S&P 100 Tick Test, Quotes, LeeReady Classification Agreement:

|  |  |  |  |
| --- | --- | --- | --- |
| S&P100 Across Stocks (4/23/2013) | Avg. Agreement Rate (Tick Test & Quotes) | Avg. Agreement Rate (Tick Test & LeeReady) | Avg. Not Classified (Quotes) |
| Average | 62.12% | 72.55% | 10.43% |
| Std | 6.78% | 6.12% | 3.16% |
| Max | 75.23% | 87.06% | 18.96% |
| Min | 46.91% | 57.71% | 1.38% |

Using S&P 100 as a population, using LeeReady Classification for SOI, the two factor model yields (for 60s):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | coef1\_two\_factor | coef2\_two\_factor | r^2\_two\_factor | p value two factor |
| Average | 0.01% | 0.01% | **1.33%** | 0.2425 |
| Std | 0.01% | 0.00% | 1.24% | 0.3070 |
| Max | 0.06% | 0.01% | 6.47% | 0.9903 |
| Min | -0.01% | 0.00% | 0.01% | 0.0000 |
| Median | 0.01% | 0.00% | **0.97%** | 0.0843 |
|  | | | p-value % < 0.05 | **36.36%** |
| p-value % < 0.1 | **53.54%** |

The individual one factor model yield, with SOI as a predictor inferior in its predictive power and p-value significance:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| coef\_sqi | r2\_sqi | p\_sqi | coef\_soi | r2\_soi | p\_soi |
| 0.01% | **1.06%** | 0.2426 | 0.00% | **0.29%** | 0.4960 |
| 0.01% | 1.22% | 0.3071 | 0.01% | 0.42% | 0.3122 |
| 0.06% | 6.45% | 0.9903 | 0.01% | 2.39% | 0.9887 |
| -0.01% | 0.00% | 0.0000 | -0.01% | 0.00% | 0.0020 |
| 0.01% | 0.70% | 0.0840 | 0.00% | 0.10% | 0.5181 |
|  | p-value % < 0.05 | **36.36%** |  | p-value % < 0.05 | **7.07%** |
|  | p-value % < 0.1 | **53.54%** |  | p-value % < 0.1 | **11.11%** |

­­­

Aggregate all S&P 100 stocks to test the predicting power of SQI.

Cross Section Results:

Quote mid-returns by scaled down by the volatility of the corresponding tick name. The volatility is calculated as the log returns of all the trading data in one day as one sample, and averaged over five days. Since almost all the volatility of the five days’ samples is smaller than the mean in magnitude (except for APPL), the averaged mean should be a good approximate for the volatility of each stock.

R^2 and intercept are stable across five days:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **DATE** | **20130423** | **20130424** | **20130425** | **20130426** | **20130429** |
| **R^2** | 0.0048 | 0.0046 | 0.0044 | 0.0034 | 0.0057 |
| **Correlation** | 0.0690 | 0.0677 | 0.0660 | 0.0587 | 0.0754 |
| ***Coefficients*** | | | | | |
| **Intercept** | 1.389E-05 | 1.393E-06 | 5.911E-06 | -1.601E-06 | 4.444E-06 |
| **X Variable 1** | 6.474E-05 | 6.135E-05 | 5.951E-05 | 5.117E-05 | 5.249E-05 |

Better Visualization:

We separate SQI values range from -1 to +1 into 20 equal length bins, (i.e. ), and then for each bin plot the average value of returns of which the corresponding SQI falls in that bin.

Regression plot for different time length from 10s to 180s is shown below. The R^2 decreases and slope increases with time length. This is consistent with Sasha’s intuition:



The slope in the univariate regression is and ; and, we observe that the R^2 decrease and slope increases as time length increase (i.e. ), where x is SQI, and y is the quote mid return. The drivers behind this behavior is shown in the following chart:

|  |  |  |
| --- | --- | --- |
| Volatility on 20130423 | | |
|  | **SQI** | **Return** |
| **10s** | 0.44 | 0.000124 |
| **20s** | 0.44 | 0.000212 |
| **30s** | 0.44 | 0.000273 |
| **40s** | 0.44 | 0.000327 |
| **50s** | 0.44 | 0.000367 |
| **60s** | 0.44 | 0.000408 |

SQI almost never vary as time length increases, while the volatility of return is monotonously increasing. As time length increases, there will be more ‘noise’ in the return, which cannot be predicted by the SQI signal. Thus shorter time length SQI has better predictive power.

When performing the same experiment on the average of individual stocks’ levels (average over S&P 100), the same pattern is confirmed:

When we perform the same analysis on signed order imbalance (SOI), the results are not consistent with SQI:

**<ADD SOI cross-sectional>**