Identifying and Comparing Change Points in Stock Market Data

Zhicheng Ji

1 Introductions

Change point analysis has been widely applied to stock market data to detect the occurrence of significant changes in stock prices. In most cases, change point can be defined as the the time point where the stock price changes from increasing to decreasing or from decreasing to increasing. Accurate predictions of change points can provide great insights for making investment decisions. For example, one can buy in stock if the trend of stock price is predicted to change from decreasing to increasing in the near future. Most of the current studies focus on developing rigorous statistical methods to identify change points for an individual stock [Andrews, 1993; Bai, 1997; Chen and Gupta, 1997]. However, in many situations one needs to compare change points for multiple stocks to discover whether there is any relationship between the trends of different stocks. An easy-to-use method to identify and compare change points for multiple stocks is lacking. To address this problem, we aim to develop a systematic method to identify change points and their confidence intervals for multiple stocks. The change points will be listed in a sequential order and a statistical method will be developed to determine whether one change point occurs earlier than the other.

2 Data

The data used in this study is the end of day prices for 4 US stock (American Express Company, AXP; Procter & Gamble Co., PG; Pfizer Inc., PFE; General Electric, GE) obtained from Quandl (https://www.quandl.com/data/EOD?filterSelection=sample&keyword=). To simplify the problem, only the stock prices in the January of 2015 are included.

3 Methods

For each stock, a smooth curve is first fitted to capture the temporal trend of the stock price. The pattern and potential change points are then identified. Finally, the confidence interval of the change points are obtained using a permutation method. The details of the methods are as follows:

1. Fit curve. Fit a Generalized Additive Model (GAM, gam function in mgcv R package) with cubic smoothing splines.
2. Identify the pattern: monotone increasing, monotone decreasing, increasing then decreasing, etc. For patterns other than monotone changes, identify the position of the change points and proceed with the following steps. The change points are defined as the stationary points in this study.
3. Identify the confidence intervals of the change points. First obtain the fitted values and residuals from model 1. Then permute the residuals, add back to the fitted values and refit the GAM to obtain the permuted pattern. Repeat the procedure 1000 times and check how many times the permuted pattern agrees with the original pattern. If less than 75% of the permuted pattern agrees with original pattern, the pattern is changed to "indefinite". Otherwise for all the permutations where the permuted pattern is the same as the original pattern, record the position of the permuted change points. The confidence interval of the change point is defined by the 0.025 and 0.975 quantile of the permuted change points.

4. Compare change points. For a give set of data, list all the change points in a sequential order. To test whether one change point occurs significantly earlier than the other, simply check whether there is overlap between the two confidence intervals. If there is no overlap, then one change point occurs significantly earlier than the other.

4 Results

We first applied our method to detect the change points and their associated confidence intervals for the four stocks. The results are shown in Figure 1.

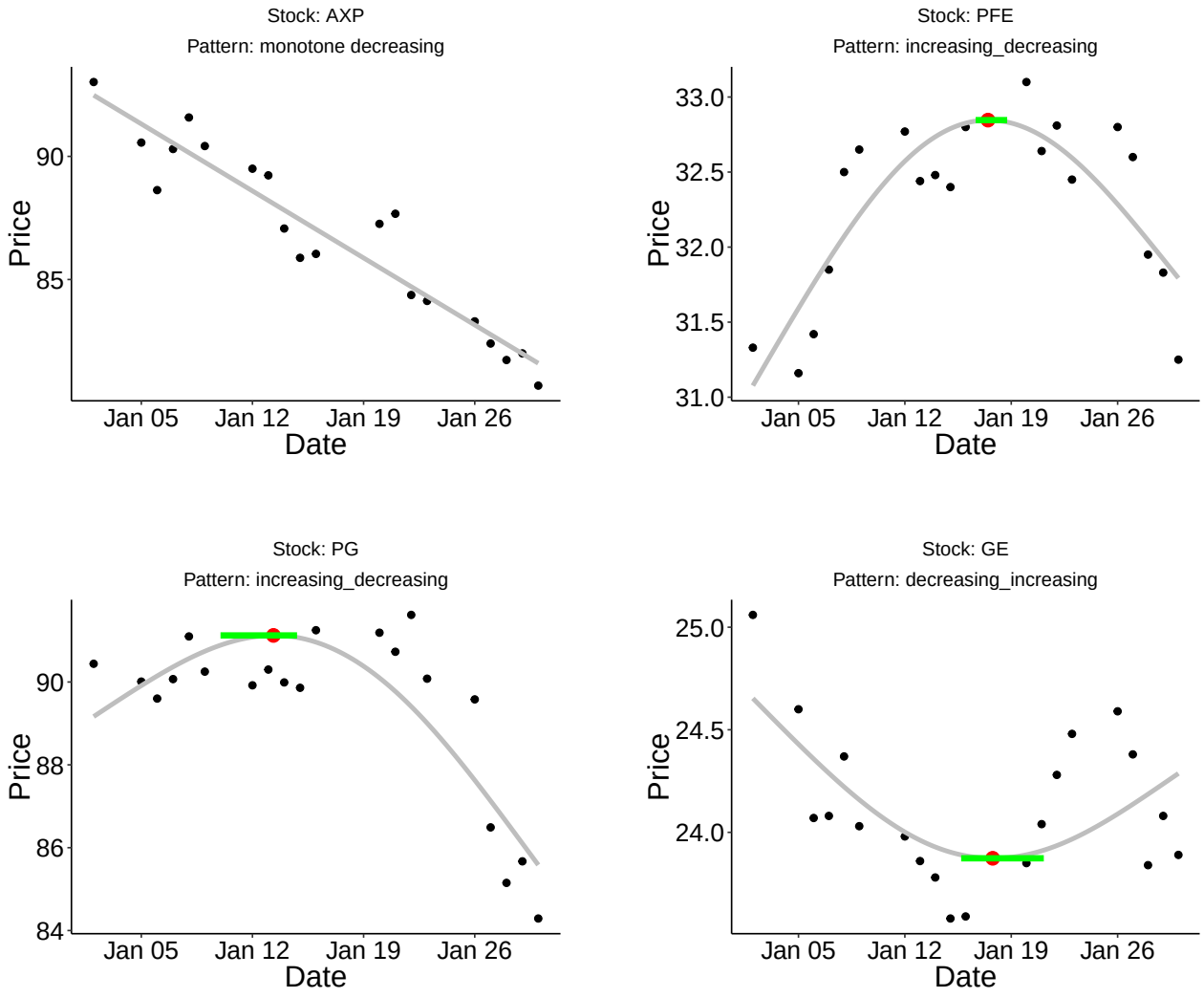

Figure 1: Chang points and confidence intervals for the four stocks. Grey lines are the fitted GAM smooth curves. Red points are the detected change points. Green line segments are the confidence intervals for the change points.

We can see that the temporal patterns and the change points of four prices are correctly identified. AXP shows a monotonically decreasing trend, GE shows a decreasing then increasing trend, and PFE and PG show increasing then decreasing trend.

For the three stocks with change points (PFE, PG, and GE), we can further compare whether one change point occurs significantly earlier than the other by checking whether their confidence intervals overlap. Figure

2 is a heatmap combing the fitted values and the change points of the three stocks. We can clearly see that the change point of PG occurs significantly earlier than those of PFE and GE, while the change points of PFE and GE do not show clear temporal order.

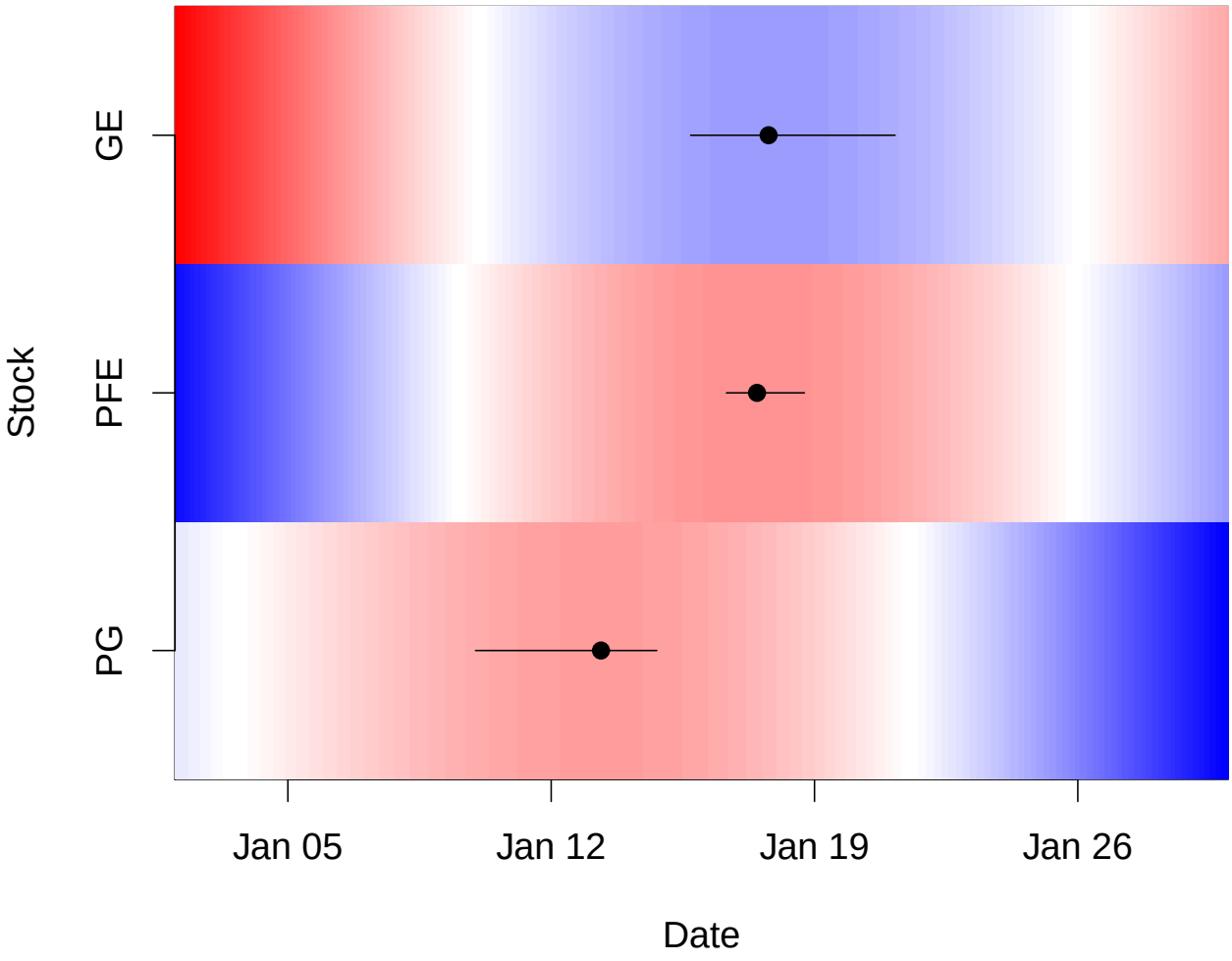

Figure 2: Heatmap showing the fitted values and the change points of the three stocks: PFE, PG, and GE. The change points and their confidence intervals are shown as black dots and line segments.

5 Discussions

In this study we develop a systematic method to detect the change points and their confidence intervals for the stock market data. We are able to order the change points in a temporal order and check whether one change point occurs significantly earlier than the other by comparing the confidence intervals. A real data example shows the method works well.

For now we are only considering comparing the change points in a fixed time window (i.e. January of 2015). In the future we need to consider multiple time windows (e.g. a sliding time window across the whole year) to see whether the observed temporal order of change points in one time window also frequently occurs in other time windows. This may provide important insights for investments: if the change point of stock A always occurs significantly earlier than that of stock B, by observing the trend of stock A we can predict the future trend of stock B and make investment decisions accordingly.

References

Andrews, Donald WK. Tests for parameter instability and structural change with unknown change point. Econometrica: Journal of the Econometric Society (1993): 821-856.

Bai, Jushan. Estimation of a change point in multiple regression models. Review of Economics and Statistics 79.4 (1997): 551-563.

Chen, Jie, and Arjun K. Gupta. Testing and locating variance changepoints with application to stock prices. Journal of the American Statistical Association 92.438 (1997): 739-747.