

Experiments in Conditioning Risk Estimates with Quantified News

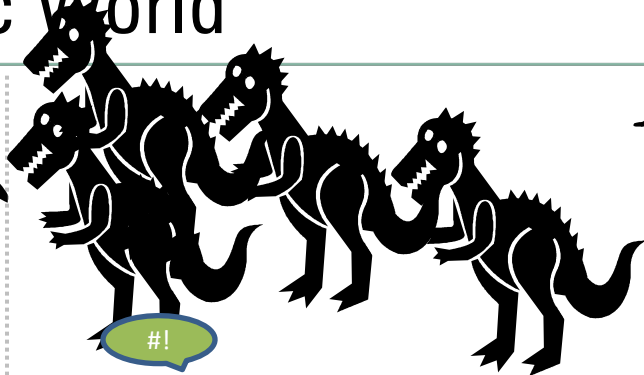
Anish R. Shah, CFA
Anish@northinfo.com
Newport Jun 7, 2013



Why?

- **Markets are dynamic**
 - 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



Conventional Risk Model



Unattainable Ideal



What's Possible



#!

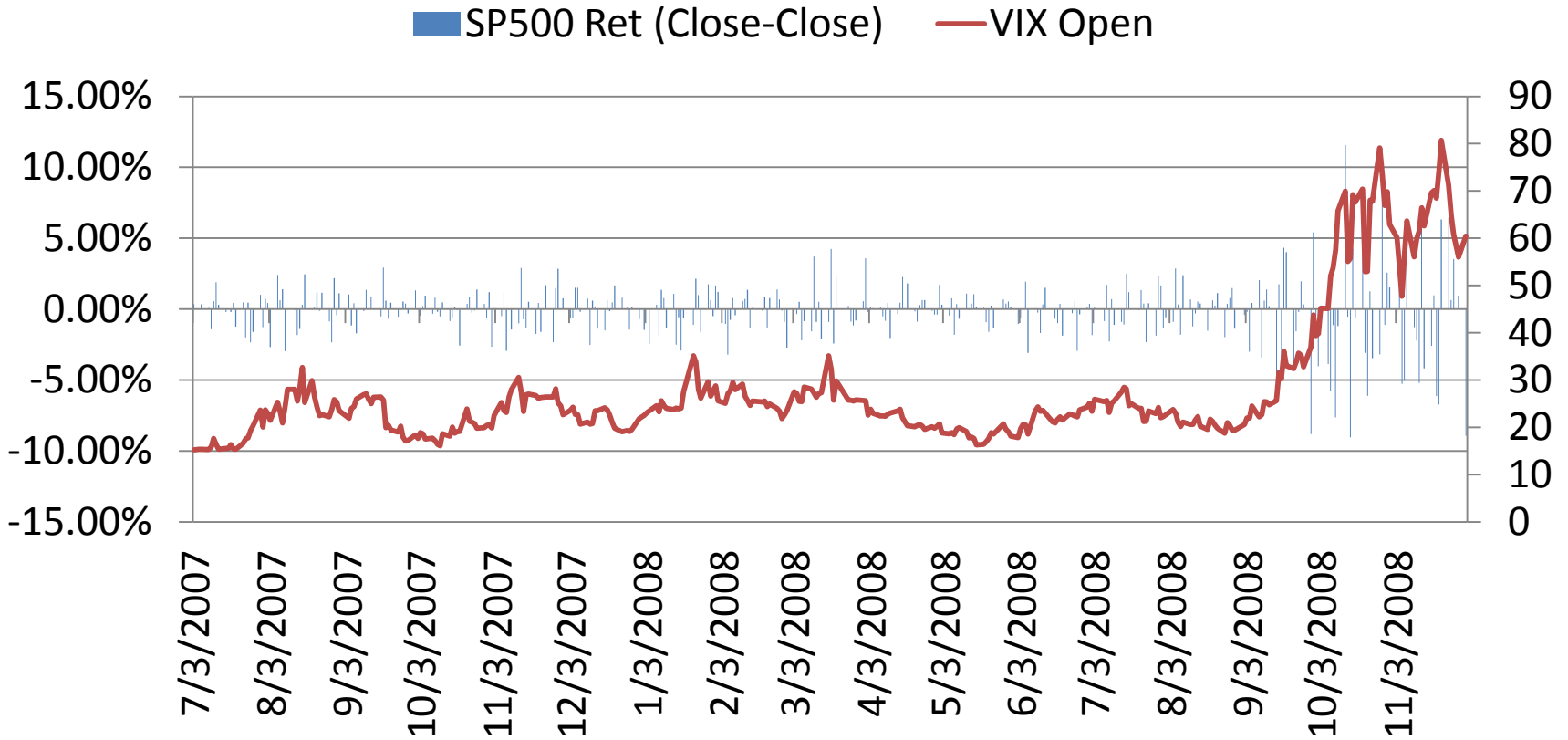
#\$@*!

#*!

#\$@*!

Phew!

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

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)

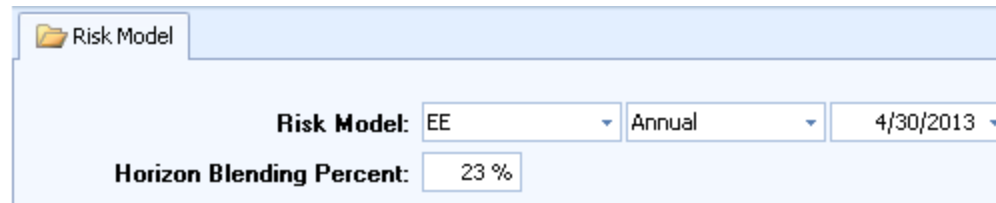
- Uses Bayesian framework for incorporating market information
 - [Short-Term Risk from Long-Term Models](#), Shah 2007
 - Adaptive Near Horizon Models (short-term versions of all Northfield models) in daily production since 2009
- **Take any risk model**, e.g. one of our models estimated monthly
- **Add 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: $fv_k \leftarrow c_k fv_k$
 - stock specific risk $ss_i \leftarrow m_i ss_i$ for stocks with implied vol information
 - $ss_j \leftarrow \mu ss_j$ for stocks w/o
- **Have forecasts from (relative to historic levels of) implied vol**
 - Individual stock variances $\{sv_i\}$
 - Portfolio variances from VIX & RVK $\{pv_i\}$
- Over the free parameters (c , m , μ), **minimize the difference between risk model predictions and forecasts**, subject to constraints
 - $(c, m, \mu) \geq$ minimum value > 0 to keep variances positive and reasonable
 - can also restrict dispersion among adjustments, e.g. $stdev(c) \leq$ some maximum spread

Horizon Blending in Northfield Optimizer (2013)

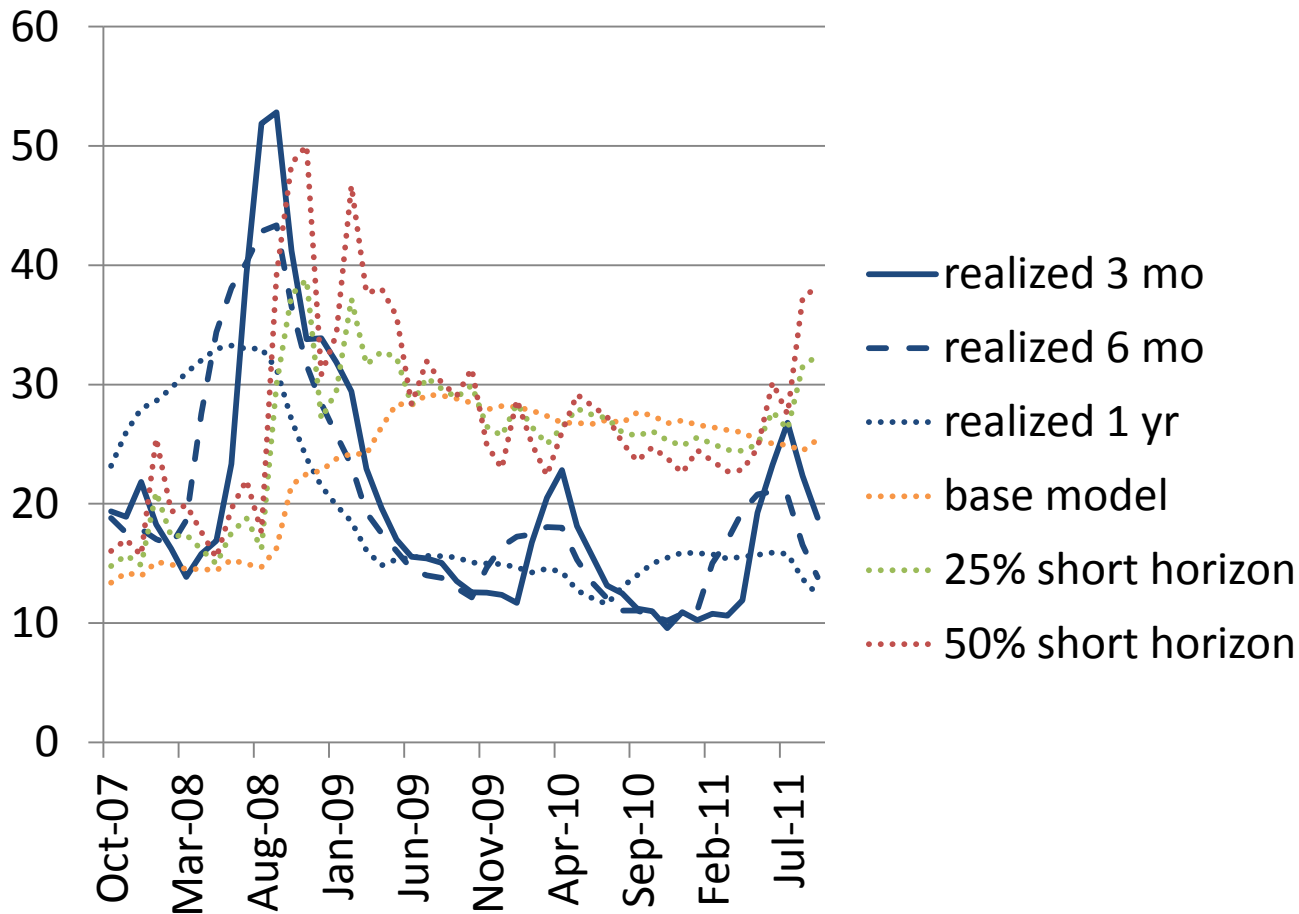
- Have long-term (traditional time series) and short-term (current information adapted) versions of all Northfield models
- Blend the two dynamically to get views of the near, intermediate, and long-term future:



The screenshot shows a software interface for the Northfield Optimizer. At the top, there is a tab labeled "Risk Model". Below the tab, there are three dropdown menus for "Risk Model": the first is set to "EE", the second to "Annual", and the third to "4/30/2013". Below these, there is a text input field for "Horizon Blending Percent" which contains the value "23 %".

- Higher blends of near horizon work better on shorter horizons
- The appropriate blend depends on the investment horizon, model, and market conditions

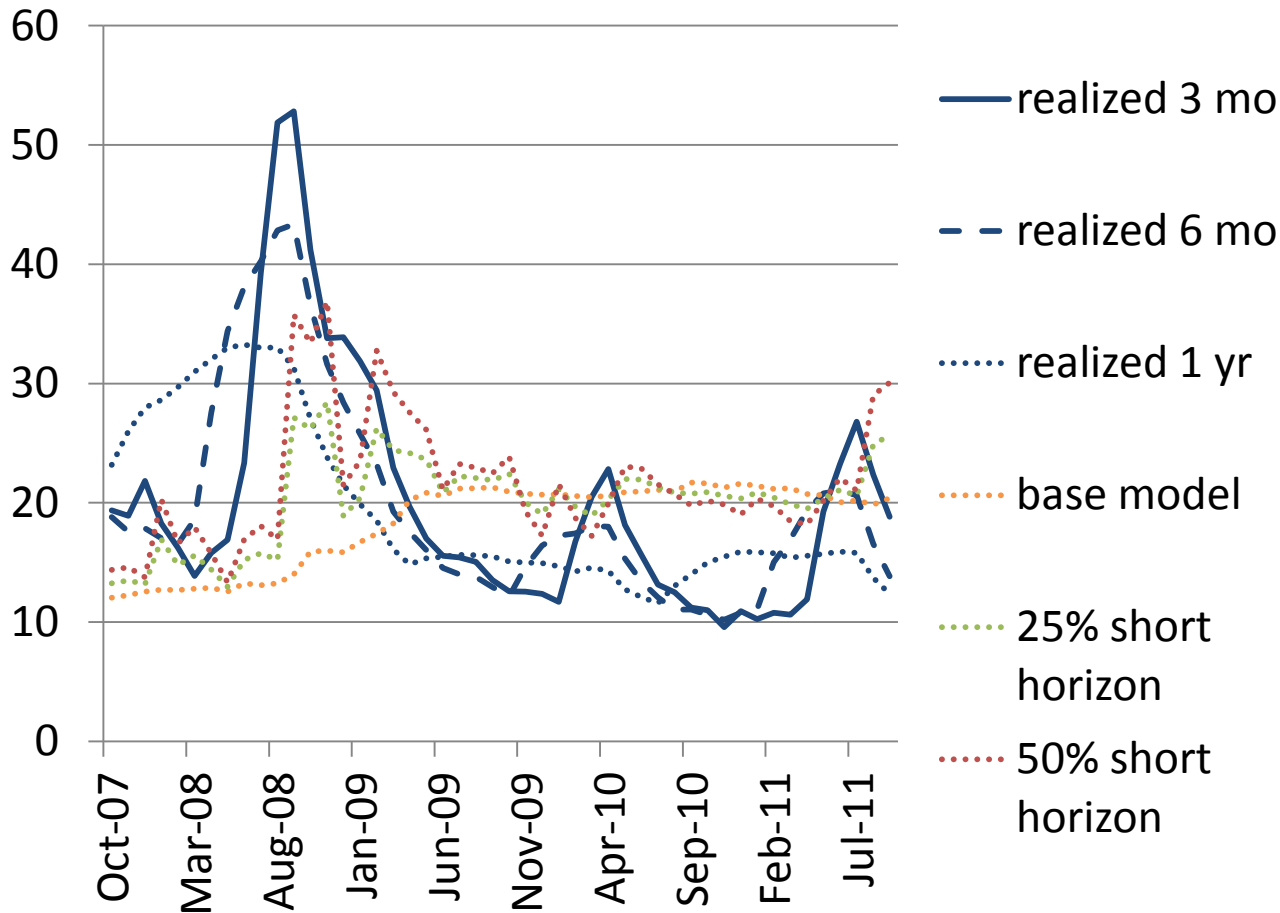
Horizon Blending in US Fundamental, 2nd Gen



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

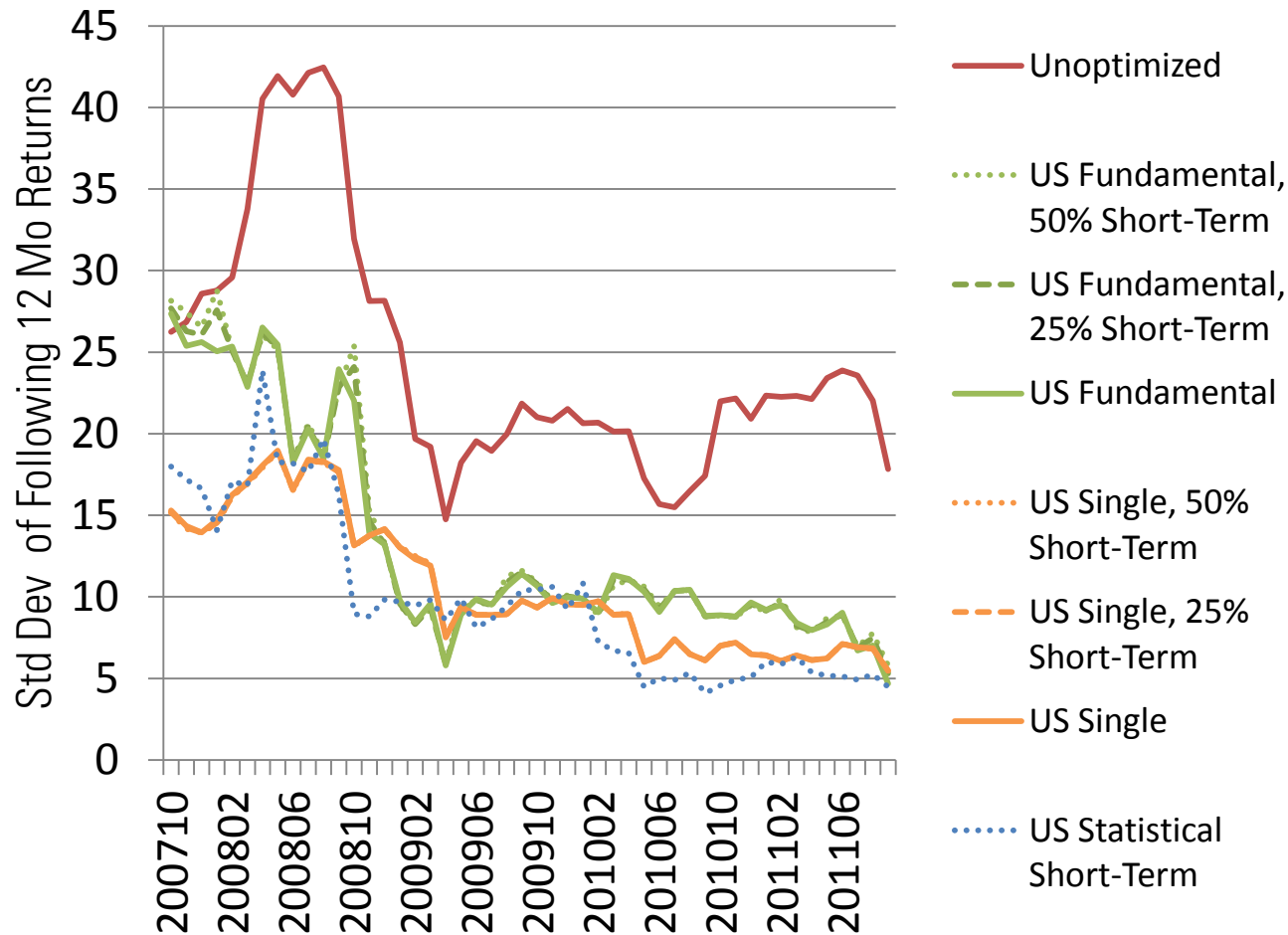
Horizon Blending in US Single Market, 2nd Gen



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.
- Similar to 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
- Predicted better than implied volatility alone

A Lot More to News

- Many vendors of quantified news – [Ravenpack](#), [Alexandria](#), [MarketPsych](#)
- 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, ...
- Time for machine learning
- Experiments ahead

Sidebar: Cross-Market Impact

- **Market Impact** = buying a stock raises its price, selling it lowers its price
- **Cross-Market Impact** = buying/selling IBM affects the price of MSFT
- It's hard enough to measure market impact. **How do you infer cross-market impact?**
- If one company's news regularly affects another's returns, you can build a proxy