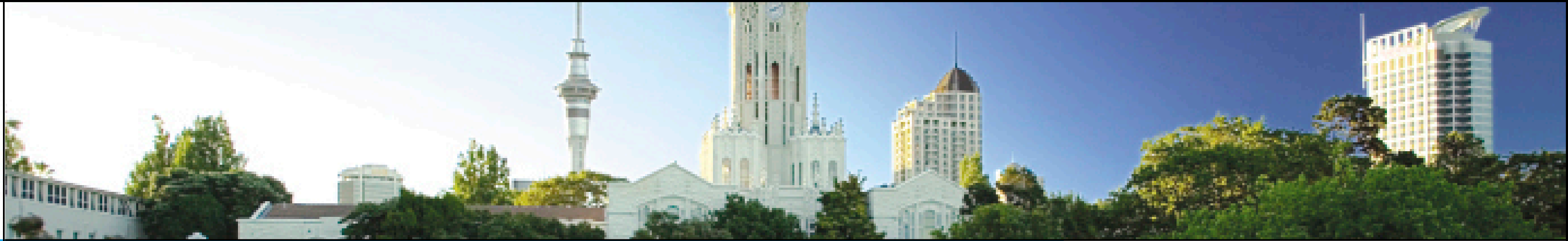
New Zealand



simario: An R Package for Dynamic Microsimulation

COMPASS Colloquium, July 2014

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

Jessica McLay

COMPASS Research Centre

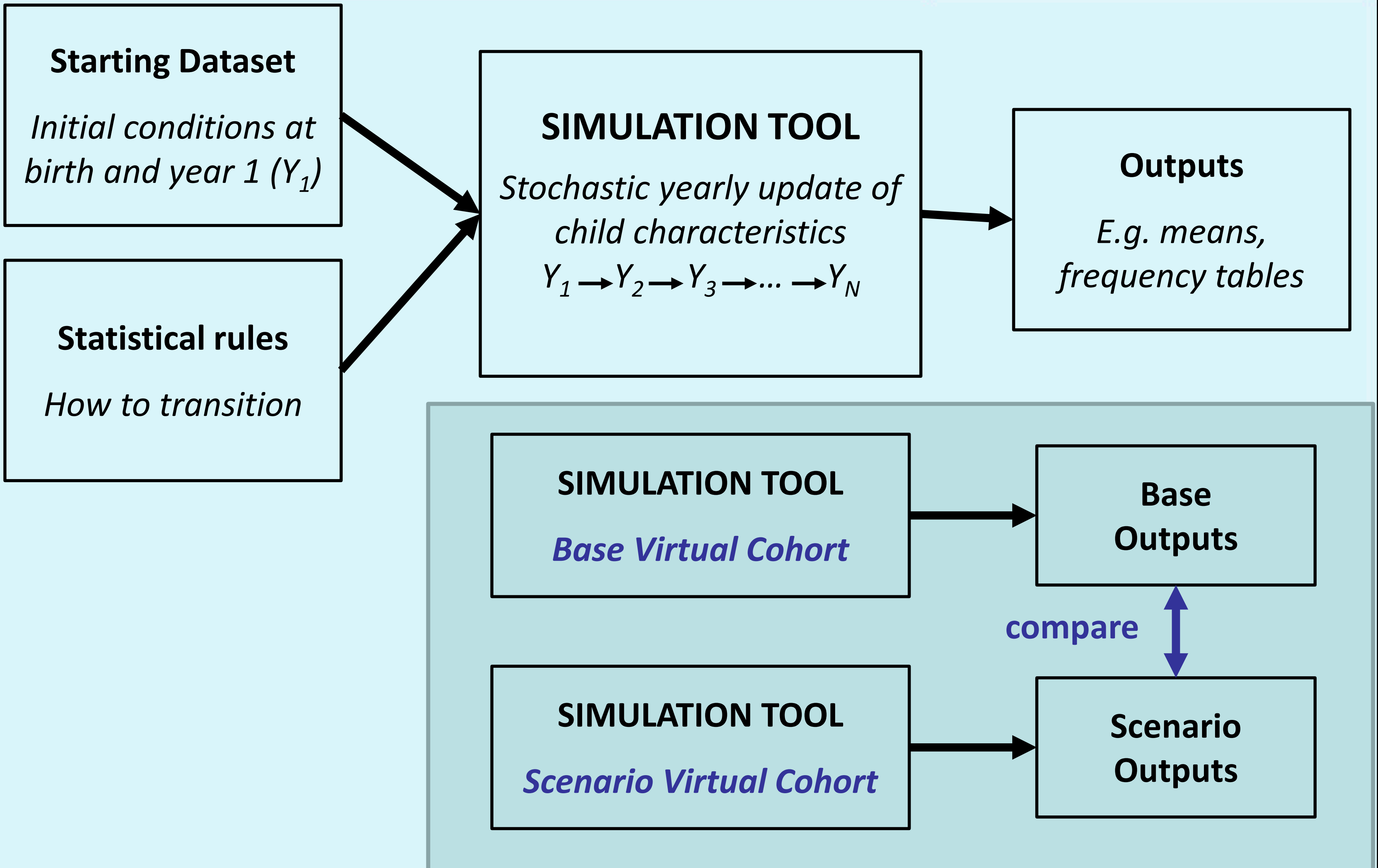
Faculty of Arts

University of Auckland

MINISTRY OF BUSINESS, INNOVATION & EMPLOYMENT
HIKINA WHAKATUTUKI

The University of Auckland

What is Dynamic Microsimulation?



Software Used for Microsimulation

- MODGEN based in C++
- C#
- SAS
- Java
- LIAM, LIAM2
- R package ‘sms’ for spatial microsimulation for small area population estimates

Why R?

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

- Open source and free
  - Anyone can install, use and further develop
  - Availability of public critique and refinement
  - Existing user base
- Designed for data analysis and manipulation
- Flexible
  - Massive 3rd party contribution
  - Libraries for most anything statistical you may want to do

The simario R package

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

- Purpose: to provide a flexible framework of functions for creating a microsimulation in R
- R package: A collection of related R functions and other R objects (e.g. a dataset)
- *Given a few csv files, use simario functions to programme your microsimulation model from start to finish, then run scenarios*
- Illustration of simario:
  - Setting up (initiation files)
  - The simulation process
  - Outputs
  - Running scenarios
  - Viewing results

Content Note:

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

- The material in this presentation is most relevant to the programmer/analyst who is developing a micrsimulation
- This level of knowledge is not required to use the MELC interrogation tool that COMPASS has developed.
- The MELC interrogation tool is a stand-alone piece of software that has been distributed to a number of government ministries and groups. It runs simario code “behind the scenes” and the users do not need any knowledge of R to use it.

Two Types of Functions

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

- Simario functions:
  - Generic stand-alone functions, no reference to objects outside of the function
- Project-specific shell functions
  - Shell/outline for the programmer to complete with details of their specific microsimulation project

Illustration: The MEL-C Microsimulation Model

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

Structural level

Child characteristics
- *gender*
- *ethnicity*

Parental characteristics
- *age at birth of child*
- *ethnicity*
- *education level*

Socio-economic position
- *SES at birth of child*
- *(single-parent status at birth)*

Other factors
e.g.:
- *Birth weight*
- *Gestational age*
- *Birth order*
- *Smoking and drinking during pregnancy*
- *breastfeeding*

Intermediate level

Family/household characteristics
e.g. single-parent status, number of children, household size

Parental employment
e.g. employed / welfare dependent

Material circumstances
e.g. accommodation type, owned/rented, overcrowding

Psychosocial factors
e.g. change of parents, change of residence

Behavioural factors
e.g. parental smoking

Outcome

Health service use
GP visits, hospital admissions, hospital outpatient attendances

Education
Reading ability

Social/Justice
Conduct disorder

The University of Auckland
New Zealand

Model Initiation

- Files needed prior to programming the simulation in R
  - Starting dataset (.csv)
  - Statistical sub-models (.csv)
  - Data dictionary file (.csv)

Starting Dataset

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

- One row per individual
- Provides the starting values from which to simulate all other variables and years
- MELC: Synthetic dataset based on the 2006 census

ID	age	gptotvis	SES	fsmoke	fhrswrk	z1overcrowd
1	1	13	2	0	40	0
2	1	2	3	0	40	0
3	1	5	3	0	40	0
4	1	6	3	20	45	1
5	1	3	1	0	45	0
6	1	13	3	0	0	0
7	1	5	1	0	55	0
8	1	6	3	0	0	0
9	1	4	1	11	37	0
10	1	6	3	0	40	1

Statistical Sub-Model

Simplified sub-model for GP visits:

Effect	ClassVal0	Estimate
Intercept		-0.103
age		-0.4211
SES	1	0.1065
SES	2	0.09175
SES	3	0
z1overcrowd	1	0.09248
fsmoke		0.003155
age*z1overcrowd	1	0.07845
age*fsmoke		0.002327
_Alpha		0.6234

E.g.: MELC Data Dictionary

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

The University of Auckland New Zealand

Varname	Description	Codings_Expr
age	Age	
SESBTH	SES at birth	c("Professional"=1, "Clerical"=2, "Semi-skilled"=3)
fsmoke	Father's cigarettes smoked per day	
fsmoke_previous	Father cigarettes smoked previous	
fhrswrk	Father's hours worked per week	
fhrswrk_previous	Father hours worked previous	
z1overcrowd	Overcrowding	
z1overcrowd_previous	Overcrowding previous	
gptotvis	GP visits	
r1stfeduc	Father's education at child's birth	c("Tertiary"=1, "Secondary"=2, "No formal qualifications"=3)

Model Initiation

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

- Complete project-specific initiation function
  - Shell function provided (e.g. `initMELC()`)
  - Point to the initiation csv files just discussed,
  - which are used to set up objects in the R environment
    - Imports statistical sub-models and starting dataset,
    - Creates empty lists and matrices that will be filled during the simulation
- Fill in other project-specific functions which are called by the initiation function
  - E.g. a "binbreaks" function that defines categorisations for continuous variables
    - number cigarettes smoked per day: 0, 1-10, 11-20, 20+

The Simulation Process

- Simulate Run 1:
  - Year 1:
    - Simulate family characteristics
    - Simulate employment
    - Simulate smoking
    - Simulate health service use
    - Simulate reading
    - Save simulated values
  - Year 2:
    - Simulate family characteristics
    - Simulate employment
    - . . .
    - Save simulated values
  - . . .
  - Year N
    - Simulate family characteristics
    - . . .
    - Save simulated values
  - Calculate summary statistics
- Simulate Run 2:
  - Year 1
  - . . .
  - Year N
  - Calculate summary statistics
- . . .
- Simulate Run M
- Calculate means of summary statistics

```
simulateRun <- function() {
    for (year in 2:NUM_YEARS) {
        simulate_family_characteristics()
        simulate_employment()
        simulate_smoking()
        simulate_health_service_use()
        simulate_reading()
        store_current_values_in_outcomes()
    }
}
```

```
for (i in 1:total_runs) {
    simulateRun()
    map_outcomes_to_run_results()
}
collate_all_run_results()
```

The Simulation Process

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

The University of Auckland New Zealand

GP visits outcomes from 1 run

Year

Individual

	1	2	3	4	5	6	7	8	9	10
1	8	5	4	9	3	12	7	3	0	0
2	5	10	7	15	3	2	2	2	0	0
3	6	5	2	0	1	1	4	2	2	0
4	5	3	3	1	7	3	3	1	0	0
5	2	2	4	4	16	7	3	1	0	0
6	6	4	0	0	1	3	3	0	0	0
7	3	1	1	1	2	2	0	0	0	1
8	5	7	10	6	6	3	1	1	3	0
9	3	2	5	5	6	1	0	0	2	0
10	5	5	5	5	2	3	1	0	0	2
11	4	4	7	3	0	0	0	5	2	0
12	3	7	5	11	5	1	0	1	3	2
13	11	9	2	6	4	3	2	0	1	0
14	4	3	2	5	16	19	7	5	5	0
15	7	7	0	2	4	6	5	0	0	0
16	1	1	3	7	9	4	0	1	4	1
17	5	4	2	5	2	1	0	0	1	0
18	2	4	4	9	6	1	1	0	5	0
19	3	4	8	4	3	2	2	3	1	0
20	5	2	3	12	9	1	4	0	1	1
21	2	2	3	3	5	2	0	0	1	2
22	7	7	6	2	4	1	3	2	4	2

The Simulation Process

run_results

GP visits means run 1

Year	1
1	6.00720
2	5.42920
3	4.13800
4	4.53000
5	4.13540
6	3.54420
7	2.23240
8	2.76500
9	2.47240
10	1.27740

GP visits means run 2

Year	1
1	6.00720
2	5.40320
3	4.06380
4	4.46580
5	4.01720
6	3.58100
7	2.21680
8	2.75340
9	2.52060
10	1.34520

run_results_collated

GP visits mean of means

Year	Mean	Lower	Upper
1	6.007200	5.920548	6.093852
2	5.416200	5.403850	5.428550
3	4.100900	4.065655	4.136145
4	4.497900	4.467405	4.528395
5	4.076300	4.020155	4.132445
6	3.562600	3.545120	3.580080
7	2.224600	2.217190	2.232010
8	2.759200	2.753690	2.764710
9	2.496500	2.473605	2.519395
10	1.311300	1.279095	1.343505

Predict and Simulate Functions

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

	predSimNorm()	predSimBinom()	predSimPois()	predSimNBinom()
Variable type	Continuous	Dichotomous	Continuous	Continuous
Type of statistical model	Linear regression model	Logistic regression model	Poisson regression model	Negative binomial regression model
Get predicted value for each individual				
Random draw from	Normal distribution	Binomial distribution	Poisson distribution	Negative binomial distribution
Other parameters	Standard deviation = residual standard error from model			Dispersion parameter

Also available: predSim 'Select' functions to select different models depending on the value of a second variable
e.g. value of overcrowding in the previous year

The University of Auckland New Zealand

Outputs (Collated Results)

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

The University of Auckland New Zealand

Means:
GP Visits

Year	Mean	Lower	Upper
1	6.015600	5.927938	6.103262
2	5.441700	5.440845	5.442555
3	4.082700	4.075575	4.089825
4	4.508400	4.500800	4.516000
5	4.114700	4.070335	4.159065
6	3.618900	3.598095	3.639705
7	2.313900	2.308865	2.318935
8	2.867700	2.822005	2.913395
9	2.584300	2.541835	2.626765
10	1.340800	1.314010	1.367590

Percentages:
Over-crowding

Year	Overcrowding (%) Mean	Overcrowding (%) Lower	Overcrowding (%) Upper
1	17.90000	16.83742	18.96258
2	19.39000	19.34250	19.43750
3	21.05000	20.88850	21.21150
4	21.37000	21.13250	21.60750
5	19.56000	19.52200	19.59800
6	23.57000	23.50350	23.63650
7	23.00000	22.81000	23.19000
8	22.71000	22.60550	22.81450
9	22.26000	21.93700	22.58300
10	20.90000	20.82400	20.97600

Outputs (Collated Results)

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

Quantiles: Hours Worked by the Father

Year	Min	10th	25th	50th	75th	90th	Max
1	0.0000000	0.0000000	40.00000...	40.00000...	50.00000...	60.00000...	98.00000...
2	0.0000000	4.5000000	35.00000...	43.00000...	50.50000...	56.00000...	82.00000...
3	0.0000000	20.50000...	36.00000...	43.50000...	50.50000...	56.50000...	78.00000...
4	0.0000000	0.0000000	35.00000...	43.00000...	50.00000...	56.05000...	80.50000...
5	0.0000000	0.0000000	33.50000...	42.50000...	50.00000...	56.00000...	77.50000...
6	0.0000000	17.00000...	34.50000...	43.50000...	51.50000...	59.00000...	90.00000...
7	0.0000000	22.00000...	35.00000...	44.00000...	52.00000...	60.00000...	87.00000...
8	0.0000000	22.00000...	35.00000...	44.00000...	52.62500...	60.00000...	87.00000...
9	0.0000000	21.50000...	35.00000...	44.00000...	52.12500...	60.00000...	87.00000...
10	0.0000000	20.00000...	35.00000...	44.00000...	52.00000...	59.50000...	87.50000...

Outputs (Collated Results)

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

The University of Auckland New Zealand

Summaries: Hours Worked by the Father

Year	Min. Mean	Min. Lower	Min. Upper	1st Qu. Mean	1st Qu. Lower	1st Qu. Upper	Median Mean	Media
1	0.00000	0.00000	0.00000	40.00000	40.00000	40.00000	40.00000	
2	0.00000	0.00000	0.00000	35.00000	35.00000	35.00000	43.00000	
3	0.00000	0.00000	0.00000	36.00000	36.00000	36.00000	43.50000	
4	0.00000	0.00000	0.00000	35.00000	35.00000	35.00000	43.00000	
5	0.00000	0.00000	0.00000	33.50000	33.02500	33.97500	42.50000	
6	0.00000	0.00000	0.00000	34.50000	34.02500	34.97500	43.50000	
7	0.00000	0.00000	0.00000	35.00000	35.00000	35.00000	44.00000	
8	0.00000	0.00000	0.00000	35.00000	35.00000	35.00000	44.00000	
9	0.00000	0.00000	0.00000	35.00000	35.00000	35.00000	44.00000	
10	0.00000	0.00000	0.00000	35.00000	35.00000	35.00000	44.00000	

Outputs (Collated Results)

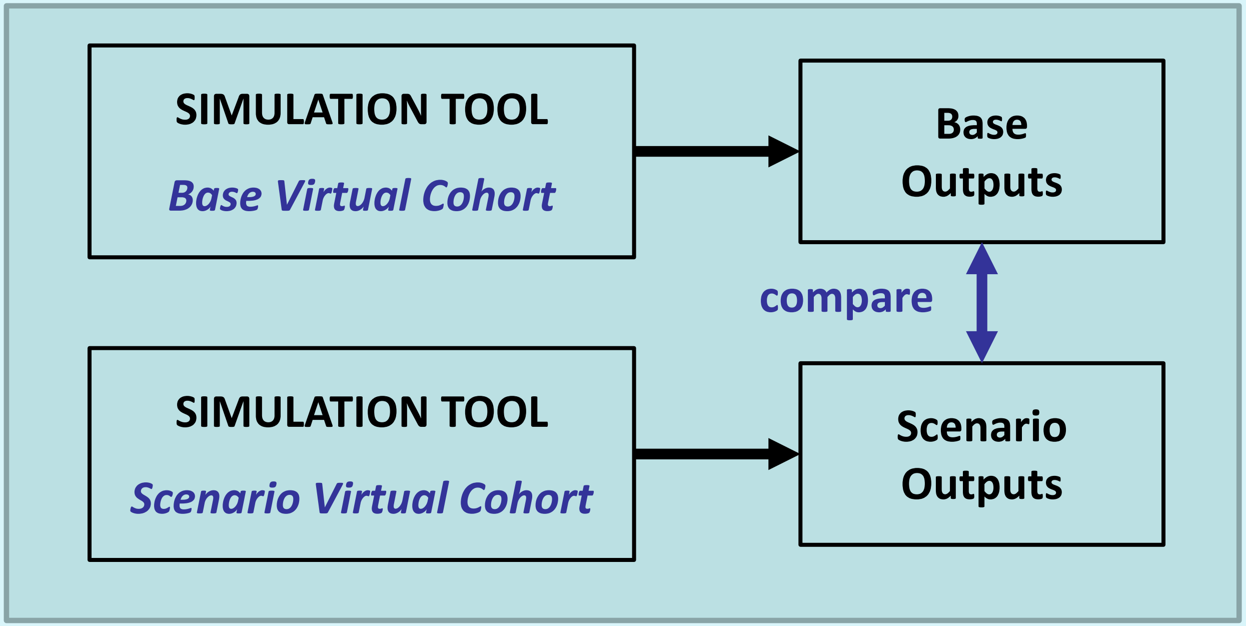
COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

Percentages for Continuous Variables: Hours Worked by the Father

Year	0	1-20	21-35	36-40	41-45	46-50	51+
1	14.16	2.84	3.64	31.06	14.38	14.02	19.90
2	9.32	0.76	14.26	14.06	18.64	17.78	25.18
3	8.98	0.52	13.56	14.94	18.68	17.66	25.66
4	11.06	0.44	13.52	14.58	17.94	17.52	24.94
5	14.56	0.68	14.16	13.16	17.24	16.26	23.94
6	9.04	1.78	17.34	12.46	15.62	14.66	29.10
7	7.20	1.86	15.72	13.28	15.78	15.74	30.42
8	7.74	1.10	16.50	12.44	14.52	15.80	31.90
9	8.20	1.64	15.24	12.90	16.06	14.74	31.22
10	8.58	1.58	16.36	12.06	15.28	15.36	30.78

Running Scenarios

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



Running Scenarios: Scenario Specification

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

Fill in cat.adjustment matrices (created by the initiation function)

Welfare

	No benefit (%)	Receiving benefit (%)
Year 1	NA	NA
Year 2	NA	NA
Year 3	NA	NA
Year 4	NA	NA
Year 5	0.85	0.15
Year 6	0.87	0.13
Year 7	0.89	0.11
Year 8	0.91	0.09
Year 9	0.93	0.07
Year 10	0.95	0.05

Father's Smoking

	0 (%)	1-10 (%)	11-20 (%)	21+ (%)
Year 1	0.925	0.029	0.032	0.014
Year 2	NA	NA	NA	NA
Year 3	NA	NA	NA	NA
Year 4	NA	NA	NA	NA
Year 5	NA	NA	NA	NA
Year 6	NA	NA	NA	NA
Year 7	NA	NA	NA	NA
Year 8	NA	NA	NA	NA
Year 9	NA	NA	NA	NA
Year 10	NA	NA	NA	NA

The University of Auckland New Zealand

Example Scenario

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

SES Distribution in the Base Simulation

	Mean (%)	Lower	Upper
Professional	24.26000	23.07185	25.44815
Clerical	41.08000	39.71633	42.44367
Semi-skilled	34.66000	33.34093	35.97907

Fill in cat.adjustment matrix with proportions for scenario

	Professional (%)	Clerical (%)	Semi-skilled (%)
SES group	0.35	0.40	0.25

The University of Auckland New Zealand

Running Scenarios: "Adjusting" Data

SES	Proportion in Base Simulation	Requested Proportions for Scenario	Number in Base Simulation	Number Needed to Match Requested Proportions	Number to Change
Professional	.252	.35	1262	1750	-488
Clerical	.399	.40	1997	Move 488 2000	-3
Semi-skilled	.348	.25	1741	1250	491

After one step:

SES	Number	Proportion	Number to Change
Professional	1262 + 488 = 1750	.350	0
Clerical	1997 - 488 = 1509	.302	-491
Semi-skilled	1741	.348	491

Move 491

After two steps:

SES	Number	Proportion	Number to Change
Professional	1750	.35	0
Clerical	1509 + 491 = 2000	.40	0
Semi-skilled	1741 - 491 = 1250	.25	0

Subgroup Scenarios

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

Year	Professional Overcrowding	Clerical Overcrowding	Semi-skilled Overcrowding
1	5.358615	13.193768	32.256203
2	9.810387	15.092502	30.986728
3	11.788953	17.672833	31.390652
4	12.448475	17.770204	32.544720
5	12.283594	15.871470	27.755338
6	13.602638	18.889971	33.929602
7	14.756801	18.257059	33.121754
8	14.344600	19.425511	34.160415
9	14.674361	18.354430	32.313907
10	13.355317	17.380721	30.929025
11	11.046991	15.968841	28.216965
12	9.480627	12.512171	22.677438
13	6.183017	11.246349	19.388344

Scenario: Reduce overcrowding **for the lowest SES group** to the level observed in the clerical group

Subgroup Scenarios: R Code

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

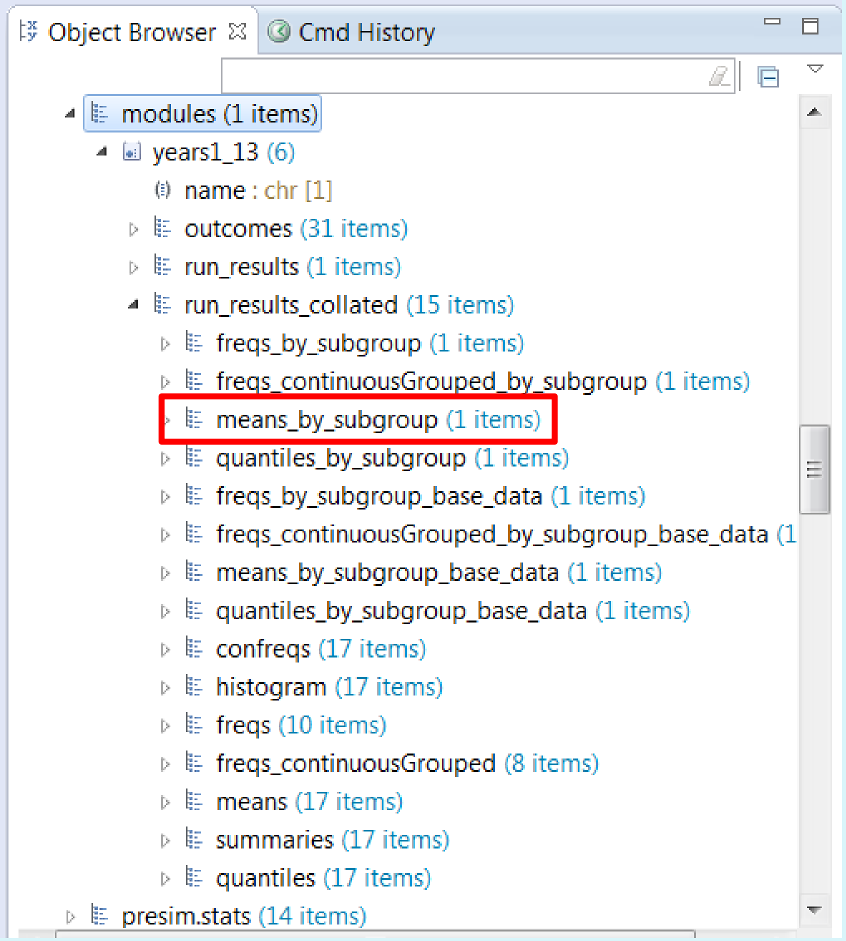
The University of Auckland New Zealand

```
subgroupExpression <- "SESLvl3==1"
setGlobalSubgroupExpression(subgroupExpression)

env.scenario$cat.adjustments$z1overcrowd[1,] <- c(.868, .132)
env.scenario$cat.adjustments$z1overcrowd[2,] <- c(.848, .152)
env.scenario$cat.adjustments$z1overcrowd[3,] <- c(.838, .162)
env.scenario$cat.adjustments$z1overcrowd[4,] <- c(.833, .167)
env.scenario$cat.adjustments$z1overcrowd[5,] <- c(.842, .158)
env.scenario$cat.adjustments$z1overcrowd[6,] <- c(.811, .189)
env.scenario$cat.adjustments$z1overcrowd[7,] <- c(.812, .188)
env.scenario$cat.adjustments$z1overcrowd[8,] <- c(.819, .181)
env.scenario$cat.adjustments$z1overcrowd[9,] <- c(.831, .181)
env.scenario$cat.adjustments$z1overcrowd[10,] <- c(.850, .169)
env.scenario$cat.adjustments$z1overcrowd[11,] <- c(.878, .150)
env.scenario$cat.adjustments$z1overcrowd[12,] <- c(.896, .122)
env.scenario$cat.adjustments$z1overcrowd[13,] <- c(.888, .104)

env.scenario$simulate()
```

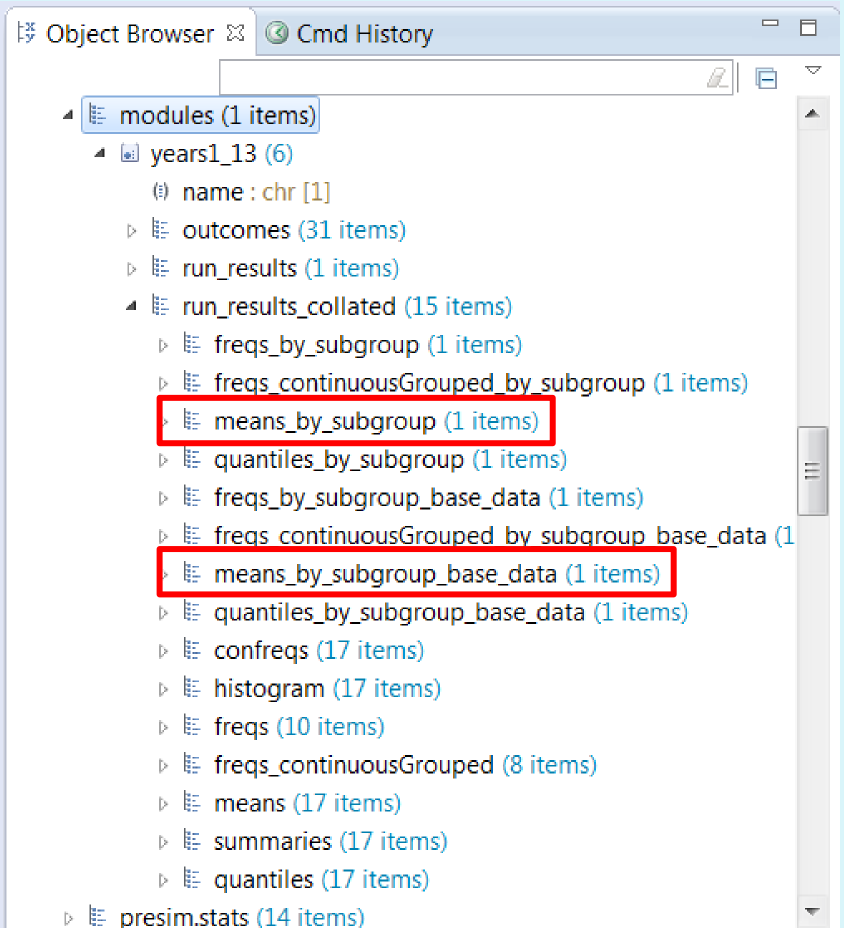
Note: This is how the programmer runs scenarios . Users using the software tool developed by COMPASS do not need to do this as there is a pretty diaglogue box for them

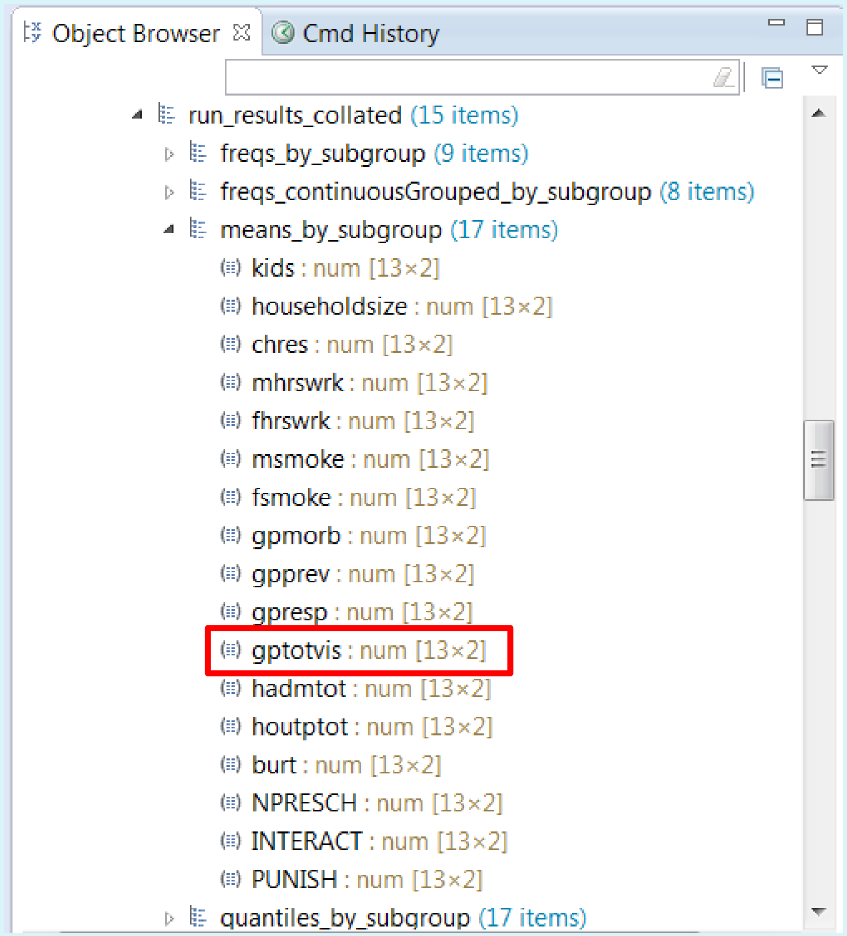
Subgroup Scenarios: Outputs

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



Subgroup Scenarios: Outputs





Subgroup Scenarios: Outputs

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

Mean Number of GP Visits in Base Simulation

Year	In subgroup Mean	In subgroup Lower	In subgroup Upper
1	6.381420	6.213516	6.549323
2	5.308713	5.292816	5.324611
3	4.146855	4.102726	4.190984
4	4.498269	4.462911	4.533627
5	4.160127	4.139022	4.181232
6	3.454414	3.436598	3.472230
7	2.167917	2.160791	2.175043
8	2.670802	2.642571	2.699033
9	2.476630	2.452510	2.500750
10	1.355164	1.320903	1.389426

Mean Number of GP Visits in Scenario Simulation

Year	In subgroup Mean	In subgroup Lower	In subgroup Upper
1	6.381420	6.213516	6.549323
2	5.385459	5.339960	5.430958
3	4.147721	4.146076	4.149365
4	4.521062	4.518869	4.523254
5	4.130121	4.023500	4.236743
6	3.406232	3.366763	3.445701
7	2.070110	2.056679	2.083540
8	2.597519	2.533107	2.661930
9	2.343047	2.325779	2.360314
10	1.168494	1.147115	1.189873

The tableBuilder() Function

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

The University of Auckland New Zealand

- Results can be viewed by
  - Looking at outputs automatically created
  - Using R manually to investigate/summarise the simulated data (which is stored for each run)
  - Using the tableBuilder() function

The tableBuilder() Function

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

Argument	Specifies:	Options / Examples
envName	Which set of simulated data to use	Base or scenario
statistic	Which statistic to calculate	frequencies, means, quintiles
variableName	The variable on which to calculate the statistic	e.g. welfare, fhrswrk, fsmoke
grpbyName	An optional variable to group the results by	e.g. childethn, SES
CI	Whether to calculate confidence intervals	TRUE or FALSE
logiset	An optional string expression that defines a group. Only data from this group will be using in calculating the specified statistics.	e.g. fhrswrk>0

The tableBuilder() Function

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

```
tableBuilder(envName="Base", statistic="means",
variableName="fhrswrk", grpbyName="FathersEd", CI=FALSE)
```

```
 Father's education at child's birth
NA          Tertiary Secondary No formal qualifications
  1         40.44891  39.88306                 36.36246
  2         42.29745  41.07103                 36.32484
  3         42.54471  42.02886                 35.58940
  4         40.72445  41.96477                 33.05138
  5         39.02646  41.36788                 30.09385
  6         41.93294  42.47657                 34.99960
  7         42.88367  42.95015                 37.16222
  8         43.14005  43.11582                 38.42112
  9         43.10036  43.07646                 37.77589
  10        43.12409  43.09333                 37.41505
  All Years 42.07092  42.19064                 35.91337
attr(,"meta")
    varname   grpby.tag
  "fhrswrk" "r1stfeduc"
```

Summary: Limitations and Disadvantages of simario

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

- Most suited to dynamic closed cohort models
  - Simulating a set group of individuals over time (no current capacity for individuals to enter or leave the simulation, births and deaths)
- Need to be confident using R
- Level of complexity to fitting all the functions together
- No current capacity for scenarios where the effect of one variable on another is changed

Summary: Advantages of simario

- Simario provides a framework for creating a microsimulation model in R
- Good for scenarios that examine the effect of changing peoples actions (e.g. smoking cessation, deciding to breastfeed)
  - (Scenarios like changing a tax rate could be incorporated by writing additional R code at the appropriate place in the simulation)
- Very flexible
  - Simulating variables
    - Currently has functions to simulate from 5 of R's many statistical distributions
    - Additional R code can be included at any point during the simulation, e.g. to simulate from a different distribution or check simulated values do not exceed an upper limit
  - Specifying outputs
  - For a given variable, can use different parameters (statistical sub-models) for different cases
  - Confident R programmers can expand and change functions to suit their own purposes

Acknowledgments

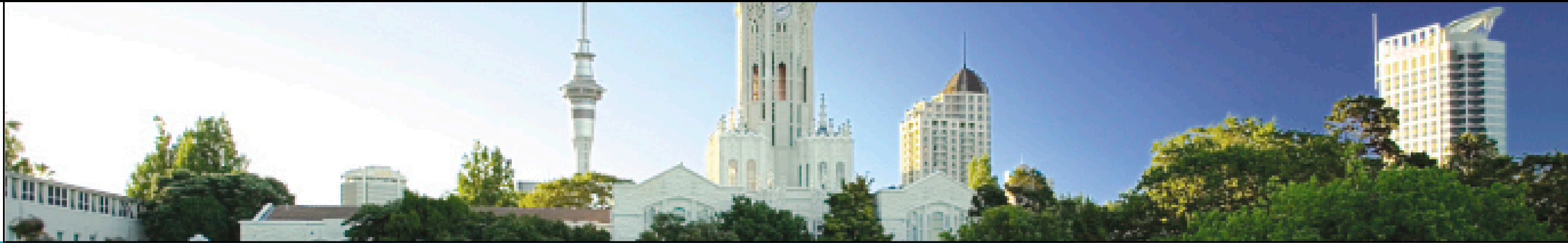
COMPASS RESEARCH CENTRE
FACULTY OF ARTS
THE UNIVERSITY OF AUCKLAND
Whare Wānanga o Tāmaki Makaurau

- Oliver Mannion
- Barry Milne
- Janet Pearson
- Mengdan Yu
- Roy Lay-Yee
- Martin von Randow
- Peter Davis

More information:

- simario to be published as an R package on CRAN for free download
- Article providing instructions on how to use simario to be published
- Code currently available on google code (search "simario")
- jessica.mclay@auckland.ac.nz

The University of Auckland New Zealand



Appendix

COMPASS RESEARCH CENTRE

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

The Simulation Process

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

Simulating Reading score: Rule from statistical model:

E[reading score] = 13.00 + .91*reading.score.previous + .07*months.breast.fed + 1.04*father.tertiary.qualification + .87*father.secondary.qualification

	Child 1
Characteristics	
Reading score at age 8	40
Number of months breast fed	12
Father's Education	Secondary
Predicted reading score at age 9	13.00 + .91*40 + .07*12 + .87 = 50.58
Random draw from a normal distribution	50.23
Reading score assigned at age 9	50

Apply Rule

Expected value

Stochastic component

The University of Auckland New Zealand