Dynamic Interpretation of Emerging Risks in the Financial Sector

Kathleen Weiss Hanley and Gerard Hoberg

February 28, 2018

Presented by Symphony Liu on October 2, 2018 in
Fin 500R: Topics in Quantitative Finance class

Agenda

- Previous research on emerging risks in the financial sector
- Central hypothesis
- Data
  - Investors
  - Banks
- Two text analytic methods
  - Latent Dirichlet Allocation (LDA)
  - Semantic Vector Analysis (SVA)
- Models and results
  - Static Model
  - Dynamic Model
  - Drill-down Model

Emerging Risks in the Financial Sector

Previous research

- Bisias, Flood, Lo, and Valavanis (2012) provide a survey of over 30 risk metrics
  - macroeconomic indicators
  - illiquidity and solvency metrics
  - the probability of financial distress
- Two primary limitations of standard computational linguistic methods
  - Difficult to interpret(network, principal components)
  - Backward-looking (liquidity mismatch, housing sector risk)

Central Hypothesis

When risk is building in the financial sector, a regression of *pairwise return covariance* on the *risk themes* will become significant and produce an elevated R^2. When no emerging risk is present, this R^2 will be close to zero.

$$Covariance_{i,j,t} = \alpha_0 + \beta_1 S_{i,j,t-1,1} + \beta_2 S_{i,j,t-1,2} + \beta_3 S_{i,j,t-1,3} + \ldots + \beta_T S_{i,j,t-1,N} + \gamma \mathbf{X_{i,j,t-1}} + \varepsilon_{i,j,t},$$

Intuition

$$\tilde{r}_{i,t} = \beta_i \tilde{F}_t + \tilde{\epsilon}_{i,t}$$

$$Cov[\tilde{r}_{i,t}, \tilde{r}_{j,t}] = \beta_i \beta_j$$

Data - Investors

- Investors

Daily return data and stock price co-movement to uncover signals about investor information

"*Prices aggregate these diverse pieces of information and ultimately reflect an accurate assessment of firm value.*"

- Bond, Edmans, and Goldstein (2012)

Comment

Central Hypothesis

$$\tilde{r}_{i,t} = \beta_i \tilde{F}_t + \tilde{\epsilon}_{i,t}$$

$$Cov[\tilde{r}_{i,t}, \tilde{r}_{j,t}] = \beta_i \beta_j$$

$$Covariance_{i,j,t} = \alpha_0 + \beta_1 S_{i,j,t-1,1} + \beta_2 S_{i,j,t-1,2} + \beta_3 S_{i,j,t-1,3} + \ldots + \beta_T S_{i,j,t-1,N} + \gamma \mathbf{X_{i,j,t-1}} + \varepsilon_{i,j,t},$$

Data – Publicly Traded Financial Institution

Criteria

- In the CRSP(Center for Research in Security Prices) and Compustat databases
- Have a SIC(Standard Industrial Classification) code in the range 6000 to 6199
- In the metaHeuristica database (LDA&SVA) and
- 10-Ks with a non-zero number of paragraphs that discusses risks

Data – Publicly Traded Financial Institution, cont.

Example:

Item 1A. Risk Factors

In the course of conducting our business operations, we are exposed to a variety of risks, some of which are inherent in the financial services industry and others of which are more specific to our own businesses. The discussion below addresses the most significant factors, of which we are currently aware, that could affect our businesses, results of operations and financial condition. Additional factors that could affect our businesses, results of operations and financial condition are discussed in Forward-looking Statements in the MD&A on page 21. However, other factors not discussed below or elsewhere in this Annual Report on Form 10-K could also adversely affect our businesses, results of operations and financial condition. Therefore, the risk factors below should not be considered a complete list of potential risks that we may face.

Any risk factor described in this Annual Report on Form 10-K or in any of our other SEC filings could by itself, or together with other factors, materially adversely affect our liquidity, competitive position, business, reputation, results of operations, capital position or financial condition, including by materially increasing our expenses or decreasing our revenues, which could result in material losses.

General Economic and Market Conditions Risk

Our businesses and results of operations may be adversely affected by the U.S. and international financial markets, U.S. and non-U.S. fiscal and monetary policy, and economic conditions generally.

Our businesses and results of operations are affected by the financial markets and general economic, market, political and social conditions in the U.S. and abroad, including factors such as the level and volatility of short-term and long-term interest rates, inflation, home prices, unemployment and under-employment levels, bankruptcies, household income, consumer spending, fluctuations in both debt and equity capital markets and currencies, liquidity of the global financial markets, the availability and cost of capital and credit, investor sentiment and confidence in the financial markets, political risks, the sustainability of economic growth in the U.S., Europe, China and Japan, and economic, market, political and social conditions in several larger emerging market countries. Continued economic challenges include under-employment, declines in energy prices, the ongoing low interest rate environment, restrained growth in consumer demand, the strengthening of the U.S. Dollar versus other currencies, and continued risk in the consumer and commercial real estate markets. Deterioration of any of these conditions could adversely affect our consumer and commercial businesses, our securities and derivatives portfolios, our level of charge-offs and provision for credit losses, the carrying value of our deferred tax assets, our capital levels and liquidity, and our results of operations. For instance, the recent sharp drop in oil prices, while likely a net positive for the U.S. economy, may also add stress to select regional markets that are energy industry dependent and may negatively impact certain commercial and consumer loan portfolios.

Our businesses and results of operations are also affected by domestic and international fiscal and monetary policy. For example, the recent rate increase by the Federal Reserve in the U.S. and continued easing at many central banks internationally impact our cost of funds for lending, investing and capital raising activities and the return we earn on loans and investments. Central bank actions can also affect the value of financial instruments and other assets, such as debt securities and mortgage servicing rights (MSRs), and their policies can affect our borrowers, potentially increasing the risk that they may fail to repay their loans. Changes in domestic and international fiscal and monetary policies are beyond our control and difficult to predict but could have an adverse impact on our capital requirements and the costs of running our business.

For more information about economic conditions and challenges discussed above, see Executive Summary – 2015 Economic and Business Environment in the MD&A on page 22.

Liquidity Risk

Liquidity Risk is the Potential Inability to Meet Expected or Unexpected Liquidity Needs While Continuing to Support our Business and Customer Needs Under a Range of Economic Conditions.

If we are unable to access the capital markets, continue to maintain deposits, or our borrowing costs increase, our liquidity and competitive position will be negatively affected.

Liquidity is essential to our businesses. We fund our assets primarily with globally sourced deposits in our bank entities, as well as secured and unsecured liabilities transacted in the capital markets. We rely on certain secured funding sources, such as repo markets, which are typically short-term and credit-sensitive in nature. We also engage in asset securitization transactions, including with the government-sponsored enterprises (GSEs), to fund consumer lending activities. Our liquidity could be adversely affected by any inability to access the capital markets; illiquidity or volatility in the capital markets; unforeseen outflows of cash, including customer deposits, funding for commitments and contingencies; increased regulatory liquidity requirements for our U.S. or international banks and their nonbank subsidiaries; or negative perceptions about our short- or long-term business prospects, including downgrades of our credit ratings. Several of these factors may arise due to circumstances beyond our control, such as a general market disruption, negative views about the financial services industry generally, changes in the regulatory environment, actions by credit rating agencies or an operational problem that affects third parties or us.

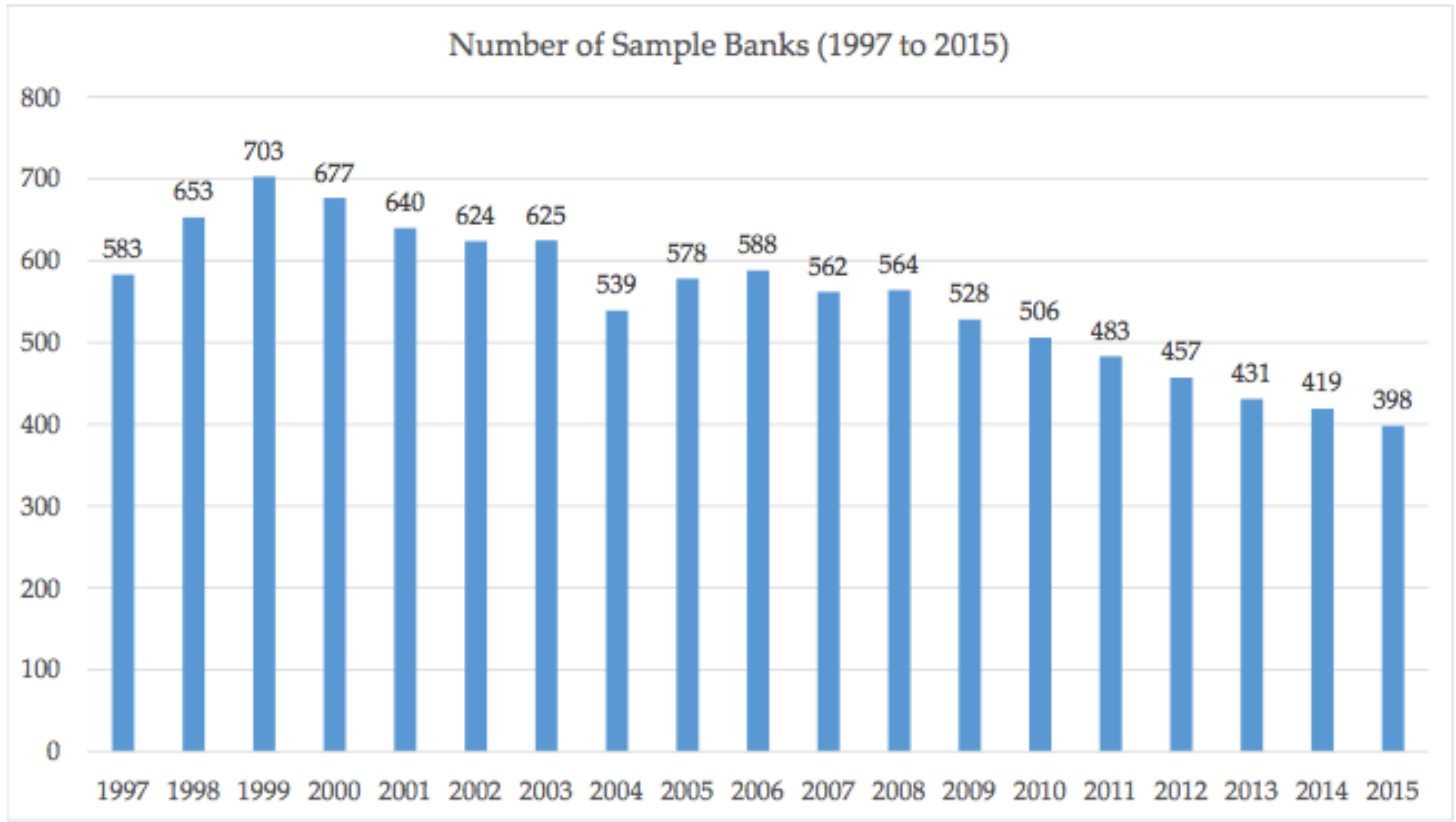
Our cost of obtaining funding is directly related to prevailing market interest rates and to our credit spreads. Credit spreads are the amount in excess of the interest rate of U.S. Treasury securities, or other benchmark securities, of a similar maturity that we need to pay to our funding providers. Increases in interest rates and our credit spreads can increase the cost of our funding. Changes in our credit spreads are market-driven and may be influenced by market perceptions of our creditworthiness. Changes to interest rates and our credit spreads occur continuously and may be unpredictable and highly volatile.

For more information about our liquidity position and other liquidity matters, including credit ratings and outlooks and the policies and procedures we use to manage our liquidity risks, see Liquidity Risk in the MD&A on page 60.

Adverse changes to our credit ratings from the major credit rating agencies could significantly limit our access to funding or the capital markets, increase our borrowing costs, or trigger additional collateral or funding requirements.

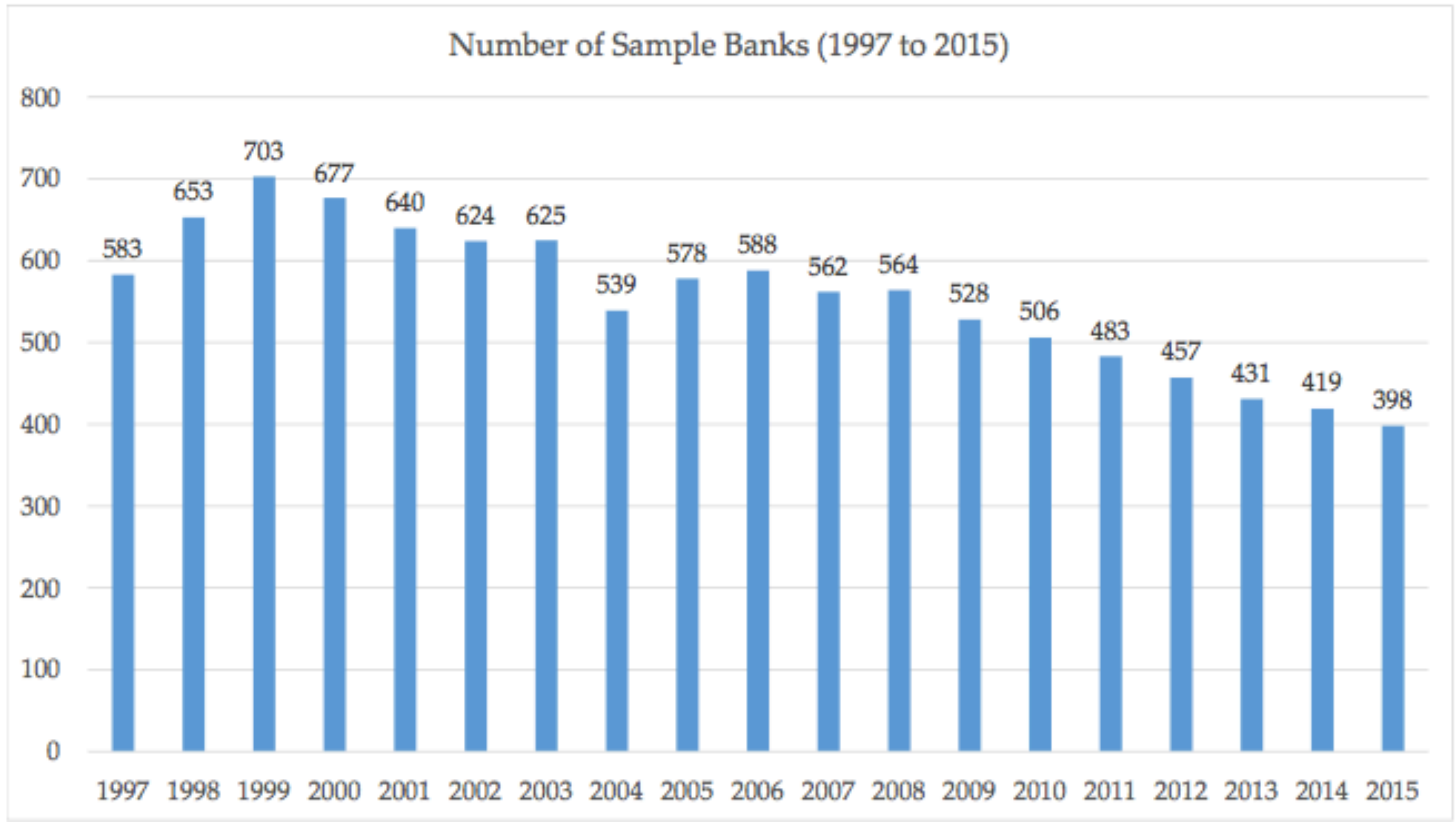
Our borrowing costs and ability to raise funds are directly impacted by our credit ratings. In addition, credit ratings may be important to customers or counterparties when we compete in

Data – Publicly Traded Financial Institution, cont.



- 10,588 bank-year observations from 1997 to 2015
- 587 publicly traded banks per year

Comment on the Data



- first 2 years
- onset of the financial crisis

Comment on the Data, cont.

- Campbell, Chen, Dhaliwal, Lu, and Steele (2014)

A positive association between specific risk factor disclosures in the 10-K and measures of risk exposures.

Text Analytic Methods Overview

- Pre-built modules

  Text processing software provided by metaHeuristica

- Latent Dirichlet Allocation (LDA)

LDA identifies a small number of verbal themes that best explain the variation in text across sample. (Similar to principal components analysis for numerical data)

LDA identifies a set of broad economic risks that occur with high prevalence

- Semantic vector analysis (SVA)

A method based on neural networks that detects semantic relatedness, to convert the output from LDA into a set of interpretable risk factors that are stable over time.

SVA maps these broad risks to specific risks that are directly interpretable

Text analytic methods - Latent Dirichlet Allocation (LDA)

- Output of metaHeuristica

The full set of paragraphs that contain discussions of risk factors for all banks in our sample from 1997 to 2015.

LDA is a dimensionality-reduction algorithm.

LDA requires only one input: the number of topics T to be generated.

- Output of LDA
  - Distribution of systematically important topics discussed by each bank in each year - 25

Text analytic methods-Latent Dirichlet Allocation (LDA), cont

LDA output in 2006

Clear interpretation

Text analytic methods-Latent Dirichlet Allocation (LDA), cont

LDA output in 2006

Unclear interpretation

Text analytic methods - Semantic Vector Analysis (SVA)

- The SVA algorithm is provided as part of the metaHeuristica program
- This is a probabilistic approach used to uncover the **semantics** of natural language

"destination" vs "last stop"

Text analytic methods - Semantic Vector Analysis (SVA), cont.

SVA is implemented in two stages:
(1) Model fitting stage: “word-to-vec” mapping two-layer neural network

(2) Mapping stage
Specific risks are mapped to vocabularies that best represent each given risk semantically

Text analytic methods - Semantic Vector Analysis (SVA), cont.

	Real Estate	
Row	Word	Cosine Dist
1	real	0.7875
2	estate	0.7875
3	foreclosure	0.4898
4	property	0.4619
5	personal	0.4563
6	physical possession	0.4539
7	foreclosed real	0.4503
8	foreclosed	0.4423

Cosine similarity:
Between the vocabulary list associated with each SVA risk theme, and the raw text of each bank's overall risk factor disclosure.

$W_{i,t}$ represents the firm i's risk factor disclosure
$T_{k,t}$ represents a vector for theme k

Firm i's loading on semantic theme k in year t as $S_{i,k,t}$ as the normalized cosine distance:

$$S_{i,k,t} = \frac{W_{i,t}}{||W_{i,t}||} \cdot \frac{T_{k,t}}{||T_{k,t}||}$$

Determination of Emerging Risks

- Central hypothesis

As investors produce information on an emerging risk that both i and j are exposed to, the stock return covariance of the pair will become abnormally high.

$$S_{i,j,l} = S_{i,l}\ S_{j,l}$$

$$Covariance_{i,j,t} = \alpha_0 + \beta_1 S_{i,j,t-1,1} + \beta_2 S_{i,j,t-1,2} + \beta_3 S_{i,j,t-1,3} + \ldots + \beta_T S_{i,j,t-1,N} + \gamma \mathbf{X_{i,j,t-1}} + \varepsilon_{i,j,t},$$

$$Covariance_{i,j,t} = \alpha_0 + \gamma \mathbf{X_{i,j,t-1}} + \varepsilon_{i,j,t}.$$

Results

Summary Statistics

Variable	Mean	Std. Dev.	Minimum	Median	Maximum	# Obs.
Panel D: Time-Series Variables						
VIX Level	20.914	7.522	11.190	20.047	51.723	76
Avg Pair Covariance	1.062	2.015	0.150	0.444	12.704	76
Avg Std Dev Monthly Returns	0.154	0.049	0.095	0.136	0.307	76
Avg Std Dev Monthly Returns (Fin. Only)	0.091	0.031	0.051	0.083	0.170	76
Econ Policy Uncertainty	110.595	33.609	63.118	103.840	215.891	76
Econ Policy Uncertainty (News Only)	116.749	38.424	52.089	111.348	235.084	76
Cov Model R^2 (Acct Vars Only)	0.081	0.063	0.005	0.056	0.279	76
Cov Model R^2 (Text Vars Only)	0.015	0.011	0.001	0.013	0.037	76

3 Models

- Static Model
- Dynamic Model
- Drill-down Model

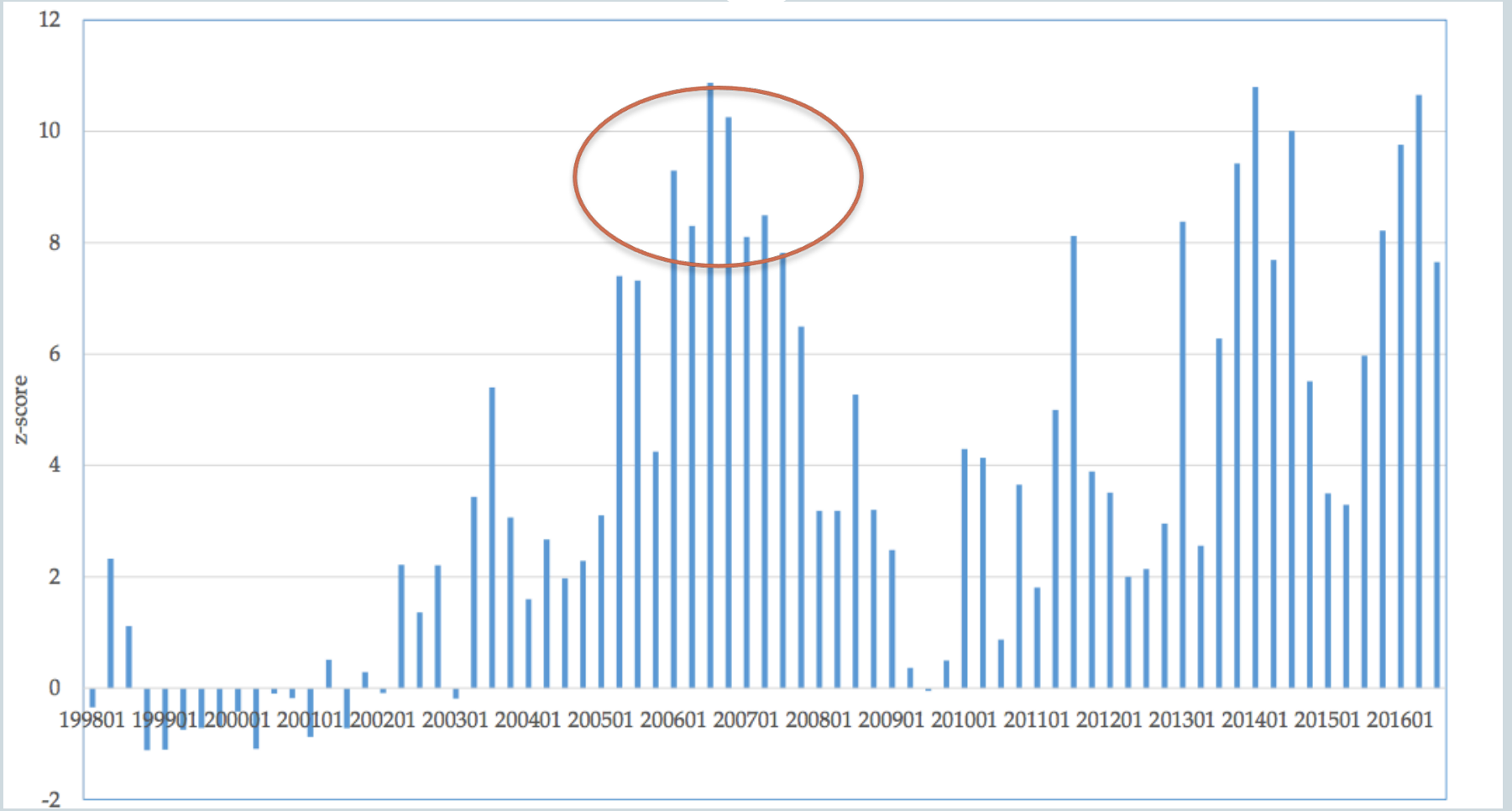
Model 1: Static SVA Model

- The “static model” specifies 31 semantic risk themes using guidance from the LDA output and the academic literature
- This allows us to assess specific risks that are generally agreed to be fundamental to the banking industry.

Model 1: Static Model, cont.

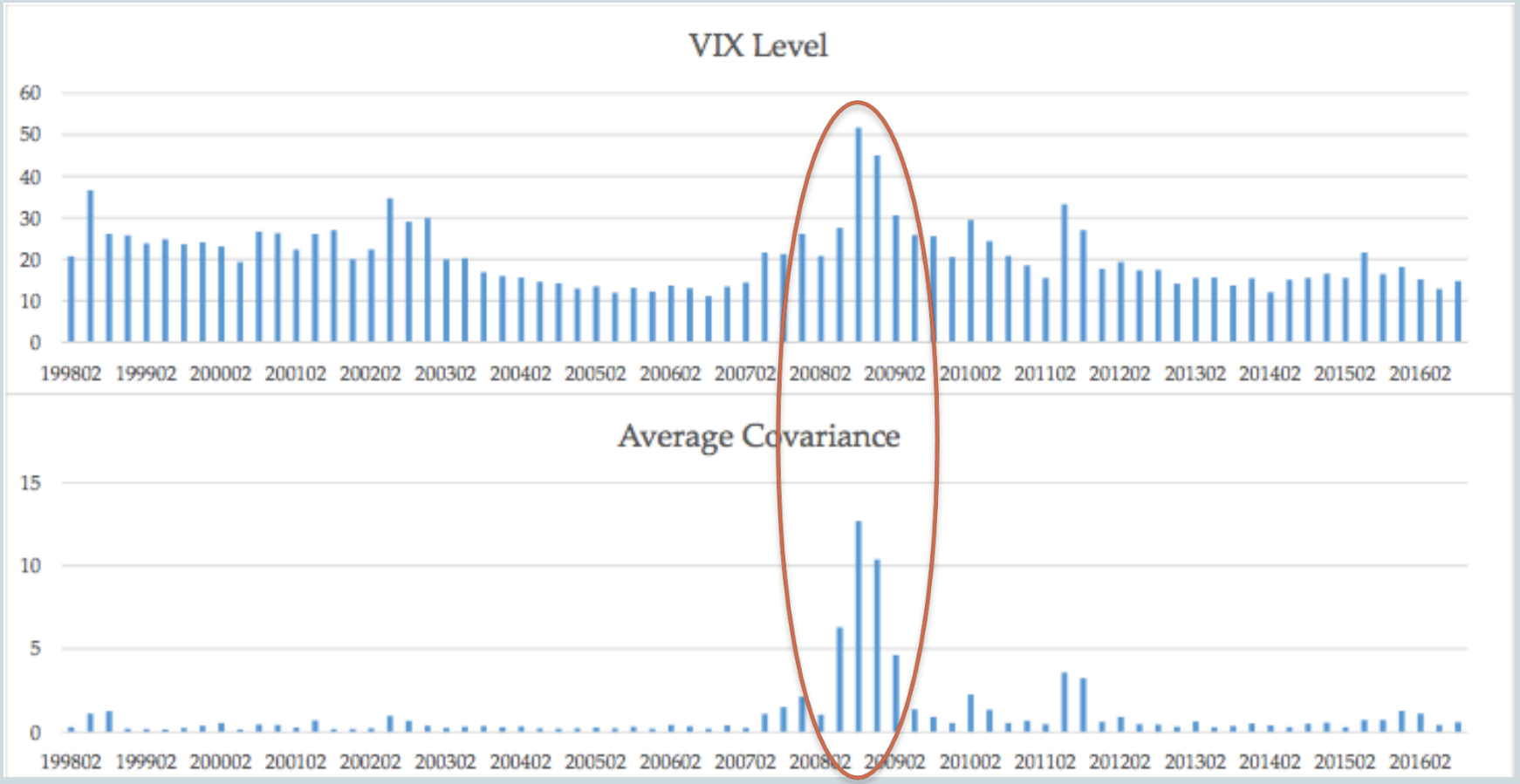
- In each quarter, we compute the marginal adjusted R^2 contribution from the 31 risks over and above a set of control variables for bank fundamentals.
- We then compute a **baseline** mean and standard deviation(1998 to 2003), and compute a z score(2004 to 2015) based on how many standard deviations the marginal adjusted R^2 is in a given quarter from the baseline mean.
- A high z score indicates the likely presence of emerging risks.

Model 1: Static Model, cont.



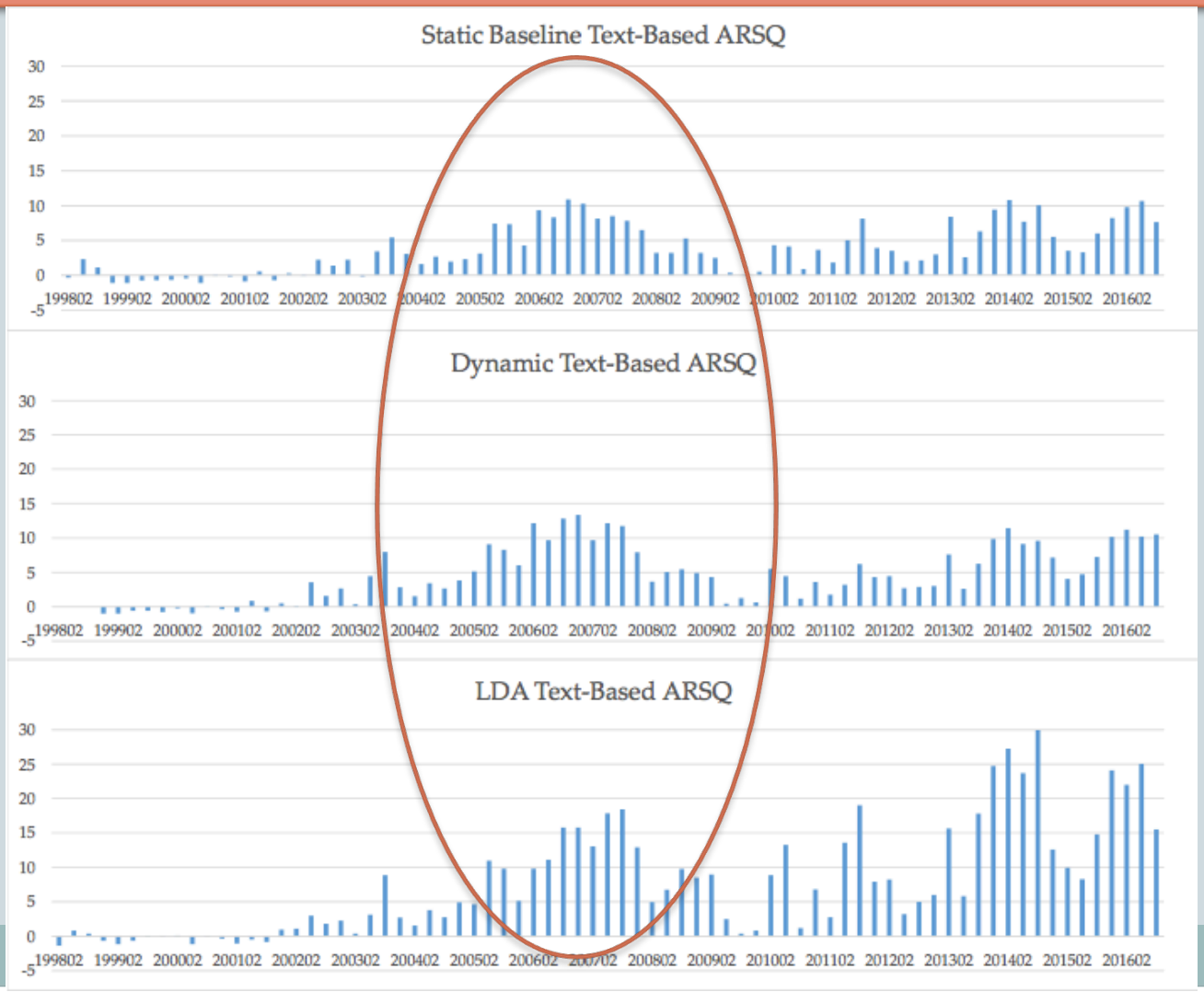
Model 1: Static Model, cont.

Emerging Risks Comparison



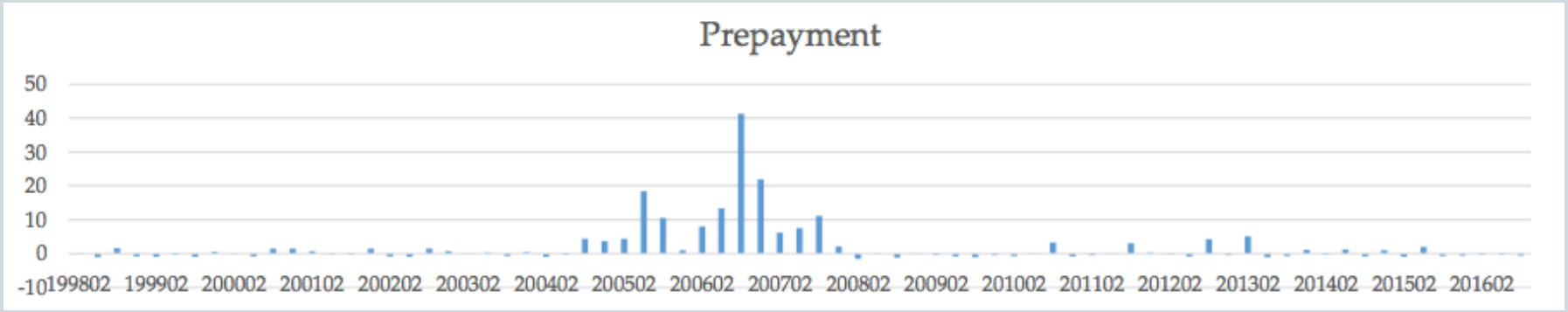
Model 1: Static Model, cont.

Emerging Risks Comparison



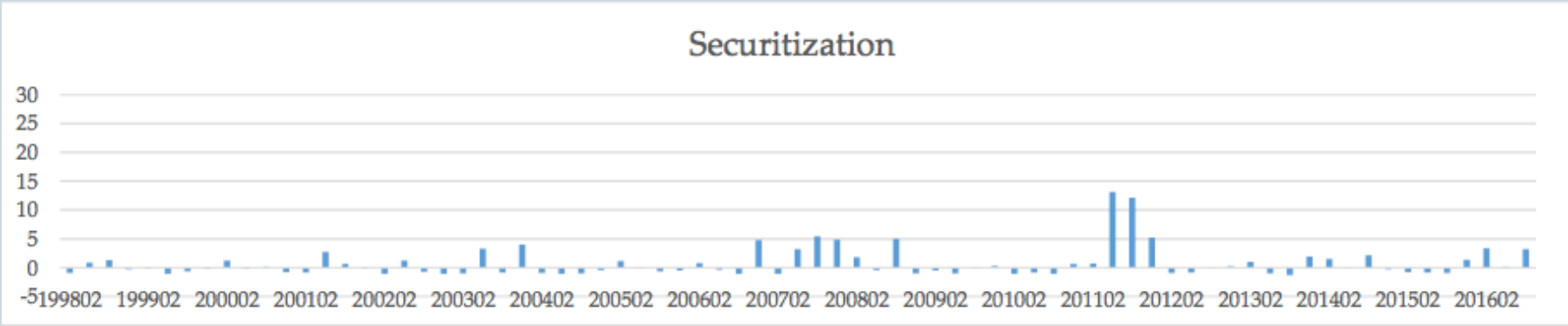
Model 1: Static Model, cont.

Time Series of Individual Emerging Risks



Model 1: Static Model, cont.

Time Series of Individual Emerging Risks



Model 1: Static Model, cont.

- Bank-specific exposures to emerging risks in each period

Emerging Risk Exposure:

Average predicted covariance bank i has with all other banks j using the main covariance model

Individual banks with greater Emerging Risk Exposure

- More negative stock returns
- More likely to fail
- Higher firm-specific volatility

Model 2: Dynamic Model

- The “dynamic model” automates the identification of important risks separately in each year.
  - LDA model that identifies 25 topics
  - Extract the 25 most probable commongrams for each topic
- This model takes no input from the researcher and can identify previously unknown or neglected risks.

Comment on Model 2: Dynamic Model

- Reveals new risks

  "We find a significant relationship between the covariance of bank returns and the theme "Puerto Rico" in 2014, well before the territory declares bankruptcy "

- Unlike the static model, the researcher **excludes** candidate **bigrams** that are not interpretable.

Model 3: Drill-down Model

- The “drill-down model” allows the flexibility to examine the channels driving a specific risk in greater detail.

Example:

Decompose the broad topic of **real estate risk** into the sub-themes such as subprime, mortgage-backed, HELOC, and foreclosure and we find that risks pertaining to many of these became elevated before the crisis period.

Key Contributions of the Model

- The methodology both dynamically measures the aggregate risk exposure of the banking industry as a whole, and also the specific underlying sources of risk.

- Models can be used in real time.