



Tactical Equity Portfolio Formation Utilizing News Analytics

By:

Peter Hafez/Director of Quant Research, RavenPack

Junqiang Xie, PhD/Quant Analyst, RavenPack

David Marra/CEO, Arianalytics Inc

Elliot Taniguchi, PhD/Data Scientist, Arianalytics Inc

April 2014

[Contact the authors](#)

This White Paper is not intended for trading purposes. The White Paper is not appropriate for the purposes of making a decision to carry out a transaction or trade. Nor does it provide any form of advice (investment, tax, legal) amounting to investment advice, or make any recommendations regarding particular financial instruments, investments or products. RavenPack may discontinue or change the White Paper content at any time, without notice. RavenPack does not guarantee or warrant the accuracy, completeness or timeliness of the White Paper. For more detailed disclaimer, please refer to the back cover of this document.

TABLE OF CONTENTS

Executive Summary	2
Introduction	3
Expected Returns Estimation Overview	4
Portfolio Construction Overview	4
Portfolio Results	7
Conclusion	11
About RavenPack	12
About Arianalytics	13

Powered by:

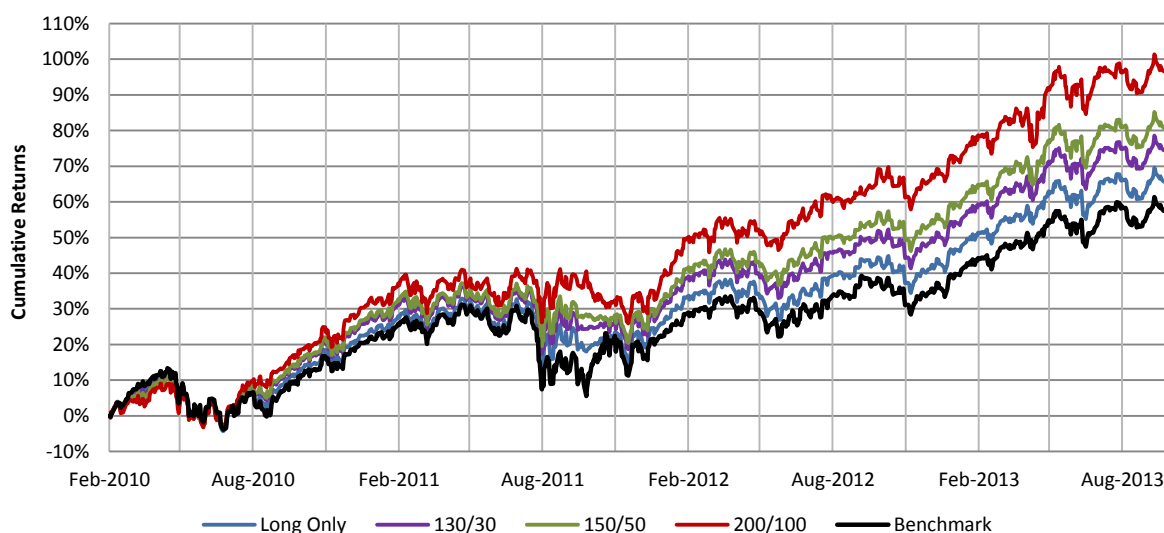
Arianalytics 

info@arianalytics.com

Executive Summary

In this study, we utilize Arianalytics' predictive financial modeling platform optimized for high-level computing and RavenPack's news analytics derived from real-time content published by Dow Jones Newswires, The Wall Street Journal, and Barron's - to construct long-only and long-short portfolios based on sector level 1-month expected returns. The constructed portfolios have been found to consistently outperform the S&P 500 on both an absolute and risk-adjusted basis.

- By constructing a **long-only sector portfolio** with a 1-month investment horizon, we improve the Sharpe Ratio by 25% to 0.79, compared to 0.63 for the S&P 500.
- Constructing a **130/30 sector portfolio**, we are able to improve the Sharpe Ratio by almost 43% to 0.90 – reducing the maximum drawdown by more than 29% to 14.0%.
- Constructing a **150/50 sector portfolio**, the Sharpe Ratio improves by almost 70% to 1.07 – reducing the maximum drawdown by more than 42% to 11.3%.



Cumulative return profiles of the news-enhanced market portfolios – net 10 bps transaction costs. The backtesting period covers February 2010 through September 2013.

SOURCE: RavenPack, Arianalytics, April 2014

About Arianalytics

Arianalytics' predictive analytics platform provides an intelligent way to extract predictive information from data and turn it into superior investment and risk management outcomes. The platform streamlines the predictive modeling process – from dataset creation to online forecasting – by automating the cleaning, organizing, and analyzing of large, complex datasets and development of dynamic prediction solutions.

About RavenPack Data

RavenPack News Analytics (RPNA) provides real-time structured sentiment, relevance and novelty data for entities and events detected in unstructured text published by reputable sources. Up to 14 years of Dow Jones newswires archive and 7 years of historical data from web publications and blogs are available for backtesting. RavenPack detects news and produces analytics data on over 33,000 listed stocks from the world's equity markets, over 2,300 financially relevant organizations, 138,000 places, 150 currencies and 80 commodities.

1. Introduction

The complexity of financial markets makes it challenging to model the impact of news on individual stock prices, sectors, or markets – especially over the long-term. Although the intraday reaction to a news event is often relatively straightforward, the impact over days and weeks become harder to model. More so when investors are bombarded with lots of new information - some of which can be extremely valuable and some merely noise. The sheer volume of news and financial data available to investors present serious analytics challenges that cannot be solved with traditional tools.

Due to the size of these ‘big data’ sets, estimating predictive models using conventional statistical means is not feasible. We therefore perform our analysis using the Arianalytics predictive financial modeling platform called “Aria”, a distributed computing platform optimized for finding statistically robust solutions from large amounts of financial (structured) and non-financial (news) data.

Catering to longer-term investors, we examine U.S. equity portfolio formation strategies utilizing one-month ahead expected return estimates derived from RavenPack News Analytics¹ and priced market factors. The underlying data files for this equity portfolio return analysis consist of daily one-month forward raw (log) return estimates for the following major U.S. industries: basic materials, consumer services, energy, financials, healthcare, industrials, consumer goods, technology, telecommunications, and utilities. These major industries are tracked by Dow Jones industry-level price indexes: DJUSBM, DJUSCY, DJUSEN, DJUSFN, DJUSHC, DJUSIN, DJUSNC, DJUSTC, DJUSTL, and DJUSUT, respectively. The expected return estimates we generate are published daily 90 minutes prior to market open. Based on our daily predictions, we construct a set of long-only and long-short strategies (including 130/30, 150/50, and 200/100); All of which are shown to outperform the long-only US market benchmark over nearly 4 years of backtesting - covering the period February 2010 through September 2013.

In the following sections, we provide an overview of how we estimate our expected returns. We describe the portfolio construction process including how the benchmarks and our long-only and long-short trading strategies are created. We evaluate our strategy results including applying different portfolio turnover restrictions and finally present our conclusions.

¹ Focusing on relevant information from Dow Jones Newswires, regional editions of the Wall Street Journal, and Barron's

2. Expected Returns Estimation Overview

The underlying data for each target security's portfolio weights, or expected return estimation, consists of a single RavenPack U.S. market file, multiple RavenPack sector files (one per sector), and their corresponding Arianalytics market and sector files.

The RavenPack data files each contain a total of approximately 7,500 candidate news analytic predictors spanning the breadth and depth of RavenPack's news analytics content. These files are created as daily aggregates of story-level news analytics across the constituents of a given sector or market. The data file contains information about overall news volume and sentiment - applying different novelty and relevance filters, cumulative sums and moving averages, various sentiment measures, and knowledge about the underlying events (e.g. layoffs, product recalls, M&A activity etc.).

The Arianalytics files each contain approximately 4,000 candidate predictors per file at a daily frequency. The predictors in these files are motivated by predictors found in the asset pricing literature that have been shown to predict the first and second moments of asset returns and changes in the investment opportunity set. These include market-wide and sector-specific state variables, CAPM and ICAPM factors, measures of hedging demands, and priced factors that include measures of risk, liquidity and uncertainty.

To derive portfolio weights and expected return estimates, Arianalytics filtered and combined each data file into a single aggregated dataset averaging over 15,000 variables per security. Due to the size of these 'big data' datasets, estimating predictive models using conventional ordinary least squares (OLS) statistical means is not feasible. Portfolio weights and expected returns are estimated using Arianalytics' predictive financial modeling platform, Aria, a distributed computing platform optimized for finding statistically robust solutions from large amounts of financial (structured) and non-financial (news) data.

3. Portfolio Construction Overview

Daily expected return estimates are converted to daily trading positions with a fixed investment period of 21 days matching the 21-day forward horizon of the estimates. The strategies evaluated in this paper utilize the direction of the expected returns to build equity portfolios. The choice of utilizing return direction (i.e. positive or negative) is primarily motivated by two considerations:

(A) *Conservative Risk Management Approach to Portfolio Construction.* Utilizing estimated return direction weakens biases that could result from excessive portfolio strategy parameter tuning and optimization, and it strengthens the linkage to live-trading results.

(B) *Objective Performance Evaluation.* Keeping trading rules simple and transparent allows the reader to more objectively evaluate results in general scenarios compared to complex/customized strategies that require additional parameters tailored to different user needs (at the possible expense of sub-optimal results).

We consider two potential portfolio construction implementation methods: interleaved and non-interleaved. The non-interleaved method picks an arbitrary investment start date and makes non-overlapping re-investments every 21 days at the market close, compounding returns. This would involve executing trades to close the previous position made 21 days prior and immediately reinvesting capital using the most recent expected returns estimates.

The interleaved method incorporates all relevant expected returns estimates into a single aggregate portfolio with daily trading and investment horizons of 21 days. The construction mechanism simply builds a single equal-weighted, aggregate portfolio using the 21 individual non-interleaved portfolios described earlier. Daily trading is required because exactly one underlying non-interleaved portfolio will expire daily and a new position will be taken. We focus on results from the interleaved portfolio construction method because it is less dependent on start-date (less timing risk), it converges to the average performance of the non-interleaved approach, and average turnover is similar.

3.1 Long-Only Equity Portfolio Detail

The long-only strategy allocates equal weight to securities with positive expected returns. More specifically, for $N_p > 0$ securities, those having strictly positive expected returns, the portfolio weight assigned to each of these securities is $100\% / N_p$. Correspondingly, the strategy assigns zero weight to all securities with non-positive expected returns. If N_p is zero, then the portfolio allocates all capital to cash.

	Case I: $N_p > 0$	Case II: $N_p = 0$
Equity Weights (Pos. Expected Returns)	$1 / N_p$	0
Equity Weights (Neg. Expected Returns)	0	0
Cash Weight	0	1

3.2 Simple Long-Short Equity Portfolio Detail

The long-short strategy has a net long market exposure and short-selling is limited to 30% of the portfolio to approximate 130/30 long-short portfolios. Three allocations are constructed which depend on expected returns. Case I, with $N_p > 0$, positive expected returns, and $N_n > 0$, negative expected returns, long and short positions have a net positive market exposure. In Case II, positive-only expected returns, the portfolio behaves as a long-only portfolio. In Case III, negative-only expected returns, the portfolio takes equal weighted short positions, holds the excess in cash, and has a net negative market exposure. This is due to the fact that the trading strategy is fundamentally based on expected return direction and not on relative returns. Similar strategies are constructed using a 50% and 100% short-selling limit to showcase the potential benefit of increasing the short exposure relative to the overall market exposure.

	Case I: $N_p > 0, N_n > 0$	Case II: $N_p > 0, N_n = 0$	Case III: $N_p = 0, N_n > 0$
Equity Weights (Pos. Expected Returns)	$(1.3) / N_p$	$1 / N_p$	0
Equity Weights (Neg. Expected Returns)	$-0.3 / N_n$	0	$-0.3 / N_n$
Cash Weight	0	0	1.3

3.3 Equal-weight, Long-Only Equity Portfolio Detail (Benchmark)

The equal-weight equity portfolio allocates an equal weight across all securities. Note that minimal trading still occurs to rebalance the portfolio daily because relative price returns differences between securities will cause the weights to drift away from equal weighting.

The benchmark portfolio can be used to evaluate the returns and investment alpha generated by the expected-returns-motivated portfolios described above. The underlying investment universe of the benchmark portfolio is exactly the same as the expected returns portfolios.

	All Cases
Equity Weights	$1.0 / (N_p + N_n)$
Cash Weight	0

3.4 Dow Jones Total Market Index Detail (Benchmark)

The reference portfolio contains only the Dow Jones Total Market Index (DJUS) which tracks the entire US equity market. DJUS represents the passive portfolio for an uninformed investor. This benchmark can be used to evaluate the degree to which the expected returns motivated portfolios deliver investment alpha over a passive, uninformed investment strategy.

	All Cases
DJUS Weight	1.0
Cash Weight	0.0

4. Portfolio Results

The table below presents the summary statistics for the RavenPack expected-returns-informed long-only, and long-short portfolios, and the benchmark equal-weight and total market portfolios. The annualized statistics are derived from daily returns following industry convention, and cover the three year period from February 2010 to September 2013 (44 months). All results are net of transactions costs, calculated at 10 basis points.

As indicated by the results in Fig 1, all RavenPack expected-returns-informed portfolios substantially outperform the reference portfolios as measured by Sharpe Ratio, Information Ratio, excess return, and drawdown. Specifically, for the long-only portfolio, we are able to improve the Sharpe Ratio by more than 25% on top of our S&P 500 benchmark – adding an average annual excess return of 1.73%. Allowing short-selling, we are able to improve the Sharpe Ratio by as much as 70% from 0.63 to 1.07 for the 150/50 portfolio – adding 6.39% to the average annual excess return.

Fig 2 presents the cumulative return and drawdown profiles for the different strategies including the equal-weighted benchmark. As can be observed, the outperformance is well-distributed over time – making our strategies an attractive alternative to the S&P 500 benchmark. The bottom panel highlights the improved drawdown profiles with a noticeable downside risk reduction in late 2011.

Fig 1: Strategy Performance Summary

Strategy	RavenPack Long-Only Portfolios	RavenPack 130/30 Long-Short Portfolios	RavenPack 150/50 Long-Short Portfolios	RavenPack 200/100 Long-Short Portfolios	(Benchmark) Equal Weight Portfolios	(Benchmark) Market Index SPX
Short-Sale Limit	NA	0.3	0.5	1	NA	NA
Maximum Drawdown	15.59%	13.95%	13.33%	11.27%	19.57%	19.39%
Max. Recovery Period	193	191	186	134	219	207
Avg. Annual Turnover	350%	727%	982%	1624%	7%	0%
Avg. Annual Excess Return (vs. Libor)	12.46%	13.92%	14.90%	17.12%	10.85%	10.73%
Stdev. Annual Excess Return (vs. Libor)	15.72%	15.46%	15.45%	16.06%	17.10%	17.35%
Sharpe Ratio (vs. Libor)	0.79	0.90	0.96	1.07	0.63	0.62
Avg. Annual Excess Return (vs. SP500)	1.73%	3.18%	4.17%	6.39%	0.12%	0.00%
Stdev. Annual Excess Return (vs. SP500)	4.90%	6.61%	7.84%	11.11%	1.57%	0.00%
Information Ratio (vs. Libor)	0.35	0.48	0.53	0.57	0.07	NaN

This figure summarizes the strategy performance of the long-only, the 130/30, the 150/50, and 200/100 portfolios as well as for the equal weighted and market-cap weighted US market benchmarks – net 10bps transaction costs. The backtesting period covers February 2010 through September 2013.

SOURCE: RavenPack, Arianalytics, April 2014

Fig 2: Strategy Cumulative Returns & Drawdown



This figure summarizes the cumulative return (top panel) and drawdown (bottom panel) profiles of the long-only, 130/30, 150/50, and 200/100 portfolios as well as for the equal weighted US market benchmark – net 10 bps transaction costs. The backtesting period covers February 2010 through September 2013.

SOURCE: RavenPack, Arianalytics, April 2014

4.1 Including Portfolio Turnover Restrictions

As previously observed, the news-based strategies hold relatively high turnover. Even though we include transaction costs in our performance statistics, for some investors high turnover presents significant implementation challenges. To address this issue, portfolio turnover is controlled by restricting trade execution to only during periods when the portfolio weights of the *model portfolio* diverges significantly away from the portfolio weights of the *trading portfolio*. The model portfolio includes the “optimal” positions driven purely by the investment strategy, while the trading portfolio represents the actual fund positions held day-to-day and including transaction costs. The divergence between the model and trading portfolio is measured by the mean squared error in the portfolio weights (excluding cash). If the divergence is greater than a pre-defined fixed threshold parameter, trades are executed to align the portfolio weights (post-transaction costs) of the trading portfolio to the model portfolio. If the fixed threshold parameter is large, trade execution is infrequent (low turnover) but the portfolio weights may diverge significantly between trades. On the other hand, if the fixed threshold parameter is small, trade execution is frequent (high turnover) and the trading portfolio closely tracks the model portfolio.

Fig 3 highlights the average excess return and Sharpe Ratio of the long-only and 130/30 portfolios without turnover restrictions as well as with a target of half the turnover². Reducing turnover comes at the cost of a lower a Sharpe Ratio. Still, even with heavy turnover restrictions, the news-based 130/30 strategy delivers a 30% outperformance on S&P 500 in terms of Sharpe Ratio, while the long-only portfolio outperforms by 20%.

Fig 3: Strategy Performance Summary with Turnover Restrictions

Annualized Statistics Net Transaction Costs (10bps)	RavenPack 130/30 Equity Portfolio (High Turnover)	RavenPack Long-Only Equity Portfolio (High Turnover)	RavenPack 130/30 Equity Portfolio (Low Turnover)	RavenPack Long-Only Equity Portfolio (Low Turnover)
Avg. Excess Return (vs. SP500)	13.9%	12.5%	13.0%	11.8%
Sharpe Ratio (vs. Libor)	0.90	0.79	0.83	0.76
Portfolio Turnover	727%	350%	380%	165%

This figure summarizes the strategy excess return and Sharpe Ratio for different levels of portfolio turnover for the long-only and 130/30 portfolios – net 10 bps transaction costs. The backtesting period covers February 2010 through September 2013.

SOURCE: RavenPack, Arianalytics, April 2014

² Note that turnover is only controlled indirectly, hence making it only an approximation for future turnover.

5. Conclusion

In this study, we utilize Arianalytics' predictive financial modeling platform called "Aria" and RavenPack's news analytics data to construct long-only and long-short portfolios based on sector level 1-month expected returns. The underlying data for each target security's portfolio weights, or expected return estimation, consists of approximately 7,500 candidate news analytic predictors spanning the breadth and depth of RavenPack's news analytic data, and approximately 4,000 candidate predictors defined by Arianalytics at a daily frequency. In total, we consider over 15,000 variables per security in our analysis.

The resulting *informed portfolios* are found to substantially outperform their U.S. market benchmarks as measured by Sharpe Ratio, Information Ratio, excess return, and drawdown. Specifically, for the long-only portfolio, we are able to improve the Sharpe Ratio by more than 25% on top of our S&P 500 benchmark – adding an average annual excess return of 1.73%. Allowing short-selling, we are able to improve the Sharpe Ratio with as much as 70% from 0.63 to 1.07 for the 150/50 portfolio - adding 6.39% to the average annual excess return. On a more traditional 130/30 portfolio, we are able to improve the Sharpe Ratio by almost 43% to 0.90 – reducing the maximum drawdown by more than 29% to 14.0% compared to 19.6% for the S&P 500.

Overall, we find evidence that including RavenPack's news analytics data on top of more traditional quant factors, we are able to construct sector-based portfolios that outperform the S&P 500 over 1-month investment horizons – not only providing positive excess returns, but also improving on the Sharpe Ratio, Information Ratio, and drawdown profile of the strategy.

As next steps, we plan to extend our research using the Arianalytics platform to individual companies, where early indications show promise, and to non-US equity indexes and stocks. We also plan to extend our efforts into other asset classes including foreign exchange and commodities using the Global Macro package of RavenPack News Analytics. Further, we intend to look at horizons both shorter and longer than one month, and to introduce more diverse trading strategies for particular investor types.

About RavenPack

Financial professionals rely on RavenPack for its speed and accuracy in analyzing large amounts of unstructured content. RavenPack's clients use news analytics to enhance returns, reduce risk or increase efficiency by systematically incorporating the effects of public information in their models or workflows. The company's clients include some of the best-performing quantitative and discretionary trading, investing and market-making firms in the world.

RavenPack News Analytics (RPNA) provides real-time structured sentiment, relevance and novelty data for entities and events detected in the unstructured text published by reputable sources. Publishers include Dow Jones Newswires, the Wall Street Journal and over 22,000 other traditional and social media sites. RavenPack News Analytics is used to enhance returns or improve efficiency by quantitative & algorithmic traders, automated market-makers, portfolio managers, risk managers and surveillance analysts. Up to 14 years millisecond time-stamped data is available for backtesting.

Global Equities

RavenPack detects news and produces analytics data on over 33,000 listed stocks from the world's equity markets. Coverage is spread across the Americas, Europe and Asia-Pacific.

Global Macro

RavenPack analyses news and delivers data on over 2,000 financially relevant organisations, 138,000 places, 150 currencies and 80 commodities.

About Arianalytics

Financial institutions rely on Arianalytics to create, test, validate and execute predictive strategies that generate better return, risk management, and research productivity outcomes – systematically and scientifically. We're using science and technology to turn data into better returns, better risk management, and higher research productivity.

The efficient extraction of predictability, aggregated across diverse data types, represents a tremendous opportunity for investors of all kinds – quantitative, fundamental, long-horizon and intra-day traders. Arianalytics helps investors seize this opportunity, today.

Our Aria Platform from Arianalytics is the world's only predictive analytics platform purpose-built for improving investing and risk management outcomes and accelerating research workflow. Collect, clean, and assemble massive datasets, quickly. Create dynamic, high-performance predictive models leveraging state-of-the-art learning algorithms and deep, objective analysis powered by hundreds or thousands of computational cores.

Arianalytics bridges the technology and science divide for investors. We offer predictive solutions that meet the needs of complex investing and trading challenges. These solutions provide investors with the most advanced predictive modeling capabilities at a fraction of the cost and time of building your own solutions.

Terms of Use

This White Paper is not intended for trading purposes. The White Paper is not appropriate for the purposes of making a decision to carry out a transaction or trade. Nor does it provide any form of advice (investment, tax, legal) amounting to investment advice, or make any recommendations regarding particular financial instruments, investments or products. RavenPack may discontinue or change the White Paper content at any time, without notice. RavenPack does not guarantee or warrant the accuracy, completeness or timeliness of the White Paper

You may not post any content from this White Paper to forums, websites, newsgroups, mail lists, electronic bulletin boards, or other services, without the prior written consent of RavenPack. To request consent for this and other matters, you may contact RavenPack at support@ravenpack.com.

THE WHITE PAPER IS PROVIDED "AS IS", WITHOUT ANY WARRANTIES. RAVENPACK AND ITS AFFILIATES, AGENTS AND LICENSORS CANNOT AND DO NOT WARRANT THE ACCURACY, COMPLETENESS, CURRENTNESS, TIMELINESS, NONINFRINGEMENT, TITLE, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OF THE WHITE PAPER, AND RAVENPACK HEREBY DISCLAIMS ANY SUCH EXPRESS OR IMPLIED WARRANTIES. NEITHER RAVENPACK NOR ANY OF ITS AFFILIATES, AGENTS OR LICENSORS SHALL BE LIABLE TO YOU OR ANYONE ELSE FOR ANY LOSS OR INJURY, OTHER THAN DEATH OR PERSONAL INJURY RESULTING DIRECTLY FROM USE OF THE WHITE PAPER, CAUSED IN WHOLE OR PART BY ITS NEGLIGENCE OR CONTINGENCIES BEYOND ITS CONTROL IN PROCURING, COMPILING, INTERPRETING, REPORTING OR DELIVERING THE WHITE PAPER. IN NO EVENT WILL RAVENPACK, ITS AFFILIATES, AGENTS OR LICENSORS BE LIABLE TO YOU OR ANYONE ELSE FOR ANY DECISION MADE OR ACTION TAKEN BY YOU IN RELIANCE ON SUCH WHITE PAPER. RAVENPACK AND ITS AFFILIATES, AGENTS AND LICENSORS SHALL NOT BE LIABLE TO YOU OR ANYONE ELSE FOR ANY DAMAGES (INCLUDING, WITHOUT LIMITATION, CONSEQUENTIAL, SPECIAL, INCIDENTAL, INDIRECT, OR SIMILAR DAMAGES), OTHER THAN DIRECT DAMAGES, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. IN NO EVENT SHALL THE LIABILITY OF RAVENPACK, ITS AFFILIATES, AGENTS AND LICENSORS ARISING OUT OF ANY CLAIM RELATED TO THIS AGREEMENT EXCEED THE AGGREGATE AMOUNT PAID BY YOU FOR THE WHITE PAPER IN THE 12 MONTHS IMMEDIATELY PRECEDING THE EVENT GIVING RISE TO SUCH CLAIM. BECAUSE SOME STATES OR JURISDICTIONS DO NOT ALLOW THE EXCLUSION OR LIMITATION OF LIABILITY FOR DAMAGES OR THE EXCLUSION OF CERTAIN TYPES OF WARRANTIES, PARTS OR ALL OF THE ABOVE LIMITATION MAY NOT APPLY TO YOU.

These Terms of Use, your rights and obligations, and all actions contemplated by these Terms of Use will be governed by the laws of New York, NY, USA and You and RavenPack agree to submit to the exclusive jurisdiction of the New York Courts. If any provision in these Terms of Use is invalid or unenforceable under applicable law, the remaining provisions will continue in full force and effect, and the invalid or unenforceable provision will be deemed superseded by a valid, enforceable provision that most closely matches the intent of the original provision.