**Selecting the right model and tree structure:**

1. Multinomial model

**Root**

Driver

Passenger

Walk

Bicycle

PT

Walk

1. Nested logit model: tree structure option 1

**Root**

Car

Driver

Passenger

Walk

Bicycle

PT

Walk

1. Nested logit model: tree structure option 2

**Root**

Car

Driver

Passenger

Walk

Bicycle

PT

Walk

Active

**LR test results for NL model option 1 versus MNL model:**

LR test result shows that NL model option 1 has a better fit than MNL data.

LL par

Model MNL -4396.69 84

Model NL -4383.83 85

Difference 12.86 1

Likelihood ratio test-value: 25.72

Degrees of freedom: 1

Likelihood ratio test p-value: 0.0000003947

**LR test results for NL model option 2 versus MNL model:**

LR test result shows that NL model option 2 has a better fit than MNL data.

LL par

Model MNL -4396.69 84

Model NL -4383.83 86

Difference 12.86 2

Likelihood ratio test-value: 25.72

Degrees of freedom: 2

Likelihood ratio test p-value: 0.0000026

**NL models option 2 versus option 1:**

Using the nesting structure as option 2, the model couldn’t estimate all the parameters and NA produced for SE and other parameters of the model. Besides, the Lambda value (dissimilarity parameter) for the active nest was above 1 (4.08). this means that model was consistent with the theoretical derivation and the NL tree structure should be rejected. The constraints were also applied to the active nest to keep the lambda value below 1. The estimated lambda was 1 which implies zero correlation among mode pairs in the active nest so the NL model is rejected.

**Variables in models**

*Socio-demographic variables:*

Some of the socio-demographic variables were reclassified to have more balanced data in their categories.

Age group:

Reclassified categories: 5-14, 15-24, 25-34, 35-44, 45-54, 55-64, 65+

Gender:

Male, Female

Household income

Reclassified categories: "less than £14999", "£15000 to £24999", "£25000 to £34999","£35000 to £49999", "more than £50000","missing/refused to respond"

Household structure”

Reclassified categories: “Households with children”, “Household without children”

Work type:

Reclassified categories: "full time", "part time/casual/volunteer", "retired", "studying full/part time", "not in work force/other"

Number of cars in households:

Reclassified categories: 0,1, 2, 3 or more

Number of bikes in households

Reclassified categories: 0,1, 2, 3, 4 or more

*Route attributes included in the model:*

Attractiveness (as a composite measure of Eye-level greenness, Shannon index, Presence of POIs, Presence of negative POIs, Crime records and Street light count along with stress link and stress junction have been included in the nested logit models. For each route type attribute (JIBE, Short and Fast) 4 models have been run to reflect the effect of travel distance in the utility functions:

1. Created a log transformation of walk and bike distance, got the average of them, and included that variable as a weighting variable in the apollo controls.
2. Used the variables that have been already multiplied by the distance (original values in the data you shared)
3. Multiplied route attributes (except stress junction) in the log-transformed distances in the utility functions
4. Multiplied route attributes (except stress junction) in the actual distance in the utility function.

As shown in table 1 among the three route types, from M1 to M4, JIBE has had the best.

Table 1. comparing models using model selection criteria

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | AIC\* | | | BIC\* | | |
| Models | JIBE | Short | Fast | JIBE | Short | Fast |
| M1 | 71509.14 | 71566.87 | 71597.29 | 71990.69 | 72048.42 | 72078.84 |
| M2 | 8770.19 | 8869.77 | 8857.11 | 9251.74 | 9351.32 | 9338.66 |
| M3 | 8705.42 | 8711.57 | 8714.48 | 9186.97 | 9193.12 | 9196.03 |
| M4 | 8770.21 | 8868.92 | 8856.53 | 9251.76 | 9350.47 | 9338.09 |

\*Smaller values show better fit.