

Multivariate Analysis

Level	Intermediate
Theme	Multivariate Analysis
Subthemes	Descriptive statistics, Construct Validity, Multivariate Regression Analysis, Data Visualization, Logistic Regression, Residuals, ANOVA

REPLICATED ARTICLE

Helbling, Marc. "Attitudes towards climate change migrants." *Climatic Change* 160.1 (2020): 89–102.

SECTION 1

Research Design

- 1** Summary of Arguments and Hypotheses
 Summarize the main arguments and all corresponding hypotheses. Critically evaluate whether these arguments provide sufficient depth to explain how the primary explanatory variables connect to the outcome under study and assess whether each hypothesis is directly aligned with the argument.
- 2** Measurement of Variables
 Discuss how the variables in the article are measured. To what extent are these measurements appropriately chosen to capture the underlying concepts in the argument?
- 3** Questionnaire Assessment
 Discuss any potential issues with the questionnaire used in the study, focusing on how well the questions and measures align with the arguments presented in the replication paper.

SECTION 2

Dataset

You will be working with the dataset `helbling_2020_class_replication.csv`. This dataset is essential for your analysis, and it is crucial to handle it properly to avoid any issues during your project.

Variable Label	Variable Name	Variable Levels
accept	Immigration acceptance	"0" - no, "1" - yes
reason	Reasons of migration	1=political, 2=economic, 3=droughts, 4=droughts/climate change, 5=sea level, 6=sea level/climate change
reason2	Reason of migration (alternative measure)	1=political, 2=economic, 3=droughts, 4=sea level

Variable Label	Variable Name	Variable Levels
muslim	Religion of migrants	1=muslim, 0=christian
higheducated	Migrant education	0=low educated, 1=high educated
agree	Agreement with arguments	1=fully yes, 2=partly yes, 3=partly no, 4=fully no
impact	Migration numbers	small=0, high=1
argu	Arguments	0=morality, 1=corrective justice, 2=capacity
eco	Index of respondent's environmentalism	1-4 (1=non-environmentalist, 4=environmentalist)
higheducation	Education level of the respondent	1=high education, 0=low education
age	Respondent's age (in years)	–
lire	Self-reported left–right position of the respondent	11-point scale 1=very right, 11=very left
left / right	–	1 - >6 on the 11-point lire scale, 0 - <6 on the 11-point lire scale. Vice versa for right variable
income	Respondent's self-reported monthly net income in Euro	1 = below €500 2 = €500 to just below €900 3 = €900 to just below €1300 4 = €1300 to just below €1500 5 = €1500 to just below €2000 6 = €2000 to just below €2600 7 = €2600 to just below €3500 8 = €3500 to just below €4500 9 = €4500 to just below €6000 10 = €6000 to just below €8000 11 = €8000 and more

The index is constructed using the following sub-questions: *v_6* – How often do you make a special effort to buy fruits and vegetables grown without pesticides or chemicals? *v_7* – How often do you cut back on driving a car for environmental reasons? *v_8* – How often do you sort glass or cans or plastic or paper and so on for recycling? *v_9* – How strong do you think the effect of climate change is on the environment? *v_10* – How strong do you think the effect of human beings' behaviour is on climate change?

1

Identify the Type of Data Collection

- Determine the data collection method used by the authors (e.g., survey, experiment, observational study).
- Assess the appropriateness of this method in relation to the research question. Consider whether it effectively captures the necessary information to address the study's objectives.

2 Analyze the Sample

- Identify the types of variables employed in the study, including independent and dependent variables, continuous (interval or ratio), or categorical (nominal, ordinal).
- Present the most relevant descriptive statistics for these variables in a table format, ensuring clarity and comprehensiveness.

3 Discuss Sample Differences and Sampling Issues

- Examine any differences within the sample that may impact the findings, such as demographic representation or response rates.
- Evaluate the accuracy and transparency of the article’s discussion regarding these sampling issues. Consider how well the authors address potential biases or limitations related to their sampling methods.

SECTION 3

Create a Multivariate Data Visualization

Feel free to get creative with your data visualization! The goal is to make something that is both accurate and informative, giving viewers a clear understanding of a key part of the article’s argument. Here are a couple of things to keep in mind:

1 Multivariate Focus

Your visualization should include multiple variables. This allows you to explore relationships and patterns that single-variable visuals just can’t capture.

2 Relevance to the Article

Make sure the visualization highlights a portion of the article’s argument. Think about what key insight you want to communicate or what story the data tells in relation to the authors’ findings.

3 Maintain Accuracy

While you can express yourself creatively, don’t let artistic flair compromise the accuracy of the data. Double-check that your visualization accurately represents the information.

4 Clarity is Key

Ensure that your visualization is easy to understand. Use clear labels, appropriate scales, and a legend if needed. A good title can help set the context.

5 Tell a Story

Think about how you can use colors, shapes, or other elements to enhance the narrative you want to convey. The best visualizations not only show data but also tell a compelling story.

SECTION 4

Analyses

Replicate tables 5 and 7 from the article (full tables are available in the Appendix). The tasks are as follows:

1 Regression Model Equations

Write down the regression model equation for one model from table 5 and one model from table 7.

2 Comparison of Findings

Present your findings alongside the author's findings in a single table for easy comparison across different models. Highlight any differences directly in the table.

3 Robustness Analyses

- In table 5 (model All), rerun the analyses using the original reason variable. Compare these results with the original findings in table 5 (All).
- Articulate an argument about how the age and income of respondents in the receiving country may affect their immigration acceptance. Add age and income variables to model All in table 5.
- Recode the 11-point income variable into a 5-point scale. Justify your choice of thresholds for the cut-off points. Rerun model All in table 5 with income as a 5-point scale and discuss any substantive differences in the effect of income on immigrant acceptance.

4 Interpretation of Findings

Interpret the findings from your replications and robustness analyses. Discuss whether you succeeded in replicating the selected results. Evaluate the external validity of the regression models in tables 5 and 7. Consider various reasons that may have affected your ability to replicate the study. Reflect on the robustness of the results to different specifications. Which results withstood robustness checks, and which did not? How were the substantive findings regarding climate change and immigration acceptance affected by the robustness checks?

5 Comparing ANOVA and OLS Regression

Using the eco (Environmentalism index) variable, construct a regression model predicting respondents' environmental stance as a function of income, education, age, and left-right ideology. Run the same model as an ANOVA. Briefly compare and interpret the findings. Plot the residuals and discuss whether any OLS assumption violations are present. Check if any ANOVA assumptions were violated as well.

SECTION 5

Final Replication Reflection

Identify which parts of the study you attempted to replicate and which remained unreplicated. Discuss whether you succeeded in replicating the selected results. Consider the various reasons that may have affected your ability to replicate the study. Reflect on the robustness of the results across different specifications. Which results withstood robustness checks, and which did not? What have you learned about the process of designing a quantitative study from your replication and robustness checks?

INSTRUCTOR'S SUPPLEMENT

In this assignment, students explore the concepts of exact replication—aiming to reproduce findings as they originally appear—and robustness testing, which involves examining findings under varied conditions or sample adjustments.

Learning Outcomes

- 1** **Conceptual Understanding**
Students will understand the differences between exact replication and robustness testing in empirical research.
- 2** **Statistical Application**
Students will apply regression analysis techniques to replicate and validate findings from existing research using statistical software (e.g., STATA).
- 3** **Data Preparation and Management**
Students will gain experience in preparing datasets, including constructing new variables from existing data through coding. The replication dataset provided may require different levels of preparation depending on students' familiarity with the material and the specific focus of the assignment. In this case, the dataset does not include the final variables; students will need to use provided code to construct these variables from existing ones. The original code is written in STATA. Instructors may choose to prepare a ready-to-use dataset based on the software used in class, or they may allow students to utilize the existing code to create the final variables and prepare the dataset for use.
- 4** **Comparative Analysis**
Students will effectively compare their analytical results with those of the original study, identifying and articulating any differences.
- 5** **Critical Thinking**
Students will analyze how demographic factors, such as age and income, influence immigration acceptance, enhancing their ability to formulate and support arguments with empirical evidence.
- 6** **Robustness Evaluation**
Students will evaluate the robustness of their findings through different model specifications and assess the impact of variable recoding on results.
- 7** **Interpretation and Reporting**
Students will interpret their findings clearly, discussing the implications for external validity and the factors that may have influenced replication success.
- 8** **OLS and ANOVA Assumptions**
Students will assess whether assumptions of ordinary least squares (OLS) and ANOVA models are met, enhancing their understanding of model fit.

9

Effective Communication

Students will communicate their methods, analyses, and findings clearly and concisely, both in written and tabular formats.