Meeting

November 2, 2023

1 Updates

This week I was focused on familiarizing myself with the data and code available in the Chen et al 2023 replication package. On the code side, I think I mostly understand the code well enough to implement their method, I just need to understand the algorithm they use to enforce the price constraint. On the data side, I realized that the data they published in the replication files is only a random sample of 1000 observations from their data. However, I have been able to locate the original data online. I will spend the next week downloading and cleaning the original data.

2 Data Processing

Chen et al mention that "To ensure stationarity of the loss experience over the long sample period, we first detrend the crop yields data with a second-order polynomial, estimated with a robust regression technique. Next, similar to Deng et al. (2007) and Harri et al. (2011), we adjust historical yields data to the 2018 level and other data heteroscedasticity." However, they don't mention what robust regression method they use, or what method they use to adjust the historical yields. This step does not appear in their code. From what I looked up, the two most popular robust regression methods are Huber regression and RANSAC. I was going to try both and see which one matches their data better.

Suppose that $y_t = f(x_t, g_t)$ is the value of our time series at time t, where x_t is the detrended value and g_t is the value of the trend at time t. From what I looked up, there are two ways to detrend a time-series after fitting a trend line. Chen et al don't mention which method they used, but the papers they mentioned in that section (Deng et al 2007 and Harri et al 2011) both use the second method, albeit by fitting a log-linear trend line instead of a second-order polynomial. Based on this, I was going to use method 2 and a second order polynomial.

Method 1: This method calculates the detrended value by removing the trend at each time step. In other words, $\hat{x}_t = y_t - \hat{g}_t$. To adjust an observation from time t to t' levels, one would first detrend that observation, and then add the trend component for t'. In other words $\hat{x}_t^{t'} = y_t - \hat{g}_t + \hat{g}_{t'}$.

Method 2: This method calculates the detrended value by dividing the time series by the trend line. In other words, $\hat{x}_t = \frac{y_t}{\hat{g}_t}$ To adjust an observation from time t to t' levels, you multiply the observation at time t by the ratio of the trend at time t' and time t. In other words, $\hat{x}_t^{t'} = y_t \frac{\hat{g}_{t'}}{\hat{g}_t}$.

So, in summary, next steps for this week are:

1. Finish downloading original data

2. Implement detrending and adjusting

3. Implement their method.