NORTHFIELD modeling financial markets worldwide

Experiments in Conditioning Risk Estimates with Quantified News

Christopher Kantos (chris@northinfo.com)

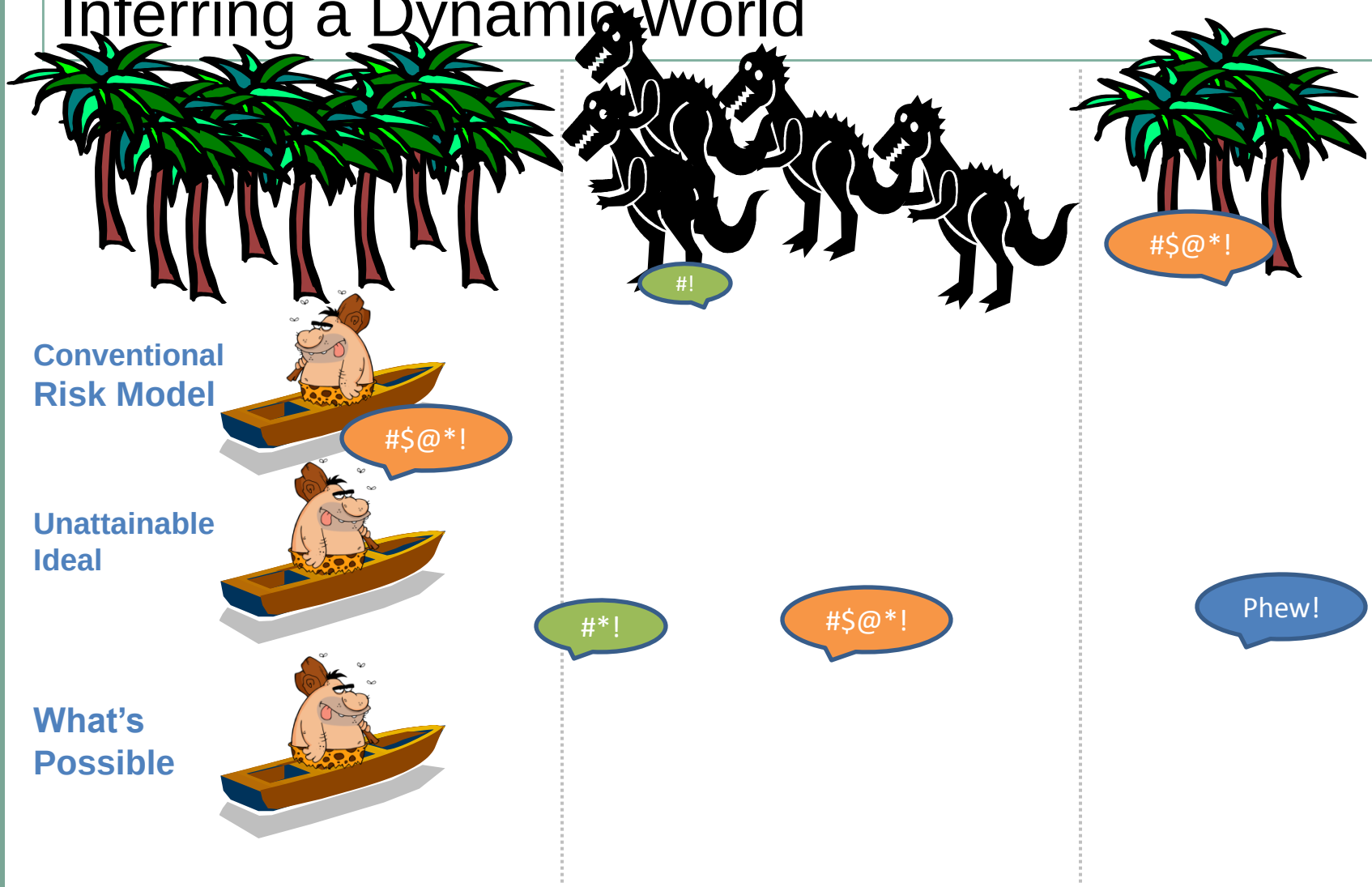
Based on research by Anish Shah, and Louis Scott

Pittsburgh, December 2014

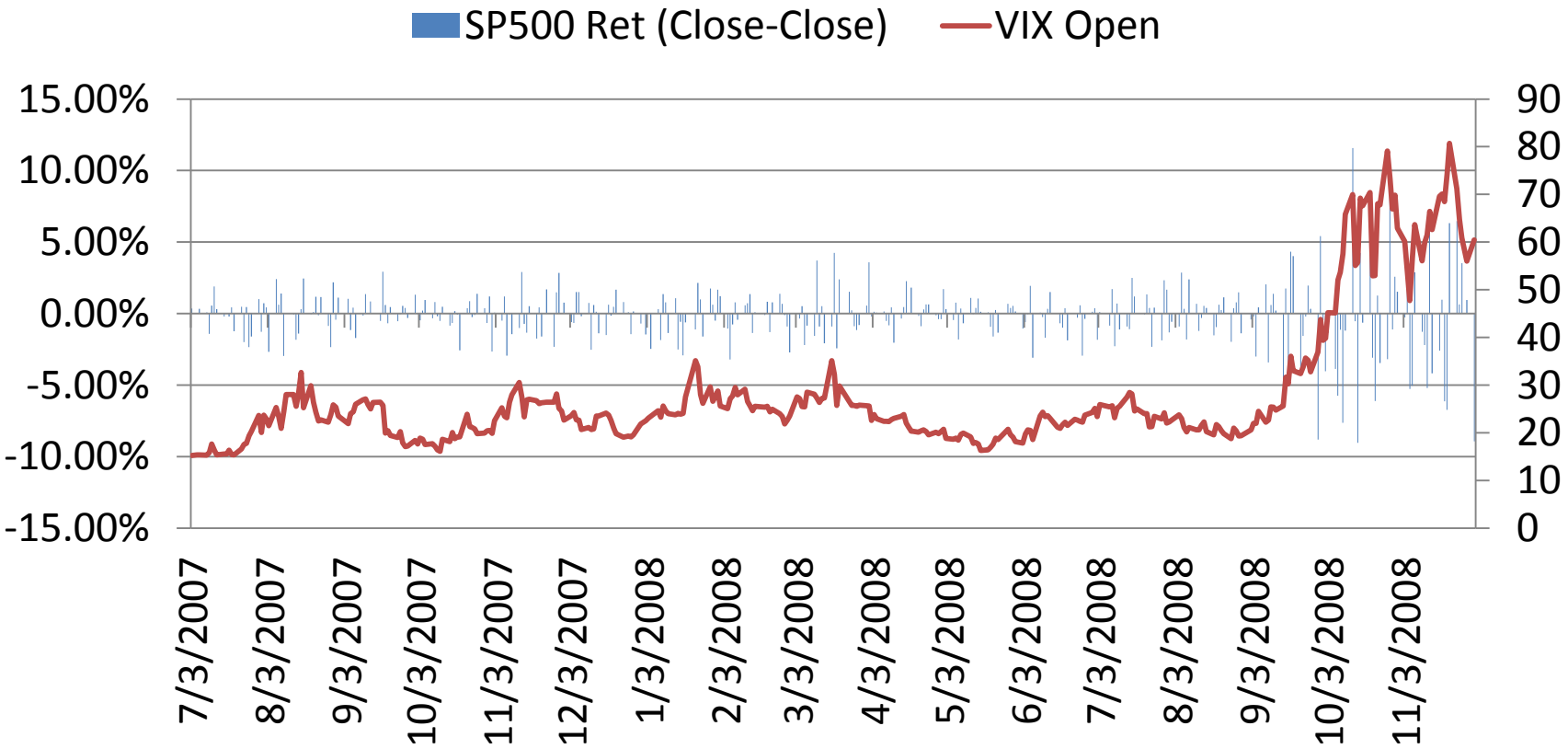
Why?

- **Markets are dynamic**
  - Most information arrives all of a sudden; effects persist
- **Conventional risk models come from a limited set of not so timely data**
  - Time series of (monthly/weekly/daily) returns
  - Financial statement data (note: updating daily adds mostly noise to ratios involving price)
- **The window used to build the model, e.g. years of monthly observations, makes changing directions slow**
  - Even under small observation intervals, e.g. daily, the window can't be sufficiently shrunk to be responsive. The model needs data points to learn structure, e.g. correlations among factors

Inferring a Dynamic World



VIX Rapidly Responds To Regime Changes (but can't see one before it happens)



What about ARCH/GARCH/… ?

- Work from time series of returns

  Future volatility forecast is adjusted considering recent error in return model

- Ignores a boatload of valuable information
  - option implied volatility
  - instantaneous volatility estimates: intra-period high/low, average ret^2 over small intervals
  - news
  - search and social network statistics
  - foreseeable patterns, e.g. higher volatility on announcement days

- Obvious: closing one's eyes doesn't make reality any tamer

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 after September 11th, 2001.

- GARCH models don't work well on announcements (e.g. earnings)
  - Market participants anticipate announcements
  - Volume and volatility dry up as investors wait for outcomes
  - Reduce volatility going into the announcement and boost it after the announcement, so they are wrong twice

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 current conditions
- 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 existing model

Our Definition of News

News is information that tells us how the world is currently different than it usually is.

This definition implies a basic context for the evaluation of any information flow

Investor Response to Information Flow

- Several early papers examined the relative market response to “news” and “announcements”. Subsequent papers came to similar conclusions

  Ederington and Lee (1996), Kwag,Shrieves and Wansley(2000),

  Abraham and Taylor (1993), Jones, Lamont and Lumsdaine (1998)

- 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
  - In general, no news is good news. The one exception is trading by corporate insiders.

Market “Surprises” That Were Anticipated?

Lets look at a precipitous decline in the implied volatility of options on LUV (Southwest Airlines) in 2001

All days in 2001 prior to September 7, the implied volatility (at the money options 45 days out) had an average of .45 with a standard deviation of .13

September 7, 2001 LUV implied = .22, September 10, LUV implied = .15

All days subsequent to September 17, average of .54 (s.d. =.18)

September 10 is in bottom 1% of the universe of US optionable stocks in implied volatility, September 17 is 91st percentile

Could this be driven by fundamentals?

Contemporaneous Information
Northfield US Equity Short-Term Model (1998)

- diBartolomeo & Warrick (2005). Making covariance based portfolio risk models sensitive to the rate at which markets reflect new information. Chapter 12 in Linear Factor Models. Knight, J. and Satchell, S., Elsevier Finance
- In daily production since 1998 (version 2 released in 2009)

- **Incorporate current, forward-looking market information** - option implied volatility
  - For stocks with liquid options, forecast vol = time series vol * **(implied/ historic implied)**

- **Glue these adjusted forecasts to a factor model**
  - Start with a statistical factor model built from the past 230 days of returns
  - Project adjusted forecasts onto the model's factor variances
    - Spreads pervasive information to those stocks without options
  - What remains is company specific (e.g. Bill Gates run over by a bus) and is applied to stock specific risk

Generalizing the Idea
Northfield Adaptive Near Horizon Models (2009)

- Uses Bayesian framework for incorporating market information
  - Short-Term Risk from Long-Term Models, Shah 2007
  - Adaptive Near Horizon Models (short-term versions of all Northfield models) in daily production since 2009

- **Take any risk model**, e.g. one of our models estimated monthly

- **Add flexibility points**, e.g make factor variances and stock-specific risks adjustable

- **Fit (now flexible) model to current conditions**
  - Volatility estimates from intraday high/low
  - Cross-sectional volatility
  - Implied volatility
  - Forecasts of volatility using other sources

- **Keeps familiar factor structure of long-term model**

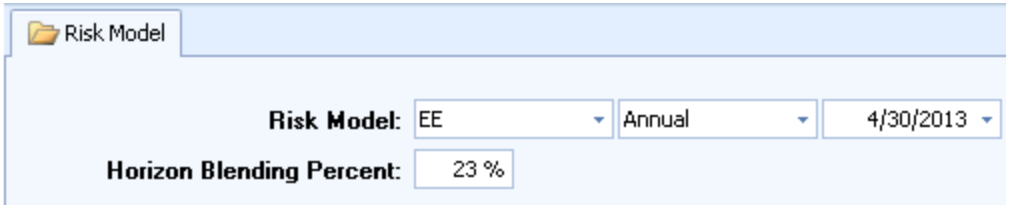
Example: 2nd Generation US Short-Term Model (2009)

- **Risk Model** = {risk factors, their variances & pairwise correlation, stock-specific variances}
- **Make the factor variances and specific risk free**
  - factor risk: $fv_k \leftarrow \mathbf{c_k}\ fv_k$
  - stock specific risk
    - $ss_i \leftarrow \mathbf{m_i}\ ss_i$ for stocks with implied vol information
    - $ss_j \leftarrow \mathbf{\mu}\ ss_j$ for stocks w/o
- **Have forecasts from (relative to historic levels of) implied vol**
  - Individual stock variances $\{sv_i\}$
  - Portfolio variances from VIX & RVK $\{pv_l\}$
- Over the free parameters (**c**, **m**, **μ**), **minimize the difference between risk model predictions and forecasts**, subject to constraints
  - (c, m μ) ≥ minimum value > 0 to keep variances positive and reasonable
  - can also restrict dispersion among adjustments, e.g. stdev(c) ≤ some maximum spread

Horizon Blending in Northfield Optimizer (2013)

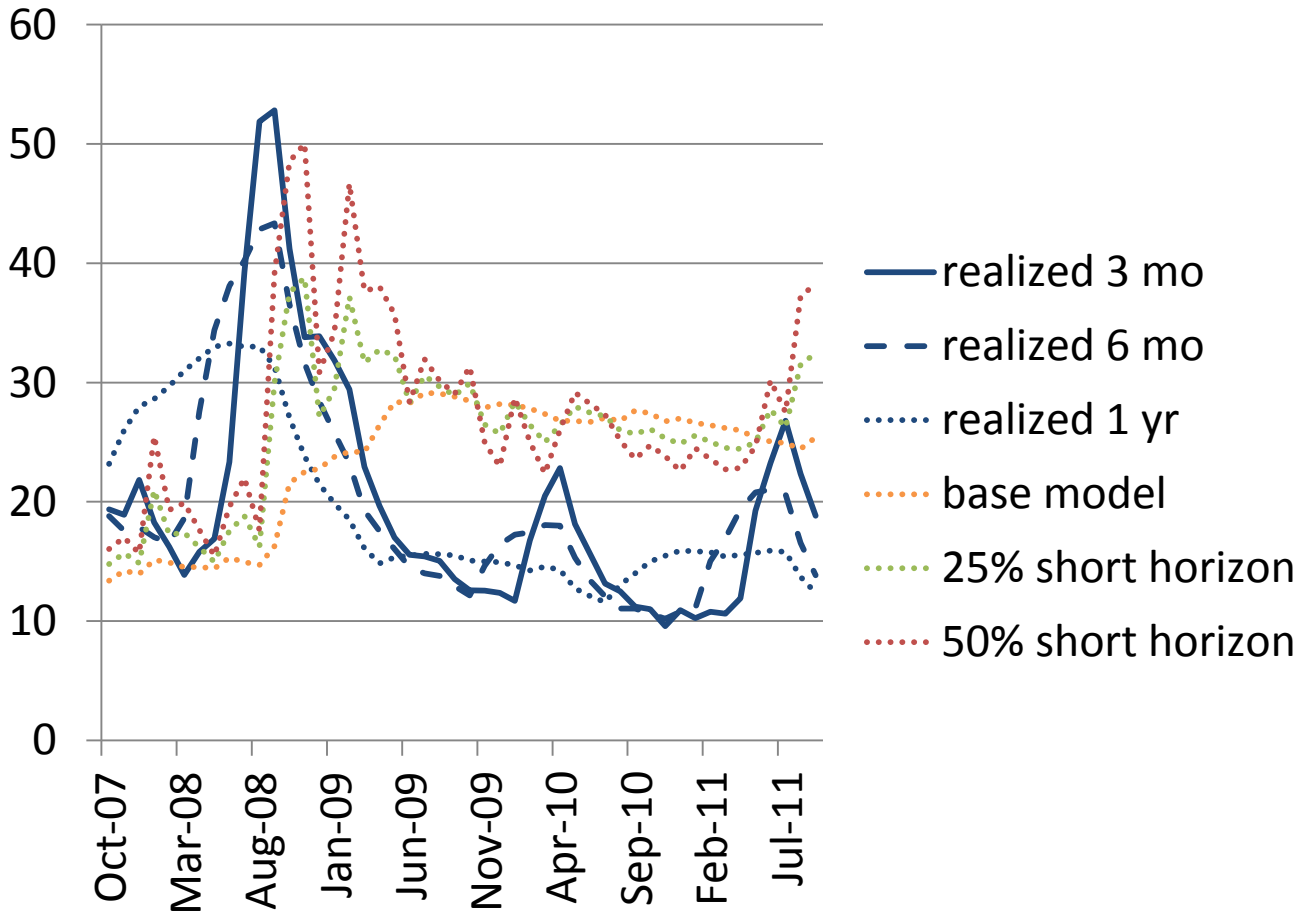
- Have long-term (traditional time series) and short-term (current information adapted) versions of all Northfield models

- Blend the two dynamically to get views of the near, intermediate, and long-term future:



- Higher blends of near horizon work better on shorter horizons
- The appropriate blend depends on the investment horizon, model, and market conditions
- Potential to create “forward” volatility estimates
  - What is the implied six month risk forecast starting six months from now?

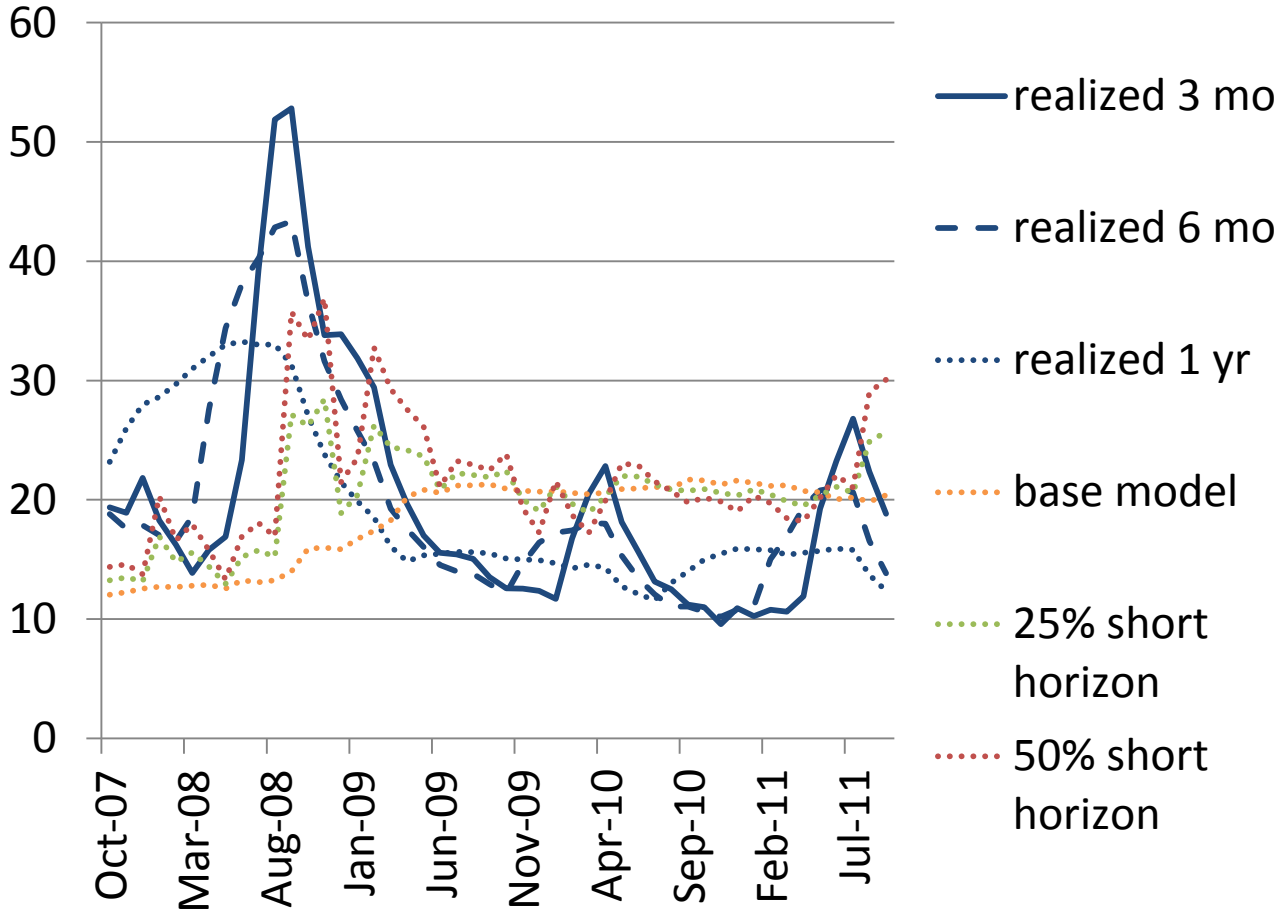
Horizon Blending in US Fundamental, 2nd Gen



Realized volatility is estimated from intraday highs/lows of the Russell 3000 over the following 3, 6, and 12 months

Model forecasts are of sqrt of cap weighted universe of ~ 3000 names

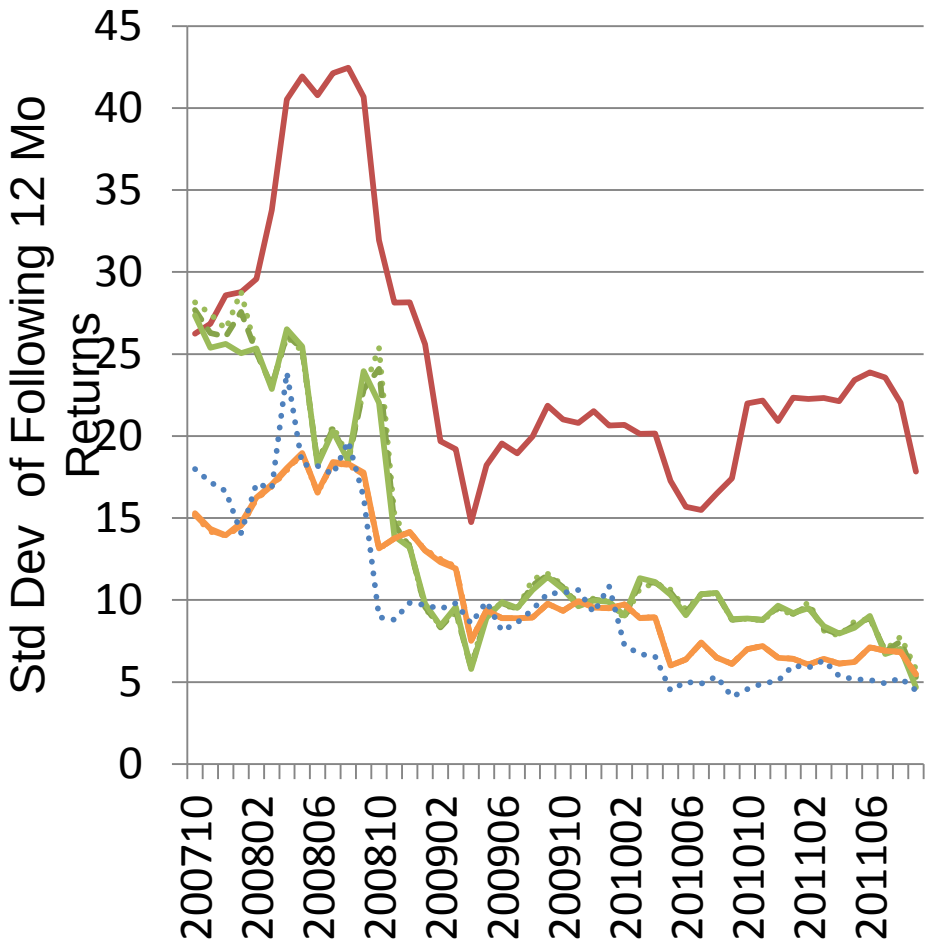
Horizon Blending in US Single Market, 2nd Gen



US Single Market is a time series model with factors:

- Market
- Sector(s)
- Bond index
- Oil
- Strength of USD
- Size spread
- Growth spread
- 5 statistical factors

Better Forecasts, Similar Optimization Results



Unoptimized is weighted by sqrt of cap

Fully invested (100% net long) portfolios optimized for minimum forecasted variance

Models are Northfield 2nd gen, not recently released 3rd gen

More Contemporaneous Information: News

- Mitra, L., Mitra, G., & Dibartolomeo, D. (2009). Equity portfolio risk estimation using market information and sentiment. Quantitative Finance, 9(8), 887-895.
- Same framework as Northfield US Short-Term Model
- A stock's contemporaneous information is a blend of
  - Current implied volatility / its recent average
  - Its intraday variation in news sentiment / its historic average
- Ravenpack data (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/nonfinancial 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, and can be applied to non-optionable assets

A Lot More to News

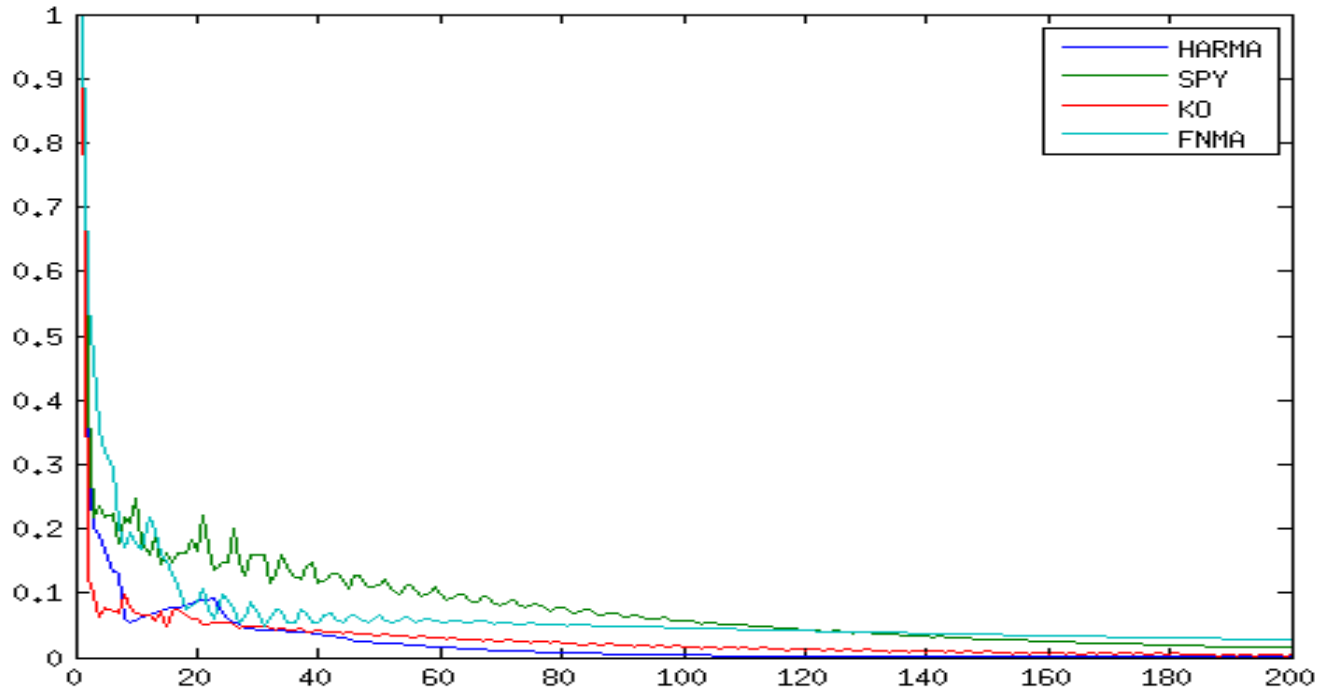
- Many vendors of quantified news –
  - Ravenpack, Alexandria, MarketPsych, Thomson Reuters, Recorded Future

- Data is extremely high dimensional

  - Each news item – subject, keywords, entities involved, sentiment, innovation, type of story (earnings, macro), …
  - Assorted entities – companies, stocks, industries, currencies, countries
  - Aggregated over time for an entity – number of stories, average sentiment, length and complexity of language

- In general, as predicted by Brown, Harlow, Tinic (1988) *no news is good news.* This is especially true for fixed income where there is little upside for bondholders when good news comes along.

Impact of Single News Events on Volatility Decays Quickly



Market Invariance and News

- Kyle, Obizhaeva, Sinha and Tuzun (2012) build on the "market invariance" concept developed in prior papers by Kyle and various co-authors
  - The invariance principle states that financial trading exists as a form of risk transfer. The rate of risk transfer among participants, called W, is the product of dollar volume and volatility
  - Makes precise predictions about relationships among numbers of trades, size of trades, bid-asked spreads and market impact costs
  - Variations in W across time and assets is explained by rescaling clock time into "business time" where time goes faster for some assets *so risk transfer per unit of business time is constant across assets*
  - A large empirical test using Reuters data confirms that "business time" speeds up and slows down proportionately to news flows

What Makes People Buy or Sell a Stock?

- They WANT to trade the stock because they believe they have 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

- *News analysis can helps distinguish whether trades are "Want To" (responses to specific information) or "Have To" and thereby understand price trends*

Distinguishing Information from Noise

- Govindaraj, Livnat, Savor and Zhou (2013)
- Find large percentage daily price changes in CRSP US universe (abs > 5%)
- Observe whether IBES analysts estimates for earnings or target price change in the subsequent five trading days
  - If there are changes in estimates, we assume the large price move was due to actual information in the market (a “Want To” trade by investors) and the trend will persist
  - If there are no changes in estimates, we assume the large price move was due to a “Have To” trade and will reverse later because there was no fundamental reason for the move
- They find about 80% of all large price moves DO NOT carry information
  - This result is robust to different magnitudes and definitions of price move
- Use of quantified news can provide immediate confirmation or negation of large price moves without having to wait five trading days
- News helps us to not over-react to large, but uninformative price moves that would otherwise pollute the risk estimates

Conclusions

- Risk is about the future not the past so models based solely on historical data must have limited value

- We can improve our forecasts of the near future by conditioning our beliefs on a range of information which is contemporaneously available such as option implied volatility and news flows

- "Quantified" news is a primary source of understanding how the state of the financial world is different than it usually is.
  - These differences are what drive investors to voluntarily transact, but not all transactions are voluntary
  - News is an excellent predictor of short term volatility
  - Early research shows that "news" driven approaches are valuable above and beyond the information embedded in implied volatility for assets where options trade