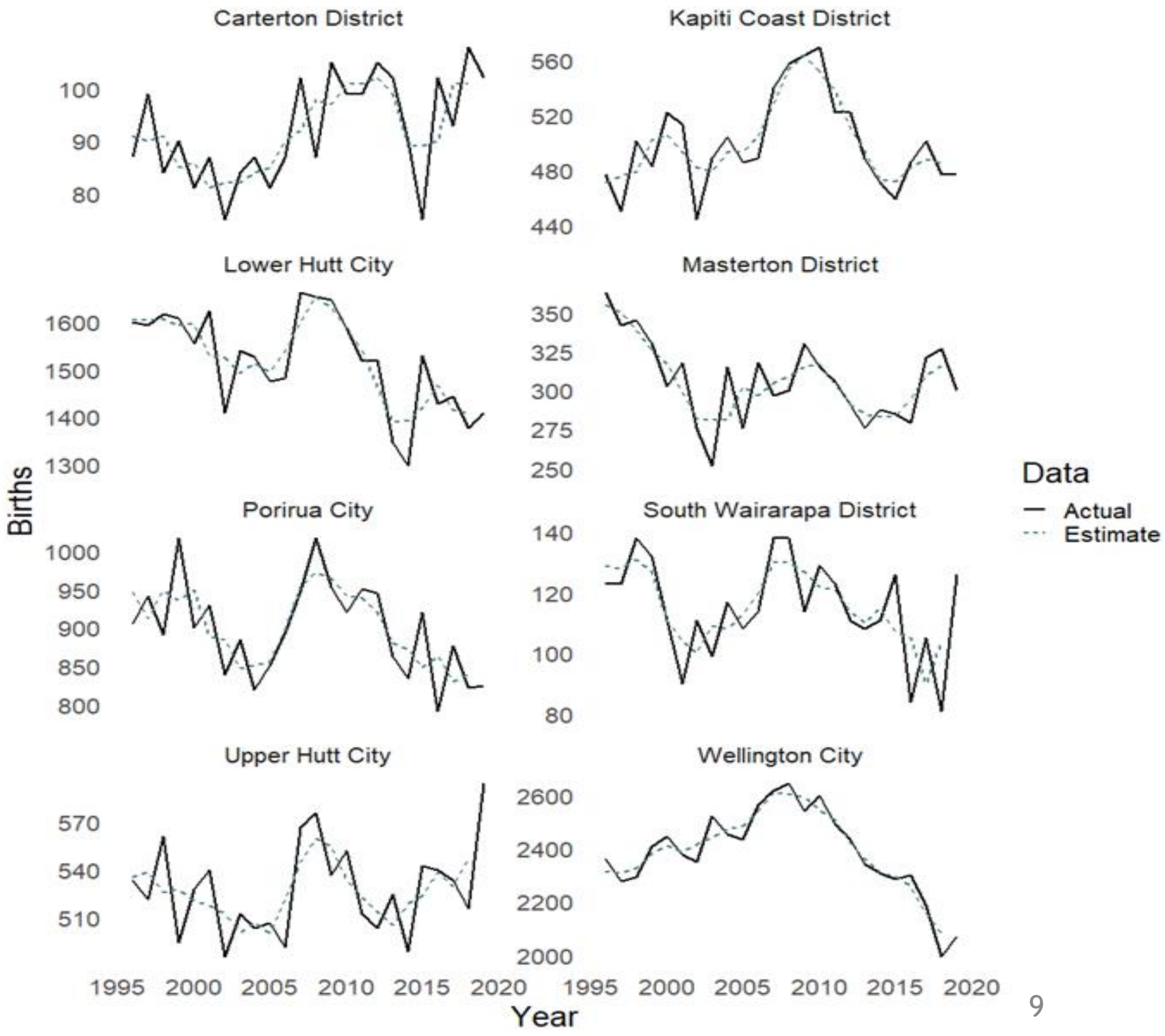


1. Adapted from R package 'Demography' by Rob J Hyndman, Heather Booth, Leonie Tickle and John Maindonald. Method based on Hyndman, R.J., Shahid Ullah, Md., 2007. Robust forecasting of mortality and fertility rates: A functional data approach. Computational Statistics & Data Analysis 51, 4942–4956. https://doi.org/10.1016/j.csda.2006.07.028

A.1.2 Fertility method, step 2

- Variations in fertility rates across districts estimated with ‘fixed effects’ (differences in averages) from a model of total birth rates, by district, based on regional fertility rates.
- The historically fitted fixed effects are assume to persist in future.
- The table below summarises district differences in births relative to the region. The Figure at right shows the fit of the model.

District	Average birth rate relative to region (=1)	Standard error
Kapiti Coast District	1.2	0.06
Porirua City	1.4	0.05
Upper Hutt City	1.1	0.06
Lower Hutt City	1.2	0.04
Wellington City	0.8	0.02
Masterton District	1.3	0.05
Carterton District	1.3	0.13
South Wairarapa District	1.3	0.13



A.2. Mortality

- Deaths (X) by district (r) by age (a) by sex (s) are forecast with

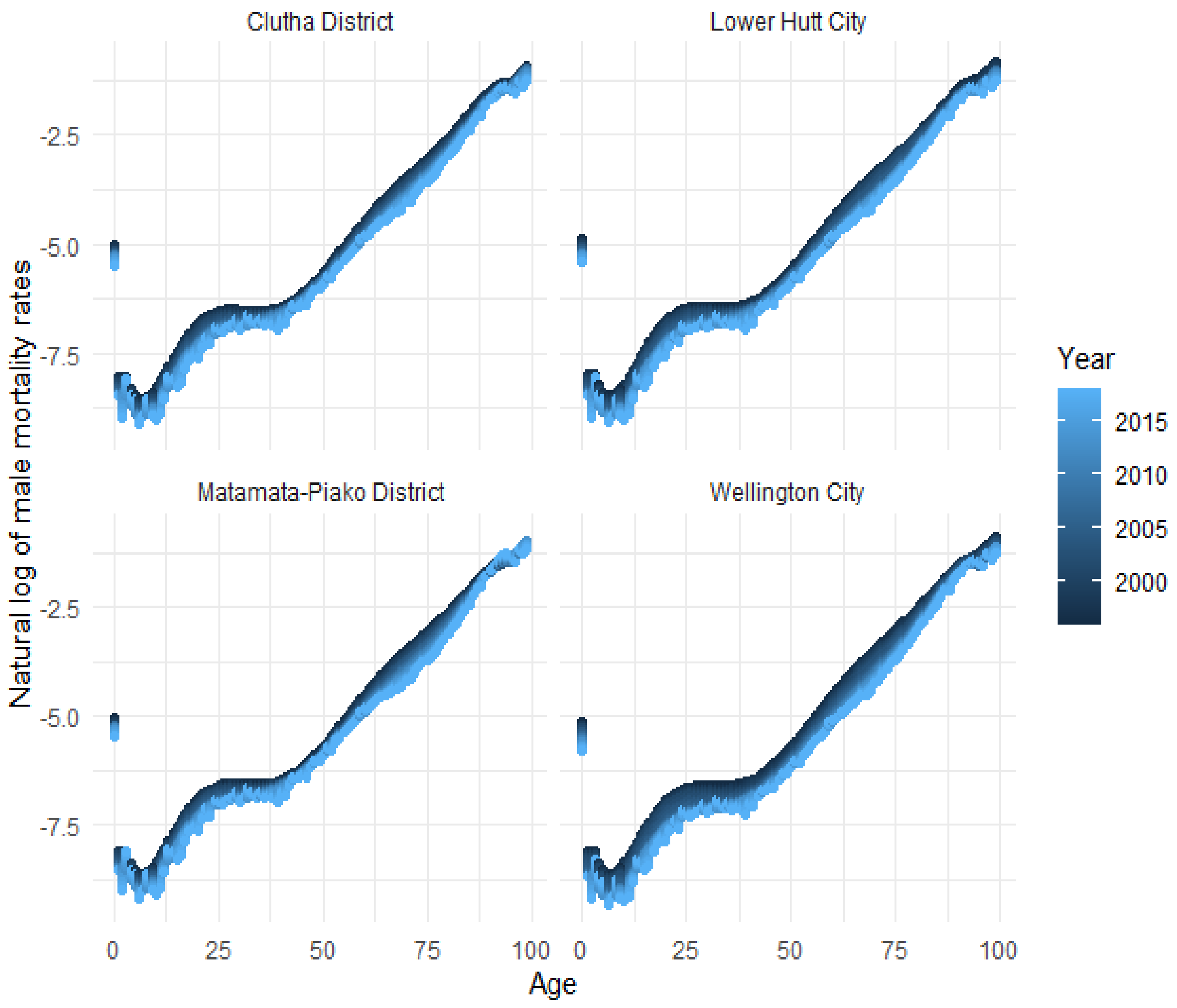
$$\mathrm{X_{r,a,s,t}} = (x_{r,a,s,t}+\epsilon_{at}).N_{r,a,s,t}$$

Where number of deaths are a function of

- the size of the population by age ($N_{r,a,s,t}$) and forecast age-specific mortality rates ($f_{r,a,t}$)
- random variation in deaths from year to year using $\epsilon_{at} \sim N(0, \sigma_{as})$, where the age- and sex-specific standard error (σ) is based on the average difference between model fitted mortality rates and observed mortality rates using a national-level model (due to significant smoothing/interpolation in district-level data).

A.2.1 Mortality data

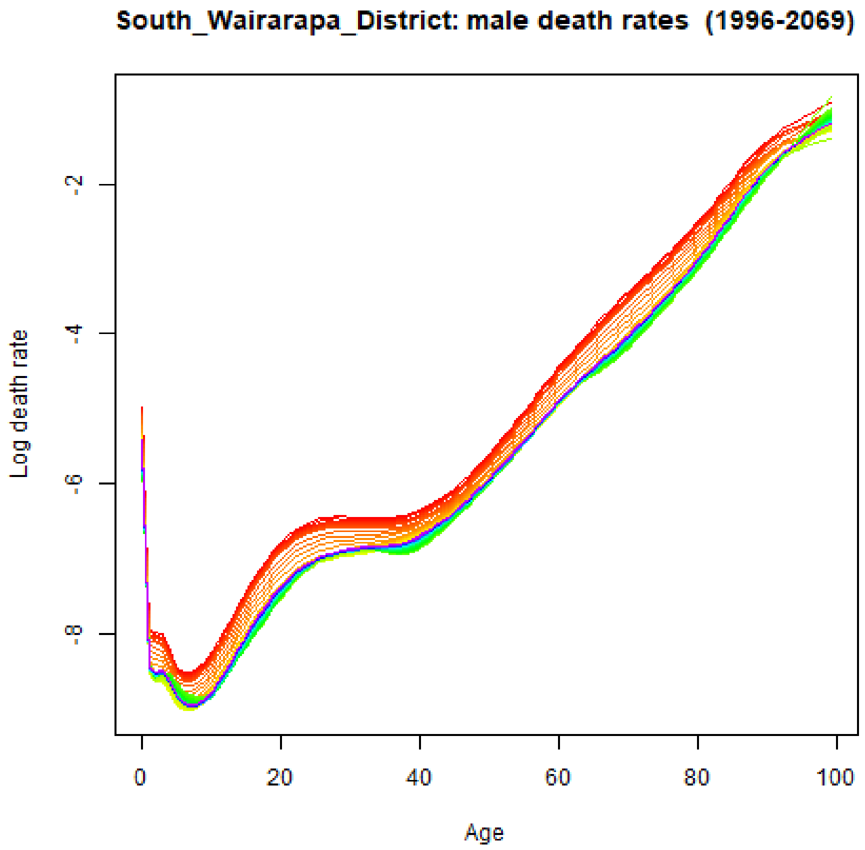
- Stats NZ national age- and sex-specific cohort life tables 1876-2018.
- Stats NZ subnational (district) age-group and sex-specific life-tables 1996, 2001, 2006, 2013. Rates for individual ages are interpolated based on splines fitted between age and log(mortality rate). The splines
- National changes (growth rates) in age-specific mortality rates are used to update the estimates of subnational age and sex-specific mortality rates (to 2018). A sample of the data is shown to the right. The updated data is much more volatile than the smoothed historical data (though this makes little difference to our models which smooth the data before estimation)



A.2.2 Mortality method

- Coherent functional demographic model used for modelling and forecasting district mortality rates.[2]

- The mortality model method is very similar to the fertility model method (a functional demographic model) with the addition that the model includes consideration of relative rates across different genders to ensure that forecasts are consistent. That is, they ensure that male and female mortality rates do not move too far apart, as reflected by historical ratios of male to female mortality rates.



2. See e.g. Woods, C., Dunstan, K., New Zealand, Statistics New Zealand, 2014. Forecasting mortality in New Zealand: a new approach for population projections using a coherent functional demographic model (Working Paper No. 14– 01), Statistics New Zealand Working Paper. Here too we use an adapted version of the R ‘Demography’. The main model function is ‘coherentfdm’.

A.3 International immigration

$$m_{d,a,t} = m^*(m^*_{t-1}, \alpha_t, \epsilon_t).\, p(a|m).\, p(r|a,m)$$

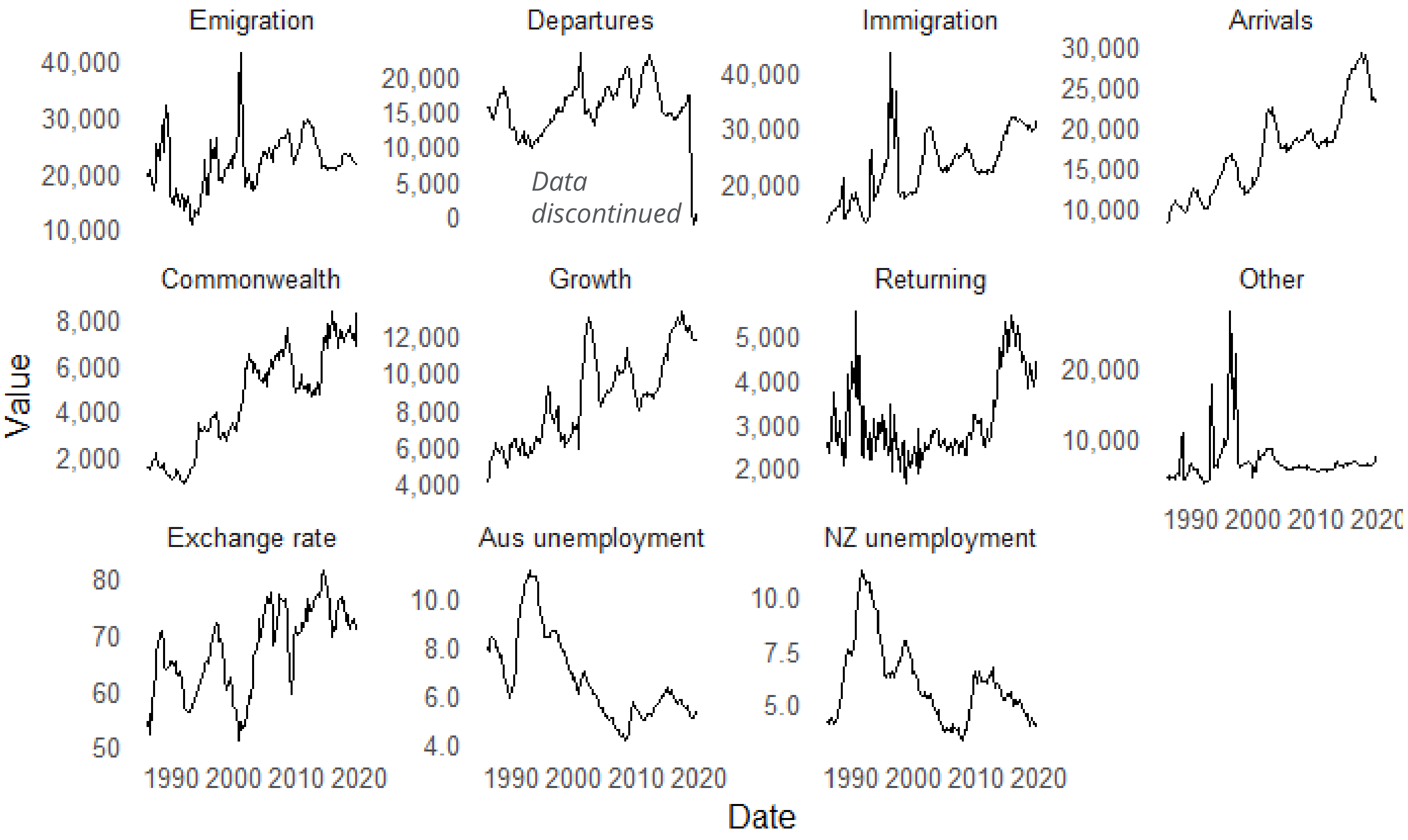
- Immigration ($m_{d,a,t}$) by district (r) and age (a) and year (t) is a function of exogenously forecast stochastic growth rates (α_t with error $\epsilon_t \sim N(0, \sigma_{\alpha_t})$) for national immigration (m^*) and fixed probabilities/shares for
  - ages of immigrants ($p(a|m)$) and
  - district destination conditional on age ($p(r|a,m)$)

The model includes arbitrary ‘shifters’ on immigration that are used to control for e.g. border closures with the immgration fixed, by year, at a chosen proportion of expected/forecast immigration inflows.

Uncertainty is modelled by varying the national immigration forecast using random selection from $\epsilon_t \sim N(0, \sigma_{\alpha_t})$.

A.3.1 International immigration, data

- The charts at right present the data used in our national immigration forecast model (this data is all pre-COVID and seasonally adjusted where appropriate)
- To incorporate, in our models, differences in immigration trends by country of origin and citizenship, we reduce the dimensions of the data using time series clustering[3] to produce 4 groups of related countries (see the second row of the chart at right) which we label, for expositional purposes, as:
  - **Commonwealth** countries (Non-NZ immigration from UK, SGP, India) (stable upward trend)
  - **Growth** countries- dominated by China (non-NZ), Australia (non-NZ) and New Zealander's returning from the UK (high growth group)
  - **Returning** New Zealanders returning from Australia, Samoa, Hong Kong
  - **Other** - numerous countries dominated by Non-New Zealand citizen movements (spiked in the 1990s, flat since)
- The migration data we use is Stats NZ's 12/16 month rule (labelled Emigration and Immigration at right) with the history of the data back cast using correlations between overlapping 12/16 month rule and permanent and long term arrivals and departures data.

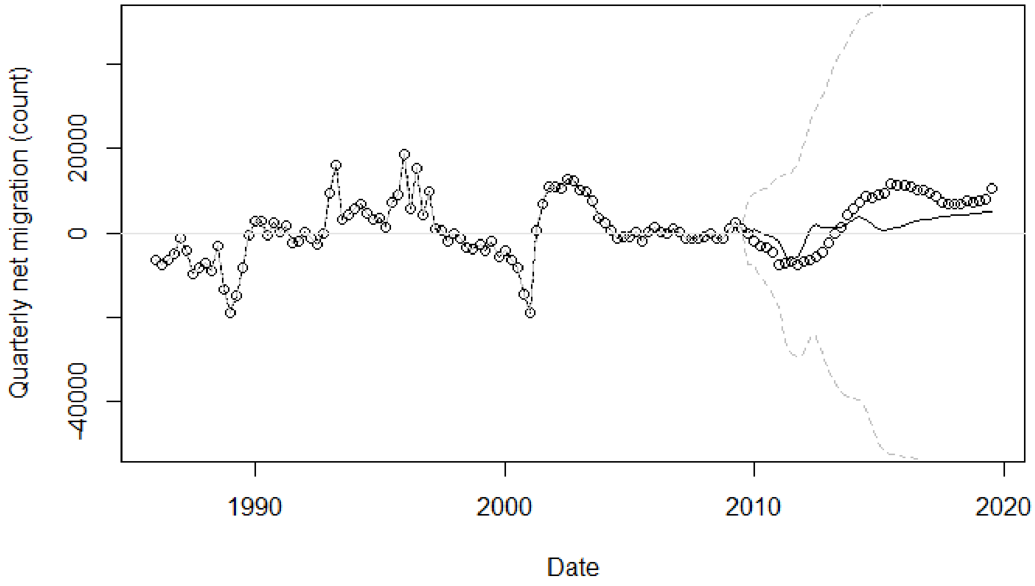


3. To do this we use the time series clustering package Tsclust and settle on the use of the logarithm of the normalized periodogram as the similarity measure. [Montero, P., Vilar, J.A., 2014. TSclust: An R Package for Time Series Clustering. Journal of Statistical Software 62, 1–43. https://doi.org/10.18637/jss.v062.i01]

A.3.2 National immigration forecast models

- Average over 5 models
  - Simple vector auto-regression model containing total immigration and emigration, 6 lags and trend and intercept terms
  - Univariate time series trend model for aggregate immigration
  - Univariate time series trend model for immigration from each of our 4 country groups
  - Simple vector error-correction model incorporating total emigration and immigration from each of our country clusters (i.e. 5 endogenous variables)
  - Vector error-correction model with macro-economic variables, incorporating total emigration, immigration from each of our country clusters, the Australian and New Zealand unemployment rates and the New Zealand trade-weighted exchange rate.
- The average over the models can be based on equal weights on each model (default in the model at right) or through calibration of weights to produce best fitting overall model (historically).
- Forecast standard error σ_{α_t} is calculated based on the lower bound of the forecast confidence interval from the 5 models and the smallest of the upper bounds from the 5 models (to minimise extreme exponential growth in immigration on the grounds that policy would be likely to prevent this).

Out of sample forecast test

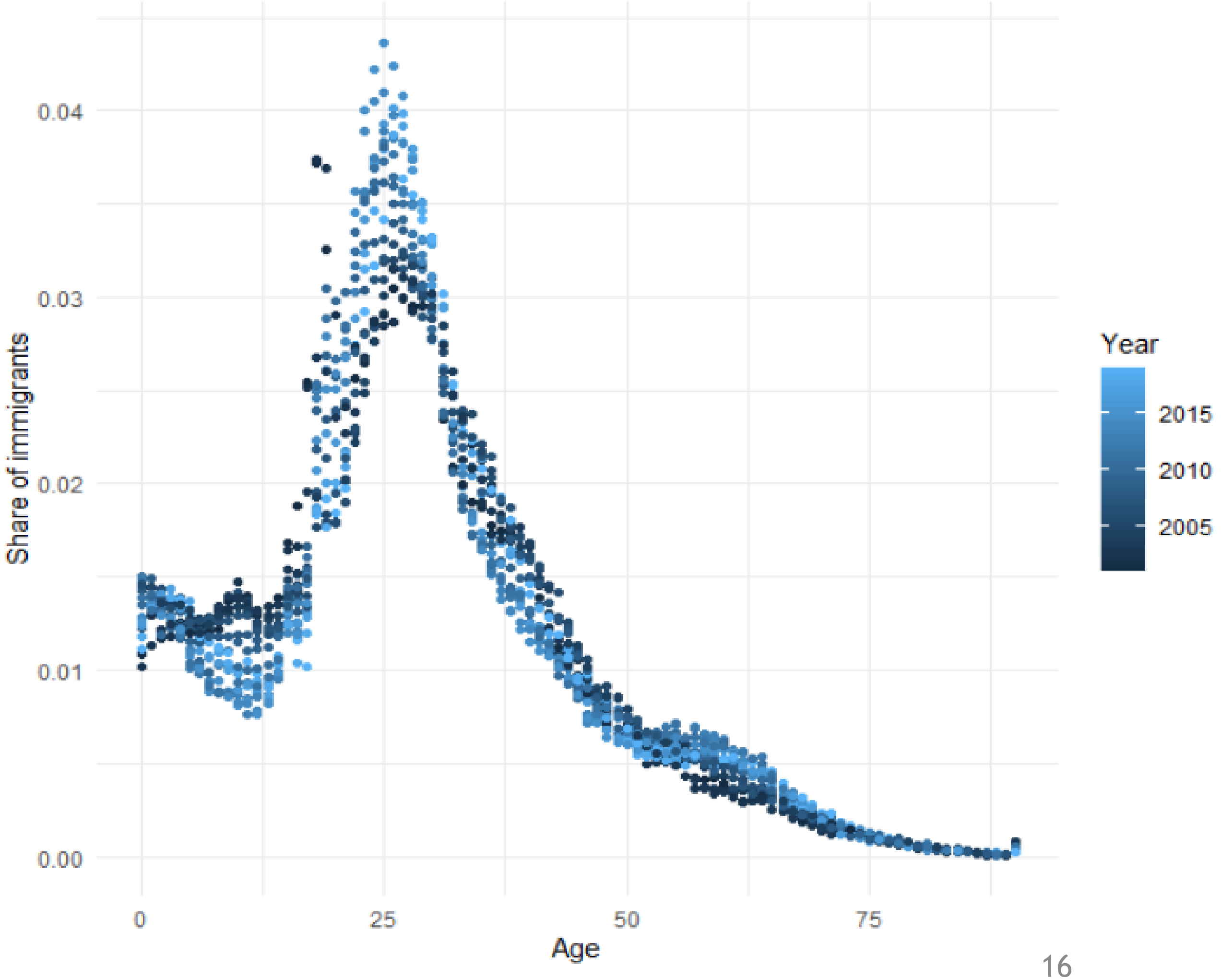


Model fitted on quarterly data 1986-2010 and forecast for 9 years 2011-2019. Forecasts of the 'growth' and 'commonwealth' clusters are very good. Forecasts of the 'returning' and other cluster consistently underpredicted immigration flows. The model predictions of unemployment rates were reasonably good. But forecasts of the real TWI were persistently low.

A.3.2 International immigration by age

- Age profile of migrants $p(a|m)$ is, by default, held constant at the most recent age-profile of immigrants, though this can be adjusted where there is evidence of a sharp change in age profile (e.g. as a consequence of border closure).
- The age profile of migrants is, broadly speaking, highly stable, although in recent years there has been a material increase in the share of migrants aged between 18 and 30 and a decline in the share of immigrants under 18.
- The sex of international immigrants is assumed to be 50% male and 50% female.

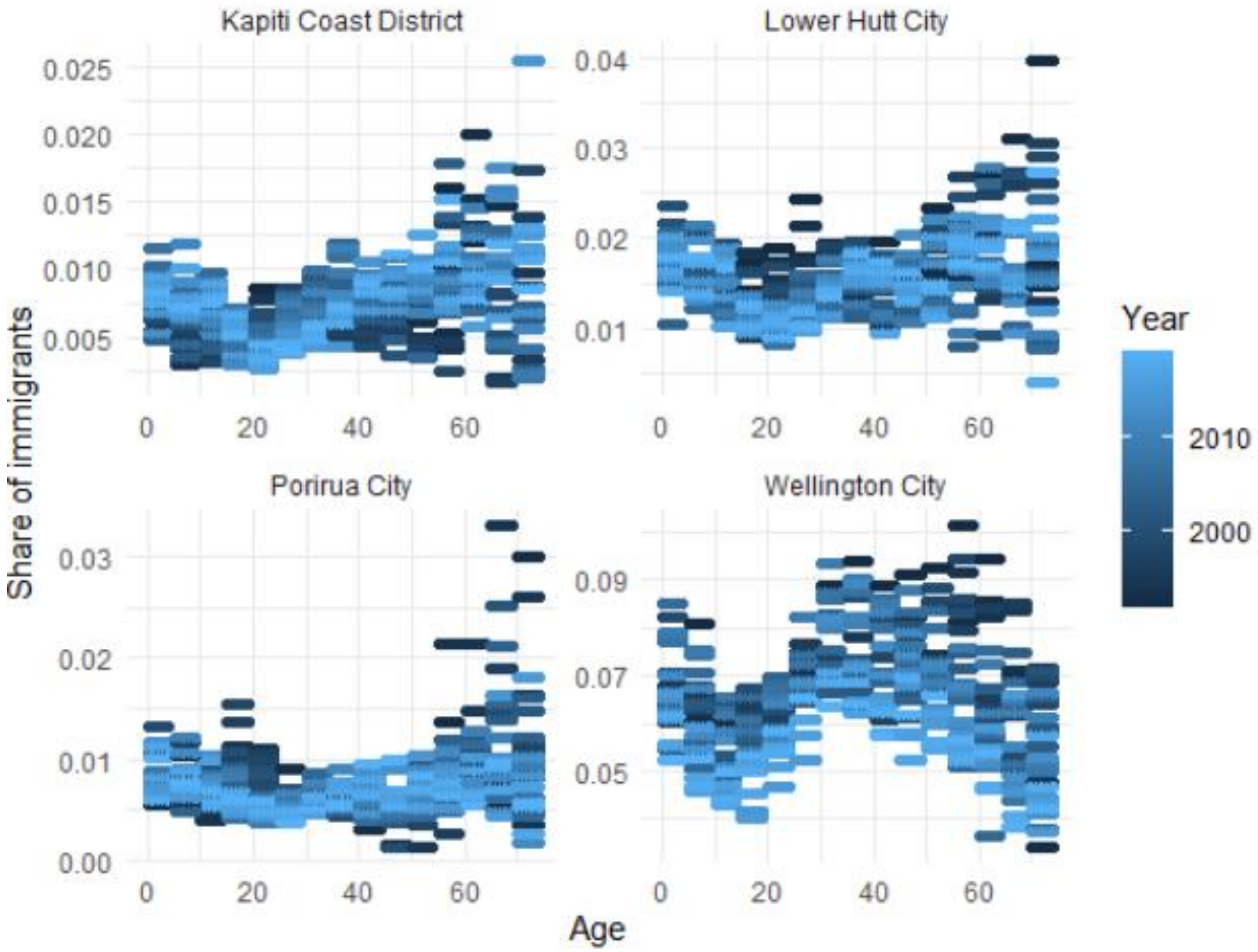
Share of international immigrants by age, 2001-2019



A.3.2 District inflows of international immigrants

- The shares of national international immigrants by age by district $p(r|a,m)$ are, by default, held constant
    - at the most recent value
    - within 5 yearly age groups.
- District shares of immigration by age are volatile – both over time and across districts. Though this is partially due to questionable data accuracy.

Share of international immigrants by age, 1992-2019



A.4 International emigration

- International emigration is assumed to be a function of district population sizes and age-distribution (i.e. population by district and age $N_{r,a}$) and stochastic age-specific propensities to emigrate ($e^{w}_{r,a,t}$).

$$E^{w}_{r,a,s,t} = N_{r,a,s,t}.\,e^{w}_{r,a,t}$$

- Age-specific propensities to emigrate are assumed to be the same for males and females
- National emigration is estimated bottom-up based on the sum of international emigration from districts.

A.4.1 International emigration, data

- There are no up-to-date indicators of international emigration by district and age since Stats NZ/government stopped collecting data on departures by district and age in 2018.
- We use Stats NZ data on permanent and long-term departures prior to 2018 to estimate international emigration rates by district.
- We adjust our models to account for persistent differences between national total permanent and long-term departures and the more robust 12/16 month rule emigration measure – based on whether travellers spend 12 of the following 16 months out of NZ
  - net migration is very similar whether measured by permanent and long-term arrivals and departures or the 12/16 month rule, however
  - arrivals overstate immigration and
  - departures understate emigration (by ~40% on average in the past 10 years).

A.4.1 International emigration model

- The core of the model is estimated mean rates/propensities of migration by age group and district
- A simple autoregressive model is fitted to rates of emigration by age-group and district, to ensure that emigration dynamics (persistence) are accounted for and so that we can estimate model errors (i.e. for stochastic simulation):

$$e^w_{r,a,t} = \mu^w_{r,a} + u_{r,a,t}$$

$$u_{r,a,t} = \rho_{r,a} u_{r,a,t-1} + \epsilon_{r,a,t}$$

$$\epsilon_{r,a,t} \sim N\left(0, \sigma_{\epsilon_{r,a,t}}\right)$$

- Where $\mu^w_{r,a}$ is the mean rate of emigration, the $u_{r,a,t}$ and $\epsilon_{r,a,t}$ are error terms (the former being structural deviation from mean and the latter being pure model error) and $\rho_{r,a}$ is the autoregressive term to be estimated.
- The means ($\mu^w_{r,a}$) and standard errors ($\sigma_{\epsilon_{r,a,t}}$) of the models are adjusted by the mean of the ratio of the national 12/16 month rule emigration to mean permanent and long-term departures (1.35), to account for under-counting of emigration using departures data.

Sample model parameters for Masterton District

Age group	Mean $u_{a,t}$	μ^w_a	ρ_a	$\sigma_{\epsilon_{r,a,t}}$
0-4 years	-0.0033	0.0160	0.0107	0.0070
5-9 years	-0.0031	0.0124	0.3457	0.0050
10-14 years	-0.0030	0.0082	0.0585	0.0025
15-19 years	-0.0040	0.0234	0.2104	0.0052
20-24 years	-0.0222	0.0716	0.2317	0.0189
25-29 years	-0.0047	0.0595	0.1696	0.0110
30-34 years	-0.0057	0.0251	0.1056	0.0090
35-39 years	-0.0033	0.0153	-0.0110	0.0058
40-44 years	-0.0045	0.0109	-0.1238	0.0041
45-49 years	-0.0004	0.0082	-0.1659	0.0042
50-54 years	-0.0023	0.0073	0.2879	0.0022
55-59 years	0.0002	0.0051	0.1231	0.0014
60-64 years	-0.0011	0.0039	-0.3877	0.0028
65-69 years	-0.0009	0.0020	-0.5447	0.0010
70-74 years	-0.0004	0.0017	0.0364	0.0015
75 years and over	-0.0011	0.0034	-0.2997	0.0025

A.5 Domestic/internal migration

- Internal emigration (E^i) between districts (r) by origin and destination (o and d) and age (a) is modelled as a function of
  - district populations ($N_{r,a,t}$)
  - emigration rates by origin-destination ($e^i_{o=r,a,t}$).

$$E^i_{r,a,s,t} = \sum_d N_{r,a,t} e^i_{o=r,d,a,t}$$

- Age-specific rates of emigration, from an origin to a destination comprise two parts:

$$e^i_{o,d,a,t} = p(E^i|r,a).p(E^i_{od}|E^i)$$

  - $p(E^i|r,a)$ = district- and age-specific probabilities/propensities that a person will emigrate
  - $p(E^i_{od}|E^i)$ = district- and age-specific probabilities/propensities that a person will migrate to a specific district (destination), given that they have chosen to emigrate
- Domestic immigration is (by definition) the sum over origins of domestic emigration to a particular destination.

$$M^i_{r,a,s,t} = \sum_o N_{r=o,a,s,t} e^i_{o=r,d,a}$$

A.5.1 Domestic migration, data

- Census domestic migration rates 1991, 1996, 2001, 2006, 2013
  - Note that these rates are remarkably stable over time, in large measure because they are dominated by predictable/stable life-cycle effects and age-specific rural-urban movements
- Data on origin-destination movements are estimates from administrative data, for 2014-2017[5]
  - Census data can be used for this purpose however it is (close to?) impossible to accurately infer age-specific and year-specific movements solely from cumulative 5-yearly population snapshots (see example in table at right).
  - the recent census data is of low quality when it comes to domestic migration origin-destination flows.

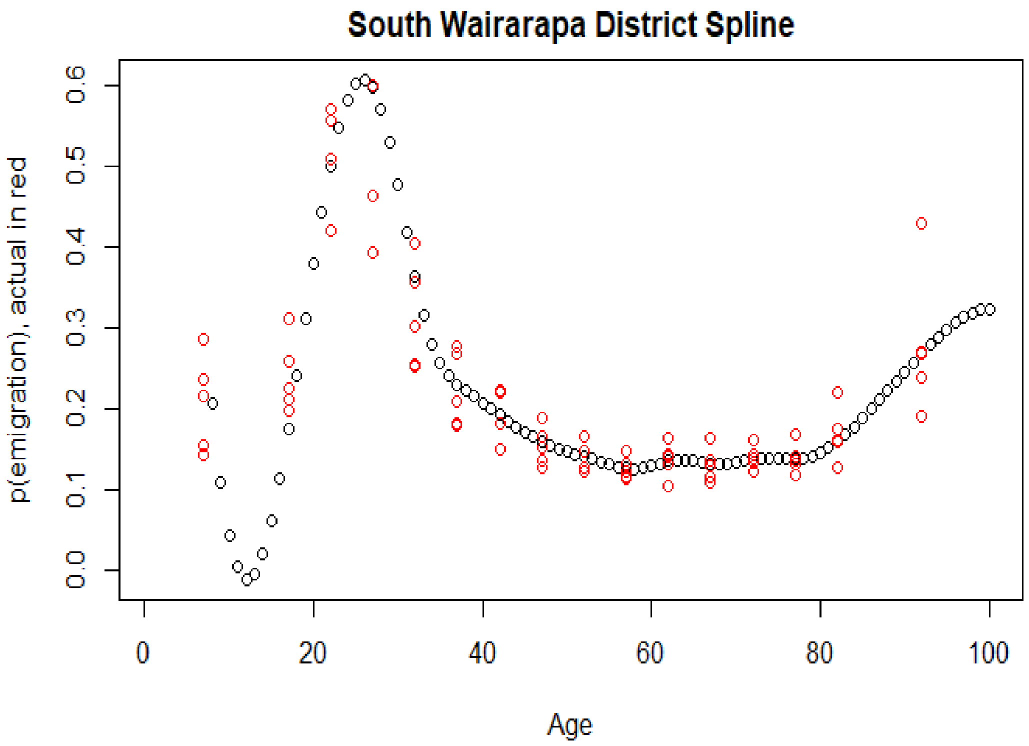
Example of differences in domestic migration data, annual vs cumulative census estimates

Female net migration in Dunedin City. 2018 annual movements assumed equal to 2017. Bottom right corner shows the numbers that would be represented in census data.

	Annual net migration (admin data)					Cumulative net migration (what census sees)				
Age	**2014**	**2015**	**2016**	**2017**	**2018**	**2014**	**2015**	**2016**	**2017**	**2018**
10	-35	-43	4	-25	-25	-35	-43	4	-25	-25
11	27	33	-3	19	19	27	-3	-45	23	-6
12	33	39	-3	23	23	33	67	-6	-22	46
13	5	7	-1	4	4	5	39	66	-2	-19
14	0	0	0	0	0	0	5	39	66	-2
15	6	6	5	4	4	6	6	11	43	70
16	8	8	7	5	5	8	14	12	16	48
17	186	168	151	115	115	186	177	165	127	130
18	884	799	720	544	544	884	985	897	710	671
19	17	15	14	10	10	17	899	999	907	720
20	-32	-28	-45	-54	-54	-32	-12	854	945	853
21	-90	-80	-128	-153	-153	-90	-112	-140	700	792
22	-95	-84	-135	-161	-161	-95	-175	-247	-301	539
23	-75	-67	-107	-127	-127	-75	-162	-281	-374	-428
24	-23	-20	-32	-38	-38	-23	-95	-194	-319	-412
Age groups:										
15-19	1,101	996	897	678	678	1,101	2,080	2,084	1,802	1,639
20-24	-315	-279	-447	-534	-534	-315	-555	-8	651	1,343

5. https://www.stats.govt.nz/reports/internal-migration-estimates-using-linked-administrative-data-201417

A.5.2 Domestic emigration model, probability of emigrating



- Simple model of mean probability of emigration by district and age with deviations around the mean

$$p\left(E^{i}_{r,a,t}\right) = \mu_{r,a} + \epsilon_{a,r,t}$$

$$\epsilon_{a,r,t} \sim N(0, \sigma_{\epsilon_{a,r,t}})$$

- A combination of splines and semi-parametric models are fitted to the natural logarithm of observed (census) emigration rates by age-group, to estimate mean age-specific rates of emigration ($\mu_{r,a,t}$). E.g. see left.
- Similarly the standard error by age is based on a spline fitted to standard errors (almost universally this shows variance increasing linearly with age)

Inspection of the spline and the semi-parametric models suggest using a combination of both. The semi-parametric model allows prediction and interpolation. While the spline provides easy to use summary information of past rates. For some areas (not very many though), the spline of best fit is linear, which is at odds with the data even if it is the best fit. So, there we use the SPM model. The SPM model is also needed to infer emigration rates at extreme ages, particularly below age 8 (given the number of observations needed to fit the first smoothed spline value). Otherwise the spline is used to characterise underlying rates.

A.5.2 Domestic emigration model, probability of emigration destination

- Simple average of shares of emigrants from an origin to all potential destination districts/destinations in New Zealand (average over the four years from 2014 to 2017 – noting that matching quality is patchy suggesting that focussing on single years of data is unwise).
- Note that this is an area of the methodology that is ripe for development of richer models e.g. with economic predictors though this would require a research project with access to Stats NZ micro-data.

E.g. probability a migrant, in a given age group, from Wellington City lands in...

Age group:

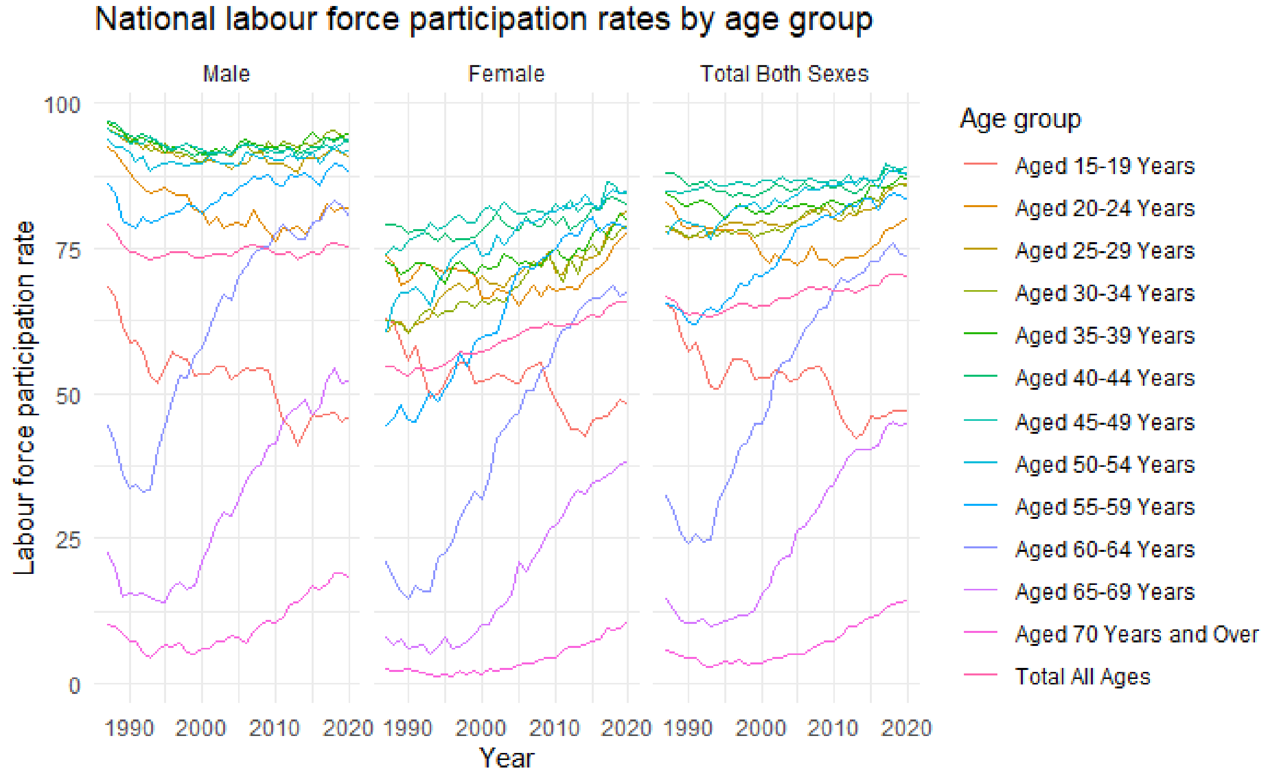
District:	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80+	All
Palmerston North	0.02	0.02	0.02	0.04	0.04	0.03	0.02	0.03	0.02	0.02	0.02	0.01	0.02	0.02	0.03	0.00	0.02	0.03
Horowhenua	0.02	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.03	0.04	0.04	0.06	0.02	0.03	0.02
Kapiti Coast	0.09	0.09	0.09	0.04	0.04	0.05	0.06	0.08	0.10	0.11	0.12	0.14	0.15	0.23	0.19	0.21	0.23	0.08
Porirua	0.19	0.18	0.19	0.08	0.08	0.10	0.12	0.14	0.14	0.14	0.14	0.12	0.11	0.09	0.11	0.12	0.17	0.12
Upper Hutt	0.05	0.04	0.05	0.02	0.03	0.04	0.05	0.05	0.04	0.05	0.04	0.04	0.03	0.03	0.05	0.04	0.04	0.04
Lower Hutt	0.18	0.16	0.14	0.10	0.12	0.18	0.21	0.18	0.18	0.16	0.14	0.17	0.11	0.11	0.14	0.23	0.16	0.15
Masterton	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.02	0.02	0.01	0.02	0.03	0.01	0.01
Carterton	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.01	0.03	0.01	0.01	0.00	0.01
South Wairarapa	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.03	0.03	0.02	0.03	0.03	0.02	0.01	0.01	0.01
Nelson	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.02	0.02	0.00	0.02	0.00	0.01
Tasman	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.01
Marlborough	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.02	0.01	0.01	0.02	0.01	0.01	0.01
Rest of NZ	0.39	0.44	0.44	0.64	0.61	0.53	0.46	0.45	0.43	0.42	0.42	0.39	0.43	0.36	0.33	0.29	0.30	0.50

E.g. probability a migrant, in a given age group, arriving in Wellington City came from...

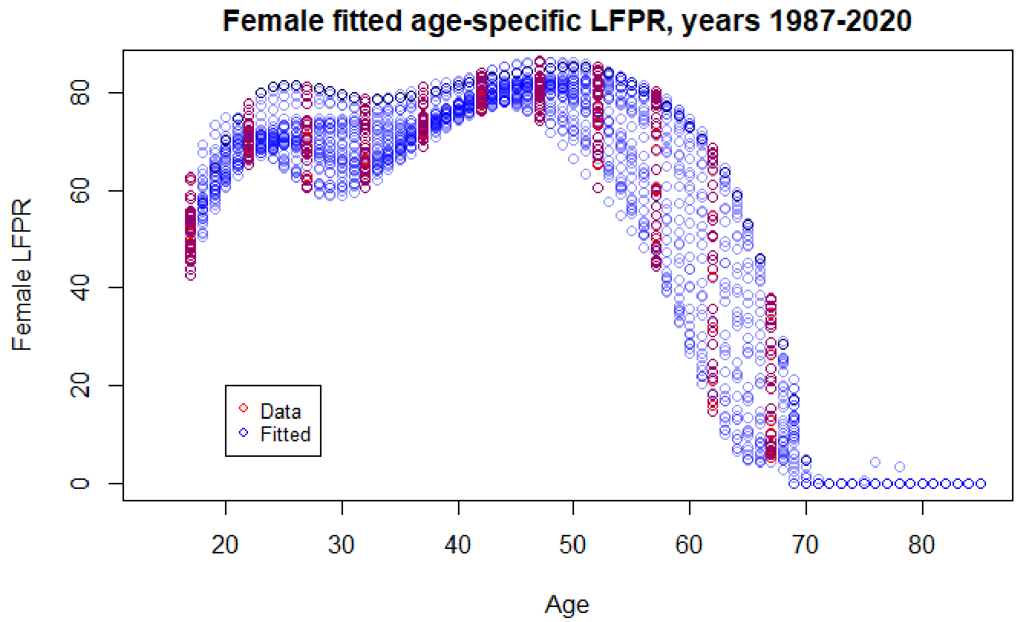
District	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80+	All
Palmerston North	0.01	0.03	0.03	0.08	0.05	0.03	0.03	0.03	0.02	0.03	0.03	0.02	0.01	0.01	0.00	0.00	0.03	0.04
Horowhenua	0.01	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.02	0.02	0.06	0.02
Kapiti Coast	0.04	0.06	0.09	0.09	0.06	0.05	0.04	0.03	0.06	0.06	0.07	0.06	0.06	0.07	0.11	0.10	0.21	0.06
Porirua	0.12	0.15	0.23	0.10	0.09	0.08	0.09	0.10	0.12	0.13	0.13	0.11	0.07	0.06	0.07	0.13	0.14	0.10
Upper Hutt	0.03	0.03	0.05	0.04	0.04	0.03	0.03	0.03	0.04	0.04	0.04	0.03	0.01	0.02	0.02	0.03	0.03	0.03
Lower Hutt	0.10	0.12	0.17	0.14	0.13	0.13	0.12	0.13	0.14	0.14	0.12	0.12	0.08	0.06	0.09	0.12	0.12	0.12
Masterton	0.01	0.01	0.01	0.04	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.02	0.00	0.01	0.01
Carterton	0.01	0.01	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01
South Wairarapa	0.00	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.00	0.02	0.01	0.01	0.02	0.02	0.01	0.03	0.01	0.01
Nelson	0.01	0.01	0.01	0.07	0.03	0.02	0.01	0.01	0.02	0.02	0.02	0.02	0.00	0.00	0.01	0.02	0.01	0.02
Tasman	0.01	0.00	0.01	0.05	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.01
Marlborough	0.01	0.01	0.01	0.03	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.00	0.01	0.01
Rest of NZ	0.30	0.38	0.48	1.44	0.71	0.54	0.44	0.38	0.39	0.43	0.37	0.33	0.25	0.23	0.27	0.27	0.24	0.58

A.6.1 Labour force, data

We use national household labour force survey (HLFS) data to examine trends in labour force participation rates

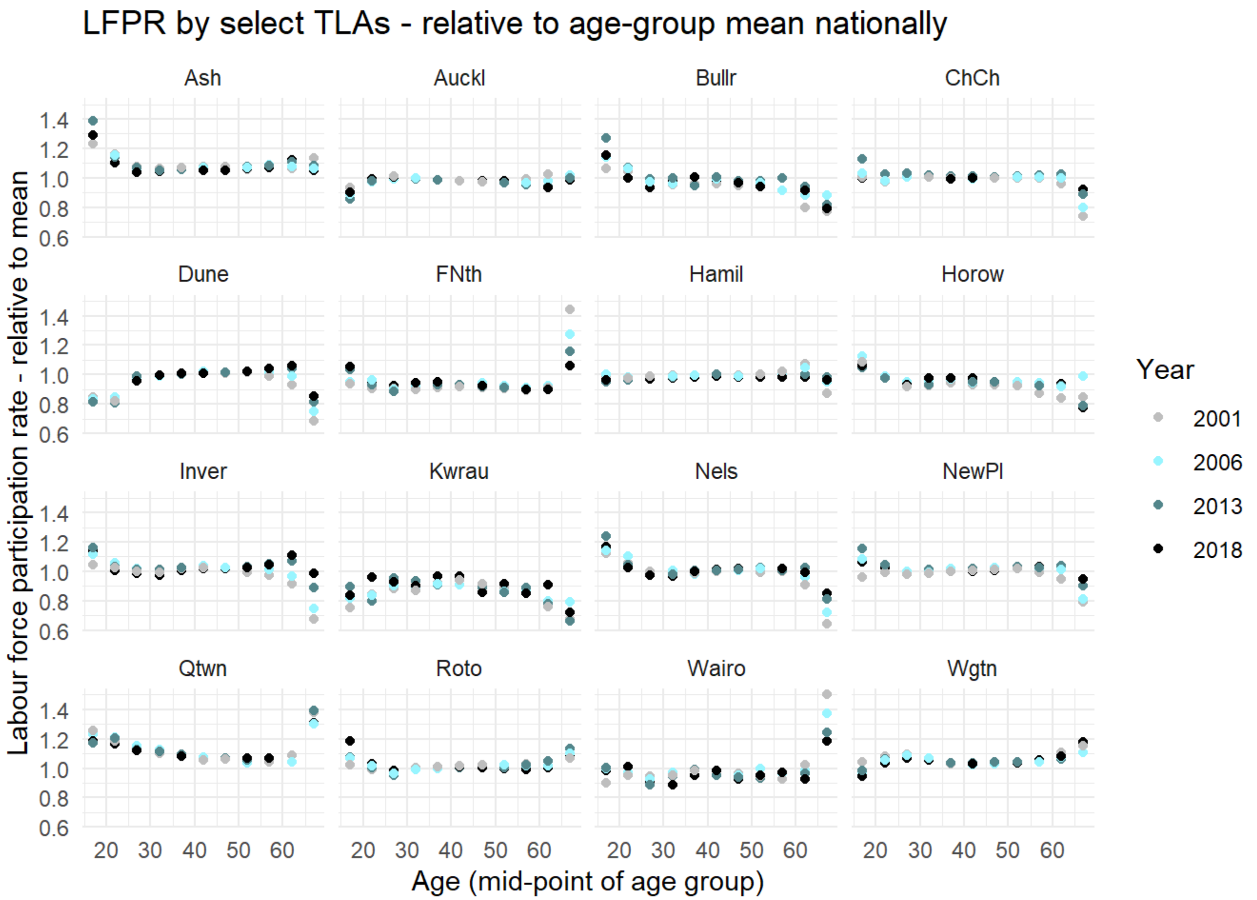


We use census data to estimate district-specific labour force participation rates. We fit splines to age-group data, to estimate age-specific labour force participation rates. This is done for both HLS data and for census data. See e.g. below for fitted national rates for females.



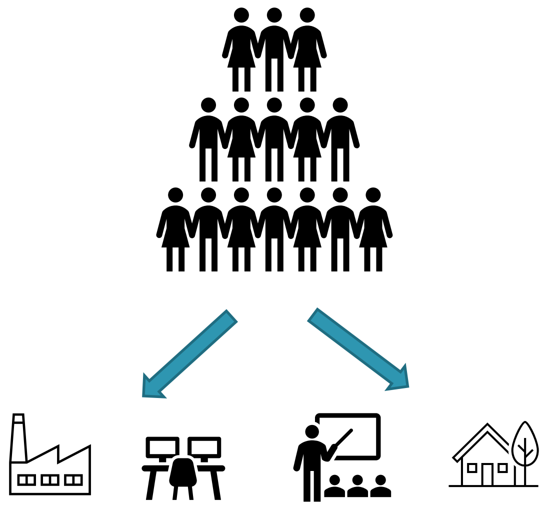
A.6 Labour force, model

- Labour force is modelled as a function of district population ($N_{r,a,s,t}$) and labour force participation rates ($l_{r,a,s,t}$) by district (r) and age (a) and sex (s) i.e. $L_{r,a,s,t} = l_{r,a,s,t}.N_{r,a,s,t}$
- Labour force participation rates (LFPR) are modelled/forecast with a top-down process (similar in spirit to our regional fertility model).
  - First we fit functional demographic models (adapted from the same methods/functions used to project fertility and mortality) to national LFPR data by sex (because national data is higher frequency than district data which is solely available from the census) and use this as the main basis for our forecasts.
  - As these are purely statistical models of trends and compositional changes we have to apply 'arbitrary' constraints to our national LFPR projection, after about 15 years, to e.g. constrain rapid growth in older age (65+) LFPRs from exceeding those of younger people (e.g. 35-40). We also constrain female labour force participation from exceeding male rates, at younger ages.
  - Then we fit smooth functions (splines) to census data (see example to right) describing the difference between districts' labour force participation rates and national rates. And we hold these relative differences in rates constant over time.



B. Work

Employment and earnings

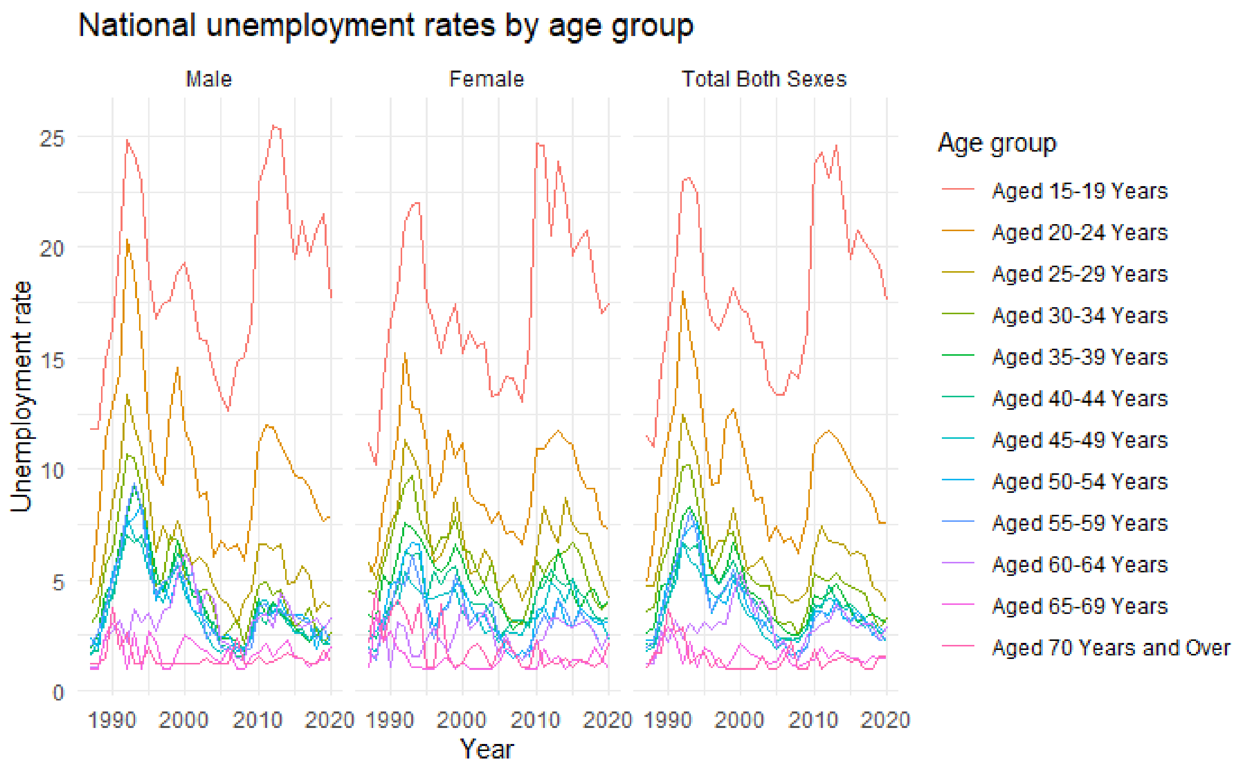


B.1 Approach

- We follow methods that are similar to those used by the Treasury for long-term fiscal modelling i.e. a growth accounting method intended to capture long term structural trends as opposed to focussing on short term cycles.
- Forecasts are based first on (external/exogenous) forecast national trends and cycles in unemployment, labour force growth and growth in multi-factor productivity (this is the growth accounting method)
- District level employment is forecast using labour force growth (as above), national unemployment rate forecasts, and district level age-specific unemployment rates relative to national rates.
- Forecasts of employment and earnings are district-specific, with no enforced adding-up constraints, with respect to national GDP.
- Earnings growth and GDP growth fall out of calculations of labour force and employment growth – that is the growth accounting method.

B.2 Employment and earnings, data

- Estimates of unemployment rates by age and by district are based on census data (as this is the only data available on unemployment by district and by age).
- National aggregate unemployment rates are measured using the Household Labour Force Survey (HLFS) as this is the conventional measure of unemployment nationally.
- HLFS and census measures of unemployment do not agree, with the census recording higher rates of unemployment than the HLFS measure. Where these data disagree we defer to census data, because of its detail at the district level.
- Earnings are defined as labour income growth, by district, and is measured using linked employer-employee data
- Historical GDP and GDP growth by district is based on Sense Partner's estimates of district GDP*

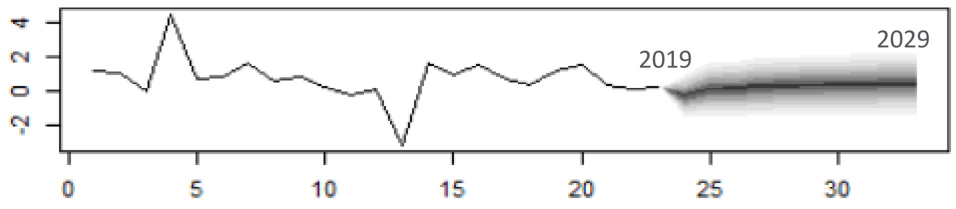


*As GDP is only an input and not a core output of these projections, we do not go into any detail here as to how our estimated are constructed. However our method for estimating district GDP is similar to the one used by MBIE. See https://www.mbie.govt.nz/business-and-employment/economic-development/regional-economic-development/modelled-territorial-authority-gross-domestic-product/mtagdp-methodology/.

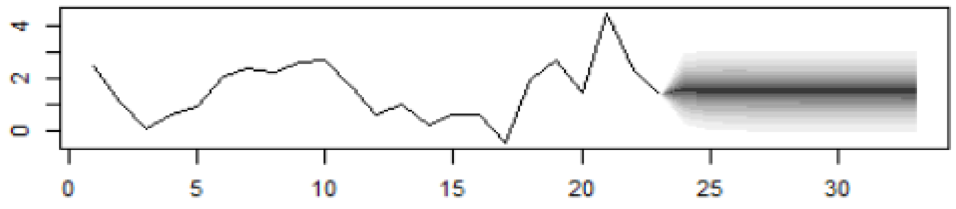
B.3 National economic growth model

- Built up from forecasts of:
  - labour force growth
  - multi-factor productivity growth
  - unemployment rates
- A simple vector auto-regression (VAR) model is used to forecast unemployment rates and multi-factor productivity (MFP), to ensure that our forecasts respect typical counter-cyclical co-variation of these variables
- Alongside the model, we adopt Treasury's (BEFU) unemployment rate forecasts for the next 3 years, given lags in the release of multi-factor productivity and therefore lags in the data that enters our VAR model.
- Thus the VAR provides
  - estimates of trend growth in MFP and unemployment conditional on labour force growth
  - forecast 95% confidence intervals (see example at right) that we use to specify the stochastic part of our model.

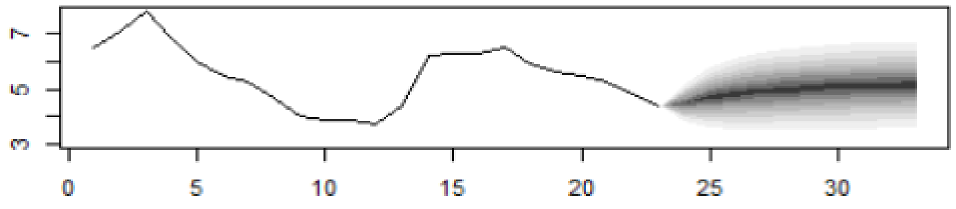
Annual multifactor productivity growth (%)



Annual labour force growth (%)

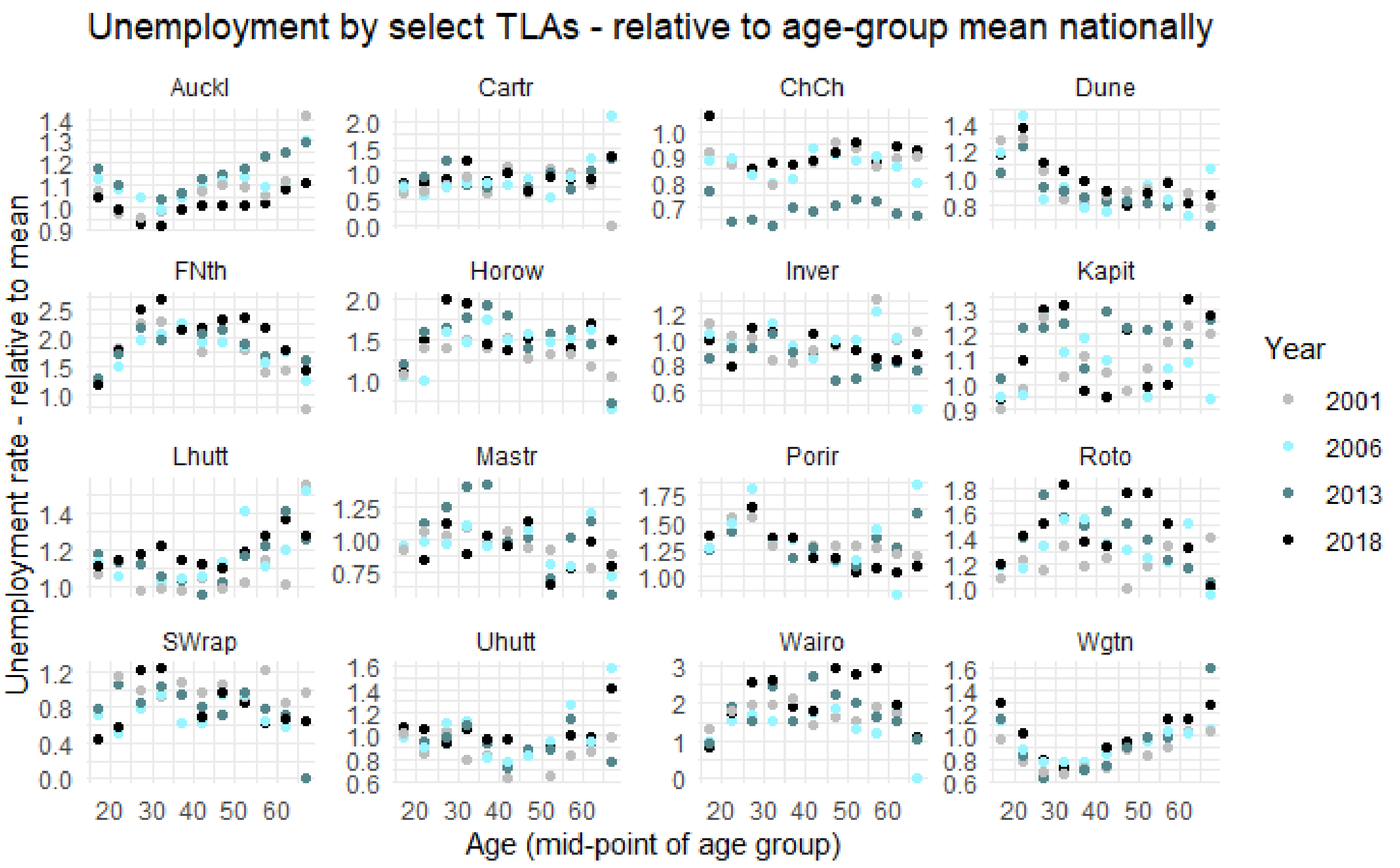


Annual average unemployment rate (%)



B.4 District employment model

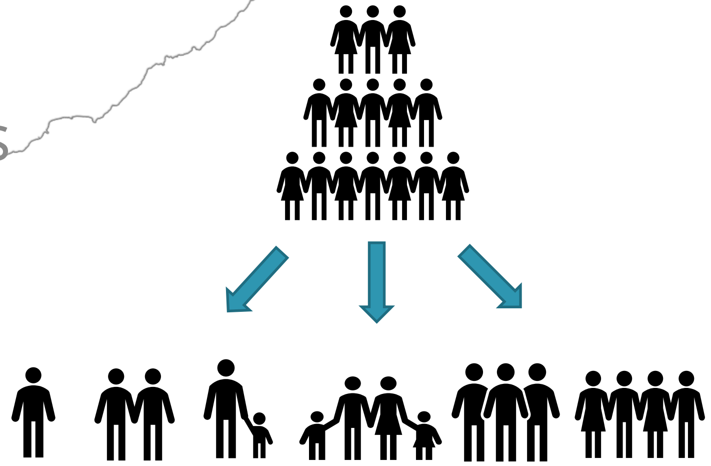
- District employment and unemployment is a function of:
  - forecast labour force
  - average national age-specific unemployment rates as a function of
    - national unemployment rates and
    - the relative rate of unemployment by age in a district (from census data), relative to the national average (see e.g. chart at right).
- The intent of this approach is to capture long term structural differences in age-specific employment rates by district, while incorporating national level trends.
- Structural differences in age-specific employment rates are quite pronounced, in part reflecting a natural 'sorting' across districts (e.g. young people with good job prospects or educational opportunities tend to leave some districts and at reasonably high rates causing high rates of unemployment amongst young people remaining in those districts while other districts have lower unemployment rates as those with good job prospects flow into the region).



GWRC districts: Cartr=Carterton, Kapit=Kapiti Coast, Lhutt=Lower Hutt, Mastr = Masterton, Porir=Porirua, SWrap=South Wairarapa, Uhutt=Upper Hutt, Wgtn=Wellington

C. Households

Living arrangements and household incomes



C.1 Household formation and income growth, approach

- We model household formation using entirely conventional methods (the same as those used by Stats NZ) involving tables of 'Living Arrangement Type Rates'. These tables describe the probability (or propensity) that an individual of a given sex and age will be in one of six living arrangements (including living in a non-private residence) and their position in those arrangements (i.e. whether a child in the family or an adult in the family). We do not model random variation in household formation.
- The living arrangement tables do not account for the number of families that are in multi-family households. This is calculated using Census data on the types of families (e.g. couples, sole-parent families, two-parent families) that typically form multi-family households).
- We infer district household income growth (at the mean) by district growth in earnings. Growth in average and median incomes by household types is adjusted using regional level observations of which households have mean and median incomes that tend to grow more quickly or more slowly than aggregate growth in earnings. This is based on simple linear regression of growth in median and mean income by household type according to the HLFS relative to growth in real earnings. The results of these regressions provide a measure of variation in the correlation between growth in earnings and growth in household income and we use this measure of variance to sample/simulate random variation in incomes.

C.2 Household formation, data and development

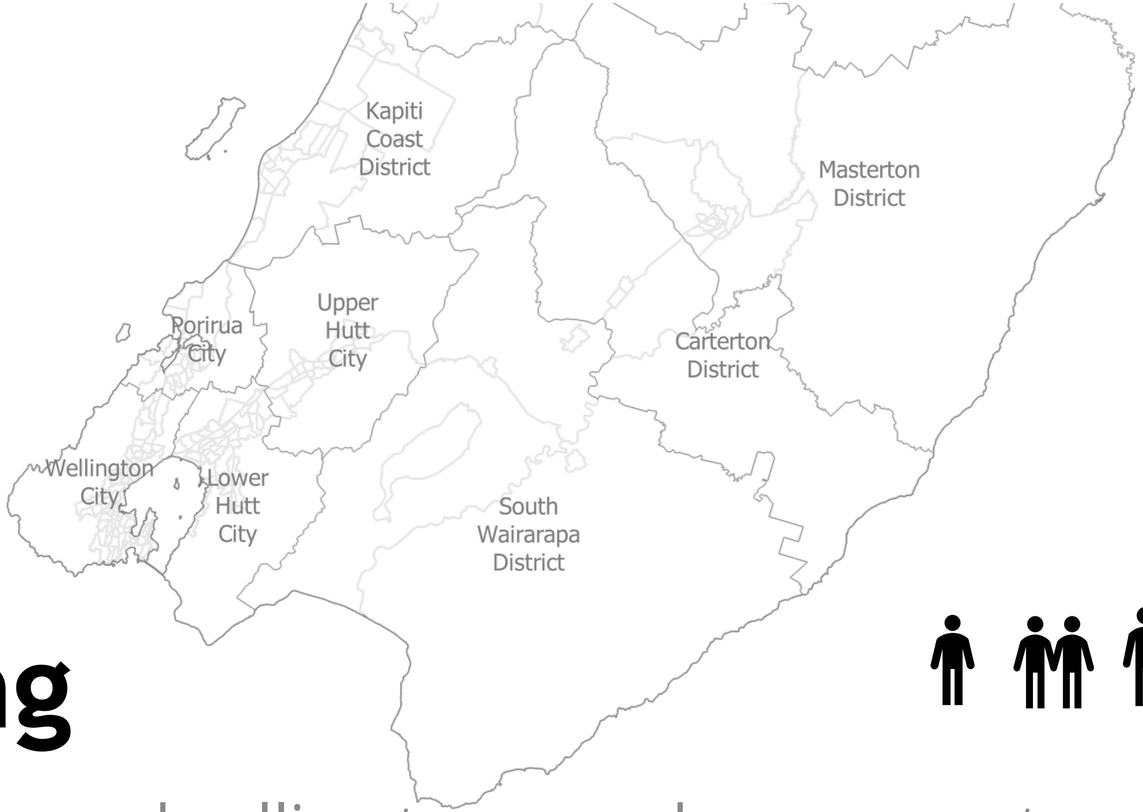
- We model seven 'household' types: sole person, sole parent, couple, two parents, multi-person, multi-family, other (non-private dwelling i.e. not a household).
- Statistics New Zealand provides us with data on living arrangement type rates by district (see illustrative examples of age-specific living arrangement rates at right). That is, the probability that a person of a given age and gender lives in a particular family or household type.
- Presently we rely on Stats NZ projections of trends in living arrangement to calibrate our forecasts. That data is currently based on the 2013 census.

National example 2013, female

	Age				
Living arrangement	**5**	**15**	**25**	**55**	**75**
Child in one-parent family	0.26	0.27	0.05	0.01	
Child in two-parent family	0.74	0.68	0.09		
One-person household			0.04	0.13	0.37
Other multiperson household			0.12	0.03	0.01
Other person in couple household		0.01	0.03	0.01	0.01
Other person in one-parent household		0.01	0.02	0.01	0.01
Other person in two-parent household		0.02	0.02	0.01	0.02
Parent in one-parent family			0.13	0.08	0.06
Partner in couple family			0.31	0.51	0.45
Partner/parent in two-parent family			0.18	0.21	0.03
Resident of non-private dwelling		0.01	0.01	0.01	0.03

Masterton District example 2013, male and female

	Age-group				
Living arrangement	**0-4**	**15-19**	**25-29**	**55-59**	**75-79**
Child in one-parent family	0.30	0.10	0.03	0.01	
Child in two-parent family	0.70	0.22	0.03		
Other living arrangement types	0.00	0.20	0.08	0.05	0.11
Parent in one-parent family	0.00	0.10	0.14	0.04	0.05
Partner in couple-without-children family		0.21	0.13	0.63	0.39
Partner/parent in two-parent family		0.14	0.51	0.09	0.02
Person in one-person household		0.05	0.07	0.19	0.44

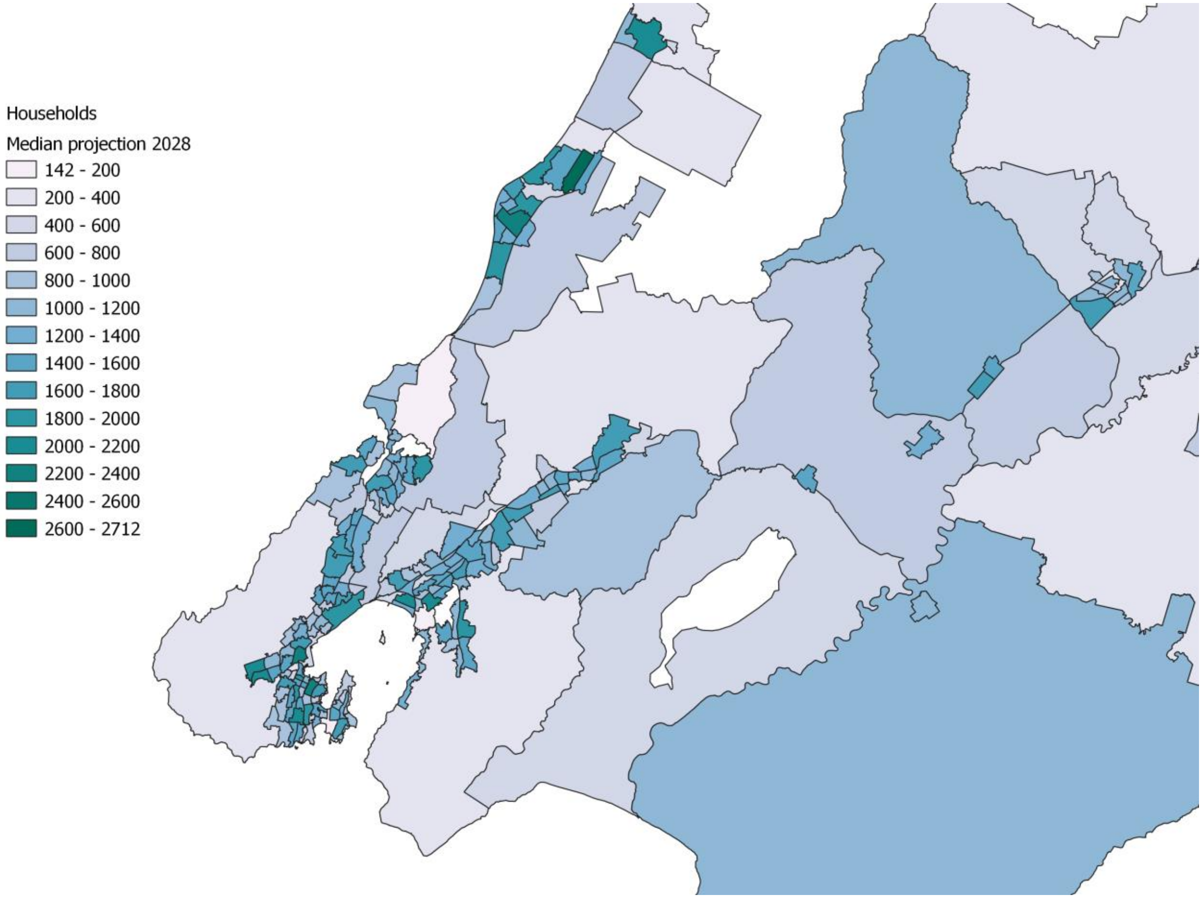


D. Housing

Household locations, dwelling types and vacancy rates

D.1 Residential location modelling methods

1. Statistical model predicts demand for houses within a district by local area (SA2) by household type and dwelling type, accounting for:
   - estimated cost of travel to employment
   - population density
   - location of similar households
2. Predicted housing demand is compared against existing stock and high-level estimates of housing development capacity:
   - based on the 2019 Housing and Business Development Capacity Assessments
   - councils' high-level assessments of the implications of recent and future policy and plan changes including initial assessments of the implications of government requirements to intensify in certain areas[1]
   - existing land zones and housing densities.
3. Excess demand, due to capacity constraints, is reallocated based on an "optimisation model" (linear programming) which accounts for:
   - land values (the lower the better)
   - excess development capacity (the more the better)
   - vacancies (the more the better)
4. We assume a long-run average <u>minimum</u> level of vacancies of:
   - 5% in typical residential areas
   - 20% in areas with high numbers of holiday homes.



1. E.g. https://planningforgrowth.wellington.govt.nz/__data/assets/pdf_file/0022/14953/PFG-Draft-Spatial-Plan-Growth-Figures-25-September.pdf

SENSE PARTNERS

DATA LOGIC ACTION