Incorporation of Quantified News into Portfolio Risk Assessment

Dan diBartolomeo
Northfield and Brunel University
London, 2009

Northfield
INFORMATION SERVICES, INC.

Motivation for the Short Term Risk Forecasts

- Risk models for asset management (as distinct from trading operations) have traditionally focused on estimating portfolio risk from security covariance over time horizons of a year or more
  - Suitable for long term investors such as pension funds
- Investment performance of asset managers is often evaluated over shorter horizons so they are interested in shorter term risk assessment. Hedge funds and other portfolios with high portfolio turnover are even stronger in this preference
- The proliferation of high frequency trading and algorithmic execution methods have created demand for very short horizon risk assessment

Northfield
INFORMATION SERVICES, INC.

A Short Chronology

- A call from Blair Hull in 1996
- diBartolomeo and Warrick (2005) in Linear Factor Models in Finance, edited by Satchell and Knight
- “Short Term Risk from Long Term Models”, Northfield research series, Anish Shah, 2007-2009
- “Equity Portfolio Risk Using Market Information and Sentiment” by diBartolomeo, Mitra and Mitra, 2009

Northfield
INFORMATION SERVICES, INC.

Simple Approach to Short Term Modeling

- The usual answer:
  - Increase the frequency of observations (daily or shorter)
  - Use a shorter sample period
  - Generally need different factors
- There are serious problems with this approach at the individual security level
  - High degree of kurtosis in return distributions (well maybe?)
  - Negative serial correlation due to short term reversal effects
  - Positive serial correlation on illiquid instruments
  - Asynchronous trading across time zones makes correlation estimation very difficult
- Address “shocks” through a GARCH process

Northfield
INFORMATION SERVICES, INC.

What's the Problem with High Frequency Data?

- Financial markets are driven by the arrival of information in the form of "news" (truly unanticipated) and the form of "announcements" that are anticipated with respect to time but not with respect to content.
- The time intervals it takes markets to absorb and adjust to new information ranges from minutes to days. Generally much smaller than a month, but up to and often larger than a day. That's why US markets were closed for a week at September 11th.
- GARCH models don't work well on announcements
  - Market participants anticipate announcements
  - Volume and volatility dry up as investors wait for outcomes
  - Reduce volatility into the announcement and boost it after the announcement, so they are wrong twice

Northfield
INFORMATION SERVICES, INC.

Some Surprising Things Appear Anticipated

- Lets look at a precipitous decline in the implied volatility of options on LUV
  - All days in 2001 prior to September 7, average of .45 with a s.d. of .13
  - September 7, LUV implied = .22
  - September 10, LUV implied = .15
  - All days subsequent to September 17, average of .54 with a s.d. of .18
  - September 10 is in bottom 1% of the universe in implied volatility, September 17 is 91st percentile
- Could this be driven by fundamentals?

Northfield
INFORMATION SERVICES, INC.

Investor Response to Information Flow

- Several papers have examined the relative market response to “news” and “announcements”
  - Ederington and Lee (1996)
  - Kwag Shrieves and Wansley(2000)
  - Abraham and Taylor (1993)
- Jones, Lamont and Lumsdaine (1998) show a remarkable result for the US bond market
  - Total returns for long bonds and Treasury bills are not different if announcement days are removed from the data set
- Brown, Harlow and Tinic (1988) provide a framework for asymmetrical response to “good” and “bad” news
  - Good news increases projected cash flows, bad news decreases
  - All new information is a “surprise”, decreasing investor confidence and increasing discount rates
  - Upward price movements are muted, while downward movements are accentuated

Northfield
INFORMATION SERVICES, INC.

Our Approach is Different

- Continue to use the existing risk models that are estimated from low frequency return observations
- Use new information that is not part of the risk model to adjust various components of the risk forecast to short-term conditions
  - Just ask yourself “How are conditions different now than they were on average during the sample period used for estimation?”
- This approach has multiple benefits
  - We sidestep almost all of the statistical complexities that arise with use of high frequency data
  - We get to keep the existing factor structure of the model so risk reporting remains familiar and intuitive
  - Since our long term and short term forecasts are based on the same factor structure, we can quickly estimate new forecasts for any length time horizon that falls between the two horizons
  - Can be applied to any of our existing models

Northfield
INFORMATION SERVICES, INC.

One Form of Working with "External Information"

- Risk estimates in our short term model of US equities have been conditioned for many years based on analysis of stock option implied volatility
  - Every day we look at the implied volatility of options on all US stocks. We keep a 30 day moving average of the ratio of implied volatility to historic volatility
  - If the implied volatility/historic ratio jumps because of an information flow to the market (e.g. Bill Gates gets run over by a bus), the specific risk of that stock is adjusted
  - If implied volatility ratio of many related stocks changes, the implied changes in factor variance are also made. Risk forecasts change even for stocks on which no options trade
  - Requires non-linear optimization process for adjustments
  - See Chapter 12, by diBartolomeo and Warrick *Linear Factor Models in Finance*, Satchell and Knight, editors (2005)

Northfield
INFORMATION SERVICES, INC.

“Variety” as External Information

- Solnik and Roulet (2000) examine the dispersion of country returns as a way of estimating correlations between markets
- Lilo, Mantegna, Bouchard and Potters use the term *Variety* to describe cross-sectional dispersion of stock returns
  - They also define the cross-sectional dispersion of CAPM alpha as *idiosyncratic variety* (noted as v(t))
  - They find that the average correlation between stocks is approximately:

$$C(t) = 1 / [1 + (v^2(t)/r_m^2(t)]$$

- diBartolomeo (2000) relates periods of high cross-sectional dispersion to positive serial correlation in stock returns (i.e. momentum strategies working)

Northfield
INFORMATION SERVICES, INC.

Other Conditioning Information

- Estimates of volatility based on high/low/open/close information instead of the dispersion of returns
  - Parkinson, Garman-Klass, Satchell-Wang, etc.
- Yield spreads for different classes of fixed income securities provide an implied default rate and the potential for large negative skew in stock returns
- Implied distribution of asset returns given the implied vols of options on market indices across strike prices
- Direct measures of information flow to investors, and investor attention that can create imbalances between supply and demand for a given stock

Northfield
INFORMATION SERVICES, INC.

What Makes People Buy or Sell a Particular Stock?

- They WANT to trade the stock
  - They believe the information that supports a valid forecast of abnormal future return
- They HAVE to trade the stock
  - They are trading to implement a change in asset allocation
  - They are trading to implement a cash versus futures arbitrage trade on a stock index
  - They are a mutual fund or ETF sponsor responding to investor cash flows in or out of the portfolio
  - They are hedge fund that is forced to transact because of a margin call
  - They are forced to cover a short position by having the stock called

Northfield
INFORMATION SERVICES, INC.

The Potential for “Have To’s”

- We can fundamentally evaluate the potential for “have to” trades
  - Index arbitrage trades only occur with index constituents and we know the open interest in futures
  - Short interest information is published
  - We know what big hedge and mutual funds have big positions in particular stocks
  - We have somewhat out of date information on full mutual fund holdings and cash flow statistics
  - We have fairly up to date information on ETF flows

Northfield
INFORMATION SERVICES, INC.

The Potential for “Want To” Trades

- Investors are responding to information, so just measure variations in the volume of information about a particular stock over time
- Judge the magnitude of information flow of news text coming over services such as Dow-Jones, Reuters and Bloomberg
  - Ravenpack and Thomson Reuters offer **real time statistical summaries** of the amount and content of text news distributed
  - Lexicons of over 2000 popular phrases are used to score the content as “good news” or “bad news”
- Judge investor attention directly by measuring the number of Google and Yahoo searches on ticker symbols

Northfield
INFORMATION SERVICES, INC.

Incorporating News Flows into Risk Assessments

- diBartolomeo, Mitra and Mitra (2009) forthcoming in Quantitative Finance
  - Follow the diBartolomeo and Warrick mathematical framework
  - Allow the conditioning information set to include both option implied volatility and variations in text news flows from Ravenpack (derived from Dow-Jones text feeds)
  - Empirical tests on Euro Stoxx 50 during January 17-23, 2008 and Dow Jones 30 stocks September 18 to 24, 2008
  - Evaluate both individual stocks, full index and financial/non-financial subset portfolios
- In all cases, inclusion of quantified news flows improved the rate of adjustment of risk estimates to time variation in volatility faster than implied volatility alone

Northfield
INFORMATION SERVICES, INC.

Incorporating Investor Attention

- Our next step will be to directly measure the degree of investor attention to a stock
  - Judge investor interest directly by measuring the number of Google and Yahoo searches on trading symbols
  - Avoid company names to eliminate product or service related searches
  - Try it yourself with Google Trends
- Da, Engleberg and Gao (2009) have already documented a strong relationship between abnormal search frequency and price momentum
- Investor attention is not always a good thing
  - Bolster and Trahan (2009) document predictable price behavior in stocks mentioned on the Jim Cramer television show
  - Clear strategy: wait two days, then short every stock mentioned positively or negatively

Northfield
INFORMATION SERVICES, INC.

Crucial Refinement

- diBartolomeo and Warrick (2005), and diBartolomeo, Mitra and Mitra (2009) both assume that the full impact of the conditioning information should applied to ex-ante risk estimates

- Shah (2008) introduces formal Bayesian framework for incorporating conditioning information into models
  - Requirement for orthogonal factors is removed
  - Non-linear optimization to “fit” the adjustments to correlated factors is even more complex
  - Introduced into Northfield “near-horizon” models in May 2009
  - Reduces noise and allows for fitting to different time horizons

Northfield
INFORMATION SERVICES, INC.

Other Differences Between Long and Short Horizon Risk

- Negative serial correlation
  - Daily overreactions & reversals, which cancel out over time, become significant e.g. under leverage
- Contagion / panic
  - Liquidity demands can drive up short-term correlations
- Transient behavior
  - A long term model intentionally integrates new phenomena slowly: Is the future like the past or are we in and concerned about a present shift?
- Lots more extreme events in the short-term
  - 3 std deviations contains less probability mass. 99% VaR is farther away from the mean

Northfield
INFORMATION SERVICES, INC.

Conclusions

- The key to good short term risk assessment is understanding how conditions now are different than they usually are
- A broad set of information other than stock characteristics and past returns are clearly useful in improving risk estimates
- Among the most useful sets of conditioning information appears to be summaries of textual news flows to investors
- A rigorous Bayesian framework should be employed to intelligently combine long term and short term information sets

Northfield
INFORMATION SERVICES, INC.