The Value of News

Vegard Høghaug Larsen and Leif Anders Thorsrud

Centre for Applied Macro- and Petroleum Economics (CAMP)
BI Norwegian Business School

NHH-UiO Macro Workshop 2015
Norges Bank, June 2
Motivation: The News-Driven Business Cycle

Expectations matter for explaining economic fluctuations.

Changes in expectations the key ingredient in the news-driven business cycle view, see e.g., Beaudry and Portier (2006, AER), Barsky and Sims (2012, AER), Blanchard et al. (2013, AER), Angeletos and La’O (2013, Econometrica) and Schmitt-Grohe and Uribe (2012, Econometrica)

- What is news? Information received today, that contain information on outcomes that might materialize in the future
- In theory models: Agents in the economy receive signals about future economic developments, and then face a signal extraction problem: What is news and what is noise?
Motivation: Identifying news shock

- A problem with testing theories for news shocks is that the news is typically not observed.

- One solution is to link news to innovations in asset prices, since asset prices should reflect all available information, *Beaudry and Portier (2006, AER)*.

- But:
  - Asset prices change for a lot of reasons, likely containing both news and noise.
  - News about what? Productivity, future policy, energy prices, etc.
  - The content of the news shock should matter for the structural interpretation.
Our solution:

We generate a measure of news, by decomposing a Norwegian business newspaper according to the topics it writes about using a Latent Dirichlet Allocation (LDA) model introduced by Blei et al. (2003):

- The LDA allows us to reduce a vast dataset of news articles into a much smaller dataset containing daily topic frequencies
- The topic frequencies can be aggregated into a time series, which give us a time series for each topic

We hypothesize: The more a newspaper write about a topic the more likely it is that this topic reflects something of importance for the economy’s future needs and developments.
We then ...

1. Evaluate the predictive power of our generated news series. This predictive step is used to filter out the topics that doesn’t have any fundamental information from those that have.

2. Combine the news topics, based on the filtering in the previous step, into an aggregate news index.

3. Finally, investigate if our news index can explain economic fluctuations using a structural VAR framework commonly applied in the news literature. In contrast to previous work we:

   1. Can explore what type of news topics that are important for the aggregate economy, i.e., what constitutes a news shock?
   2. Separate news from noise shocks, i.e., if news shocks are innovations to our news index, what is unexpected innovations to asset prices? We call it noise.
Main results:

1. The LDA decomposition of the business newspaper deliver news topics that are easily classified and interpretable.

2. Many news topics predict macroeconomic outcomes. Especially noteworthy is the news topics ability to predict asset prices.

3. Using our aggregate news index, we find that news shocks generate large and persistent aggregate fluctuations and a permanent increase in productivity and output in line with the existing literature.
   
   ▶ Among the most important topics contributing to the shocks are *Funding, IT/Startup, Oil production, and Monetary policy*.

4. Noise shocks cause only an initial boom in output, and then a contraction.

5. Asset prices are almost fully explained by news and noise shocks together.
   
   ▶ Implication: Using asset prices to identify news shocks likely mixes together the effects of news and noise shocks.
Step 1: Topic extraction

Data:

- Todays Business (TB), Norway’s biggest business newspaper.
- Database is Retrievers Atekst: 25 years of newspaper articles (almost 500,000)

The LDA takes a set of articles as input and return two sets of distributions:

- One set of distributions over words, one distribution for each topic
- One set of distributions over topics, one distribution for each article
Topic 14: A distribution over words
Topic 14: A distribution over words

Funding
A representation of the whole corpus
The *Funding* news series

Note: Each individual topic time series is transformed to year-on-year growth rates and standardized. The Business cycle series is the standardized value of Hodrick Prescott (HP) filtered GDP, with the smoothing parameter set to 40000. Because the topics are not sign identified, the business cycle estimate is reported in absolute value.
Step 2: Identifying fundamental topics

How do we then link these frequency measures of topics to economic outcomes?

We run predictive regressions of the form $ARX_p$, where $X$ represents the topics – one $ARX$ for each topic. The predictive regressions are estimated as Latent Threshold Models (LTMs). The time varying parameter in the LTM allows us to identify the sign of the topic relative to an outcome variable. The threshold dynamics allow for topics to be important for parts of our estimation sample and also avoid overfitting. The predictive content of a model is evaluated based on the posterior odds ratio $PO_{ij} = \frac{p(y|M_i)}{p(y|M_j)} = \frac{p(y|ARX)}{p(y|AR)}$. 

Larsen and Thorsrud

The value of news

June 2, 2015
Step 2: Identifying fundamental topics

1. How do we then link these frequency measures of topics to economic outcomes?

2. We run predictive regressions of the form $ARX(p)$, where $X$ represents the topics – one ARX for each topic.
Step 2: Identifying fundamental topics

1. How do we then link these frequency measures of topics to economic outcomes?

2. We run predictive regressions of the form $ARX(p)$, where $X$ represents the topics – one $ARX$ for each topic.

3. The predictive regressions are estimated as Latent Threshold Models (LTMs):
   - The time varying parameter in the LTM allows us to identify the sign of the topic relative to an outcome variable.
   - The threshold dynamics allow for topics to be important for parts of our estimation sample and also avoid overfitting.
Step 2: Identifying fundamental topics

1. How do we then link these frequency measures of topics to economic outcomes?

2. We run predictive regressions of the form $ARX(p)$, where $X$ represents the topics – one ARX for each topic.

3. The predictive regressions are estimated as Latent Threshold Models (LTMs):
   - The time varying parameter in the LTM allows us to identify the sign of the topic relative to an outcome variable.
   - The threshold dynamics allow for topics to be important for parts of our estimation sample and also avoid overfitting.

4. The predictive content of a model is evaluated based on the posterior odds ratio:

$$PO_{ij} = \frac{p(y|M_i)}{p(y|M_j)} = \frac{p(y|ARX)}{p(y|AR)}.$$
Relative marginal likelihood – $\ln P_{ij} > 1$
Step 3: The aggregate news index

The aggregate news index is constructed as follows

\[ N_t = \sum_{i=1}^{T} w_i b_{i,t} n_{i,t-1}, \quad \text{where} \quad w_i = \frac{p(y|M_i)}{\sum_{i=1}^{T} p(y|M_i)} \]

We base our news index on asset prices: Not revised, commonly used, forward looking, and should contain all fundamental information
Step 3: The aggregate news index

The aggregate news index is constructed as follows

\[ N_t = \sum_{i=1}^{T} w_i b_{i,t} n_{i,t-1}, \text{ where } w_i = \frac{p(y|M_i)}{\sum_{i=1}^{T} p(y|M_i)} \]

We base our news index on asset prices: Not revised, commonly used, forward looking, and should contain all fundamental information
Step 4: News and the business cycle

![Graphs showing the relationship between news and the business cycle.](image-url)
Step 4: News and the business cycle

C

Y

TFP \ TFP^a

News C Y

News C Y

TFP Osebx

TFP Osebx

TFP BCI

TFP BCI

TFP News

TFP News

TFPA Osebx

TFPA Osebx

TFPA BCI

TFPA BCI

Larsen and Thorsrud

The value of news

June 2, 2015
What constitutes a news index shock

Average contribution to the aggregate news shocks
If news shocks are innovations to our news index, what is unexpected innovations to asset prices? We call it noise

We can uncover the structural shocks ($e^\text{NewsIndex}$ and $e^\text{Noise}$) by the following system (left hand side is the reduced form VAR residuals):

\[
\begin{bmatrix}
  u_t^{\text{NewsIndex}} \\
  u_t^{\text{TFP}} \\
  u_t^Y \\
  u_t^\pi \\
  u_t^{\text{R3MR}} \\
  u_t^{\text{OSEBX}} \\
\end{bmatrix}
=
\begin{bmatrix}
  a_{11} & 0 & 0 & 0 \\
  a_{21} & a_{22} & 0 & 0 \\
  \vdots & \vdots & \ddots & 0 \\
  a_{n1} & a_{n2} & \ldots & a_{nn} \\
\end{bmatrix}
\begin{bmatrix}
  e_t^{\text{NewsIndex}} \\
  \vdots \\
  e_t^{\text{Noise}} \\
\end{bmatrix}
\]
News index shock

TFP / $TFP^a$

$\pi$

$Y$

R3MR

The value of news

Larsen and Thorsrud

June 2, 2015
Noise shock

TFP / $TFP^a$

$\pi$

R3MR

Y

Noise Shock

Larsen and Thorsrud
Summing up

1. The LDA decomposition of the business newspaper deliver news topics that are easily classified and interpretable.

2. Many news topics predict macroeconomic outcomes. Especially noteworthy is the news topics ability to predict asset prices.

3. Using our aggregate news index, we find that news shocks generate large and persistent aggregate fluctuations and a permanent increase in productivity and output in line with the existing literature.
   - Among the most important topics contributing to the shocks are Funding, IT/Startup, Oil production, and Monetary policy.

4. Noise shocks cause only an initial boom in output, and then a contraction.

5. Asset prices are almost fully explained by news and noise shocks together.