REMOVING SIMULATION NOISE IN ACTIVITYSIM

JAN ZILL, VEITCH LISTER CONSULTING

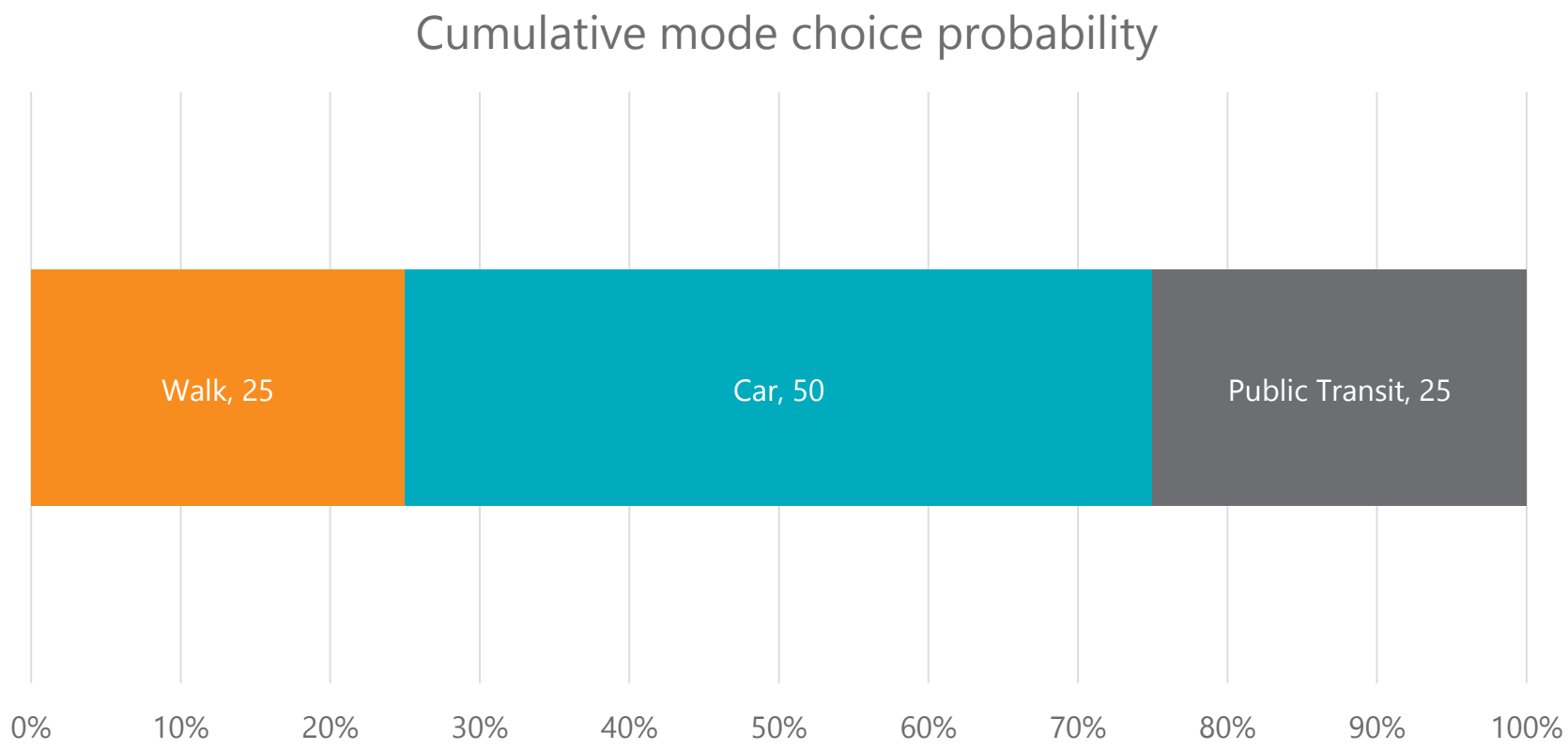
PRESENTATION TO ACTIVITYSIM CONSORTIUM

1 March 2023

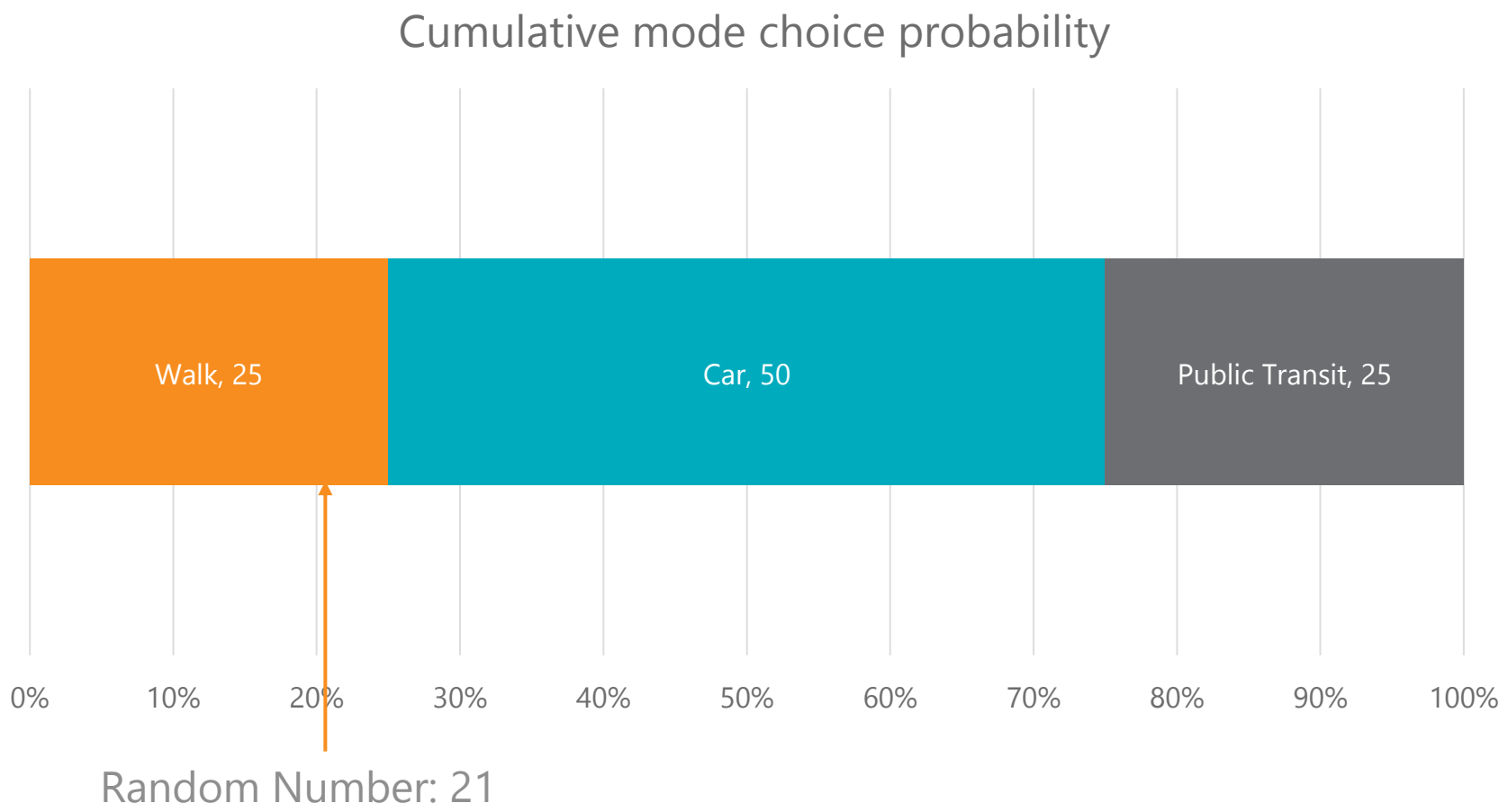
Simulation noise in micro-simulated models

- Transport models are often based on random utility theory - multinomial logit, nested logit
- Models with small number of alternatives (e.g. trip-based): Can calculate all probabilities for all alternatives (e.g.: 100.23 car trips between O1 and D2 in AM for purpose white collar work)
- ActivitySim: too many alternatives, computationally not feasible (person A drives to work at 8am, lunch break out of office at 12.30pm, home at 5pm, goes shopping with dependant at 7pm)
- Micro-simulation necessary: Turning choice probabilities into discrete outcomes

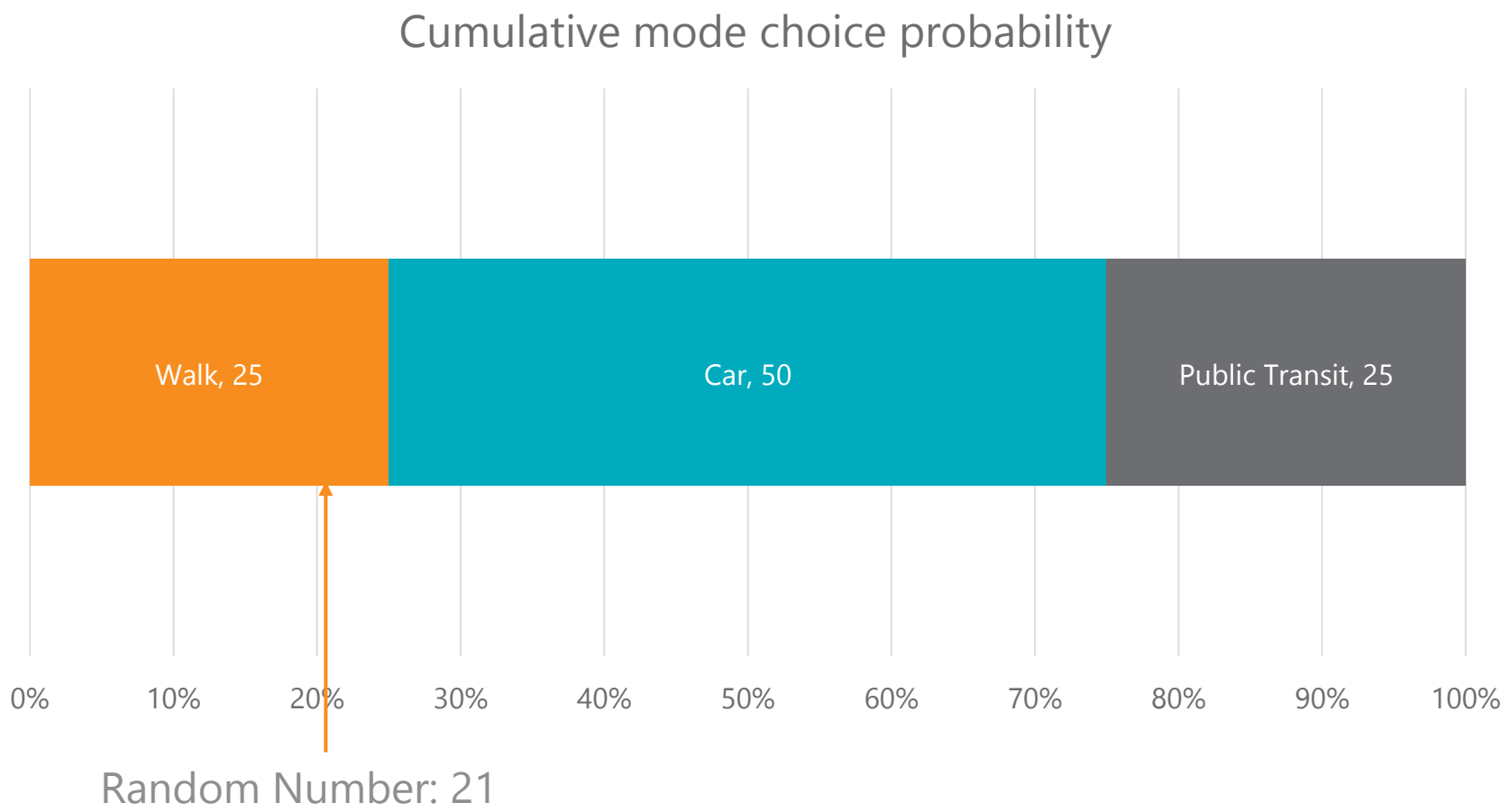
Making choices and treatment of randomness in ActivitySim - example



Making choices and treatment of randomness in ActivitySim - example

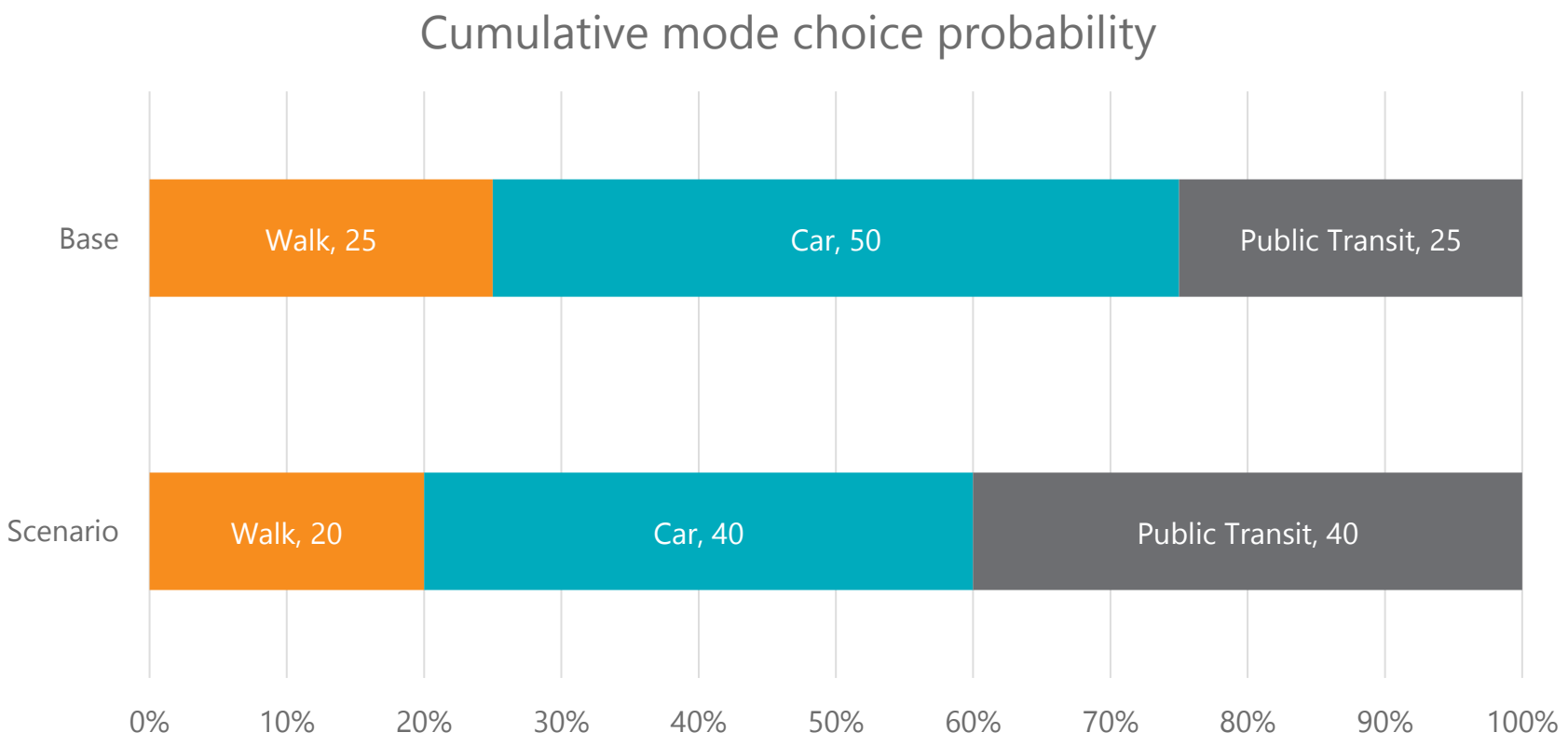


Making choices and treatment of randomness in ActivitySim - example

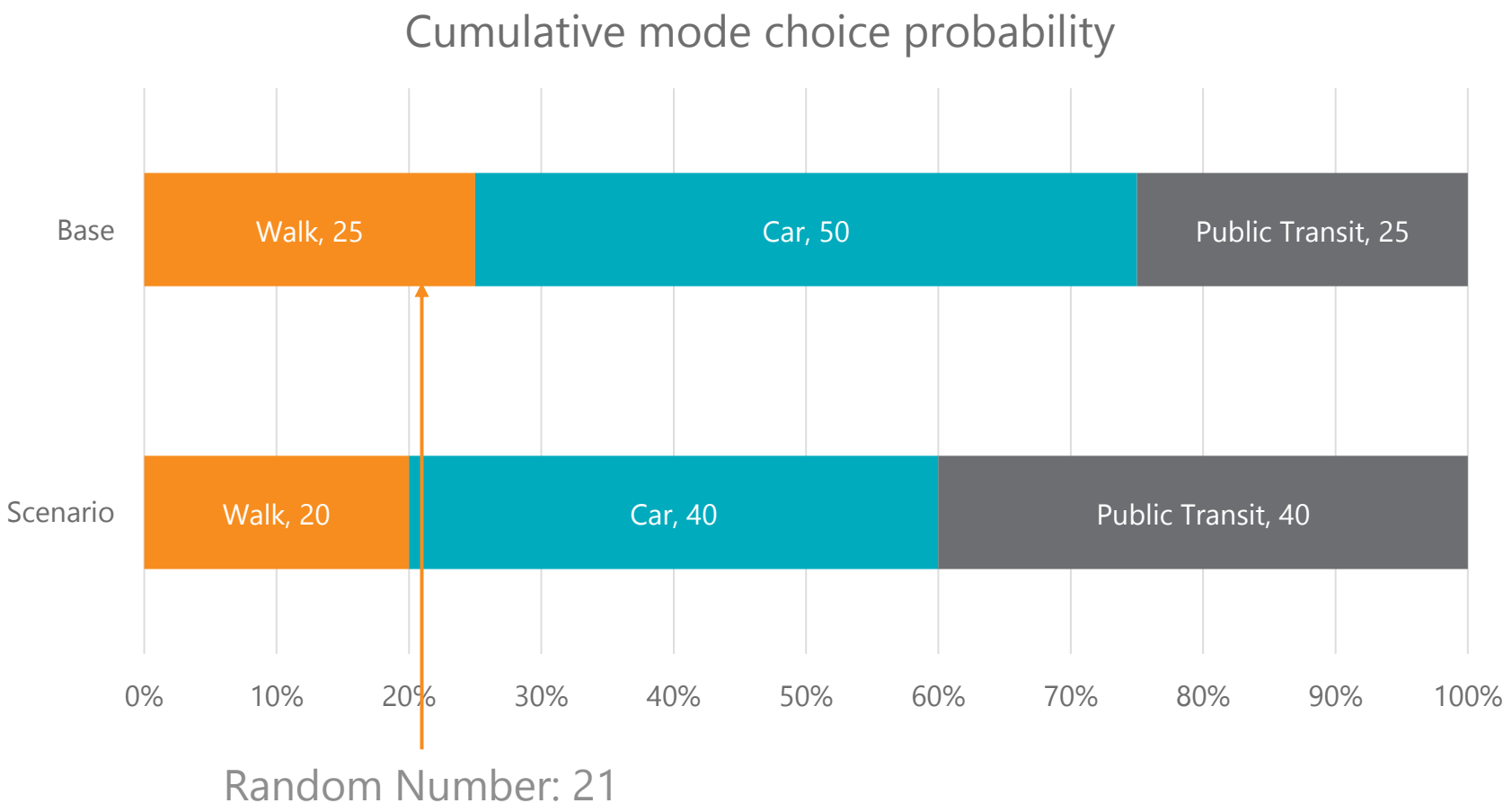


Fix seed for reproducibility and consistency across scenarios

Example: Choices across scenarios in ActivitySim – PT upgrade

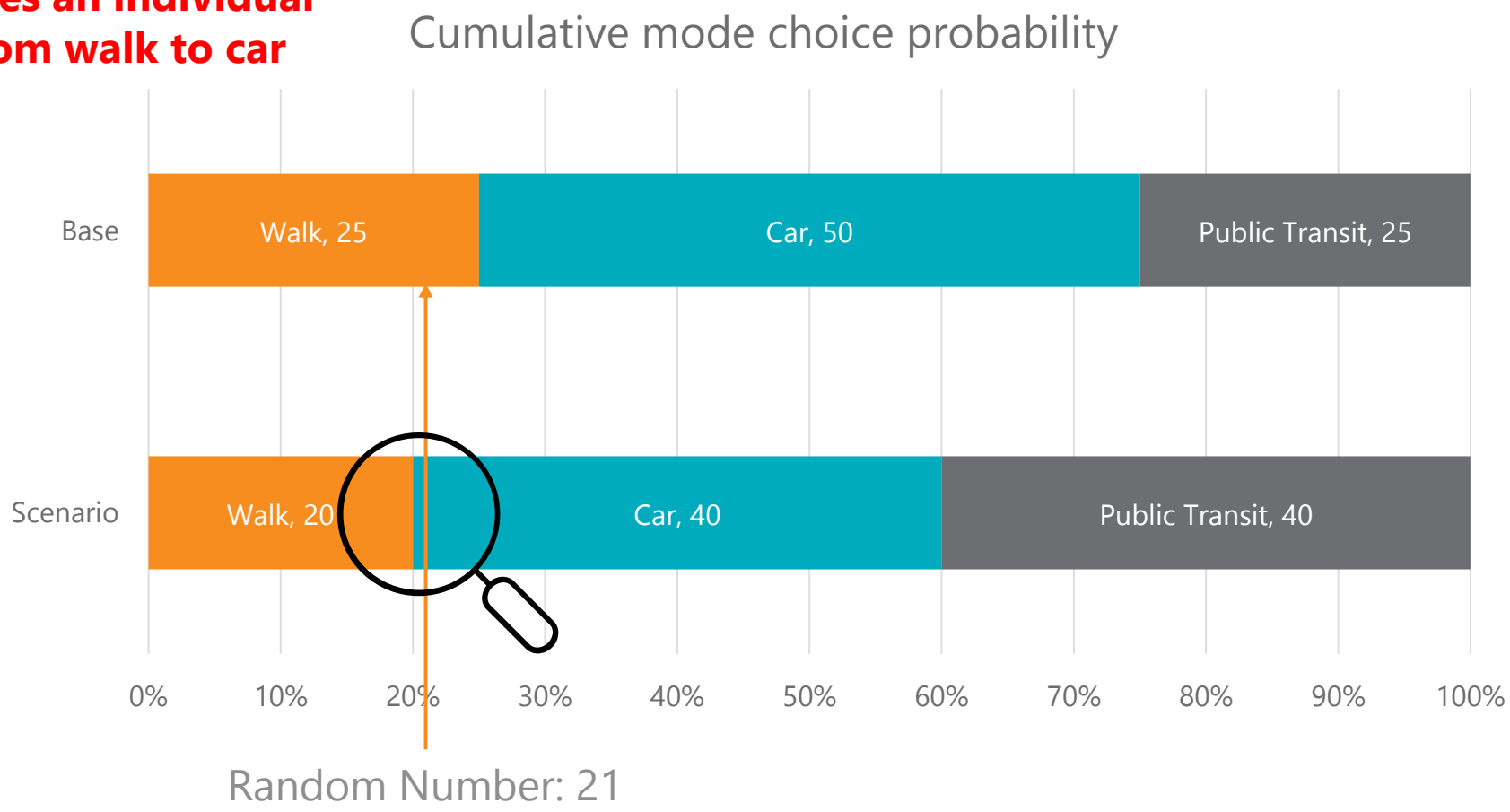


Example: Choices across scenarios in ActivitySim – PT upgrade



Example: Choices across scenarios in ActivitySim – PT upgrade

An upgrade to the PT network causes an individual to switch from walk to car



Let's take a step back – where do probabilities come from?

Random Utility Theory

Each chooser (i) chooses alternative (j) that maximizes their utility

The utility of each alternative (U_{ij}) is composed of a deterministic part (V_{ij}), which is a function of variables the modeler can observe, and an "error" term which is a random variable representing unobserved utility (ε_{ij}):

$$U_{ij} = V_{ij} + \varepsilon_{ij}$$

Choice probabilities are obtained by integrating out random variable

Let's take a step back – where do probabilities come from?

Random Utility Theory

Each chooser (i) chooses alternative (j) that maximizes their utility

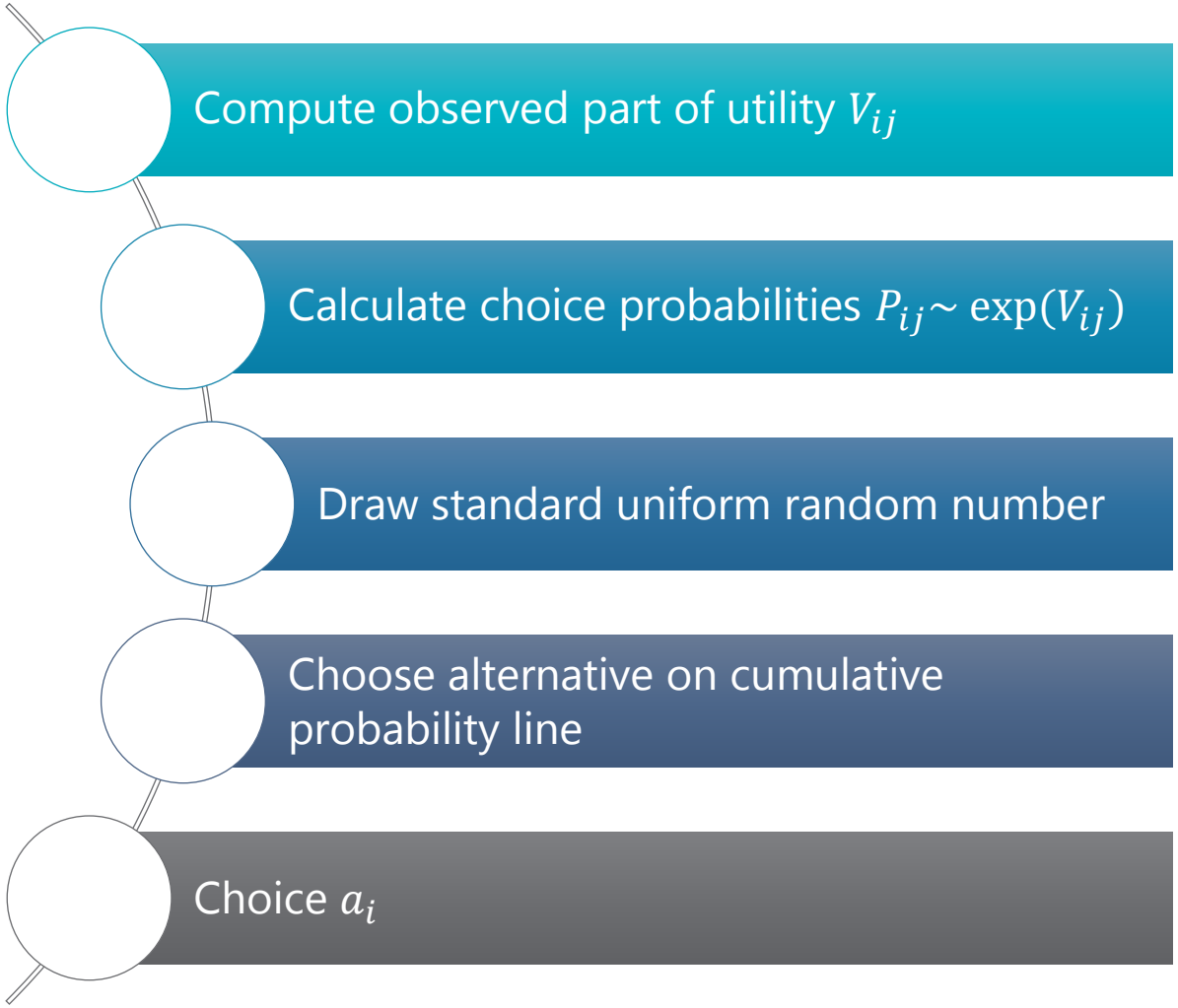
The utility of each alternative (U_{ij}) is composed of a deterministic part (V_{ij}), which is a function of variables the modeler can observe, and an "error" term which is a random variable representing unobserved utility (ε_{ij}):

$$U_{ij} = V_{ij} + \varepsilon_{ij}$$

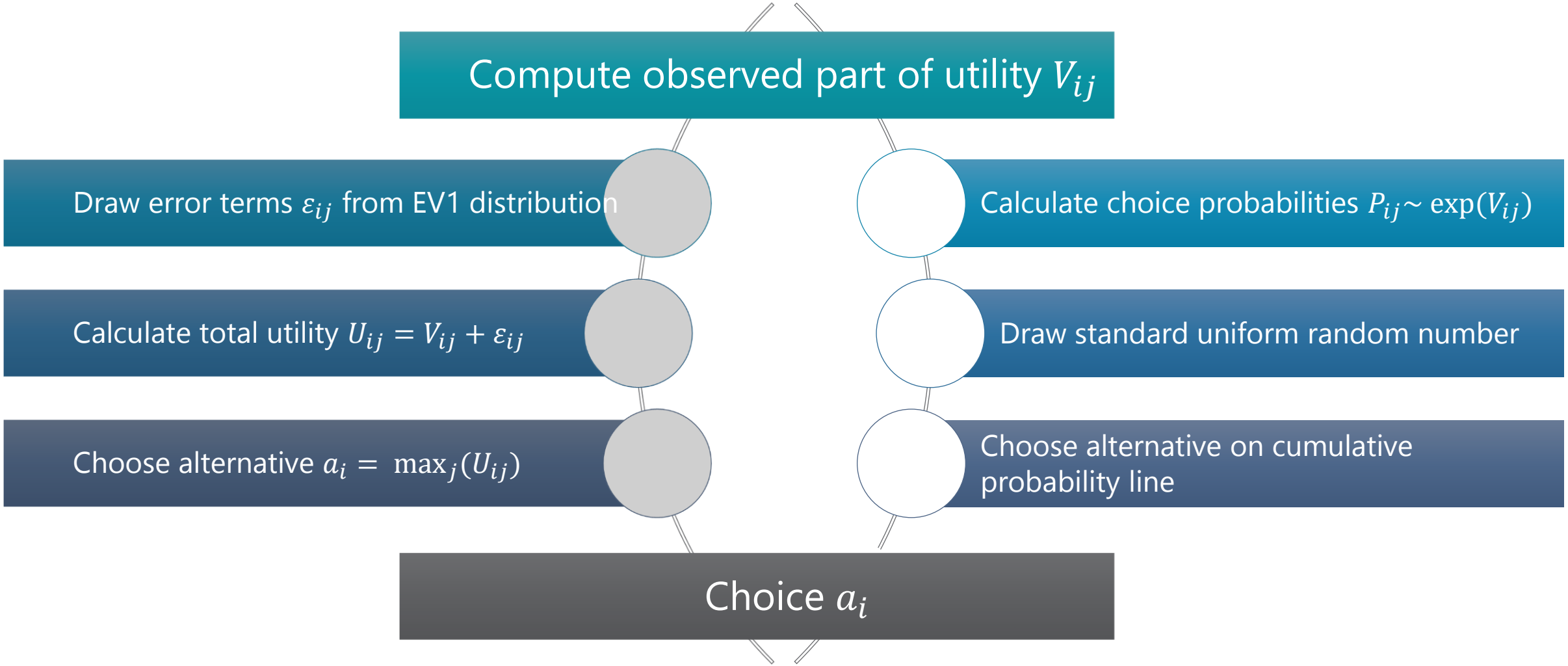
Draw directly

Choice probabilities are obtained by integrating out random variable

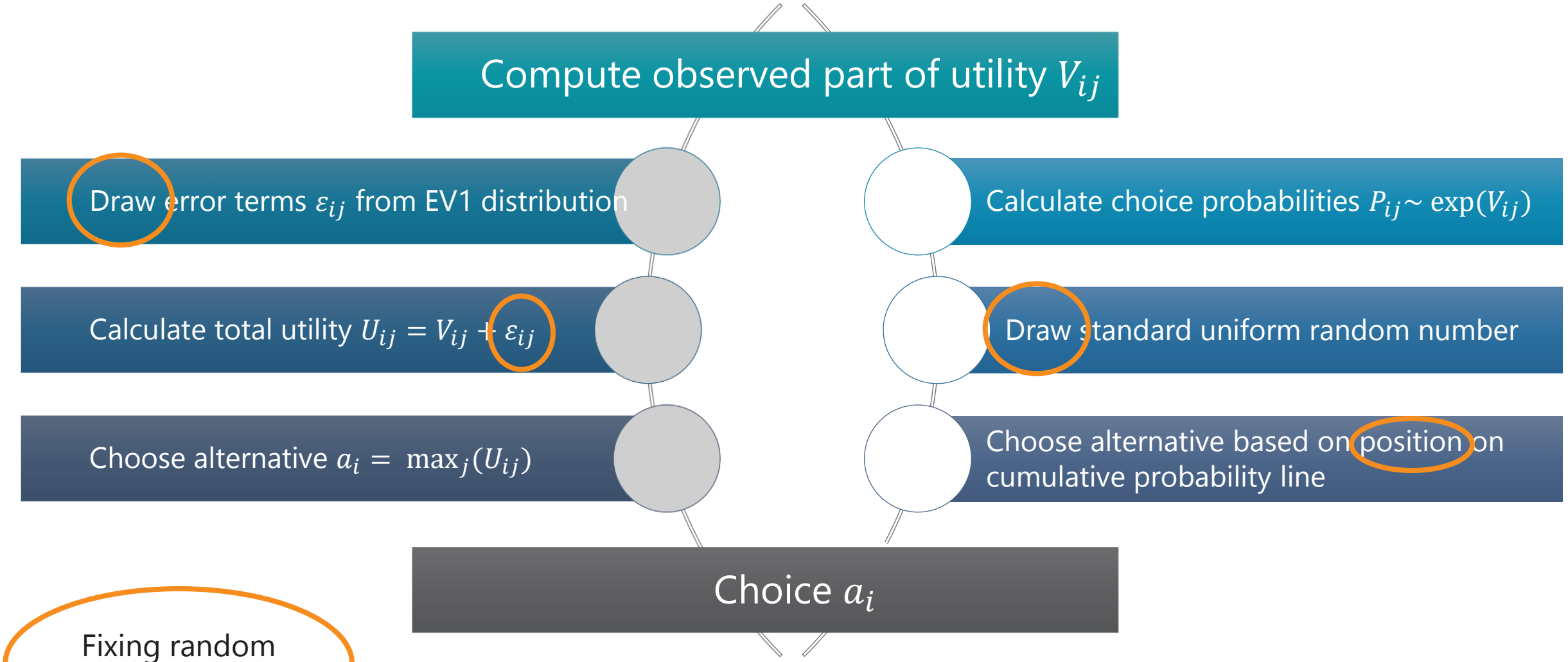
Making discrete choices – current method



Making discrete choices – two possible ways

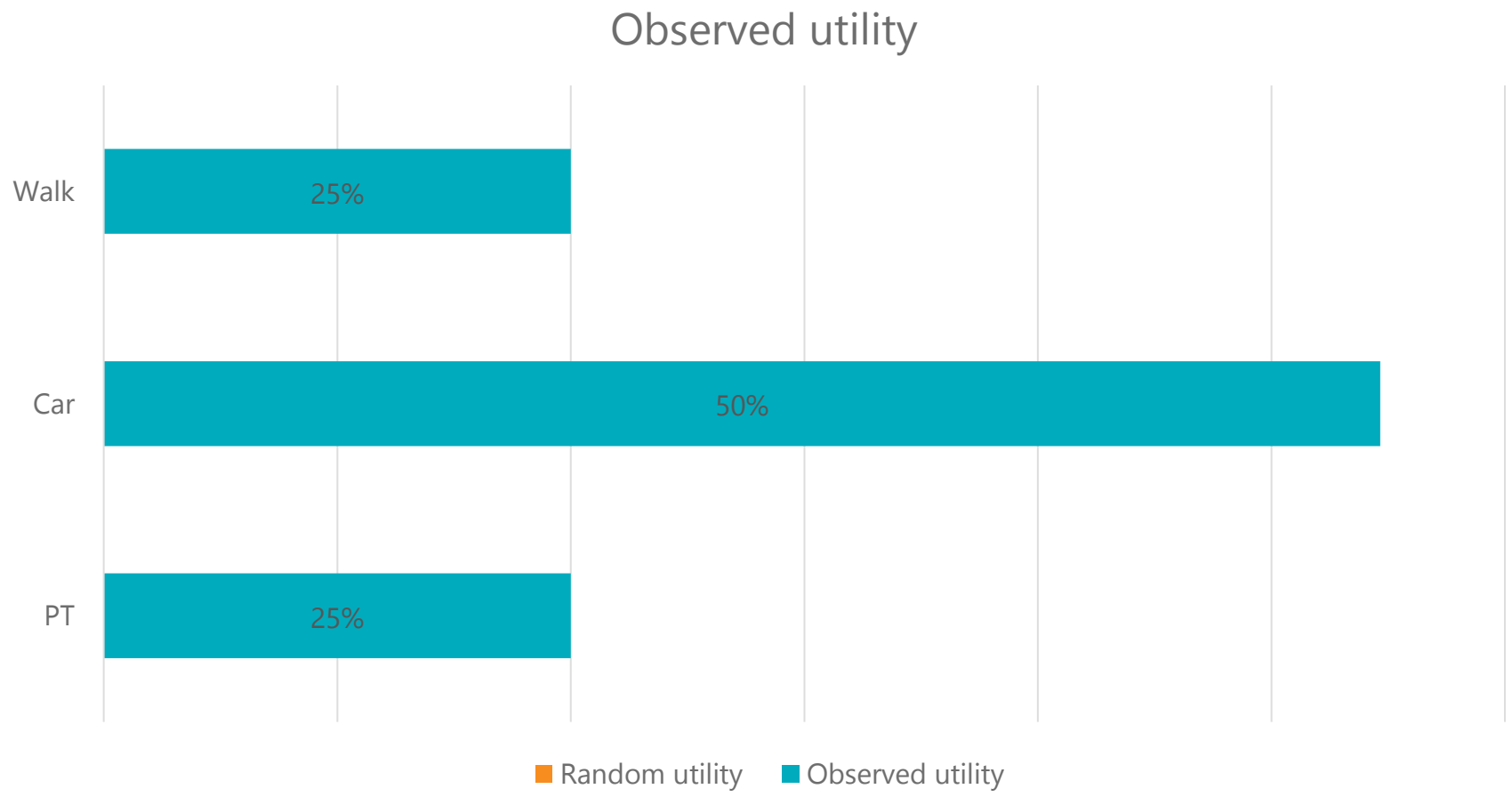


Making discrete choices – two possible ways

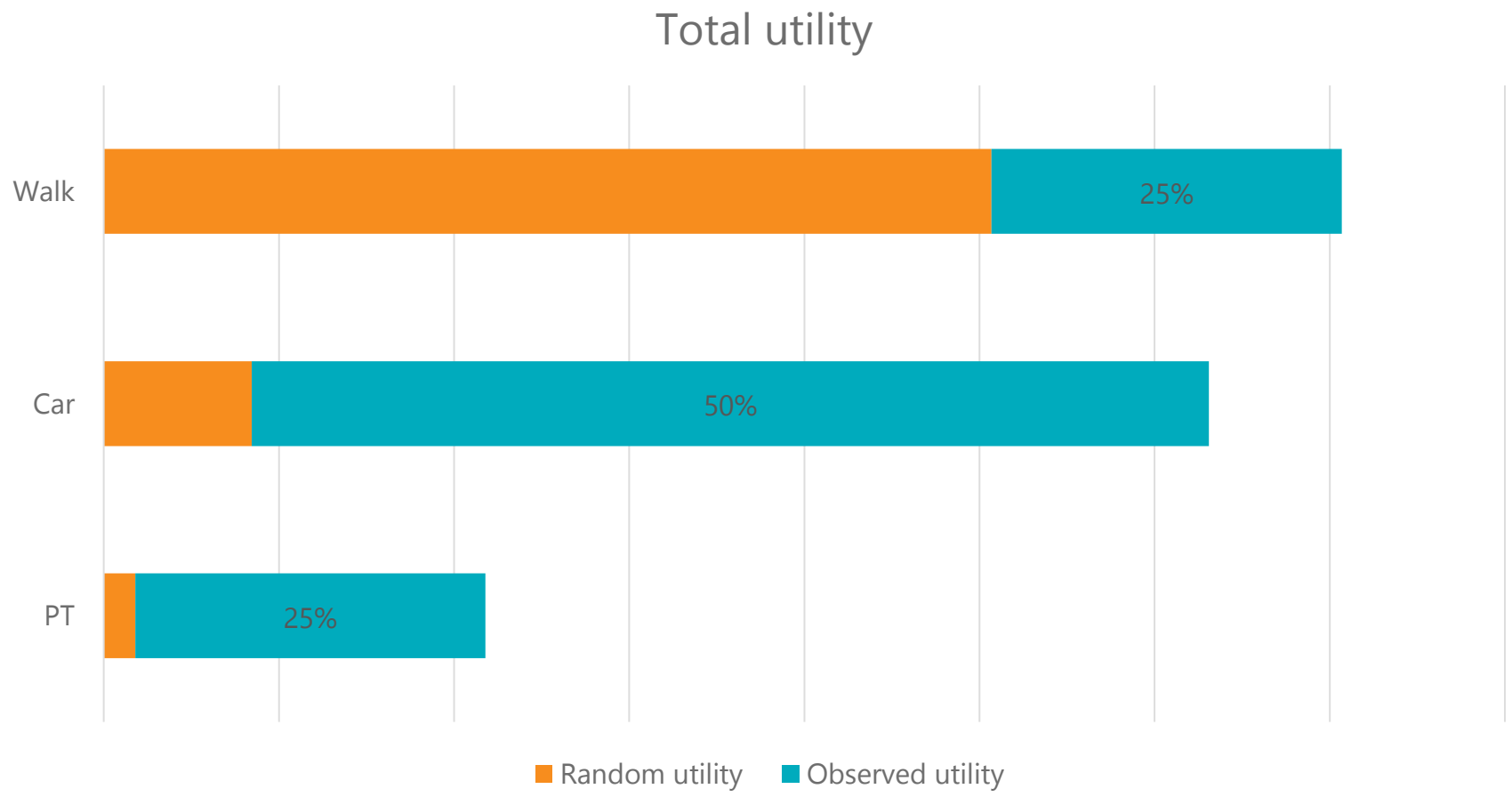


Fixing random seed

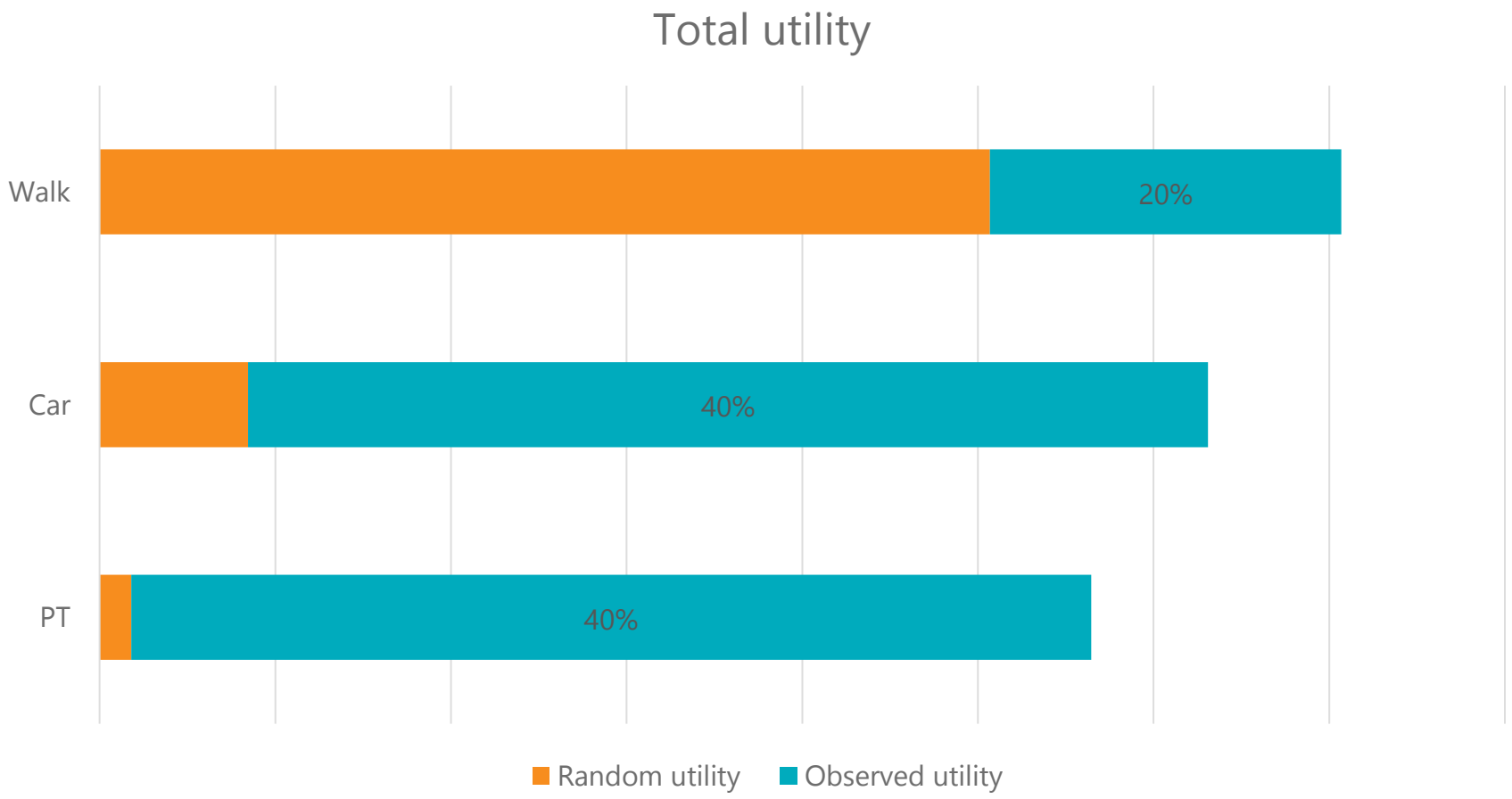
Mode choice example with random utility fixed



Mode choice example with random utility fixed



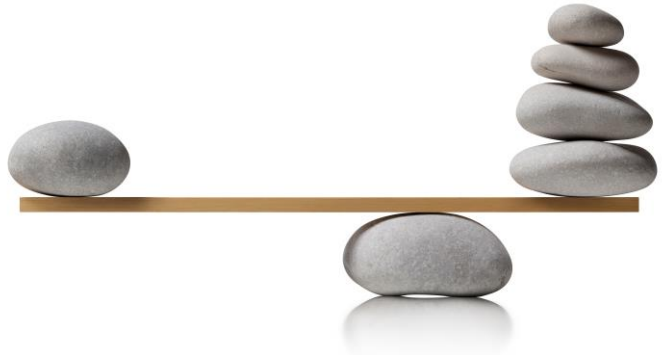
PT improvement scenario



But does it matter in practice?

Scenario comparisons

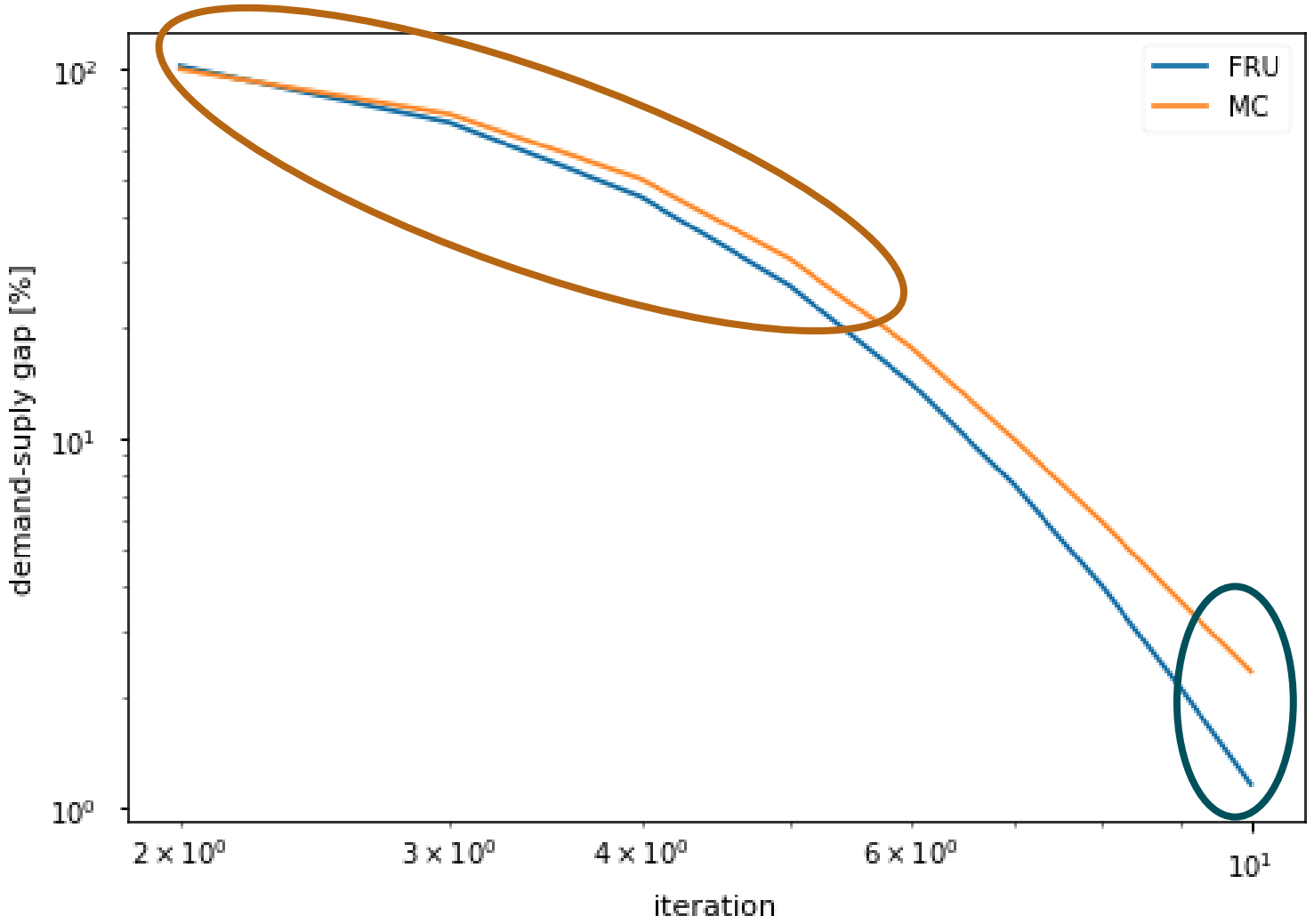
- Mode choice example from before
- Make model comparisons more meaningful – who switches due to deterministic utility change?



Overall model equilibration

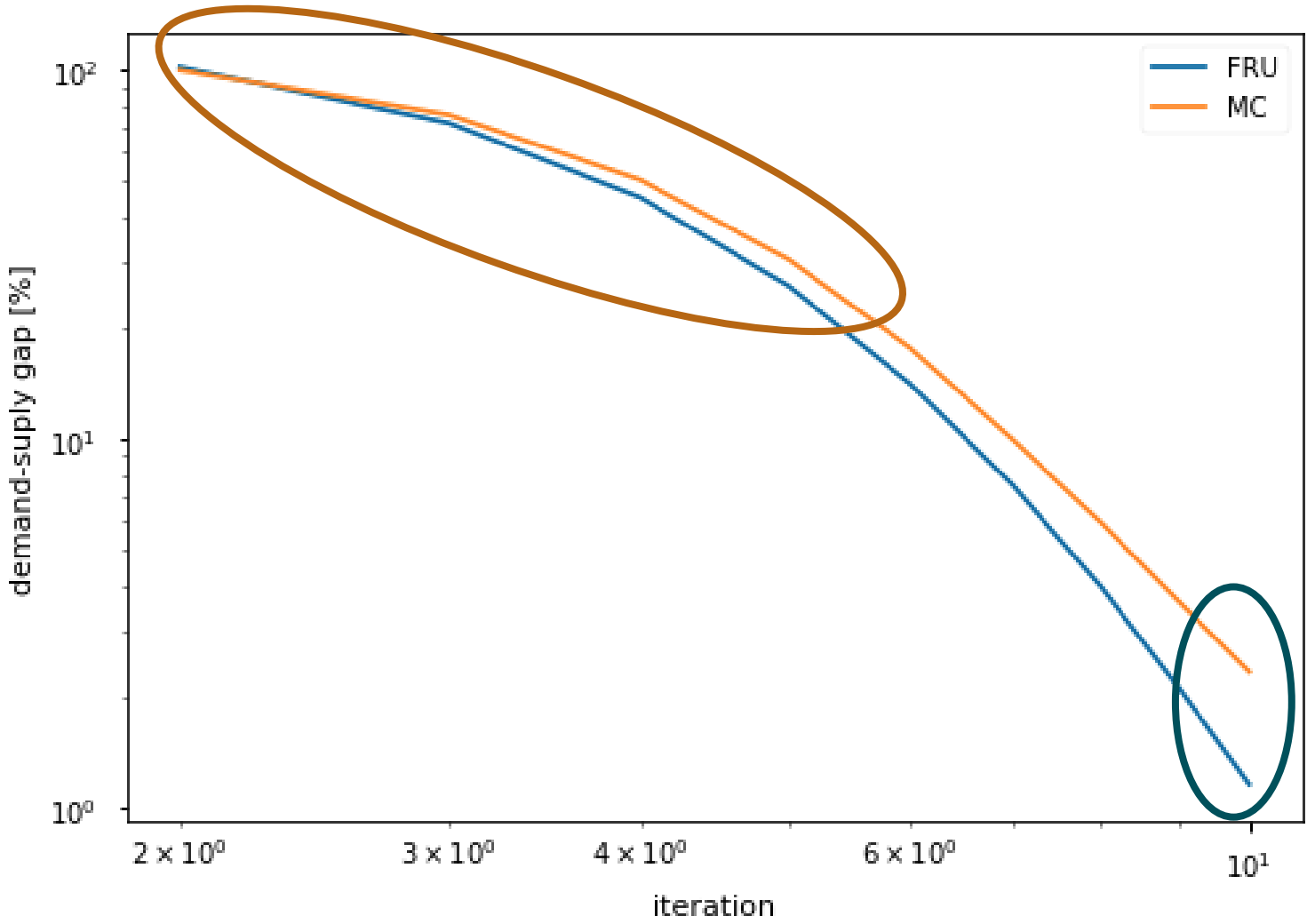
- Feedback between demand and assignment
- Assignment requires numerical solution -> noise feedback

Equilibration



Note: this plot is on a log-log scale.

Equilibration



Convergence limit in practice?

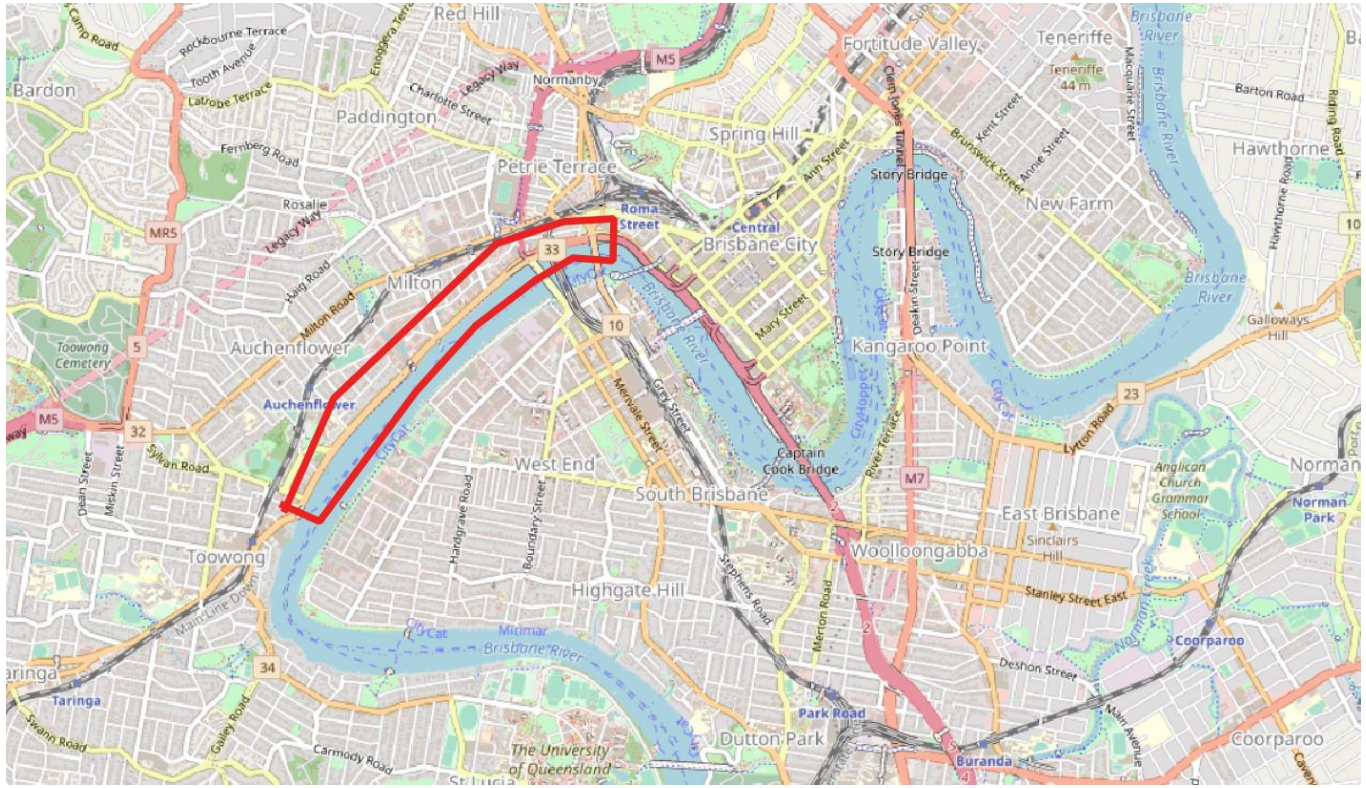
Note: this plot is on a log-log scale.

Scenario comparison

Increase capacity along Coronation Drive by two lanes – adding capacity for about 2000 vehicles/hour in both directions

Coronation Drive

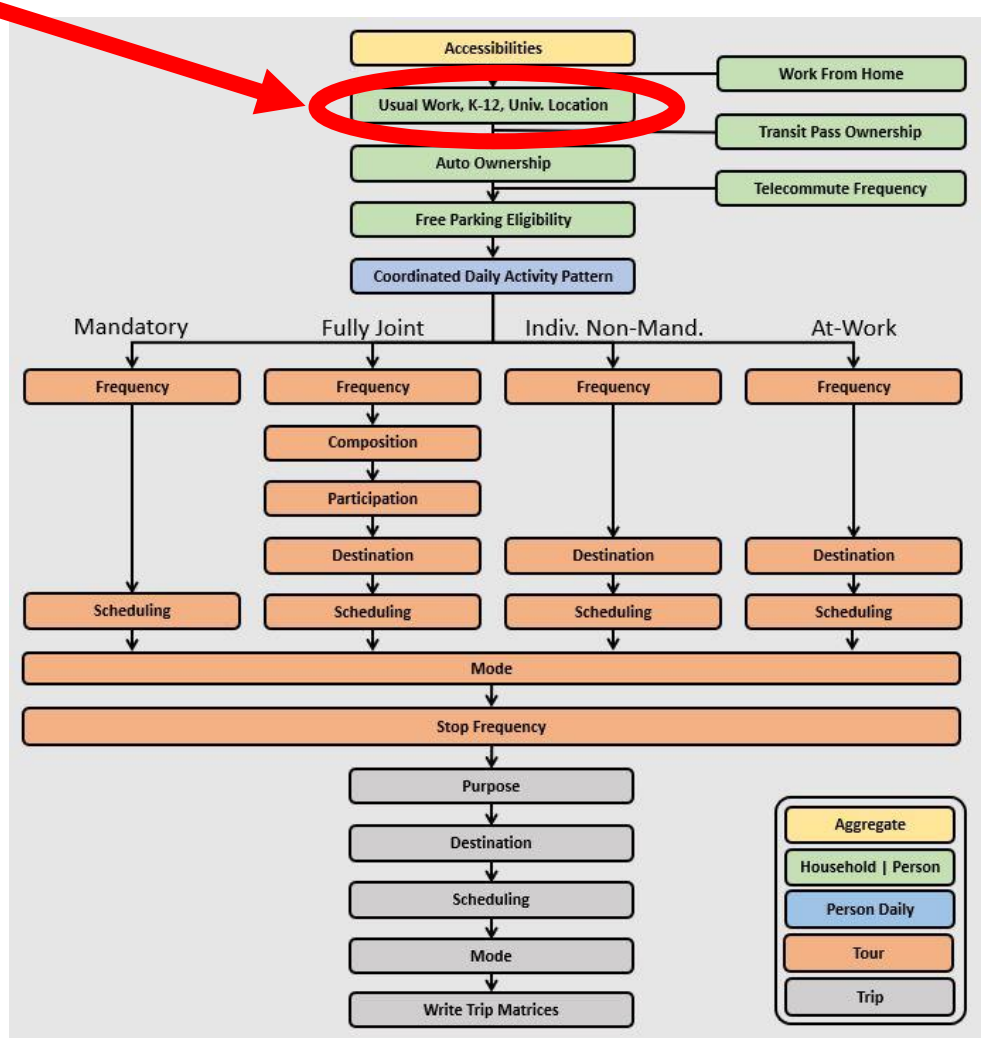
- Heavily congested in both directions during peak times
- Plays several roles: feeds the CBD, UQ (largest university campus in Brisbane) to its south, and forms part of a CBD by-pass route. Additionally, due to a lack of bridges in the south, it connects the west of Brisbane to the south
- Serves many high-frequency buses
- Rail alternative exists primarily for CBD access, but not high frequency



Work location choice difference

How many workers choose different work location zone in scenario compared to base (note: home location is fixed)

Choice method	Work location TAZ change
Current ActivitySim	10,971
Frozen random utility	1821



Work location choice difference

Choice method	Sum of absolute differences (SA4)	Sum of absolute differences (SA3)	Sum of absolute differences (SA2)	Sum of absolute differences (TAZ)
Current Asim	720	1925	5842	10,971
Frozen random utility	700	1263	1682	1821
Ratio	1.03	1.52	3.47	6.02

Australian Statistical Areas:

- SA2 has 3k to 25k people (10k on average)
- SA3 has 30k to 130k people
- SA4 has 100k to 300k people

Current status

- Implemented on fork of ActivitySim at https://github.com/janzill/activitysim/tree/janzill/utility_based_choices_with_sharrow
- Simple switch turns functionality on/off

HOWEVER

- Longer runtimes (per demand model run)
- Nested logit models currently have some edge cases

Current status

- Implemented on fork of ActivitySim at https://github.com/janzill/activitysim/tree/janzill/utility_based_choices_with_sharrow
- Simple switch turns functionality on/off

HOWEVER

- Longer runtimes (per demand model run)
    - Some profiling leads us to believe a lot of this is due to implementation of random generator in ActivitySim – re-seeding for each draw (chooser and alternative) instead of keeping generators in memory. Can trade off memory for runtime. Not investigated yet.
    - Overall runtime depends on number of demand-supply iterations, and these reduce as we saw
- Nested logit models currently have some edge cases

Conclusions

- The Frozen Random Utility approach leads to demonstrably more stable results
- This should improve the interpretability of model results
- It might prove necessary for better convergence (economic analysis)
- In application, it increases runtime, but this may be offset by the need to run fewer model cycles
- Papers: ATRF 2022 (https://australasiantransportresearchforum.org.au/frozen-randomness-at-the-individual-utility-level) for technical derivation and ETC 2022 (https://aetransport.org/past-etc-papers/conference-papers-2022?abstractId=7583&state=b) for results presented here

Thank you!

Scenario comparison for different seeds

- Second iteration work location choice with five different seeds each for FRU/MC and base and increased scenario.
- As can be seen, results are stable at all aggregation levels.

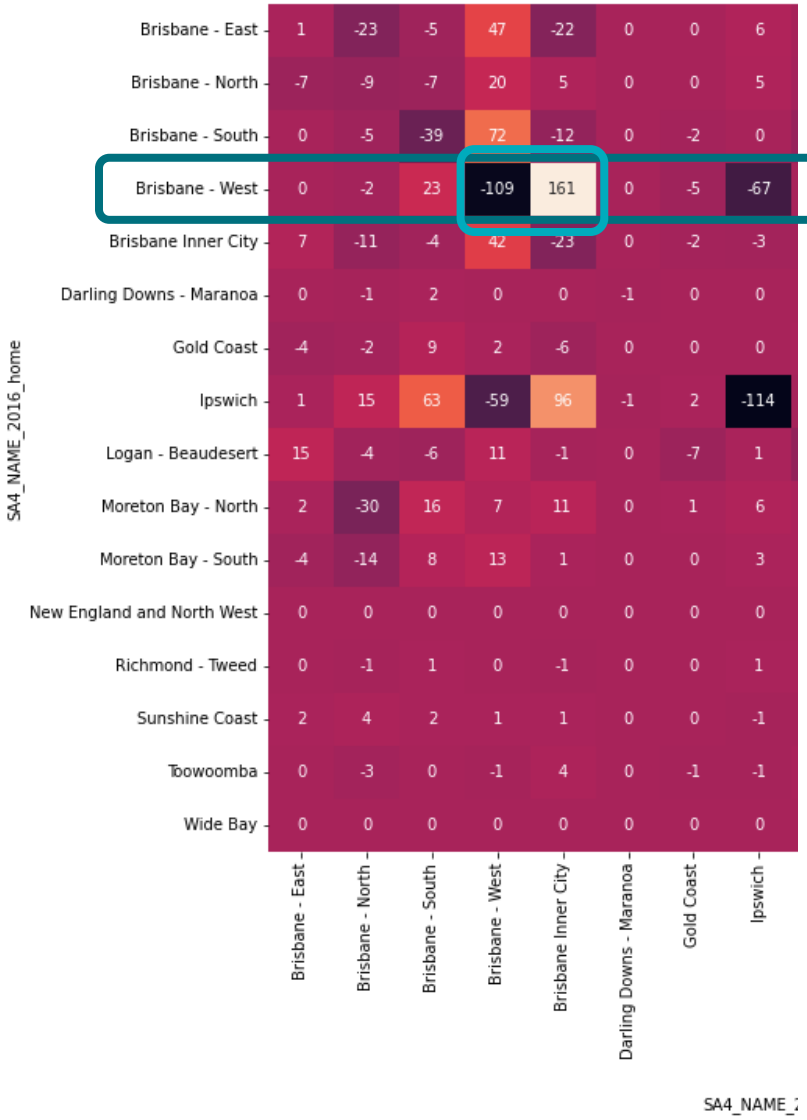
	method	seed	sa4	sa3	sa2	zonal_agg	zonal_individual
0	fru	0	721	1185	1586	1733	1734
1	fru	379	687	1198	1561	1707	1707
2	fru	9999	767	1193	1578	1729	1732
3	fru	123456	752	1199	1592	1712	1714
4	fru	67549036	776	1227	1589	1744	1746
5	mc	0	801	2037	6568	<NA>	12657
6	mc	379	833	2072	6641	12383	12781
7	mc	9999	804	1996	6479	12202	12552
8	mc	123456	820	1995	6637	<NA>	12872
9	mc	67549036	771	2062	6668	<NA>	12944

Nested logit model technical details

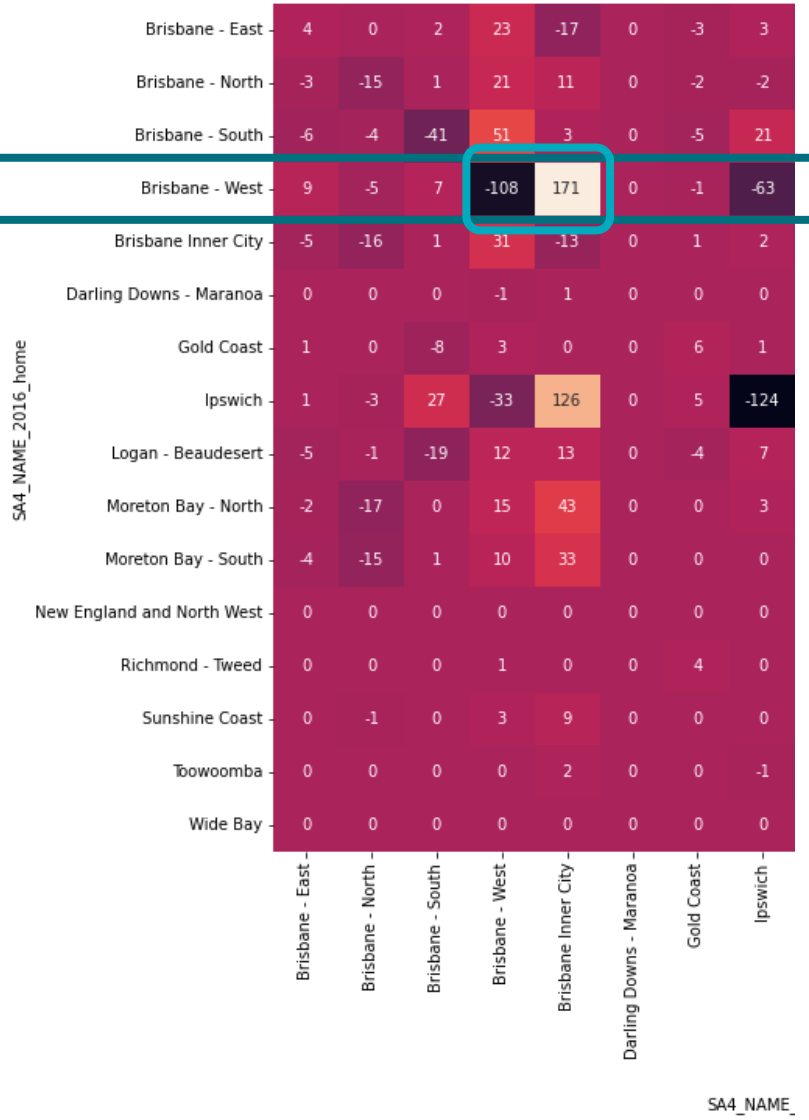
- Can write as product of conditional and marginal probabilities such that random term at each level is EV1
- Uses logsums (expected maximum utility); there can be edge cases where response to scenario is not consistent with RUM at individual level (nest switching)
- Recently, A. Galichon showed how to avoid this; involves distribution which generally does not have closed-form representation – however, its Laplace transform does (https://doi.org/10.1017/S026646662000047X)
- It is known how to do draw from this distribution numerically, however we have not implemented this (yet)

SA4 to SA4 changes

Naïve inversive method



Frozen random utility



Each table shows the difference in work location outputs (home SA4 x work SA4) between the base case and project case.

At this level, the results are quite similar.

BRISBANE
Level 5, 200 Mary Street,
Brisbane, QLD 4000
T: +61 7 3870 4888

SYDNEY
Level 6, 46 Kippax Steet,
Surry Hills, NSW 2010
T: +61 2 9051 2423

LONDON
49 Greek Street,
London, W1D 4EG
T: +44 7813 320553

MELBOURNE
Level 10, 565 Bourke Street,
Melbourne, VIC 3000
T: +61 3 9602 5200

PERTH
Level 12, 197 St Georges
Terrace, Perth, WA 6000
T: +61 8 6388 2830

PLANWISELY

SITEWISELY
Choose your site wisely

veitchlister.com

vlc