Econ 512 Lab 3
High Frequency Data

Eric Zivot

Due: May 24, 2002.

1 Exercises

1. Modification of Tsay (2001), chapter 5, exercise 2. .Let P_t be the observed market price of an asset, which is related to the fundamental value of the asset P_t^* via

$$P_t = P_t^* + I_t \frac{S}{2}$$

where $S = P_a - P_b$ is the bid-ask spread, and $\{I_t\}$ is a sequence of *iid* binary random variables (order type indicators) such that $I_t = 1$ with probability 0.5 and $I_t = -1$ with probability 0.5. Assume that ΔP_t^* for a Gaussian white noise process with mean zero and variance 1. Suppose that the bid-ask spread is two ticks.

   (a) What is the lag 1 autocorrelation of the price change series ΔP_t when the tick size is \$1/8?

   (b) What is the lag 1 autocorrelation of the price change series ΔP_t when the tick size is \$1/16?

   (c) What is the implication for the lag 1 serial correlation due to bid-ask bounce of the change to decimal reporting of prices?

2. Modification of Tsay (2001), chapter 5, exercises 4 and 5. The file `mmm9912-dtp.dat` contains the transactions data of the stock of 3M Company in December 1999. There are three columns: day of the month, time of transaction in seconds from midnight, and transaction price. Transactions that occurred after 4:00 p.m. Eastern time are excluded. You will find the examples in the file `econ512highFreq.ssc` helpful for this problem.

   (a) In S-PLUS, create a `timeSeries` object containing the data for 3M. Plot the prices over the first three days of trading.

(b) Create a new `timeSeries` containing the trade day, durations between trades in seconds and the price changes. Make sure to eliminate negative durations due to the overnight period. Compute summary statistics for the durations and price changes and comment. Plot histograms for the durations and price changes and comment. Compute the sample ACF for the durations and price changes and comment. Is there any evidence for the lag 1 autocorrelation induced by the bid-ask bounce?

(c) Is there a diurnal pattern in the durations for the 3M stock? To answer this question, you may want to create a new time series containing the average durations in 5 minute buckets and a new time series containing the number of transactions in 5 minute buckets.

(d) Use the data to construct an intraday 5 minute log return series (*Hint*: this is easy to do using the S-PLUS function `aggregateSeries`). Investigate the serial correlation properties of the 5 minute returns.

(e) There are 77 5 minute returns in a normal trading day. Some researchers suggest that the sum of squares of the intraday 5-minute returns can be used as a measure of daily volatility. Apply this approach and calculate the daily volatility of the log return of 3M stock in December 1999.(*Hint*: the S-PLUS function `aggregateSeries` can do this very easily) Discuss the validity of such a procedure to estimate daily volatility.

3. Modification of Tsay (2001), chapter 5, exercise 6. The file `mmm9912-adur.dat` contains an adjusted intraday trading duration of 3M stock in December 1999. The adjustment is done as follows. There are 39 10-minute time intervals in a trading day. Let d_i be the average of all log durations for the ith 10-minute interval across all trading days in December. Define an adjusted duration as $t_j/\exp(d_i)$, where j is in the ith 10-minute interval. Note that more sophisticated methods can be used to adjust the dirurnal pattern of trading duration. Here a simple local average is used.

   (a) In S-PLUS, create a `timeSeries` object containing the data for 3M. Is there a diurnal pattern in the adjusted duration series? Why?

   (b) Let x_i denote the adjusted duration at time t_i. A result from Engle and Russell (1998) is that quasi-maximum likelihood estimates of the parameters of an EACD(1,1) model for x_i may be be computed using standard GARCH software as follows: Use $\sqrt{x_i}$ as the dependent variable and set the mean of the conditional mean equation equal to zero. In `S+FinMetrics`, this may be accomplished using the function `garch` as follows

```
> eacd.fit = garch(sqrt(x)~-1, ~garch(1,1))
```

where `x` is a `timeSeries` containing the adjusted durations. Note, when computing standard errors make sure to compute the robust or sandwich standard errors.