

Question 1: Correcting the constants

Consider the endogenous sampling strategy where the strata are defined based on car availability and the chosen alternative. The proportion of individuals in each group is reported in Table 1 for the population, and in Table 2 for the sample.

Choice	Car availability	
	Yes	No
Public transportation	0.234751	0.05092
Car	0.649256	0.0
Slow modes	0.064707	0.000366

Table 1: Proportion of each group in the population

Choice	Car availability	
	Yes	No
Public transportation	0.2	0.2
Car	0.2	0.0
Slow modes	0.2	0.2

Table 2: Proportion of each group in the sample

We have performed an experiment based on a synthetic population, such that we know the true model, and the true value of the parameters. We have extracted 200 samples using the sampling strategy described above, and have estimated using Exogenous Sampling Maximum Likelihood (ESML) 200 sets of parameters. Table 3 reports, for the alternative specific constants, their true value, the estimated value (that is, the mean across the 200 realizations), the empirical standard deviation of the 200 realizations, and the t-test, calculated as the difference between the true and the estimated value, divided by the standard deviation. It is used to test the hypothesis that the estimated value is equal to the true value.



Mathematical Modeling of Behavior (2024-2025)



Michel Bierlaire.

TAs: Cloe Cortes, Pavel Ilinov, Evangelos Paschalidis, Anne-Valérie Preto, Negar Rezvany

Sampling — November 19, 2024

	True	Estimated	StdDev	t-test
ASC_CAR_FRENCH	-2.396417	-4.628909	1.220144	-1.829696
ASC_CAR_GERMAN	-3.507716	-5.827157	0.975760	-2.377059
ASC_PT_FRENCH_CARAVAIL	-2.361225	-3.613808	0.824435	-1.519323
ASC_PT_FRENCH_NOCARAVAIL	4.620998	-0.627619	0.957465	-5.481784
ASC_PT_GERMAN_CARAVAIL	-2.343931	-3.686178	0.480577	-2.792990
ASC_PT_GERMAN_NOCARAVAIL	2.452726	-2.543666	0.340305	-14.682115

Table 3: Estimation results

The values of the t-statistics show that the estimated values of the parameters are significantly different from the true values.

1. Why is it the case?
2. Using Tables 1 and 2, calculate the corrections that must be applied to the estimated values to recover the true values.
3. Calculate the t-statistics for the corrected values.