Hobo Economicus Replication Package README

Overview

The instructions, data, and code in this replication package generate all analyses required to reproduce the results presented in the text, tables, and figures of the manuscript using Stata. The replicator may expect the code to run for less than 5 minutes.

Data Availability and Provenance

- ☐ This paper does not involve analysis of external data (i.e., no data are used or the only data are generated by the authors via simulation in their code).

Statement about Rights

- ☑ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.
- ☑ I certify that the author(s) of the manuscript have permission to provide the data contained within this replication package.

Summary of Availability

- ☑ All data **are** publicly available.
- ☐ Some data **cannot be made** publicly available.
- ☐ **No data can be made** publicly available.

Data on number of panhandlers at each metro station visit

Data on the number of panhandlers at each station visit were collected by the authors. See Section 3 of the paper for a detailed description of the data collection procedure and construction of variables. The data will be available in The Economic Journal repository.

Datafile: `metrovisitslog.csv`

Data on friendliness of passersby at each metro station

Data on the friendliness of passersby at each station were collected by the authors. See Section 3 of the paper for a detailed description of the data collection procedure and construction of variables. The data will be available in The Economic Journal repository.

Datafile: `friendliness.csv`

Data on panhandling receipts

Data on panhandling receipts at each station were collected by the authors. See Section 3 of the paper for a detailed description of the data collection procedure and construction of variables. The data will be available in The Economic Journal repository.

Datafile: `receipts.csv`

Data on number of passersby for each metro station visit

Data on the number of passersby is measured by total number of metro riders exiting a given station each month. These data are publicly available from the Washington Metropolitan Area Transit Authority (WMATA) through a Public Access to Records Policy (PARP) request. See WMATA's Public Access to Records and Privacy Policies for further details. For annual data on the number of passersby, visit their Ridership Data Portal.

We submitted our request for the monthly number of riders exiting each station for our sample stations and sample months by email to PARP@wmata.com. See Section 3 of the paper for a detailed description of the metro stations and months/years included in our sample.

Datafile: `passersby.csv`

Data on metro station characteristics

Data on the presence of a homeless service provider near each metro station were collected using Google Maps. The datafile includes for each station the address of the nearest homeless service provider and the walking distance in minutes to this service (as of November 5, 2021).

Data on the presence of a homeless shuttle stop near each metro station were collected using DC Human Services and Google Maps. Open Data DC regularly updates and publicly provides the scheduled homeless shuttle routes and stops, along with other data in the DC Data Catalog. The datafile includes for each station the address of the nearest homeless shuttle stop and the walking distance in minutes to this shuttle stop (as of November 5, 2021).

See Section 3 of the paper for a detailed description of the construction of variables using these data.

Datafile: `service.csv`, `shuttle.csv`

Code for data cleaning and analysis

Code for data cleaning and analysis is included in the replication package. It will be available in The Economic Journal repository.

Data and Codebook Files

Data file	Source	Provided
`data\metrovisitslog.csv`	Authors	Yes
`data\friendliness.csv`	Authors	Yes
`data\receipts.csv`	Authors	Yes
`data\passersby.csv`	WMATA	Yes
`data\service.csv`	DC Human Services, Google Maps	Yes
`data\shuttle.csv`	Google Maps	Yes
`data\codebook.csv`	Authors	Yes

Computational requirements

Software Requirements

Stata: code was last run with version 16 on November 6, 2021

- `labmask` (requires Stata version 7, package updated November 23, 2013)
- `outreg` (requires Stata version 10.1, package updated September 18, 2015)
- `stripplot` (requires Stata version 8.2, package updated July 11, 2021)

The program "0_setup.do" will setup the environment and should be run once.

Memory and Runtime Requirements

Approximate time needed to reproduce the analyses on a standard 2021 desktop machine:

- ☑ <10 minutes
- ☐ 10-60 minutes
- ☐ 1-8 hours
- ☐ 8-24 hours
- ☐ 1-3 days
- ☐ 3-14 days
- ☐ > 14 days
- ☐ Not feasible to run on a desktop machine, as described below.

The code was last run on an Intel Core i7 laptop with 16GB of RAM, 1 TB of fast local storage. Computation took 2 minutes.

Description of code

- Code in `code\1_dataprep.do` will reformat/combine datasets and generate all variables required for the analysis.
- Code in `code\2_analysis.do` will generate all results presented in the tables, figures, and main body of the article. Where applicable, output files are called appropriate names (`table5.txt`, `figure1a.png`) that easily correlate with the manuscript.

Instructions to Replicators

Follow steps below in sequence:

1. Download the zipped folder "replicate." Extract entire compressed folder "replicate" into chosen subdirectory.
2. Open do-file `code\0_setup.do` in Stata. Manually edit `code\0_setup.do` to adjust the default path
3. Run `code\0_setup.do` once on a new system to set up the working environment
4. Run `code\1_dataprep.do`
5. Run `code\2_analysis.do`

List of analysis results and code files

The provided code reproduces:

- ☑All numbers provided in text in the paper
- ☑All tables and figures in the paper
- ☐ Selected tables and figures in the paper, as explained and justified below.

Figure/Table #	Replication file	Line Number	Output file
Table 1	`code\2_analysis.do`	3	table1.txt
Table 2	`code\2_analysis.do`	21	table2.doc
Table 3	`code\2_analysis.do`	33	table3.txt
Table 4	`code\2_analysis.do`	50	table4.txt
Table 5	`code\2_analysis.do`	64	table5.txt
Figure 1, Panel A	`code\2_analysis.do`	81	figure1a.png
Figure 1, Panel B	`code\2_analysis.do`	88	figure1b.png
In-Text	`code\2_analysis.do`	94	intext.txt

References

Cox, N. (2003). 'STRIPPLOT: Stata module for strip plots (one-way dot plots).' *Statistical Software Components* S433401, Boston College Department of Economics, revised 11 July 2021.

Cox, N. (2008). 'Speaking Stata: Between tables and graphs.' *The Stata Journal*, vol. 8(2), pp. 269-289.

DC Data Catalog (2018). *Homeless Facility Shuttle Stops*. http://data.dc.gov/ (accessed: 16 July 2018).

Gallup. J. (1999). 'OUTREG: Stata module to write estimation tables to a Word or TeX file.' *Statistical Software Components* S375201, Boston College Department of Economics, revised 18 September 2015.

WMATA (2016-2019). Washington Metropolitan Area Transit Authority, https://www.wmata.com/about/records/ (accessed: October 2016 to April 2019).

Acknowledgements

This ReadMe file follows the template from Social Science Data Editors, prepared by Lars Vilhuber, Miklos Kóren, Joan Llull, Marie Connolly, Peter Morrow.