Project 5
Due February 23, 2016

This project concerns both jump regression and diffusive regression. Use the same 5-min data on the ETF SPY (tracks the S&P 500 Market Index), 2007–2015 on the course website. Merge these data with your stock data. Note, the merge must be exact or subsequent research output will be corrupted. Be sure to delete the data for the two days of the flash crashes: 2013-04-23 and 2015-05-06.

1. (Jump Regression, discontinuous volatility). In class we gave the formulas for the asymptotic standard error of the jump regression OLS estimator $b_{OLS}$ as a mixed normal. When the local variance is discontinuous, then inference has to be done differently. The user needs to estimate the left and and right variances of the residual, $\hat{c}_{-t_p}^e$, $\hat{c}_{+t_p}^e$ and do the simulation business. Draw $Z_{1p}$, $Z_{2p}$, $\kappa_p$ from $N(0,1)$, $N(0,1)$, $U(0,1)$, and compute

$$\tilde{b} = b_{OLS} + \frac{1}{\sum_{p=1}^{P_n} (\Delta Z_{1p}^n)^2} \sum_{p=1}^{P_n} \Delta Z_{1p}^n \left( \sqrt{\kappa_p \hat{c}_{-t_p}^e Z_{1p}} + \sqrt{(1 - \kappa_p) \hat{c}_{+t_p}^e Z_{1p}} \right)$$

(1)

Repeat 1000 times to get $\tilde{b}_1, \ldots, \tilde{b}_{1000}$, and use the quantiles to get 95% confidence intervals for you full sample jump regression estimates.

A Briefly describe the logic underlying the procedure in (1).

B For both stocks, compute the confidence intervals described above. Do they differ much from those computed under the assumption continuous volatility?

2. When you run jump regressions over sub-periods it is very important to use the full sample to compute the time-of-day (tod) and detect the jumps. The sub-period estimates of the tod will be very noisy, and you can find different jump times when doing it sub-interval by sub-interval instead of full sample.

A For one of your stocks, use data for each year 2007, 2008, ..., 2015 as a sub-period and compute the jump betas and the 95% confidence intervals as above year-by-year.

B Because of the unusual error structure above, formal tests of equality across sub-periods are a computational nuisance. Nonetheless, by just comparing the confidence intervals, can you see evidence or not for year-by-year variation the this stocks jump beta?

3. Now we want to undertake diffusive beta estimation. You will need to first compute the
continuous parts $r^c$ of your two stock and the market returns. For simplicity use $\alpha = 4.0$ across everything. Use $k_n = 11$ so there are 7 intervals per day.

A For each stock, use the blocking method to compute the most efficient year-by-year diffusive beta estimates as per class notes.

B Compute the asymptotic standard errors. Actually, if you follow the class notes exactly, then these standard errors will automatically be in applied form. (See note below.)

4. Now the big question: How do the jump betas by year from 2 above compare to the diffusive betas of 3? A formal test is too hard, but a check of whether the jump beta lies in the 95% confidence interval for the diffusive beta (and vice versa) would be informative. This issue is an unaddressed research question; do your best to summarize the evidence.

NOTE: A bit more on diffusive volatility estimation: Recall we had $\text{Var}(b_j) = V_{bj} = (1/k_n)c_{ee,j}/c_{zz,j}$. If we use weights $w_j$ then

$$\text{Var}(b_w) = \frac{\sum_{j=1}^{M} w_j^2 V_{bj}}{(\sum_{j=1}^{M} w_j)^2}$$

and this will automatically be the "applied form" of the asymptotic variance. To see this, consider the weights $w_j = 1$ as per estimator-2 of (Li et al., 2016). Then

$$\text{Var}(b_w) = \frac{\sum_{j=1}^{M} V_{bj}}{M} = \frac{1}{Mk_n} \frac{1}{M} \sum_{j=1}^{M} c_{ee,s}/c_{zz,s} ds$$

Now $nobs = nT$ and $nT = Mk_n$ (up negligible edge effects). Then

$$\text{Var}(\sqrt{nobs}(b_w - \beta)) = \frac{1}{M} \sum_{j=1}^{M} c_{ee,j}/c_{zz,j} \rightarrow \int_0^{nT} \frac{c_{ee,s}}{c_{zz,s}} ds$$

is the theoretical asymptotic variance exactly in accordance with the expression (Li et al., 2016) for $\Sigma(w_2)$ of Subsection 4.2. The same can be shown for the other estimators but the algebra is slightly more involved.

References