Project 5: Local Variance and Jump Regressions

Instructions

Project 5 is due on October 18th by 10:00 pm. This is a hard deadline, so no exceptions. You must push your local repository back to GitHub before the deadline. Your repository must contain:

- The Matlab code you used to complete the project;
- A script named `main.m` file that generates all required plots;
- A `report.pdf` file with your answers to the project questions. The report must also contain an Appendix with the code used to solve the project;
- All plots in the report must be self-contained. Self-contained means that a reader who only sees your figure (image and caption, but not the surrounding text) can understand what you are plotting. This translates to all plots having axis titles, correct units on the axis, and a caption that summarizes what is plotted.

This project makes use of stock data. Refer to the Data page for instructions on how to download the data and which files to download (requires Duke login). You must complete all exercises for both of your stocks using the data at the 5-minutes sampling frequency, unless stated otherwise.

You can obtain the repository for this project by clicking **on this link**.

Questions

The purpose of this project is to understand how to estimate the spot volatility (also called local volatility) and estimate jump betas. You will also conduct valid inference for the jump betas, which in turn depends on being able to estimate the spot volatility of residuals.

Exercise 1 - Local Variance

The purpose of this exercise is to estimate the local variance under the assumption that the variance process is continuous. And then verify whether a pattern in volatility is indeed present in the data.

A.

The local variance process c_t is a process defined over $[0, T]$. While we cannot estimate the value of c_t for all times $t \in [0, T]$, we can estimate the value of c at intervals containing t. Remember that we index the interval which contains t by i_t:

$$i \in \mathbb{N} : t \in ((i-1)\Delta_n, i\Delta_n]$$

Write a Matlab function that computes the local variance estimator for c_t. First, assume that the variance process is continuous. Hints:

- Start by coding the case where you want to estimate the local variance in the middle of the day;
- Extend the function to estimate the local variance at the beginning or end of the day. You can use i_t and n to determine whether you are at the beginning or end of the day. You might want to use the `max` or `min` function here.
- Check that your function still works for the middle of the day.

Select a day in your sample and estimate the local variance for all t of that day, using $k_n = 11$. Plot and interpret your results.

B.

Use the function you created before to estimate the local variance for all times and days in the sample. Then, compute the average local variance across time for each interval:

$$\bar{\hat{c}}_i \equiv \frac{1}{T} \sum_{t=1}^{T} \hat{c}_{i_t}$$

for $i = 1, 2, \ldots, n$.

Plot the values of $\bar{\hat{c}}_i$. Do you observe any pattern in the plot? Interpret. How does it relate to the time of day factor?

Exercise 2 - Jumps in the Variance

A possible way to interpret jump returns is to relate jumps to the release or incorporation of some unexpected information. If this unexpected information causes a jump in the returns, it is plausible that it would also lead to a jump in the volatility. For example, very bad news might make investors more uncertain, which would also increase the volatility of prices.

The purpose of this exercise is to analyze what happens to the volatility when there are jumps in returns.

A.

To begin, identify the time of all jumps in your stocks. That is, run the jump detector and obtain all intervals where we believe a jump return occurred. Let's index these intervals by i_t.

Estimate the local variance at the intervals you identified. Now, the assumption that the variance process is continuous is not plausible, and we need to work under the assumption that the variance process is discontinuous. Therefore, for each jump you will obtain two spot volatility estimates, one using returns from the left of the interval, and another using returns from the right of the interval.

Create a plot containing the following:

- The estimate of the volatility before the jump: $\hat{c}_{i_t}^{-}$
- The estimate of the volatility after the jump: $\hat{c}_{i_t}^{+}$
- The magnitude of the jump return: $\left|r_{t,i_t}^{d}\right|$

Interpret the results. Is there any indication the volatility jumps when there are jumps in returns? Is there any relation between the moves in the volatility and the magnitude of the jump returns?

Hints:

- Separate the plots of the volatility from the plot of the jump return. Use a single figure with two plots. On the top you can plot the two spot volatility estimates, on the bottom you can plot the absolute value of the returns (a bar plot will make things visually clear).
- How can we interpret the magnitude of a jump return?

B. (Optional, PhD Required)

Use the bootstrap to compute 99% confidence intervals for $\hat{c}_{i_t}^{-}$ and $\hat{c}_{i_t}^{+}$. To do so, draw random samples with replacement from the returns around each of the intervals where there was a jump return. Use the new draws to obtain new estimates of $\hat{c}_{i_t}^{-}$ and $\hat{c}_{i_t}^{+}$. Repeat the process 1000 times. Then construct the confidence intervals for both spot variances.

Repeat the plot from the previous question, but now include the confidence intervals. Compute the number of times the confidence intervals for $c_{i_t}^{-}$ and $c_{i_t}^{+}$ do not intersect (indicating a jump in the spot volatility). Interpret the results.

Exercise 3 - Jump Regression (Jump Beta)

The main idea of jump regressions is the opposite of what was done with the realized beta. Instead of discarding the jump data and using only the diffusive returns, to run jump regressions we discard the diffusive returns and keep only the data points at which the market index jumped.

Suppose the data are (X_1, X_2) where X_1 is the market index, as given by the S&P500 index (SPY), and X_2 is one of your stocks. Notice that if one of your original stocks is SPY then you need to download a third stock for this project. If this is your case, indicate which alternative stock you are going to use.

The jump regression model is:

$$\begin{aligned} \Delta X_{2,t} &= \beta \Delta X_{1,t} + \Delta J_{2,t} \\ \Delta X_{1,t} \Delta J_{2,t} &= 0, \forall t \in [0, T] \end{aligned}$$

where T is fixed. This model is the CAPM from finance applied to the jump moves.

We do not observe the jumps directly, but we can identify the intervals over which the market jumps. The index of the market jump intervals are:

$$\mathcal{I}'_n \equiv \left\{ i : |\Delta_i^n X_1| > \alpha \Delta_n^{0.49} \sqrt{\tau_i BV_t} \right\}$$

The list of integers $\mathcal{I}'_n$ is just a list of the subset of the nT intervals where the market is detected to have jumped. There are P_n elements in $\mathcal{I}'_n$ (the detected market jumps), which we denote by i_p for $p = 1, 2, \ldots, P_n$.

A.

Run the jump detection on the market index to locate the jump indices i_p, using $\alpha = 5$. Report the number of jumps per year and the average magnitude of the jumps each year.

B.

Make a scatter plot of the stock returns at the jump times, that is $\Delta_{i_p}^n X_2$ against the market jump returns $\Delta_{i_p}^n X_1$. Does a linear regression appear plausible?

C.

The OLS jump regression estimator is:

$$\hat{\beta} \equiv \frac{\sum_{p=1}^{P_n} \Delta_{i_p}^n X_1 \Delta_{i_p}^n X_2}{\sum_{p=1}^{P_n} (\Delta_{i_p}^n X_1)^2}$$

Compute the OLS jump beta for each of your stocks. Report the estimated values and interpret.

D.

Add to the scatter plot a regression line $y = x\hat{\beta}$ for $1.3 \times \min_p \Delta_{i_p}^n X_1 \leq x \leq 1.3 \times \max_p \Delta_{i_p}^n X_1$. Use a fine grid for the values of x. The 1.3 just extends the line over the domain so that the plot looks better.

E.

Given the jump beta estimate, $\hat{\beta}$, compute the local residuals: $\hat{E}_t = X_{2,t} - \hat{\beta} X_{1,t}$. Estimate the local variance of the residual process using the local variance estimator at the jump times. Use $k_n = 11$ to compute $\hat{c}_{e,t_p}$ for $p = 1, 2, \ldots, P_n$. Compute:

$$\hat{V}_\beta = \frac{\sum_{p=1}^{P_n} (\Delta_{i_p}^n X_1)^2 \hat{c}_{e,t_p}}{(\sum_{p=1}^{P_n} (\Delta_{i_p}^n X_1)^2)^2}$$

Report the value of $\sqrt{\Delta_n \hat{V}_\beta}$.

F.

Create a 95% confidence interval for the jump beta. Report the confidence interval.

G.

Imagine you work in a hedge fund and have an account holding $200 million worth of a particular stock. You are concerned about a possible market jump in the near future. Suppose you can trade market index futures (SPY futures). How can you hedge your position using futures? Based on the confidence intervals what is the range of values you need to short sell in order to hedge your position?

H.

One of the underlying assumptions is that the jump beta is constant. Let's test whether that is plausible. Split your sample into two periods:

- Period 1: 2007-2011
- Period 2: 2012-2017

Estimate the jump beta for each of the periods: $\hat{\beta}_1$ and $\hat{\beta}_2$. Compute $\widehat{V}_\beta$ for both periods: $\widehat{V}_{\beta_1}$ and $\widehat{V}_{\beta_2}$.

I.

It is possible to create a 95% confidence interval for the difference of the jump betas:

$$CI(\beta_1 - \beta_2) = \hat{\beta}_1 - \hat{\beta}_2 \pm 1.96\sqrt{\Delta_n \hat{V}_{\beta_1 - \beta_2}}$$

We need to compute $\hat{V}_{\beta_1 - \beta_2}$. The fluctuations in the jump betas are generated by the diffusive parts of the process, and so are independent across the two periods. Thus, we can compute $\hat{V}_{\beta_1 - \beta_2}$:

$$\hat{V}_{\beta_1 - \beta_2} = \hat{V}_{\beta_1} + \hat{V}_{\beta_2}$$

Compute the confidence interval above. Does the confidence interval contain the number 0? Interpret your findings.