Technical Appendix

Thesis topic: "Socio-economic determinants of return migration intentions.
A case study from Bulgaria"

Master's programme: Applied Econometrics and Economic Modelling

Student: Mariya Hristova

Faculty number: 11250

February, 2023

Data Description

The files include the original questionnaire in Bulgarian; a data dictionary
with all of the variables and descriptions; a csv with the mapping between
the variables and the original questions.

- **Appendix_MariyaHristova_11250_Feb2023_questionnaire.doc**

- **Appendix_MariyaHristova_11250_Feb2023_data_description.xlsx**

- **Appendix_MariyaHristova_11250_Feb2023_data_mapping.csv**

Data Sets

The files include the raw data in a .sav and in a .csv format; an intermediate
dataset used only for data exploration part; the cleaned data set after the
data preparation step; a final dataset used for modelling after the feature
selection step.

- **Appendix_MariyaHristova_11250_Feb2023_raw_data.sav**

- **Appendix_MariyaHristova_11250_Feb2023_raw_data.csv**

- **Appendix_MariyaHristova_11250_Feb2023_data_for_exploration.csv**

- **Appendix_MariyaHristova_11250_Feb2023_data_clean.csv**

- **Appendix_MariyaHristova_11250_Feb2023_data_for_model.csv**

Data preparation and exploration

The files include Python Jupyter Notebooks, used for exploring the data, data cleaning and preparation steps.

- **Appendix_MariyaHristova_11250_Feb2023_data_prep.ipynb**

- **Appendix_MariyaHristova_11250_Feb2023_data_exploration.ipynb**

Feature selection and modelling

The files include Python Jupyter Notebooks, used for feature selection and statistical modelling.

- **Appendix_MariyaHristova_11250_Feb2023_feature_selection.ipynb**

- **Appendix_MariyaHristova_11250_Feb2023_modelling.ipynb**

Graphical analysis

A Jupyter notebook for creating the visualizations used in the thesis.

- **Appendix_MariyaHristova_11250_Feb2023_vizualization.ipynb**