

## Replication Package for Innovation Booms, Easy Financing, and Human Capital Accumulation

Johan Hombert, Adrien Matray

### Reproducibility certificate

All the results in the paper have been reproduced by the Certification Agency for Scientific Code And Data ([cascad-CASD](#)) using the code in this replication package. The reproducibility certificate is available at <https://doi.org/10.5281/zenodo.7969175>.

### Data sources

French administrative data are made available to researchers by CASD (link to [Secure Data Access Centre](#)). The databases from CASD used in the paper are:

1. FICUS 1994-2007 <https://www.casd.eu/source/statistique-structurelle-annuelle-dentreprises-issue-du-dispositif-suse/>
2. FARE 2008-2016 <https://www.casd.eu/source/statistique-structurelle-annuelle-dentreprises-issue-du-dispositif-esane/>
3. DADS Panel tous salariés 2015 <https://www.casd.eu/source/panel-tous-salaries/>
4. DADS Postes 1994-2015 <https://www.casd.eu/source/base-tous-salaries-fichier-postes/>
5. Créations d'Entreprises 1993-2018 <https://www.casd.eu/source/creations-dentreprises/>
6. Stocks Entreprises et Etablissements 1993-1999 <https://www.casd.eu/source/stocks-dentreprises-et-etablissements/>
7. Stocks Entreprises 2000-2008 <https://www.casd.eu/source/stocks-dentreprises/>
8. Enquête LIFI 1994-2007 <https://www.casd.eu/source/enquete-liaisons-financieres-entre-societes/>

The other databases used in the paper are:

9. Eurofidai: stock price information to construct the over-valuation proxies. The data can be purchased from Eurofidai (see <https://www.eurofidai.org/>).
10. Current Population Survey: for the appendix on the US. The data can be downloaded from IPUMS: <https://cps.ipums.org/cps/>

### Instructions for replication of results using CASD data

All the results in the paper based on CASD data (except those using the over-valuation proxies constructed using Eurofidai data) can be replicated by researchers with access to the above databases from CASD using the code provided in this replication package and going through the following steps:

1. The master file is the Stata dofile `_master.do` in the folder CASD.
2. In the master file, input path to the project directory (line 8 in the code) et to the directories containing the CASD data (line 11).
3. Run the Stata code in the master file up to line 49.
4. Run the SAS code in the file `extract_dads.txt`.
5. Continue running the Stata code in the master file up to line 78.
6. Run the SAS code in the file `extract_fare.txt`.

7. Place the file `stock_price_firm.dta` constructed using Eurofidai data in the folder `rawdata`.
8. Run the rest of the Stata code in the master file until the end of the file. The code produces all the tables and figures in the paper, except for Figures 1a and C.1 which are described below.

#### Instructions for replication of results using Eurofidai data

The file `eurofidai.do` in the folder `Eurofidai` contains the Stata code, which starts from the raw data downloaded on the Eurofidai platform and produces Figure 1a.

#### Instructions for replication of results using CPS data

The file `cps.do` in the folder `CPS` contains the Stata code, which starts from the raw data downloaded from IPUMS and produces Figure C.1.

The data are downloaded from IPUMS in the following way:

1. Go to: <https://cps.ipums.org/cps-action/variables/group>
2. Click on "Select samples". Download the data decade by decade for decades 1990s, 2000s and 2010s (see screenshot) and select all the variables.
3. Save the files as `cps1990s.dta`, `csp2000s.dta`, and `csp2010s.dta`.