**Instructions for LatentGold 3-step Model**

1. **Step 1: Identifying the optimal number of clusters** 
   1. Run the model with 1-6 clusters

The attributes to be selected are: "Work onboard"; '"Spare time onboard"; "Getting ready onboard"; '"Meal onboard"; "Work stationary"; "Spare time stationary"; "Sleep stationary"; "Getting ready stationary"; "Meal stationary"; "DifferenceWorkTrip"; "DifferenceHomeTrip";

A screenshot of a computer

Description automatically generated

* 1. Compare the outcomes and select the best number of clusters

The models are compared by looking at their BIC value, number of parameters, class error, and bivariate residuals (see section 4.2 for more details). We select 5 clusters as the optimal number.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. **Step 2: Producing the classification model**
   1. Open source file: Activities\_Travel\_Differences.sav

A white board with a black border

Description automatically generated

* 1. Go to Model -> Cluster

A computer screen shot of a computer

Description automatically generated

* 1. Select Indicators (attributes) and set them all to “continuous”

The attributes to be selected are: "Work onboard"; '"Spare time onboard"; "Getting ready onboard";

'"Meal onboard"; "Work stationary"; "Spare time stationary"; "Sleep stationary"; "Getting ready stationary"; "Meal stationary"; "DifferenceWorkTrip"; "DifferenceHomeTrip";

A screenshot of a computer

Description automatically generated

* 1. Set the number of clusters to 5

A screenshot of a computer

Description automatically generated

* 1. Go to ClassPred, select Classification- Posterior, and give a name to the output file (which will contain the posterior probabilities)

A screenshot of a computer

Description automatically generated

* 1. Select the covariates and press Keep

The covariates to be selected are: "Gender"; '"Age\_group"; "Income\_group"; “Work\_type”, “travelFrequency”; “travelMode”; “travelTime”; “isCarOwner”, “travelTimeBehaviorChange”; “avFrequency”, “hasMotionSickness”; “newTechnologyAdaption”; “selfDrivingKnown”; “considersSelfDrivingCar”; “DailyTimePressure”; “isWorkPossibleInCar” ; “surveyDifficulty”; “Automationlevel”; “step1Time”; “step4Time” ; “Planner1Fragments”; “CopyCurrentDay”.

A screenshot of a computer

Description automatically generated

* 1. Go to Technical, keep the default values except the following:

Random sets= 3000

Iterations =300

A screenshot of a computer

Description automatically generated

* 1. Estimate Model

A screenshot of a computer screen

Description automatically generated

Disclaimer: A characteristic of LatentGold is that when the indicators are continuous and have values ​​between 0 and 1, then the LL will always be positive (residual variance will be less than 1 by definition). Due to the attributes in our dataset having values below 1 as they were divided over the travel duration, LatentGold produces a model with a positive LL. In order to remedy this, we were advised by the LatentGold developers to adjust the scale to 0-100 by multiplying by 100. They have confirmed that the cluster solution remains the same. The values produced in the “Indicators” section are the same only multiplied by 100. In the analysis phase, we divide them by 100 in order to facilitate the analysis.

* 1. You can save the results and definition by going to File -> Save results or Save definition

A screenshot of a computer

Description automatically generated

1. **Step 3: Adding the covariates and running the 3 step analysis** 
   1. Open the classification model output (Classification\_model.sav), and select step3 in the Model tab

A computer screen shot

Description automatically generated

* 1. Drag the posterior classifications (clu#1, clu#2, clu#3, clu#4, clu#5) to the Posteriors pane. Select “Covariates” in the dialog box Analysis

A screenshot of a computer

Description automatically generated

* 1. Drag the external covariates of interest to the Covariates pane. We had initially added all covariates and systematically excluded the non-significant ones based on the Wald statistic produced. Set them all to numeric

The covariates found to be significant in our final model are: "Gender"; “Education”, “travelTime”; “avFrequency”, “considersSelfDrivingCar”; “DailyTimePressure”; “isWorkPossibleInCar”’ “step1Time”; “Planner1Fragments”; “CopyCurrentDay”.

A screenshot of a computer

Description automatically generated

* 1. Select the classification type. Because our indicators are continuous, select BCH to reduce biases

A screenshot of a computer

Description automatically generated

* 1. Next click on the ‘Output’ Tab, here you can set the coding to ‘effect coding’ for this, the type of standard errors, and the output to be reported

A screenshot of a computer

Description automatically generated

* 1. Next click on the ‘Technical’ Tab. Select ‘Include All’ for Missing Values

A screenshot of a computer

Description automatically generated

* 1. Click on ‘Estimate’

A screenshot of a computer

Description automatically generated

* 1. You can save the results and definition by going to File -> Save results or Save definition

A screenshot of a computer

Description automatically generated