

Read Me

Ryan Chahrour*

Boston College

Kristoffer Nimark[†]

Cornell University

Stefan Pitschner [‡]

Uppsala University

May 5, 2021

Abstract

Data and replication instructions for *Sectoral Media Focus and Aggregate Fluctuations*.

*Department of Economics, Boston College, Chestnut Hill, MA 02467, U.S.A. *e-mail*: ryan.chahrour@bc.edu

[†]Department of Economics, Cornell University, Ithaca, NY 14853, U.S.A. *e-mail*: pkn8@cornell.edu

[‡]Department of Economics, Uppsala University, Uppsala, Sweden. *e-mail*: pitschner@gmail.com,

Data availability and Provenance:

1. The BEA input-output tables were downloaded from the BEA ([https://apps.bea.gov/industry/iTables Static Files/AllTablesIO.zip](https://apps.bea.gov/industry/iTables%20Static%20Files/AllTablesIO.zip)). The relevant table is called `IOUse_After_Redefinitions_PR0_1997-2017_Summary.xlsx`, which we have renamed `IOUse.xlsx`. A copy of these data are included in this archive.
2. The BEA/BLS integrated multi-factor productivity tables were downloaded from the BLS (<https://www.bls.gov/mfp/indprod11.xlsx>). A copy of this data is included in this archive.
3. Aggregate GDP and hours, as well as all sectoral employment series, were downloaded from the FRED database maintained by the Federal Reserve Bank of St. Louis (<https://fred.stlouisfed.org>). The aggregate variables used are gross domestic product (FRED code GDP), the GDP deflator (GDPDEF), non-institutional population over 16 (CNP16OV), hours worked in the non-farm business sector (HOANBS). Sectoral employment series used are listed in the file `/cnp_sectors.xlsx`, in the tab labelled “CNP Sectors to CES”. The data were download in Matlab using the DataStream toolbox, using the commented portions of the `load_basic_data.m` and `load_sectoral_emp.m` commands included in the replication files. A copy of this data is included in this archive.
4. The measure of news heard comes from the Survey of Consumers at the University of Michigan (<https://data.sca.isr.umich.edu/charts.php>), Table 23. A copy of this data is included in this archive.
5. The measures of economics expectations come the Survey of Professional Forecasters are available from the Philadelphia Federal Reserve, and were downloaded from <https://www.philadelphiafed.org/-/media/frbp/assets/surveys-and-data/survey-of-professional-forecasters/historical-data/mediangrowth.xlsx>. A copy of this data is included in this archive.
6. The real-time measures of GDP (we use the third-release) are also available from the Philadelphia Fed, and were downloaded from https://www.philadelphiafed.org/-/media/frbp/assets/surveys-and-data/real-time-data/data-files/xlsx/routput_first_second_third.xlsx. A copy of these data are included in this archive, renamed as `routput.xlsx`.

7. All news data are obtained from Factiva (<https://www.factiva.com>). All search definitions are included in the corresponding Excel files included in this archive.
8. The news linktable is used to merge company tags from Factiva to their sectors. For each company tag, the source of the link is indicated in the column “match source”.

Dataset List:

Data file / Folder	Source	Notes	Provided?
cnp_sectors.xlsx	Authors correspondances	Based on Atalay (2018)	Yes
BEA_io/IOUse.xlsx	BEA	Input-output calibration	Yes
BLS_multifactor/indprod11.xlsx	BEA/BLS	TFP timeseries	Yes
output_files/gdp_hrs_raw.mat	St. Louis FRED	Basic aggregates	Yes
output_files/sector_labor_raw.mat	St. Louis FRED	Sectoral labor timeseries	Yes
SVAR_data/Michigan/redbk23.xls	Mich. Survey of Cons. Expect.	News-heard measure	Yes
SVAR_data/SPF/medianGrowth.xlsx	Philadelphia Fed.	Median growth forecasts	Yes
news/inputs/raw_company_counts/	Factiva	Quarterly counts of company names	Yes
news/inputs/raw_ngram_counts/	Factiva	Quarterly counts of n-gram sectors	Yes
news/inputs/linktable.xlsx	Factiva/Compustat/Manual	Links company names to sectors	Yes
news/inputs/multifactor_output.xlsx	BEA/BLS	Macro data used in news analysis	Yes

Instructions for Replication of the Macro Analysis:

The replication code is organized according to each task. Helper functions that are reused several times across directories are stored in directories labeled `helper_functions` and `ts_box`.

All computational steps can be performed at once by calling:

```
>> run_all
```

from the root directory. Re-computing all results from scratch takes around 45 minutes.

Step-by-step instructions for replicating each figure in the paper, with or without using saved results, are below. All `>>cd` commands assume you begin in the root `/public_replication_code` directory. All output files are saved in the `output_files` directory.

Software:

This code was last run/tested in Mac OSX 10.14 using Matlab R2020a. The code requires Matlab's Symbolic Toolbox and Parallel Computing Toolbox. Original download of FRED data also used Matlab's Datafeed toolbox. No additional toolboxes are needed to replicate the results here.

The files have also all been run on Red Hat Linux Cluster, `Linux 3.10.0-1062.el7.x86_64`.

Hardware:

These programs have no special hardware requirements. Running the full replication code takes about 45 minutes on a 4-core 2018 iMac.

Setting up & data transformation

- Before running any code, add all auxiliary files to the path by calling:

```
>> setup
```

from the root directory.

- To load and compile the BEA input-output tables:

```
>> cd BEA_io  
>> import_use_table
```

These steps save the required variables to `/output_files/I029.mat`.

- To load and compile the BEA/BLS productivity timeseries:

```
>> cd BLS_multifactor  
>> import_bls_data
```

These steps save the required variables to `/output_files/tfp.mat`.

- To load and transform the aggregate historical US data mat-file from the original series downloaded from FRED:

```
>> cd gdp_stats  
>> load_basic_data
```

These steps save the required variables to `/output_files/gdp_hrs.mat`.

- To load the data used in the SVAR analysis (including sectoral labor, nowcast errors, RGDP growth, and the news-heard index):

```
>> cd SVAR_data
>> load_all
```

These steps save the required variables to `/output_files/var_data.mat`.

Main results

- To estimate the sectoral news weights used in the news selection function:

```
>> cd news_selection_function
>> nsf_calibrator
```

These steps save the required variables to `/output_files/omega.mat`.

- To solve the model for a range of ν .

```
>> cd model_solver
>> run_nu
```

These steps save the required variables to `/output_files/model_solution.mat`.

- To reproduce figures 11-15:

```
>> cd model_solver
>> make_figures
```

These steps save the figures to `out_files/*.pdf`. This command also prints to screen all numbers related to model solution that are reported in the text.

SVAR exercise

- To perform the SVAR exercise:

```
>> cd SVAR_approach
>> SVAR
```

These steps use results saved in `/output_files/svar_output.mat`.

- To reproduce Figures 16-18:

```
>> cd SVAR_approach
>> make_figures
```

These steps save the figures to `out_files/*.pdf`. This command also prints to screen all numbers related to this exercise that are reported in the text.

Instructions for Replication of the News Analysis:

Hardware and Software:

The code was written in Python 3 and was last run/tested in Mac OSX 10.15. The code requires the following Python packages: matplotlib, numpy, os, pandas, regex, statsmodels. Versions as of May 2021 are recommended. The code has no special hardware requirements.

Setup and Running the Replication Scripts:

All replication scripts are located in the folder news/scripts. All input files and raw data are provided in the folder news/inputs. All outputs generated by the replication scripts are written to the folder news/outputs.

To replicate the figures and regressions, first copy all files to a local folder on your computer. Then set this folder as the working directory of the script you want to run. This is done in line 6 of each script. The following table summarizes what each script does.

Filename	Description
news/scripts/prep_news_data.py	creates news datasets from the raw inputs
news/scripts/regressions.py	replicates the regressions shown in Table 2
news/scripts/plots_news_vs_macro.py	creates Figures 6 and 8
news/scripts/plots_top_companies_top_sectors.py	creates Figure 7
news/scripts/plots_news_time_series.py	creates Figures 9, A7, A8 and A9
news/scripts/plots_ngram_comparison.py	creates Figure 10
news/scripts/plots_matched_and_unmatched_tags	creates Figure A1
news/scripts/plots_top_companies_all_sectors.py	creates Figures A2, A3, A4, A5 and A6