

## ULRA Application Essay

Junior year internships in quantitative economic research helped me recognize that I wanted my senior thesis to be a data-intensive, rigorous economic analysis like those published in academic economic journals. When I embarked on my thesis, however, I had no idea how important the resources of the Hesburgh library system would be for transforming my thesis from an exciting possibility to actual economic analysis. Writing my thesis would not have been possible without the resources of the library system, nor as enjoyable without the consultation of its staff.

My senior thesis analyzes copper market price expectations, which are an integral component for macroeconomic modeling and forecasting. Building on my advisor's similar work in the oil market, I conduct a forecasting horserace between existing price expectations models in the literature to find the most accurate estimate. My analysis produces an improved expectations estimate, and its most significant implication for the literature is apparent heterogeneity in "risk pricing factors"—the variables to pay attention to in order to accurately measure price expectations—across commodity markets. Going forward, I hope to extend the work to other base metals and prepare the article for submission to economics journals.

My many hours spent in the Mendoza library began in May 2017, when I started the certification course for using the Bloomberg terminals. A leading financial database, the Bloomberg terminals provided my core historical macroeconomic, financial, and commodity data. Each of the models in my forecasting horserace is based on a collection of predictor variables. Methodically reconstructing each of the 30 variables meant I spent the much of the first two months of the school year collecting and processing data of diverse frequencies, types, and sources. The Bloomberg terminals were my primary data source, but I also leveraged the library's access to CRSP, Datastream, the Global Financial Database (GFD), and digital libraries like JSTOR. Most critically for my thesis's success, many visits to the library flowed into conversations with the BIC staff, most often Pete Pietraszewski.

During data collection, Mendoza library resources increased my ability to locate relevant data. My foundational dataset is composed of "generic" futures prices, which are constructed from COMEX exchanges price data. By carefully inspecting its Bloomberg documentation, however, I discovered that the generic copper data was constructed differently than the oil data used by my

advisor. The copper data was constructed using prices from only five COMEX-designated “active contract” months, instead of all twelve months. This discrepancy forced me to reason through the implications of constructing the generic data differently. The five-month method would mismatch my data, which had monthly frequencies for every other variable. This discrepancy in construction would invalidate my regression analysis, so I needed to base the generic data on every month.

After I consulted Pete about adjusting Bloomberg’s construction of the generic copper data, she put me on the phone with the Bloomberg Help Desk using her librarian ID. The first technician I spoke with said that because the CME only designates five active contract months, the copper data simply could not be constructed based on twelve months. However, because the oil data is also from the CME, I knew that there must be way to do so. I called back. After two more phone calls with the Bloomberg Help Desk, I was able to pinpoint which obscure Bloomberg function (GFUT) needed to be adjusted to construct the generic copper data based on “serial,” or monthly, contracts. This allowed me to correctly construct the foundational dataset for my analysis of the copper market.

Critically evaluating data was another skill I developed using the Mendoza library resources. Combining data from six different sources increased my fluency in summarizing and checking data for consistency, and one particularly important example was my copper spot price data. At first, I used spot data from GFD, which I combined with Bloomberg futures. However, while evaluating the data from GFD and Bloomberg in tandem, I uncovered a contradiction. Economic theory holds that in a perfect market the spot price should follow the futures curve. But the spot price was did not follow the futures curve. With Pete’s guidance, I expanded my checks to other major copper spot price datasets, none of which were in line with the futures curve either.

However, by searching JSTOR for papers mentioning copper spot prices, I discovered copper spot market volume is exceptionally low, since commercial buyers use futures markets. This lack of liquidity means the market is not deep enough to produce a reliable price. Thus, I determined that the front-month futures contract, which is traded more frequently, would proxy for the spot price. Newly curious about the commodity literature’s data sources, I dove back into the data appendices of the JSTOR papers to investigate their data reliability. What I learned is that finding reliable spot prices is a recurring problem in the literature, and that a significant strain in the literature already proxies spot prices with front-month futures. My data evaluation uncovered an

issue with the spot price data (which, unbeknownst to me at the time, was a common problem) and then helped me arrive on my own to a solution. This confirmed using the front-month proxy, and taught me to proactively and relentlessly investigate the accuracy and consistency of data.

In addition to its six data sources, my thesis ended up citing 48 papers. Many of my other materials were articles or datasets accessed online through Mendoza's data and journal subscriptions. The CRC's GitHub version control workshop I attended gave me the principles to methodically systemize my data storage and analysis. That same workshop taught me how to code in Python, whose syntax is very similar to MATLAB. The skills I honed using the MATLAB and Stata software on the library system's computers allowed me to execute my quantitative analysis. Looking back, Mendoza's library resources were essential to my project—not only its data, articles, and software, but principally the excellent librarians like Pete who taught me how to be a researcher.