

Meeting

November 8, 2023

1 Updates

This week I focused on processing the data and I finished understanding their code. I implemented the detrending and adjusting procedure they described, and I think I now fully understand how to implement their method. However, while processing the data I came across some issues.

1. First, the data they use provides the average bushels per acre in each county in Illinois. However, in their paper they don't specify how they compute the loss from this. It's not in the code they released either. The data they shared in the replication package has a positive loss in every year, except for one year in which the loss was 0. I think they took the year with the highest yield as the baseline, and measured loss of other years relative to that year. So, $\ell_t = y_{\max} - y_t$. Where ℓ_t is loss in year t , y_{\max} is the maximum yield observed in the data, y_t is yield in year t .
2. I couldn't get the number of observations to match, the data has 9 more observations than stated in the paper.
3. They didn't specify the robust regression method they used to detrend the data, and neither of the ones I tried seemed to match their data.

I will define loss as I described above and continue with implementing their method. The next steps for this week will be:

- Implement the NN method they describe.
- Make a copy of the project on the server or Google Colab so that we can have access to a GPU.

Once I finish implementing their method, I will compile all of my questions and draft an email to ask the original authors about them.