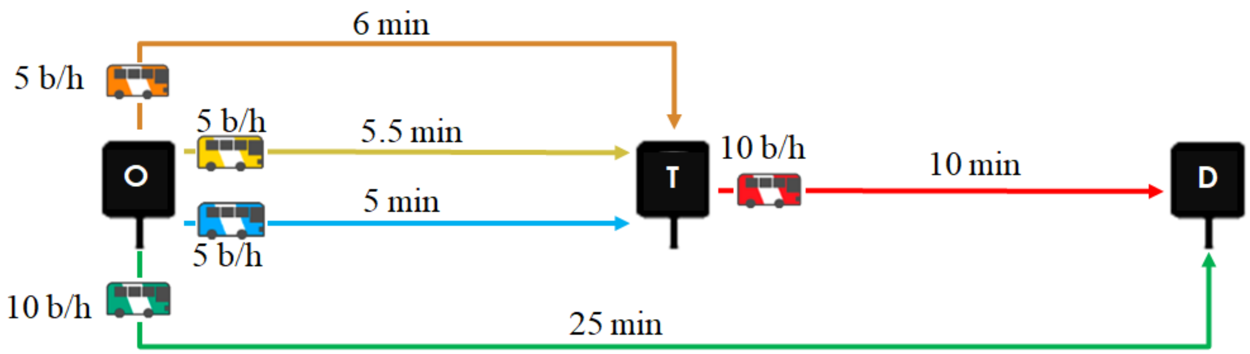
Observing and understanding route choice behaviour of public transport passengers from smart card data

Jacqueline Arriagada[a], Marcela Munizaga[a], Angelo Guevara[a] and Carlo Prato[b]

[a] Universidad de Chile, [b] University of Queensland

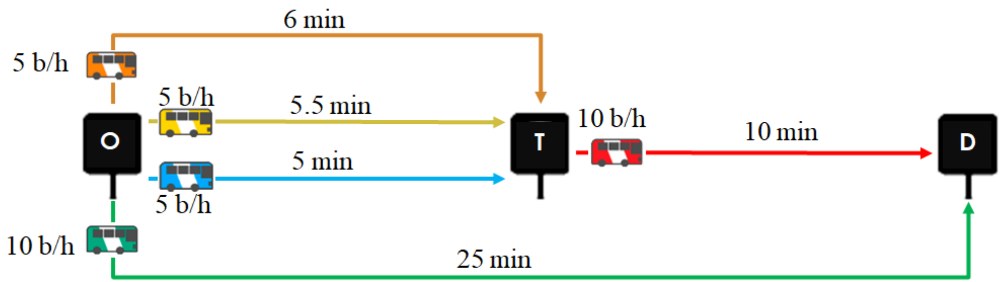
July 9, 2019

Route choice problem



Strategy : “A set of rules, when applied, allows the traveller to reach his or her destination” Spiess and Florian (1988)

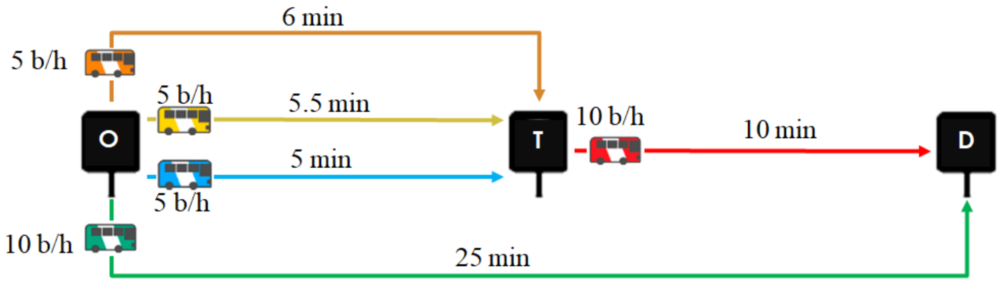
Route choice strategies



Simple strategy: "Take line blue to stop T; transfer to line red and exit at stop D"

	Waiting time [min]	Travel time [min]	Expected travel time [min]
O -  - T -  - D	12 + 6 = 18	6 + 10 = 16	34
O -  - T -  - D	12 + 6 = 18	5.5 + 10 = 15.5	33.5
O -  - T -  - D	12 + 6 = 18	5 + 10 = 15	33
O -  - D	**6**	**25**	**31**

Route choice strategies



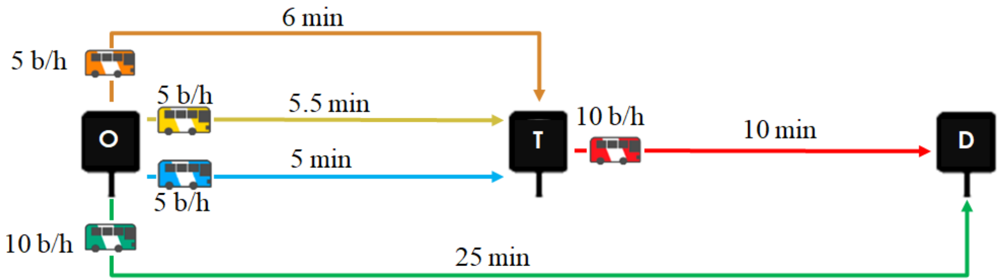
Aggregate strategy: "Take the next bus among lines orange, yellow and blue, exit at stop T; transfer to line red and exit at stop D"

	Waiting time [min]	Travel time [min]	Expected travel time [min]
O- orange/yellow/blue -T- red -D	**4 + 6 = 10**	**5.5 + 10 = 15.5**	**25.5**
O - green - D	6	25	31

$$5.5 = \sum_{i \in CL} P(i) * t_i \qquad P(i) = \frac{f_i}{\sum_{j \in CL} f_j}$$

Chiriqui and Robillard (1975); Raveau and Muñoz (2014)

Route choice strategies



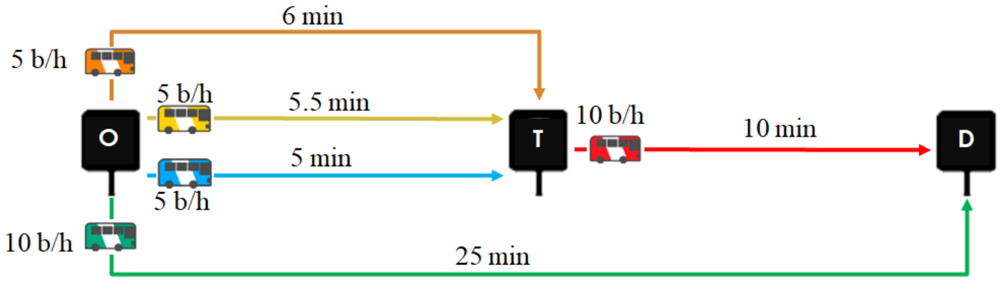
Hyperpath: set of path that minimize the total expected travel time (Spiess and Florian, 1988; Nguyen and Pallotino, 1998)

$$InitialWaitingTime = \frac{60}{25} = 2.4min$$

$$ExpectedTravelTime = 2.4 + \frac{5}{25} * 22 + \frac{5}{25} * 21.5 + \frac{5}{25} * 21 + \frac{10}{25} * 25$$

$$ExpectedTravelTime = 25.3$$

Route choice strategies

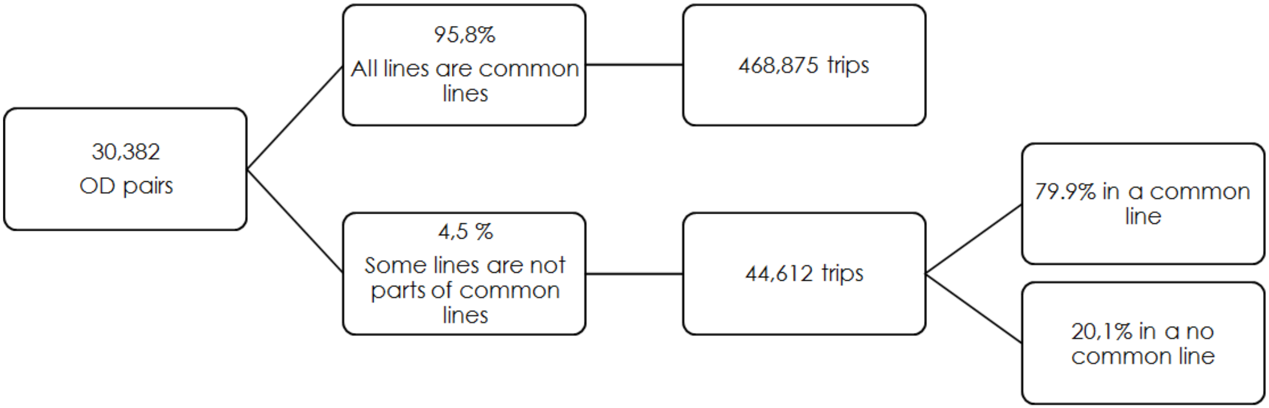


- Simple strategy or elemental alternatives are used in RUM models
- Aggregate strategy was incorpored in a RUM by Raveau and Muñoz (2014)
- Hyperpath are widly used in Transit assignment models

Data studio: Santiago

- Payment option: smart card
- 300 bus services, all of them in both directions
- 7 metro lines (more than 100 km)
- +11.000 bus stops
- 3 millon of passengers per week

Common line analysis



Are people taking the first line or service that arrive to the stop?

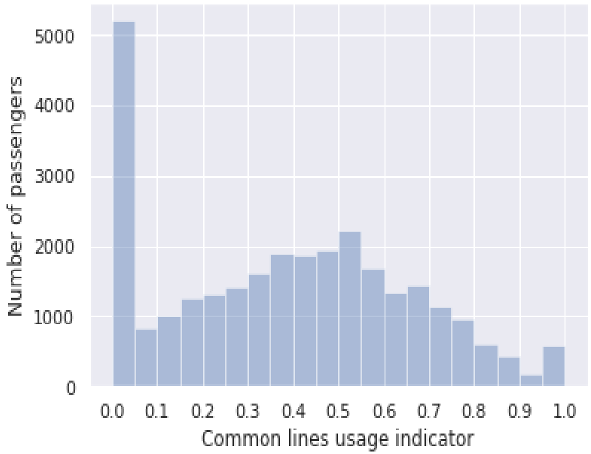
Common line analysis

$$q_p = \sum_{r=1}^{R_p} \sum_{l=1}^{L_r} \frac{\pi_{lrp}}{\sum_{r=1}^{R_p} L_r}$$

$$\pi_{lrp} = \frac{|ET_{lrp} - OT_{lrp}|}{max(ET_{lrp}, OT_{lrp})}$$

$$\mathrm{ET}_{lrp} = \frac{f_l}{\sum_{i \in CL} f_i} * \sum_{i=1}^{L_r} OT_{irp}$$

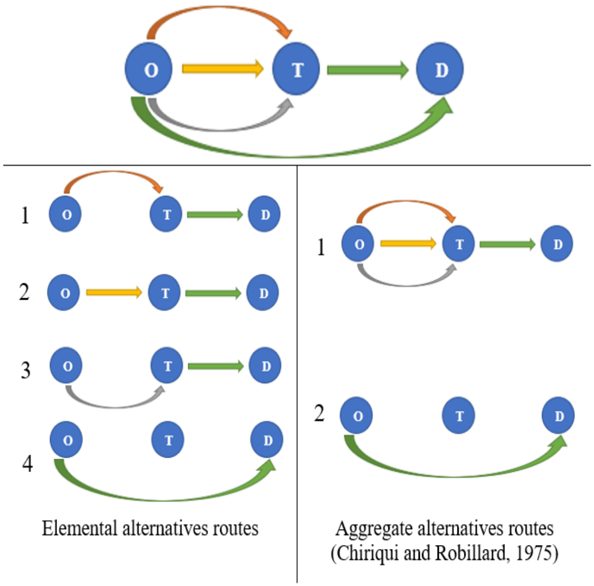
$$\mathrm{q}_p \in [0,1]$$



ET_{lrp}: expected trips of passenger p in line r at OD pair r

OT_{lrp}: observed trips of passenger p in line r at OD pair r

Path size logit model



$$V_j = \sum_m \beta_m TI_m + \sum_c \beta_c TR_m + \beta_{CR} CR_j + \beta_{PS} PS_j$$

$$CR_{ls} = \frac{\sum_{e \in l} Occupancy_e}{\sum_{e \in l} Capacity_e}$$

$$CR_{js} = \frac{\sum_{s \in l} CR_{ls}}{L_j}$$

$$PS_j = \sum_{a \in j} \frac{l_a}{L_j} \ln \frac{1}{M_{an}}$$

Source: Bovy et al. (2009)

Path size logit model

	Elemental alternatives		Aggregate alternatives	
Description	Estimated coefficient	Rates of substitution	Estimated coefficient	Rates of substitution
$\beta_{BusTravelTime}$	-0.067 (-47.0)	1	-0.067 (-41.2)	1
$\beta_{MetroTravelTime}$	-0.068 (-35.9)	1.01	-0.073 (-34.8)	1.09
$\beta_{InitialWaitingTime}$	-0.099 (-83.4)	1.48	-0.133 (-56.5)	1.99
$\beta_{TransferWaitingTime}$	-0.083 (-24.5)	1.24	-0.077 (-11.1)	1.15
$\beta_{TransferWalkingTime}$	-0.320 (-46.4)	4.78	-0.099 (-10.3)	1.48
$\beta_{TransferBusBus}$	-1.070 (-20.6)	15.97	-0.771 (-13.1)	11.51
$\beta_{TransferBusMetro}$	-0.84 (-21.8)	12.54	-0.771 (-17.6)	11.51
$\beta_{TransferMetroBus}$	-0.92 (-6.7)	13.73	-0.995 (-7.0)	14.85
$\beta_{BusCrowding}$	-0.85 (-18.8)	12.69	-0.858 (-9.6)	12.8
$\beta_{PathSizeTerm}$	-0.073 (-2.6)		0.030 (0.5)	
	Obsv: 154,335 $\mathcal{L}(\hat{\beta})$= -161000.6 $\bar{\rho}^2$= 0.046		Obsv: 154,335 $\mathcal{L}(\hat{\beta})$= -35978 $\bar{\rho}^2$= 0.083	

Latent class model

$$P_p(l) = P(l|CL) * P_p(CL) + P(l|\overline{CL}) * P_p(\overline{CL})$$

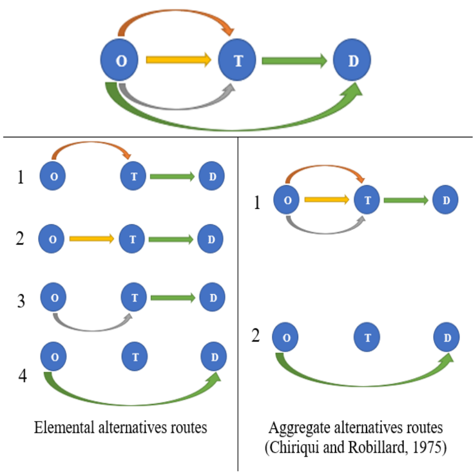
$$P_p(\overline{CL}) = \frac{\exp\left(\beta_{\overline{CL}} + \beta_{q_p} * q_p\right)}{1 + \exp\left(\beta_{\overline{CL}} + \beta_{q_p} * q_p\right)}$$

$$q_p = \sum_{r=1}^{R_p} \sum_{l=1}^{L_r} \frac{\pi_{lrp}}{\sum_{r=1}^{R_p} L_r}$$

$$\pi_{lrp} = \frac{|ET_{lrp} - OT_{lrp}|}{max(ET_{lrp}, OT_{lrp})}$$

Latent class model

$$P_p(l) = P(l|CL) * P_p(CL) + P(l|\overline{CL}) * P_p(\overline{CL})$$



$$\mathrm{P}(1|\overline{CL}) = \frac{\exp(V_1)}{\exp(V_1)+\exp(V_2)+\exp(V_3)+\exp(V_4)}$$

$$\mathrm{P}(1|CL) = \frac{\exp(V_1)}{\exp(V_1)+\exp(V_2)} * \frac{f_{orange}}{f_{orange}+f_{yellow}+f_{gray}}$$

Latent class model

	Class 1: Elemental alternatives		Class 2: Aggregate alternatives	
Description	Estimated coefficient	Rates of substitution	Estimated coefficient	Rates of substitution
$\beta_{BusTravelTime}$	-0.090 (-41.2)	1	-0.058 (-19.4)	1
$\beta_{MetroTravelTime}$	-0.085 (-29.5)	0.944	-0.068 (-17.8)	1.172
$\beta_{InitialWaitingTime}$	-0.095 (-53.3)	1.056	-0.095 (-22.5)	1.638
$\beta_{TransferWaitingTime}$	-0.067 (-12.9)	0.74	-0.093 (-7.5)	1.60
$\beta_{TransferWalkingTime}$	-0.484 (-37.9)	5.38	-0.037 (-2.2)	0.64
$\beta_{TransferBusBus}$	-0.850 (-10.7)	9.44	-0.994 (-8.5)	17.14
$\beta_{TransferBusMetro}$	-0.964 (-14.6)	10.71	-0.682 (-9.3)	11.76
$\beta_{TransferMetroBus}$	-2.005 (-6.7)	22.28	-0.479 (-1.9)	8.26
$\beta_{BusCrowding}$	-1.410 (-20.5)	15.67	-0.299 (-2)	5.16
$\beta_{PathSizeTerm}$	-0.374 (-8.0)		0.323 (3.4)	
$\beta_{\overline{CL}}$		-3.329 (-11.2)		
β_{q_p}		10.571 (-12.14)		
N°Observations:		154,335		
$\mathcal{L}(\hat{\beta})$		-160311.9		
$\bar{\rho}^2$		0.029		

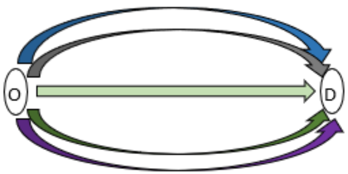
Comparison of models

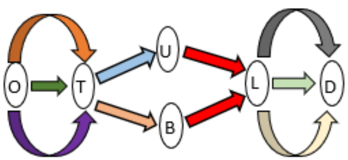
First Preference Recovery (FPR): proportion of the cases in which the model assigns the maximum probability to the chosen alternative (Ortúzar, J. de D., 1982).
Expected recovery (ER): Probability average of the chosen alternative (Ortúzar, J. de D., 1982).

Model	FPR	ER
Path Size Logit (elemental alternatives)	0.465	0.464
Path Size Logit (aggregate alternatives)	0.459	0.468
Latent class	0.468	0.471

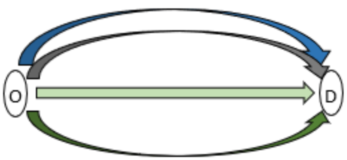
Hyperpaths

Example

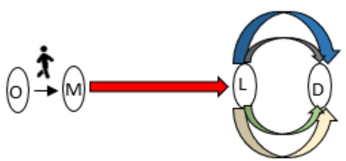




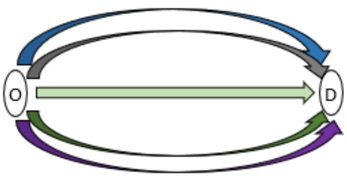
Hyperpath, 18 paths



Hyperpath with transfer penalization, 4 paths

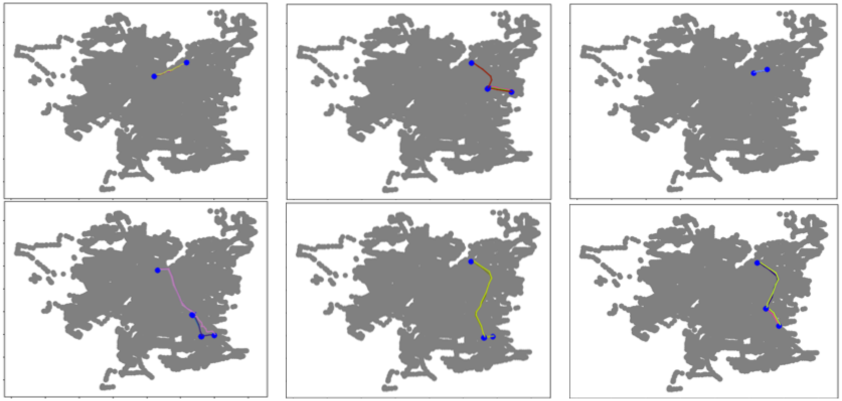


Hyperpath with waiting time penalization, 4 paths



Hyperpath with transfer and waiting time penalization, 5 paths

Hyperpaths



Indicator	N°Trips	hyperpath	hyperpath T	hyperpath WT	hyperpath T-WT
ER	776	0.223	0.310	0.224	0.329
FPR	776	0.294	0.320	0.294	0.351

T: transfer penalty, WT: waiting time penalty, T-WT: transfer and waiting time penalty

Conclusions

Conclusions

- Smart card data allows to study passenger behaviour and estimate route choice models
- There are two groups of passengers: (1) elemental alternatives and (2) aggregate alternatives
- Metro travel time and waiting time is more burdensome for individuals in group (2).
- Walking time is more burdensome for individuals in group (1).
- It is necessary to penalize waiting time and transfer time to construct the optimal hyperpath