5th set of SAS assignments

Price Discovery with cross-listed stocks

1. Data preparation

The SAS files `abytsx` and `abyny` contain quote data on the stock ABY for the Toronto Stock Exchange (TSX) and the New York Stock Exchange (NYSE). The SAS file `fx` includes quotes on the CAD/USD exchange rate.

1a) Read the three files into three temporary SAS datasets. Format the time and date variables, if necessary. Then use bid and ask quotes to create midquotes and keep only the time and date and midquote variables in each data set. Sort all three data sets and delete duplicate timestamps (i.e., quote changes that were recorded at the same time) using the `nodupkey` option in `proc sort`.

1b) Use `proc expand` to get the real time data in equally spaced intervals of 10 minutes within each trading day. This means that the current value is kept and written forth until there is a new value in the data set or, if there are several observations within a ten minute interval only the last observation is kept.

1c) Use `proc sql` to merge all three datasets together (Hint: instead of the `join` option use a `where` command and use three datetime variables to merge the datasets). Then convert the NYSE midquote into Canadian Dollars.

1d) Choose data for one day only and plot the TSX midquote and the converted NYSE midquote in a SAS graph.

2. Estimation

We want to estimate the following regressions:

$$r_t^{TSX} = \boldsymbol{\alpha}_{TSX}(midq_{t-1}^{TSX} - midq_{t-1}^{NY}) + \boldsymbol{\beta}_1 r_{t-1}^{TSX} + \boldsymbol{\beta}_2 r_{t-1}^{NY} + \varepsilon_{1t}$$

$$r_t^{NY} = \boldsymbol{\alpha}_{NY}(midq_{t-1}^{TSX} - midq_{t-1}^{NY}) + \boldsymbol{\beta}_3 r_{t-1}^{TSX} + \boldsymbol{\beta}_4 r_{t-1}^{NY} + \varepsilon_{2t}$$

2a) First create the three variables on the right hand side:
The first new variable is the lagged (!) difference between the log midquotes in both markets $(log(midq_{t-1}^{TSX}) - log(midq_{t-1}^{NY}))$. The second and third variable are the first lags of the log returns on the TSX (r_{t-1}^{TSX}) and on the NYSE (r_{t-1}^{NY}).

2b) Then delete overnight returns:
Create an indicator variable that equals one for the first observation of a day and zero otherwise. Then use this variable to delete the first observation of each day from the dataset.

2c) Estimate the two regressions. Merge the two resulting datasets with the estimated pa-

rameters together and caluclate the adjustment coefficient ratios:

$$Adj^{TSX} = \frac{|\alpha^{TSX}|}{|\alpha^{TSX}|+|\alpha^{NY}|} \text{ and } Adj^{NY} = \frac{|\alpha^{NY}|}{|\alpha^{TSX}|+|\alpha^{NY}|}$$

Interpret the results.

3. Turn your program code into a macro that keeps the input dataset and the frequency when expanding the data flexible. Call the macro using different frequencies and compare the results.