News and narratives in financial systems: exploiting big data for systemic risk assessment


*Bank of England, **University College London

Presentation to 3rd Annual MIT-GCFP Conference on “Causes of and Policy Responses to the U.S. Financial Crisis: What Do We Know Now that the Dust Has Settled?”

Cambridge (MA), 29 September 2016

The views expressed in this presentation are those of the speaker and should not be thought to represent those of the Bank of England, Monetary Policy Committee members, or Financial Policy Committee members
Motivation

• Policy: when to deploy time-varying macroprudential policies?
  – need to identify emerging exuberance

• Theory: narratives and emotions as key drivers of economic and financial activity (eg Keynes, 1936; Akerlof & Shiller, 2009):
  – within the context of Knightian uncertainty, agents act by gaining conviction through the use of narratives – such conviction narratives (Chong & Tuckett, 2014) must have emotional support: excitement about gain, suppressing doubt and anxiety about loss
  – narratives can spread ‘systemically’ via social networks or media (Shiller, 2000) and precipitate ‘consensus’

• Empirical: growing text-based analysis linked to sentiment:
  – economic policy uncertainty (Baker et al, 2016) and asset prices (eg Loughran and McDonald, 2011; Tetlock, 2007, 2011; Soo, 2013)
This Paper

• Exploits big text data to investigate the effect of narratives on the economy and financial system

• Aim to get a quantitative lens on market news and intelligence for systemic risk assessment:
  – can text-based measures of shifts in the relative balance between excitement and anxiety be useful as an early indicator?
  – can we gauge the extent of consensus to yield further insight?

• We provide evidence of increasing narrative consensus high in excitement and lacking anxiety prior to the crisis

• Key contributions:
  – theoretical filter to text-based analysis
  – focus on systemic risk: exploring the role of market intelligence
  – financial system data sources, including an internal BoE source
Outline

• Data

• Text-based analysis of sentiment

• Gauging consensus in narratives

• Summary and further work
## Data

<table>
<thead>
<tr>
<th>Data</th>
<th>Range</th>
<th>Description</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal Market Commentary</td>
<td>January 2000 through July 2010</td>
<td>Daily comments on market events</td>
<td>MCDAILY</td>
</tr>
<tr>
<td>Broker Circulars (Macro view)</td>
<td>January 2008 through June 2013</td>
<td>Low volume prior to June 2010. Primarily weekly economic research reports</td>
<td>BROKER</td>
</tr>
</tbody>
</table>
Relative Sentiment – Methodology

- Relative Sentiment Shifts
  - Theoretically motivated (and validated) word dictionaries are used
  - Ordinary English words

- Excitement/Anxiety word samples ~ 150 words each
  - Amaze, amazed, amazes, amazing, attract, attracted, attraction, etc.
  - Anxiety, anxious, avoid, avoids, bother, bothers, bothered, etc.

- Relative sentiment metric = (# excitement - # anxiety) / # characters
Sentiment – Results (Internal Market Commentary)
Biggest component of sentiment increase in mid-2000s is anxiety (red)
Largely correlated with RTRS (green) and BROKER (red)
Comparing with other metrics (1)

VIX (yellow)
MCDAILY (black)

2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

Dotcom High
Dotcom Low
BNP Paribas
Northern Rock
Case-Shiller peak
Bear Stearns
Lehman/AIG
First Greek Rescue

Anxiety
Excitement

BANK OF ENGLAND
UCL
Comparing with other metrics (2)

Correlations between relative sentiment series and common measures of sentiment, ignoring signs (-1 is t-1)

<table>
<thead>
<tr>
<th></th>
<th>MCD</th>
<th>RTRS</th>
<th>BRO</th>
<th>VIX</th>
<th>MCI</th>
<th>EPU</th>
<th>BoEU</th>
<th>CDS</th>
<th>PMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCD</td>
<td>1</td>
<td>0.59</td>
<td>-</td>
<td>0.65</td>
<td>0.26</td>
<td>0.43</td>
<td>0.54</td>
<td>0.67</td>
<td>0.38</td>
</tr>
<tr>
<td>RTRS</td>
<td>-</td>
<td>1</td>
<td>0.71</td>
<td>0.37</td>
<td>0.54</td>
<td>0.61</td>
<td>0.52</td>
<td>0.71</td>
<td>0.51</td>
</tr>
<tr>
<td>BRO</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>0.57</td>
<td>0.66</td>
<td>0.06</td>
<td>0.60</td>
<td>0.23</td>
<td>0.42</td>
</tr>
</tbody>
</table>

MCD (-1) - - - - 0.65 0.27 0.41 0.61 0.63 0.43
RTRS (-1) - - - - 0.37 0.58 0.63 0.67 0.69 0.57
BRO (-1) - - - - 0.65 0.87 0.01 0.76 0.22 0.42

Note: *p<0.05; **p<0.01
### Granger causality

Wald test p-values of Granger-causality from the relative sentiment shift series

<table>
<thead>
<tr>
<th>RSS Series</th>
<th>MCI</th>
<th>VIX</th>
<th>BoEU</th>
<th>EPU</th>
<th>CDS</th>
<th>PMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTRS EXC-ANX</td>
<td>0.005**</td>
<td>0.28</td>
<td>4e-06**</td>
<td>0.3</td>
<td></td>
<td>0.0002**</td>
</tr>
<tr>
<td>RTRS EXC</td>
<td>0.032*</td>
<td>0.044*</td>
<td>0.0013**</td>
<td>0.03*</td>
<td>0.05*</td>
<td>0.05*</td>
</tr>
<tr>
<td>RTRS ANX</td>
<td>0.003**</td>
<td>0.56</td>
<td>7e-05**</td>
<td>0.1</td>
<td></td>
<td>0.0004**</td>
</tr>
<tr>
<td>MCDAILY EXC-ANX</td>
<td>0.5</td>
<td>0.09</td>
<td>6e-05**</td>
<td>0.05*</td>
<td>0.09</td>
<td>0.06</td>
</tr>
<tr>
<td>MCDAILY EXC</td>
<td>0.8</td>
<td>0.44</td>
<td>0.13</td>
<td>0.85</td>
<td>0.57</td>
<td>0.57</td>
</tr>
<tr>
<td>MCDAILY ANX</td>
<td>0.8</td>
<td>0.38</td>
<td>0.001**</td>
<td>0.06</td>
<td>0.12</td>
<td>0.33</td>
</tr>
<tr>
<td>BROKER EXC-ANX</td>
<td>2e-11**</td>
<td>0.18</td>
<td>0.92</td>
<td>0.6</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>BROKER EXC</td>
<td>0.022*</td>
<td>0.84</td>
<td>0.77</td>
<td>0.43</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>BROKER ANX</td>
<td>3e-05**</td>
<td>0.12</td>
<td>0.72</td>
<td>0.68</td>
<td>0.03*</td>
<td></td>
</tr>
</tbody>
</table>

Note: *p<0.05; **p<0.01
Impact on the Wider Economy (1)

• Explore the impact on economic activity using a simple VAR of the UK economy from 1996-2015:

\[
\begin{bmatrix}
rss_t \\
GDP_t \\
L_t \\
CPI_t \\
r_t \\
credit_t
\end{bmatrix}
= A_1
\begin{bmatrix}
rss_{t-1} \\
GDP_{t-1} \\
L_{t-1} \\
CPI_{t-1} \\
r_{t-1} \\
credit_{t-1}
\end{bmatrix}
+ A_2
\begin{bmatrix}
rss_{t-2} \\
GDP_{t-2} \\
L_{t-2} \\
CPI_{t-2} \\
r_{t-2} \\
credit_{t-2}
\end{bmatrix}
+ \varepsilon_t
\]

- \(rss\): quarterly relative sentiment shift series for the UK (RTRS),
- \(GDP\): quarterly level of GDP,
- \(L\): quarterly level of employment in hours worked,
- \(CPI\): seasonally adjusted level of the consumer price index,
- \(r\): level of Bank Rate
- \(credit\): an indicator of credit conditions
Impact on the Wider Economy (2)

- VAR structure follows Haddow et al (2013) estimating the impact of *uncertainty* on the UK economy, in which they showed uncertainty had a significant negative impact (left)
- Relative sentiment has a similar but opposite impact (right) – however, not significant
Impact on the Wider Economy (3)

- Including both uncertainty and RSS we notice RSS impacts uncertainty but not vice versa, confirming Granger-causality
- Anxiety $\rightarrow$ “perceived uncertainty” $\rightarrow$ growth?
Measuring Consensus in Narratives

• We attempt to quantify ‘consensus’ in Reuters, by measuring
  – the number of narratives at a given moment
  – the ‘size’ of each such narrative

• Analyse the entropy (dispersion) of the distribution of topics

• Automatic topic detection: cluster stories into distinct groups; each cluster treated as a topic

• Method yields a distribution of documents over topic clusters
  – eg 100 articles about sovereign debt, 300 about oil etc.
Reflects news content – does not explicitly model opposing views or capture market consensus, but market consensus may reflect what people read:

- “The history of speculative bubbles begins roughly with the advent of newspapers. [...] Although the news media… present themselves as detached observers of market events, they are themselves an integral part of these events. Significant market events generally occur only if there is similar thinking among large groups of people, and the news media are essential vehicles for the spread of ideas.” (Shiller, 2000)
Summary

• We have explored a measure of relative sentiment shifts and narrative consensus in a variety of financial market data sources

• Metrics seem useful for both high & low frequency developments
  – evidence from text-sources of pre-crisis belief in a new paradigm?

• Potential use for systemic risk assessment

• Demonstrate value of theoretical filter for big data text-based analysis
Further Work

• Development of sentiment series and consensus, including enhanced identification and visualisation of topics and narratives

• Macroeconomic applications, including to forecasting & nowcasting

• More generally, big data and text-based analysis are key elements of the BoE’s One Bank Research Agenda
  – Scottish referendum tweets Bank Underground blog
  – supervisory letters; online vacancy postings; Agency reports; gauging central bank credibility
Reserve slides
Topic modelling: UK property
Granger causality: reverse direction, to sentiment

<table>
<thead>
<tr>
<th>RSS Series</th>
<th>MCI</th>
<th>VIX</th>
<th>BoEU</th>
<th>EPU</th>
<th>CDS</th>
<th>PMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTRS EXC-ANX</td>
<td>0.29</td>
<td>0.093</td>
<td>0.022</td>
<td>0.57</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>RTRS EXC</td>
<td>0.39</td>
<td>0.22</td>
<td>0.038</td>
<td>0.62</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>RTRS ANX</td>
<td>0.03*</td>
<td>0.0013**</td>
<td>0.21</td>
<td>0.85</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>MCDAILY EXC-ANX</td>
<td>0.95</td>
<td>0.39</td>
<td>0.58</td>
<td>0.18</td>
<td>0.89</td>
<td>0.49</td>
</tr>
<tr>
<td>MCDAILY EXC</td>
<td>0.94</td>
<td>0.61</td>
<td>0.47</td>
<td>0.76</td>
<td>0.32</td>
<td>0.03*</td>
</tr>
<tr>
<td>MCDAILY ANX</td>
<td>0.92</td>
<td>0.52</td>
<td>0.69</td>
<td>0.08</td>
<td>0.98</td>
<td>0.95</td>
</tr>
<tr>
<td>BROKER EXC-ANX</td>
<td>0.94</td>
<td>0.16</td>
<td></td>
<td>0.73</td>
<td>0.97</td>
<td>0.41</td>
</tr>
<tr>
<td>BROKER EXC</td>
<td>0.22</td>
<td>0.084</td>
<td></td>
<td>0.72</td>
<td>0.001**</td>
<td>0.72</td>
</tr>
<tr>
<td>BROKER ANX</td>
<td>0.72</td>
<td>0.33</td>
<td></td>
<td>0.98</td>
<td>0.78</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Note: *p<0.05; **p<0.01
Predictions of Michigan Consumer Sentiment using BROKER RSS
Consensus – Methodology

• A machine learning approach to automatic topic detection
  – We cluster documents after mapping them to vectors
• Vectors created from word/document occurrence statistics
  – Create a word by document frequency matrix
  – Remove uninformative words - with extreme (low/high) frequency
• Each column (document) is a vector of word counts
  – But dimensionality is too high and vectors are sparse
  – Use principal component analysis (PCA) to reduce the dimension
  – In the new lower dimensional space, the latent factors are more like ‘topics’ - words and documents can correlate over the ‘topic factors’
• Assume that each document belongs to a single topic
• Can now cluster all documents and treat clusters as topics