Instructions to Use the Transportation Modeling Package

MDCEV Forecasting Model with no Outside Good

MDCEV Forecasting Model with no Outside Good

- This is an addendum to help understand how to use the R script for running the MDCEV Forecasting Model with no Outside Good, and with no mixing. The input dataset specifications, a description of the test data, and the outputs from the test data are also discussed in this documentation.
- For further details, please refer to Bhat (2007) "The Multiple Discrete-Continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extensions," Technical paper, Department of Civil Engineering, The University of Texas at Austin, available at http://www.caee.utexas.edu/prof/bhat/MDCEV.html

MDCEV Forecasting Model with no Outside Good: R Scripts

- In the package, there are four scripts that are provided for forecasting an MDCEV model with no outside good:
  - **mdcev_alpha_forecast.r and mdcev_gamma_forecast.r:** These are the scripts containing the actual MDCEV code based on alpha and gamma profiles. Please do not modify these scripts.
  - **Run_mdcev_alpha_forecast.r and Run_mdcev_gamma_forecast.r:** These R scripts contain the settings and parameters that need to be specified for forecasting the MDCEV model for the two configurations. Users can update the model parameters and settings using this script.

MDCEV Forecasting Model with no Outside Good: Preparing the Input Data

- The input data to the MDCEV forecasting R-code should be in the form of a Comma Separated Values (CSV) file. The input data **should** include the following columns with the variables names given below:
  - A column of 1s with a variable name "uno" (Mandatory)
  - A column of 0s with a variable name "sero" (Mandatory)
  - Case ID: A column of observation numbers (or case numbers) from 1 to the number of observations in the data. The name of this column should be "caseid" (Mandatory)

MDCEV Forecasting Model with no Outside Good: Preparing the Input Data

- Dependent variables: As many columns as the number of alternatives, with each column containing the expenditure amount for each alternative. (Mandatory)
- Price variables: As many columns as the number of alternatives, with each column containing the price per unit consumption for each alternative. This data is required only when there is price variation across goods. (Optional)
- Weights: One column of estimation-weights to be applied to each decision-maker. This data is required only when working with weighted data. (Optional)
- Explanatory variables: One column for each explanatory variable. (Mandatory)

MDCEV Forecasting Model with no Outside Good: Running the Model Using the R Script

- **Step 1:** Based on the desired configuration, open the R-code "Run_mdcev_alpha_forecast.r" or "Run_mdcev_gamma_forecast.r".
- **Step 2:** Update the path of the source code (mdcev_alpha_forecast.r or mdcev_gamma_forecast.r) based on your chosen configuration, and the path and name of the input data as well as the random numbers file (here, halton.csv) as required.
- **Step 3:** Change the settings of the MDCEV forecasting model in the R-code as required.
- **Step 4:** In the independent variable section, update the coefficients of the chosen explanatory variables based on the estimation model or another source.
- The data dictionary for the dataset used here is as follows.

MDCEV Forecasting Model with no Outside Good: Running the Model Using the R Script

Data dictionary for "testnout.csv" data

Variable Name	Description
config	Utility specification configuration, possible values: 1, 4, 5, 6, or 7
alp0to1	1 if you want the Alpha values to be constrained between 0 and 1; 0 otherwise
price	1 if there is price variation across goods; 0 otherwise
nc	Number of alternatives (in the universal choice set) including outside goods
po	Position of pointer to case number in data set
ivuno	Position of UNO variable (i.e., the column of ones) in data set
ivsero	Position of SERO variable (i.e., the column of zeros) in data set
nrep	Number of sets of error term Halton draws over which you want to simulate the unobserved heterogeneity
halton_startrow	Starting rows of Halton Sequence data for loading
def	DEPENDENT variables (i.e., the consumption quantities for each alternative - NOT consumption expenditures for each alternative) (number of labels = number of alternatives).
fp	PRICE variables (number of labels = number of alternatives). First good (i.e., outside good, or the first of the outside goods if there are several) should be specified as a numeraire good with price one (UNO). Provide all UNO variables if there is no price variation

MDCEV Forecasting Model with no Outside Good: Running the Model Using the R Script

Data dictionary for "testnout.csv" data (Cont.)

Variable Name	Description
ivmt	• Definition of INDEPENDENT variables • First 'numout' goods are those that are always consumed (i.e., the outside goods), and the first good is numeraire • In the following specification, ivmt[[1]], ivmt[[2]], ivmt[[3]] contain independent variable specifications (on right hand side) for baseline utility (PSI) for alternatives 1, 2, and 3 • Add a row for ivmt[[4]] if there is a 4th alternative, another additional row for ivmt[[5]] if there is a 5th alternative, ... (number of rows = number of alternatives) • Number of columns = Number of variables including alternative specific constants; consider the first alternative as base
ivdt	• In the following specification, ivdt[[1]], ivdt[[2]], ivdt[[3]] contain input data specifications (on the right hand side) for satiation parameters (alphas) for alternatives 1, 2, and 3 • Add a row for ivdt[[4]] if there is a 4th alternative, another additional row for ivdt[[5]] if there is a 5th alternative,.... (number of rows = number of alternatives) • Number of columns = Number of alternatives; Note that you can also add individual-specific variables below, so that satiation varies across individuals • However, you will then have to translate outputs to compute actual alpha parameters • This code is written to provide you with the alpha parameters directly for the case when there is no variation in alpha across individuals

MDCEV Forecasting Model with no Outside Good: Running the Model Using the R Script

Data dictionary for "testnout.csv" data (Cont.)

Variable Name	Description
ivgt	• In the following specification, ivgt[[1]], ivgt[[2]], ivgt[[3]] contain input data specifications (on the right hand side) for translation parameters (gammas) for alternatives 1, 2, and 3 • Add a row for ivgt[[4]] if there is a 4th alternative, another additional row for ivgt[[5]] if there is a 5th alternative,....(number of rows = number of alternatives) • Number of columns = Number of alternatives; Note that you can also add individual-specific variables, so that gamma varies across individuals • However, you will then have to translate outputs to compute actual gamma parameters • This code is written to provide you with the gamma parameters directly for the case when there is no variation in gamma across individuals • Since gamma=0 for the outside goods, the first "numout" columns of the "ivgt" vectors will be "sero"
ivmtc	The corresponding estimated coefficients of variables ivmt
ivdtc	The corresponding estimated coefficients of variables ivdt
ivgtc	The corresponding estimated coefficients of variables ivgt

MDCEV Forecasting Model with no Outside Good: Running the Model Using the R Script

- **Step 5:** Select all and run (Ctrl+A and Ctrl+R).

Note 1: Detailed information regarding the settings of the MDCEV model is available within the run forecast R code.

Note 2: In alpha profile, coefficient values for the corresponding satiation parameters (ivdt) are set, as opposed to ivgt parameter in gamma profile.

MDCEV Forecasting Model with no Outside Good: Results from the Test Data – "testnout.csv"

- The output files of the forecasting model for the two configurations are provided as a separate attachment.