UPDATE:
An Automated Trading Strategy Using Earnings Date Revisions (data from Wall Street Horizon)

By: The Deltix Quantitative Research Team

Introduction

This study is an update of the research we first published in November 2015. At that time, we had data from January 2006 to September 2015. In this study, we have data from January 2006 to March 2017.

The study examines whether revisions of earnings announcement dates are a source of information for generating alpha in the universe of S&P500 companies.

Wall Street Horizon (WSH) provides dates for future company earnings announcements as snapshots published daily at 16:00 Eastern Time (ET). For each company, the future date on which earnings are announced is published by the company or inferred by WSH well before the actual announcement of the earnings per share (EPS). Sometimes companies change a previously published or inferred earnings date and it is such changes that we assess as opportunities to generate alpha through trading.

The WSH earnings announcement data set was loaded into Deltix TimeBase, the time-series data warehouse, for the period January 3, 2006 to March 31, 2017. Candidate trading strategies were developed, tested and refined in Deltix QuantOffice.

Daily Earnings Date Snapshots

The record features a set of fields:

Stock Symbol	Char(255)	This is the company's ticker symbol.
Next ED	Date	This field indicates the next earnings date.
Time of Day	Char(50)	This field indicates the time of the announcement ("Before Market", "During Market", "After Market" or "Unspecified").
Next ED Quarter	Char(2)	This is the quarter for the next earnings announcement (Q1, Q2, Q3 or Q4). Note that this is the company's financial quarter, not necessarily the calendar quarter. Relates to fiscal year.
EType	Char(10)	This is the state of the earnings date confirmation ("V" for Verified as a firm date or "T" for Tentative or 'I" for Inferred (WSH forecast)
Fiscal year	Char(4)	This is the reporting fiscal year, relates to Next ED Quarter.

An example of this data as loaded into TimeBase is shown below:

- OPXT
- OPY
- ORA
- ORAN
- ORB
- ORBC
- ORBK
- ORBT
- ORBT.PK
- ORC
- ORCC
- ORCD
- ORCH
- ORCI
- ORCL
- ORCT
- ORCT.OB
- OREX
- ORGN
- ORGN.PK
- ORH
- ORI
- ORIG
- ORIT
- ORITD
- ORLY

Instrument Type	Time	EPSSnapshot				
		EPSDate	TimeOfDay	EType	FiscalQuarter	FiscalYear
EQUITY	02/11/2009 16:00:00	03/25/2009 23:00:00		I	Q3	2,009
EQUITY	02/12/2009 16:00:00	03/25/2009 23:00:00		I	Q3	2,009
EQUITY	02/13/2009 16:00:00	03/25/2009 23:00:00		I	Q3	2,009
EQUITY	02/17/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	02/18/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	02/19/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	02/20/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	02/23/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	02/24/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	02/25/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	02/26/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	02/27/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	03/02/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	03/03/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	03/04/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	03/05/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	03/06/2009 16:00:00	03/25/2009 23:00:00	After Market	T	Q3	2,009
EQUITY	03/10/2009 15:00:00	03/17/2009 23:00:00	After Market	V	Q3	2,009
EQUITY	03/11/2009 15:00:00	03/17/2009 23:00:00	After Market	V	Q3	2,009
EQUITY	03/12/2009 15:00:00	03/17/2009 23:00:00	After Market	V	Q3	2,009
EQUITY	03/13/2009 15:00:00	03/17/2009 23:00:00	After Market	V	Q3	2,009
EQUITY	03/16/2009 15:00:00	03/17/2009 23:00:00	After Market	V	Q3	2,009
EQUITY	03/17/2009 15:00:00	03/17/2009 23:00:00	After Market	V	Q3	2,009

Basis of Research Using Wall Street Horizon data

The purpose of the research described in this paper is to determine if there are opportunities to generate alpha in US equities using WSH data as a basis for market movement prediction on the day close prior to an EPS announcement (the holding period is about 17.5 hours).

We show how with the use of a regression model, we can exploit WSH EPS date snapshots to generate excess returns.

In this research, we not only use official data issued by a company itself but also rely on WSH forecasts (designated "I" in data attribute EType) as a significant data source.

Multinomial Logistic Regression

In statistics, logistic regression ("logit regression") predicts the probabilities of the different possible outcomes of a categorically distributed dependent variable based on one or more predictor variables (features) which may be real-valued, binary-valued, categorical-valued, etc.

The probabilities describing the possible outcomes of a single trial are modeled, as a function of the explanatory (predictor) variables, using a logistic function. Frequently logistic regression is used to refer specifically to the problem in which the dependent variable is binary and problems with more than two categories are referred to as multinomial logistic regression.

Multinomial logistic regression is known by a variety of other names, including multiclass LR, multinomial regression, softmax regression, multinomial logit, maximum entropy (MaxEnt) classifier and conditional maximum entropy model.

As in other forms of linear regression, multinomial logistic regression uses a linear predictor function *f(k,i)* to predict the probability that observation *i* has outcome *k*, of the following form:

$$f(k,i) = \beta_{0,k} + \beta_{1,k}x_{1,i} + \beta_{2,k}x_{2,i} + \cdots + \beta_{M,k}x_{M,i}$$

where $\beta_{m,k}$ is a regression coefficient associated with the *m-th* explanatory variable and the *k-th* outcome. The regression coefficients and explanatory variables are normally grouped into vectors of size M+1, so that the predictor function can be written more compactly as:

$$f(k,i) = \beta_k \cdot x_i$$

where β_k is the set of regression coefficients associated with outcome *k*, and x_i (a row vector) is the set of explanatory variables associated with observation *i*.

The unknown parameters in each vector β_k are found using iteratively reweighted least squares (IRLS).

"Regressors" is an N-by-P design matrix with N observations on P predictor variables:

$$X = \begin{bmatrix} x_{0,0} & \cdots & x_{0,P-1} \\ \vdots & \ddots & \vdots \\ x_{N-1,0} & \cdots & x_{N-1,P-1} \end{bmatrix}$$

"Regressands" is an N-by-K matrix, where Regressands*(i,j)* is the number of outcomes of the multinomial category j for the predictor combinations given by Regressors.Rows*(i)*:

$$Y = \begin{bmatrix} y_{0,0} & \cdots & y_{0,K-1} \\ \vdots & \ddots & \vdots \\ y_{N-1,0} & \cdots & y_{N-1,K-1} \end{bmatrix}$$

The result β is a *(P+1)*-by-*(K-1)* matrix of estimates, where each column corresponds to the estimated intercept term and predictor coefficients, one for each of the first *(K-1)* multinomial categories. The estimates for the *K-th* category are taken to be zero:

$$\beta = \begin{bmatrix} \beta_{1,0} & \cdots & \beta_1, P \\ \vdots & \ddots & \vdots \\ \beta_{K-1,0} & \cdots & \beta_{K-1,P} \end{bmatrix}$$

Multinomial logistic regression is implemented by QuantOffice's FinMath numerical library.

Model Implementation

Firstly, we need to categorize company behavior prior the EPS announcement. We will consider two independent variables:

- EPS date shift: we will denote it by ΔD_{EPS};
- Total days from last EPS date update: we will denote it by ΔU_{EPS}.

Each of them can be categorized using baskets for its values; in our simple case we will consider two baskets for each of them:

- ΔD_{EPS}: < 0 or >0;
- ΔU_{EPS}: < 1/2 of quarter or ≥ 1/2 of quarter.

45 days value is taken as a boundary for half of the quarter.

The dependent variable will be the stock's return:

$$Return = Ln\left(\frac{Price\ after\ EPS\ announcement}{Price\ before\ EPS\ announcement}\right)$$

Companies make their EPS announcements after the market close so we take the day close price prior to the EPS announcement and the day open price after the announcement date to calculate the return. The dependent variable can be categorized similarly into baskets of values. In our case we will take two baskets: Return <0 and >0, which makes our regression model into a logistic regression.

The following algorithm is applied:

1. Collect data for companies' behavior and returns;
2. Calculate number of possible cases to form the matrix of regressors; in our simple case there will be 4 cases and matrix will contain 4 rows:

$$X = \begin{bmatrix} \Delta \mathrm{D_{EPS}} > 0 & \Delta \mathrm{U_{EPS}} < 45 \\ \Delta \mathrm{D_{EPS}} < 0 & \Delta \mathrm{U_{EPS}} < 45 \\ \Delta \mathrm{D_{EPS}} > 0 & \Delta \mathrm{U_{EPS}} \geq 45 \\ \Delta \mathrm{D_{EPS}} < 0 & \Delta \mathrm{U_{EPS}} \geq 45 \end{bmatrix}$$

3. Calculate number of outcomes for each row;
4. Calculate estimates matrix using QuantOffice' FinMath numerical library;
5. Compare the estimates and decide which row provides the highest probability for each of the dependent variable outcomes. For example, we get the following estimates matrix:

$$\beta = \begin{bmatrix} 0.49552 & 0.50448 \\ 0.55181 & 0.44819 \\ 0.44748 & 0.55252 \\ 0.50376 & 0.49624 \end{bmatrix}$$

   The first column indicates probabilities for a positive return, second column for a negative one. We can see that the highest probability of positive return is for row 2 and negative return is for row 3 of matrix X.

6. We will also choose the significance level equal to 10%, which means that one outcome must be at least 10% more probable than the other one for a single row of independent variables.
7. Recalculate estimates every 90 days (duration of a quarter) for a sliding window of 1500 observations, where each observation represents data pertaining to an EPS announcement for a particular stock.

If we look at the report with categories that were chosen as most probable by the logistical regression method, we can see the following:

- Positive returns most likely happen for the case

  $\Delta \mathrm{D_{EPS}} < 0$, $\Delta \mathrm{U_{EPS}} < 45$;
- Negative returns most likely happen for the case

  $\Delta \mathrm{D_{EPS}} > 0$, $\Delta \mathrm{U_{EPS}} \geq 45$.

Possible reasons for such results are the following:

- If a company shifts the EPS date up in the second half of the quarter, it is more likely to report positive news as it already has positive information to report;
- Conversely, if a company delays its EPS date to have a significant time span before the EPS announcement (not less than 45 days), most likely negative news will be reported.

Trading Strategy

We can now implement our findings as a trading strategy:

1. On day close prior to the EPS announcement

   Open **long position** if company falls into categories with highest probability of positive return

   $\Delta D_{EPS} < 0$, $\Delta U_{EPS} < 45$.

   Open **short position** if company falls into categories with highest probability of negative return

   $\Delta D_{EPS} > 0$, $\Delta U_{EPS} \geq 45$.

2. Close positions on next day open.

3. We calculate money value of the position as follows:

$$MV = \begin{cases} BetSize, & if\ BetSize * n \leq TradingCapital \\ \dfrac{TradingCapital}{n}, & if\ BetSize * n > TradingCapital \end{cases}$$

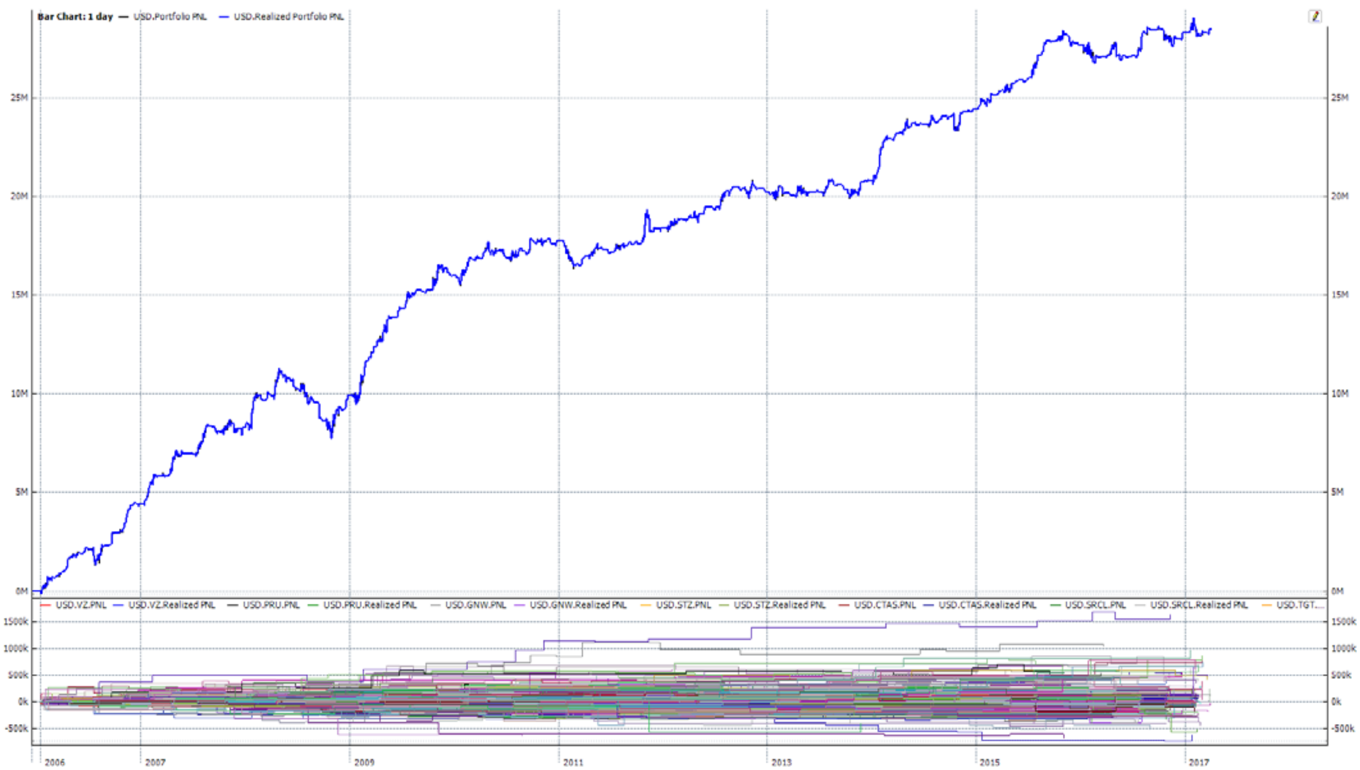
   where n is the number of positions to be opened on a particular day.

   We designed another version of the strategy: dollar-neutral strategy, where portfolio is hedged with the SPY ETF in order to exclude market movement impact.

Results

This approach is back-tested for S&P500 equities over the period from Jan 3, 2006 to March 31, 2017 for TradingCapital = $10M. The QuantOffice summary reports from the back-tests are shown below.

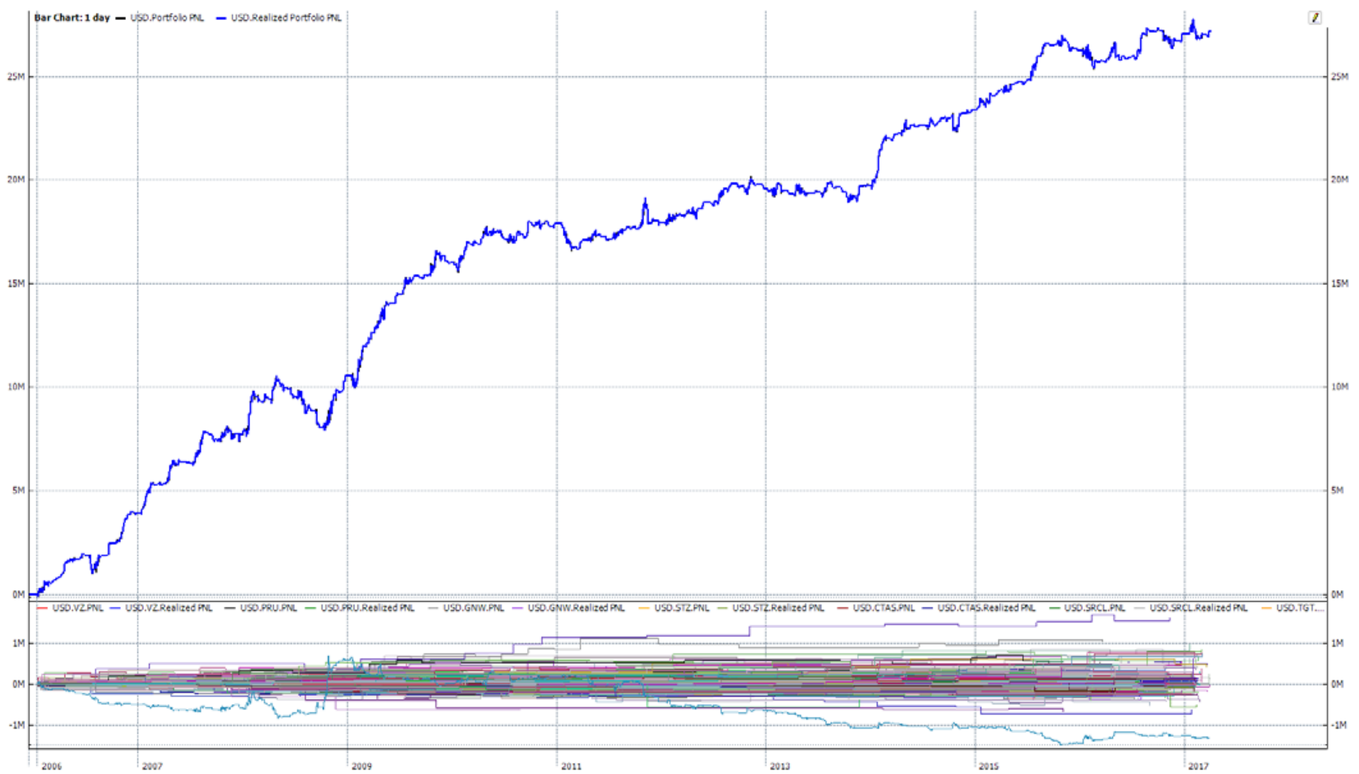
Non-hedged Strategy

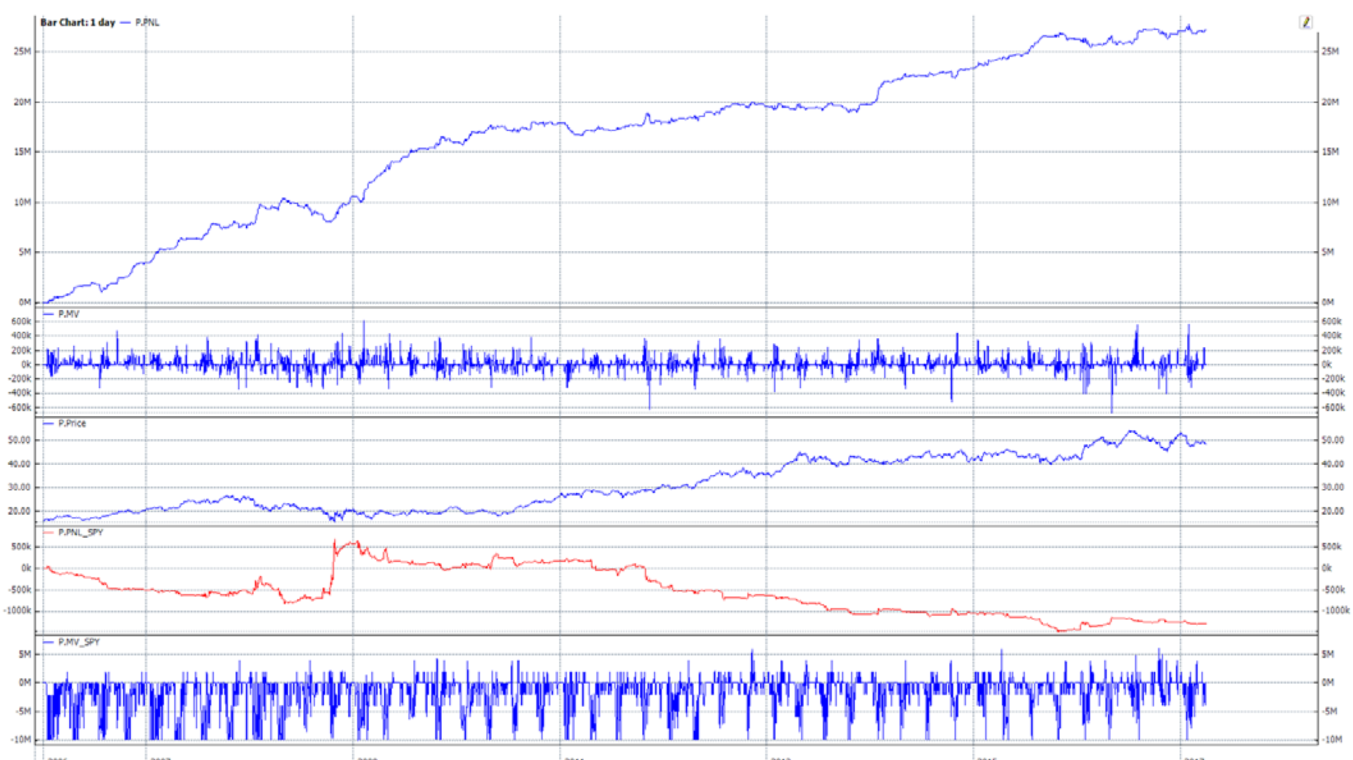


Parameter	All Trades	Long Trades	Short Trades
Net Profit/Loss	28,490,708.43	26,139,744.08	2,350,964.35
Total Profit	128,390,239.50	103,679,078.40	24,711,161.10
Total Loss	-99,899,531.07	-77,539,334.32	-22,360,196.75
Cumulated Profit %	284.91 %	261.40 %	23.51 %
Max Drawdown	-3,432,710.53	-2,497,417.48	-1,537,397.70
Max Drawdown %	-16.20 %	-11.94 %	-12.24 %
Max Drawdown Date	10/24/2008	10/24/2008	11/13/2013
Drawdown Days Percent	79.00 %	79.56 %	89.02 %
Max Drawdown Duration	323	535	1092

CAGR	13.17 %	12.51 %	1.96 %
Sharpe Ratio	1.96	1.96	0.31
Annualized Volatility	6.72	6.38	6.40
Sortino Ratio	3.40	3.38	0.48
UPI	0.32	0.41	0.03
Information Ratio	1.73	1.76	0.30
Optimal f	29.14	30.72	4.77
Historical VaR 95% 1D	-121837.43	-105739.16	-50850.75
Historical CVaR 95% 1D	-213707.62	-196393.82	-107976.92
Theoretical VaR 95% 1D	-144271.49	-130104.10	-72053.83
Theoretical CVaR 95% 1D	-393390.10	-355014.47	-189690.17
All Trades #	5595	4597	998
Profitable Trades Ratio	0.54	0.55	0.50
Winning Trades #	3010	2512	498
Losing Trades #	2585	2085	500
Average Trade	5092.17	5686.26	2355.68
Average Winning Trade	42654.56	41273.52	49620.81
Average Losing Trade	-38645.85	-37189.13	-44720.39
Avg. Win/ Avg. Loss Ratio	1.10	1.11	1.11
Average Profit per Share	0.10	0.11	0.04
Max Conseq. Winners	15	16	9
Max Conseq. Losers	10	9	13
Average Trade Holding Time	00.19:54:59	00.19:49:30	00.20:20:15

Dollar-neutral Strategy Hedged by SPY





Parameter	All Trades	Long Trades	Short Trades
Net Profit/Loss	27,188,623.14	26,350,141.60	838,481.54
Total Profit	140,656,919.19	104,840,603.36	35,816,315.83
Total Loss	-113,468,296.05	-78,490,461.76	-34,977,834.29
Cumulated Profit %	271.89 %	263.50 %	8.38 %
Max Drawdown	-2,414,151.38	-2,453,068.02	-2,271,205.34
Max Drawdown %	-11.82 %	-11.74 %	-18.81 %
Max Drawdown Date	10/13/2008	10/24/2008	11/13/2013
Drawdown Days Percent	80.13 %	80.30 %	97.42 %
Max Drawdown Duration	429	323	1977
CAGR	12.81 %	12.57 %	.74 %
Sharpe Ratio	1.99	1.97	0.09
Annualized Volatility	6.45	6.38	7.96
Sortino Ratio	3.46	3.40	0.14
UPI	0.36	0.42	0.01
Information Ratio	1.72	1.77	0.09
Optimal f	30.79	30.92	1.17
Historical VaR 95% 1D	-115853.27	-106018.13	-65394.04
Historical CVaR 95% 1D	-203318.97	-196590.47	-124480.87
Theoretical VaR 95% 1D	-138541.86	-130331.89	-84571.37
Theoretical CVaR 95% 1D	-377669.51	-355730.02	-221547.40
All Trades #	7712	5095	2617
Profitable Trades Ratio	0.48	0.52	0.40
Winning Trades #	3699	2640	1059
Losing Trades #	4013	2455	1558

Average Trade	3525.50	5171.76	320.40
Average Winning Trade	38025.66	39712.35	33820.88
Average Losing Trade	-28275.18	-31971.67	-22450.47
Avg. Win/ Avg. Loss Ratio	1.34	1.24	1.51
Average Profit per Share	0.08	0.10	0.01
Max Conseq. Winners	12	16	8
Max Conseq. Losers	15	14	13
Average Trade Holding Time	00.18:25:42	00.18:55:07	00.17:28:25

Consolidated Report

Parameter	Un-hedged	Hedged by SPY
Net Profit/Loss	28,490,708.43	27,188,623.14
Total Profit	128,390,239.50	140,656,919.19
Total Loss	-99,899,531.07	-113,468,296.05
Cumulated Profit %	284.91 %	271.89 %
Max Drawdown	-3,432,710.53	-2,414,151.38
Max Drawdown %	-16.20 %	-11.82 %
Max Drawdown Date	10/24/2008	10/13/2008
Drawdown Days Percent	79.00 %	80.13 %
Max Drawdown Duration	323	429
CAGR	13.17 %	12.81 %
Sharpe Ratio	1.96	1.99
Annualized Volatility	6.72	6.45
Sortino Ratio	3.40	3.46

UPI	0.32	0.36
Information Ratio	1.73	1.72
Optimal f	29.14	30.79
Historical VaR 95% 1D	-121837.43	-115853.27
Historical CVaR 95% 1D	-213707.62	-203318.97
Theoretical VaR 95% 1D	-144271.49	-138541.86
Theoretical CVaR 95% 1D	-393390.10	-377669.51
All Trades #	5595	7712
Profitable Trades Ratio	0.54	0.48
Winning Trades #	3010	3699
Losing Trades #	2585	4013
Average Trade	5092.17	3525.50
Average Winning Trade	42654.56	38025.66
Average Losing Trade	-38645.85	-28275.18
Avg. Win/ Avg. Loss Ratio	1.10	1.34
Average Profit per Share	0.10	0.08
Max Conseq. Winners	15	12
Max Conseq. Losers	10	15
Average Trade Holding Time	00.19:54:59	00.18:25:42

Conclusion

We presented an approach in which WSH earnings (EPS) announcement date daily snapshots can serve as a strong predictive factor of stock price directionality. We developed a trading strategy that implements an algorithm based on this approach.

As the result, we found that the most probable positive returns are for shifts where the EPS announcement date is brought forward in the second half of the quarter; and negative returns for delays in the EPS announcement date when reported in the first half of the quarter.

Further, we developed a dollar-neutral version of the strategy to exclude market return impact. This approach demonstrates that the generated return is independent of market movement and as such represents alpha return.

In back-testing this strategy on stocks in the S&P500, back-testing shows that the strategy has an average Sharpe Ratio of 1.97 over the period 2006-2017 with average profit per share of 8 cents.

About Deltix

Deltix is a leading provider of software and services for quantitative research, algorithmic and automated systematic trading. Deltix software enables a complete straight through processing environment for the development and deployment of closely-integrated alpha generation and/or execution strategies. Deltix has won twelve industry awards since 2012. Deltix is headquartered in Natick, Massachusetts, and has offices in New York, Minsk and St. Petersburg, Russia. For more information, please visit http://www.deltixlab.com.

About Wall Street Horizon

Wall Street Horizon provides institutional investors and traders with an ever expanding set of forward-looking and historical corporate event datasets including earnings dates, dividend dates, options expiration dates, splits, spinoffs and a wide variety of investor-related conferences. With access via machine-readable feeds or Enchilada, its easy-to-use online application, the company's data is widely recognized for its unmatched accuracy and timeliness. For more information, please visit www.WallStreetHorizon.com or email them at info@wallstreethorizon.com.