Replication code for

"Heterogeneous Choice Sets and Preferences"

by Levon Barseghyan, Maura Coughlin, Francesca Molinari, and Joshua C. Teitelbaum

1. The empirical analysis reported in Sections 4-5 makes use of proprietary data. Access to the data can be requested from the authors (see the attached agreement).

2. All replication files are written for MATLAB. The version used was R2020a.

3. Before running Steps 1-7, one needs to specify the proper path to the data folder. This will be self-explanatory once access to the data is granted.

4. Simulations do not require access to the data and can be run on any machine. They do call on parallel processing within MATLAB. The replicator may need to edit lines 118-121 (pasted below) in **Replication_Simulations_02.m** to force MATLAB into proper (depending on the machine used by the replicator) parallel cluster; or the replicator can just create Local12 parallel cluster profile.

```
if N_menu > 15
delete(gcp('nocreate'))
parpool Local12
end
```

5. The code is licensed under a **BSD** license. See **LICENSE.txt** for details.

6. The part of the code that uses the Kaido, Molinari, and Stoye (2019) method uses parts of the code described in ``Calibrated Projection in MATLAB: Users' Manual'' by Hiroaki Kaido, Francesca Molinari, Jörg Stoye, and Matthew Thirkettle. The latest version of that code can be found at https://github.com/MatthewThirkettle/calibrated-projection-MATLAB. Redistribution and use of that code are permitted on the terms and conditions of the LICENSE pertaining thereto, which can also be found at https://github.com/MatthewThirkettle/calibrated-projection-MATLAB.

The scripts that call for data

The scripts below must be run in the order presented here. Once the replicator runs them all, all Figures and Tables pertaining to the results in Sections 4-5 and the Supplement will be produced.

1. **Replication_Main_01.m:** This file reads in the proprietary data and creates the various analysis samples based on covariates. Empirical choice frequencies are computed and bootstrapped and model implied choice shares are computed across a fine grid of risk aversion.
2. **Replication_Main_02_dimension2.m:** This file computes Andrews and Shi (2013) confidence sets for the (mean,variance) of risk aversion and Kaido, Molinari, and Stoye (2019) confidence intervals for the mean of risk aversion.
3. **Replication_Main_02_dimension3.m:** This file computes Andrews and Shi (2013) confidence sets and Kaido, Molinari, and Stoye (2019) confidence intervals for the distribution of choice set size.
4. **Replication_Logit.m:** This file computes Andrews and Shi (2013) confidence sets for the main analysis sample used in Figure 5.1 using the corresponding procedure from Replication_Main_01 and Replication_Main_02_dimension2.
5. **Replication_ComparatorModels.m:** This file estimates the point identified uniform random and alternative-specific random comparison models and computes confidence intervals.
6. **Replication_Main_03a.m:** This file computes the descriptive statistics and confidence intervals in Table 4.1, Table 5.1, Table 5.2, Table S3.1, Table S3.2, Table S3.3, and Figure S3.2.
7. **Replication_Main_03b.m:** This file creates Figure 5.1 and Figure S3.1.

Simulations

A. **Replication_Simulations_01.m:** This file runs the calculations and plots the simulation figures for the large feasible set simulations in Table 6.1 and Figures 6.1 and 6.2.
B. **Replication_Simulations_02.m:** This file runs the calculations and plots the simulation figures for the high-dimension ν simulations in Table 6.1 and Figure 6.3.