**README**

**for**

**“INTRINSIC INFORMATION PREFERENCES AND SKEWNESS”**

**by Yusufcan Masatlioglu, Yesim Orhun, Collin Raymond**

openicpsr-190641

**OVERVIEW**

The replication material in this repository includes the datasets and all necessary scripts to replicate the results in tables, figures and text of the main manuscript and the associated appendices. The code in this replication package constructs the results from 5 data sources using Stata. The package contains 30 do files, of which three are main files. The code was last run using Stata (version 16.0) on MacOS. The replicator should expect the code to run under 5 minutes on a standard (2023) desktop machine.

**DATA AVAILABILITY AND PROVENANCE STATEMENTS**

The data used in this paper has been collected by the authors and is available in the replication package. The data comes from three lab experiments, one online experiment and an online survey. The set of instructions, stimuli and sample selection criteria for these studies are included in the Appendix of the manuscript.

### Statement about Rights

I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

### Data Availability

All data used in the manuscript are publicly available. The experimental and survey data used to support the findings of this study have been deposited in the OpenICPSR repository (openicpsr-190641).

Dataset List

There are 5 input datasets, each corresponding to a study in the paper: AD\_input.dta, Exp1\_input.dta, Exp2\_input.dta, Exp3\_input.dta, IQ\_input.dta. These data are used to generate 5 final datasets used by the analysis codes: Exp1.dta, Exp2.dta, Exp3.dta, alzheimer.dta, IQdata.dta. The data generating codes are listed in sections below. A codebook for the data used in analyses is included at the end of this document. All data are also labeled in the datasets, and data preparation codes are annotated for convenience.

**COMPUTATIONAL REQUIREMENTS**

Software Requirements

* STATA 16
* Stata package grc1leg

Memory and Runtime Requirements

Approximate time needed to reproduce the analyses on a standard (2023) desktop machine is less than 5 minutes. The code was last run on a 12-core Intel-based desktop with MacOS version 13.4.

**DESCRIPTIONS OF CODE**

## The code folder has 30 do-files. The file 0\_setup.do will install additional packages used by the programs and allows the replicator to define the root path. It needs to be run only once. There are three main do files: 1\_data\_setup.do, 2\_main.do and 3\_appendix.do.

* The file “1\_data\_setup.do” runs all required programs that prepare the datasets for analysis. The input data and the programs “1\_data\_setup.do” calls, and the datasets it generates are listed below (organized by study name).

|  |  |
| --- | --- |
| **Data Preparation Code** | **Dataset Created** |
| Code/Exp1\_dataprep.do | Data/Output/Exp1.dta |
| Code/Exp2\_dataprep.do | Data/Output/Exp2.dta |
| Code/AD\_dataprep.do | Data/Output/alzheimer.dta |
| Code/IQ\_dataprep.do | Data/Output/IQdata.dta |
| Code/Exp3\_dataprep.do | Data/Output/Exp3.dta |

* The file “2\_main.do” calls all programs that generate all tables and figures in the main body of the article. It also generates the statistics mentioned in the text.
* The file “3\_appendix.do” will generate all tables and figures in the online appendix.

## INSTRUCTIONS TO REPLICATORS

* Edit 0\_setup.do to indicate the local pathways and run once on a new system to set up the working environment. Then, run 1\_data\_setup.do to generate final data used in analyses. These do files need to be run before 2\_main.do or 3\_appendix.do.
* Run 2\_main.do to generate results in the main manuscript.
* Run 3\_appendix.do to generate results in the appendix.

## LIST OF TABLES AND PROGRAMS

## The provided code reproduces all numbers provided in text in the paper and all tables and figures in the paper and appendix.

|  |  |  |
| --- | --- | --- |
| **Figure/Table #** | **Program** | **Line Number** |
| Figure 1 | doesn’t require computations | - |
| Figure 2 | doesn’t require computations | - |
| Figure 3a | Figure3.do | 49 |
| Figure 3b | Figure3.do | 51 |
| Figure 4a | Figure4.do | 88 |
| Figure 4b | Figure4.do | 47 |
| Figures A1-A9 | screenshot | - |
| Figure B1 | illustration | - |
| Figure E1 | FigureE1.do | 21-22 |
| Figure E2 | illustration | - |
| Table 1 | table1.do | 25-106 |
| Table 2 | table2.do | 16-61 |
| Table 3 | table3.do | 19-55 |
| Table 4 | table4.do | 38-54 |
| Table A1 | tableA1.do | 18-43 |
| Table A2 | tableA2.do | 18-174 |
| Table A3 | tableA3.do | 19-176 |
| Table B2 | tableB2.do | 20-47 |
| Table B3 | tableB3.do | 25-183 |
| Table B4 | tableB4.do | 60-134 |
| Table C1 | doesn’t require computations | - |
| Table C2 | tableC2.do | 9-63 |
| Table C3 | tableC3.do | 12-71 |
| Table C4 | tableC4.do | 12-117 |
| Table C5 | tableC5.do | 9-118 |

|  |  |  |
| --- | --- | --- |
| **In-text numbers** | **Program** | **Line Number** |
| p. 10, footnote 10 | Exp1\_text.do | 5 |
| p. 11, 1 | Exp1\_text.do | 12-17 |
| p. 11, 2 | Exp1\_text.do | 24-27 |
| p. 11, footnote 12 | Exp1\_text.do | 32-54 |
| p. 11, 3 | Exp1\_text.do | 63-71 |
| p. 11, 4 | Exp1\_text.do | 78-83 |
| p. 12 | Exp1\_text.do | 91-96 |
| p. 15, 1 | Exp2\_text.do | 10-12 |
| p. 15, footnote 16 | Exp2\_text.do | 22 |
| p. 15, 2 | Exp2\_text.do | 31-40 |
| p. 16, 1 | Exp2\_text.do | 46-47 |
| p. 16, 2 | Exp2\_text.do | 53 |
| p. 16, 3 | Exp2\_text.do | 58 |
| p. 18, 1 | AD\_text.do | 12-15 |
| p. 18, 2 | AD\_text.do | 21-28 |
| p. 18, 3 | AD\_text.do | 37-40 |
| p. 19, 1 | AD\_text.do | 49-52 |
| p. 19, 2 | AD\_text.do | 58-59 |
| p. 19, 3 | AD\_text.do | 66-67 |
| p. 30 | AD\_text.do | 71 |
| p. 21, 1 | IQ\_text.do | 14 |
| p. 21, 2 | IQ\_text.do | 19 |
| p. 21, 3 | IQ\_text.do | 27-31 |
| p. 21, 4 | IQ\_text.do | 37-42 |
| p. 21, 5 | IQ\_text.do | 64-69 |
| p. 22 | IQ\_text.do | 85-89 |
| Online Appendix p. 34 | IQ\_text.do | 94-95 |

**CODEBOOK**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name: Exp1.dta**  Description: Experiment 1 | | Obs: 700  Variables: 16 | |
| **Variable Name** | **Description** | | **Levels** |
| choicemajor | Choice Corresponds to the Option Preferred by Majority | | 0 (No) / 1 (Yes) binary |
| infostrength | Preference Strength for Chosen over Unchosen Option | | 0 – 10 numerical |
| Info[X] | Would switch choice for [X] cents, where X=1, 5, 10, 15, 20, 25, 30, 35, 40, 50 | | 0 (No) / 1 (Yes) binary |
| wta\_min | MCTS (min compensation to switch) from chosen option to rejected option | | Numerical (0.1 to 50.1) |
| infoprem | Information Premia. Takes on positive values when chosen option corresponds to the option chosen by the majority, negative otherwise. | | Numerical (-50.1 to 50.1) |
| wave | Experimental Period | | Summer 2025 (1) or Winter/Spring 2017 (2) |
| treatment | The choice problem presented to the participant | | T1 – T10 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Name: Exp2.dta**  Description: Experiment 2 | | Obs: 250  Variables: 26 | |
| **Variable Name** | **Description** | | **Levels** |
| early | Chose (1, 1) ≻ (0.5, 0.5) in Q1 | | 0 (No) / 1 (Yes) binary |
| early\_pref | Preference Strength for Chosen over Unchosen Option in Q1 | | Numerical (0 – 10) |
| pos\_extreme | Chose (0.5, 1) ≻ (1, 0.5) in Q2 | | 0 (No) / 1 (Yes) binary |
| pos\_extreme\_pref | Preference Strength for Chosen over Unchosen Option in Q2 | | Numerical (0 – 10) |
| pos\_slight | Chose (0.3, 0.9) ≻ (0.9, 0.3) in Q3 | | 0 (No) / 1 (Yes) binary |
| pos\_slight\_pref | Preference Strength for Chosen over Unchosen Option in Q3 | | Numerical (0 – 10) |
| pos\_inter | Chose (0.6, 0.9) ≻ (0.9, 0.6) in Q5a | | 0 (No) / 1 (Yes) binary |
| pos\_slight\_pref | Preference Strength for Chosen over Unchosen Option in Q5a | | Numerical (0 – 10) |
| abit\_early | Chose (.55, .55) ≻ (0.5, 0.5) in Q5b | | 0 (No) / 1 (Yes) binary |
| abit\_early\_pref | Preference Strength for Chosen over Unchosen Option in Q5b | | Numerical (0 – 10) |
| Q4A\_C1 | (Among Information non-Avoiders in Condition 1) Chose (0.3, 0.9) ≻ (0.76, 0.76) in Q4a | | 0 (No) / 1 (Yes) binary |
| Q4A\_C1\_pref | Preference Strength for Chosen over Unchosen Option in Q4a (Condition 1) | | Numerical (0 – 10) |
| Q4A\_C2 | (Among Information non-Avoiders in Condition 2) Chose (0.1, 0.95) ≻ (0.67, 0.67) in Q4a | | 0 (No) / 1 (Yes) binary |
|  | Preference Strength for Chosen over Unchosen Option in Q4a (Condition 2) | | Numerical (0 – 10) |
| Q4B\_C1 | (Among Information non-Avoiders in Condition 1) Chose (0.3, 0.9) ≻ (0.55, 0.55) in Q4b | | 0 (No) / 1 (Yes) binary |
| Q4B\_C1\_pref | Preference Strength for Chosen over Unchosen Option in Q4b (Condition 1) | | Numerical (0 – 10) |
| Q4B\_C2 | (Among Information non-Avoiders in Condition 2) Chose (0.5, 1) ≻ (0.66, 0.66) in Q4b | | 0 (No) / 1 (Yes) binary |
| Q4B\_C2\_pref | Preference Strength for Chosen over Unchosen Option in Q4b (Condition 2) | | Numerical (0 – 10) |
| monot | "Indicator for preference ordering that reflects a consistent preference for informativeness across Q1 and Q4 answers" | | 0 (No) / 1 (Yes) binary |
| monot\_avoid | "(Among Information Avoiders) Indicator for preference ordering that reflects a consistent preference for informativeness across Q1 and Q4 answers" | | 0 (No) / 1 (Yes) binary |
| monot\_taker | "(Among Information Non-Avoiders) Indicator for preference ordering that reflects a consistent preference for informativeness across Q1 and Q4 answers" | | 0 (No) / 1 (Yes) binary |
| condition | The set of questions presented to the participant Condition 1.  Q1. Option 1: (1, 1) vs. Option 2: (.5, .5)  Q2. Option 1: (1, .5) vs. Option 2: (.5, 1)  Q3. Option 1: (.9, .3) vs. Option 2: (.3, .9)  Q5a. Option 1 (.9, .6) vs. Option 2 (.6, .9)  Q5b. Option 1: (.55, .55) vs. Option 2: (.5, .5)  Condition 2.  Q1. Option 1: (.5, .5) vs. Option 2: (1, 1)  Q2. Option 1: (.5, 1) vs. Option 2: (1, .5)  Q3. Option 1 (.9, .6) vs. Option 2 (.6, .9)  Q5a. Option 1: (.9, .3) vs. Option 2: (.3, .9)  Q5b. Option 1: (.5, .5) vs. Option 2: (.55, .55) | | 1 or 2 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Name: Exp3.dta**  Description: Experiment 3 (Appendix only) | | Obs: 232  Variables: 15 | |
| **Variable Name** | **Description** | | **Levels** |
| majorchoice | Choice Corresponds to the Option Preferred by Majority | | 0 (No) / 1 (Yes) binary |
| infostrength | Preference Strength for Chosen over Unchosen Option | | 0 – 10 numerical |
| Info[X] | Would switch choice for [X] cents, where X=1, 5, 10, 15, 20, 25, 30, 35, 40, 50 | | 1 (No) / 2 (Yes) binary |
| prior | Treatment: Prior Level | | 10 or 90 |
| school | Experimental Location | | Amherst / UM |
| condition | The choice problem presented to the participant | | C1, C2, C3 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Name: alzheimer.dta**  Description: Alzheimer’s Disease Study | | Obs: 626  Variables: 43 | |
| **Variable Name** | **Description** | | **Levels** |
| gender | Gender | | 0 "male" 1 "female" |
| age | Age | | Numerical |
| age\_death | Expected age at death | | Numerical |
| risk\_learn | Chose to get the test to learn if has the risky allele or not | | 0 (No) / 1 (Yes) binary |
| safe\_learn | Chose to get the test to learn if has the safe allele or not | | 0 (No) / 1 (Yes) binary |
| exact\_learn | Chose to get the test to learn the exact combination of alleles | | 0 (No) / 1 (Yes) binary |
| risk\_pay0 | Would get the test for the risky allele if it were free | | 0 (No) / 1 (Yes) binary |
| risk\_paid[$X]  (X=5, 10, 15, 25, 25, 50) | Would get the test for the risky allele if were paid $X | | 0 (No) / 1 (Yes) binary |
| risk\_pay[$X]  (X=5, 10, 15, 25, 25, 50) | Would pay $X to get the test for the risky allele | | 0 (No) / 1 (Yes) binary |
| safe\_pay0 | Would get the test for the safe allele if it were free | | 0 (No) / 1 (Yes) binary |
| safe\_paid[$X]  (X=5, 10, 15, 25, 25, 50) | Would get the test for the safe allele if were paid $X | | 0 (No) / 1 (Yes) binary |
| safe\_pay[$X]  (X=5, 10, 15, 25, 25, 50) | Would pay $X to get the test for the safe allele | | 0 (No) / 1 (Yes) binary |
| exact\_pay0 | Would get the test for the exact combination of alleles if it were free | | 0 (No) / 1 (Yes) binary |
| exact \_paid[$X]  (X=5, 10, 15, 25, 25, 50) | Would get the test for the exact combination of alleles if were paid $X | | 0 (No) / 1 (Yes) binary |
| exact \_pay[$X]  (X=5, 10, 15, 25, 25, 50) | Would pay $X to get the test for the exact combination of alleles | | 0 (No) / 1 (Yes) binary |
| risk\_wtp | Maximum willingness to pay for the risky allele test | | Numerical |
| safe\_wtp | Maximum willingness to pay for the safe allele test | | Numerical |
| exact\_wtp | Maximum willingness to pay for test for the exact combination of alleles | | Numerical |
| no\_switching | Indicates that the participant only switched once (as expected) in the list-price elicitation method | | 0 (No) / 1 (Yes) binary |

|  |  |  |  |
| --- | --- | --- | --- |
| **Name: IQdata.dta**  Description: IQ Test Experiment | | Obs: 600  Variables: 11 | |
| **Variable Name** | **Description** | | **Levels** |
| gender | Gender | | 1 "female" 2 "male" |
| age | Age | | Numerical |
| education | Highest education level attained | | 1 "Less than high school" 2 "High school graduate" 3 "Some college" 4 "2 year degree" 5 "4 year degree" 6 "Professional or Master's degree" 7 "Doctorate" |
| No\_info | Rank of the “No Information” Option among 4 Options | | 1—4 numerical, 1 “top rank”, 4 “lowest rank” |
| certain\_info | Rank of the “Most Informative” Option among 4 Options | | 1—4 numerical, 1 “top rank”, 4 “lowest rank” |
| pos\_skew | Rank of the “Positively Skewed” Option among 4 Options | | 1—4 numerical, 1 “top rank”, 4 “lowest rank” |
| neg\_skew | Rank of the “Negatively Skewed” Option among 4 Options | | 1—4 numerical, 1 “top rank”, 4 “lowest rank” |
| full | Prefer Most Info > No Info | | 0 (No) / 1 (Yes) binary |
| pos | Prefer Pos Skew > No Info | | 0 (No) / 1 (Yes) binary |
| neg | Prefer Neg Skew > No Info | | 0 (No) / 1 (Yes) binary |
| avoid | Prefer No Info > Most Info | | 0 (No) / 1 (Yes) binary |