README – Replication package for Blow and Crawford, "Nonparametric Analysis of Reference-Dependent Preferences"

Oct 22, 2026 · @Laura

Overview

This package reproduces every table, figure and computed number in the main paper and Supplemental Appendix of Laura Blow and Vincent P. Crawford, "Nonparametric Analysis of Reference-Dependent Preferences", from the raw data. All code is in R. One file runs everything: `code/00_run_all.R`.

- **Default run (shipped results): under a minute.** The package ships the estimation results and the Table A5 sweep results, so the default run rebuilds every exhibit from them and checks key numbers against the printed paper.
- **Full run from scratch: about 3 days** on a standard laptop: about 37 hours for the estimation and about 37 hours for the Table A5 sweep. Both can be interrupted and resumed.

Figures 1-5 and A1-A2 are theoretical schematics drawn by hand; no data or code produces them.

Data availability and provenance

The analysis uses two raw data files, both publicly available and both included in `data_raw/`. The taxi data are posted in the American Economic Association's replication archives (Farber 2008a; Crawford and Meng 2011a), and we include a copy with attribution. The CPI data are published by the US Bureau of Labor Statistics and are in the public domain.

File	Contents	Source	Format	Included
`taxi015.dta` (and `taxi015.csv`)	Trip-level records for 21 New York City taxi drivers	Henry S. Farber; replication archives of Farber (2008) and Crawford and Meng (2011)	Stata, 4.4 MB; CSV copy	Yes
`CPI.NY.NJ.PA.Urban.All.Items.csv`	Monthly consumer price index, 1999-2001	US Bureau of Labor Statistics, series CUURS12ASA0	CSV	Yes

Taxi data. Henry Farber's trip sheets for 21 drivers, the sample of Farber (2005). The dataset is publicly posted in Farber's (2008) replication package and again in Crawford and Meng's (2011), both now distributed through openICPSR (DOIs in References). `taxi015.dta` is the authors' copy of this dataset. The code aggregates trips to driver-days and uses the variables `id`, `date`, `fare`, `tripmin`, `brmin`, `waitmin`, `rainfall`, `snowfall`, `shift`, `mintemp` and `maxtemp`. `taxi015.csv` is a plain-text copy of `taxi015.dta`, provided because the Stata format is proprietary; the code reads the `.dta` file.

CPI data. Consumer Price Index for All Urban Consumers (CPI-U), all items, New York-Newark-Jersey City, NY-NJ-PA, monthly, not seasonally adjusted (BLS series CUURS12ASA0), months from 1999 to 2001, downloaded from https://data.bls.gov/timeseries/CUURS12ASA0 (file dated June 2018). The code divides each driver-day's earnings by the index for that month, normalised to the sixth month in the file.

Computational requirements

Software. R 4.3.1 on Windows 11 (64-bit), with these packages (versions used):

Package	Version	Used for
lpSolve	5.6.19	the linear programs (estimation, Table A5)
foreign	0.8-85	reading `taxi015.dta`
lubridate	1.9.3	dates (merging the monthly CPI)
fmsb	0.7.5	radar plots (Figures 6-9)

Install them once with `install.packages(c("lpSolve", "foreign", "lubridate", "fmsb"))`. `00_run_all.R` stops with this instruction if any is missing, and records the session details in `output/sessionInfo.txt`.

`03_figures_6_9.R` draws with `bmp(type = "windows")`, which exists only on Windows. On macOS or Linux, change it to `type = "cairo"`.

Hardware used. A laptop with an Intel Core i7-1165G7 (2.80 GHz, 4 cores, 8 threads) and 16 GB of RAM. The package needs under 100 MB of disk space.

Run times on that laptop:

Step	Time
Default run: scripts 02-09 using the shipped results	under a minute
`01_estimation.R` from scratch	about 37 hours (Hours and Earnings about 20 hours, Earnings only and Hours only about 8 hours each)
`08_table_A5.R` from scratch	about 37 hours (the five grid settings take 2-18 hours each)

Description of the code and folders

The package has three folders. `00_run_all.R` runs the scripts in `code/` in numbered order; each script can also be run on its own after setting the path at its top.

```
Replication package/
  README.pdf
  code/          00_run_all.R and scripts 01-10
  data_raw/      taxi015.dta, taxi015.csv, CPI.NY.NJ.PA.Urban.All.Items.csv
  output/
    data/        estimation results and tables (.Rdata, .csv)
      tuning/    Table A5 sweep results
    pics/        figures (.bmp)
    run_all_log.txt, sessionInfo.txt   (written by each run)
```

Script	What it does	Self-check
`00_run_all.R`	Master: checks the folder and packages, runs 01-10, writes the log and session details	
`01_estimation.R`	All estimation: for each driver and model, the verdict on the actual choices, the proximity and the area rate (1,000 random datasets), with and without additive separability. Runs only if `RUN_ESTIMATION = TRUE`	
`02_table_A1.R`	Table A1: descriptive statistics and Afriat efficiency index by driver	against the printed table (168 cells)
`03_figures_6_9.R`	Figures 6-9: radar plots of proximities and Selten measures	
`04_figures_A4_A13.R`	Figures A4-A13: empirical CDFs of proximities and Selten measures	
`05_tables_A2_A3.R`	Tables A2 and A3: proximities and Selten measures by driver	
`06_table_A4_figures_A14_A15.R`	Table A4 (paired t-tests) and Figures A14-A15	

Script	What it does	Self-check
`07_sections_5_5_5_6.R`	Numbers in Supplemental Appendix Sections 5.5 (unresolved linear programs) and 5.6 (linear program vs exact WARP test)	Section 5.6 counts against the printed ones
`08_table_A5.R`	Table A5: recovery of WARP-consistent draws by grid setting. Runs only if `RUN_TABLE_A5 = TRUE` (the default)	Section 5.6 setting against the printed counts
`09_driver_8.R`	The check behind the explanation of driver 8 in Section 7.5: which days each model uses, and WARP on those days	
`10_figure_A3.R`	Figure A3: each driver's budget lines and chosen bundles, from the raw data	

Shipped results. `output/data/` contains the estimation results, one file per dimension of reference dependence: `Areas (Both) ...`, `Areas (Earnings) ...` and `Areas (Hours) ...`, each ending `NestedGeneralU.intervals_24_D1000_LPt20_scaled_slope1e-04_final.Rdata`. `output/data/tuning/` contains the Table A5 sweep, one file per grid setting plus the WARP verdicts. These are what the paper reports. They let the default run finish in under a minute; the from-scratch runs regenerate them.

Instructions to replicators

Default run (under a minute).

1. Install R and the four packages listed under Computational requirements.
2. Unzip the package. Open `code/00_run_all.R` and set `project_root`, near the top, to the folder that contains `code/`, `data_raw/` and `output/`. Use forward slashes, even on Windows.
3. Run the whole file (in RStudio: Source). Leave the switches at their defaults: `RUN_ESTIMATION <- FALSE`, `RUN_TABLE_A5 <- TRUE`.

4. Figures are written to `output/pics/` and tables to `output/data/`. Everything printed is also saved in `output/run_all_log.txt`. The self-checks print, among other lines, "ALL 168 CELLS MATCH the printed Table A1" and "Section 5.6 reproduced exactly".

Full run from scratch (about 3 days).

1. Move the three shipped `Areas (...)` files out of `output/data/` and move the `output/data/tuning/` folder out of the package, for example into a backup folder.
2. In `00_run_all.R`, set `RUN_ESTIMATION <- TRUE` (and keep `RUN_TABLE_A5 <- TRUE`).
3. Run the whole file. It runs the estimation (about 37 hours), then the Table A5 sweep (about 37 hours), then all the exhibits.

If a run is interrupted (a forced restart, a flat battery), run `00_run_all.R` again unchanged. The estimation and the Table A5 sweep both save their results after every driver and resume from the last finished driver; the driver in progress starts again. Each restart writes a fresh `run_all_log.txt`. Keep the computer plugged in and set not to sleep.

List of exhibits and the code that produces them

Every exhibit and computed number in the paper maps to one script. Figures are in `output/pics/`, tables in `output/data/`; numbers quoted in the text are also printed in `output/run_all_log.txt`. `[tag]` stands for `intervals_24_D1000_LPt20_scaled_slope1e-04_final`.

Main paper

Exhibit	Script	Output file
Figures 1-5	none (hand-drawn schematics)	
Figure 6	`03_figures_6_9.R`	`ProximityRadarPlotExpectations_NestedGeneralU.bmp`
Figure 7	`03_figures_6_9.R`	`SeltenMeasuresRadarPlotExpectations_NestedGeneralU.bmp`
Figure 8	`03_figures_6_9.R`	`ProximityRadarPlotExperience_NestedGeneralU.bmp`
Figure 9	`03_figures_6_9.R`	`SeltenMeasuresRadarPlotExperience_NestedGeneralU.bmp`

Exhibit	Script	Output file
Section 7.5: driver-by-driver comparisons	`05_tables_A2_A3.R`	`TablesA2_A3_experience.[tag].Rdata`
Section 7.5: mean differences and p-values	`06_table_A4_figures_A14_A15.R`	`TableA4_paired_tests.[tag].Rdata`
Section 7.5: explanation of driver 8	`09_driver_8.R`	printed in the log

Supplemental Appendix

Exhibit	Script	Output file
Figures A1-A2	none (hand-drawn schematics)	
Figure A3	`10_figure_A3.R`	`FigureA3_budget_lines.bmp`
Table A1	`02_table_A1.R`	`Table_A1_descriptives.csv` (and `.Rdata`)
Tables A2-A3	`05_tables_A2_A3.R`	`TablesA2_A3_experience.[tag].Rdata`
Table A4	`06_table_A4_figures_A14_A15.R`	`TableA4_paired_tests.[tag].Rdata`
Figure A4	`04_figures_A4_A13.R`	`ProximityNeoclassical.dist.bmp`
Figure A5	`04_figures_A4_A13.R`	`SeltenMeasuresNeoclassical.dist.bmp`
Figure A6	`04_figures_A4_A13.R`	`ProximityRefDep.dist.bmp`
Figure A7	`04_figures_A4_A13.R`	`SeltenMeasuresRefDep.dist.bmp`
Figure A8	`04_figures_A4_A13.R`	`ProximityRefDepModelKinds.dist.bmp`
Figure A9	`04_figures_A4_A13.R`	`SeltenMeasuresRefDepModelKinds.dist.bmp`
Figure A10	`04_figures_A4_A13.R`	`ProximityRefDepModelForms.dist.bmp`
Figure A11	`04_figures_A4_A13.R`	`SeltenMeasuresRefDepModelForms.dist.bmp`
Figure A12	`04_figures_A4_A13.R`	`Neoclassical.v.RefDepProximity.dist.bmp`
Figure A13	`04_figures_A4_A13.R`	`Neoclassical.v.RefDepSeltenMeasures.dist.bmp`
Figure A14	`06_table_A4_figures_A14_A15.R`	`FigureA14_expectations_differences.bmp`

Exhibit	Script	Output file
Figure A15	`06_table_A4_figures_A14_A15.R`	`FigureA15_experience_differences.bmp`
Section 5.5: unresolved linear programs	`07_sections_5_5_5_6.R`	`unresolved_bound_summary.csv`
Section 5.6: linear program vs WARP	`07_sections_5_5_5_6.R`	`section5_6_validation_counts.csv`
Table A5	`08_table_A5.R`	`tuning/Table_A5.Rdata`

Notes on reproducibility

The random datasets are generated once, from a fixed seed (`set.seed(731969)`), so every run uses the same 1,000 datasets per driver.

Re-running the estimation from scratch reproduces exactly all WARP results, the verdicts on the drivers' actual choices, the proximities (Table A2), Table A1, Table A5 and the Section 5.6 counts. Results that depend on area rates (the Selten measures in Table A3 and the figures that plot them, and the paired tests in Table A4) can differ very slightly. The cause is the linear-program solver (lpSolve 5.6.19). On a small number of numerically ill-conditioned random datasets it stops without a verdict, mostly with a numerical failure (status 5), and the code records the dataset as unresolved rather than as a pass. Whether a given borderline program fails in this way is not perfectly repeatable: it can vary between runs, machines, operating systems and package versions. Raising the solver's time limit does not help, because these failures occur within seconds, not at the limit.

When a dataset that is unresolved in one run is resolved in another and passes, the area rate for that driver and model changes by 0.001 (one dataset in 1,000), and the Selten measure by the same amount. A from-scratch rerun on a different machine may therefore change a few Selten measures in Table A3 by 0.001 and the paired tests in Table A4 in the fourth decimal place, with no change in sign or significance. The number of unresolved solves in Section 5.5 also varies between runs: the run reported in the paper gave 421 (Both dimension), and the shipped results files, from a second from-scratch run, give 375. All such differences lie within the worst-case bound in Section 5.5, which treats every unresolved dataset as passing: no area rate rises by more than 2.9 percentage points and no reported comparison reverses.

References

The data citations below also appear in the paper's references.

- Crawford, Vincent P., and Juanjuan Meng. 2011. "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income." *American Economic Review* 101 (5): 1912-1932.
- Crawford, Vincent P., and Juanjuan Meng. 2011a. "Replication data for: New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income." Nashville, TN: American Economic Association; Ann Arbor, MI: Inter-university Consortium for Political and Social Research (openICPSR). https://doi.org/10.3886/E112450V1
- Farber, Henry S. 2005. "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers." *Journal of Political Economy* 113 (1): 46-82.
- Farber, Henry S. 2008. "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers." *American Economic Review* 98 (3): 1069-1082.
- Farber, Henry S. 2008a. "Replication data for: Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers." Nashville, TN: American Economic Association; Ann Arbor, MI: Inter-university Consortium for Political and Social Research (openICPSR), distributed 2019-10-12. https://doi.org/10.3886/E113241V1
- U.S. Bureau of Labor Statistics. "Consumer Price Index for All Urban Consumers (CPI-U): All Items in New York-Newark-Jersey City, NY-NJ-PA, not seasonally adjusted," series CUURS12ASA0. Accessed June 2018. https://data.bls.gov/timeseries/CUURS12ASA0