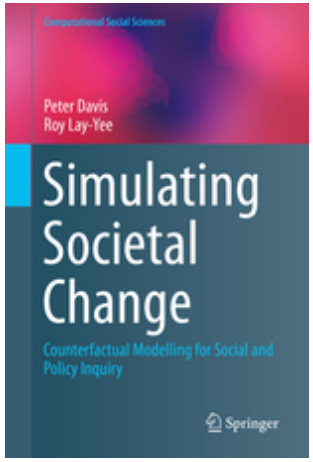
SociaLab:

A Census-based simulation tool for policy inquiry



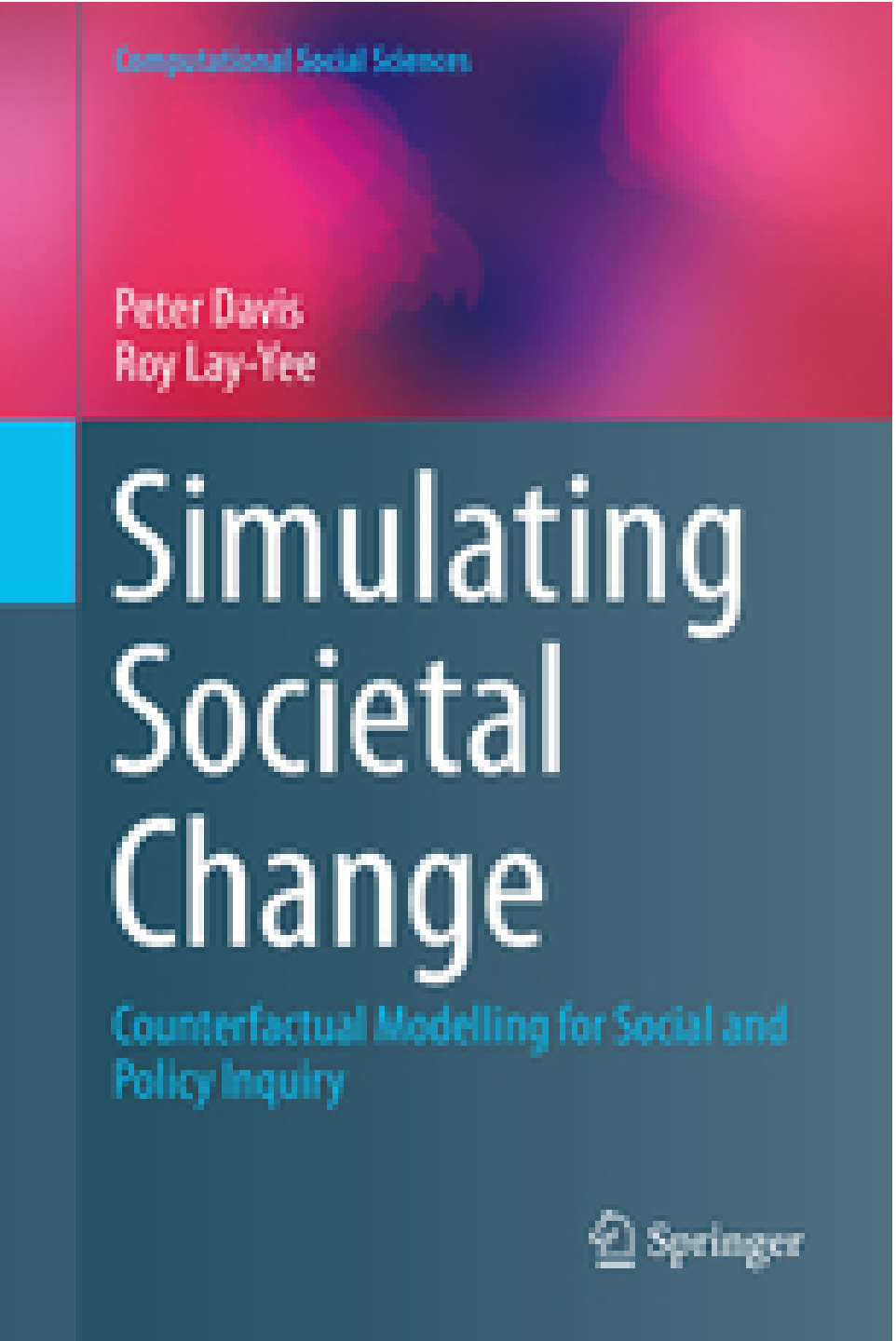
COMPASS Seminar Series

5th April, 2019

Peter Davis and Roy Lay-Yee

Department of Statistics, COMPASS Research Centre

University of Auckland



- Features a full-scale, realistic, working simulation model of society based on demographic and social information and transitioning through time
- Contains a comprehensive description of the construction of the working model, together with details of a novel open-source micro-simulation method that will facilitate transfer, application and learning across sites
- Includes worked examples of key policy and substantive questions tested with the simulation model against real data

Otamanewa Island, Manukau Harbour Sunrise – Bryan Lay Yee

Outline

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- **Foundational - Peter**
  - Inspiration, Objectives
  - Background, Framework
  - Counterfactuals
- **Operational - Roy**
  - Data, Statistical analysis, Simulation
  - Software
- **Aspirational - Peter**
  - Results
  - Strengths and limitations
  - Future

The Inspiration

COMPASS RESEARCH CENTRE

FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND

Whare Wānanga o Tāmaki Makaurau

- New Zealand
  - 1890-1920 a “social laboratory”
  - 1980-2010 a “transformational period”
- Canada
  - The Social Policy Simulation Database and Model
  - OpenM++ open source microsimulation platform

Three Objectives

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- To construct a “whole-of-society” simulation model of New Zealand over the period 1981-2006 using microdata from the longitudinal 5-yearly Census

- To formulate and test policy counterfactuals about a period of far-reaching change

- To develop an Inquiry tool – SociaLab – that is both interrogable and visual

The Background

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- Research programme in simulation at COMPASS
  - Marsden (2005) – residential segregation; partnership
  - HRC – primary care (2005), balance of care (2009)
  - MBIE – early life course (2009), knowledge laboratory (2013)
  - RSNZ, James Cook (2015) – “social laboratory”
  - TEC, Te Punaha Matatini CoRE (2015) – complexity science
  - MSD, Ernst and Young (2016) – vulnerable children investment

- Developments in data access at SNZ
  - Longitudinal Census (NZLC)
  - Remote access data facility (DataLab)
  - Integrated Data Infrastructure (IDI)

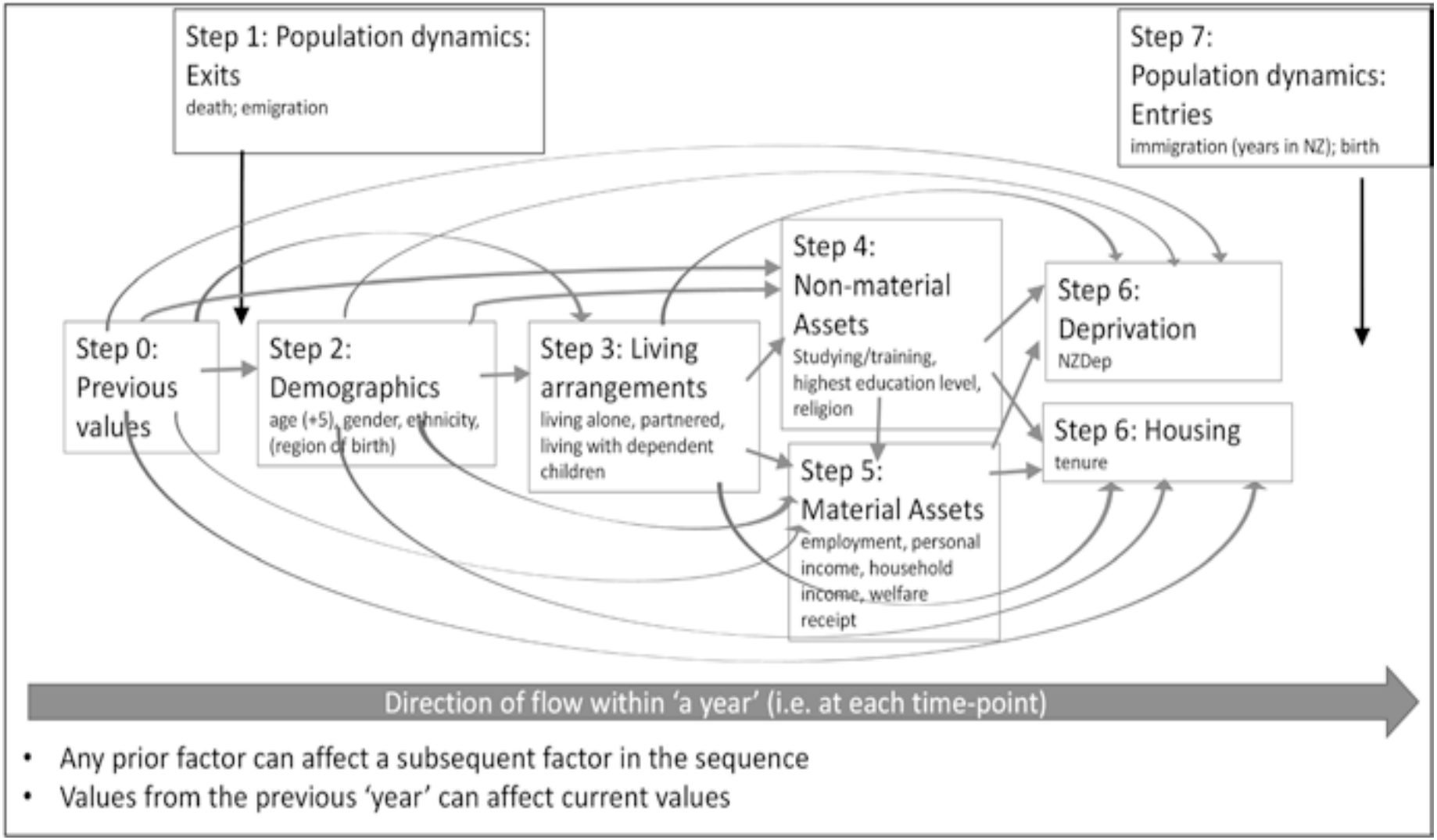
Conceptual Model

COMPASS RESEARCH CENTRE

FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND

Whare Wānanga o Tāmaki Makaurau

Simulation framework – at each time point



- Any prior factor can affect a subsequent factor in the sequence
- Values from the previous 'year' can affect current values

Operational Detail

COMPASS RESEARCH CENTRE

FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

Simulation framework - showing variables simulated

Step	Variables simulated	Age
0	Previous values	All
1	*Population dynamics – Exits*: death; emigration	All
2	*Demographics*: age (+5); gender (*time-invariant*); ethnicity (*time-invariant*); region of birth (*time-invariant*)	All
3	*Living arrangements*: retain three separate variables conditional on "living alone". Living alone – if yes, then partnered = no and living with dependent children = no. If not living alone, then partnership status (y/n); living with dependent children, i.e. age <15 or <18 if in full-time education or training (y/n)	0–14 (never living alone, nor partnered, nor living with dependent children) 15+
4	*Non-material assets*: in full-time education or training; education (highest level) [personal factors]; religion [household factor]	0–14 (household factors only)
5	*Material assets*: employment; personal income (CPI-adjusted); welfare receipt [personal factors]; household income (CPI-adjusted) [household factor]	10–14 15+
6	*Standard of living*: deprivation; housing tenure [household factors]	
7	*Population dynamics – Entries*: immigration (years in NZ: born in NZ/longer-term immigrant/recent immigrant<5 years); birth (new-born in dwelling, aged 0–4)	All; Women 15–49

"Seven Ages"

(**All the world's a stage**,
As You Like It, W. Shakespeare, First Folio,1623)

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- Early Childhood – health & thriving
- Childhood and Youth – education and readiness for life
- Young Adulthood – gaining & keeping employment
- Later Adulthood – settling into stable partnership
- Middle Adulthood – successfully raising families
- Older Life – retirement and successful ageing
- Later Life – the risks of dependency

The Framework: Early-life trajectories

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

Census	Age	*Living arrangements*	*Education*	*Employment*	*Housing*
1981	5	*Family of origin*	*At school*	*NA*	*NA*
1986	10				
1991	15				
1996	20	*Live alone/with others*	*Study-training*	*Employed or Unemployed or Home-maker*	*Own home or Rent*
2001	25	*Partnering*			
2006	30	*Having children*			

The Framework: Mid-life trajectories

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

Census	Age	*Living arrangements*	*Education*	*Employment*	*Housing*
1981	35	*Live alone/with others* *Partnering* *Having children*	*Study-training*	*Employed* *or* *Unemployed* *or* *Home-maker*	*Own home* *or* *Rent* *or* *Institution*
1986	40				
1991	45				
1996	50				
2001	55				
2006	60				

The Framework: Late-life trajectories

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

Census	Age	*Living arrangements*	*Education*	*Employment*	*Housing*
1981	65	*Live alone/with others* *Partnering* *Having children*	*Study-training*	*Employed* *or* *Unemployed* *or* *Home-maker* *or* *Retired*	*Own home* *or* *Rent* *or* *Institution*
1986	70				
1991	75				
1996	80				
2001	85				
2006	90				

"Capitals"

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- Material
  - Employment
  - Income
- Non-material
  - Education (human)
  - Religion (cultural)
  - [Social]
  - [Functional/health]

The Counterfactuals

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- "What If?" counterfactual scenarios
  - The liberalisation of immigration
  - Early childhood education, in-work family support
  - The "baby boomer" generation
  - The availability of life-course assets/capital
  - Future projections

Karamatura Stream, Waitakare Ranges – Bryan Lay Yee

Methods

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- Data preparation
  - Harmonise Longitudinal Census data series
  - Missing data imputation (using MICE method)
  - Supplement with data on “exits” and “entries”
- Statistical analysis (regression)
  - Use inter-censal data to estimate transitions
- Simulation – reproduces Census parameters
- Interrogation software
  - base model vs. scenarios with adjusted settings

Imputation

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

Imputation models for 'starting sample': Adults (15+)

Outcome	Type of model	Significant predictors ($p < 0.05$)
Partnership	Logistic	Age, gender, NZ European/Other ethnicity, birth region, living alone, living with children, in study/training, education, employment, welfare receipt, personal income, household income, deprivation, housing tenure
Education	Ordinal	Age, Māori ethnicity, Pacific ethnicity, NZ European/ Other ethnicity, birth region, years in NZ, living alone, partnership, living with children, in study/training, religion, welfare receipt, personal income, deprivation, housing tenure
Employment	Multinomial	Age, gender, Māori ethnicity, birth region, years in NZ, partnership, religion, welfare receipt, personal income
Welfare receipt	Logistic	Age, gender, Māori ethnicity, birth region, living alone, partnership, living with children, in study/training, education, employment, welfare receipt, personal income, household income, deprivation, housing tenure
Personal income	Linear	Age, gender, NZ European/Other ethnicity, living alone, partnership, living with children, in study/training, education, employment, welfare receipt, household income, housing tenure
Household income	Linear	Age, gender, birth region, living alone, partnership, living with children, in study/training, education, religion, employment, welfare receipt, personal income, deprivation, housing tenure
Deprivation	Ordinal	Māori ethnicity, Pacific ethnicity, NZ European/Other ethnicity, living alone, partnership, in study/training, education, welfare receipt, household income, housing tenure
Housing tenure	Logistic	Age, gender, Māori ethnicity, Pacific ethnicity, Asian ethnicity, NZ European/Other ethnicity, birth region, years in NZ, living alone, partnership, living with children, in study/training, education, employment, welfare receipt, personal income, household income, deprivation

Starting Sample *(1981)*

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

Pair = 0, Year = 1	Time-invariant	Time-variant	Categorisation
Age		(incremental)	raw
Gender	y		male/female
Ethnicity	y		binaries: European-&-other, Maori, Pacific, Asian
Number of years in NZ		(incremental)	categories: 'born in NZ', 5+ years, 0-4 years
Country of birth	y		region: NZ, Pacific, Asia, Europe, Americas, Middle East/Africa
New-born (in dwelling) (age 0-4)		y	yes/no
Living alone		y	yes/no
Partnership status		y	partnered-married (yes/no)
Living with dependent children		y	yes/no
Studying (in full-time education/training)		y	yes/no
Education (Highest level)		y	no qualification, school, post-school, tertiary
Religion		y	none, Christian, Other
Income (personal)		y	NZD - Consumers-Price-Index-adjusted to 2013 value
Income (household)		y	NZD - Consumers-Price-Index-adjusted to 2013 value
Employment		y	employed, unemployed, not in labour force
Welfare receipt		y	yes/no (income-tested benefits only)
Deprivation (area-based)		y	NZDep quintiles
Housing tenure		y	own / not own home

Predictive Equations with stochastic element

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- Example: Probability of 'being partnered' at age 25-34 in 1986 (derived from logistic regression)

$$P(\text{Outcome} = \text{Yes} \mid x_1, \ldots, x_n) = \frac{e^{\alpha + \beta_1 x_1 + \ldots + \beta_n x_n}}{1 + e^{\alpha + \beta_1 x_1 + \ldots + \beta_n x_n}}$$

where x1,…,xn denote significant predictors ($p<0.05$), β1,…,βn denote their coefficients, and α is the intercept

- Main predictors: census-pair (1981-86), previous partnership status (in 1981), age, gender, ethnicity, income, religion, beneficiary, deprivation
- Stochastically assign 'being partnered' or 'not' – random number compared to probability (from predictive equation)

COMPASS RESEARCH CENTRE

FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND

Whare Wānanga o Tāmaki Makaurau

Simulation Framework – *across time points*

BASE *(Starting Sample)*

SIMULATED

1981 → 1986 → 1991 → 1996 → 2001 → 2006

→ direction of flow across 'years'

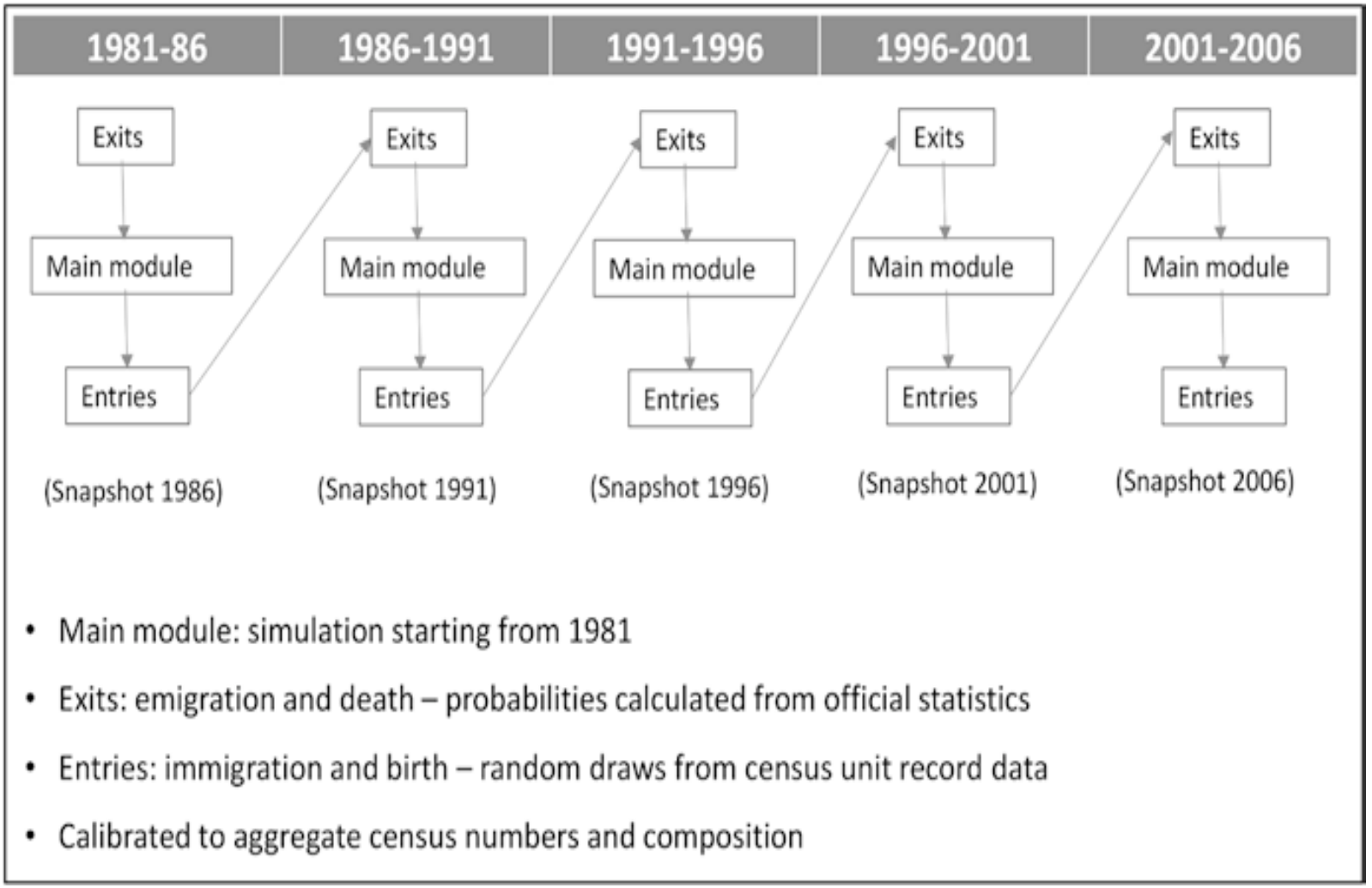
"Exits" and "Entries"

COMPASS RESEARCH CENTRE

FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND

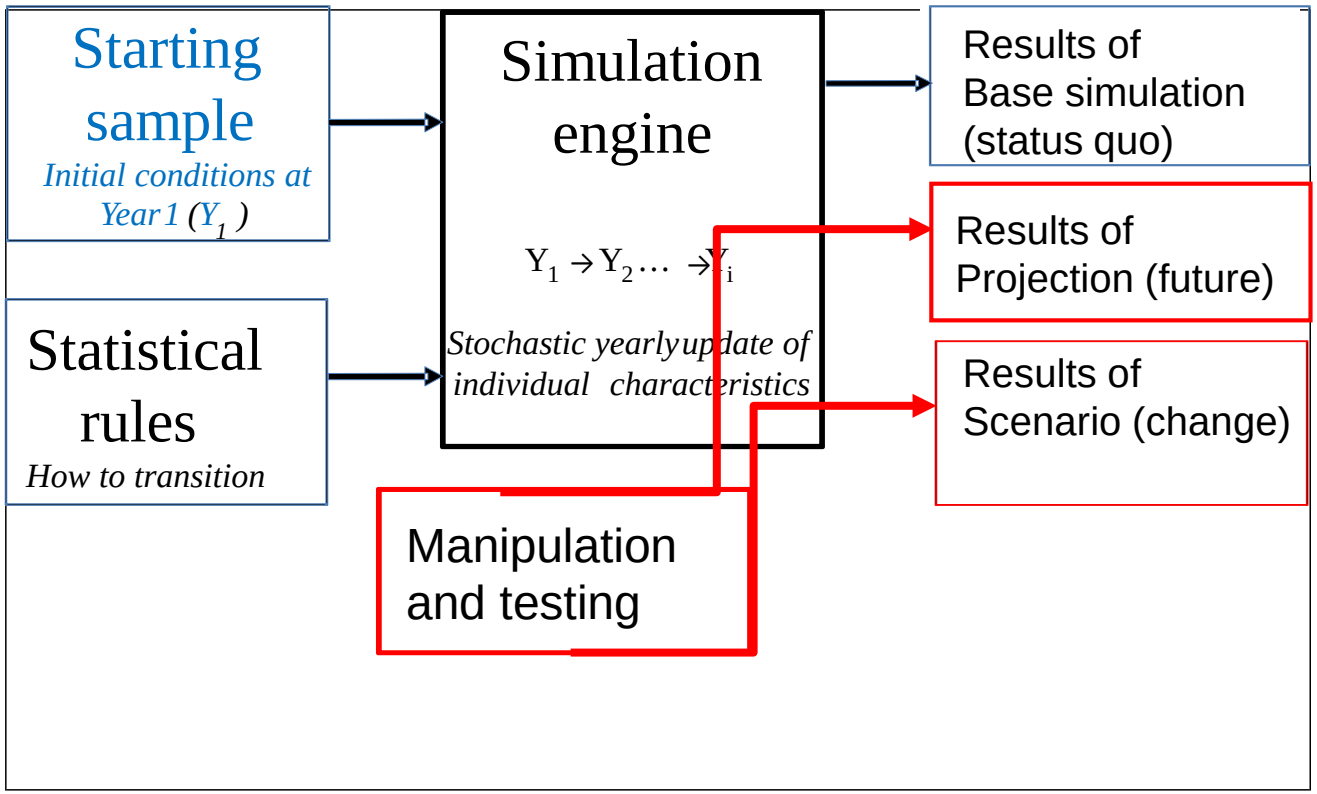
Whare Wānanga o Tāmaki Makaurau

Simulation schedule



- Main module: simulation starting from 1981
- Exits: emigration and death – probabilities calculated from official statistics
- Entries: immigration and birth – random draws from census unit record data
- Calibrated to aggregate census numbers and composition

Simulation engine - inputs and outputs



Simario programmed in R
Shiny web interface

The Software

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- Developed in **R** language (open source)
- **Simario** - to read data and implement simulation
  - source code at https://github.com/kcha193/simarioV2
- **Shiny** web-based application - user-friendly tool for interrogation and visualisation
  - source code at https://github.com/kcha193/SociaLabShiny
  - application at https://compassnz.shinyapps.io/SociaLabShiny
- We endeavour to deposit as much as possible in the public domain
- We hope to share our software with both developers and end-users in research and policy communities

Summing up microsimulation

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

1. Conceptualisation

1.1 Design simulation to mimic individual transitions through life course

2. Data preparation

2.1 Build base file. 1% sample of >3 million Census 1981 = 30,174

2.2 Harmonise 1981-2006 data series to generate usable inter-censal pairs

2.3 1% sample of 11.4 million individual inter-censal pairs = 110,000

2.4 Statistical analysis of inter-censal pair data to derive predictions of changes in individual states and behaviour through life course

3. Implementation

3.1 Starting with base sample, apply predictive equations progressively from 1981 to 2006 – in effect reproducing Census over time

3.2 Check these synthetic data against actual Census to make sure that we are reproducing it accurately "from the bottom up"

4. Application

4.1 Design and test scenarios by varying relevant factors in data and probabilities

Mt Taranaki, view from Tongaporutu Beach – Bryan Lay Yee

The Results

(of counterfactuals, for overall population)

COMPASS RESEARCH CENTRE

FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND

Whare Wānanga o Tāmaki Makaurau

- The liberalisation of immigration
  - we alter proportion of immigrants to that prevalent before liberalisation (i.e. decrease) – result is: decreased income
- Early childhood education, in-work family support
  - we alter proportion of mothers in employment to that before these programs (i.e. decrease) – little impact
- The “baby boomer” generation
  - we alter employment level among “baby boomer” women to resemble that of women from previous generation (i.e. decrease) – results are: increased welfare dependency, decreased income (esp. for women)
- The availability of life-course assets/capital
  - we alter education level in earlier years to that in a recent year (i.e. increase) – results are: decreased welfare dependency, increased employment and income (esp. for Māori and women)

The Results
(of forward projections)

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- Projection of current 'base' trends into the future
  - demographic ageing
  - increasing ethnic diversity
  - rising immigration
  - changing pattern of living arrangements departing from traditional norm
  - increasing secularisation
  - development of a more highly educated workforce
  - greater participation in paid employment
  - stabilising of levels of dependence on welfare benefits
  - rising average incomes
  - lower levels of deprivation
  - declining home ownership

Discussion

COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- Strengths
  - Dynamic microsimulation model
  - Whole of society, linked Census data
  - Inquiry tool for scenario testing – SociaLab

- Limitations
  - Census data “thin”, linkage rates incomplete
  - Counterfactual scenario tests not striking
  - Not fully delivered on analytical framework

The Future

COMPASS RESEARCH CENTRE

FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND

Whare Wānanga o Tāmaki Makaurau

- Include household structure
- Individual transitions through further censuses
- Use more administrative data (Integrated Data Infrastructure)
- Richer data (beyond Census)
- Deliver on analytical framework
- Open source tool and data for all (e.g. NGOs)

SociaLab: A Census-based simulation tool for policy inquiry

QUESTIONS, COMMENTS, DISCUSSION

Sunset, Awhitu Peninsula – Bryan Lay Yee