FINM 35000 Problem Set 2: Monetary Policy and Inflation

Instructor: Joanna Harris, TA: Lisheng Su

Due October 28, 2022, 6PM

Submission instructions: please submit one copy of the assignment per group. The submission should include a PDF or Word document with the written results, tables and figures and a separate code file (in a programming language of your choosing). Please make sure to write the names of all group members at the top of the writeup.

Supplementary Files: the data files and papers that you will need for this assignment can be found at https://www.dropbox.com/sh/2q97auk8o531upy/AACSGfndpkNu4EfJzvDhMQYXa?dl=0

1 Measuring Hawkish/Dovish Tone of FOMC Statements

1. Scrape the text of the FOMC statements from January 2000 to present. You will need to use https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm for 2017-2022 and https://www.federalreserve.gov/monetarypolicy/fomc_historical_year.htm for 2000-2016. See Appendix A for hints on scraping.

   (a) How many statements do you obtain?

   (b) Provide summary statistics (mean, standard deviation, minimum, first quartile, median, third quartile, maximum) for the number of words in each statement.

2. Use the methodology described in section 3.1 (pages 4-8) of Tadle (2022) to measure the tone of each speech.[1] Download the Fed Funds Effective Rate from https://fred.stlouisfed.org/series/DFF. Plot both both the statement tone and the Fed Funds Effective Rate over time.

3. Comment on the Tadle (2022) methodology. What do you like about it? What are its shortcomings?

4. Describe and implement a different way to measure hawkish/dovish tone of FOMC statements.[2] How does your alternative measure address some of the shortcomings in the Tadle (2022) method? What is the correlation between the Tadle (2022) measure and your measure?

5. Redo the plot from problem 2, adding your tone measure. Your plot should include the Tadle (2022) measure, your measure and the Fed Funds Effective Rate.

6. Complete this problem ***twice***: once with the Tadle (2022) measure of hawkishness and once with your measure. The steps will guide you through using the Fama and MacBeth (1973) procedure to estimate the monetary policy risk premium in industry returns data.[3] In particular:

[1] See Appendix B for a walkthrough of this measure. I have provided an Excel file with the terms from Tables 1 and 2.

[2] You can use a topic model, use BERT, try to improve on the dictionary based method, etc. Get creative! I will give you some examples about how NLP techniques are used for monetary policy in the October 19 lecture, so don't worry if this is new to you.

[3] Remember from last time that you have to replace -99.99 and -999 with null values.

(a) Estimate an AR(1) model from the hawkishness data. Compute the residual. We will call this the "text-based monetary policy shock."[4]

(b) Use the value-weighted returns from the daily industry returns file for this problem. For each of the 49 industries, regress returns on the day of the Fed announcement on the text-based monetary policy shock. Create a table with three columns: column 1 has the industry name, column 2 has the OLS regression coefficient and column 3 has the p-value for that coefficient. Sort the table from largest to smallest coefficient. To be concrete, the regression you are running at this stage is:

$$R_{it} = \alpha_i + \beta_i \Delta HAWK_t + \varepsilon_{it}$$

for each of the 49 industries, indexed by i. Time t here indexes Fed announcement days and $\Delta HAWK_t$ is the text-based monetary policy shock at time t.

(c) Comment on the ordering of the industries. Is it in line with what you would have expected?

(d) Now turn to the monthly returns data.[5] Again use the value-weighted returns. Separately for each month, regress returns of each industry on its "beta" from step (b). To be concrete, for each month, indexed by T, you are running the following regression:

$$R_{i,T} = \gamma_T + \lambda_T \beta_i + \eta_{iT}$$

(e) What is the average λ across all months? This is the risk premium associated with holding assets exposed to monetary policy risk. Comment on how its sign can be interpreted.

(f) What is the standard deviation of λ? Use this to compute the t-statistic. Is the risk premium significantly different from zero at the 10%, 5% or 1% level?

2 Reading Response

Note: This section is optional and can be submitted for bonus points.

Read Aruoba and Drechsel (2022) and answer the following questions.

1. Describe the natural language processing technique that the authors use. What do you like about it? What are its shortcomings?

2. Summarize (in your own words) the findings discussed in Section 4.

3. What questions do you have after reading the paper?

[4] There are a number of things wrong with estimating the shock this way, but it is the most straightforward way for the sake of this problem set. First, one should estimate the AR(1) in rolling windows so that the estimation only uses information available that is available prior to the time of each shock. Second, the expectation should also reflect macro information that has been released since the last meeting.

[5] Alternatively, you can compute cumulative returns for each month using the daily data. Do whichever one you find easier.

3 Appendix A: Hints for Scraping in Python

Note: you do not have to take these hints, as it is very possible that some of you are more proficient with web scraping than I am and will know a better way.

1. Use the BeautifulSoup package to extract the html from the website. Load this package by running the code: `from bs4 import BeautifulSoup`.
2. Use the command `soup.findall('a',href=True)` to get a list of all the URL links on the web page.
3. Inspect the URLs to figure out what the format of the links to statement files look like and narrow your list of links to only include these.
4. Iterate through the links to the statements and again use the `BeautifulSoup` package to extract the text.
5. Inspect the text to remove extraneous parts from the start and end.
6. Make sure to remove HTML tags.
7. You will need the date of the statement to merge the text to the returns. It can be found in the URL. For example, https://www.federalreserve.gov/newsevents/pressreleases/monetary20220126a.htm was released on Jan. 26, 2022, which is at the end of the URL.

4 Appendix B: Description of the Tadle (2022) Methodology

Follow these steps to implement the Tadle (2022) methodology:

1. Split the statement into sentences using `sent_tokenize` from the `nltk` package in Python (or an equivalent in the programming language of your choice).
2. Remove punctuation and capitalization.
3. Remove all sentences that do not contain any keywords (defined as those from the hawkish or dovish lists).
4. For each sentence that remains, count the number of positive and negative keywords, adjusting for negation terms. "When a positive term is in the proximity (that is, if they occur after three words or less) of a negation term, then its effect is counted as negative. On the other hand, a negative term is counted as positive if it immediately follows a negation term"
5. Use equation (1) on page 7 of Tadle (2022) to assign sentiment scores to each sentence. Note footnote 13 for how to handle sentences that contain both hawkish and dovish keywords.
6. Aggregate across sentences to get the document-level score for each statement following equation (2) on page 7 of Tadle (2022).

References

Aruoba, S. B. and Drechsel, T. (2022). Identifying monetary policy shocks: A natural language approach.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, 81(3):607–636.

Tadle, R. C. (2022). Fomc minutes sentiments and their impact on financial markets. *Journal of economics and business*, 118:106021.