Modeling Demand for Ridesourcing as a Feeder for High Capacity Mass Transit Systems: A Case Study of the Planned Beirut BRT

Najib Zgheib, Maya Abou Zeid, Isam Kaysi

Workshop on Discrete Choice Models
Lausanne, 2019

Outline

- ✓ Introduction
- ✓ Research Objectives
- ✓ Experimental Design
  - ✓ Stated Preference Design
  - ✓ Data Collection
- ✓ Demand Modeling
- ✓ Future Tasks

Introduction

- ✓ **Ridesourcing is emerging as a main mode in the travel industry.**
- ✓ **Contradicting claims about its relation to transit**

CONS

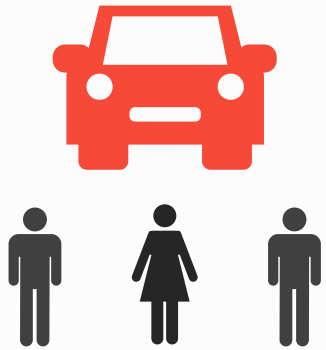
- ❑ Cut the shares of traditional modes.
- ❑ Complaints from taxi drivers due to unfair competition (insurance, driver license, depreciation, …).
- ❑ Claims about safety concerns and inadequate background checks for drivers.

PROS

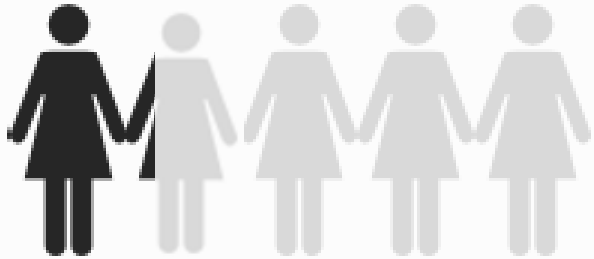
- ❑ Complementing transit by enlarging its catchment area.
- ❑ Replacing low usage transit lines.
- ❑ The city of Centennial, Colorado teamed up with Lyft to provide free rides to and from light rail stations (Shen et al., 2017).

Congestion in Lebanon

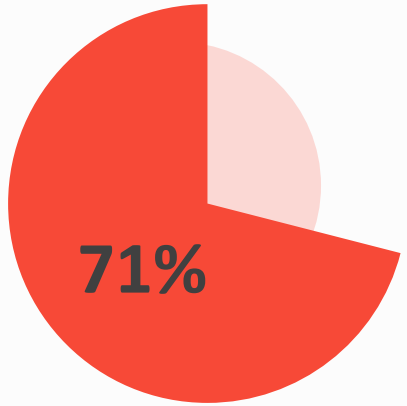
Lebanon suffers from severe congestion on the entrances to Beirut for diverse reasons



- High reliance on private cars.
- Car ownership of around 1 car per 3 persons (MoE/UNDF/GEF, 2015)



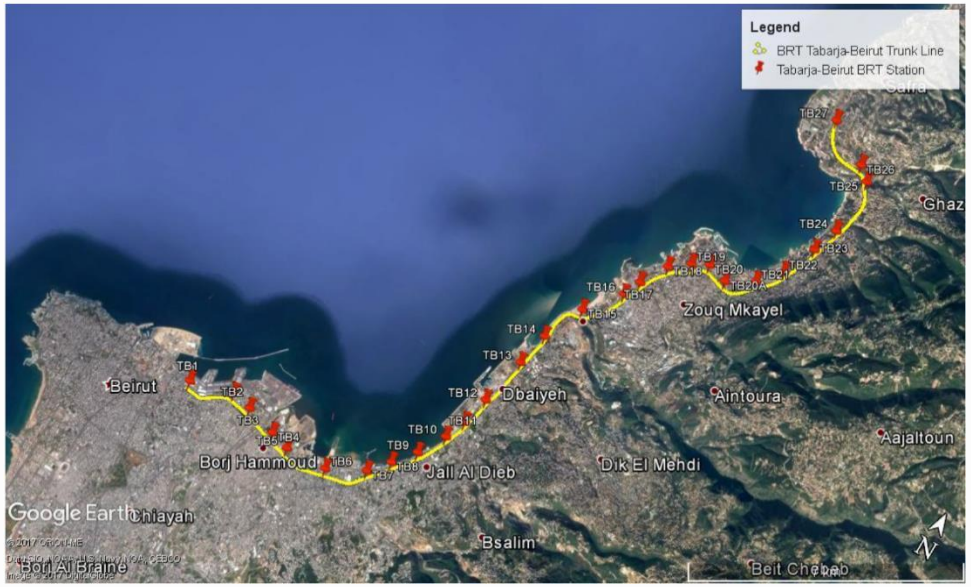
- Vehicle occupancy is low at 1.2 passengers per car on average (MoE/UNDP, 2015)

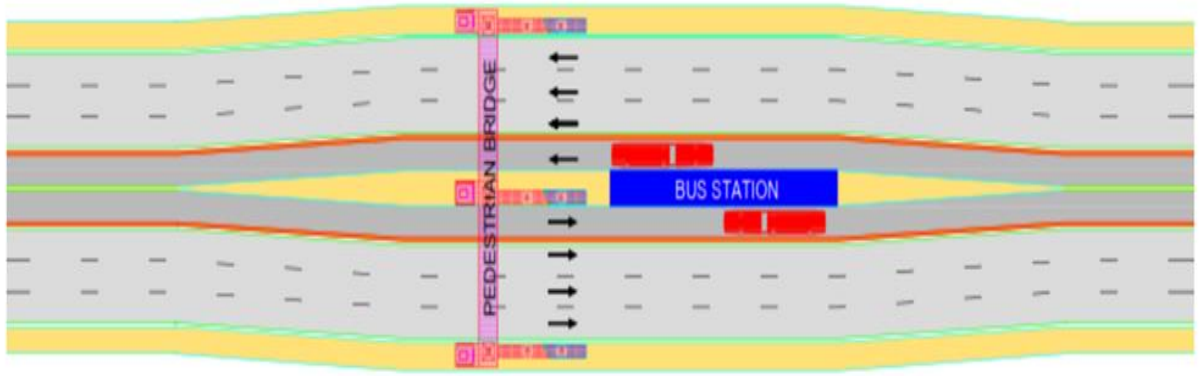


- 71% is the share of private cars from all trips in Greater Beirut.
- Jitneys/taxis serve 19% and buses/vans serve 10% (Kaysi et al., 2010)

Proposed Solution

- The World Bank recently proposed a $300 M Bus Rapid Transit (BRT) system from Tabarja (North) to Charles Helou.
- The proposed BRT has the following characteristics:
  - ✓One dedicated lane per direction at the middle of the coastal highway.
  - ✓Stations located along the highway at intervals of around 1 km.
  - ✓Tickets sold online and at stations (for one trip, 5 trips, 10 trips, daily pass,…).
  - ✓BRT buses follow exact schedules, with short headways (2 - 3 minutes).
- A successful integration of the BRT largely depends on accessibility and first-mile-last-mile connections.



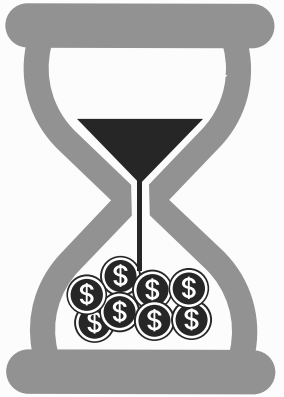


Outline

- ✓ Introduction
- ✓ Research Objectives
- ✓ Experimental Design
  - ✓ Stated Preference Design
  - ✓ Data Collection
- ✓ Demand Modeling
- ✓ Future Tasks

Research Objectives

- ✓ Develop a framework for the assessment of demand for ridesourcing as a feeder for high capacity transit services.
- ✓ Develop a case study based on the planned Beirut BRT with the following aims:
  - Determine the share that will shift from private cars to the BRT.
  - Estimate the modal split among feeder modes.
- ✓ Test the impact of several policies on overall demand such as:
  - Increasing parking prices.
  - Subsidizing ridesourcing fares for trips that start/end at a BRT station.

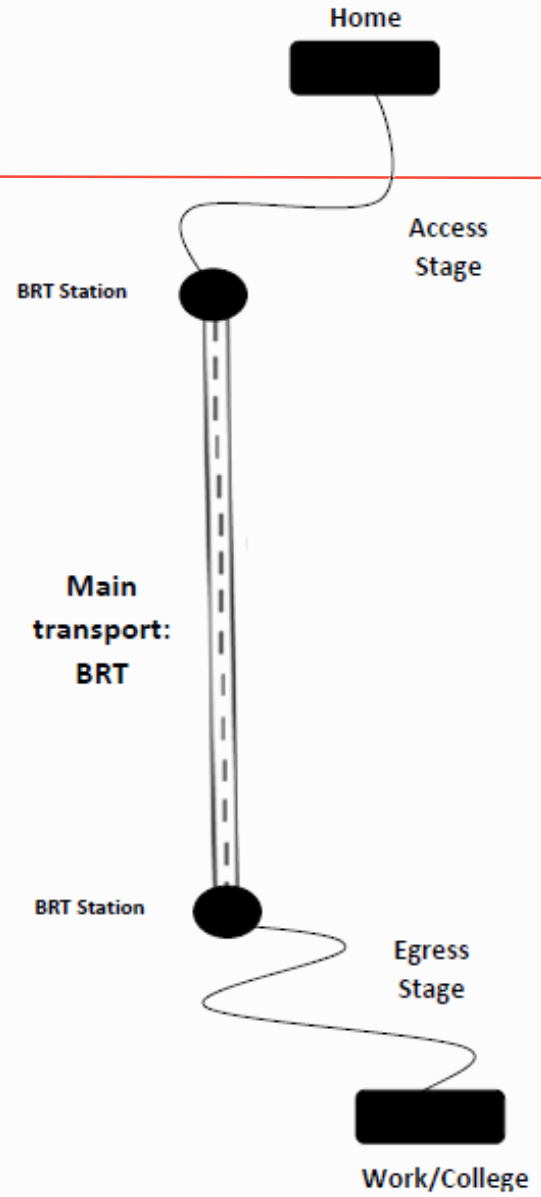


Outline

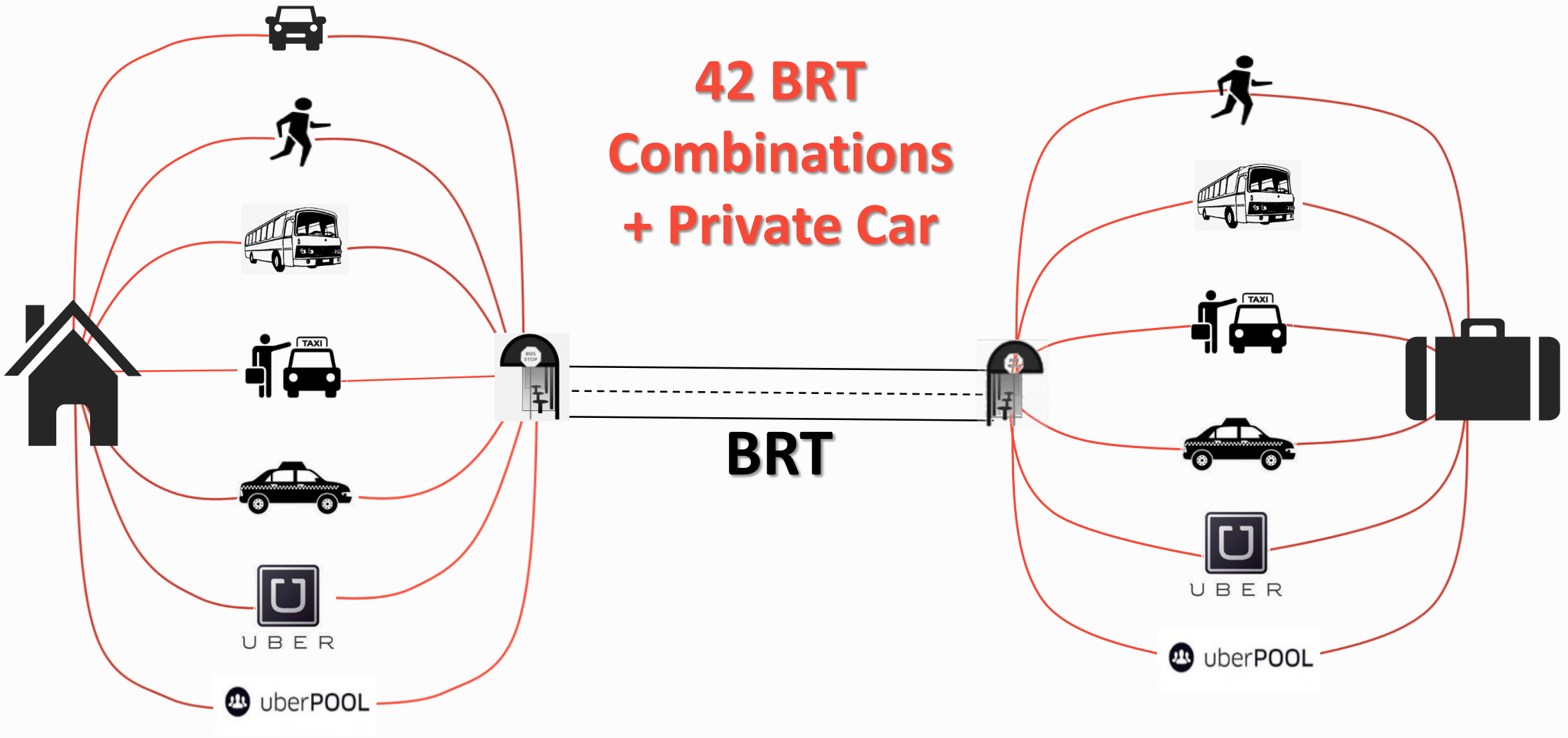
- ✓ Introduction
- ✓ Research Objectives
- ✓ Experimental Design
  - ✓ Stated Preference Design
  - ✓ Data Collection
- ✓ Demand Modeling
- ✓ Future Tasks

BRT Trip Representation

- A typical BRT trip will be divided into three main stages:
  - Access stage: travel from home to the BRT station.
  - Main transport: travel using the BRT from one station to another.
  - Egress stage: travel from the BRT station to the desired destination.
- Access and egress modes are selected independently from each other.

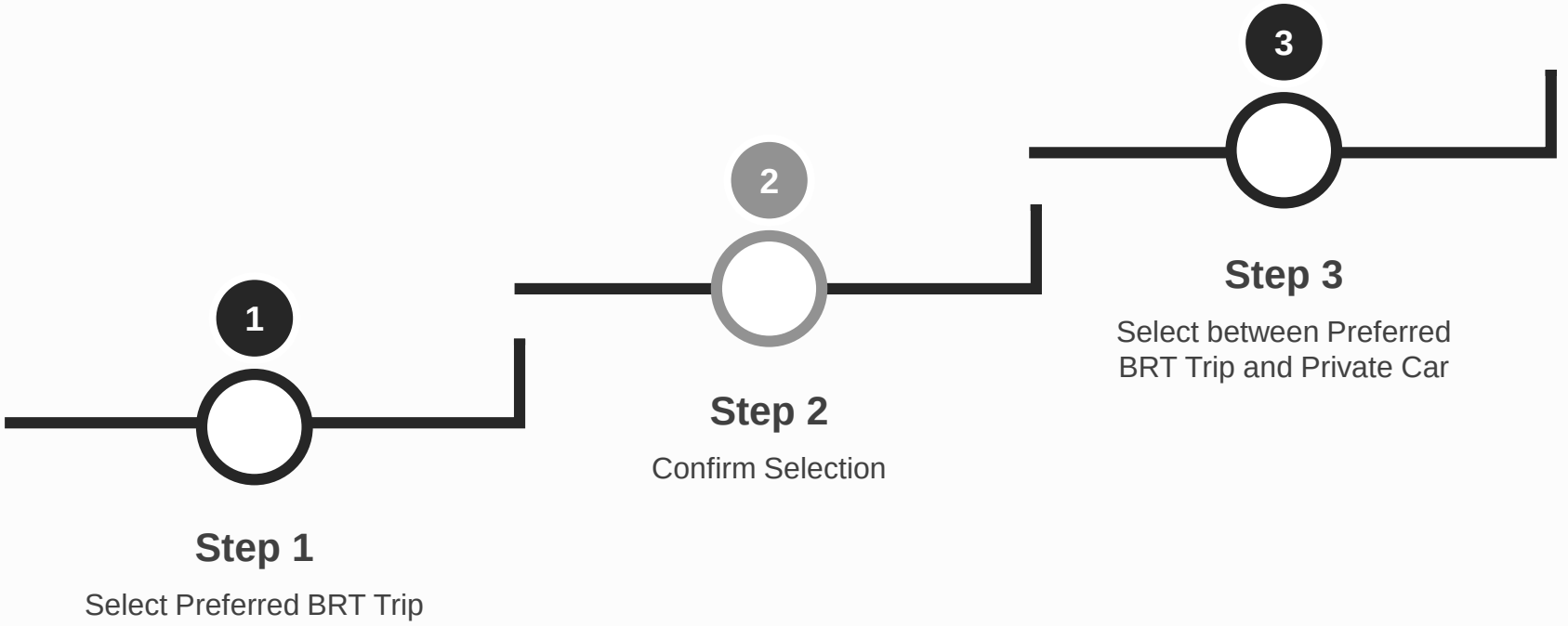


BRT Trip Combinations



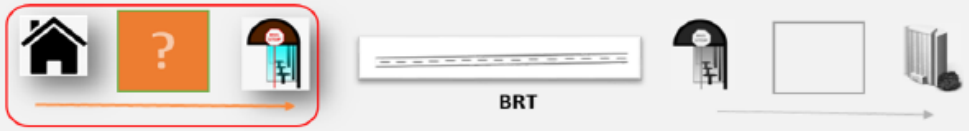
Stated Preference Design

- A sequential choice experiment is developed to simplify the choice process.
- Selection is staged over 3 different steps in each scenario.



Step 1: Preferred BRT Trip

ACCESS MODE



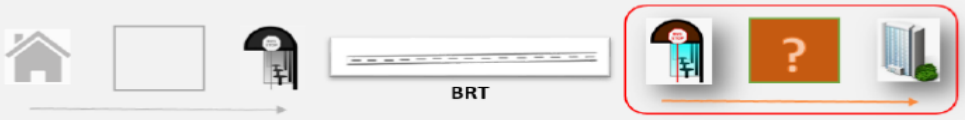
Available Access Modes

		Park & Ride	Walk	Bus	Taxi	Service	Ridesourcing (private)	Ridesourcing (shared)
	In-Vehicle Travel Time (min)	3	-	4	3	4	3	4
	Walking Time (min)	2	12	2	-	1	-	-
	Waiting Time (min)	-	-	6	5	4	3	4
	Fuel Cost (LL)	1000	-	-	-	-	-	-
	Daily Parking Cost (LL)	2000	-	-	-	-	-	-
	Fare (LL)	-	-	1000	5000	2000	3500	2500
SELECTION								
							X	

MAIN TRANSPORT: BRT

	In-Vehicle Travel Time (min)	21
	Waiting Time (min)	4
	Fare (LL)	3000

EGRESS MODE



Available Egress Modes

		Walk	Bus	Taxi	Service	Ridesourcing (private)	Ridesourcing (shared)
	In-Vehicle Travel Time (min)	-	9	6	7	5	6
	Walking Time (min)	12	2	-	2	-	-
	Waiting Time (min)	-	2	4	2	4	3
	Fare (LL)	-	2000	6000	2000	3000	2000
Selection							
		X					

Step 2: Choice Confirmation

You Selected: Ridesourcing (private) + BRT + Walking		
	In-Vehicle Travel Time (min)	24
	Walking Time (min)	12
	Waiting Time (min)	7
	Fuel Cost (LL)	0
	Daily Parking Cost (LL)	0
	Fare (LL)	6500
Confirm Your Selection		
Go Back to Step 1		

Step 3: Selection between Preferred BRT Trip and Private Car

Overall Trip			
		Ridesourcing (private) + BRT + Walking	Private Car
	In-Vehicle Travel Time (min)	24	45
	Walking Time (min)	12	5
	Waiting Time (min)	7	0
	Fuel Cost (LL)	0	4000
	Daily Parking Cost (LL)	0	5000
	Fare (LL)	6500	
Selection			

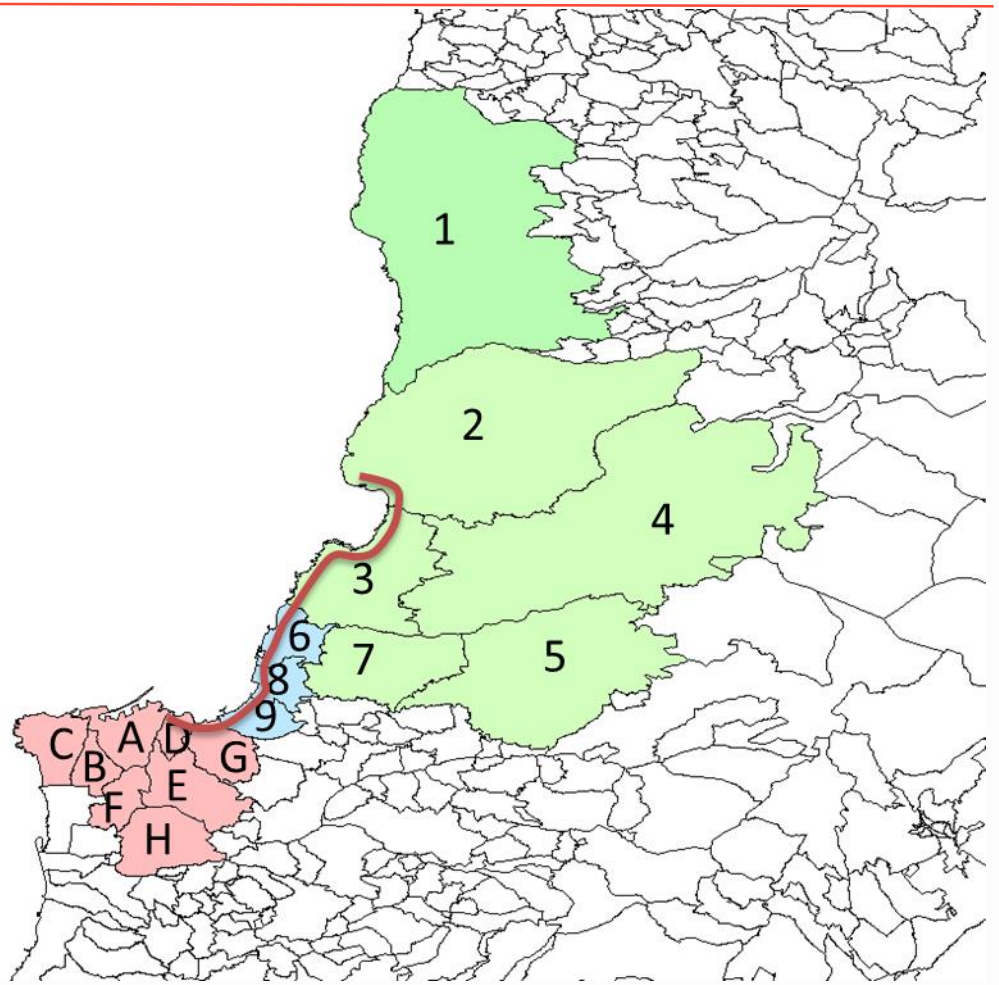
<u>Note:</u> Levels are generated randomly from a set of 4 values for each variable

Outline

- ✓ Introduction
- ✓ Research Objectives
- ✓ Experimental Design
  - ✓ Stated Preference Design
  - ✓ Data Collection
- ✓ Demand Modeling
- ✓ Future Tasks

Study Area

- The study area is divided into 9 origin zones (1 to 9) and destination zones (A to H).
- Zones 6, 8, and 9 can also serve as destinations for trips originating in zones 1, 2, or 3.
- Origin zones cover regions expected to generate most demand.
- As for destinations, the following zones are included:
  - Municipal Beirut
  - Major employment zones in Greater Beirut Area in proximity of BRT alignment



Sampling Frame

- Stratified Random Sampling based on the exogenous variable:

$$\frac{IVTT_{car}}{IVTT_{BRT}+2\times IVTT_{Acc}}$$

- The variable serves as measure of attractiveness of the BRT.
- Sample size of 400 (3 scenarios per individual).

Stratum	Interval for $\frac{IVTT_{car}}{IVTT_{BRT}+2\times IVTT_{Acc}}$	Zones	Share of Sample (%)
1	[0, 1[	5	10
2	[1, 1.5[	1, 2, 4	35
3	[1.5, 2[	3, 7	35
4	[2, ∞[	6, 8, 9	20

Stratum	# Responses	Zone	Population	%Pop in Stratum	# Responses
1	40	5	92827	100.00%	40
2	140	1	40000	36.70%	51
		2	17639	16.18%	23
		4	51354	47.12%	66
3	140	3	96370	83.83%	117
		7	18594	16.17%	23
4	80	6	8875	16.11%	13
		8	23671	42.96%	34
		9	22556	40.93%	33

Outline

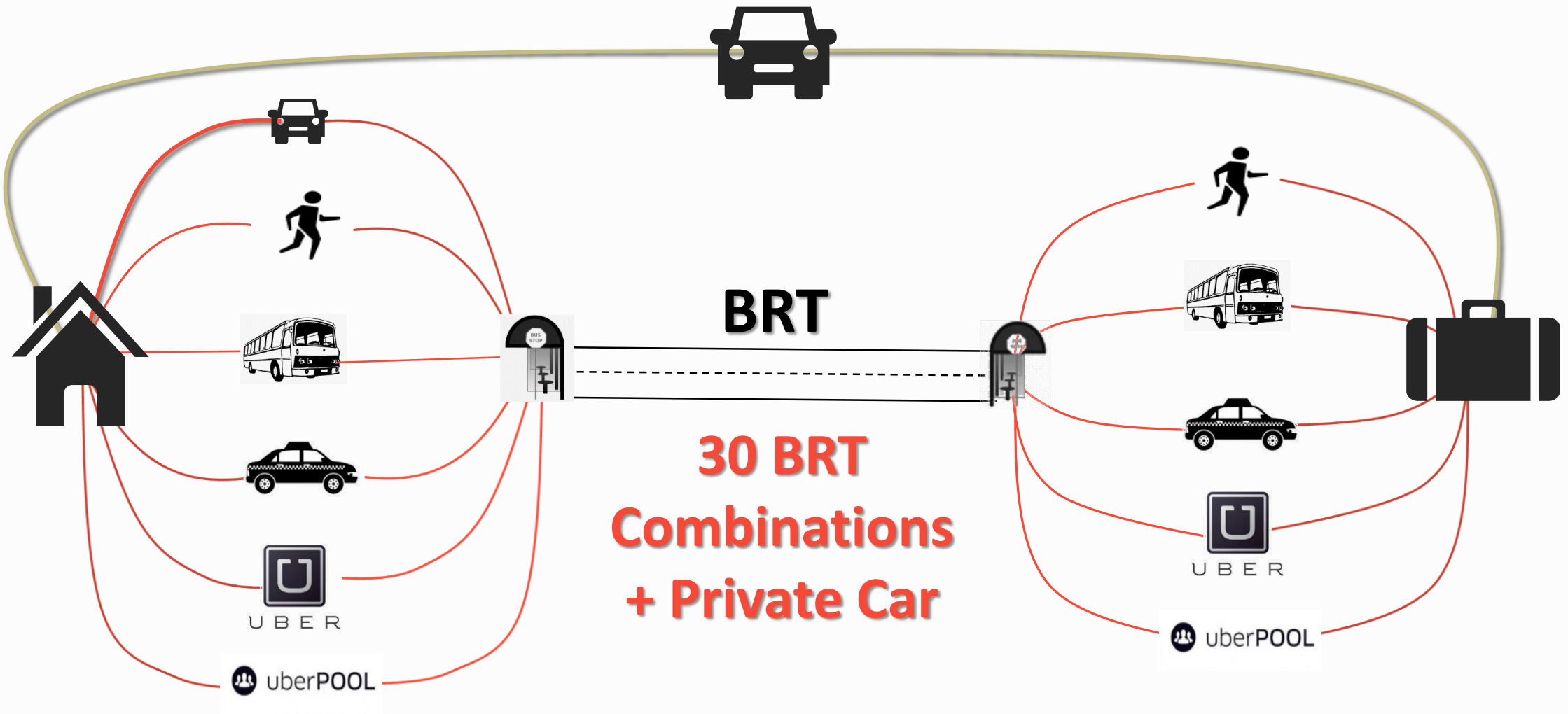
- ✓ Introduction
- ✓ Research Objectives
- ✓ Experimental Design
  - ✓ Stated Preference Design
  - ✓ Data Collection
- ✓ Demand Modeling
- ✓ Future Tasks

Modeling Specifications

- ✓ Alternatives
- ✓ Alternative Specific Constants
- ✓ Variables
- ✓ Market Segmentation for Car Time
- ✓ Nested Logit
  - ✓ Systematic Heterogeneity (Interacting cost with income)
  - ✓ Random Heterogeneity (Defining cost coefficient as lognormal)
- ✓ Cross Nested Logit
  - ✓ Systematic Heterogeneity (Interacting cost with income)
  - ✓ Random Heterogeneity (Defining cost coefficient as lognormal)

Note: All model specifications will be translated to mixed logit for estimation

Alternatives



Alternative Specific Constants

ACCESS MODE

- $ASC_{ParkAcc}$
- $ASC_{WalkAcc}$
- ASC_{BusAcc}
- ASC_{SerAcc}
- $ASC_{RidePriAcc}$
- $ASC_{RideShaAcc}$

$$V_{Park-BRT-Ser} = ASC_{ParkAcc} + ASC_{SerEgr} + \dots$$

EGRESS MODE

- $ASC_{WalkEgr}$
- ASC_{BusEgr}
- ASC_{SerEgr}
- $ASC_{RidePriEgr}$
- $ASC_{RideShaEgr}$

- ❑ **ASC_Car is set at zero**
- ❑ **Model is unidentifiable**
- ❑ **Set ASC_Acc = ASC_Egr for every mode.**

Alternative Specific Constants

ACCESS MODE

- ASC_{Park}
- ASC_{Walk}
- ASC_{Bus}
- ASC_{Ser}
- $ASC_{RidePri}$
- $ASC_{RideSha}$

EGRESS MODE

- ASC_{Walk}
- ASC_{Bus}
- ASC_{Ser}
- $ASC_{RidePri}$
- $ASC_{RideSha}$

$$V_{Park-BRT-Ser} = ASC_{Park} + ASC_{Ser} + \dots$$

$$V_{Ser-BRT-Ser} = 2ASC_{Ser} + \dots$$

Variables

The following variables are included in the models:

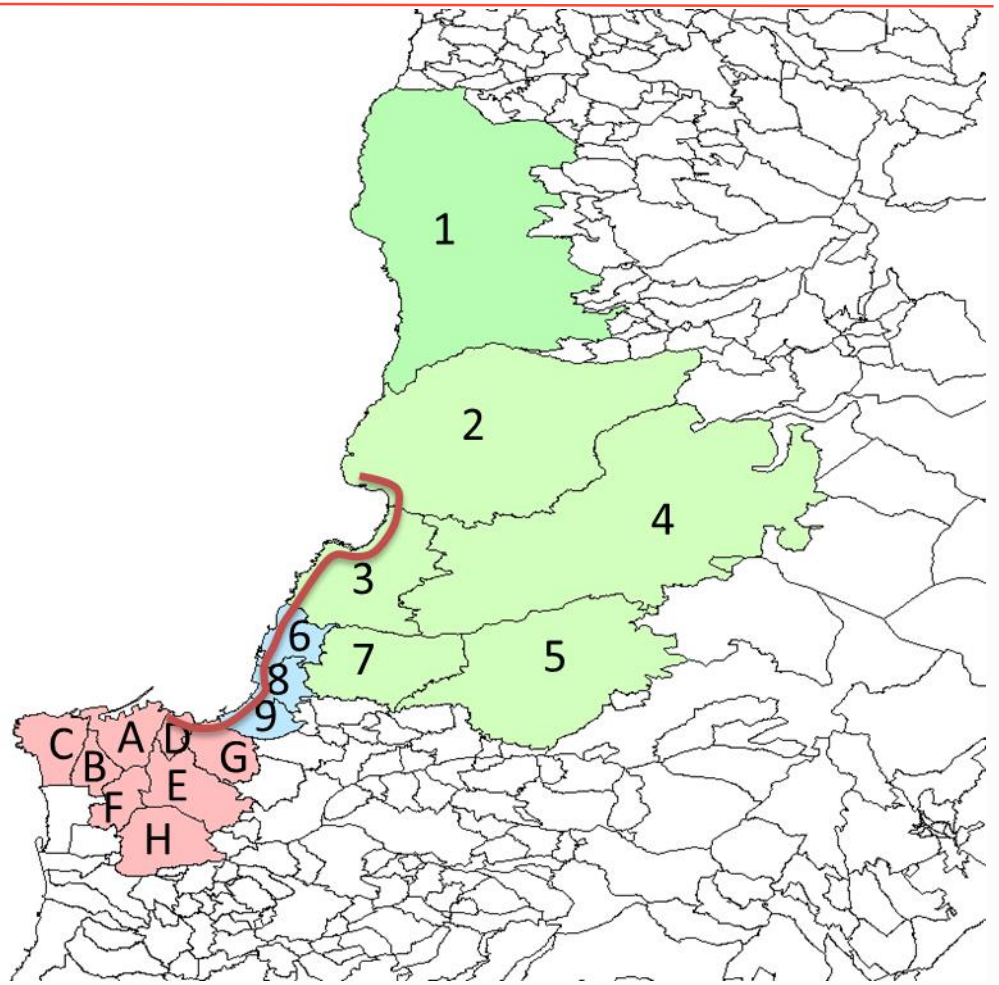
- $CarTime_n$: Time, in h, for travel by car for individual "n".
- $BRTTime_n$: Time, in h, for travel in the BRT for individual "n".
- $FeederTime_{i,j,n}$: Time, in h, for access and egress travel if access mode "i" and egress mode "j" are selected for individual "n".
- $Cost_n$: Overall trip cost, in 1000s of LBP*, for individual "n", defined for all alternatives.

*1 USD = 1500 LBP

- Age_n: Age, in years, for individual "n".
- $Female_n = \begin{cases} 1, if\ individual\ \text{"n"}\ \text{is a}\ Female \\ 0, otherwise \end{cases}$
- $RideUser_n = \begin{cases} 1, if\ individual\ \text{"n"}\ used\ ridesourcing\ previously \\ 0, otherwise \end{cases}$

Market Segmentation for Car Travel Time

- Wrong time coefficient obtained due to large discrepancies in car travel time between zones.
- Four different market segments defined as follows:
  1. Stratum 1: Zones 1, 2, and 4.
  2. Stratum 2: Zones 3 and 7.
  3. Stratum 3: Zones 6, 8, and 9.
  4. Stratum 4: Zone 5 (Eliminated).
- The three segments reflect sensitivity to car travel time based on trip duration:
  1. Stratum 1: long trip (> 90 min typically).
  2. Stratum 2: medium trip (60-90 min typically).
  3. Stratum 3: short trip (30-60 min typically).



Approaches to Randomize Value of Time

The following models are suggested to make VOT variable over alternatives:

Model 1: Systematic Heterogeneity

- Cost is divided by family income.
- The following variables are introduced:
  - ✓ Fam_Income_n: Monthly family income, in Million LBP, for individual "n".
  - ✓ $M\_Income_n = \begin{cases} 1, if\ missing\ income\ for\ "n" \\ 0, otherwise \end{cases}$

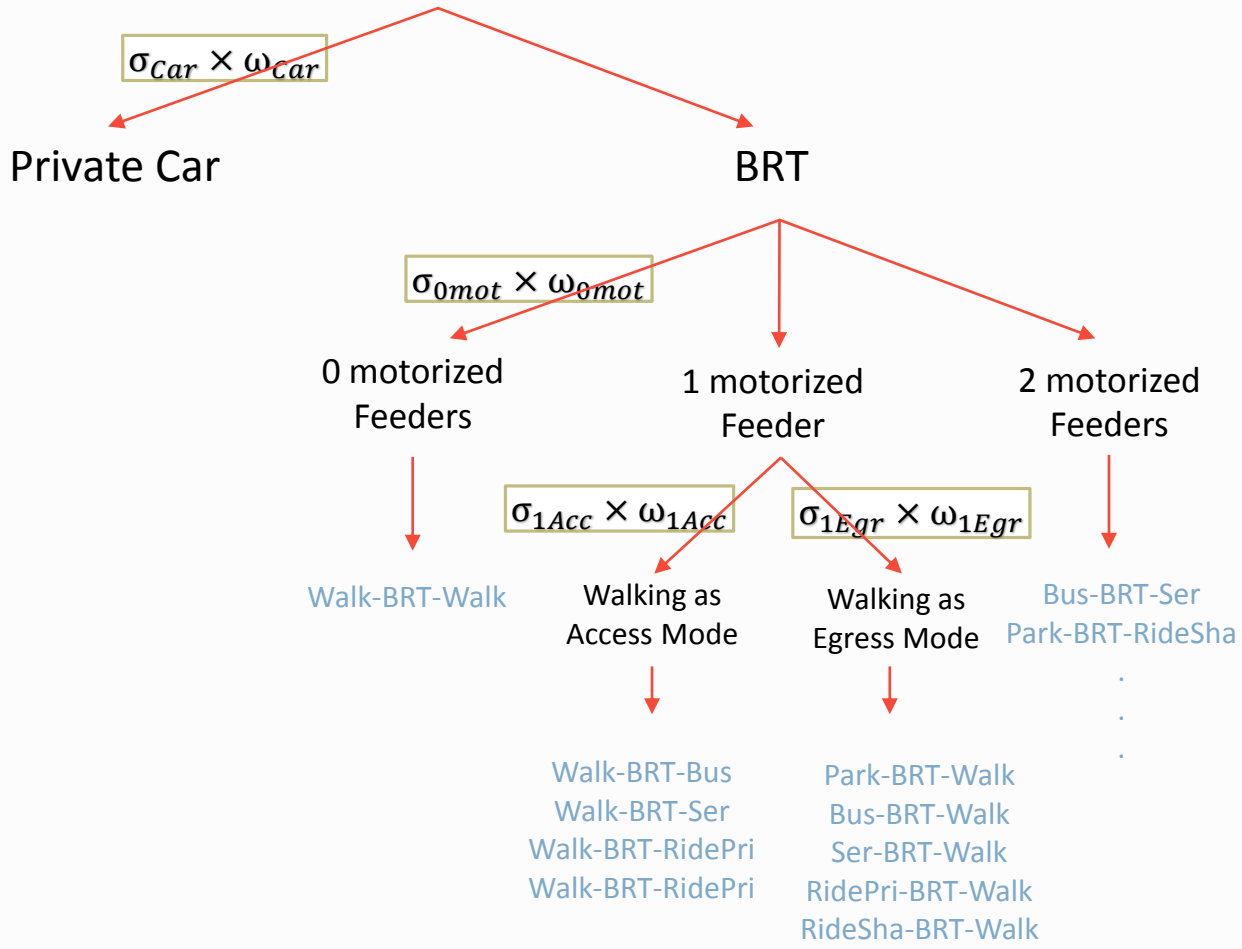
Model 2: Random Heterogeneity

- Lognormal distribution is adopted to keep cost coefficient negative for all individuals.
- $\beta_{Cost} = -\exp(\mu_{Cost} + \sigma_{Cost} \times \omega_{Cost})$
- μ_{Cost} & σ_{Cost} are parameters to be estimated.
- ω_{Cost} ~ N(0,1)

Nested Logit Specification

- ✓ Random components are added to account for correlation in unobserved variables between alternatives of the same nest.
- ✓ Sub-nesting for BRT is based on the number of motorized feeder modes.

- ✓ All "σ"s are parameters to be estimated.
- ✓ All "ω"s ~ N(0,1).



Nested Logit Estimation Results (Part 1 of 2)

Model 1: Systematic Heterogeneity

Variable	Coefficient	P-value
ASC_{Park}	1.09	0.50*
ASC_{Walk}	-0.559	0.74*
ASC_{Bus}	-0.866	0.60*
ASC_{Ser}	-0.504	0.76*
$ASC_{RidePri}$	-2.00	0.22*
$ASC_{RideSha}$	-0.211	0.90*
$CarTime1$	-1.18	0.21*
$CarTime2$	-2.16	0.05
$CarTime3$	-4.54	0.01
$BRTTime$	-2.85	0.23*
$FeederTime$	-3.45	0.00

Model 2: Random Heterogeneity

Variable	Value	P-value
ASC_{Park}	1.57	0.63*
ASC_{Walk}	-0.636	0.86*
ASC_{Bus}	-0.551	0.87*
ASC_{Ser}	0.113	0.97*
$ASC_{RidePri}$	-1.34	0.68*
$ASC_{RideSha}$	0.236	0.94*
$CarTime1$	-1.34	0.25*
$CarTime2$	-2.10	0.09*
$CarTime3$	-5.03	0.01
$BRTTime$	-2.81	0.33*
$FeederTime$	-4.73	0.00

All estimations performed on Python Biogeme using simulated maximum likelihood with 2000 draws

Nested Logit Estimation Results (Part 2 of 2)

Model 1: Systematic Heterogeneity

Variable	Coefficient	P-value
Age (Specific to Car Utility)	0.176	0.01
$Cost/Income$	-1.03	0.00
$Cost \times M\_Income$	0.129	0.28*
$Female\ (Specific\ to\ Car\ Utility)$	-2.76	0.06*
$RideUser\ (Specific\ to\ Ridesourcing\ Modes)$	0.660	0.12*
σ_{Car}	7.59	0.00
σ_{0mot}	4.79	0.00
σ_{1Acc}	-3.05	0.00
σ_{1Egr}	2.77	0.00

Model 2: Random Heterogeneity

Variable	Coefficient	P-value
Age (Specific to Car Utility)	0.202	0.27*
μ_{Cost}	-1.58	0.00
σ_{Cost}	-1.22	0.07
$Female\ (Specific\ to\ Car\ Utility)$	-2.90	0.10*
$RideUser\ (Specific\ to\ Ridesourcing\ Modes)$	0.707	0.12*
σ_{Car}	8.30	0.02
σ_{0mot}	5.14	0.11*
σ_{1Acc}	-3.53	0.00
σ_{1Egr}	2.9	0.00

All estimations performed on Python Biogeme using simulated maximum likelihood with 2000 draws

Cross-Nested Specifications

ACCESS MODE

 $\sigma_{Park} \times \omega_{Park}$

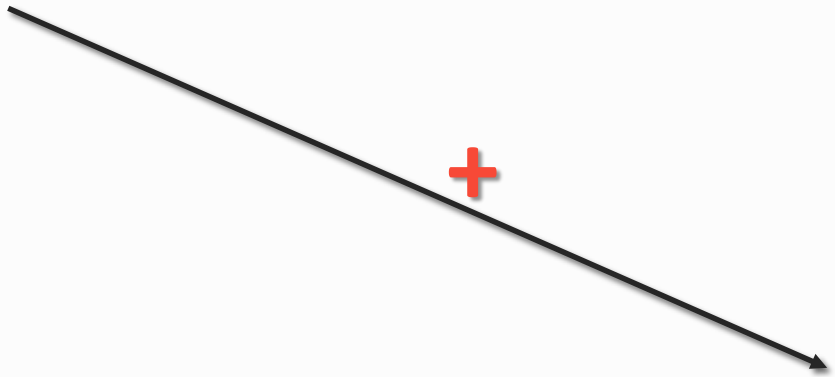
 $\sigma_{Walk} \times \omega_{Walk}$

 $\sigma_{Bus} \times \omega_{Bus}$

 $\sigma_{Ser} \times \omega_{Ser}$

 $\sigma_{RidePri} \times \omega_{RidePri}$

 $\sigma_{RideSha} \times \omega_{RideSha}$



$V_{Park-BRT-Ser} = \ldots + \sigma_{Park} \times \omega_{Park} + \sigma_{Ser} \times \omega_{Ser}$

$V_{Ser-BRT-Ser} = \ldots + 2\sigma_{Ser} \times \omega_{Ser}$

EGRESS MODE

 $\sigma_{Walk} \times \omega_{Walk}$

 $\sigma_{Bus} \times \omega_{Bus}$

 $\sigma_{Ser} \times \omega_{Ser}$

 $\sigma_{RidePri} \times \omega_{RidePri}$

 $\sigma_{RideSha} \times \omega_{RideSha}$

Cross Nested Logit Estimation Results (Part 1 of 2)

Model 3: Systematic Heterogeneity

Variable	Coefficient	P-value
ASC_{Park}	1.27	0.30*
ASC_{Walk}	0.401	0.75*
ASC_{Bus}	-0.987	0.46*
ASC_{Ser}	-1.35	0.29*
$ASC_{RidePri}$	-2.46	0.09*
$ASC_{RideSha}$	-0.0799	0.95*
$CarTime1$	-1.18	0.24*
$CarTime2$	-1.90	0.09*
$CarTime3$	-5.49	0.00
$BRTTime$	-3.43	0.22*
$FeederTime$	-4.03	0.00

Model 4: Random Heterogeneity

Variable	Value	P-value
ASC_{Park}	1.65	0.21*
ASC_{Walk}	0.143	0.91*
ASC_{Bus}	-1.04	-0.78*
ASC_{Ser}	-0.877	0.50*
$ASC_{RidePri}$	-1.77	0.25*
$ASC_{RideSha}$	0.0942	0.94*
$CarTime1$	-1.04	0.32*
$CarTime2$	-1.83	0.11*
$CarTime3$	-5.23	0.00
$BRTTime$	-4.45	0.11*
$FeederTime$	-4.39	0.00

All estimations performed on Python Biogeme using simulated maximum likelihood with 2000 draws

Cross Nested Logit Estimation Results (Part 2 of 2)

Model 3: Systematic Heterogeneity

Variable	Coefficient	P-value
Age (Specific to Car Utility)	0.233	0.00
$Cost/Income$	-1.64	0.00
$Cost \times M\_Income$	0.0463	0.78*
$Female\ (Specific\ to\ Car\ Utility)$	-2.60	0.02
$RideUser\ (Specific\ to\ Ridesourcing\ Modes)$	1.34	0.01
σ_{Car}	-8.63	0.00
σ_{Park}	-3.72	0.00
σ_{Walk}	2.43	0.00
σ_{Bus}	1.14	0.00
σ_{Ser}	2.44	0.00
$\sigma_{RidePri}$	1.44	0.00
$\sigma_{RideSha}$	-1.06	0.00

Model 4: Random Heterogeneity

Variable	Coefficient	P-value
Age (Specific to Car Utility)	0.232	0.00
μ_{Cost}	-0.729	0.00
σ_{Cost}	-0.696	0.00
$Female\ (Specific\ to\ Car\ Utility)$	-2.97	0.01
$RideUser\ (Specific\ to\ Ridesourcing\ Modes)$	1.39	0.01
σ_{Car}	-8.74	0.00
σ_{Park}	-3.69	0.00
σ_{Walk}	2.49	0.00
σ_{Bus}	1.25	0.00
σ_{Ser}	2.21	0.00
$\sigma_{RidePri}$	1.37	0.00
$\sigma_{RideSha}$	-1.32	0.00

All estimations performed on Python Biogeme using simulated maximum likelihood with 2000 draws

Models Comparison

Model	$L(\hat{\beta})$	AIC	BIC	VOT_{Car1}	VOT_{Car2}	VOT_{Car3}	VOT_{BRT}	VOT_{Feed}
1	-1281.84	2603.682	2702.926	1.14*Inc (3.42)	2.10*Inc (6.30)	4.41*Inc (13.23)	2.77*Inc (8.31)	4.54*Inc (13.62)
2	-1258.82	2557.64	2656.885	3.09	6.50	15.57	8.70	14.64
3	-1180.47	2406.958	2521.090	0.72*Inc (2.16)	1.16*Inc (3.48)	3.35*Inc (10.05)	2.09*Inc (6.27)	2.46*Inc (7.38)
4	-1168.81	2383.630	2497.762	1.42	2.59	7.40	6.30	6.21

- VOT is in USD/h; Income is in 1000 USD .
- Median VOT is reported for models 2 and 4 (obtained through simulation).
- For models 1 and 3, the value in brackets corresponds to the VOT for the median family income ($3,000/month based on sample).
- Cross nested model with lognormal cost coefficient seems to best fit the data so far.

Future Tasks

- **Model Estimation**
  - ✓Latent Variables
  - ✓Outlier Analysis
  - ✓Generic vs Alternative Specific Variables
  - ✓Socio-Economic Variables

- **Forecasting/Policy Analysis:**
  - ✓Share that will shift from private cars to the BRT.
  - ✓Modal split among feeder modes.
  - ✓Change in the share of BRT if parking prices are increased.
  - ✓Change in BRT share and modal split of feeders if ridesourcing fare is reduced.

My Questions

- Other approaches for models with large number of alternatives.
- Improvement of ASC significance (or different configuration).
- Improvement of main variables significance (Car Time, BRT Time).
- Sequential choice experiment and its effects on modeling.
- Advantages of sequential nested logit.

THANK YOU