Quantathon Challenge at the University of Toronto
March 18-20, 2016

Summary
You are managing a long-short portfolio in 100 stocks with daily rebalancing. In Part 1 you are implementing a strategy provided to you and your returns are computed using closing prices each day. In Part 2 you are using open-to-close returns each day and you are provided a portfolio framework with a set of parameters that must be estimated by you. In Part 3 you are using the framework in Part 2 but the direction of your trades are constrained each day as per our instructions below. In Part 4 you are free to develop a more general portfolio strategy. The deliverables for each part are listed after Part 4. Please read the entire document before proceeding.

The Data
Daily data on 100 stocks is posted on Google Drive in the form of a single CSV file with the following content: The first column has the date in format yyyymmdd and then for each of the 100 stocks, you will find columns with:

1) The opening price (SO)
2) The intraday high price (SH)
3) The intraday low price (SL)
4) The closing price (SC)
5) The trading volume (number of shares traded) during the day (TVL)
6) The trade-direction indicator variable (IND)

Stock prices have been adjusted for stock splits and other corporate events and include dividends. The data provided corresponds to actual stocks but their tickers are not identified in the file. You are *not allowed* to use data other than that provided in the file.

Part 1: Implement a Specified Strategy

You are in charge of managing a long-short portfolio of 100 stocks. The strategy is defined as follows: For each day and for each stock compute the portfolio weight in stock j at the close of day t-1 using

(1) W1(t,j) = - (1/N)*[RCC(t-1,j) - AvrRCC(t-1)], for t=3,4 ..,T.

where T is the number of days in the sample, N is the number of stocks, and where AvrRCC(t-1) is the equal-weighted average close-to-close return across the N stocks on day t-1. Note that the weights, W1(t,j), sum to zero each day. The daily close-to-close return for stock j is computed as follows:

(2) RCC(t,j) = SC(t,j)/SC(t-1,j)-1, for t=2,3, ..,T.

where SC(t,j) denotes the closing price for stock j on day t. The realized close-to-close return on the portfolio is

(3) $RP1(t) = [\Sigma_j\{W1(t,j)*RCC(t,j)\}] / [\Sigma_j\{|W1(t,j)|\}]$,

where $\Sigma_j\{\cdot\}$ denotes the sum across stocks of the argument inside $\{\cdot\}$ and $|\cdot|$ denotes the absolute value. Your task is to implement the long-short strategy defined above.

Part 2: Optimize Performance over a set of Parameters

For each stock compute close-to-open return on stock j defined by

(4) $RCO(t,j) = SO(t,j)/SC(t-1,j)-1$

Where SO(t,j) denotes the opening price of stock j on day t. For each stock, also compute open-to-close return defined by

(5) $ROC(t,j) = SC(t,j)/SO(t,j)-1$

Finally, compute the open-to-open return for each stock defined as

(6) $ROO(t,j) = SO(t,j)/SO(t-1,j)-1$

You also need to compute the following range-based proxy for variance on day t

(7) $RVP(t,j) = [1/(4*\ln(2))]*[\ln(SH(t,j)) - \ln(SL(t,j))]^2$

where $\ln(\cdot)$ denotes the natural logarithm, SH(t,j) denotes the intraday high and SL(t,j) the intraday low of the stock price.

In Part 2 you are buying or selling at the opening price and then unwinding the position at the closing price on the same day corresponding to equation (5). Your portfolio weights must be of the form

(8) $W2(t,j) = a1*[RCC(t-1,j) - AvrRCC(t-1)]/N + a2*[ROO(t,j) - AvrROO(t)]/N$

$+ a3*[ROC(t-1,j) - AvrROC(t-1)]/N + a4*[RCO(t,j) - AvrRCO(t)]/N$

$+ a5*[TVL(t-1,j) / AvrTVL(t-1,j)]*[RCC(t-1,j) - AvrRCC(t-1)]/N$

$+ a6*[TVL(t-1,j) / AvrTVL(t-1,j)]*[ROO(t,j) - AvrROO(t)]/N$

$+ a7*[TVL(t-1,j) / AvrTVL(t-1,j)]*[ROC(t-1,j) - AvrROC(t-1)]/N$

+ a8*[TVL(t-1,j) / AvrTVL(t-1,j)]*[RCO(t,j) - AvrRCO(t)]/N

+ a9*[RVP(t-1,j) / AvrRVP(t-1,j)]*[RCC(t-1,j) - AvrRCC(t-1)]/N

+ a10*[RVP(t-1,j) / AvrRVP(t-1,j)]*[ROO(t,j) - AvrROO(t)]/N

+ a11*[RVP(t-1,j) / AvrRVP(t-1,j)]*[ROC(t-1,j) - AvrROC(t-1)]/N

+ a12*[RVP(t-1,j) / AvrRVP(t-1,j)]*[RCO(t,j) - AvrRCO(t)]/N

TVL(t-1,j) denotes the trading volume on day t-1, AvrTVL(t-1,j) refers to the average TVL for stock j from max(1, t-200) to t-1. Similarly, AvrRVP(t-1,j) refers to the average RVP for stock j from max(1, t-200) to t-1. Note the timing of each variable.

Your open-to-close portfolio returns must be computed as

(9) RP2(t) = [Σ_j{W2(t,j)*ROC(t,j)}] / [Σ_j{|W2(t,j)|}],

Your task is to find a vector of time-invariant coefficients, {a1, a2,…, a12}, that maximize the Sharpe-ratio (average return over standard deviation of returns) for RP2. The Sharpe-ratio corresponding to your coefficients will ultimately be evaluated by us on an out-of-sample period that you do not have access to.

Part 3: Impose Restrictions on Trade Direction

Consider the trade-direction indicator variable, IND(t,j), in the data set. If IND(t,j) = +1, then you can buy but you cannot sell stock j on day t. If IND(t,j) = -1, then you can sell-short but you cannot buy stock j on day t.

We now define FILL3(t, j) which indicates if you were able to be filled on (that is, execute) your desired W3(t,j).

(10) FILL3(t, j) = 1,if W3(t,j)*IND(t,j) $\geq$ 0, otherwise FILL3(t, j) = 0.

Please provide a new set of time-invariant coefficients {b1, b2,…, b12} corresponding to equation (8) for this market environment. Your portfolio returns are now defined by

(11) RP3(t) = [Σ_j{FILL3(t,j)*W3(t,j)*ROC(t,j)}] / [Σ_j{|FILL3(t,j)*W3(t,j)|}]

If [Σ_j{|FILL3(t,j)*W3(t,j)|}] = 0 on day t then set RP3(t) = 0. You will again be evaluated using the Sharpe-ratio on an out-of-sample period that is not available to you.

Part 4: Generalized Portfolio Weights

In this part you can freely specify a W4(t,j) matrix *without* relying on equation (8) while still using daily rebalancing.

We now define FILL4(t, j) which indicates if you were able to be filled on (that is, execute) your desired W4(t,j).

(12) FILL4(t, j) = 1,if W4(t,j)*IND(t,j) ≥ 0, otherwise FILL4(t, j) = 0.

You will be evaluated using the Sharpe-ratio for RP4(t):

(13) $RP4(t) = [\Sigma_j\{FILL4(t,j)*W4(t,j)*ROC(t,j)\}] / [\Sigma_j\{|FILL4(t,j)*W4(t,j)|\}]$

If $[\Sigma_j\{|FILL4(t,j)*W4(t,j)|\}] = 0$ on day t then set RP4(t) = 0. You will again be evaluated using the Sharpe-ratio on an out-of-sample period that is not available to you.

If you get stuck on Part 2 or 3 you can still move on to Part 4.

Deliverables for Each Part

Please provide separate CSV and PDF files for each Part with the following information. Your CSV should contain a single header row. Use the value 99 for missing observations. Please provide the information in the order listed below:

1) The date in format yyyymmdd as in the input data file.
2) The time series of the long-short return, RP#(t).
3) The time series of the cumulative long-short return in natural logarithms defined as $CumR(t) = \ln[\Pi_{3,t}\{1+RP\#(\tau)\}]$, where $\Pi_{3,t}\{\cdot\}$ denotes the product across τ from day 3 through day t.
4) The time series of $\Sigma_j\{|W\#(t,j)|\}/N$, or for Parts 3 and 4: $\Sigma_j\{|FILL\#(t,j)*W\#(t,j)|\}/N$.
5) The time series of $[\Sigma_j\{W\#(t,j)\}]/[\Sigma_j\{|W\#(t,j)|\}]$ or for Parts 3 and 4: $[\Sigma_j\{FILL\#(t,j)*W\#(t,j)\}]/[\Sigma_j\{|FILL\#(t,j)*W\#(t,j)|\}]$.
6) The W#(t,j) for all stocks and days. Let each column correspond to a stock and let each row correspond to a day.

For Part 2 and 3 please provide separate CSV files with a single header row and with the following row of numbers: {a1, a2, …, a12} and {b1, b2, …, b12}, respectively.

Please provide a separate PDF file for each Part with time series plots of

- The long-short return, RP#(t).
- The cumulative long-short return in natural logarithms
- $\Sigma_j\{|W\#(t,j)|\}/N$ or for Parts 3 and 4: $\Sigma_j\{|FILL\#(t,j)*W\#(t,j)|\}/N$.
- $[\Sigma_j\{W\#(t,j)\}]/[\Sigma_j\{|W\#(t,j)|\}]$ or for Parts 3 and 4:

$$[\Sigma_j\{FILL\#(t,j)*W\#(t,j)\}]/[\Sigma_j\{|FILL\#(t,j)*W\#(t,j)|\}].$$

In the same PDF file please include the following:

- The average and standard deviation of daily log returns, defined as ln(1+RP#(t), and the annualized Sharpe ratio (SR) of the return on the strategy. The SR can be computed as the average daily return divided by the standard deviation of the daily returns. Annualize by multiplying the SR by the square root of 252.
- The skewness and excess kurtosis of the returns on the strategy.
- The maximum drawdown: The largest cumulative loss incurred during the sample. Report the number of days and the peak-to-trough cumulative return during the drawdown period.
- The correlation of the strategy return with the return on the corresponding equal-weighted (long-only) portfolio of the 100 stocks.
- The parameters {a1, a2, …, a12} for Part 2, and {b1, b2, …, b12} for Part 3 along with a brief description on how you obtained the values. For Part 4 include the formula that you used to compute your W(t,j), including any parameter values, and a brief description of how you arrived at the formula and the parameter values.

In the PDF please also comment briefly on the strategy performance. Does the strategy perform well in your view? What do you think is causing its performance?

Summary of Deliverables

By 5PM Sunday March 20 please email to erinp@wil.com the following:

CSV File:	data_part1.team_###.csv
CSV File:	data_part2.team_###.csv
CSV File:	coeff_part2.team_###.csv
CSV File:	data_part3.team_###.csv
CSV File:	coeff_part3.team_###.csv
CSV File:	data_part4.team_###.csv
PDF File:	description_part1.team_###.pdf
PDF File:	description_part2.team_###.pdf
PDF File:	description_part3.team_###.pdf
PDF File:	description_part4.team_###.pdf

where ### denotes your team number. Sample CSV files are posted on Google Drive. Please pay attention to the format of these files. Your data should be comma separated.

Presentations

Should your team be chosen then on March 28 you will do a 7-10 minute presentation followed by a question and answer session. Add judges will be given the PDF documents and no further handouts are allowed at the presentation.

Questions can be emailed to erinp@wil.com by 2pm on Saturday, March 19. Based on the questions received, we will address select questions and concerns, and email one response to all teams on Saturday evening. All questions will be kept anonymous.