README file

This README file documents replication files for:

> Richard Blundell, Joel Horowitz and Matthias Parey (2020): "Estimation of a Heterogeneous Demand Function with Berkson Errors," forthcoming *Review of Economics and Statistics*.

Replication files: The replication files contain code for Matlab (m-files) and Stata (do-files), as well as data in Matlab format (mat-files) and Stata format (dta-files). Instructions on how to run these files are found below (see Instructions).

Data sources: Note that data sources (1)-(3) below are provided with the code, while data sources (4)-(8) can be downloaded via the links provided.

1. The household demand data are derived from the 2001 National Household Travel Survey (NHTS, see `https://nhts.ornl.gov/download.shtml`), and supplemented with the data on distance to the Gulf of Mexico from Blundell, Horowitz and Parey (2012).

   The relevant variables in the data file `data_BHP2_trim.mat` are the following:

Variable name	Description
`log_q`	log quantity demanded per year [gallons/year]
`log_p`	log price per gallon [dollar/gallon]
`log_y`	log income [dollar/year]
`state_fips`	FIPS code of state where household is located
`distance_oi1000`	distance to Gulf of Mexico [1000km]

2. The file `berkson_sd_state.mat` provides the state-level Berkson standard deviation.

3. The file `state_draw.dta` indicates the observations from the station data drawn for the analysis in Table A.3.

Additional data sources are the following:

4. Yilmazkuday (2017), Supplemental material.

   File: `https://www.tandfonline.com/doi/suppl/10.1080/00036846.2017.1340575/suppl_file/raec_a_1340575_sm1109.xlsx`

   See also: `https://www.tandfonline.com/doi/suppl/10.1080/00036846.2017.1340575`

5. Missouri Census Data Center: 2015 Crosswalk ZIP to ZCTA.

   File: `https://mcdc.missouri.edu/data/corrlst/Zip_to_ZCTA_crosswalk_2015_JSI.csv`

   See also: `https://mcdc.missouri.edu/cgi-bin/uexplore?/data/corrlst`

6. U.S. Gazetteer Files, 2015.

   File: `https://www2.census.gov/geo/docs/maps-data/data/gazetteer/2015_Gazetteer/2015_Gaz_counties_national.zip`

   See also: `https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.2015.html`

7. 2010 ZIP Code Tabulation Area (ZCTA) Relationship File Record Layouts.

   File: `https://www2.census.gov/geo/docs/maps-data/data/rel/zcta_county_rel_10.txt`

   See also: `https://www.census.gov/programs-surveys/geography/technical-documentation/records-layout/2010-zcta-record-layout.html`

8. US Intercensal Population by County and State 1970-on.

   File: `https://www.nber.org/census/popest/county_population.dta`

   See also: `https://www.nber.org/research/data/us-intercensal-population-county-and-state-1970`

(Note that some of these data files are provided as zip files and need to be extracted from the zip archives.)

Software: The software used is (i) Matlab, together with the NAG Toolbox for Matlab, and (ii) Stata. Note that for producing the maps, Stata requires the command `maptile` to be installed, with the geographies U.S. counties 2014 (`county2014`) and U.S. states (`state`).

Instructions: Select a directory and set up the following sub-directories: `/code/` for the m-files and do-files provided, `/input/` for the data files referenced in points (1)-(8) above, `/output/` to collect intermediate files, and `/results/` for resulting tables and figures. (These paths can be adjusted in the beginning of the three files indicated in steps 1-3 below.)

1. Run `run1_transfer_station_data.m` using Matlab.
2. Run `run2_prepare_station_data.do` using Stata.
3. Run `run3_master.m` using Matlab.

Output files: The resulting tables are stored in the folder `/results/` as text files (`table*.txt`). The resulting figures are stored in the folder `/results/` as pdf files (`figure*.pdf`).