**TABLE OF CONTENTS**

Executive Summary 2

Introduction 3

Data Description 4

Methodology 6

Results 11

Conclusions 20

References 30

*Special thanks to Factset for providing the supply chain data, and Russell Investments for providing the index data:*

C:\Users\phafez.DEVELOPMENT\Documents\FactSetlogo.png



**The RavenPack Data Science Team:**

**Peter Hafez**

*Chief Data Scientist*

**Francesco Lautizi**

*Data Scientist*

**Stefan Duprey**

*Data Scientist*

**Jose A. Guerrero-Colón**

*Data Scientist*

**Mads Koefoed**

*Data Scientist*

[**Contact**](mailto:research@ravenpack.com?subject=Research%20Question)



**Trading intraday**

**Ravenpack structured news feed:**

**A complete intraday event study**

**November 2016**

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 Terms of Use in back cover of this document.

[**www.ravenpack.com**](http://www.ravenpack.com)

**New York – 535 Fifth Ave., 4th Floor, New York, NY 10017 | TEL: +1 (646) 277-7339 |** [**research@ravenpack.com**](mailto:research@ravenpack.com?subject=Research%20Question)

**®RavenPack Quantitative Research 2014 – All Rights Reserved. No duplication or redistribution of this document without written consent**

# Executive Summary

In the race to generate alphas and avoid overcrowded standard factors, investors must identify and extract value from every possible data source.

We show in this paper that Ravenpack news database is an alternative source of alpha apart from standard fundamental data built factors.

In this study, we tackle the problematic of intraday trading reaction to news under the angle of big data analytics for finance from Ravenpack Big Data analytics.

More specifically, we detail:

* By using Ravenpack unique real time automated and pre-processed news data feed based on linguistic pattern recognition we exploit Ravenpack entity detection and event taxonomy classification to build an intraday minute level market reaction profile database and find distinct responses in returns, volatility and trading volumes due to news arrivals. We extract abnormality in minute returns, volatility and volume around the event for the whole Russell 1000 universe.
* By ranking our market reactions according to three dimensions: post-event return, volatility and event frequency, we exhibit how Ravenpack big data deep taxonomy cumulated to its large premium sources scope allows us to characterize finely the market event reaction and thus build accordingly the most profitable customized trading strategy.
* By using Ravenpack analytics aggregated over all its sources, we show how Ravenpack computed event relevance, novelty and sentiment for company-specific news is crucial to filter out noise and to identify significant effects and that sentiment indicators have predictability for future price trends.

**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. Publishers include Dow Jones Newswires, Barron’s, the Wall Street Journal and over 19,000 other traditional and social media sites. Over 15 years of Dow Jones newswires archive and 8 years of historical data from web publications and blogs are available for backtesting. RavenPack detects news and produces analytics data on over 40,000 listed stocks from the world's equity markets, over 2,500 financially relevant organizations, 138,000 places, 150 currencies and 80 commodities.

# 1. Introduction

Previous research has demonstrated (see [2]) that news signal tend to be incorporated in prices quickly, making it advantageous to trade news intraday.

We focus here in this paper on specific intraday market reaction to a Ravenpack equity-specific news flow: we follow Hautsch and Gross-Klussmann (see [1]) methodology to build intraday abnormal return and volatility at the minute level.

The paper is organized as follows:

Section 2 describes the different data sources that we use. Section 3 will detail the overall methodology used to compute our abnormal return, volatility and volume profiles. Section 4 will highlight how to use Ravenpack analytics version 4.0 taxonomy and analytics to filter the most relevant signals and get to the most profitable event driven strategy.

An appendix will detail some of the numerous results shown here to avoid overcrowding our main paper.

# 2. Data Description

In this section, we provide an overview of the datasets that we use in our research.

We describe here RavenPack Big Data’s event taxonomy, which we use to define our events and its analytical metadata which proves to be of great value to ascertain the quality of a profile.

## Ravenpack Big Data Analytics

News event are hard to detect and collect into a unified data base and analyze in a systematic fashion. Ravenpack does provide you with real time structured content fetched from premium quality financial news sources.

We build our event trading profile library using Ravenpack Big Data Analytics (RBDA) taxonomy, event detection and entity detection.

Ravenpack big data analytics possess all the required characteristics to build our market event reaction database:

* **Ravenpack big data taxonomy:**

Ravenpack automated event detection and assignment to its hierarchical taxonomy allows you to dive to the most granular level and detect the slightest market moving event for a company. The fine granularity of Ravenpack company specific event taxonomy will prove fundamental to build the most profitable calibrated event driven trading strategy.

We here want to emphasize that Ravenpack Big Data analytics nearly triples the number of market-moving events we detect (up to 6753 categories), capturing most of the nuances in language towards a tradable company specific news. This feature will prove mostly relevant to finely assess market reaction and make the most comprehensive event study.

* **Ravenpack entities:**

Ravenpack news sources cover spans more than 41500 companies, 138500 places, 160 currencies and 85 commodities starting from 2000. We will focus our event studies on Russell 1000 universe for its high liquidity.

* **Historical data:**

Up to 16 year of millisecond time-stamped data available for market reaction profile construction. To evaluate our profiles, we will consider RavenPack data spanning a period of 10 years covering 2007 through 2016. This historical availability is a key feature to help you build and backtest event driven strategy.

* **Ravenpack real time feed:**

All news article are assessed within milliseconds of receipt and the resulting data is immediately pushed to the user, which is an essential component of intraday news trading: processing velocity. We are collocated with our news feed provider and electronics markets thus guaranteeing to process and deliver the news analytics to a machine readable structured content within 200-250 milliseconds form the moment we receive the news

As we see in many profiles, a significant jump happens at the very event minute. The sooner the reaction the greater the return you might capture: being timely is the essence of a good event-driven trading strategy.

* **Ravenpack analytics:**

Ravenpack applies entity and event specific relevance, novelty and sentiment scores to all events using proprietary techniques. This paper will show the importance of those analytics to assess the market reaction: we will see how Ravenpack analytics like news novelty or the more novel a news is, the more the market will move. The more relevant the event is, the more the market will move. The higher our sentiment is, the more the market will move.

* **Ravenpack premium sources:**

And last but not least, the power of analytics, velocity, and refined taxonomy would be nothing without Ravenpack broad scope of high quality financial news provider. Each of those sources focus on certain specific and brings its expertise which when aggregated through Ravenpack brings the great quality our data has always shown.

We will split up our sources study in two different specific types: Dow Jones news, Premium pack and news content.

We will show that Ravenpack Premium pack (containing Benzinga, The Fly, MT\_Newswire, FX\_Street\_News, Alliance News, Ticker report and SleekMoney) brings great value through their specific expertise, especially their exclusive scoop to our M&A, analyst ratings and earnings groups. This result will be highlighted later when ranking our profiles per sources. We will see that those sources add real value for analyst ratings and merger and acquisition.

## Trading volume and Pricing Data

As we focus on intraday trading patterns, we use minute data provided by Algoseek for our Russell 1000 universe. Those minutes contain standard Open-High-Low-Close prices and trading volumes. We derive from them minutes return, volatility and volume.

# 3. Intraday event minute study methodology

## 3.1 Abnormal returns definition

Abnormal returns are the crucial measure to assess the impact of an event. The general idea of this measure is to isolate the effect of the event from other general market movements. The abnormal return is defined as the difference of the realized return and the expected return given the absence of the event.

The expected return is usually called "normal return" in academic literature and is by definition unconditional on the event.

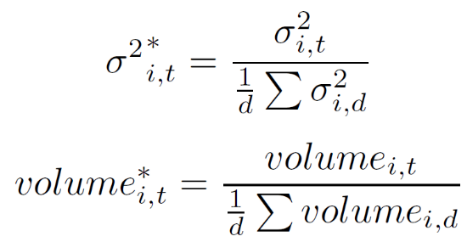
This normal return is to be modelled and conditioned on a separate information set, whose definition might vary.

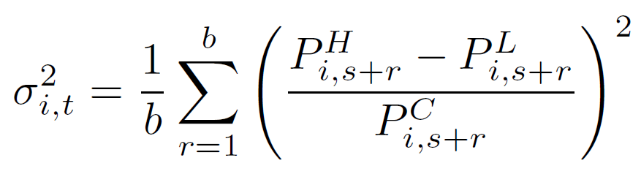
The definition of this past pre-event asset returns set leads to various models for the normal return.

## 3.2 Abnormal volume and volatility

One other very important aspects of intraday high frequency trading (others than to which extent high-frequency movements affect returns) are volatility, liquidity, trading sizes, imbalances and spread. We will here focus mainly on volatility (as a proxy for liquidity, size and spread)

and trading volume at the minute level.

We define abnormal volume and volatility as :

where

defines the minute volatility with b being the number of minutes in an aggregated bucket, d is the number of days in the estimation period and H,L,C refer to high, low and close price of the specific minutes inside the bucket.

Volatility and volume are crucial factors to assess that the market reacts to news independently of any direction and to get a sense of the scalability of any given strategy.

All volatility and liquidity variables exhibit pronounced intraday trading patterns. To capture the seasonality and to account for an "abnormality" measure, we define our abnormality as the ratio of the bucket measure divided by the averaged same quantity over the estimation period for all the same minutes of the day.

## 3.2 Normal return model

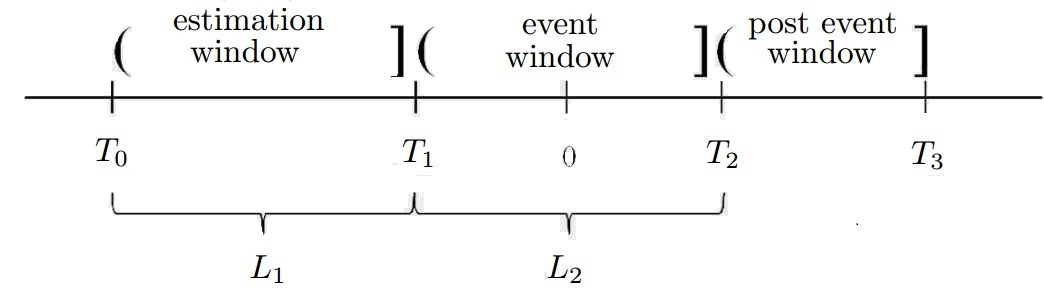
A substantial feature of an event study is the choice of the appropriate normal return model.

All models contain parameters that need to be estimated, even the simpler constant mean return model.

The time period over which those parameters are estimated is commonly denoted as the estimation window.

Since the normal return is the expected return in absence of the event, overlapping event and estimation windows should be avoided.

Otherwise normal return model parameters are estimated from returns affected by the event.

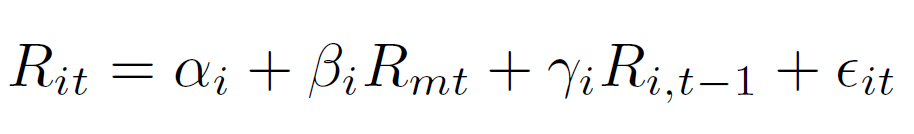
We will here adopt the most common approach by restricting the estimation window (T1) to the time period prior to the event window (T2).

**FIGURE 2: Estimation and event window definition.** L1 is a 6 months, L2 +/- minutes around the event.

*Source: RavenPack, November 2016*

T1 will span over 3 months before the event and T2 size will be 180 minutes spread around the event and we will restrict ourselves to 5 minutes OHLC event bucket data and all quantity will be computed by bunch of 5 minutes bucket.

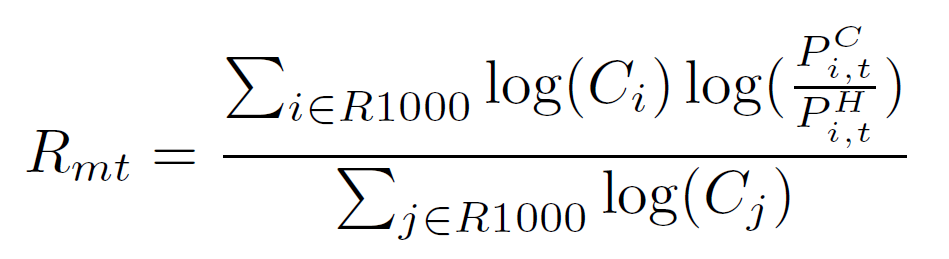
Many market models exist only for daily event studies (CAPM, multifactor models ...). Those models complexity is based on the factors they incorporate in addition to standard betas (market factor) like for instance more advanced Fama-French factors. Those factors do not translate easily to intraday high-frequency model as they are built on a daily frequency based on fundamental data.

We here follow the methodology from Gross-Klussmann and Hautsch in [1]. We assume the following market model for normal returns:

Where t denotes the underlying 5 minutes bucket, Rmt is the market return, whose definition is previously exposed for the Russell 1000 universe and Rit is the bucket return for stock I and minute bucket t. To capture return dynamics on high frequencies, we also include lagged returns from the previous bucket.

## 3.3 Russell 1000 minute level market index

We construct our own Russell 1000 market index at the minute level by aggregating a log capitalization weighted mean over all the Russell 1000 constituents minutes returns for all traded minutes of the day.

The log return of our market index is computed as:

We end up with intraday minute values for our R1000 reference index, which will the basis to compute our beta market factor and compute our abnormal returns.

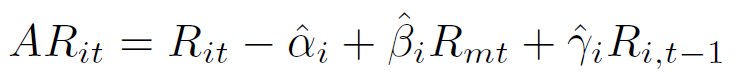
## 3.3 Model parameters calibration

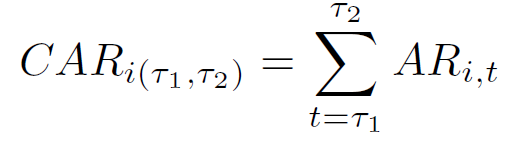
Our normal return model parameters are estimated based on a complete minute-level return time series without including the event windows for each R1000 stock.

The size of the estimation windows is chosen to be three months. The longer the estimation the more stable our betas and gammas will be.

We restrict ourselves to 3 months mainly due to computational issues and simply regress the times series to compute our estimated alpha and beta for each of our stocks, to which we will append a ^. Let's note here that we tested many different market models (constant market model, daily betas market model) and observe always the same abnormality patterns.

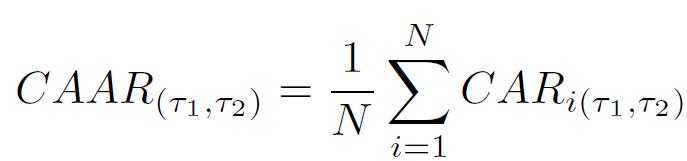
## 3.4 Return profile

Our model parameters being calibrated, we compute the abnormal returns:

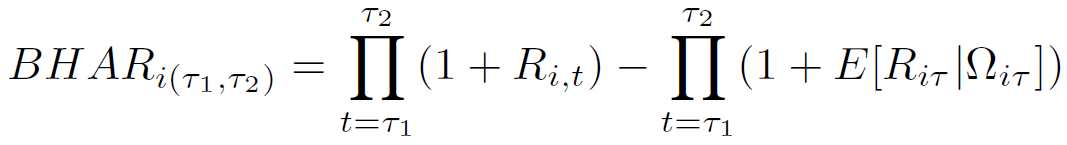
Cumulating those abnormal returns across time yields the cumulative abnormal return:

This measure is the one we have chosen to assess our abnormal return in the event window.

We will show the cross sectional average cumulated abnormal return around event news minute with 180 minutes padding bucketed by 5 minutes.

For cumulative abnormal returns the cross-sectional average (CAAR) is computed as:

We will center our CAAR around the event windows and make it null for the very bucket containing the timestamp of Ravenpack captured event, so that the potential returns we could make if entering the market in the minute after the news will be displayed past zero (t>0).

Let's just mention here some other measures we tested: the buy-and-hold abnormal return (BHAR), which is defined as the difference between the realized buy-and-hold return and the normal buy-and-hold return

## 3.5 Statistical significance

In our search for profitable trading profiles, we will use Ravenpack analytics to its full extent to filter per source, news novelty, entity relevance, and so on, but as we filter our events, the observations count thus lessens and to properly discriminate among those distinct cumulative trading profiles, one must devise a proper statistical significance measure to ponder the potential average compounded return possibly tradable.

Many of those statistical tests have been extensively used and detailed in academic literature. We have tested most of them and retained only the most robust one to rank our profiles: the generalized Corrado rank test. We here follow once again [2] path.

We use the generalized cumulative Corrado test centered on the event: the whole methodology being detailed in Appendix A. This is the best way to assess statistically the event impact over minute’s returns both after and before Ravenpack event timestamp. We will see that the behavior of those tests really depends on the news soft or hard type. The more unexpected the news is, the bigger the statistical significance post event is. The more anticipated the news is for hard scheduled news, the bigger the pre event statistical significance is.

# 4. Results

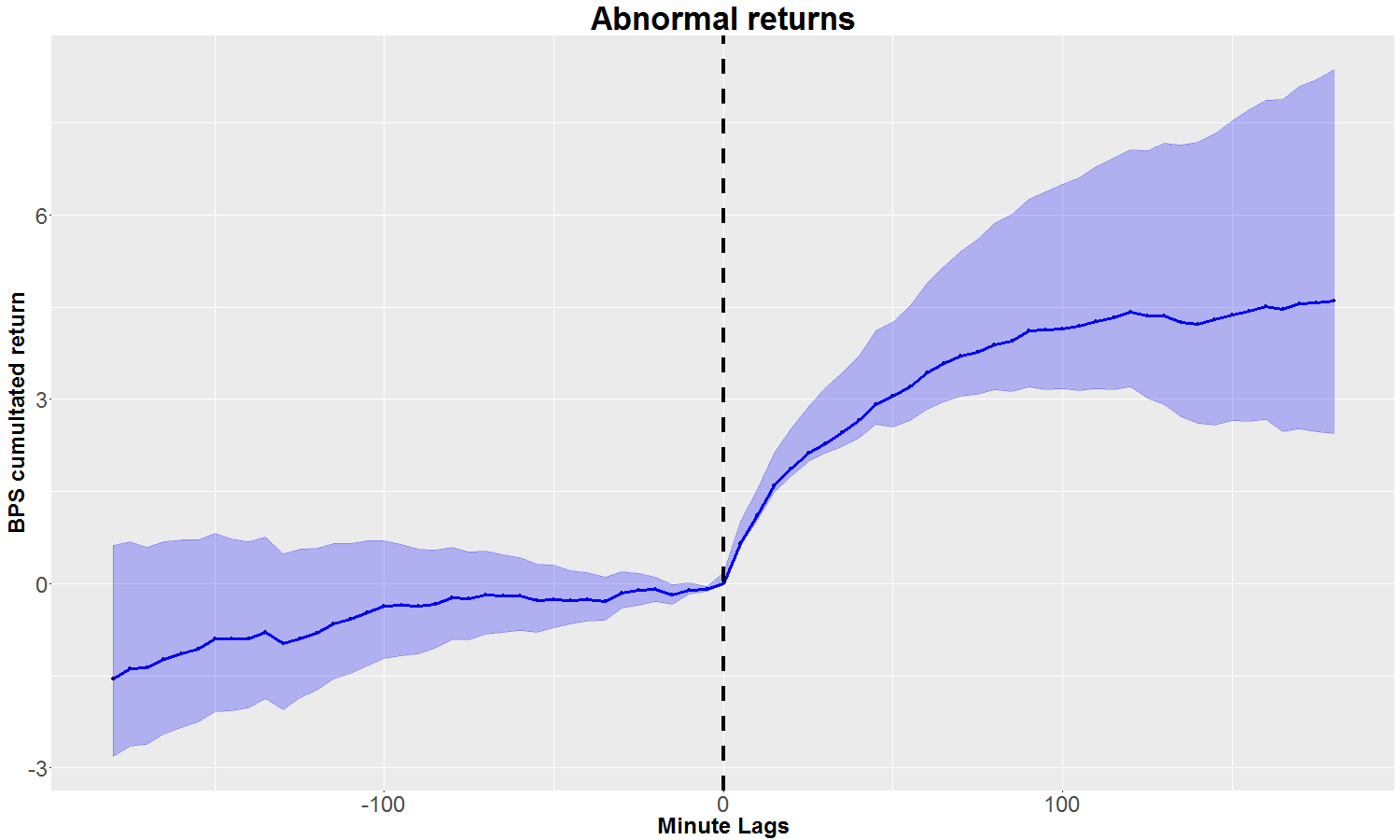
Having described our methodology, in this section, we evaluate the quality of our profiles. We begin by detailing the overall average behavior at Ravenpack most granular level.

Let’s here once again emphasize the scope of the study: we average overall Ravenpack big data company specific categories (up to 6753) over the last ten years. For each category, we build profiles for a Cartesian grid of Ravenpack analytics: novelty, event relevance, similarity gap and source type. We end up here with more than 100 000 market reaction profiles. The average profile is aimed at giving us some raw overall information.

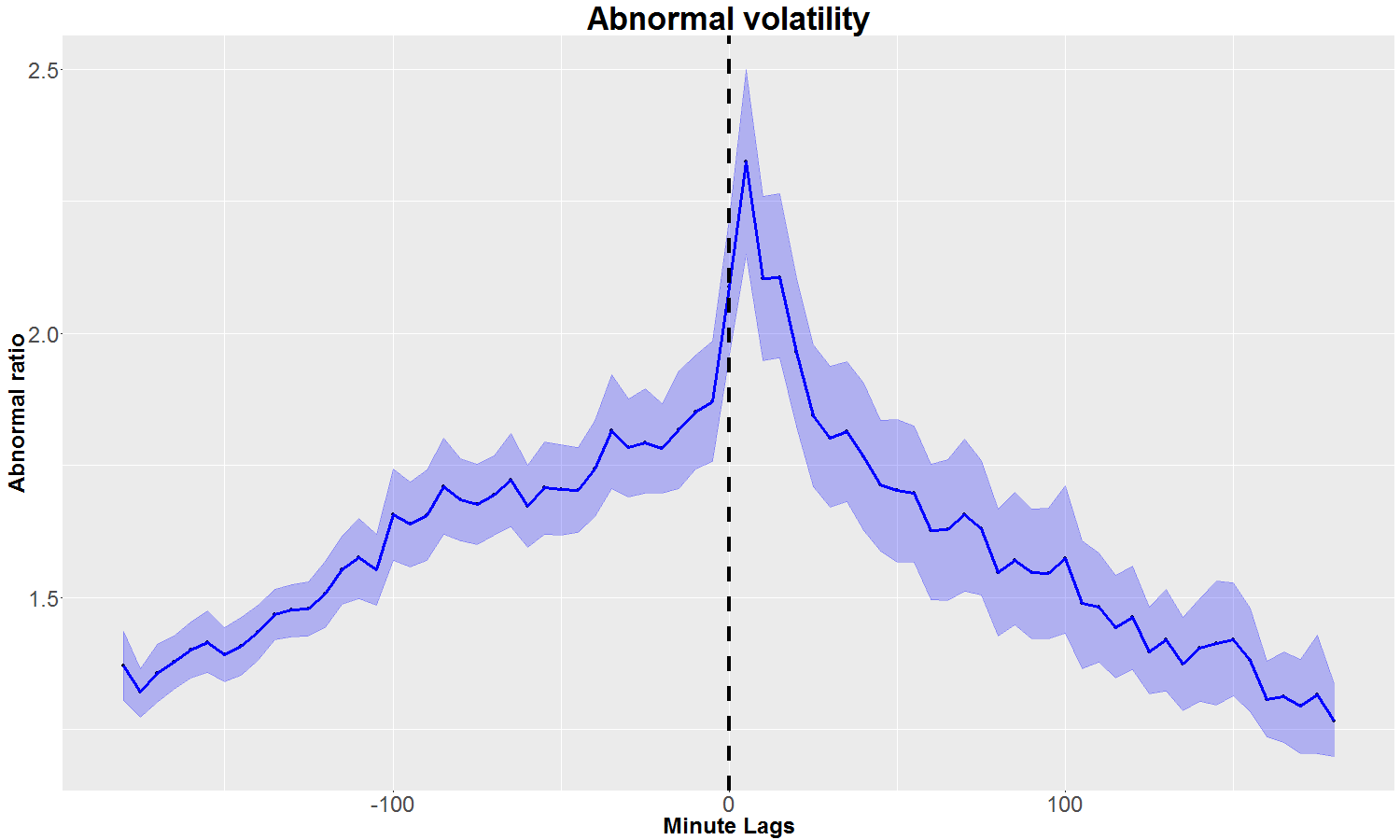
# 4.1 Average profile behavior

We here show the average market reaction behavior.

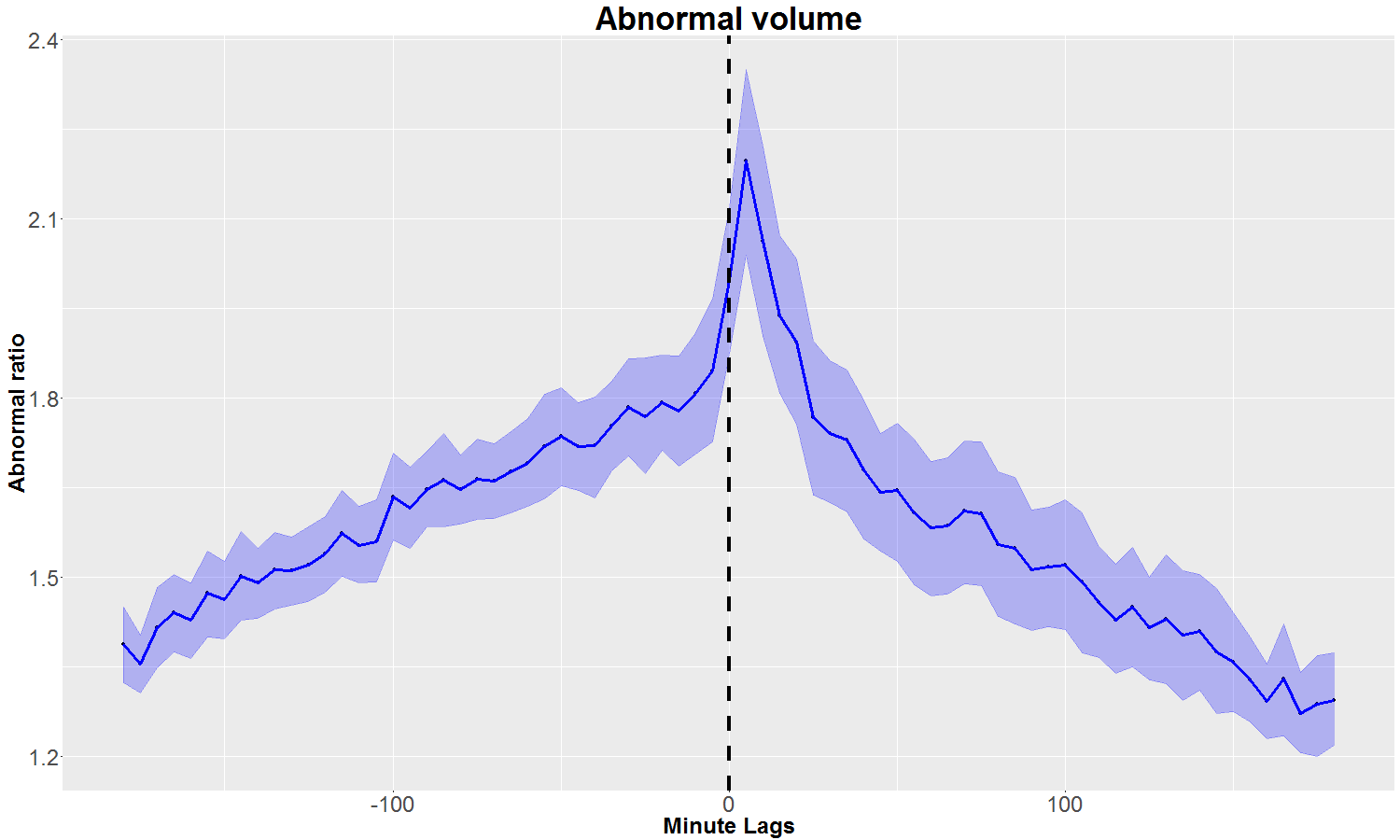
To avoid overfitting, we here average only for all taxonomy events with the same filtering methodology: similarity gap greater than one day, high entity and event relevance. We compute the average profile overall Ravenpack sources. We here apply no specific pondering on the profiles according to their frequency. We will see in the next paragraph a way to filter category profiles according to different metrics.

**Average abnormal returns, volatility and return at the category level:**

**FIGURE 4: Average cumulative abnormal returns across all category profiles.**  This average profile is computed over all Ravenpack sources filtering for similarity gap greater than one day and high event/entity relevance. Negative sentiment categories have been reversed. The event window spans between -180/180 minutes around the event.

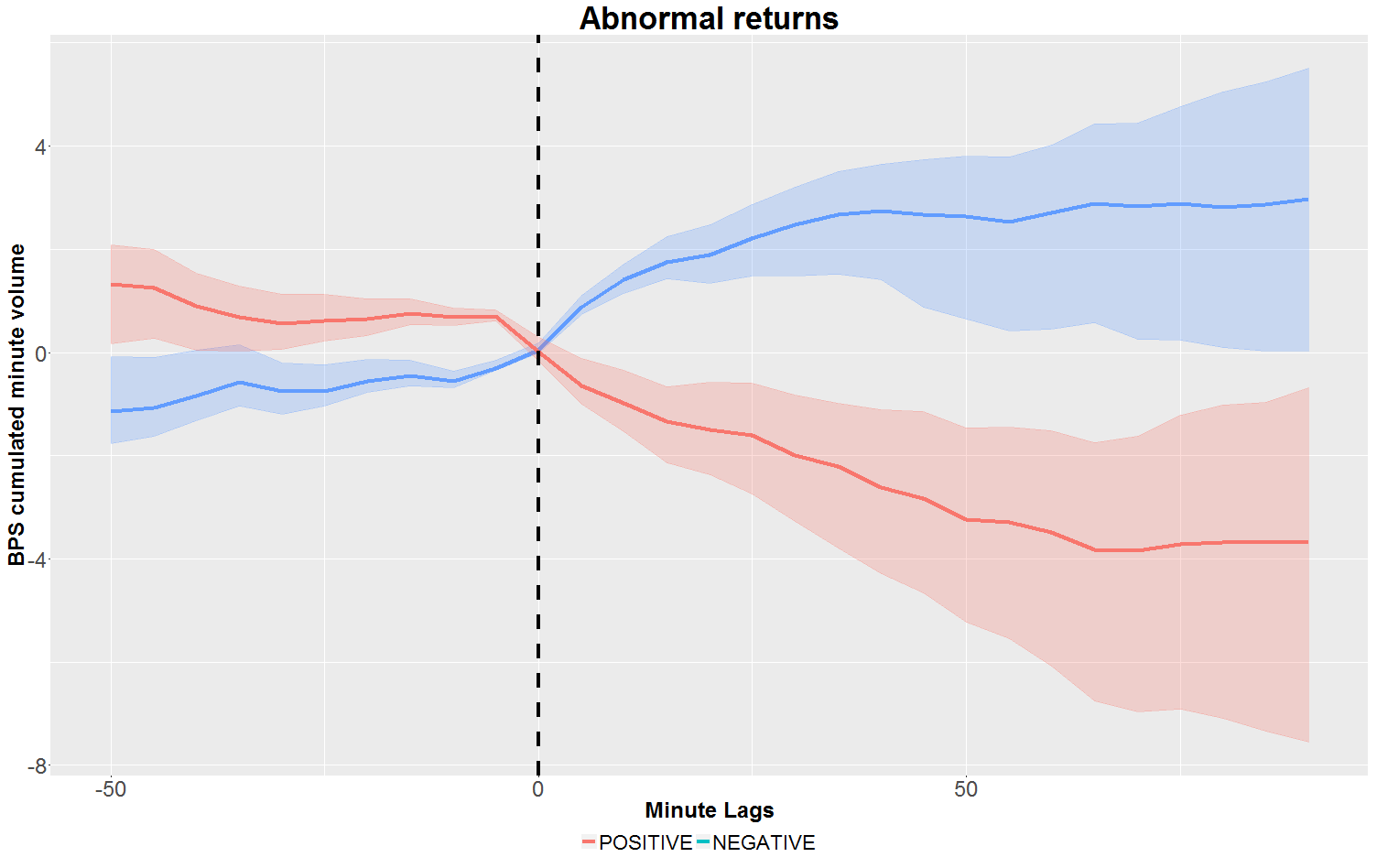


**FIGURE 5: Average abnormal volatility across all category profiles.** This average profile is computed over all Ravenpack sources filtering for similarity gap greater than one day and high event/entity relevance



**FIGURE 6: Average abnormal volume across all category profiles.** This average profile is computed over all Ravenpack sources filtering for similarity gap greater than one day and high event/entity relevance

# 4.2 Ravenpack sentiment

Ravenpack analytics provides each flagged event with a sentiment score, whose sign gives the direction of the prediction and the absolute value the magnitude of the news. We use here only the sentiment direction by splitting our incoming news in a positive or negative bucket.

**FIGURE 7: Average abnormal return split by positive versus negative sentiment.** This average profile is computed over all Ravenpack sources filtering for similarity gap greater than one day and high event/entity relevance

Let’s notice here that we found back a finance market well known stylized fact: financial market tends to react in a stronger way to negative news, showing an asymmetry to the negative side.

# W:\OutputData\sduprey\NAR-326\PAPER_PICTURES\BEST_PROFILES\card_post_ranked_returnbigdataf_significance_group_average_profile_ci.png4.3 Statistical significance

**FIGURE 7: Average Corrado statistical profile across all categories.** This average profile is computed over all Ravenpack sources filtering for similarity gap greater than one day and high event/entity relevance

We here show the average generalized Corrado rank test across all category profiles with the same common filtering as before to avoid overfitting. The generalized Corrado test is cumulated from both sides of the event.

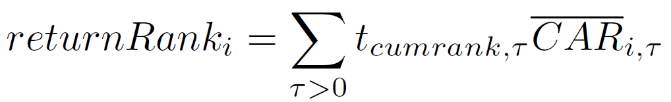
The Corrado test clearly argues for an event happening and that Ravenpack analytics clearly flags a novel market moving event at the right time.

Similar results are found at the group level. It exhibits the same patterns for returns, volatility and volume, but slightly less emphasized. As groups are more frequent, the signal tends to be less strong, with a lower expected cumulated abnormal return and surge in abnormal volatility. Using Ravenpack full taxonomy granularity up to the category is the best way to create the most profitable event driven strategies. Nevertheless trading at the group level can still be used for very frequent statistical arbitrage. We display the results for the group level in the appendix.

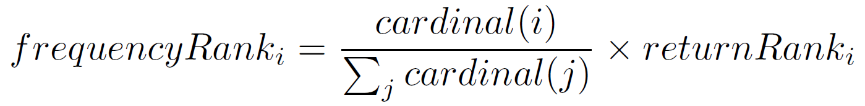
# 4.3 Investigating our profiles

# 4.3.2 Ranking profiles

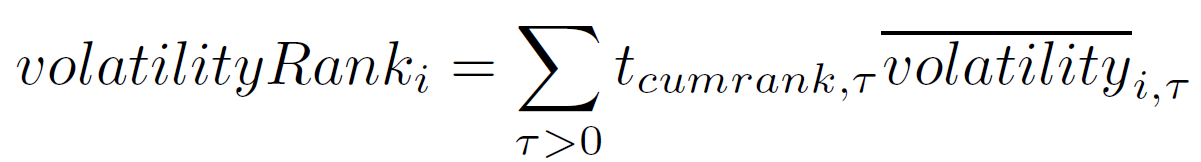
Ranking a profile is an arduous tasks. First the value of a profile is only to be asserted according to the trading strategy you want to apply in the aftermath of the event. We here devise three ranking metrics to ascertain the quality of a profile. Those metrics are chosen to fit to different type of strategies.

* **ABSOLUTE RETURN** :

“Absolute return” just favors absolute statistically confident returns: the best ranking categories for that metric are the categories you must focus on to get the most profitable event driven strategies.

* **FREQUENCY RETURN** :

“Frequencyreturn” ponders the confident absolute return with the frequency of the category. You here get the perfect categories for statistical arbitrage, which show both confident return and frequency.

* **VOLATILTY** :

“Volatility” favors only post event volatility without considering return. This metrics is devised to deal with category with a more sophisticated reaction (reversal behavior, uncertainty). This metric highlights categories where you should rather device volatility based strategy to benefit from the event.

For a deeper investigation of our average market reaction profile, we will now average our profiles only among the first decile of our categories sorted according to the three different methodologies.

We know that this opens up for overfitting, but the purpose is to highlight interesting profile types. Future research will try to focus on in-sample vs. out-sample return predictions.

to all categories as previously done for the average profile

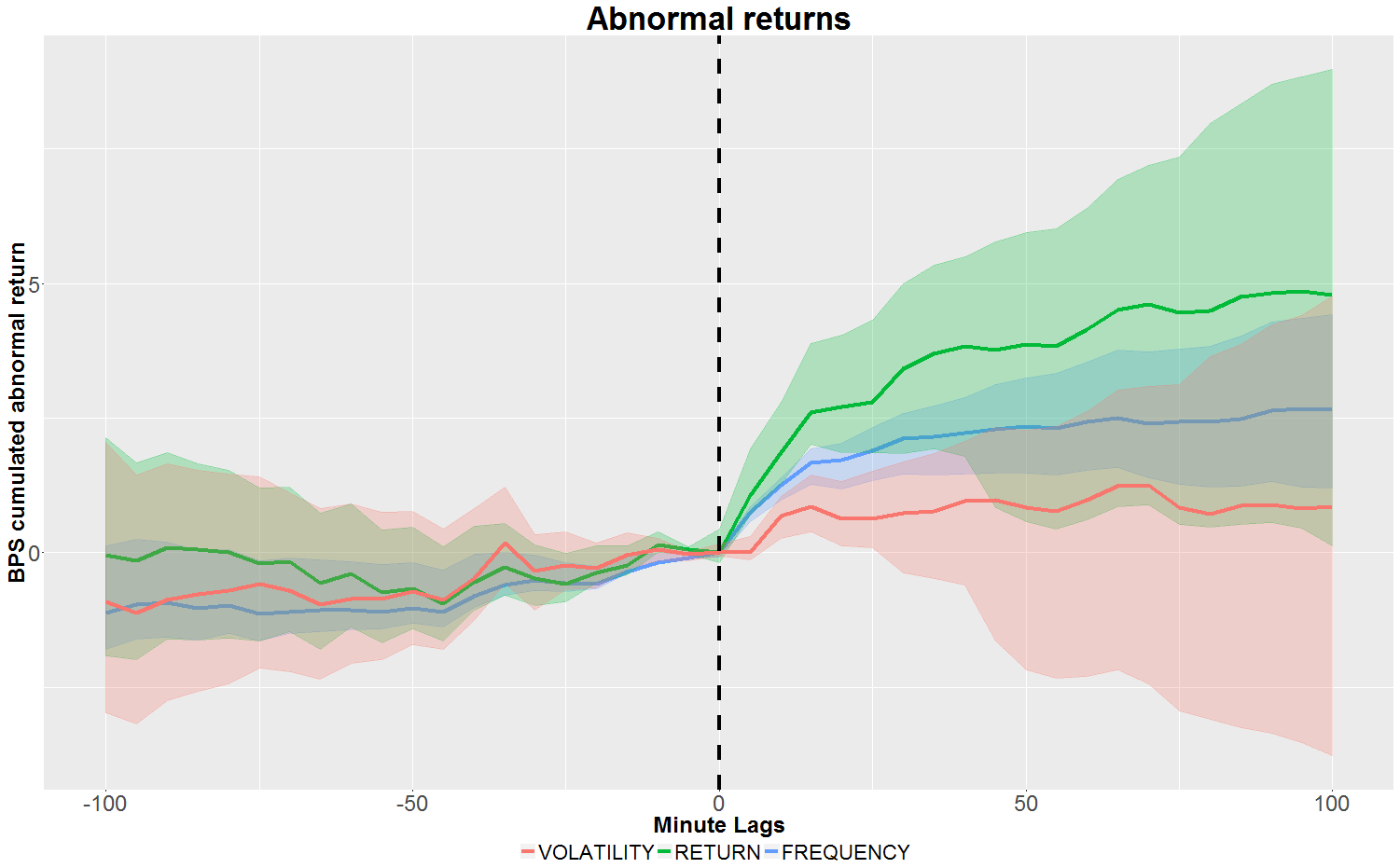
We will keep consistently with red color for the volatility type, green for the pure confident return and blue for both return and frequency.

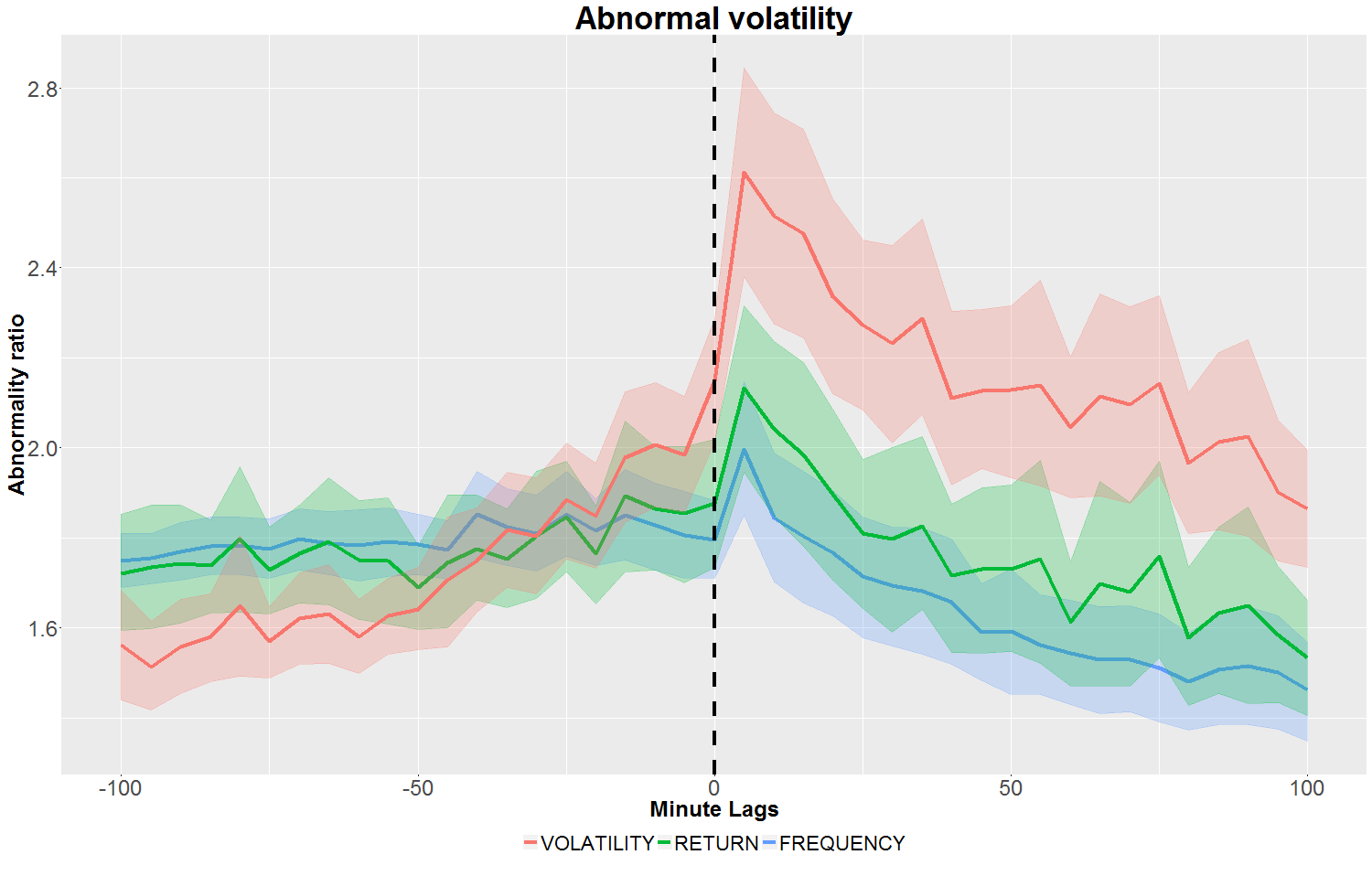
We will also investigate Ravenpack source effect and split our results for two different Ravenpack sources types: Dow Jones, Premium pack sources, which are the most prone to intraday reactions.

Web content sources being the results of frequent crawling, market reaction are less point-in-time sensitive. Nevertheless Web content is still a very good source to catch the trend of abnormal return for a stock.

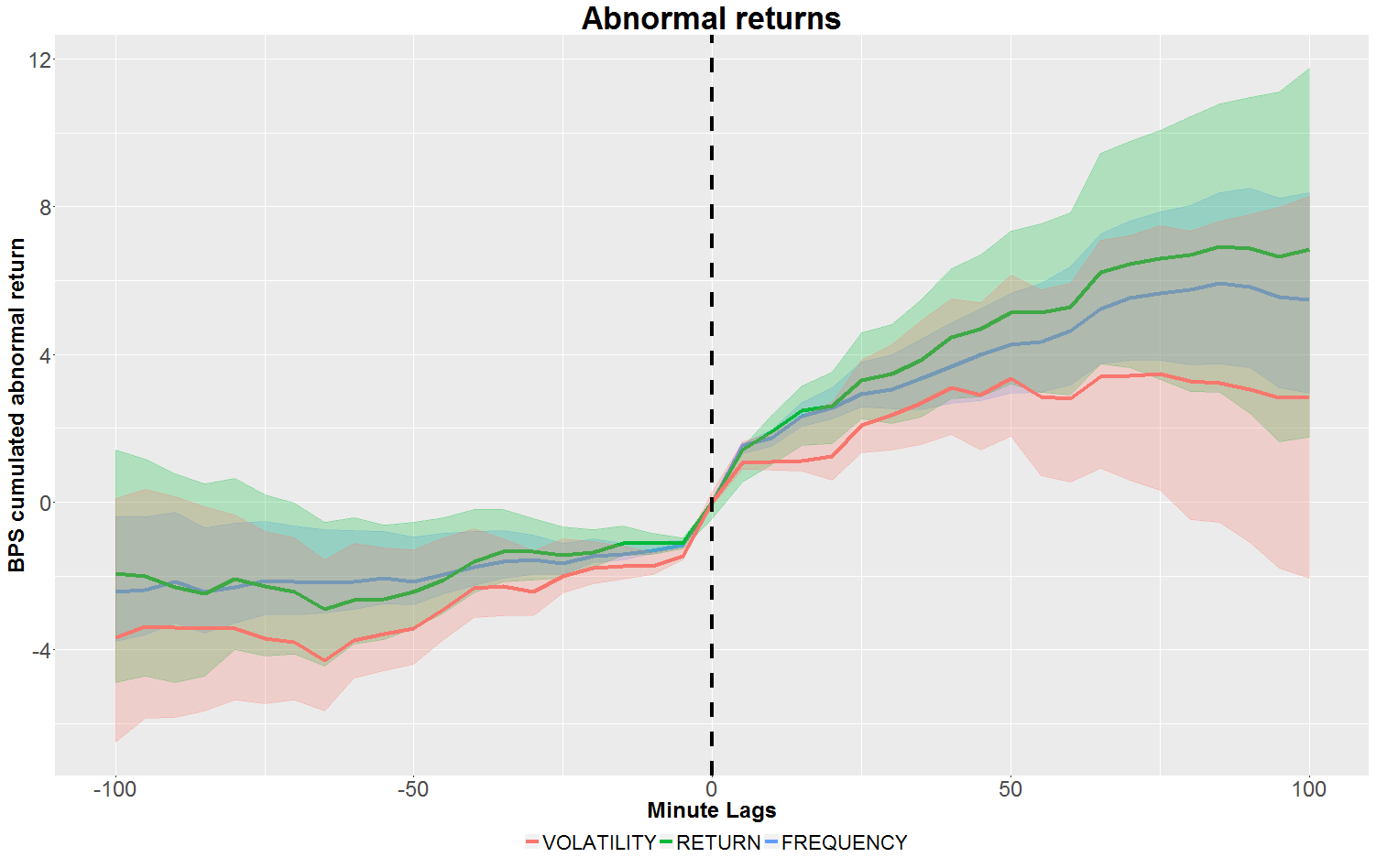
We will see here that Dow Jones wire and Ravenpack premium pack both are sources generating a strong market reaction in abnormal return and volatility surge.

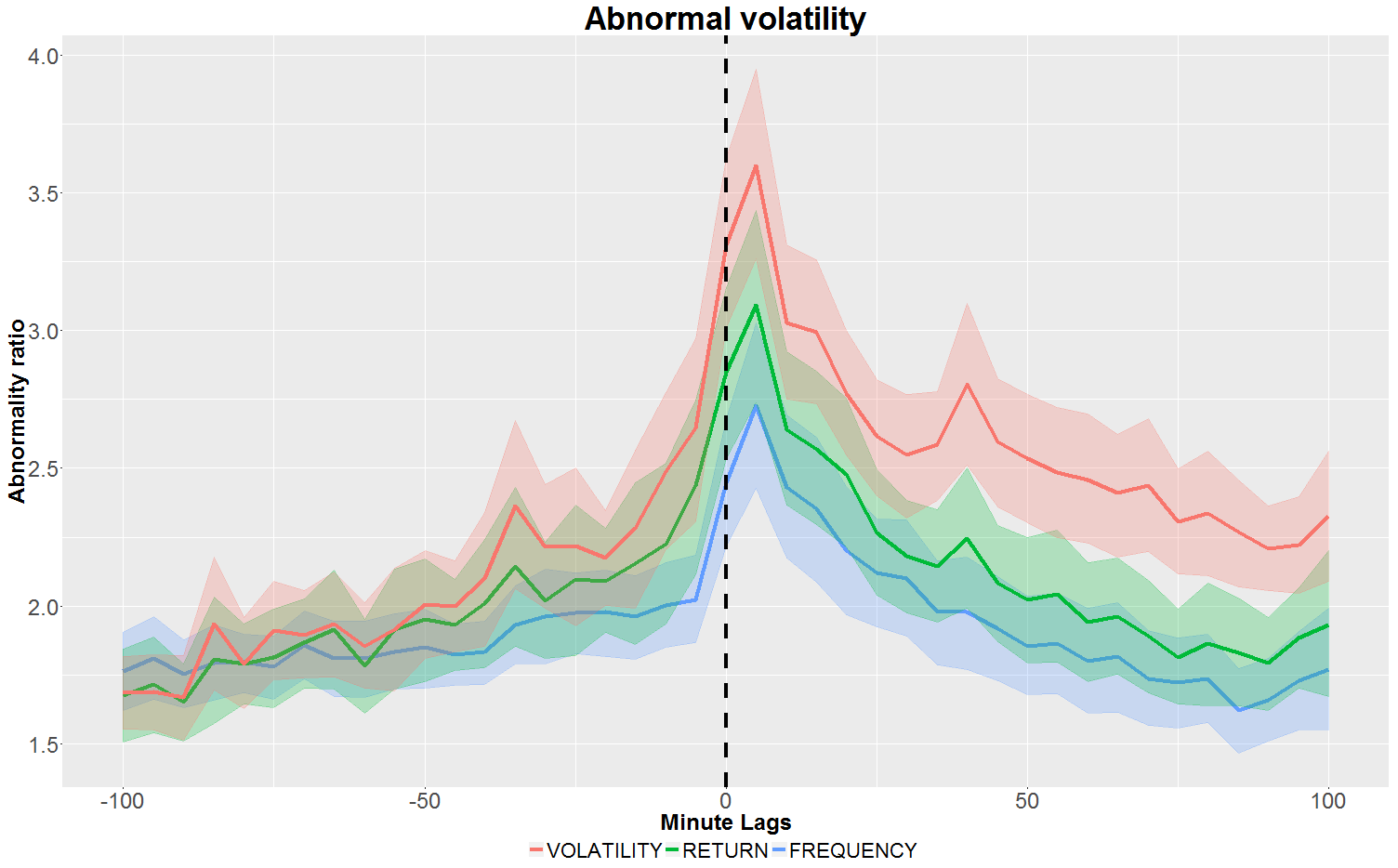
The premium pack even seemingly slightly over performs in absolute returns metrics.

**Dow Jones news feed:**

 **FIGURE 7: Average abnormal return across all category profiles for Dow Jones source average for across the first 100 ranked categories according to three metrics. RED favors volatility, BLUE frequent return, and GREEN absolute return.** This average profile is computed filtering for similarity gap greater than one day and high event/entity relevance

**FIGURE 8: Average abnormal volatility across all category profiles for Dow Jones wire average for across the first 100 ranked categories according to three metrics. RED favors volatility, BLUE frequent return, and GREEN absolute return.** This average profile is computed filtering for similarity gap greater than one day and high event/entity relevance

**Ravenpack premium sources:**

**FIGURE 9: Average abnormal return across all category profiles for PREMIUM PACK sources averaged across the first 100 ranked categories according to three metrics. RED favors volatility, BLUE frequent return, and GREEN absolute return.** This average profile is computed filtering for similarity gap greater than one day and high event/entity relevance

**FIGURE 10: Average abnormal volatility across all category profiles for PREMIUM PACK sources averaged across the first 100 ranked categories according to three metrics. RED favors volatility, BLUE frequent return, and GREEN absolute return.** This average profile is computed filtering for similarity gap greater than one day and high event/entity relevance

# 4.3 Using Ravenpack analytics to refine our profiles

We will now investigate how Ravenpack analytics can help you filter out sharper market response.

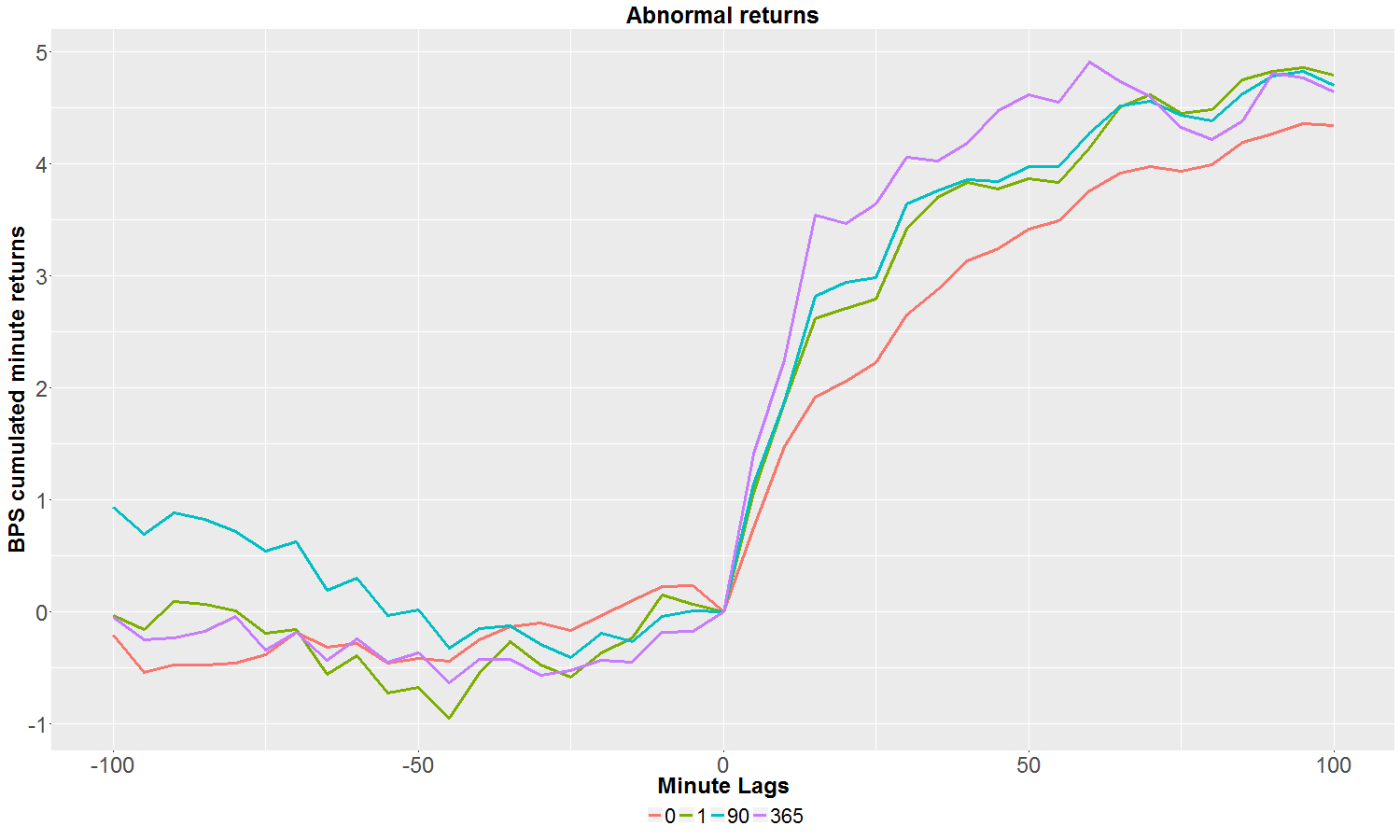
We will here investigate mainly two very important dimensions: Ravenpack novelty metrics to assess how novel a news is and Ravenpack event relevance to assess the strength of an event.

Ravenpack analytics 4.0 has significantly improved both analytics.

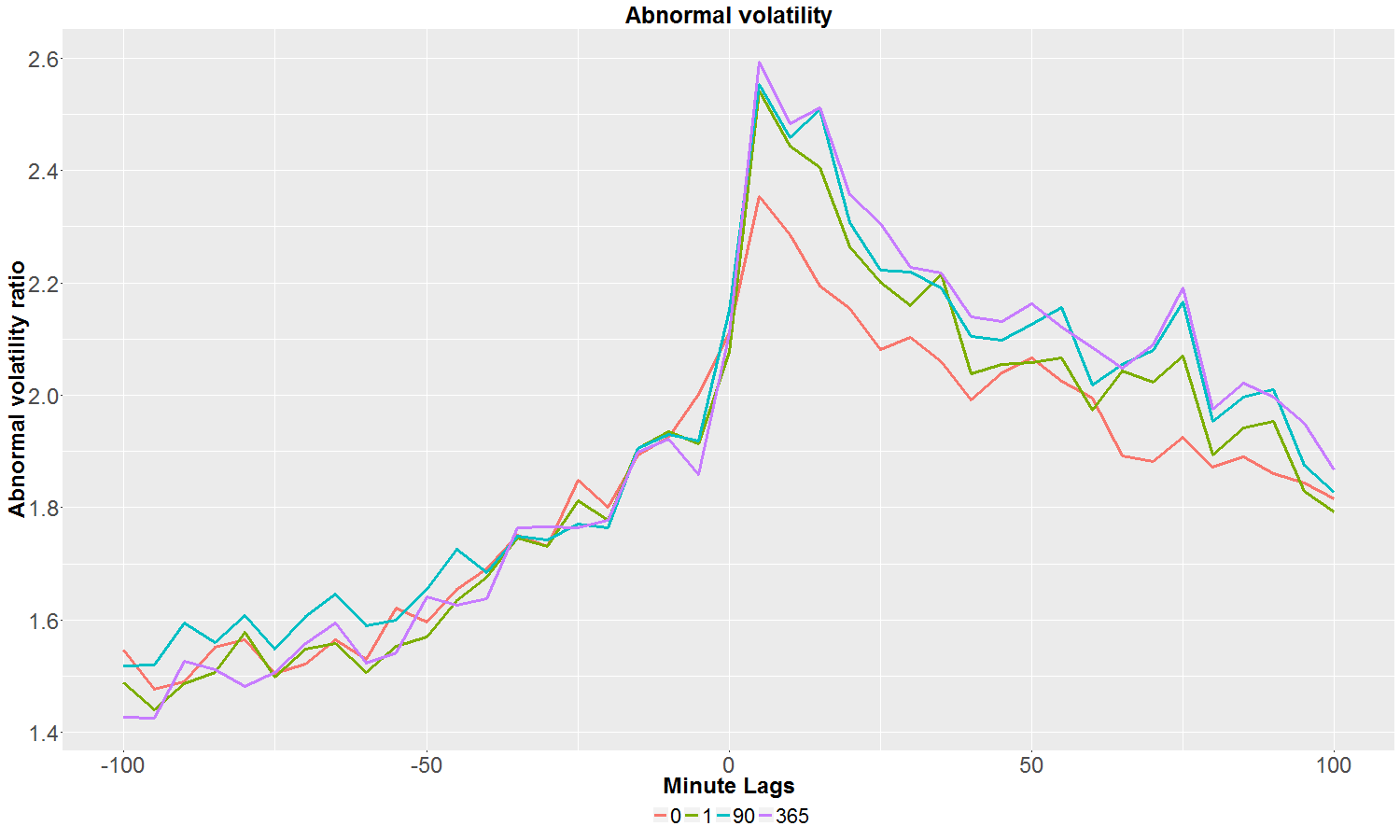
Ravenpack computes a similarity metrics for a specific event for an entity across all its sources. The 4.0 analytics version has increased the look back window from 100 to 365 days.

The following results expose results for Dow Jones and Premium pack type showcasing how the similarity gap filtering affects the profile.

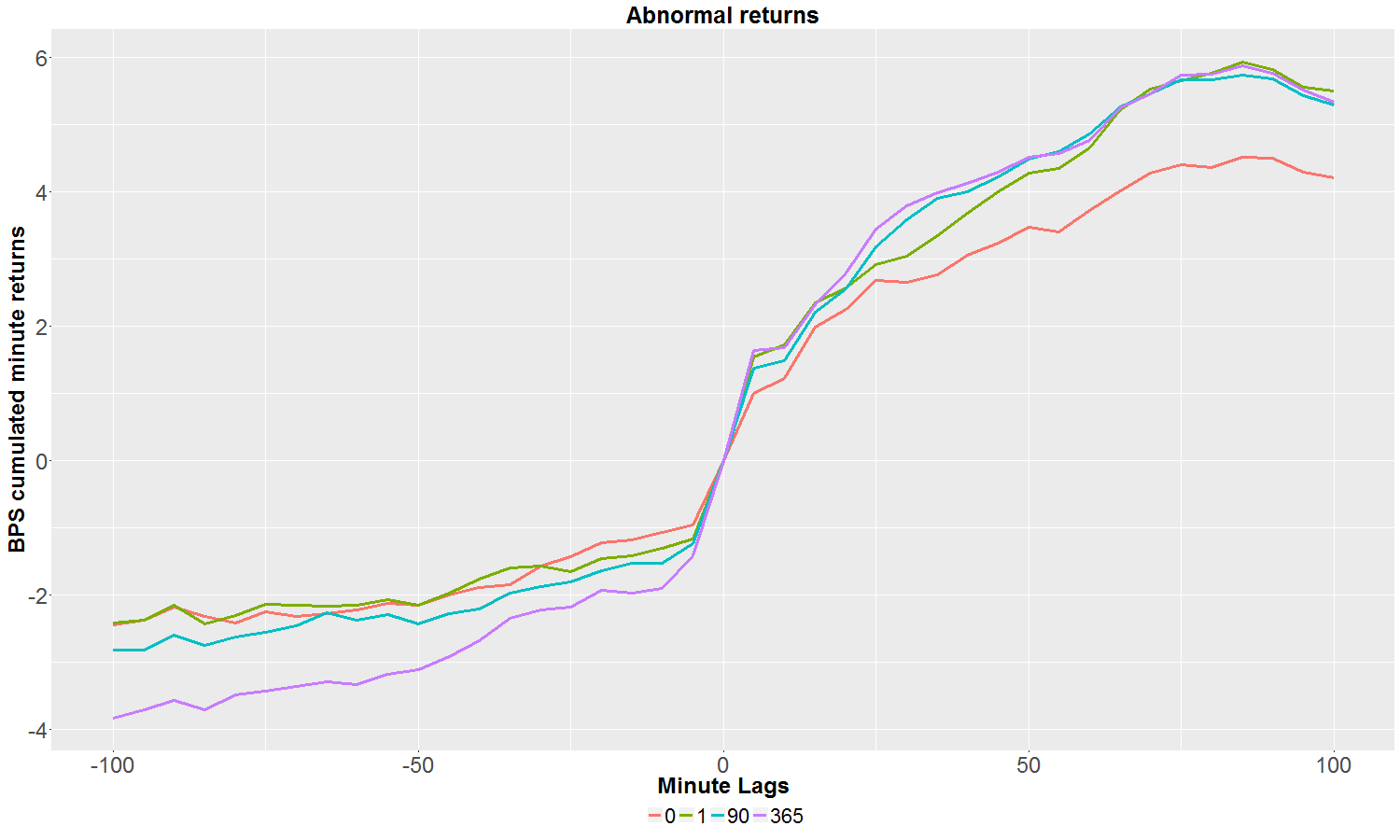
# 4.3.1 Ravenpack novelty assessment

**Dow Jones news feed**:

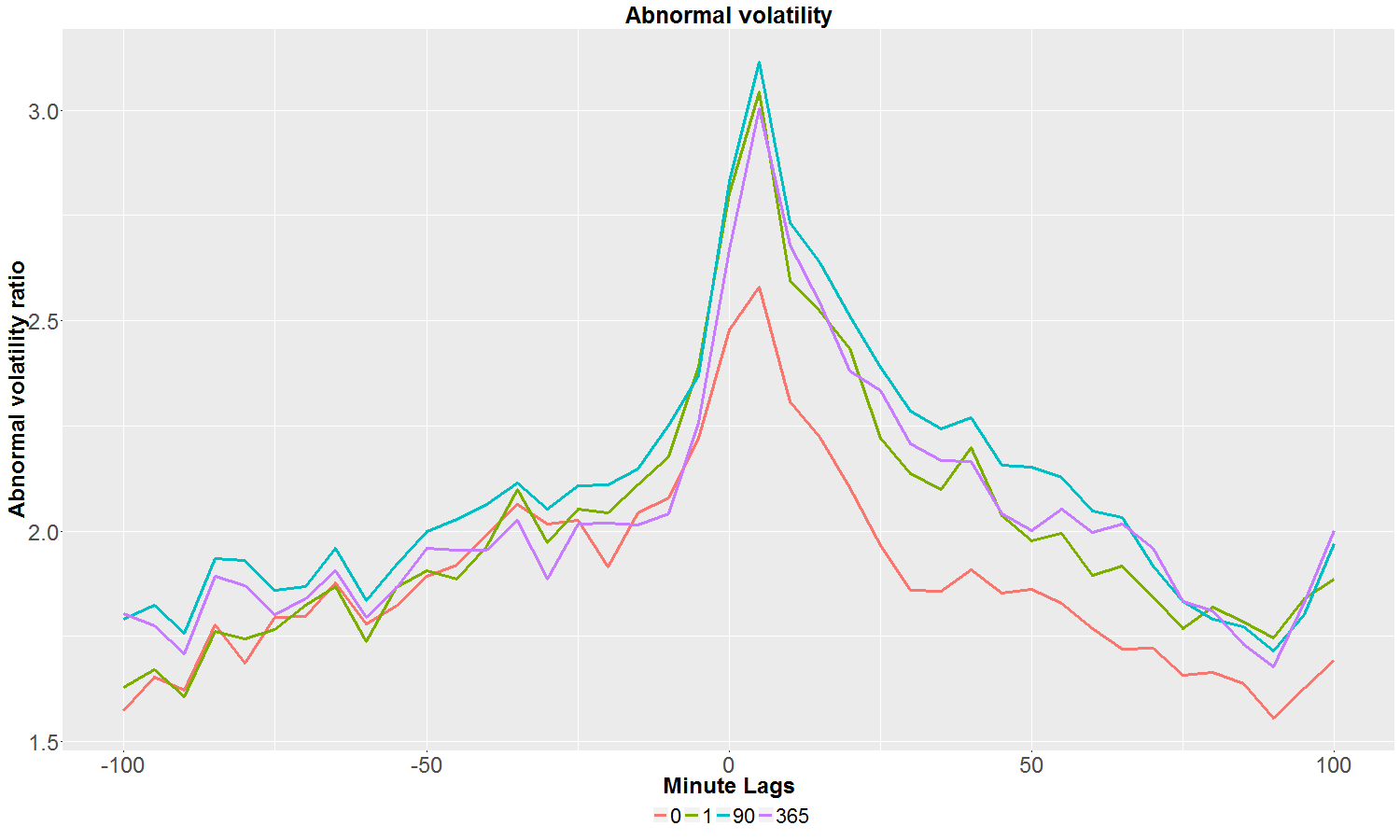
**FIGURE 13: Similarity gap filtering effect on abnormal return for Dow Jones wire average profile.** This average profile is computed filtering for high event/entity relevance



**FIGURE 14: Similarity gap filtering effect on abnormal volatility for Dow Jones wire average profile.** This average profile is computed filtering for high event/entity relevance

**PREMIUM PACK:**

**FIGURE 15: Similarity gap filtering effect on abnormal return for PREMIUM PACK average profile.** This average profile is computed filtering for high event/entity relevance

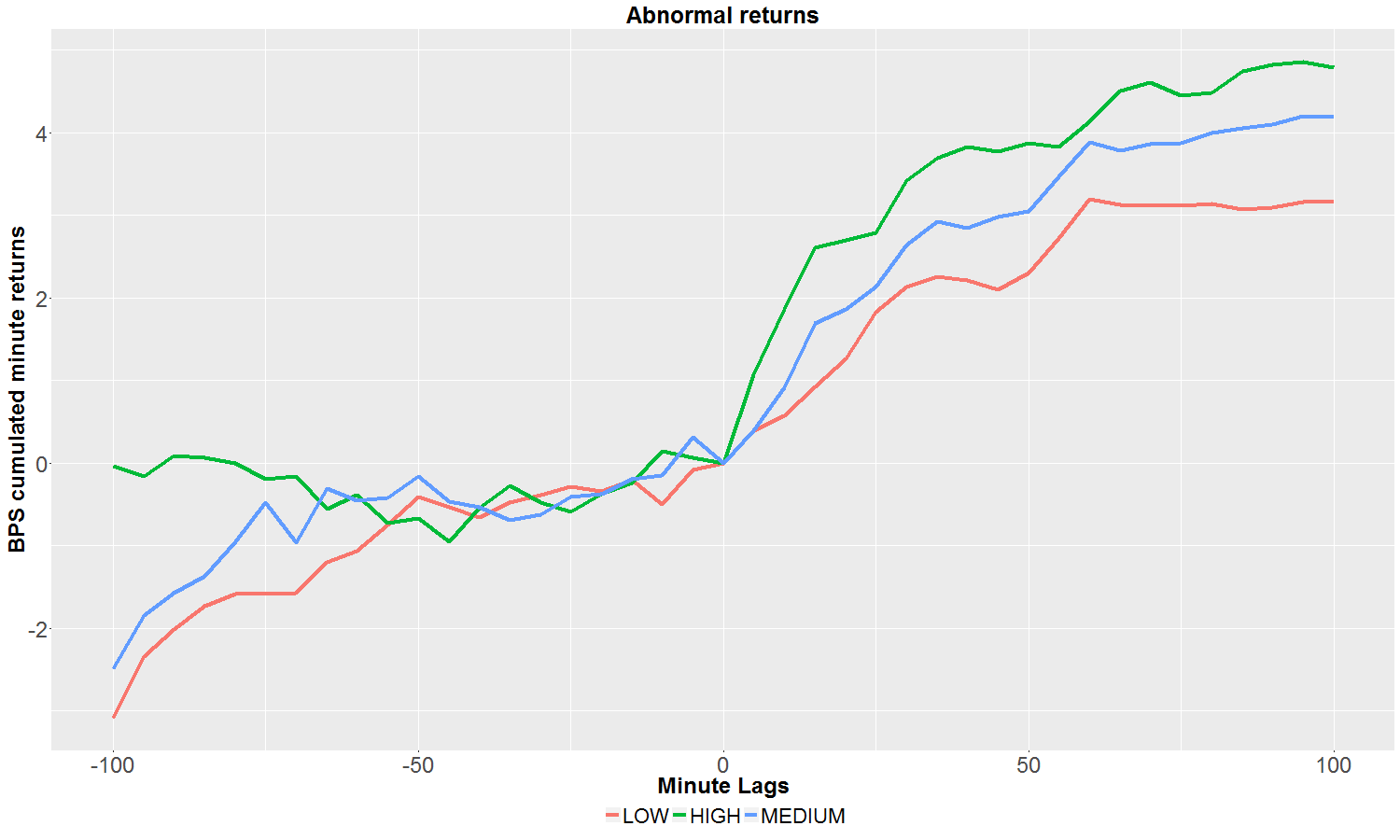
****

**FIGURE 16: Similarity gap filtering effect on abnormal volatility for Dow Jones wire average profile.** This average profile is computed filtering for high event/entity relevance

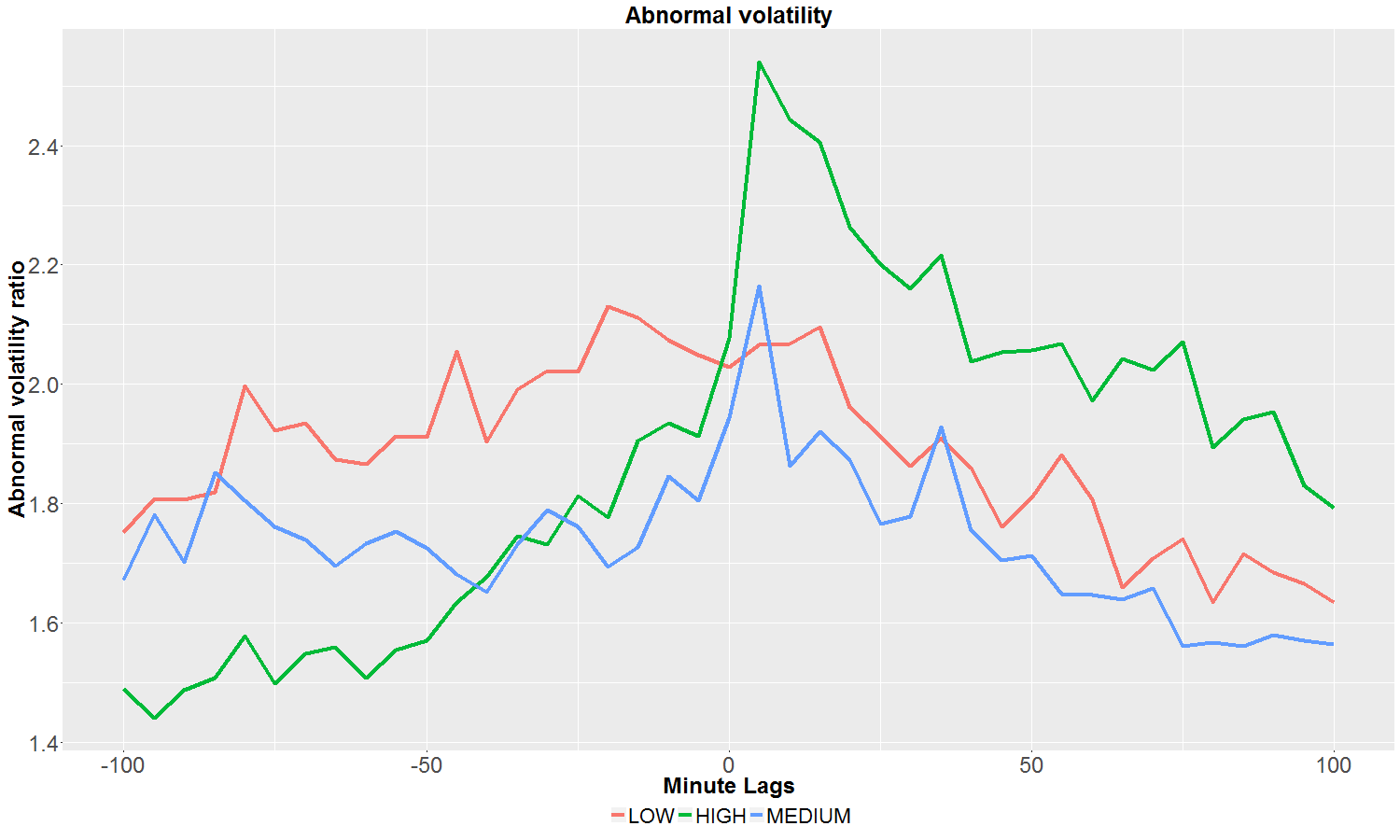
We here clearly see that the more novel a news is, the more the market will move in term of abnormal return and volatility surge. It is notable to see here that even for news without novelty a slight market reaction occurs. This is due to the fact that repeated news tend to still deeper instill their message into stocks prices at a time of globalization where investors do not live in the same time zone.

# 4.3.2 Ravenpack event relevance

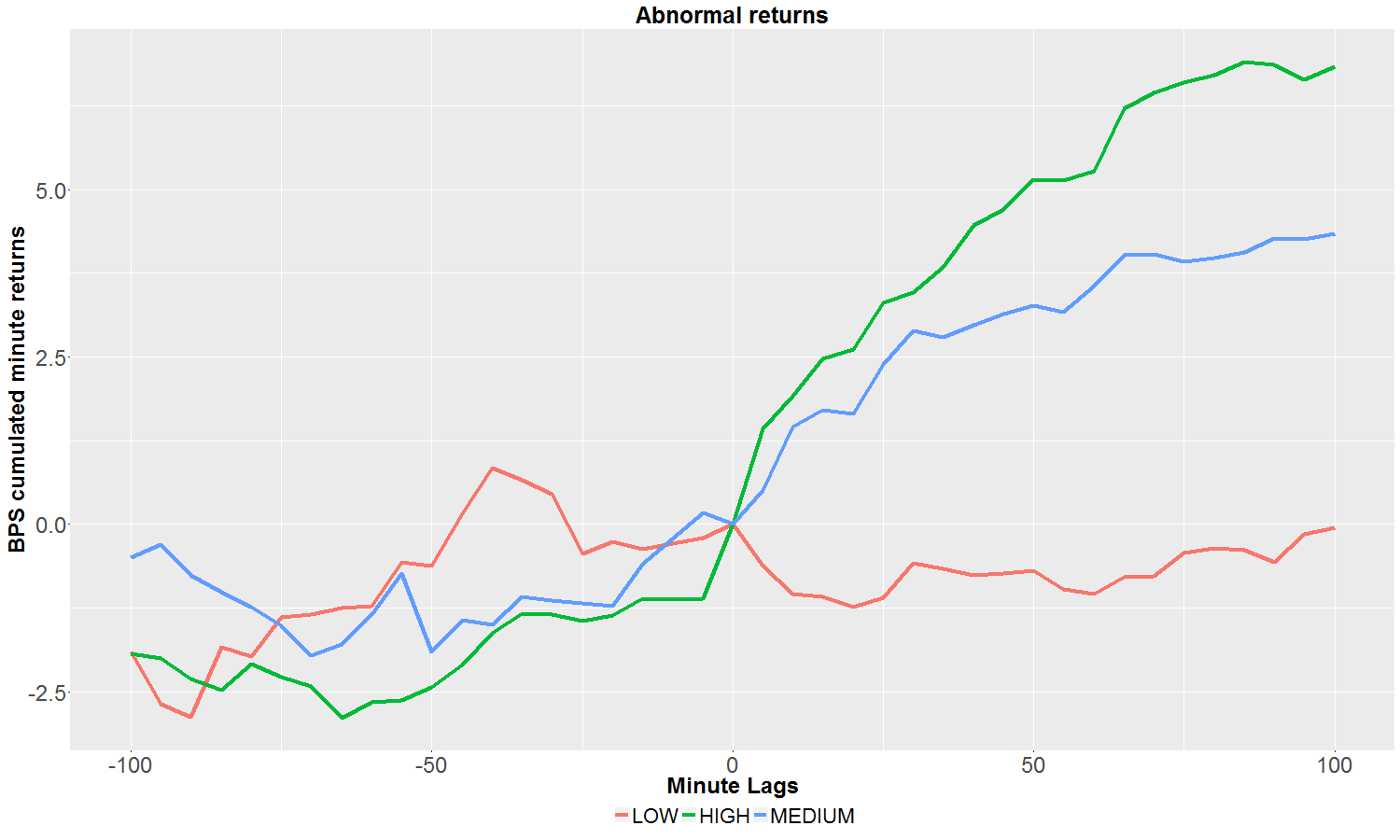
Ravenpack developed a news score that indicates how relevant an event is within a story. The scores fluctuate 0-100. We bucket those values as high (>=90), middle (>=70 and <90) and low (<70). Our high confidence means that the event for the specific entity has been matched in the headline. The following results expose results for Dow Jones and Premium pack type showcasing how the event relevance filtering affects the profile.

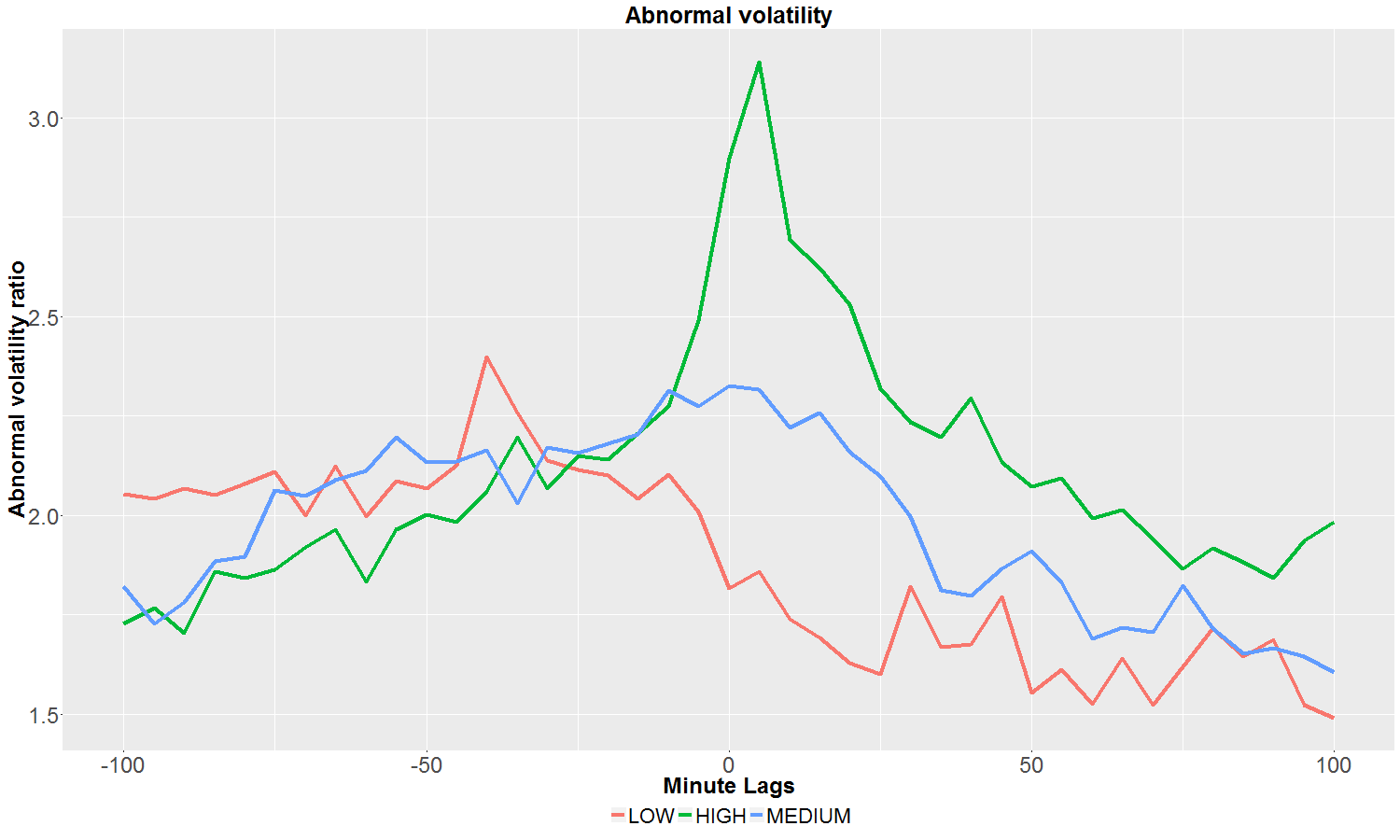
**Dow Jones wire news feed**:

**FIGURE 17: Event relevance effect on abnormal return for Dow Jones average profile.** This average profile is computed filtering for similarity gap greater than one day

****

**FIGURE 18: Event relevance effect on abnormal volatility for Dow Jones average profile.** This average profile is computed filtering for similarity gap greater than one day

**PREMIUM PACK:**

**FIGURE 19: Event relevance effect on abnormal return for PREMIUM PACK average profile.** This average profile is computed filtering for similarity gap greater than one day

**FIGURE 20: Event relevance effect on abnormal volatility for PREMIUM PACK average profile.** This average profile is computed filtering for similarity gap greater than one day

It is clear that both novelty measures and event relevance matter in our profile selection.

The lower the event relevance, the less steep is the abnormal return jump and volatility surge are.

Nevertheless it is important to notice here that those news still matter for the specific company in the sense that the abnormal return is not a white noise but clearly trends in the direction of the news sentiment. So those profiles are more momentum driven profiles, from which you can still extract value from intraday trading. It emphasizes the importance of Ravenpack Big Data analytics where all observation possess value. Even low relevance/less novel entries still possess value for intraday trading. The market profiles account then more as a quantification of the trend following momentum still providing great value for intraday trading.

# 4.4 Ranking groups and category

The last step of our investigation is to dig always deeper in Ravenpack taxonomy and investigate now all categories in a comparative table.

Now that we have established how Ravenpack big data analytics filtering mattered, we will present with a comprehensive ranking for all groups and entities according to the FREQUENCY metric and the VOLATILITY one. We restrict ourselves to those two metrics and the first 30, but more comprehensive results will be displayed in the appendix.

Note here the presence of our PREMIUM pack sources which provide their value and expertise for analyst ratings and merger and acquisitions.

Notice also the presence of our web content for more soft economic categories which demonstrate the value of the web content for those topics.

# W:\OutputData\sduprey\NAR-326\PAPER_PICTURES\TABLES\ALLcard_post_ranked_returnclean_bigdata_best_groups_html.png4.4.1 Group ranking: FREQUENCY RETURN

# W:\OutputData\sduprey\NAR-326\PAPER_PICTURES\TABLES\ALLcard_post_ranked_returnclean_bigdata_best_categories_html.png4.4.2 Category ranking: FREQUENCY RETURN

# W:\OutputData\sduprey\NAR-326\PAPER_PICTURES\TABLES\ALLvolatility_correctionbigdata_best_groups_html.png4.4.5 Group ranking: VOLATILITY

# W:\OutputData\sduprey\NAR-326\PAPER_PICTURES\TABLES\ALLvolatility_correctionclean_bigdata_best_categories_html.png ranking:VOLATILITY

# W:\OutputData\sduprey\NAR-326\PAPER_PICTURES\ALL\card_post_ranked_return\PREMIUM_PACKanalyst-ratingscard_post_ranked_return1best_return.png4.4 Displaying the first three best groups and their 2 best categories for each

# 5. Conclusion

We have here shown that Ravenpack Big Data analytics is the most compelling tool to build a market reaction database.

We have shown how its analytics helps you filter out noise and get to the stronger intraday signal.

This study is a big data study as our financial spans the R1000 universe for the past ten years at the minute level with the requirement of 3 months historic data to compute our abnormal return and volatility.

We have done that in a very general way for all categories generally and shown a comprehensive ranking.

The next step would be to focus on single specific categories and use Ravenpack most advanced analytics to improve them: Ravenpack big data release provides us with new dimensions to do exactly that: Fact level tagging, event temporal assessment and category specific analytics (like our earnings type) will even more refine our profiles to give you a statistical edge.

Then when market reaction is calibrated to the finest level, the next will be to build an event trading strategy based on those response calibration. This will be the next step of our research.

# 6. References

[1] Axel Gross-Klussmann and Nikolaus Hautsch (2011), “When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions”, Journal of Empirical Finance

[2] Quantamentals (September 2012), “Quantifying events”

[3] Peter Hafez and Jose A. Guerrero-Colón (March 2016), “Earnings Sentiment Consistently Outperforms Consensus”, RavenPack Quantitative Research

[4] Eugene F.Fama, Lawrence Fisher, Michael C. Jensen, Richard Roll, “The Adjustment of Stock Prices to New Information”, International Economic Review, Vol 10

[5] Nikolaus Hautsch (2008), “Capturing common components in high frequency time series: a multivariate stochastic multiplicative error model”, J. Econ.Dyn. Control 32

[6] Event study metrics methodology book, http://eventstudymetrics.com/

**Appendix A: Generalized Cumulative Corrado Test**

**Corrado rank test:**

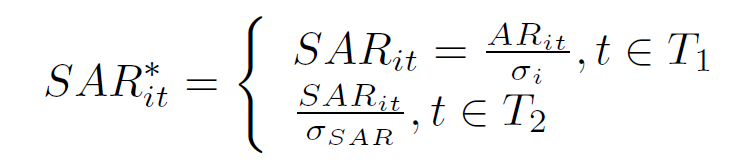
In our search for profitable trading profiles, we will use Ravenpack analytics to its full extent to filter per source, news novelty, entity relevance, and so on, but as we filter our events, the observations count lessens and to properly discriminate among those distinct cumulative trading profiles, one must devise a proper statistical significance measure to ponder the potential average compounded return possibly tradable.

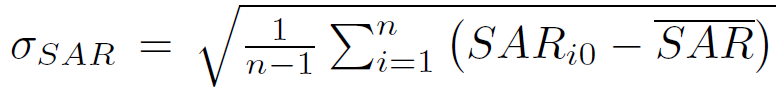
Many of those statistical tests have been extensively used and detailed in academic literature. We have tested most of them and retained only the most robust one to rank our profiles: the generalized Corrado rank test.

The Corrado rank test statistic is based on the rank of the abnormal returns standardized by their standard deviation over both the estimation window and event window combined T=T\_1+T\_2.

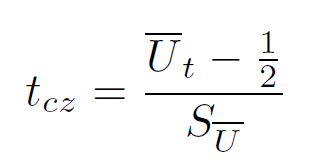
**Generalized Corrado rank test:**

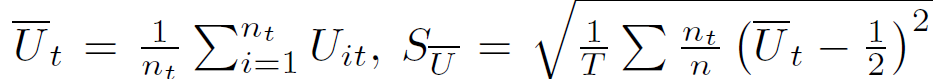
In order to account for the possible event induced volatility (following Boehmer, Mucumeci and Poulsen (1991)), the generalized Corrado test restandardize the SARs with the cross sectional standard deviation of the SAR defining the generalized standardized abnormal return.



Where

Is the standard deviation of the event minute standardized abnormal return (the over line meaning average cross sectional standardized abnormal return on the event minute).

The Corrado and Zivney (1992) test statistic (CZ) is then:

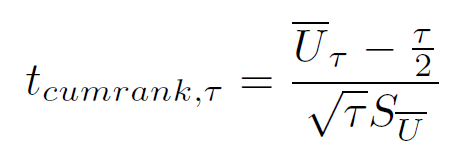
Where

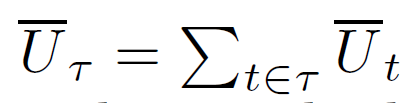
Are the cross-sectional average over the nt number of valid returns on bucket t.

It is notable (and computationally intensive!) that the generalized CZ test uses both the estimation period and the event period observations.

**Cumulative generalized Corrado rank test centered on the event:**

In order to test mean return on event windows longer than one bucket, which is our case with our 180 minutes padding, (following Cowan and Campbell and Wasley (2003)) suggest simply to aggregate the ranks over the window (CUM RANK) :



Where

Is just the cumulated sum of each minute bucket Corrado rank.

We use this methodology centered on the event: the cumulative rank generalized Corrado test in the event widow is computed cumulatively around the event based on the distance to the precise minute bucket when the event happened.

## 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 [research@ravenpack.com](mailto: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. 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.