Econ 280 Replication Project: Slope of the Phillips Curve

Martha Fiehn

December 6, 2024

For the final part of the replication project, I wanted to re-run the analyses including the data for a state that the authors had dropped. Ultimately, I was unable to find the missing data in the files provided, so I will instead detail the process of searching for it.

In their detailing of the State-Level Price Index Construction Footnote 17 on p.1326 of Hazell et al. (2022) reads "We drop Arizona because of anomalous trends that we have not been able to investigate due to COVID-19 related access restrictions at the BLS." The footnote is to a sentences that indicate the authors were "able to approximate the official BLS data very closely [...] even for the pre-1988 period when we rely on the micro-data recovered by Nakamura et al. (2018), which likely have greater measurement error."

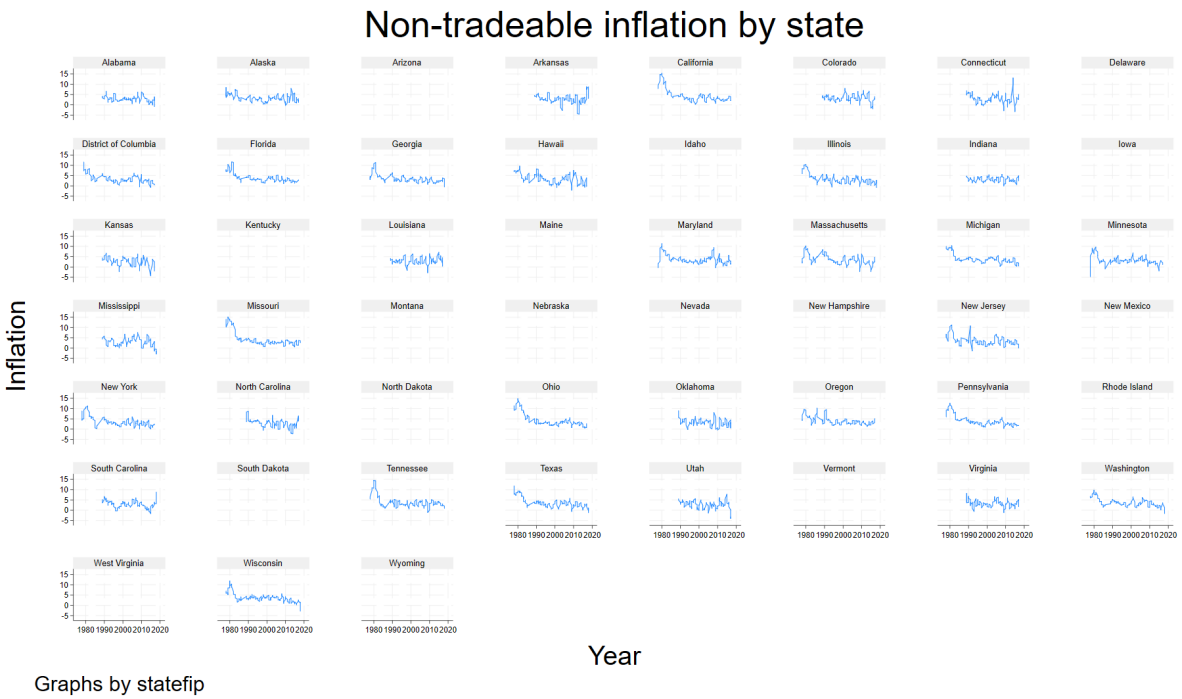
Deciding how to handle missing or anomalous data is one of the most interesting parts of data manipulation, and in some cases, can be where the bodies are hidden. Therefore, I was naturally curious.

In the paper's README, it lists the following variables for the main estimation:

- year
- quarter
- date
- mean_une: mean unemployment rate
- qt bartik_sa: seasonally adjusted tradeable demand spillover instrument
- state: name of the state
- statecode: encoded variable for state
- constant: a variable filled with ones used for the specifications without fixed effects
- infl_reg: non-tradeable inflation
- rp: relative price of non-tradeables

Since the paper references dropping Arizona in the section about constructing a state-levelprice index and that their main empirical analysis focuses on nontradeables, the *infl_reg* variable seems like the most useful starting point. Figure 1 shows the non-tradeable inflation by state. Interestingly, the data is not only missing for Arizona, but for Delaware, Idaho, Iowa, Kentucky, Montana, Nebraska, Nevada, New Hampshire, New Mexico, North Dakota, Rhode Island, South Dakota, Vermont, West Virginia, and Wyoming as well. The paper discusses missing prices, but it appears to be missing prices for *individual products* rather than entire states. I am not sure what to make of this.

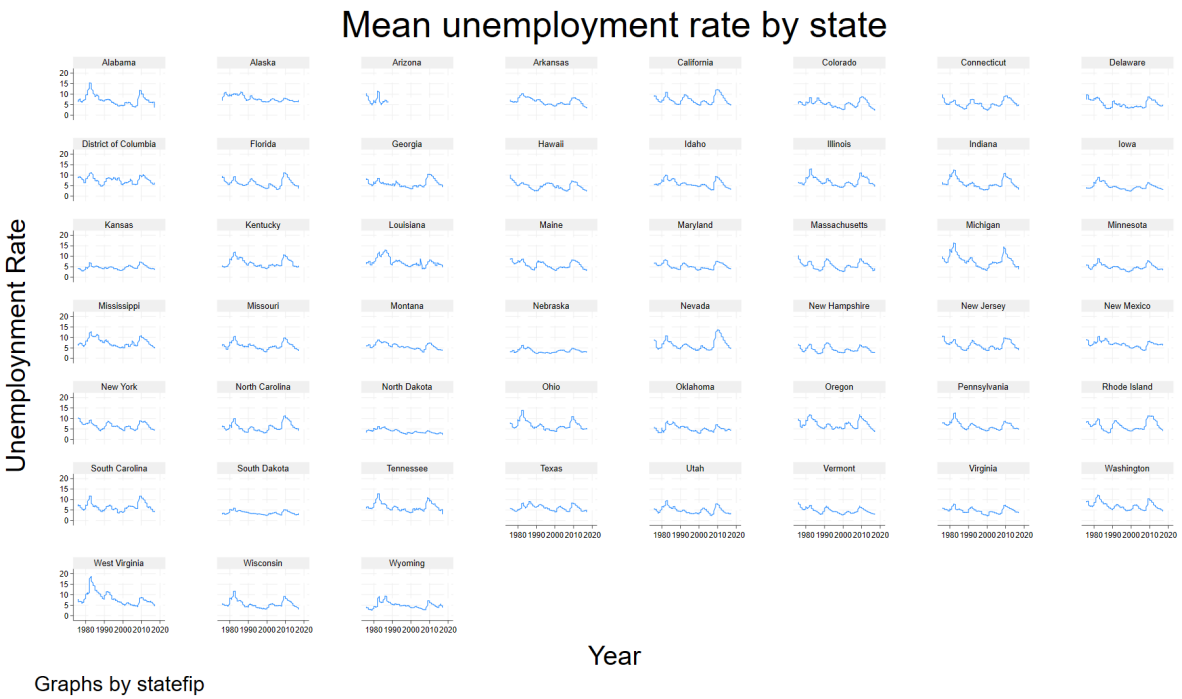
FIGURE 1



The mean unemployment rate is another key variable in the dataset. Figure 2 shows the state-wide evolution of this variable. Here, Arizona is the only state with clearly missing data. Since both the price index and the unemployment rate is calculated using BLS data, perhaps it is this variable to which the footnote refers.

Irrespective of the variable, I aimed to identify where in the code the data for Arizona was dropped so that I could "undo" that step and re-run the analyses. So I went through the various .do files and .dta files. It is very possible that I missed something, but I did not find the place where this occurs, and I did not find any data files that contain the Arizona data. Next, I decided to look at the replication package for Nakamura et al. (2018) since it was mentioned in the sentence associated with the footnote. The cleaning code was in SAS, which is not a language I am familiar with. I did briefly scan the .txt version of the code for anything that looked like dropping a state and was not successful. I also found no mention of Arizona in the paper itself, which suggests that the data dropping occurred only in Hazell et al. (2022).

FIGURE 2



If I were working on a project and really needed the data, my next steps would be to:

1. Carefully run every section of the replication code to see if there's anything I missed
2. Inspect the data files even more thoroughly
3. Familiarize myself with SAS and go through the cleaning code carefully
4. Look for the original BLS data the authors used (i.e., go straight to the source)
5. Email the corresponding co-author of the paper to learn more about their process. It is, for instance, that certain data were not included in the replication package due to limits on data sharing.

I am sure there is a perfectly innocent explanation (including user error) that explains the missing Arizona data, and I'd love to have been able to spend more time investigating!

References

Hazell, J., Herreño, J., Nakamura, E., & Steinsson, J. (2022, August). The Slope of the Phillips Curve: Evidence from U.S. States. *The Quarterly Journal of Economics*, *137*(3), 1299–1344. doi: 10.1093/qje/qjac010

Nakamura, E., Steinsson, J., Sun, P., & Villar, D. (2018, November). The Elusive Costs of Inflation: Price Dispersion during the U.S. Great Inflation*. *The Quarterly Journal of Economics*, *133*(4), 1933–1980. doi: 10.1093/qje/qjy017