**The Significance of Data-Sharing Policy**

This document presents the files used in the article “The Significance of Data-Sharing Policy” by Askarov et al. *Journal of the European Economic Association*. There are two sets of files: (1) files used to produce results in the main paper and the Online Supplement and (2) intermediate files used to handle the raw data.

**1. Main paper and Online Supplement**

*Command file*

The file **Mandatory data-sharing 30 Aug 2022.do** contains all the Stata (version 16) commands used in the main article and in the Online Supplement. To run these commands, you will need to install: (1) **did\_imputation** and (2) **event\_plot**. Also, update: (1) **reghdfe** and (2) **ftools**.

*Main paper & supplement data files*

|  |  |
| --- | --- |
| File: | Used to produce: |
|  | *Main paper* |
| **Mandatory data-sharing 30 Aug 2022.dta** | Tables 2, 3, 4 and 5. Figures 1 (except panel D), 2, 3 (except panel D), 5 & 6. |
| **Time series.dta** | Figure 1, panel D & Figure 3, panel D |
| **Data shared.dta** | Figure 4 |
|  |  |
|  | *Online Supplement* |
| **Mandatory data-sharing 30 Aug 2022.dta** | Tables S1, S2, S3, S4 (Columns (1) & (2)), S5, S6, S7, & S8. Figures S3-S8. |
| **Using only top 31.dta** | Table S4, Column (3) |
| **With outliers included 30 Aug.dta** | Table S4, Column (4) |
| **Stacked data 30 Aug.dta** | Table S10 |
| **Analysis of shares July 2022.dta** | Table S11 |
| **Data-sharing journals in event time.dta** | Figure S2 |
| **New submissions AER.dta** | Figure S7 |
| **Selection bias.dta** | Figure S8 |

*Power of pre-trends test*

The CSV files (**sigma tstat2.csv**, **beta tstat2.csv, beta\_true tstat2.csv, sigma ess.csv, beta\_true ess.csv** and **beta ess.csv**) are used to conduct the pre-trends power and likelihood ratio tests reported in footnote 38 of the main paper. These tests re conducted via <https://jonathandroth.shinyapps.io/pretrendspowerapp/> (accessed 11th July 2022).

**2. Intermediate data files**

The files in the folder **Datasharing Code**, present the Python code used to identify outliers and leverage points, to calculate UWLS meta-averages and to calculate the associated ESS estimates. These files also produce Table 1 of the main paper. We used Python version 3.6. See README.md.

|  |  |
| --- | --- |
| *Intermediate data files* | |
| **All data.xlsx** | This file contains the raw data from which we calculate outliers, leverage points, and UWLS. We then use the associated estimates from top journals for the analyses reported in the main paper and Online Supplement as documented above. |