NORTHFIELD modeling financial markets worldwide

December 2017

Risk Systems That Read®

Northfield Information Services Research

Dan diBartolomeo

After a series of research projects going back to 1997, Northfield is about to commercially introduce risk models augmented with quantified news flows for investors. News conditioned versions of our "near horizon" models will be available to Northfield clients starting in December 2017. *We believe this will be the biggest step forward in risk modeling for asset management since the creation of the multi-factor risk model in the 1970s.*

News and Conditional Risk Models

Almost all available risk models are "unconditional." They are based on a sample of past history that is deemed relevant, possibly giving more weight to recent observations, or assuming a simple trend in volatility (e.g. GARCH). We recently surveyed industry practice and found models based on sample periods ranging from as short as sixty trading days to more than twenty years. Once the sample period is determined, the *heroic assumption is made that the future will be like the past.*

This process omits everything we know about the present, and how the present is different from the past average conditions of the sample period. Using the information about the present to adjust the risk estimates has been standard in some Northfield models since 1997 and in all models since 2009.

For our purposes, "News" is the set of information coming to investors that tell us how the present is different from the past. This definition implies that routine information affirming the "status quo" is not news irrespective of how it is delivered. We also recognize the extensive literature showing that investors respond differently to "announcements" (time of information release anticipated) than to "news" where both the content and timing are a surprise. Finally, we must be selective about what we define as news, as only a minority of large asset price moves are a direct response to investors responding to news. There are a lot of "information-less" trades (see Livnat, et. al. 2013).

What is obvious from a visit to any investment firm is that investors go to great effort to receive and analyze financial news as it comes to public awareness. Based on their interpretation of the content of the news, investors revise their views on asset prices, and transact assets accordingly. *It should be very intuitive that risk assessments should also respond to news.*

Risk Systems That Read®

News conditioned risk assessments will be immediately useful in a number of applications. Of greatest interest to active managers are the areas of alpha estimation, portfolio risk and optimization for high turnover portfolios, especially hedge funds. Another area of importance will be algorithmic trading including "high frequency" and optimal execution algorithms, especially those like the Northfield trade scheduling algorithm which are framed as multi-period optimization problems.

We also believe this new approach also holds a lot of promise for daily updating of fixed income credit risk and derivative counterparty risk. A new "white paper" describing our methods will be published this summer and a presentation on use of this approach for credit risk was done June 16, 2016 at Boston QWAFAFEW.

We also believe that asset owners and fund of funds will find the enhanced models useful for monitoring risk of their hedge fund positions. Many fund managers and asset owners also need short horizon risk assessments for regulatory compliance purposes (e.g. UCITS). While there is more work to be done, our research to date suggests that a single large news event may change forecast risk levels by 10 to 15% at time horizons as far out as one year.

Previous Literature and Research

In 1997, Northfield pioneered the use of security level information on daily changes in implied option volatility. The methodology was subsequently published in diBartolomeo and Warrick (2002, 2005), A draft can be found at http://www.northinfo.com/Documents/534.pdf.

This process uses percentage changes in option implied volatility to condition (adjust) volatility estimates of individual stocks. Our approach uses a regression method to separate adjustments to security level volatility into adjustments factor volatility and security specific risks, which allows for partial adjustment of risk for non-optional stocks. For example, if there is a big jump in the implied volatility of one stock it is assumed to security specific but if most stocks in an industry see a jump in implied volatility most of the change would be applied to the related industry factor.

An example of the benefit of such a process was that very intuitive results were obtained when markets reopened after September 11th, 2001. Based on September 17 opening option data, we estimated that a portfolio of airline stocks had a 55% increase in volatility, while a 35% increase was forecast for a portfolio of property/casualty insurers. Portfolios consisting of stocks of consumer staples producers such as food and tobacco companies showed no change.

The primary limitation of the option approach was that only a small fraction of individual securities around the world have liquidly traded options. As such, the approach was only partially effective (factor portion of risk) for most securities. To address this problem, we proposed to use quantified text news in diBartolomeo, G. Mitra, L. Mitra (2009). This paper followed the analytical structure of diBartolomeo and Warrick (2005) but replaced option implied volatility with measures of quantified news flow and sentiment. Empirical tests were conducted on high liquidity stocks (Dow Jones 30 and EuroStoxx 50 names) for the hypothesis that news flow metric could effectively predict next day intra-day volatility. Our findings were that *news driven metrics were more efficient predictors of changes*

in volatility than metrics based on changes in option implied volatility and that the predictive contribution from news flows was statistically significant even when using option data and news together. We believe that news metrics work better than implied volatility because option markets have trading costs so the changes in implied volatility are muted compared to changes in the beliefs of investors.

Beyond the published research, we had two separate teams of MIT graduate students conduct their own research projects over two years. The key findings were that unsurprisingly, changes in one day security volatility was highly statistically significantly associated with changes in prior news flows at the individual stock level. More than a dozen functional forms of the relationship were tested on a large sample of hundreds of stocks over several hundred trading days. Very intuitively, the predictive power of news flow changes decayed rapidly for the most liquid stocks and more slowly for less liquid names. The impact of news also decayed more quickly for firms with more public recognition (e.g. Apple or Google).

Another important piece of research is that of Kyle, Obizhaeva, Sinha and Tuzun (2012). This works shows that a theoretically predicted relationship between the frequency of news articles on companies, and the volatility of their stocks was fit almost perfectly by the empirical data over hundreds of companies and many years. They use a particularly clever construct suggesting that a function of stock volatility and trading volume across stocks is constant when the rate of time passage is defined in "numbers of articles" which they call "business time." The prescribed functional form is a power function that also includes an "expected changes in a trading volume" component. Related papers from Kyle show how this structure can be used to predict bid/asked spreads and more generally trading costs

Northfield internal research has continued from 2013 to the present. All of the foregoing research suggests a *multiplicative* relationship between news flow and security volatility. During this period we also tested an *additive* functional form (H1: tomorrow's volatility goes up when a threshold value of news flow is hit today). This is an easy structure to fit in an existing factor model as a dummy variable. The factor exposure is 1 if there was "enough" news today and zero otherwise. To test this hypothesis, we estimated a GARCH style model (see Heston 1993 for details) using historic sample periods as short as twenty-two trading days (roughly one month). The main data for this work had 1.7 million data points (stocks * days). We tested feeds from multiple providers of news flow analytics. It was very comforting to observe that data from all providers worked to a highly statistically significant degree. T stats ranged from 7 to 9, which is "off the scale" in terms of statistical significance.

The additive construct is somewhat close to the idea of using news flow (or some metric thereof such as "sentiment") as one of the factors in the risk model. We reject this formulation for a number of reasons. The first and most important is that is misses the profound basis of the entire process. We agree with Kyle, et. al. that one can alternatively describe changes in asset volatility over time as time itself speeding up and slowing down. When there is a lot of news time is passing quickly, so volatility seems high when measured in clock time. When there is very little information coming to investors, time is passing slowly when measured in clock time. This concept was previously explored in Haug (2004).

The second reason for the rejection is that if we define some metric of news as a factor in the model, we must be able to estimate periodic returns to that factor. For example, we might believe that positive returns would arise from the arrival of "good news" and that negative returns would arise from the "bad news" or (perhaps perversely) the opposite. Nevertheless, for these effects to be factors rather than security specific effects, the preponderance of securities must react in the same way at each moment in time. We can find no statistically significant evidence of such effects over relevant time horizons in the literature.

The third reason is that by using news as a factor in the risk model we eliminate the possibility of conditioning the other factors on the model. Let's consider a hypothetical case of violent civil unrest in Nigeria which endangers production from local oil refineries. Such events would result in stories on the oil companies potentially impacted in the financial press. While capturing the company specific effects of such news would be a more useful improvement in portfolio risk it would depend on conditioning any oil related factors in the risk model. The process should properly reflect this development on energy markets as a whole, including all firms that either produce or consume large amounts of oil related energy.

The Northfield internal research effort in 2015 was focused on independent replication of the 2013-2014 research to confirm our findings on different vendors and to perform additional tests to compare effectiveness of multiplicative versus additive functional forms. We are now in the process of finalizing analytical details and production processes of "risk systems that read" for July introduction of daily updates of most Northfield models and delivery platforms.

Ingredients to the Conditioning Process

We are using Dow-Jones news feeds as summarized by Alexandria Investment Research. This combination gives us both very wide coverage and amazingly low latency times. Quantitative summarization of each article takes no more than 30 milliseconds and each summary is delivered to Northfield in real time. The Alexandria system "reads" both English and Japanese. In a recent 90 day period, Northfield received more than 210,000 article summaries associated with more than 21,000 different companies (as measured by ISIN IDs). Alexandria also provides similarly structured data on thousands of "non-company" topics such as countries, industries and commodities.

Each article summary provides numerous metrics including the "sentiment of the article" (good news/neutral/bad news). Brown, Harlow, Tinic (1988) provides a framework where the economic impact of bad news is stronger than that of good news. Another important indicator is "relevance." For example, an article about Verizon may peripherally mention Apple because Verizon sells iPhones. In this case, the relevance metric would be high for Verizon but lower for Apple. Other important factors include "novelty" (has similar news been previously reported recently?) and "event type" (news about a merger is probably more important than the announcement of a routine dividend payment). Northfield then compiles "today to date" aggregations of the article count on each subject.

Put most simply, Northfield is comparing the "*importance weighted* amount of news flow" over the last few days to a long term average base level of news flow for each company mentioned in an article. The Alexandria metrics are used to weight the importance of individual articles in the overall count. The functional form is multiplicative and

exponential so our conditional forward estimate of risk is our unconditional forward estimate of risk times some scalar (with default value one) which is raised to a power (less than one). The scalar is derived from the average amount of news flow in recent days in comparison to a long-term average of news flows on the same company. Scalars can be above or below one.

We also aggregate the news flow in various combinations of sectors, countries and geographic regions. This aggregate data can then be applied to smaller firms that by themselves have sparse, if any, news flow. For example, a high degree of news about large automobile manufacturers will be indicative of events that will also impact smaller firms that are parts to car makers.

The computation also captures multiple aspects of time decay. How long ago did the news take place? For example, there may have been a spike in news volume three days ago which will still be important, but less important than if the spike in news volume occurred today. How fast will investors notice the events? For high volume, US liquid names, the impact of news events will decay a lot faster for an obscure firm with no analyst coverage. Separately we consider the impact of time decay based on the forward risk horizon. For example, if we are trying to forecast intra-day volatility for tomorrow, an increase in news volume will have more impact than if we are trying to forecast average daily volatility over the next ten trading days.

Separation of Factor and Specific Risk

diBartolomeo and Warrick (2005) shows how adjustments to the security volatilities can be "fed back" into the model to adjust factor variances and volatility estimates for stocks on which no options are traded in Equations 7 through 9. The same process was used for news flow in diBartolomeo, Mitra, Mitra (2009). If the risk model factors are close to orthogonal, the distinction into factor and specific risks can be reliably estimated with a simple regression. If the model factors are not orthogonal you can use a non-linear optimization process (used in Northfield models since 2009) or generate an orthogonal transform of the factors, estimate using a regression, and then translate the factors back to the original basis. The more complex non-linear optimization is used at Northfield because it allows for the imposition of important boundary conditions such as maintaining specific risk at the security level to always be positive, factor volatilities to always be positive, and the implied security level covariance matrix (see diBartolomeo 1999) must have the mathematical property of being positive semi-definite. A recent paper detailing the various procedures is Shah (2015).

Conclusions

After a lot of research and eighteen years of experience in a similar framework, we are moving ahead commercially with daily conditioning of Northfield risk models based on our proprietary measures of news flow. There is little doubt that the strategies of most investors involve some form of response to financial news as it comes forward. *It is implausible that estimation of security and portfolio risk should somehow ignore this very obvious and elemental fact.*

References

diBartolomeo Dan and Sandy Warrick. 2005. Making Covariance-based Portfolio Risk Models Sensitive to the Rate at which Markets Reflect New Information. In: Knight J. and Satchell S. editors. Linear Factor Models in Finance. Oxford. Butterworth-Heinemann, p. 249-259.

diBartolomeo, Dan, Gautam Mitra and Lila Mitra. "Equity Portfolio Risk Estimation Using Market Information and Sentiment", *Quantitative Finance*, December 2009.

diBartolomeo, Dan. 2011. "Using News As a State Variable in Assessment of Financial Market Risk" In: G. Mitra and L. Mitra, Editors. The Handbook of News Analytics in Finance, Wiley, p. 247-254.

diBartolomeo, Dan. 2011. "News Analytics in Risk Management Framework for Asset Managers",. In: G. Mitra and L. Mitra, Editors. The Handbook of News Analytics in Finance, Wiley, p. 315-318.

diBartolomeo, Dan. "Credit Risk Assessment of Corporate Debt Using Sentiment and News", in G. Mitra and X. Yu Editors, "Handbook of Sentiment Analysis in Finance", Albury Books, 2016.