**Sub-project 1:**

* Classification Agreement Revisited (Agreement Ratio):

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| AMZN, 4/23 | Tick Rule | Tick Rule (2 periods back) | Quotes | LeeReady | Not classified (2 period back Tick) | Total # of Classifications |
| Tick Rule | 100.00% | 70.44% | 72.19% | 75.22% | 0.00% | 22983 |
| Tick Rule (2 periods back) | 70.44% | 100.00% | 47.57% | 53.50% | 29.56% |  |
| Quotes | 72.19% | 47.57% | 100.00% | 96.97% | 3.03% |  |
| LeeReady | 75.22% | 53.50% | 96.97% | 100.00% | 0.00% |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| AMZN, 4/24 | Tick Rule | Tick Rule (2 periods back) | Quotes | LeeReady | Not classified (2 period back Tick) | Total # of Classifications |
| Tick Rule | 100.00% | 73.65% | 72.45% | 76.77% | 0.00% | 16370 |
| Tick Rule (2 periods back) | 73.10% | 100.00% | 54.56% | 56.07% | 26.90% |  |
| Quotes | 73.65% | 54.56% | 100.00% | 96.88% | 3.12% |  |
| LeeReady | 76.77% | 56.07% | 96.88% | 100.00% | 0.00% |  |

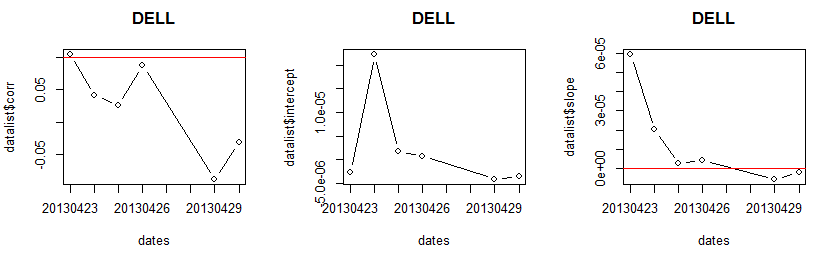
* There was a bug in two-period back tick rule, so we redid all SOI measures using quotes with 30ms delay
* Expand SOI and SQI measures for all S&P 100 (exclude BRK.B):

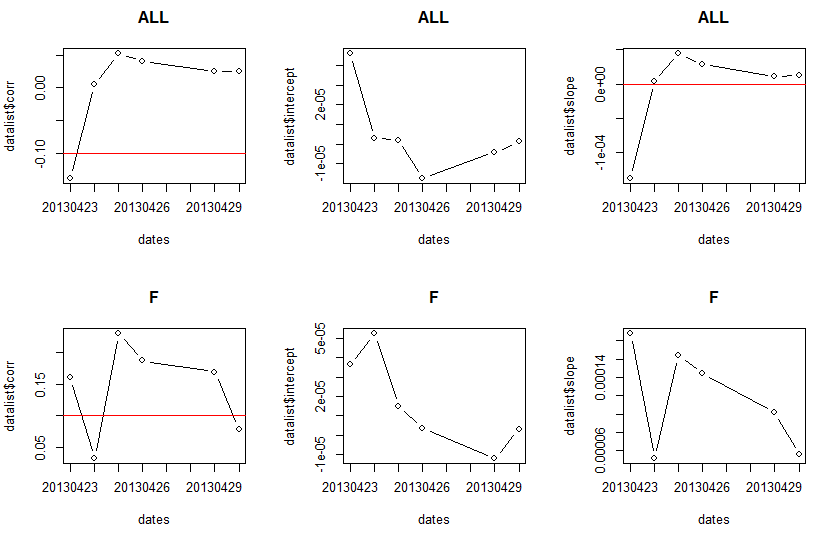
The results (both SOI & SQI) show consistent correlation of 0.1 – 0.2 ( ) for top results.

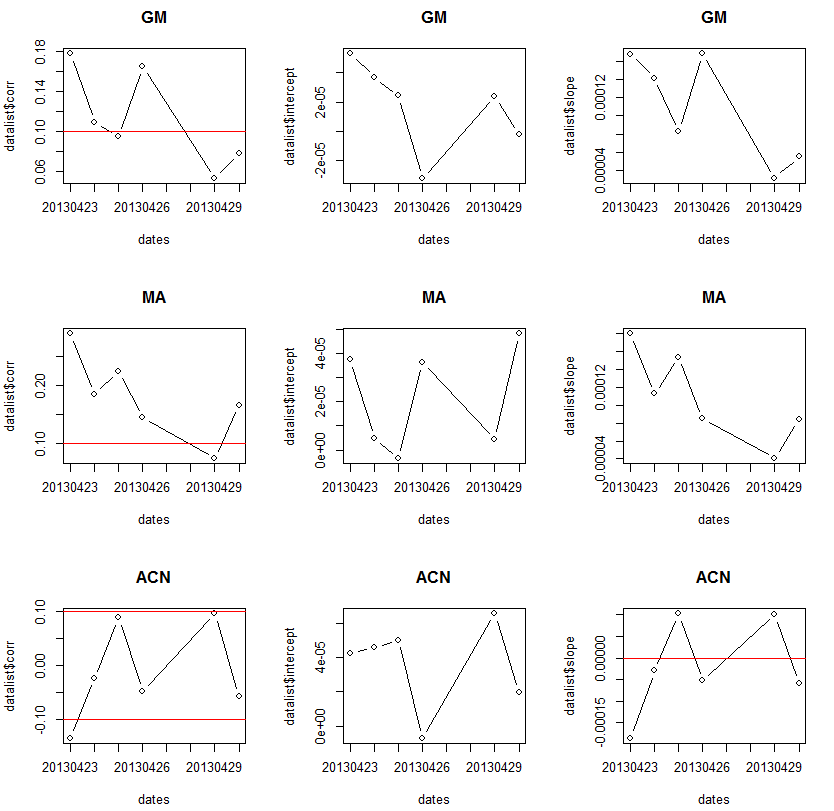
This contains 4/23/13 run for SOI forecast (using tick):

<https://cornell.box.com/s/1xlth90fmyz1bbrbey3o>

Stability of regression coefficients for TOP names:



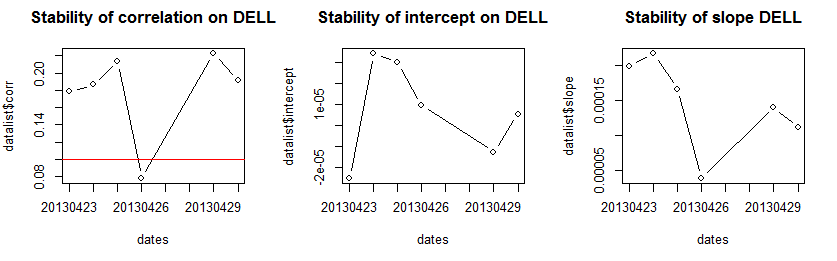


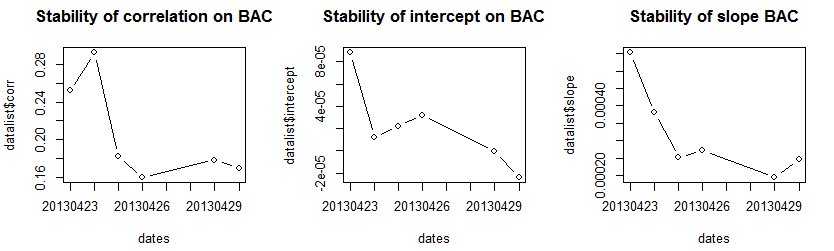


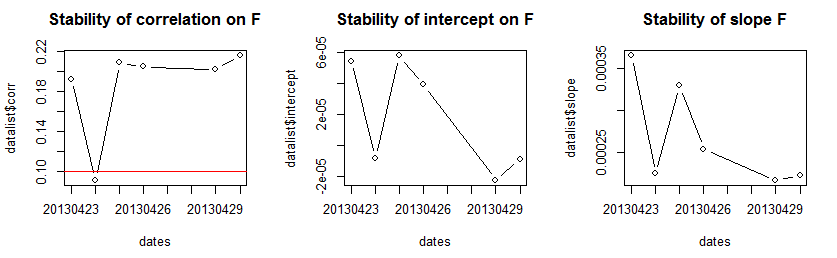
This contains 4/23/13 run for SQI forecast:

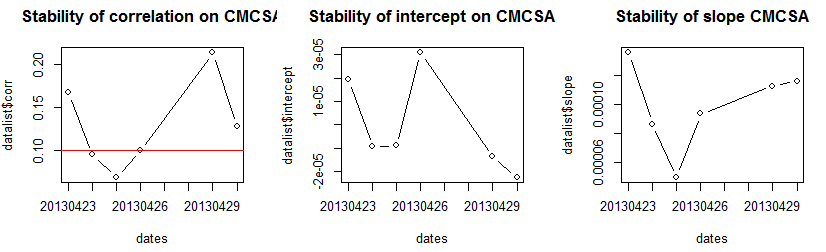
<https://cornell.box.com/s/it3he3d2expghjk2qcyj>

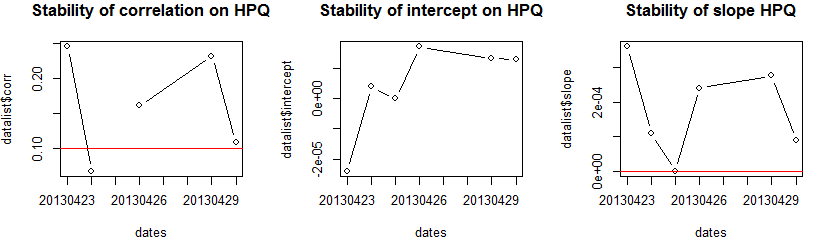
Coefficient stability tests for top correlation measures:











We will discuss the predictors’ effectiveness at tomorrow’s discussion.

* Concurrent regression (SOI, rerun using quotes classification for S&P 100):

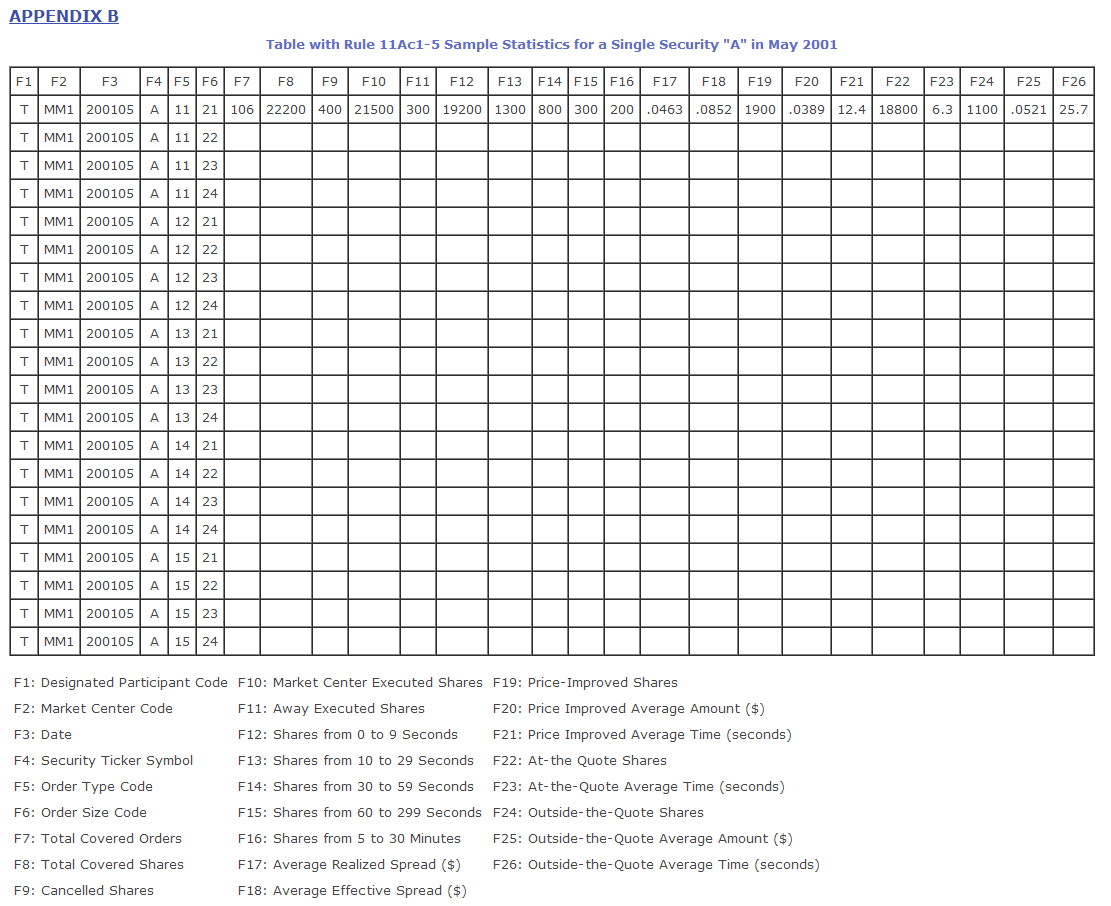
<https://cornell.box.com/s/px1aqo332pf1ndt31m77>

Consistent significance with R^2 = 0.2 – 0.3

* Two Factor Model (using SOI & SQI) and determine if both predictors have significance: still refactoring code in order for it to happen.

**Sub-project 3:**

Data Description



Reference: <http://www.sec.gov/interps/legal/slbim12b.htm>

Order Type Code (F5) Explanation:

11: Market orders

12: Market limit orders

13: Inside the quote limit orders

14: At the quote limit orders

15: Near the quote limit orders

Reference: <http://www.sec.gov/interps/legal/slbim12a.htm#q1> (footnote 6)

Summary Statistics:

*“What fraction of executed volumes (field F10) in each symbol come with market orders (code 11) vs marketable limit orders (code 12) and other limit orders (codes 13, 14, 15)?”*

*“What fraction of executed volumes in each symbol are executed with price improvement (F19) vs those executed at the quote (F22) or outside the quote (F24).”*

For UBS’s 09/13, the summary statistics for all the ratios are calculated, and the file can be retrieved from:

<https://cornell.box.com/s/8b2zfaqh9an5moh0l6h7>

A sample output is looks like, this file includes all the ratio for:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Ticker | Mkt order Ratio | Lim order ratio | Other lim order ratio | Price impr. Ratio | At Quotes Ratio | Outside Quotes Ratio |
| A | 0.4206966 | 0.464097 | 0.1152061 | 0.53056388 | 0.459855447 | 0.099498298 |
| AA | 0.3502736 | 0.416053 | 0.2336732 | 0.81575019 | 0.344683523 | 0.002646194 |
| AA PR | 0.6753247 | 0.324675 | 0 | 0.25974026 | 0.519480519 | 0.220779221 |
| AACC | 0.5929867 | 0.370927 | 0.0360864 | 0.75462163 | 0.975883318 | 0.135966662 |

*“Are the executed volumes in each symbol a consistent fraction of all D volume per symbol?”*

We will provide this answer later. We are still cleaning the data as some stocks traded on NYSE is not included in the .h5 files.