News Implied Volatility and Disaster Concerns

Asaf Manela
Washington University in St. Louis

Alan Moreira
Yale University

November 2015

Motivation

2 minute intro to Asset Pricing for non-financial economists

- Price is expectation of discount factor m times future payoff x

$$P_{it} = E_t \left[m\left(s_{t+1}\right) x_i\left(s_{t+1}\right) \right]$$

- One could assume m is iid ($\Rightarrow$ constant expected returns)
  - Implies no predictability in stock returns
  - Efficient Markets Hypothesis (Fama, 1970)
- But prices move too much compared with future dividends and returns are predictable (Shiller, 1981)
  - m distribution and risk premia must be time-varying
- Modern AP models derive $m\left(s\right)$ to fit many "stylized facts"
  - Stochastic volatility, rare disasters, Knightian uncertainty, ...
- First-order business cycle effects (Gilchrist-Zakrajsek, 2012)

Motivation

2 minute intro to Asset Pricing for non-financial economists

- Price is expectation of discount factor m times future payoff x

$$P_{it} = E_t \left[m \left(s_{t+1} \right) x_i \left(s_{t+1} \right) \right]$$

- One could assume m is iid ($\Rightarrow$ constant expected returns)
  - Implies no predictability in stock returns
  - Efficient Markets Hypothesis (Fama, 1970)
- But prices move too much compared with future dividends and returns are predictable (Shiller, 1981)
  - m distribution and risk premia must be time-varying
- Modern AP models derive $m(s)$ to fit many "stylized facts"
  - Stochastic volatility, rare disasters, Knightian uncertainty, ...
- First-order business cycle effects (Gilchrist-Zakrajsek, 2012)

Motivation

2 minute intro to Asset Pricing for non-financial economists

- Price is expectation of discount factor m times future payoff x

$$P_{it} = E_t \left[m\left(s_{t+1}\right) x_i \left(s_{t+1}\right) \right]$$

- One could assume m is iid ($\Rightarrow$ constant expected returns)
  - Implies no predictability in stock returns
  - Efficient Markets Hypothesis (Fama, 1970)
- But prices move too much compared with future dividends and returns are predictable (Shiller, 1981)
  - m distribution and risk premia must be time-varying
- Modern AP models derive $m\left(s\right)$ to fit many “stylized facts”
  - Stochastic volatility, rare disasters, Knightian uncertainty, ...
- First-order business cycle effects (Gilchrist-Zakrajsek, 2012)

Motivation

2 minute intro to Asset Pricing for non-financial economists

- Price is expectation of discount factor m times future payoff x

$$P_{it} = E_t \left[m \left(s_{t+1} \right) x_i \left(s_{t+1} \right) \right]$$

- One could assume m is iid ($\Rightarrow$ constant expected returns)
  - Implies no predictability in stock returns
  - Efficient Markets Hypothesis (Fama, 1970)
- But prices move too much compared with future dividends and returns are predictable (Shiller, 1981)
  - m distribution and risk premia must be time-varying
- Modern AP models derive $m(s)$ to fit many "stylized facts"
  - Stochastic volatility, rare disasters, Knightian uncertainty, ...
- First-order business cycle effects (Gilchrist-Zakrajsek, 2012)

Motivation

2 minute intro to Asset Pricing for non-financial economists

- Price is expectation of discount factor m times future payoff x

$$P_{it} = E_t \left[m \left(s_{t+1} \right) x_i \left(s_{t+1} \right) \right]$$

- One could assume m is iid ($\Rightarrow$ constant expected returns)
  - Implies no predictability in stock returns
  - Efficient Markets Hypothesis (Fama, 1970)
- But prices move too much compared with future dividends and returns are predictable (Shiller, 1981)
  - m distribution and risk premia must be time-varying
- Modern AP models derive $m(s)$ to fit many “stylized facts”
  - Stochastic volatility, rare disasters, Knightian uncertainty, ...
- First-order business cycle effects (Gilchrist-Zakrajsek, 2012)

Our Goal

- Measure uncertainty about the future over a long history
- What types of uncertainty drive aggregate stock market risk premia?
- Starting point: time-variation in topics covered by business press reflects evolution of investors' concerns
- Our approach: estimate a news-based measure of uncertainty based on co-movement between front-page coverage of the *Wall Street Journal* and options-implied volatility (VIX)
  - News-implied volatility index (NVIX)
  - Use a machine learning technique (support-vector regression)
- NVIX has two useful features for our purposes
  1. Long-time series (1890–2009)
  2. Interpretable variation

Our Goal

- Measure uncertainty about the future over a long history
- What types of uncertainty drive aggregate stock market risk premia?
- Starting point: time-variation in topics covered by business press reflects evolution of investors' concerns
- Our approach: estimate a news-based measure of uncertainty based on co-movement between front-page coverage of the *Wall Street Journal* and options-implied volatility (VIX)
  - News-implied volatility index (NVIX)
  - Use a machine learning technique (support-vector regression)
- NVIX has two useful features for our purposes
  1. Long-time series (1890–2009)
  2. Interpretable variation

Our Goal

- Measure uncertainty about the future over a long history
- What types of uncertainty drive aggregate stock market risk premia?
- Starting point: time-variation in topics covered by business press reflects evolution of investors' concerns
- Our approach: estimate a news-based measure of uncertainty based on co-movement between front-page coverage of the *Wall Street Journal* and options-implied volatility (VIX)
  - News-implied volatility index (NVIX)
  - Use a machine learning technique (support-vector regression)
- NVIX has two useful features for our purposes
  1. Long-time series (1890–2009)
  2. Interpretable variation

Our Goal

- Measure uncertainty about the future over a long history
- What types of uncertainty drive aggregate stock market risk premia?
- Starting point: time-variation in topics covered by business press reflects evolution of investors' concerns
- Our approach: estimate a news-based measure of uncertainty based on co-movement between front-page coverage of the *Wall Street Journal* and options-implied volatility (VIX)
  - News-implied volatility index (NVIX)
  - Use a machine learning technique (support-vector regression)
- NVIX has two useful features for our purposes
  1. Long-time series (1890–2009)
  2. Interpretable variation

Results Summary

- News-implied volatility (NVIX) captures well the disaster concerns of the average investor over this longer history
  - Peaks during world wars, financial crises, times of policy-related uncertainty, and stock market crashes
- 1945–2009 US Post-war sample:
  - High NVIX is followed by above average stock returns
    - Even controlling for contemporaneous and forward-looking measures of stock market volatility
- Wars (47%) and government policy (23%) coverage explains most of the time variation in risk premia
- 1890–2009 sample includes Depression and two World Wars:
  - High NVIX predicts high future returns in normal times
  - Rises just before transitions into economic disasters
- Consistent with recent theories emphasizing time-varying rare disaster risk

Results Summary

- News-implied volatility (NVIX) captures well the disaster concerns of the average investor over this longer history
  - Peaks during world wars, financial crises, times of policy-related uncertainty, and stock market crashes
- 1945–2009 US Post-war sample:
  - High NVIX is followed by above average stock returns
    - Even controlling for contemporaneous and forward-looking measures of stock market volatility
- Wars (47%) and government policy (23%) coverage explains most of the time variation in risk premia
- 1890–2009 sample includes Depression and two World Wars:
  - High NVIX predicts high future returns in normal times
  - Rises just before transitions into economic disasters
- Consistent with recent theories emphasizing time-varying rare disaster risk

Results Summary

- News-implied volatility (NVIX) captures well the disaster concerns of the average investor over this longer history
  - Peaks during world wars, financial crises, times of policy-related uncertainty, and stock market crashes
- 1945–2009 US Post-war sample:
  - High NVIX is followed by above average stock returns
    - Even controlling for contemporaneous and forward-looking measures of stock market volatility
- Wars (47%) and government policy (23%) coverage explains most of the time variation in risk premia
- 1890–2009 sample includes Depression and two World Wars:
  - High NVIX predicts high future returns in normal times
  - Rises just before transitions into economic disasters
- Consistent with recent theories emphasizing time-varying rare disaster risk

Results Summary

- News-implied volatility (NVIX) captures well the disaster concerns of the average investor over this longer history
  - Peaks during world wars, financial crises, times of policy-related uncertainty, and stock market crashes
- 1945–2009 US Post-war sample:
  - High NVIX is followed by above average stock returns
    - Even controlling for contemporaneous and forward-looking measures of stock market volatility
- Wars (47%) and government policy (23%) coverage explains most of the time variation in risk premia
- 1890–2009 sample includes Depression and two World Wars:
  - High NVIX predicts high future returns in normal times
  - Rises just before transitions into economic disasters
- Consistent with recent theories emphasizing time-varying rare disaster risk

Results Summary

- News-implied volatility (NVIX) captures well the disaster concerns of the average investor over this longer history
  - Peaks during world wars, financial crises, times of policy-related uncertainty, and stock market crashes
- 1945–2009 US Post-war sample:
  - High NVIX is followed by above average stock returns
    - Even controlling for contemporaneous and forward-looking measures of stock market volatility
- Wars (47%) and government policy (23%) coverage explains most of the time variation in risk premia
- 1890–2009 sample includes Depression and two World Wars:
  - High NVIX predicts high future returns in normal times
  - Rises just before transitions into economic disasters
- Consistent with recent theories emphasizing time-varying rare disaster risk

Rare Disaster Asset Pricing

- Theory: Rietz (1988), Barro (2006), Gabaix (2012), Gourio (2008, 2012), Wachter (2013)
  - Disaster probability process is a key unobserved input
- Empirical: Backus-Chernov-Martin (2011), Bollerslev-Todorov (2011), Bates (2012), Kelly-Jiang (2014)
  - Focus on relatively short samples
  - Silent about the underlying drivers of disaster concerns

News Implied Volatility

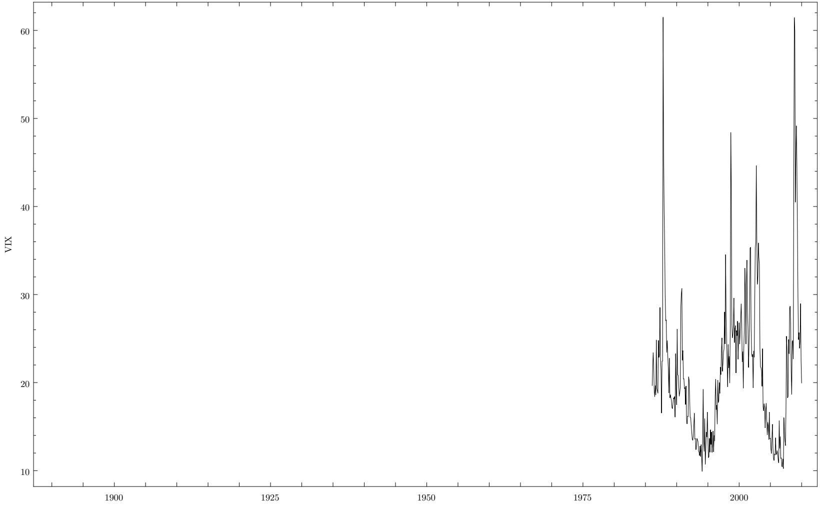
- Assumption: business press word choice provides a good and stable reflection of average investor's concerns
  - Reputation maximizing news firm observes real-world events and chooses what to emphasize in its report
  - Theoretical and empirical support
    - Gentzkow-Shapiro (2006), Tetlock (2007), Manela (2014)
- Asset pricing theory suggests options implied volatility (VIX) predicts stock market returns as it measures
  - Expected stock market volatility (Merton, 1973)
  - Variance risk premium (Drechsler-Yaron, 2011)
  - Probability of large disaster events (Gabaix, 2012; Gourio, 2012; Wachter, 2013)

News Implied Volatility

- Assumption: business press word choice provides a good and stable reflection of average investor's concerns
  - Reputation maximizing news firm observes real-world events and chooses what to emphasize in its report
  - Theoretical and empirical support
    - Gentzkow-Shapiro (2006), Tetlock (2007), Manela (2014)
- Asset pricing theory suggests options implied volatility (VIX) predicts stock market returns as it measures
  - Expected stock market volatility (Merton, 1973)
  - Variance risk premium (Drechsler-Yaron, 2011)
  - Probability of large disaster events (Gabaix, 2012; Gourio, 2012; Wachter, 2013)

News Implied Volatility

VIX (VXO) is available only recently, 1986-present



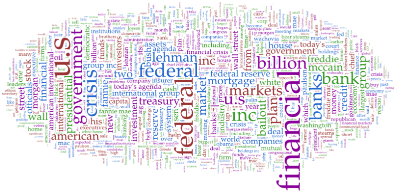
Our Data

We have news, front-page titles and abstracts of the *Wall Street Journal*, 1890-2009

Date	Title	Abstract
2008-09-16	AIG Faces Cash Crisis As Stock Dives 61%	American International Group Inc. was facing a severe cash ...
2008-09-16	AIG, Lehman Shock Hits World Markets ...	The convulsions in the U.S. financial system sent markets ...
2008-09-16	Business and Finance	Central banks around the world pumped cash into money ...
2008-09-16	Keeping Their Powder Dry: Draft Boards ...	The Selective Service System has the awkward task of ...
2008-09-16	Old-School Banks Emerge Atop New ...	Banks are heading "back to basics – to, if you like, the core ...
2008-09-16	World-Wide	Thailand's ruling party chose ousted leader Thaksin's ...

$$VIX_t - \overline{VIX} = w_0 + \mathbf{w} \cdot \mathbf{x}_t + \upsilon_t$$

September 2008:



Raw word frequencies

Weighted word frequencies

News Implied Volatility

Support Vector Regression Avoids Overfitting

- SVR regression estimates $\mathbf{w}$, a $K \gg T$ vector of coefficients

$$VIX_t - \overline{VIX} = w_0 + \mathbf{w} \cdot \mathbf{x}_t + \upsilon_t \qquad t = 1 \ldots T \qquad (1)$$

- $\mathbf{w}$ is restricted to be a weighted-average of regressors
- Only the weights α_t of *support vectors* are non-zero

$$\hat{\mathbf{w}}_{SVR} = \sum_{t \in train} \alpha_t \mathbf{x}_t \qquad (2)$$

- Support vectors are word usage vectors of months that are "important" in the train sample
  - Benefit: Reduces an infeasible problem $O(K)$, to a feasible one $O(T)$
  - Benefit: Method has been shown to predict well out-of-sample
  - Cost: SVR cannot concentrate on $\mathbf{x}_t$ subspaces or do standard inference

News Implied Volatility

Support Vector Regression Avoids Overfitting

- SVR regression estimates $\mathbf{w}$, a $K \gg T$ vector of coefficients

$$VIX_t - \overline{VIX} = w_0 + \mathbf{w} \cdot \mathbf{x}_t + \upsilon_t \qquad t = 1 \dots T \qquad (1)$$

- $\mathbf{w}$ is restricted to be a weighted-average of regressors
- Only the weights α_t of *support vectors* are non-zero

$$\hat{\mathbf{w}}_{SVR} = \sum_{t \in train} \alpha_t \mathbf{x}_t \qquad (2)$$

- Support vectors are word usage vectors of months that are "important" in the train sample
  - Benefit: Reduces an infeasible problem $O(K)$, to a feasible one $O(T)$
  - Benefit: Method has been shown to predict well out-of-sample
  - Cost: SVR cannot concentrate on $\mathbf{x}_t$ subspaces or do standard inference

News Implied Volatility

Support Vector Regression Avoids Overfitting

- SVR regression estimates $\mathbf{w}$, a $K \gg T$ vector of coefficients

$$VIX_t - \overline{VIX} = w_0 + \mathbf{w} \cdot \mathbf{x}_t + \upsilon_t \qquad t = 1 \dots T \qquad (1)$$

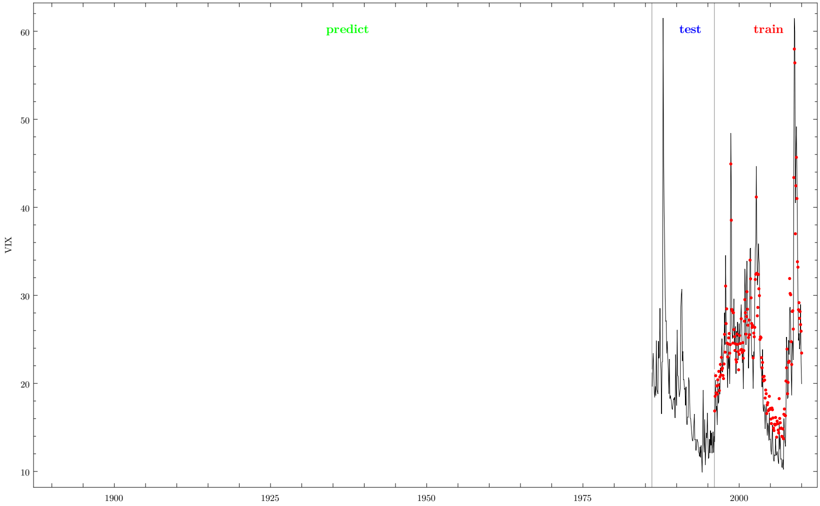
- $\mathbf{w}$ is restricted to be a weighted-average of regressors
- Only the weights α_t of *support vectors* are non-zero

$$\hat{\mathbf{w}}_{SVR} = \sum_{t \in train} \alpha_t \mathbf{x}_t \qquad (2)$$

- Support vectors are word usage vectors of months that are "important" in the train sample
  - Benefit: Reduces an infeasible problem $O(K)$, to a feasible one $O(T)$
  - Benefit: Method has been shown to predict well out-of-sample
  - Cost: SVR cannot concentrate on $\mathbf{x}_t$ subspaces or do standard inference

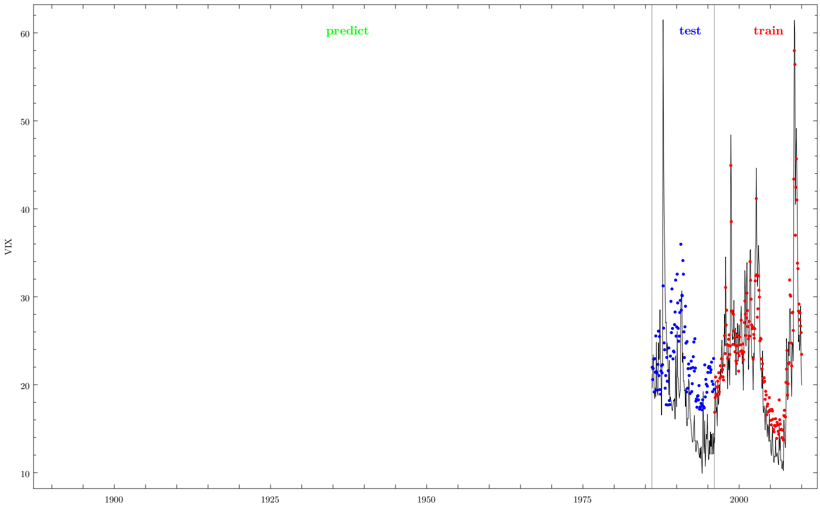
News Implied Volatility

Support Vector Regression: $VIX_t - \overline{VIX} = w_0 + \mathbf{w} \cdot \mathbf{x}_t + \upsilon_t$



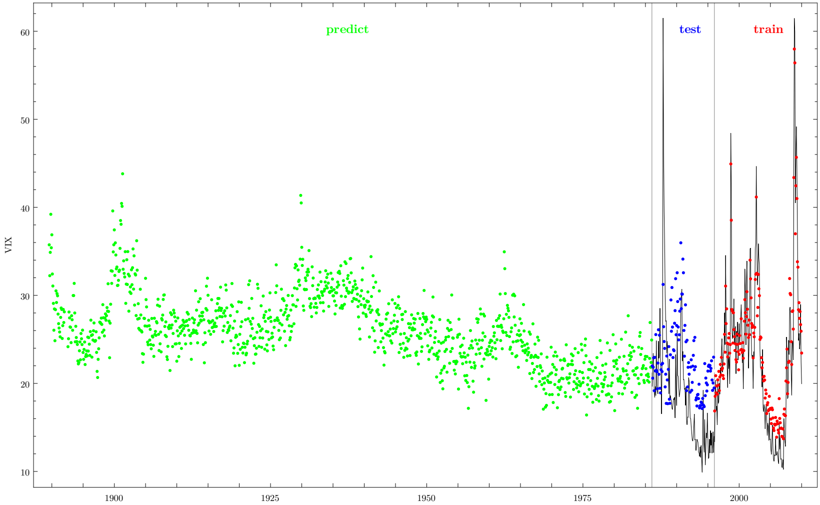
News Implied Volatility

Out-of-sample Fit: $RMSE\,[test] = 7.52$ ($R^2\,[test] = 0.34$)



News Implied Volatility

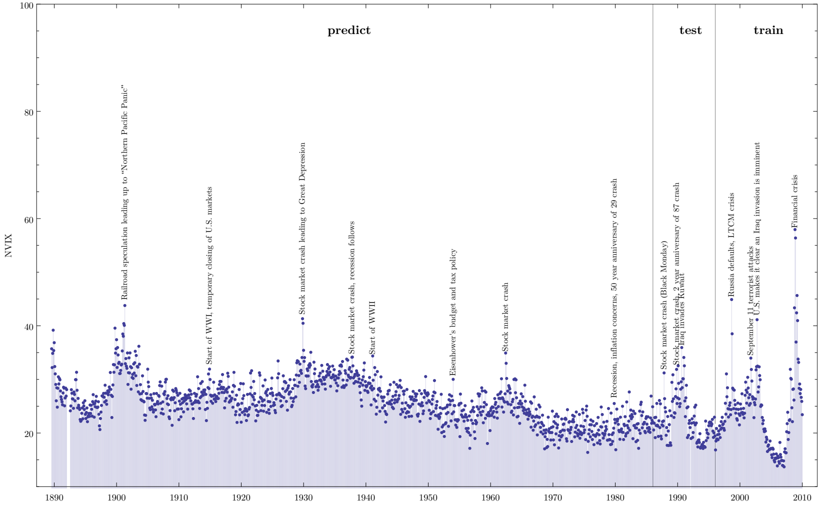
Fig. 1: NVIX captures well the fears of the average investor over this long history



NVIX interactive chart with word clouds available on Asaf Manela's website

Is NVIX a Reasonable Proxy for Uncertainty?

Fig. 2: NVIX peaks during stock market crashes, times of policy-related uncertainty, world wars and financial crises



Word-choice Stability and Measurement Error

- Common concern: meaning of certain words or phrases used by the press may change considerably over our long sample
  - e.g. "Japanese navy" in 1940s vs. today
- Wish to quantify the increase in measurement error from moving back in time
  - But VIX is unavailable before 1986
  - We use realized volatility (a blood-related cousin)
- Find that our predictive ability over long sample is quite stable
  - Out-of-sample RMSE increases from 9.6 to 10.9 percent volatility moving from *test* to *predict* subsample (Table 2)
  - SVR is designed to and seems to avoid overfitting
    - Even in 1890 WSJ was written in English ...

Word-choice Stability and Measurement Error

- Common concern: meaning of certain words or phrases used by the press may change considerably over our long sample
  - e.g. "Japanese navy" in 1940s vs. today
- Wish to quantify the increase in measurement error from moving back in time
  - But VIX is unavailable before 1986
  - We use realized volatility (a blood-related cousin)
- Find that our predictive ability over long sample is quite stable
  - Out-of-sample RMSE increases from 9.6 to 10.9 percent volatility moving from *test* to *predict* subsample (Table 2)
  - SVR is designed to and seems to avoid overfitting
    - Even in 1890 WSJ was written in English ...

Word-choice Stability and Measurement Error

- Common concern: meaning of certain words or phrases used by the press may change considerably over our long sample
  - e.g. "Japanese navy" in 1940s vs. today
- Wish to quantify the increase in measurement error from moving back in time
  - But VIX is unavailable before 1986
  - We use realized volatility (a blood-related cousin)
- Find that our predictive ability over long sample is quite stable
  - Out-of-sample RMSE increases from 9.6 to 10.9 percent volatility moving from *test* to *predict* subsample (Table 2)
  - SVR is designed to and seems to avoid overfitting
    - Even in 1890 WSJ was written in English ...

Alternative Text-based Analysis Approaches

- We use Support Vector Regression (SVR) to overcome the large dimensionality of the words space
- Our approach lets the data speak
- Kogan et al (2009) use SVR to predict firm-specific volatility using 10-Ks
- Two alternative approaches suggested by previous literature:
  1. Create topic-specific compound search statement and count the resulting number of articles
     - e.g. Baker-Bloom-Davis (2013) search for articles containing the term 'uncertainty' or 'uncertain', the terms 'economic' or 'economy' and one additional term such as 'policy', 'tax', etc.
  2. Classifies words into word lists that share a common tone and count all occurrences of words in the text belonging to a particular word list
     - e.g. Loughran-McDonald (2011) develops a negative word list, along with five other word lists, that reflect tone in financial text and relate them to 10-Ks filing returns

Alternative Text-based Analysis Approaches

- We use Support Vector Regression (SVR) to overcome the large dimensionality of the words space
- Our approach lets the data speak
- Kogan et al (2009) use SVR to predict firm-specific volatility using 10-Ks
- Two alternative approaches suggested by previous literature:
  1. Create topic-specific compound search statement and count the resulting number of articles
     - e.g. Baker-Bloom-Davis (2013) search for articles containing the term 'uncertainty' or 'uncertain', the terms 'economic' or 'economy' and one additional term such as 'policy', 'tax', etc.
  2. Classifies words into word lists that share a common tone and count all occurrences of words in the text belonging to a particular word list
     - e.g. Loughran-McDonald (2011) develops a negative word list, along with five other word lists, that reflect tone in financial text and relate them to 10-Ks filing returns

Alternative Text-based Analysis Approaches

- We use Support Vector Regression (SVR) to overcome the large dimensionality of the words space
- Our approach lets the data speak
- Kogan et al (2009) use SVR to predict firm-specific volatility using 10-Ks
- Two alternative approaches suggested by previous literature:
  1. Create topic-specific compound search statement and count the resulting number of articles
     - e.g. Baker-Bloom-Davis (2013) search for articles containing the term 'uncertainty' or 'uncertain', the terms 'economic' or 'economy' and one additional term such as 'policy', 'tax', etc.
  2. Classifies words into word lists that share a common tone and count all occurrences of words in the text belonging to a particular word list
     - e.g. Loughran-McDonald (2011) develops a negative word list, along with five other word lists, that reflect tone in financial text and relate them to 10-Ks filing returns

Return Predictability

- Models with time-varying risk premia suggest that times when risk is relatively high would be followed by above average stock market returns
  - Time-varying volatility (Merton, 1973)
  - Time-varying disaster risk (e.g. Gabaix, 2012)
- Prescribe a regression of excess stock returns on lagged forward-looking risk measured by $NVIX^2$
- First focus on post-war period (quality data, no disasters)

NVIX Predicts Post-War Stock Market Returns

Tbl 3: $\sigma\left(NVIX^2\right)$ change means 3.4 pp higher annualized excess return next year

$$r^e_{t\rightarrow t+\tau} = \beta_0 + \beta_1 NVIX^2_{t-1} + \epsilon_{t+\tau}$$

τ months		1945-2009	1945-1995	1986-2009
1	β_1	0.15	0.33**	0.09
	$t(\beta_1)$	[1.04]	[2.21]	[0.58]
	R^2	0.37	0.74	0.28
6	β_1	0.18***	0.39***	0.11
	$t(\beta_1)$	[2.59]	[3.72]	[1.44]
	R^2	2.56	4.91	1.93
12	β_1	0.16***	0.28***	0.10
	$t(\beta_1)$	[3.27]	[2.79]	[1.64]
	R^2	3.50	4.78	2.99
24	β_1	0.14***	0.19**	0.11**
	$t(\beta_1)$	[3.55]	[2.17]	[2.13]
	R^2	5.12	4.26	6.13
	Obs	779	611	287

Drill-down into Predictability

Disentangle several types of uncertainty potentially in NVIX

- Time-varying volatility does not explain these results
  - NVIX coefficients and significance hardly change with $Variance_t$ controls (Table 4)
  - Why?

$$VIX_t^2 = Variance_t + RiskAdjustment_t$$

  - Newspaper does a good job filtering out volatility part
- Horse races with financial predictors
  - NVIX captures additional information relative to variance-based measured of VIX, credit spreads, or price/earnings ratio (Table 5)
- Alternative measures of uncertainty focused on tail risk
  - NVIX captures concerns about large and infrequent macroeconomic disasters (Table 6)

Drill-down into Predictability

Disentangle several types of uncertainty potentially in NVIX

- Time-varying volatility does not explain these results
  - NVIX coefficients and significance hardly change with $Variance_t$ controls (Table 4)
  - Why?

$$VIX_t^2 = Variance_t + RiskAdjustment_t$$

  - Newspaper does a good job filtering out volatility part
- Horse races with financial predictors
  - NVIX captures additional information relative to variance-based measured of VIX, credit spreads, or price/earnings ratio (Table 5)
- Alternative measures of uncertainty focused on tail risk
  - NVIX captures concerns about large and infrequent macroeconomic disasters (Table 6)

Drill-down into Predictability

Disentangle several types of uncertainty potentially in NVIX

- Time-varying volatility does not explain these results
  - NVIX coefficients and significance hardly change with $Variance_t$ controls (Table 4)
  - Why?

$$VIX_t^2 = Variance_t + RiskAdjustment_t$$

  - Newspaper does a good job filtering out volatility part
- Horse races with financial predictors
  - NVIX captures additional information relative to variance-based measured of VIX, credit spreads, or price/earnings ratio (Table 5)
- Alternative measures of uncertainty focused on tail risk
  - NVIX captures concerns about large and infrequent macroeconomic disasters (Table 6)

Origins of Uncertainty Fluctuations

- What were investors worried about?
- Text-based measure allows us to study which concerns drive risk premia
- Content analysis
  - Classify words into five broad categories
  - Rely on Princeton's widely used WordNet project

Categories Total Variance Share

Tbl 8: *Stock Market* words explain half the variation in NVIX, *War* words explain 6%

Category	% of Variance	n-grams	Top n-grams
Government	2.59	83	tax, money, rates, government, plan
Intermediation	2.24	70	financial, business, bank, credit, loan
Natural Disaster	0.01	63	fire, storm, aids, happening, shock
Stock Market	51.67	59	stock, market, stocks, industry, markets
War	6.22	46	war, military, action, world war, violence
Unclassified	37.30	373988	u.s, washington, gold, special, treasury

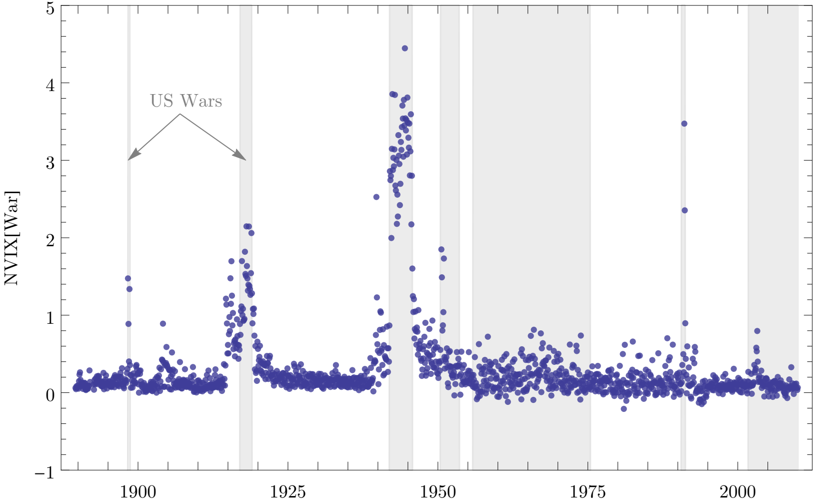
Which Concerns Drive Risk Premia Variation?

Risk premia decomposition strongly supports the time-varying rare disaster risk model

- Risk premia decomposition (Table 9):
  - *War* words explain 47% of risk premia variation
  - *Government* words explain 23%
  - Other categories are insignificant
- About half the variation in risk premia is unequivocally about disaster concerns

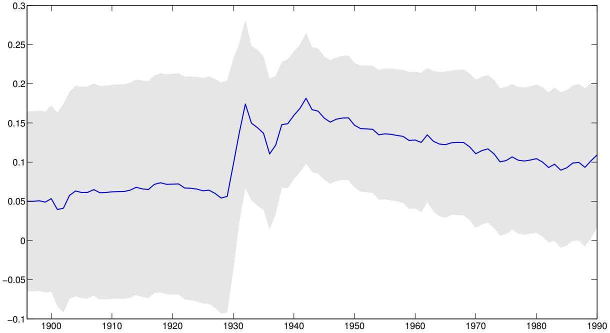
NVIX due to War-related Words

Fig 3b: Captures well not only whether the US was engaged in war, but also the degree of concern about the future prevalent at the time



Predictability Coefficients Starting in Year X until 2009

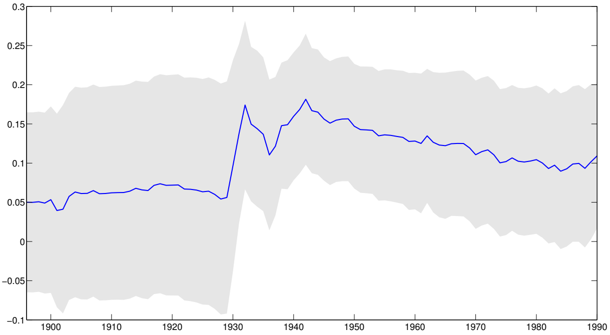
Fig 4: Inclusion of Great Depression or WWII has a large impact on our estimates



- Two plausible explanations could attenuate predictability
  1. Disaster realizations
  2. Long-lasting disaster periods (Nakamura et al, 2013)
- We fit a structural model to filter disaster states

Predictability Coefficients Starting in Year X until 2009

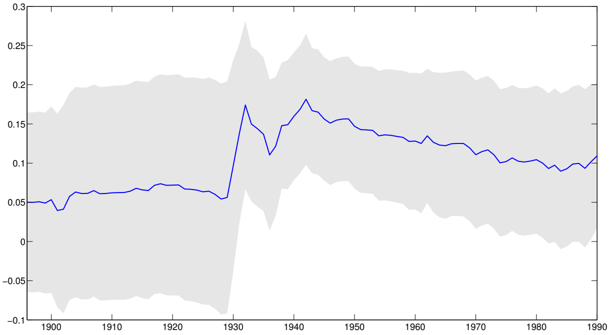
Fig 4: Inclusion of Great Depression or WWII has a large impact on our estimates



- Two plausible explanations could attenuate predictability
  1. Disaster realizations
  2. Long-lasting disaster periods (Nakamura et al, 2013)
- We fit a structural model to filter disaster states

Predictability Coefficients Starting in Year X until 2009

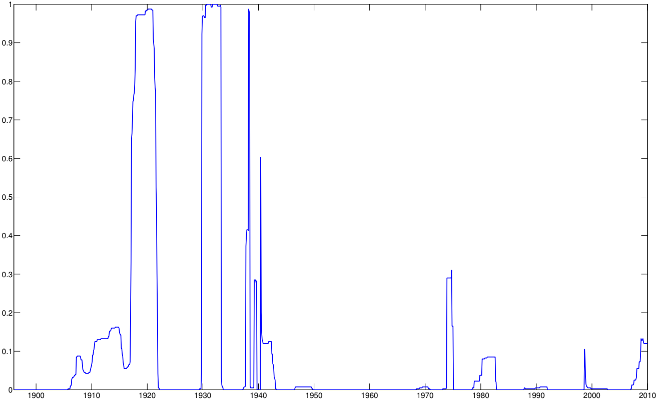
Fig 4: Inclusion of Great Depression or WWII has a large impact on our estimates



- Two plausible explanations could attenuate predictability
  1. Disaster realizations
  2. Long-lasting disaster periods (Nakamura et al, 2013)
- We fit a structural model to filter disaster states

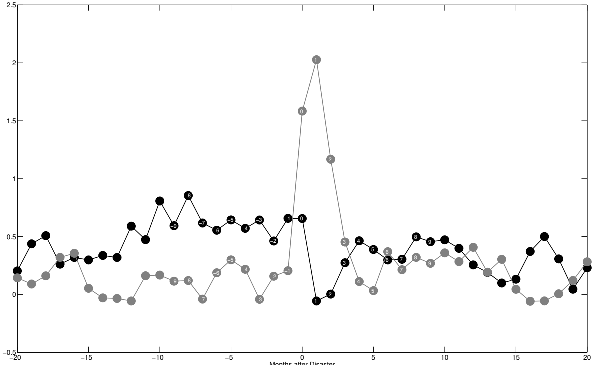
Filtered Probability that the Economy is in a Disaster State

Fig 5: disasters identified from consumption data, but timing from stock market returns



Disaster Predictability

Fig 6: NVIX is consistently above average up to a year before disaster, but variance-based measures are not



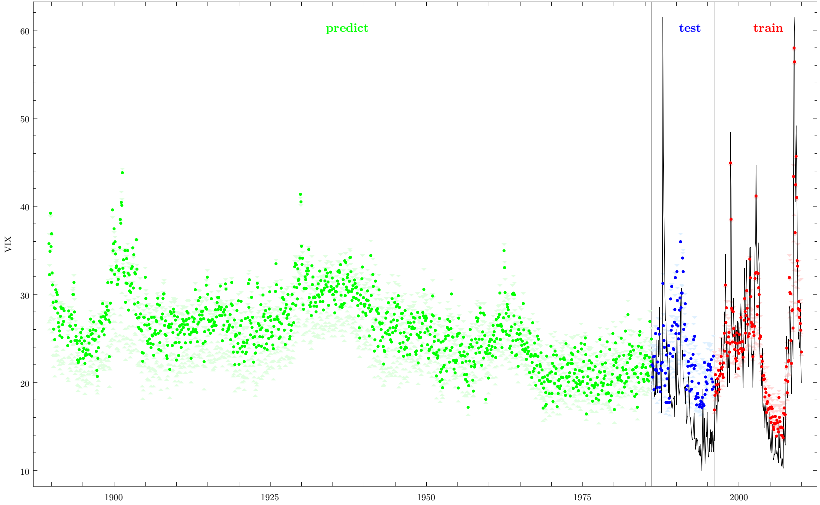
- Mechanically attenuates return predictability
- Return predictability reemerges in full sample when conditioning on non-disaster states (Table 11)

Conclusion

- We propose a text-based method to extend options-implied measures of uncertainty back to 1890
  - NVIX is plausibly related with concerns about rare disasters
  - Out-of-sample fit is stable over the long sample
- NVIX predicts returns and large economic disasters
  - Predictability results largely driven by war related concerns
- Strong evidence in new data for an asset pricing model with time-varying disaster concerns
- A step forward in applying text analysis to answer difficult economic questions
  - Content analysis is promising avenue for future research

News Implied Volatility

Fig. 1: Estimation is not sensitive to randomizations of the *train* subsample



News-Implied *Realized Volatility*

Tbl 2: SVR predictive ability over long sample is quite stable

Subsample	$RMSE$ SVR	R^2 SVR	$RMSE$ Reg	R^2 Reg	Correlation
train	3.35	0.68	2.64	0.93	0.96
test	9.60	0.27	9.09	0.20	0.45
predict	10.91	0.38	8.49	0.16	0.40

Stochastic Volatility Does Not Explain these Results

Tbl 4: NVIX coefficients and significance hardly change with $E_t\left[Var\right]$ controls

$$r^e_{t\to t+\tau} = \beta_0 + \beta_1 NVIX^2_{t-1} + \beta_2 EVAR_{t-1} + \epsilon_t$$

τ		(1)	(2)	(3)	(4)	(5)
1	β_1	0.21	0.21	0.23	0.21	0.26
	$t(\beta_1)$	[1.59]	[1.47]	[1.6]	[1.64]	[1.62]
	R^2	0.55	0.46	0.52	0.49	0.48
6	β_1	0.19**	0.22***	0.24***	0.23***	0.27**
	$t(\beta_1)$	[2.51]	[2.64]	[2.91]	[2.93]	[2.44]
	R^2	2.57	2.75	3.01	2.87	2.94
12	β_1	0.17***	0.19***	0.21***	0.20***	0.26**
	$t(\beta_1)$	[3.15]	[2.77]	[2.98]	[2.92]	[2.39]
	R^2	3.56	3.75	4.19	4.14	4.36
24	β_1	0.15***	0.17***	0.19***	0.21***	0.30***
	$t(\beta_1)$	[3.32]	[2.79]	[2.8]	[2.98]	[2.67]
	R^2	5.18	5.51	6.27	7.35	8.67
	Obs	779	778	778	778	778
	EVAR Model R^2	9.21	25.53	25.87	28.22	31.83

Horse Races with Financial Predictors

Tbl 5: NVIX captures additional information relative to variance-based measured of VIX, credit spreads, or price/earnings ratio

$$r^e_{t\to t+\tau} = \beta_0 + \beta_1 NVIX^2_{t-1} + \sum_{j=2}^{N} \beta_j X_{j,t-1} + \epsilon_{t+\tau}$$

τ		(1)	(2)	(3)	(4)	(5)
1	β_1	0.15	0.20	0.21	0.19	-
	$t(\beta_1)$	[1.04]	[1.45]	[1.43]	[1.32]	-
	R^2	0.37	0.45	0.51	0.85	0.49
6	β_1	0.18***	0.22***	0.22***	0.21**	-
	$t(\beta_1)$	[2.59]	[2.64]	[2.63]	[2.42]	-
	R^2	2.56	2.73	3.51	5.34	3.33
12	β_1	0.16***	0.19***	0.19***	0.18***	-
	$t(\beta_1)$	[3.27]	[2.78]	[2.79]	[2.62]	-
	R^2	3.50	3.72	4.47	8.80	6.22
24	β_1	0.14***	0.17***	0.17***	0.15***	-
	$t(\beta_1)$	[3.55]	[2.82]	[2.82]	[3.01]	-
	R^2	5.12	5.49	5.49	16.46	12.99
	Obs	779	779	779	779	779
Controls						
$NVIX^2_{t-1}$		yes	yes	yes	yes	no
$E\left[VIX^2_{t-1} \mid VAR\right]$		no	yes	yes	yes	yes
$Creditspread_{t-1}$		no	no	yes	yes	yes
$(\frac{P}{E})_{t-1}$		no	no	no	yes	yes

Alternative Measures of Uncertainty Focused on Tail Risk

Tbl 6: NVIX captures concerns about large and infrequent macroeconomic disasters

$$r^e_{t\to t+\tau} = \beta_0 + \beta_1 \widehat{X_{t-1}} + \beta_2 EVAR_{t-1} + \epsilon_{t+\tau}$$

τ	X :	VIX^2	VIX premium	LT	Slope
1	β_1	0.21	0.42***	1.39*	128.21*
	$t(\beta_1)$	[1.47]	[2.62]	[1.82]	[1.93]
	R^2	0.46	1.34	0.43	0.53
6	β_1	0.22***	0.18**	1.33**	80.13**
	$t(\beta_1)$	[2.64]	[2.14]	[2.02]	[1.98]
	R^2	2.75	1.60	2.22	1.40
12	β_1	0.19***	0.12*	1.26**	57.19*
	$t(\beta_1)$	[2.77]	[1.87]	[2.45]	[1.73]
	R^2	3.75	1.67	3.51	1.53
24	β_1	0.17***	0.11**	0.82*	54.65**
	$t(\beta_1)$	[2.79]	[2.20]	[1.70]	[2.33]
	R^2	5.51	2.34	3.15	2.39
	Obs	779	779	779	779

Risk Premia Decomposition, 12-months Horizon

Tbl 9: *War* words explain 47% of risk premia variation, *Government* explains 23%

	1945–2009	1896–1945	1896–2009
Government	4.22***	-0.57	2.54**
	[2.90]	[0.26]	[2.12]
	(57.18)	(0.57)	(23.19)
War	3.03**	3.76***	3.63***
	[2.32]	[2.65]	[4.37]
	(13.54)	(59.99)	(47.45)
Intermediation	0.70	1.19	1.38
	[0.40]	[0.52]	[0.97]
	(1.49)	(3.09)	(6.8)
Stock Markets	-0.73	-2.78	-1.07
	[0.24]	[1.09]	[0.58]
	(0.16)	(23.44)	(4.09)
Natural Disaster	1.08*	-0.28	1.04
	[1.70]	[0.15]	[1.54]
	(5.88)	(0.05)	(3.87)
R^2	9.12	6.52	6.33
Obs	779	588	1367