

Problem Set #2 : Time-Dependent Active List
--

WARNING: This problem set is very long. You have one week to do it. It is due on Thursday 25 October 2012 before class (10am). If you start two days before you'll never make it. You should start today.

1) Download from the intranet the variable `allstocks` that was given as the solution to Problem Set 1. It contains the universe of stocks we will study. Do not take the version of `allstocks` that you produced: everybody must start from the same universe. Let n denote the number of stocks in this universe.

2) Download the following data from DataStream for all the stocks in the universe and for all the days from January 1st, 1997 to December 31st, 2002. Let T denote the number of weekdays in this time range. Every variable is a large matrix of dimension $(T \times n)$:

- a) `price`: unadjusted share price in local currency,
- b) `tri`: total return index in Euros,
- c) `volume`: daily volume expressed in Euros,
- d) `book`: book value of equity in Euros (this will only change once per year, but still you must have a value for it every day so the matrix has the same size as the other matrices),
- e) `cap`: total market capitalization in Euros.

As a by-product, you will also obtain from DataStream a list of all the weekdays from January 1st, 1997 to December 31st, 2002. Transform it into the format `dd-mmm-yyyy` and store it in a variable called `myday`.

3) Download the database `TickSummary.mat`. It contains median monthly bid-ask spreads for a large list of stocks that has a substantial overlap with the universe in `allstocks`. Use it to build a bid-ask spread matrix for our universe. Fill in the missing data with `NaN` (Matlab value for Not-a-Number).

4) Using the IBES tickers associated with each DataStream `DScode`, pull down the history of analyst recommendation revisions for every stock. Construct a matrix of dimension $(T \times n)$ named `rec` that contains the total number of analyst upgrades minus the total number of analyst downgrades for any given stock on any given day. Most of the entries of this matrix will be zero.

5) Construct a boolean matrix `isactivenow` of dimension $(T \times n)$ that contains the time-dependent active list. The entry is equal to 1 if the stock is active that day, and 0 if it is inactive. The criteria to be applied are as follows:

- a) The stock must have valid price and total return index for at least 90% of the past 252 business days
- b) The stock must not be dead. A dead stock is a stock whose unadjusted share price in local currency stays the same for 10 days or more.
- c) The stock must have valid volume and total return index for at least 90% of the past 21 business days.
- d) The stock must have valid book value of equity and valid total market cap today.
- e) The stock must have valid bid-ask spread last month.
- f) The stock must have had at least one analyst recommendation revision in the past.

6) Further refine the matrix `isactivenow` by applying two liquidity filters. For a stock to be active, *both* liquidity filters must be satisfied. Do this stock by stock.

- a) At the beginning of each month, a stock passes the first liquidity threshold if:
 - it was active the previous month and its bid-ask spread is less than 1%
 - or it was inactive the previous month and its bid-ask spread is less than 0.80%
- b) At the beginning of each month, a stock passes the second liquidity threshold if:
 - it was active the previous month and its average daily volume over the past month was more than 1 million euros per day
 - or it was inactive the previous month and its average daily volume over the past month was more than 1.2 million euros per day.

Problem Set output: You must e-mail T.A. Kyle Matoba a Matlab database containing the variables:

1. `allstocks` (1×n)
2. `myday` (T×1)
3. `price` (T×n)
4. `tri` (T×n)
5. `volume` (T×n)
6. `book` (T×n)
7. `rec` (T×n)
8. `isactivenow` (T×n)

Grading: You will be judged by how complete and correct these variables are, especially `isactivenow`. To get `isactivenow` right you must get all the others right.

VERY IMPORTANT: Due to the limited time that the T.A. can devote to this class, every student is allowed at most one e-mail question to the T.A. – so use it wisely! In addition, the T.A. will hold one office hour which will be devoted to this problem set.