Replication Package for

Biases in Information Selection and Processing: Survey Evidence from the Pandemic,"

by Ester Faia, Andreas Fuster, Vincenzo Pezone, and Basit Zafar

This package includes all the datasets and codes needed for the replication of the results presented in "Biases in Information Selection and Processing: Survey Evidence from the Pandemic," by Ester Faia, Andreas Fuster, Vincenzo Pezone, and Basit Zafar. All the analyses are executed using the software Stata.

The folder "replication_package" includes all the raw data needed for the analysis. They are the following:

- full_survey_1.xlsx and full_survey_2.xlsx are the raw survey data downloaded from the Qualtrics platform. The first row of each file has the question identifier and the second row reports the actual question asked to participants. The rest of the file includes the actual answers. To ease the import of the data into Stata we manually eliminate the second row of the files and work with the Excel files survey_1.xlsx and survey_2.xlsx, also included in the folder;
- bls_ur.xlsx is the county-level unemployment rate downloaded from the Bureau of Labor Statistics website (see https://www.bls.gov/lau/);
- zipTOcounty.dta is a crosswalk linking ZIP codes to counties (see the replication package of Abel and Fuster, AEJ 2021, available at the link https://www.openicpsr.org/openicpsr/project/116461/version/V1/view);
- infections_population_alldates.dta has county-level COVID-related infections and deaths made freely available by The New York Times (https://github.com/nytimes/covid-19-data) and imported in Stata.

The folder also includes three additional subfolders. The folder "codes" has the Stata do-files needed to replicate the analysis in the main text. The user should add a suitable directory at the beginning of each do-file before running them. It includes the following do-files:

- cleaning_survey_1.do and cleaning_survey_2.do import the raw survey answers and generate the two datasets data_wave1.dta and data_wave2.dta;
- tables.do produces tables 1 through 7 (in .tex format);
- figures.do produces figures 1 and 2 (in .png format);
- main.do runs the four do-files listed above in the correct order.

The subfolders "tables" and "figures" store tables and figures, respectively.

The data dictionary that follows presents definitions for the main variables used in the paper, as well as the corresponding names in the do-files.

Variable Name (Data)	Variable Name (text)	Variable Definition
worried_health	Worried - Health	A measure, on a 1 to 10 scale, of how worried the participant is about the health consequences of the pandemic.
worried_economy	Worried - Economy	A measure, on a 1 to 10 scale, of how worried the participant is about the economic consequences of the pandemic.
worried	Worried H+E	The sum of Worried - Economy and Worried - Health.
worried_h_e	Worried H-E	The difference between Worried - Health and Worried - Economy.
pc_10	PC1	The first principal component of Worried - Health, Worried - Economy, Pr. Deaths > 200k, Pr. UR > 20%, Forecast Deaths, Forecast UR, and Pr. Infection (all measured in prior stage).
support_sofar	Lockdown Support	An evaluation of lockdown measures taken so far, from 1("The measures went much too far") to 5 ("The measures did not go nearly far enough").
pref_seen	Preferred	Dummy equal to 1 if the participant reads the article she had ranked as preferred (Survey 1).
pess, p_h, p_e	Pessimistic	Dummy equal to 1 if the participant reads the article with the pessimistic headline. In Survey 1, when individuals who chose the health or the economy domain are pooled, we distinguish between a "Pessimistic - Health" (p_h) and a "Pessimistic - Economy" (p_e) dummy.
pess_pref	Preferred × Pessimistic	Interaction term between Preferred and Pessimistic
p_deaths_prior, p_deaths_posterior, p_deaths_rev	Pr. Deaths > 200k	The participant's estimate of the probability that the number of Covid-related deaths will be higher than 200,000 by the end of the year (prior, posterior, and revision).
p_ue_prior, p_ue_posterior, p_ue_rev	Pr. UR > 20%	The participant's estimate of the probability that the unemployment rate will be higher than 20%. by the end of the year (prior, posterior, and revision)..
p_inf_prior, p_inf_posterior, p_inf_rev	Pr. Infection	An estimate of the probability that the participant will be infected by the coronavirus by the end of the year (prior, posterior, and revision).
revealed	Revealed Before	A dummy equal to 1 if the participant is revealed the source of each article before making the choice (Survey 2).
rev_after	Revealed After	A dummy equal to 1 if the participant is revealed the article source after having made the choice (Survey 2).
pess_rev	Pessimistic × Revealed After	Interaction term between Pessimistic and Revealed After
pess_lib	Pessimistic × Liberal	Interaction term between Liberal and Pessimistic.

lib_rev	Liberal × Revealed After	Interaction term between Liberal and Revealed After.
lib_rev_bef	Liberal × Revealed Before	Interaction term between Liberal and Revealed Before.
lib_rev_pess	Liberal × Revealed After × Pessimistic	Interaction term between Liberal, Pessimistic, and Revealed After.
worried_revealed, pc_revealed, support_sofar_revealed	Worried × Revealed Before, PC × Revealed Before, Lockdown Support × Revealed Before	Interaction term between Revealed Before and Worried, PC, and Lockwdown Support, respectively.
sign_up	Pr. Contact Tracing (CT) App	The likelihood that the participant will voluntarily sign up for a contact tracing app (prior, posterior, and revision).
informativeness	Informativeness	An assessment of the informativeness of the article read on a scale from 1 to 7.
reliability	Reliability	An assessment of the reliability of the article read on a scale from 1 to 7.
headline_ok	Headline Check	A dummy equal to 1 if the participant correctly identifies, among four headlines, the one corresponding to the article just read.
money_spent	Amount Spent for Advertising	The dollar amount that the participant is willing to spend to advertise the news source on Facebook, from a $50 budget.
support	Support	A measure of support for future lockdown policies from 1 ("No intervention, just let the virus run its course") to 5 ("Stricter stay-home orders than this time, even if the economic cost is very large").
donation_who	Money to WHO	The dollar amount the participant is willing to donate to the World Health Organization, as opposed to FreedomWorks, out of a budget of 20 dollars (Survey 1).
college	College	A dummy equal to 1 if the participant has a bachelor, master, or doctoral degree.
retired	Retired	A dummy equal to 1 if the participant is a retiree or an early retiree.
unemployed	Unemployed	A dummy equal to 1 if the participant is unemployed.
employment_loss_hh	Employment Loss in HH	A dummy equal to 1 if somebody in the participant's household has suffered an employment loss related to the pandemic.
income_scale	Income	A measure of household income that takes values 0 ("Less than $10,000"), 1 ("$10,000 - $19,999"), 2 ("$20,000 - $29,999"), 3 ("$30,000 - $39,999"), 4 ("$40,000 - $49,999"), 5 ("$50,000 - $59,999"), 6 ("$60,000 - $69,999"), 7 ("$70,000 - $79,999"), 8 ("$80,000 - $89,999"), 9 ("$90,000 - $99,999"), 10 ("$100,000 - $149,999"), 11 ("$150,000 - $199,999"), 12 ("$200,000 or more").
health	Health	Self-assessed overall health, from 1 ("poor") to 5 ("excellent").
health_conditions	Nr Health Conditions	The number of conditions the participant suffers or has suffered from, between (i) Lung disease such as chronic bronchitis or emphysema; (ii) Diabetes or high blood sugar; and (iii) A heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems. The list of conditions is from Alcott et al (2020).

already_infected	Already Infected	A dummy equal to 1 if the participant indicates that she has been tested positive for COVID-19 or that she is very likely to have been infected.
liberal	Liberal	A dummy equal to 1 if the participant defines herself as "very liberal/democrat," "liberal/democrat," or "leaning liberal/democrat," and 0 otherwise.
age_c	Age	The participant's age.
n_old	Nr People above 65 in HH	Number of individuals older than 65 in the household.
female	Female	A dummy equal to 1 if the participant is a woman.
married	Married	A dummy equal to 1 if the participant is married.
race_black, race_asian, hispanic	Black, Asian, Hispanic	Dummies corresponding to the participant's race/ethnicity.
own_house	Own Primary Residence	A dummy equal to 1 if the participant owns her primary residence.
stock	Hold Stocks	A dummy equal to 1 if the participant owns stocks.
health_insurance	Health Insurance	A dummy equal to 1 if the participant owns a health insurance.
risk_loving	Willingness to Take Risks	Self-reported willingness to take risks, from 1 ("Absolutely unwilling to take risks") to 7 ("Fully prepared to take risks").
discount	Willingness to Wait	Self-reported willingness to give up something that is beneficial today in order to benefit more from that in the future, from 1 ("Completely unwilling to give up") to 7 ("Very willing to give up").
trust	Trust in People	Self-reported trust in other people, from 1 ("Need to be very careful") to 7 ("Most people can be trusted").
severity_lockdown	Stay at Home Order	A dummy equal to 1 if there is a stay at home order in the area where the participant lives.
cases_pc	Cases Per Capita	Number of Covid cases per capita multiplied by 1 million, measured the day before the participant takes the survey.
deaths_pc	Deaths Per Capita	Number of Covid-related deaths per capita multiplied by 1 million, measured the day before the participant takes the survey.
ur10	County UR	Unemployment rate at the county level as of April, 2020, divided by 10.
health_order, econ_order	Order fixed effects	Dummy equal to 1 if the participant is shown the headline in the health/economy domain first.
date	Date fixed effects	Date in which the survey is taken.
region	Region fixed effects	Macro-region where the participant is taking the survey from.
p_lib_cons	Surprise	sum of the percent chance that the individual assigns to the article being from CNN, The New York Times, MSNBC, and The Washington Post, minus the sum of the percent chance that the individual assigns to the article being from Breitbart, Fox News, and The Wall Street Journal.
no3		Dummy equal to 1 if participants who are shown the article source before making the choice are excluded.
sample_10		Dummy equal to 1 if participants who are in the bottom 10% in terms of the time spent on the questions prior to reading the article are excluded.