

Econometric Benchmarking Replication

The following instructions guide interested parties in replicating PEG's results.

Install R, Stata, and the Required Packages

The model data processing and transformations are coded in R. If the user does not already have R installed, it can be downloaded for free at <https://cran.rstudio.com/>. PEG recommends R Studio as the integrated development environment for R, which can also be downloaded for free at <https://www.rstudio.com/products/rstudio/#Desktop>. Additionally, the packages "foreign", "gtools", "dplyr", and "openxlsx" need to be downloaded and installed into R if they are not already. In R Studio, the user can install these packages by going to Tools > Install Packages and searching for the required package. Once R and the required packages are installed, the contents of the folder that contains this README file, along with the confidential database, should be saved to a directory. For the remainder of this document, such a directory will be referred to as [dir].

Specify the Path Variables

Find the path variables under the heading "Define Paths": dataPath, pegIDmapPath, pegDataOutPath, funPath. They are currently written as:

```
dataPath = "[dir]/ HQTDB (Confidential).xlsx"
```

```
funPath = "[dir]/SingleEqFUNS.R"
```

```
pegDataOutPath = "[dir]/EconometricData_Transformed.xlsx"
```

and should now be written to reflect the user's directory path, noting that on a Windows machine, R requires forward slashes instead of backward slashes for separating directories in the file path.

Execute the R Script

After the path variables are specified, click the "source" button in the top-right corner to execute the script. If there are no user-errors, the output file will have been written to the location specified in the "pegDataOutPath". It contains the data with all imputations and transformations which can be imported to STATA to estimate the models.

Use STATA to perform model estimation:

Edit the file HQT_OLSmodels.do to replace the [dir] in line 12 with the filepath where the R output was saved on the user's computer, and replace the [dir] in line 70 with the filepath where the user wishes to store the benchmark scores.

Execute the .do file using Tools>Execute(do)

After the score results are exported to the Excel file, the user will need to change the cell formatting in Excel to Number in the "diff" columns and increase the number of decimal places displayed to view the score details. The models can be viewed in the STATA window and pasted into a text file if the user wishes.