READ ME FILE

Title: Start Spreading the News: News Sentiment and Economic Activity in Australia

Authors: Kim Nguyen and Gianni La Cava

Description

This 'read me' file details the replication files for RDP 2020-08. The Data folder contains csv and excel files. The Codes folder contains Python scripts, R files and Stata's do files that are used to produce the results presented in the paper and the appendices.

If you make use of any of these files you should clearly attribute the authors in any derivative work.

Folder structure

The main folder contains the rdp-2020-08-read-me.pdf and the xlsx file 'rdp-2020-08-graph-data.xlsx', which provides the publicly available data for the figures as printed in the RDP in excel format.

It also contains the following folders:

Data

The four files in the data folder are the primary inputs to the analysis files:

A. weekly_data.csv – this contains the weekly News Sentiment Index and News Uncertainty Index.

B. monthly_data.csv – this contains the following monthly time series:

   1) News Sentiment Index

   2) News Uncertainty Index

   3) ABS Change in Unemployment Rate

C. mp_data folder – this contains the following time series:

   1) Monetary Policy News Sentiment Index – daily

   2) Changes in cash rate

   3) Interest rates and yields – money market (current and historical)

   4) Board meetings dates

   5) ASX all ordinaries price index

   (Note: More details about the use of these series are contained in the Stata code 'Daily Cash Rate and News Sentiment – Setup.do')

D. alternative_nsi.csv – this contains the News Sentiment indices using alternative methods:

   1) using Harvard General Inquirer (GI) Dictionary

   2) using Hu and Liu (2004)'s lexicon

   3) using Shapiro, Sudhof and Wilson (2017)'s fixed effect model

   4) using Machine Learning

E. The following time series are used in the analysis but cannot be released publicly:

   1) ANZ – Roy Morgan Consumer Sentiment Index (weekly and monthly)

   2) NAB Business Condition Index

3) NAB Business Confidence Index

4) NAB Capital Expenditure Index

5) Bloomberg Economic News

6) Monetary Policy shocks series

Codes

This folder contains the files you can use to reproduce the analysis, including:

1. The following Python codes clean and do textual analysis on raw news data to generate time series of NSI and NUI: dna_clean.py, dna_sentiment.py, dna_categories.py, dna_timeseries.py.

2. main_LP.R – this code produces main local projection and smooth local projection results presented in Section 5 and Subsection 7.1.

3. lproj.R – this code creates the LP and SLP functions to estimate impulse responses and the confidence intervals. Adapted from Jordà (2005) and Barnichon and Brownlees (2019).

4. robustness_LP.R – this code produces the local projection and smooth local projection results presented Appendix A.

5. robustness_VAR.do – this code produces the VAR impulse responses presented in Section 7.3 and Appendix B.

6. To produce the local projection and smooth local projection results presented in Section 6, run the following codes:

    a. For setting up the data, run Daily Cash Rate and News Sentiment – Setup.do

    b. For estimating the local projections, run Daily Cash Rate and News Sentiment – LP.do

    c. For estimating the smoothed local projections, run Daily Cash Rate and News Sentiment - Smoothed LP.do

Software versions:

1. Python 3.6.6 and Anaconda 1.9.12 (with Spyder 3.3.4)

2. R for Windows 3.6.1 and R Studio 1.2.5001

3. Stata MP 16 (64 bit)

References

Barnichon R and C Brownlees (2019), 'Impulse Response Estimation by Smooth Local Projections', *The Review of Economics and Statistics*, 101(3), pp 522–530.

Jordà Ò (2005), 'Estimation and Inference of Impulse Responses by Local Projections', *The American Economic Review*, 95(1), pp 161–182.

Hu M and B Liu (2004), 'Mining and Summarizing Customer Reviews', in R Kohavi, J Gehrke, W DuMouchel and J Ghosh (eds), *KDD-2004 Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, The Association for Computing Machinery, New York, pp 168–177.

Shapiro AH, M Sudhof and D Wilson (2017), 'Measuring News Sentiment', Federal Reserve Bank of San Francisco Working Paper 2017-01, rev March 2020.

3 February 2021