**READ ME**

**Folder Contents:**

The replication folder contains the following do-files:

- regfile_constrn.do: Construction of regressionfile.dta used for generating the tables and figures used in the paper. This do-file details out the creation of the main variables used in the paper and the filters applied to publicly available Compustat data.
- regressionfile_pseudodata. This is a pseudo-random sample of the regressionfile.dta that can be used to run TablesADM_RFS.do. Note that this is pseudo data so it will not generate the tables in the paper.
- TablesADM_RFS.do: Generates all the tables used in the paper.
- FiguresADM_RFS.do: Generates all the figures used in the paper.

**Data Provenance:**

The following publicly available datasets were used in the paper, "*The Rise of Star Firms: Intangible Capital and Competition*" by Ayyagari, Demirguc-Kunt, and Maksimovic.

1> Compustat from WRDS:
Section 2.2-2.4 in the paper defines the data cleaning procedure on Compustat. The main filters used were:
   - Restrict companies to those incorporated in the US (fic="USA")
   - Industrial format (indfmt="INDL")
   - Missing NAICS codes
   - Drop Public Administration (SIC 90-99), Finance, Real Estate and Insurance (SIC 60-69), Utilities (SIC 49)
   - Replaced negative values as missing for INTAN, XRD, SALE, COGS, XOPR, CAPX, AM, DP, GDWL, XSGA, RDIP
   - Drop if Sales (SALE) or Total Assets (AT) is missing

2> Peters and Taylor (2017)'s measures of Total Q and Intangible Capital from WRDS

We use the Peters and Taylor definition of intangible capital. They do not recommend using the measure (a) Before 1975 (b) For firms with missing or non-positive book value of assets/sales or (c) For firms with <$5 million in physical capital (Compustat variable ppegt) measured in 1990 dollars.

3> Hoberg and Maksimovic (2022)'s measures of product life cycle downloaded from Hoberg-Maksimovic Product Life Cycles Data Repository (http://faculty.marshall.usc.edu/Gerard-Hoberg/HobergMaxLifeCycles/index.html )

4> Kogan et al. (2017)'s measure of Patent Market Value. Available from the online Appendix at https://data.nber.org/data-appendix/w17769/

5> GDP Deflator from U.S. Bureau of Economic Analysis, Gross Domestic Product: Implicit Price Deflator [A191RI1A225NBEA], retrieved from FRED, Federal Reserve Bank of St. Louishttps://fred.stlouisfed.org/series/A191RI1A225NBEA, July 21, 2023.

6> Chinese Import Shock as constructed in Autor, Dorn, and Hanson (2013) but at 4-digit NAICS Level.

TableA9 in the paper provides detailed variable definitions for all the variables used in the paper.