This file is the "**ReadMe file**" for the posted programs (in Stata and Matlab) that generate the results reported in "**Consumer Price Search and Platform Design in Internet Commerce**" by Dinerstein, Einav, Levin, and Sundaresan.

All of the estimation data for the paper is available online. We include 9 datasets:

- **Data/rawsearchdatascrubbed.dta** contains the anonymized/scrubbed Halo Reach search data used for estimating our model.

- **Data/rawlistingsdatascrubbed.dta** contains the anonymized/scrubbed Halo Reach listings data used for estimating our model.

- **Data/products_data.csv** contains the product-level statistics used for Table 1.

- **Data/ab_data.dta** contains the statistics from the A/B experiment used for Figures 6a and 6b.

- **Input/searchdata.csv** contains the anonymized/scrubbed Halo Reach search data read into Matlab. It is also created in **databuild.do**

- **Input/listingsdata.csv** contains the anonymized/scrubbed Halo Reach listings data read into Matlab. It is also created in **databuild.do**

- **Input/qualitydata.csv** contains the platform data used to estimate our platform model in Matlab. It is also created in **databuild.do**

- **Input/activelistingsfreq0.csv** contains the distribution of the number of actives sellers in the before period and is read into Matlab. It is also created in **databuild.do**

- **Input/activelistingsfreq1.csv** contains the distribution of the number of actives sellers in the after period and is read into Matlab. It is also created in **databuild.do**

The available data allows interested researchers to replicate the analysis.

We posted 3 wrapper code files:

- **databuild.do** is a Stata do file, which calls the code that generates the csv files read into Matlab for estimation and the associated datasets. It is in the "code > stata" folder. The csv files are saved in the "input" folder as they are Matlab input while the Stata files are in the "data" folder. This data build cannot be run without the underlying raw data, but we provide all of the relevant code that was used to generate the estimation data.

- **wrapper_matlab.m** is a Matlab m file, which calls the code that generates Figures 1, 2, A1 and Tables 3, 4, 5. It also creates output that is used in Tables 7, 8, 9. It also calls the code that produces the results referred to in Section 6.3. The script and function m files it calls are in the same "code > matlab" folder.

- **replicate_tables_figures_stata.do** is a Stata do file, which contains the code that generates Figures 5a, 5b, 6, 7, 8, 9, 10 and Tables 1, 2, 6. We provide the output, which matches the tabs in our tables and figures file: "**Source of Tables and Most Figures.xlsx**." Note that this Stata wrapper should be run after the Matlab code. This file is in the "code > stata" folder.