



**Imposing Domain Expertise
on Sentiment Analysis**

Evan A. Schnidman Ph.D.

Forest or Trees?



Data Without Domain Expertise



Domain Expertise Without Data



Data is Everywhere

- ▶ Data is not limited to numerical.
- ▶ Examples in Finance:
 - Corporate Communications
 - Regulatory Filings
 - Central Bank Communications
- ▶ Need to know things not easily vectorized
- ▶ Dimension reduction to impose structure and unlock text as data



Imposing Structure

- ▶ Expertise allows us to impose structure on otherwise messy data, like text.



Traditional Sentiment Analysis



- ▶ **Good Buzzword minus Bad Buzzword == Sentiment**



- ▶ **Domain expertise allows for much more refined analysis**
- ▶ **Not a pure data science solution**
- ▶ **Time for experts to embrace tech and data science to utilize experts!**

Example: Central Banks

- ▶ **Central Bank communications are complex and important**
- ▶ **Focus today is Federal Reserve**



Traditional Fed Watching



The Briefcase Watch



Not Much has Changed



Experts in “Fed Speak”

Failed Attempts

- ▶ Experts are biased and fail to be comprehensive
- ▶ Simple text analysis dictionaries don't work for Fed Speak and other complex language
- ▶ Ex. “modest” v. “moderate”

Necessary Components

- ▶ Must use expertise to train the system based on whole communications
- ▶ Market response matters (Hawkish v. Dovish)



Enough documents
can eliminate bias



+

Expertise allows scaling
based on whole documents



=

End result is whole
communications scored
in orderly fashion



Qual Turned Quant

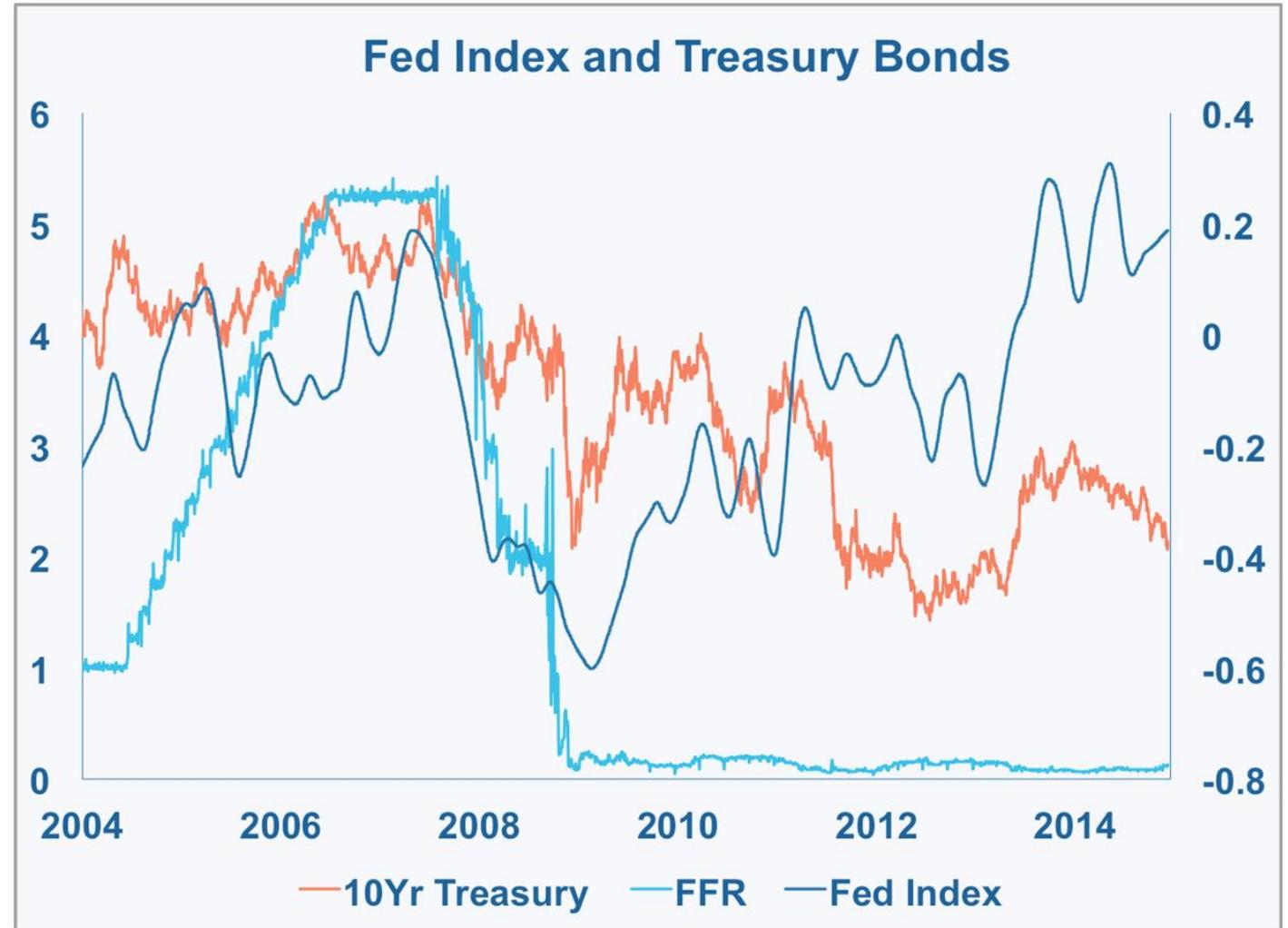


Resulting Data

- ▶ Comprehensive
- ▶ Unbiased
- ▶ Quantitative
- ▶ Fast

Many Possible Uses

- ▶ Eliminate post-hoc hedging on CB policy
- ▶ Forecast based on established correlations
- ▶ Add as a signal in multifactor model

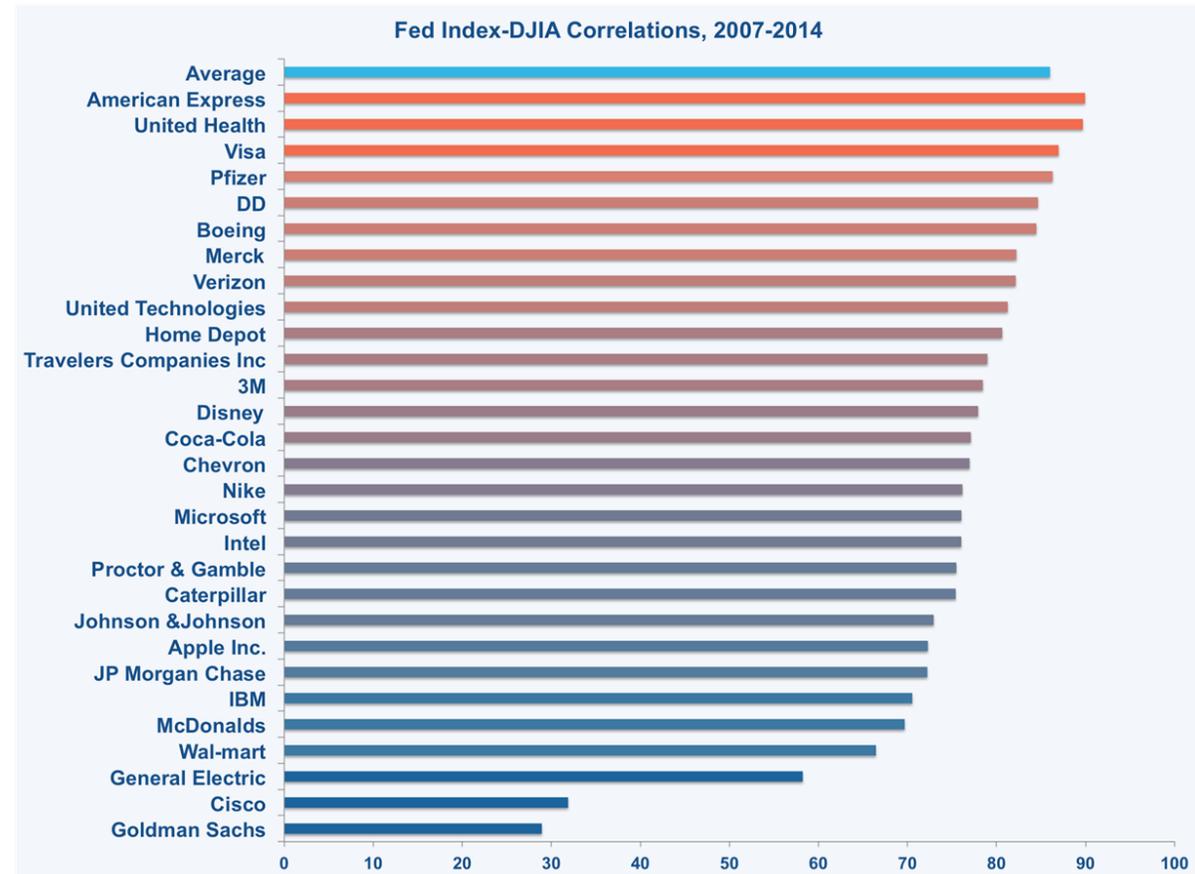
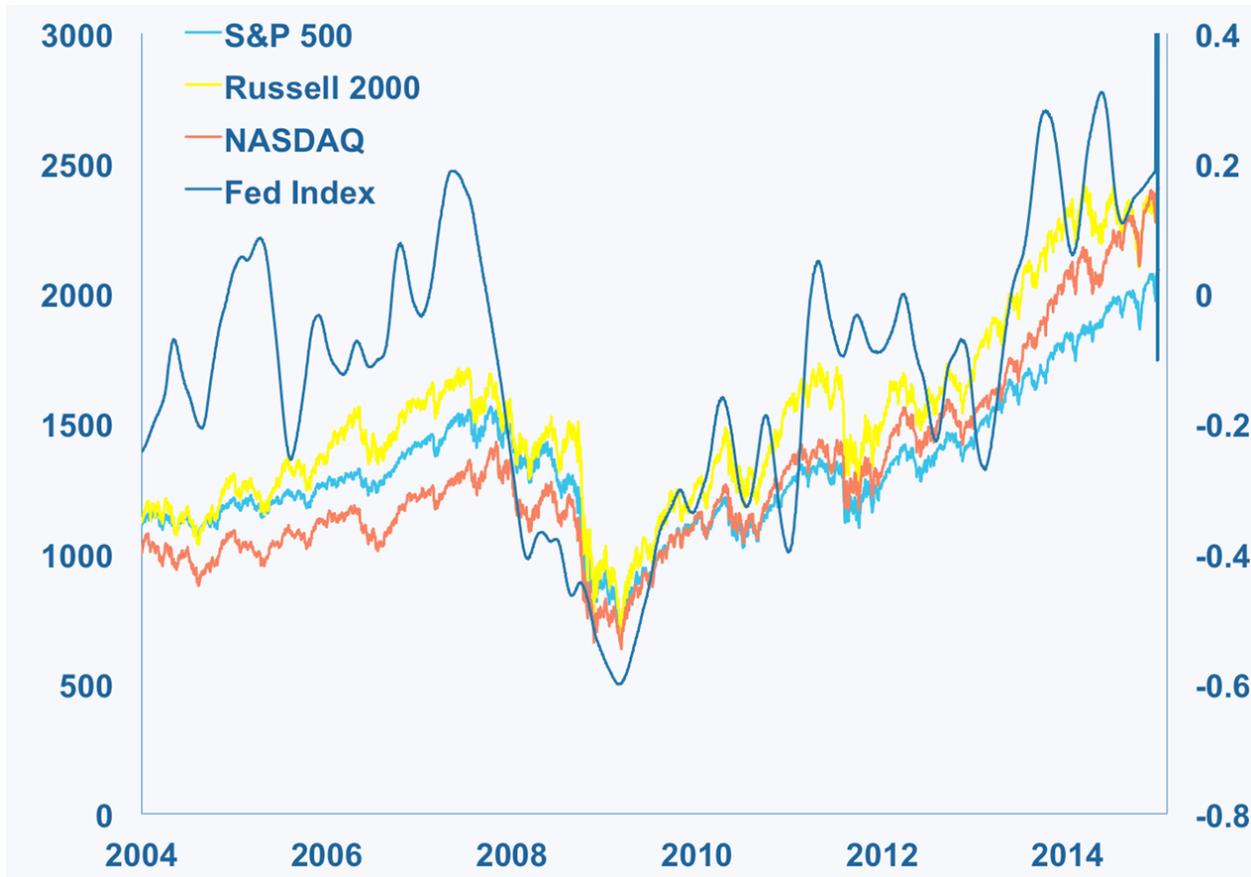


Trend matters more than value!

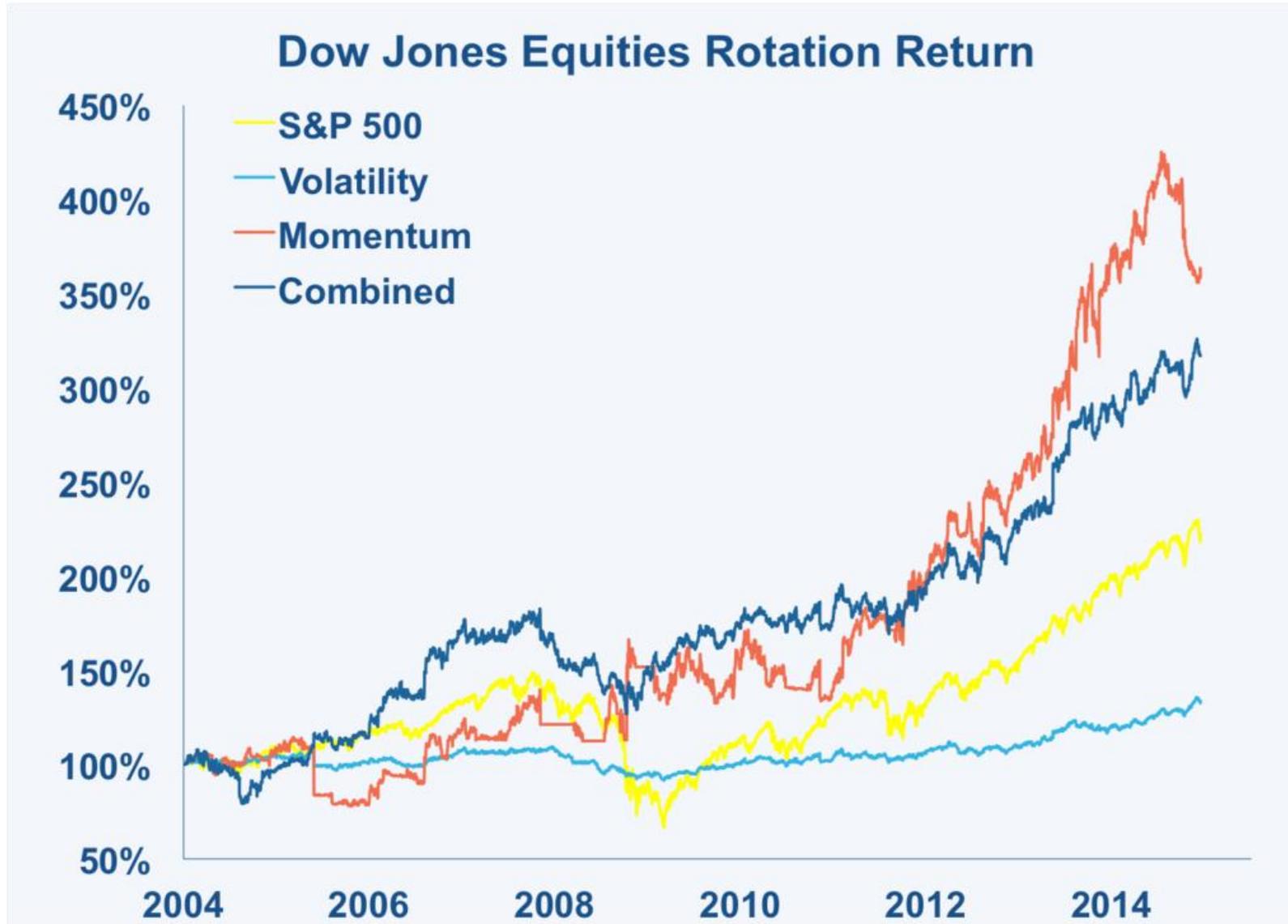
Equity Market Correlations



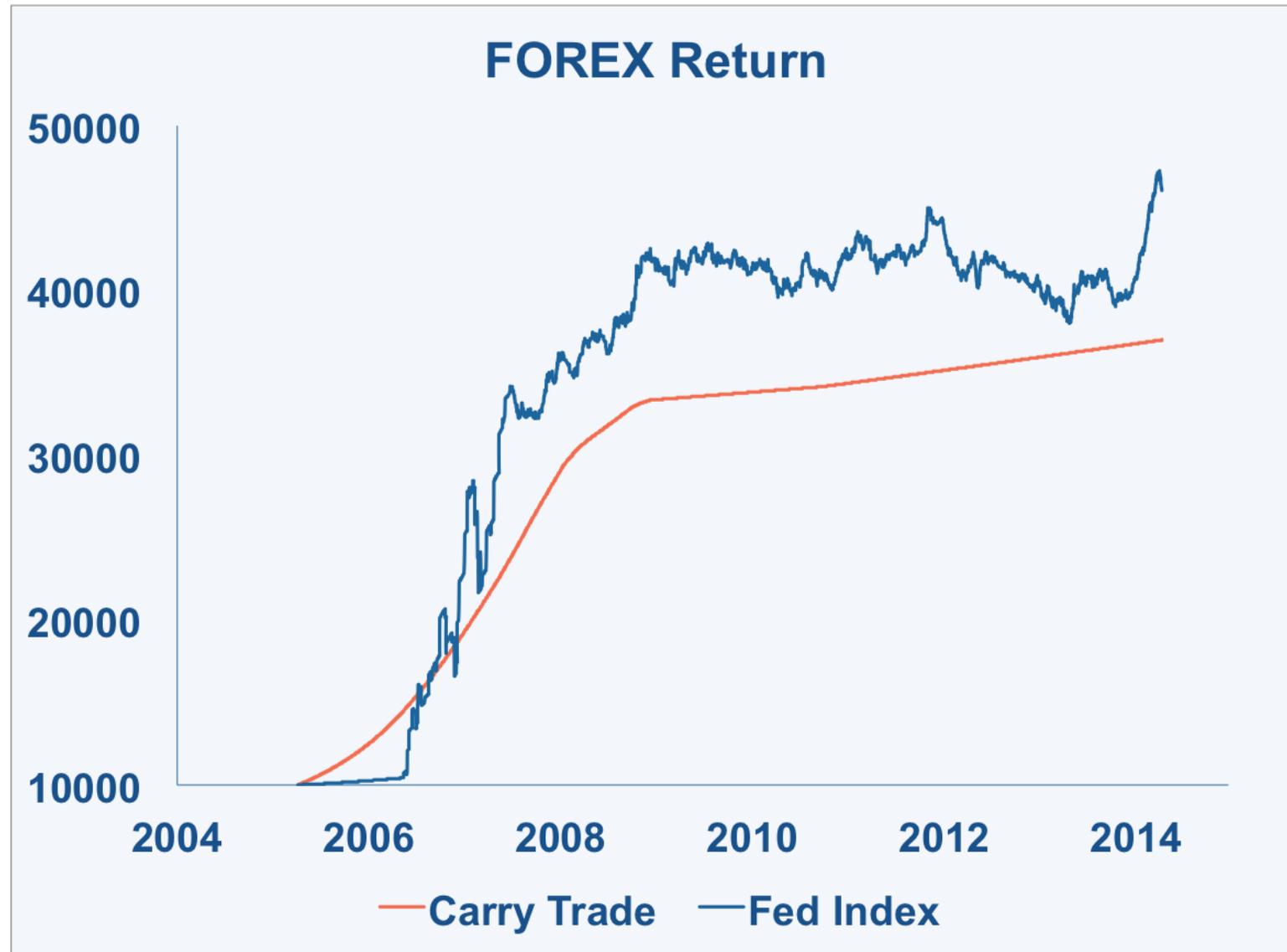
► Correlation across asset classes, not just Fixed Income



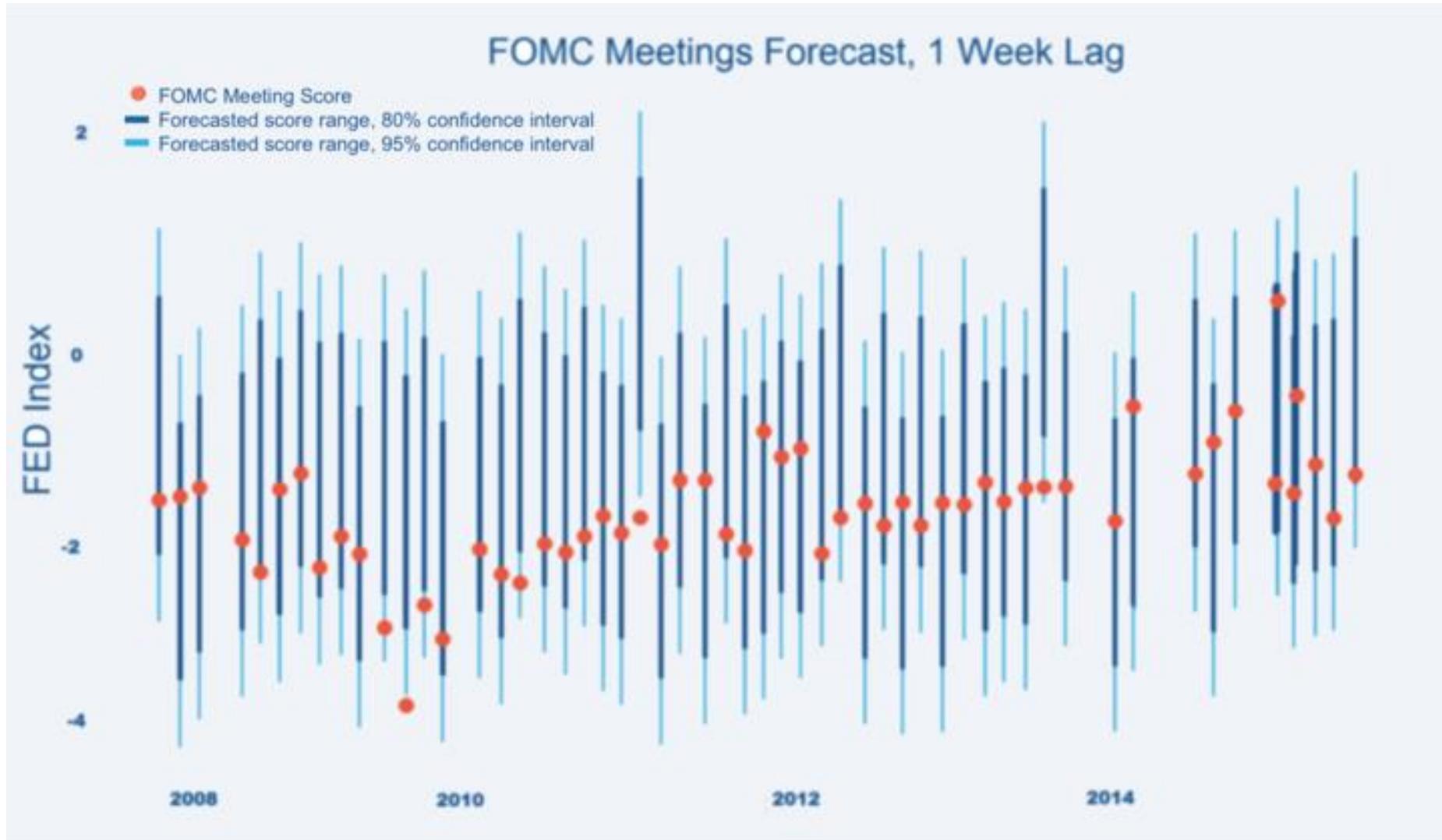
Equity Portfolio Simulation



FOREX Portfolio Simulation

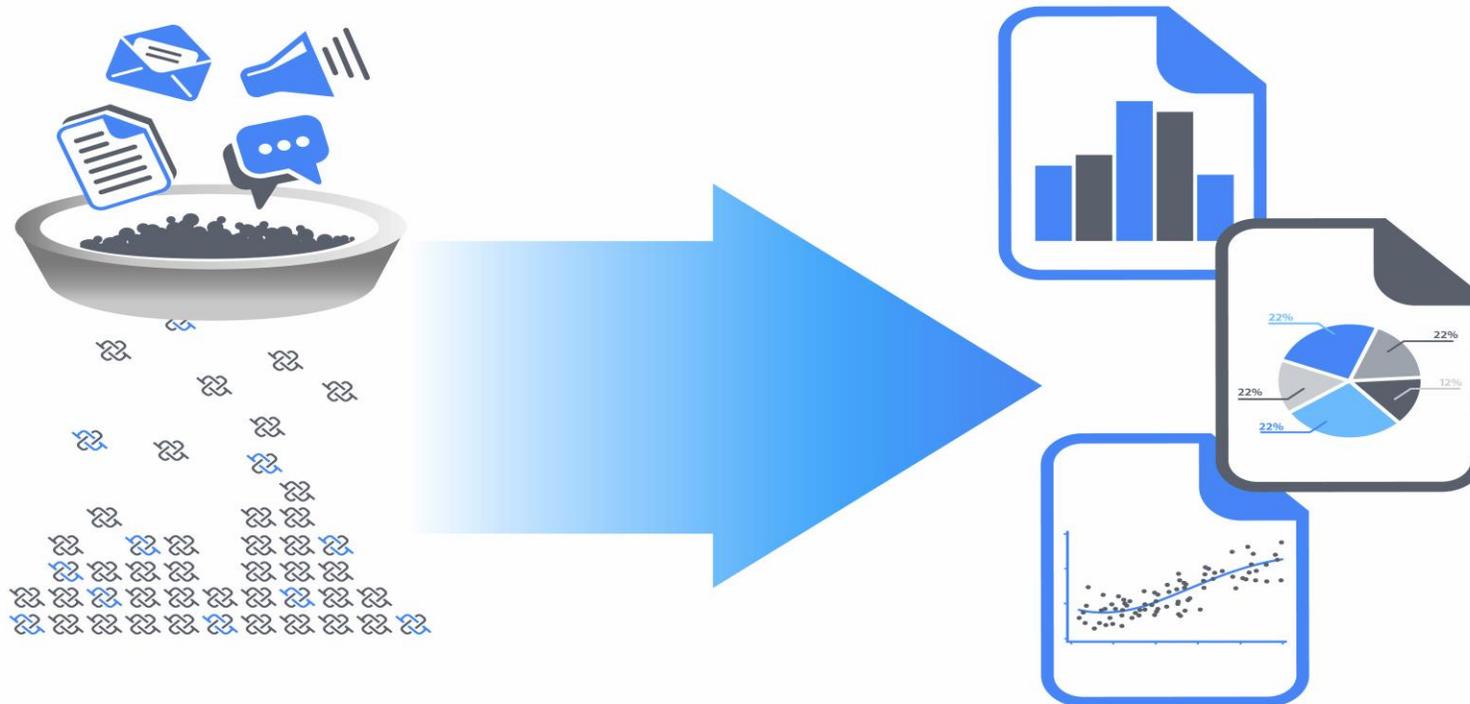


Leading Indicator



Other Applications

- ▶ Method translates across wide variety of financially important texts
- ▶ Regulatory and shareholder documents for individual equities
- ▶ Other regulatory information (Dodd-Frank, FDA, EPA etc.)



Data + Domain Expertise





**Imposing Domain Expertise
on Sentiment Analysis**

Evan A. Schnidman
evan@pratttle.co