

Kayle Liao

May 2022

The Thesis Library Research Award Essay

I wanted to start my essay by saying this — without Eric Lease Morgan's help of the Center for Digital Scholarship, I would not have been able to complete my Honors Economics thesis. I used some graphs generated from the software Eric introduced me to put on my thesis poster and I won second place poster award for my thesis. My thesis research started when I was formulating my research question for my thesis. I wanted to do something about measuring political pressure and evaluating the effect of political pressure on the Federal Reserve. The biggest challenge I faced upfront was not being able to measure the unquantifiable political pressure. Per my thesis advisor Eric Sims' advice, I came to the Center for Digital Scholarship on the second floor of Hesburgh Library to ask for help. I had the pleasure of meeting Eric Morgan, one of the best data scientists Hesburgh Library has to offer. At our first meeting in September 2022, Eric and I explored datasets that capture political pressure and opportunities to measure political pressure. Eric was very resourceful to tell me about available datasets out there capturing political pressure. I communicated with him my research interest and my motivation behind writing this thesis. I told Eric that former President Donald Trump had made negative comments on the Federal Reserve and monetary policy during his term and I was intrigued to explore how politicians' speeches impact how monetary policy is conducted. Eric is a great listener and quickly broke down my task into counting and tabulating the number of times certain words are mentioned in the presidential speeches. Before deciding on using presidential speeches, we considered other sources of political pressure index. I brought up the idea of using Twitter as a venue to look at political pressure. However, Eric had a realistic assessment and said

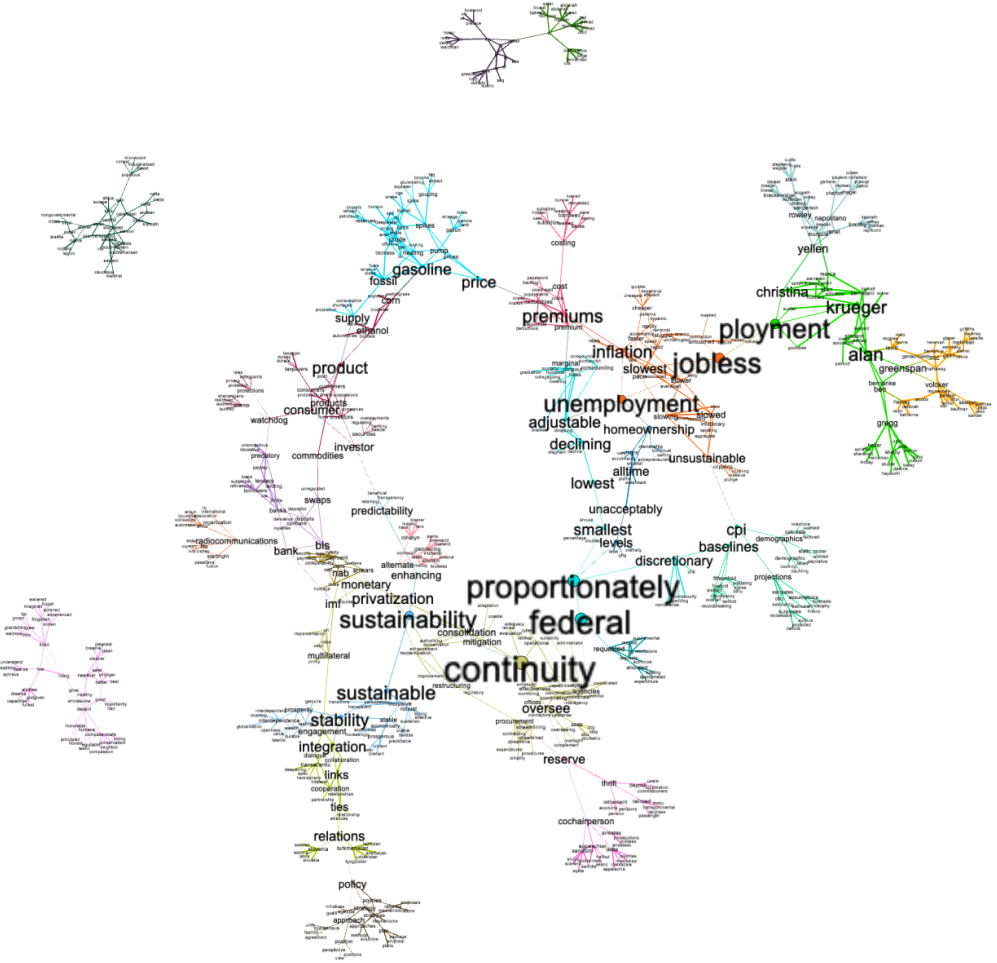
that it would be too time-consuming to filter out the tweets that contain the idea of monetary policy and the federal reserve. At the end of our discussion, we concluded that presidential speeches would be the prime source of my data because speech records are readily available on the reserve.gov website. I'm very thankful to Eric for helping me identify my data source as an experienced data scientist.

What differentiates my thesis from most other Economics theses is that I need to construct my own dataset because there isn't a readily available political pressure index. I knew at the end of my first meeting with Eric Morgan that my senior year would be a rewarding year learning data science and visualizing abstract social science concepts through coding. He opened me to a brand new world of data science with various terms. My favorite one is "lexicons", which refers to the keywords I'm using to reflect the idea of "monetary policy" and "Federal Reserve". For example, some of my initial lexicons include "inflation", "unemployment rate", and "Janet Yellen."

I still vividly remember Eric telling me that this would be an "iterative" process — we would work together throughout my senior year to construct the "political pressure index" that I will later use to put into my economic regression model to regress against the two main macroeconomic variables that are inflation and unemployment. I want to give credit to Eric Morgan for his fantastic idea of starting a pilot study for constructing the political pressure index. I was able to learn what "text-mining" was through Eric Morgan — it is the data science tool to count and tabulate the number of times some words are mentioned in the text. To allow Python to process the texts I was using, I needed to download the presidential speeches in xml format. Eric helped write the Python programming to count and tabulate the "lexicons" and trained me to

connect to the computer in downtown south bend in order to run his Python program when I have an updated version of my lexicons.

Eric Morgan was also very helpful in my search for appropriate lexicons that best capture the idea of monetary policy and the federal reserve. At the beginning of my thesis research, I was very concerned about missing out on words that presidents use to refer to economic terms like unemployment rate and inflation. Because what if some president refers to inflation as “gas prices” whereas another president may say “cost of living.” To refine my list of lexicons, Eric introduced a mathematical graphing tool called Gephi to me to map out the words in relation to one another in the context they are mentioned. Gephi is great in that it tells me what other words are mentioned when, for example, “inflation” is mentioned in the presidential speeches. Below is a graph I generated through Gephi using the list of lexicons.



I was able to identify more lexicons by using Gephi. As Eric said in September 2022, it truly was an “iterative” process even when it came to the process of refining my lexicons. Every time I updated my list of lexicons, I reran the Gephi visualization tool to generate a graph like the above. By working with Eric throughout the past year, I learned not just cutting-edge data science to turn the “intangible” like political pressure into “tangible” data series but also the spirit of doing academic research. Research requires patience and repetition and is a reiterative process that takes time. In the end, I was able to use my political pressure index as a variable in my economic regression and reached my conclusion that the Fed deviates from targeting inflation to unemployment when the political pressure is high.

To conclude, my thesis journey was not only a demonstration of my intellectual curiosity and interdisciplinary understanding but also a testament to the Hesburgh Library’s excellent staff and services.

Figure 3 Political Pressure Index

