Discussion of

Interpreating Big-Data in the Macro Economy: A Bayesian Mixed Frequency Estimator

by David Kohns and Arnab Bhattacharjee

Daniele Bianchi

School of Economics and Finance
Queen Mary University of London

2019 Modelling with big data and machine learning

Why should we care?

A decision maker $\mathcal{D}$ wants to predict in "real-time" the state of the economy y based on an information set $\boldsymbol{X}$.

A standard approach is

$$y_t = \boldsymbol{\beta}' \boldsymbol{X}_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \nu_t)$$

Why should we care?

A decision maker $\mathcal{D}$ wants to predict in "real-time" the state of the economy y based on an information set $\boldsymbol{X}$.

A standard approach is

$$y_t = \beta' \boldsymbol{X}_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \nu_t)$$

The information in $\boldsymbol{X}$ may be incomplete:

- Macroeconomic information is often delayed, i.e., publication lags, and only available on a lower frequency basis.
- Instead, alternative data, i.e., big data, may be more granular.

Why should we care?

A decision maker $\mathcal{D}$ wants to predict in "real-time" the state of the economy y based on an information set $\boldsymbol{X}$.

A standard approach is

$$y_t = \beta' \boldsymbol{X}_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \nu_t)$$

The information in $\boldsymbol{X}$ may be incomplete:

- Macroeconomic information is often delayed, i.e., publication lags, and only available on a lower frequency basis.
- Instead, alternative data, i.e., big data, may be more granular.

Question:

- Can we exploit big data to assess the state of the economy?... without losing economic intuition

Why should we care?

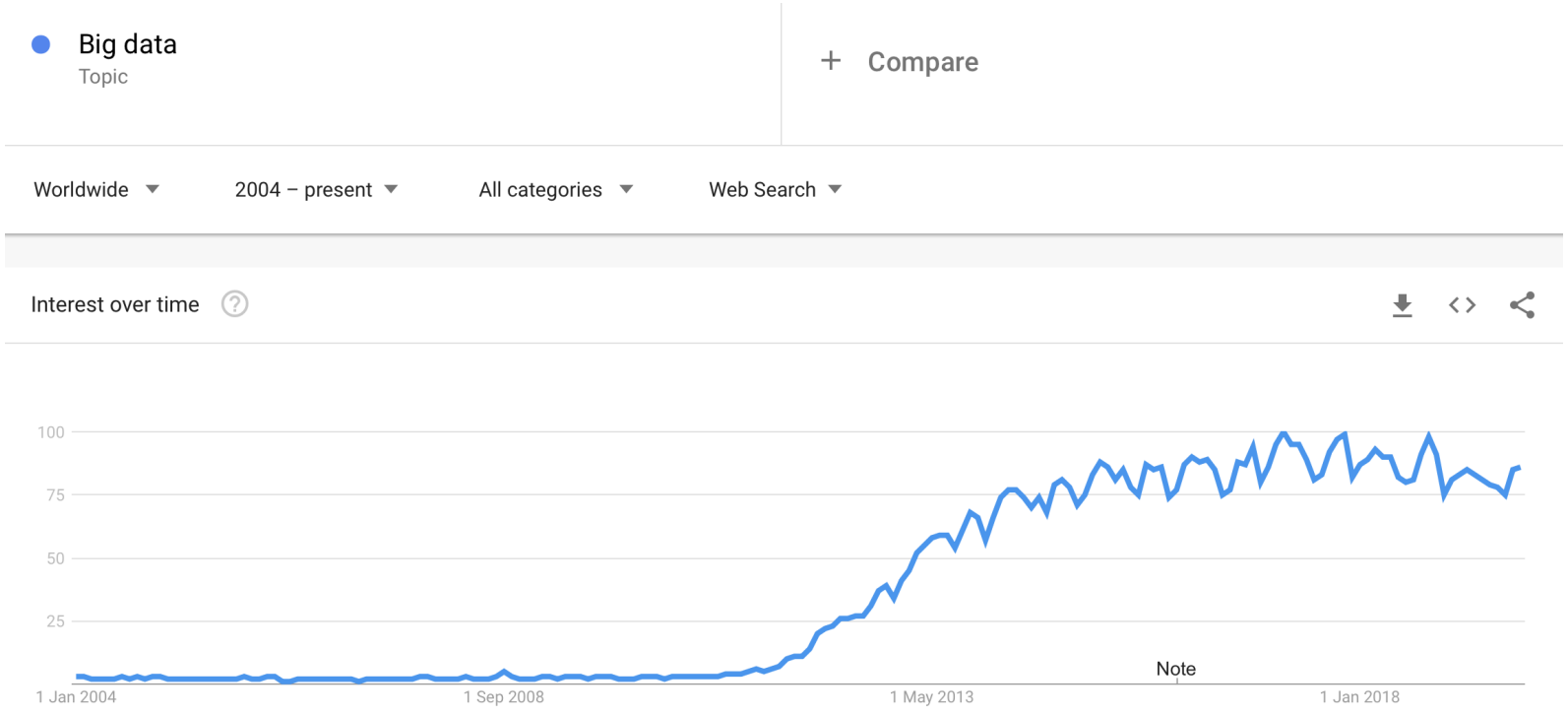
'Only by analyzing numerous times series, each of restricted significance, can business cycles be made to reveal themselves definitively enough to permit close observation.'

Burns and Mitchell (1946), *Measuring Business Cycles*

Why should we care?

'Only by analyzing numerous times series, each of restricted significance, can business cycles be made to reveal themselves definitively enough to permit close observation.'

Burns and Mitchell (1946), *Measuring Business Cycles*



This Paper

Few issues in combining Big Data with macroeconomic variables:

- Data are sampled at different frequencies.
- Asynchronous data publications.
- Outliers, trends, cycles and low signal-to-noise ratio.

This Paper

Few issues in combining Big Data with macroeconomic variables:

- Data are sampled at different frequencies.
- Asynchronous data publications.
- Outliers, trends, cycles and low signal-to-noise ratio.

The authors build on Scott and Varian (2014) to "nowcast" variables of interest

$$y_t = \mu_t + \boldsymbol{\beta}' \boldsymbol{x}_t + \epsilon_t,$$
$$\mu_t = \mu_{t-1} + \omega_t$$

i.e., linear local-level model + regression

This Paper

Few issues in combining Big Data with macroeconomic variables:

- Data are sampled at different frequencies.
- Asynchronous data publications.
- Outliers, trends, cycles and low signal-to-noise ratio.

The authors build on Scott and Varian (2014) to "nowcast" variables of interest

$$y_t = \mu_t + \boldsymbol{\beta}' \boldsymbol{x}_t + \epsilon_t,$$
$$\mu_t = \mu_{t-1} + \omega_t$$

i.e., linear local-level model + regression

This paper:
The predictors in $\boldsymbol{x}_t$ could be at $\neq$ frequencies (U-MIDAS, Foroni and Marcellino 2014).

My Comments:

Comment 1: Google Search Volume index.

Comment 2: Spike-and-Slab prior.

Comment 1: Google Search Volume Index

- Google search query data offers a quasi-real time view into the topics of greatest interest to the public.

Comment 1: Google Search Volume Index

- Google search query data offers a quasi-real time view into the topics of greatest interest to the public.
- This could be useful when searches are directly related to the variables of interest, for e.g.,
  - "file for unemployment" correlates with unemployment rate.
  - "Sales/shopping" correlates with monthly retail sales.

Comment 1: Google Search Volume Index

- Google search query data offers a quasi-real time view into the topics of greatest interest to the public.

- This could be useful when searches are directly related to the variables of interest, for e.g.,
  - "file for unemployment" correlates with unemployment rate.
  - "Sales/shopping" correlates with monthly retail sales.

Comments:

- What is the causal nexus between searching for "real GDP growth" and the actual growth rate?

Comment 1: Google Search Volume Index

- Google search query data offers a quasi-real time view into the topics of greatest interest to the public.

- This could be useful when searches are directly related to the variables of interest, for e.g.,
  - "file for unemployment" correlates with unemployment rate.
  - "Sales/shopping" correlates with monthly retail sales.

Comments:

- What is the causal nexus between searching for "real GDP growth" and the actual growth rate?
- Where is the true source of information?

Comment 1: Google Search Volume Index

- Google search query data offers a quasi-real time view into the topics of greatest interest to the public.

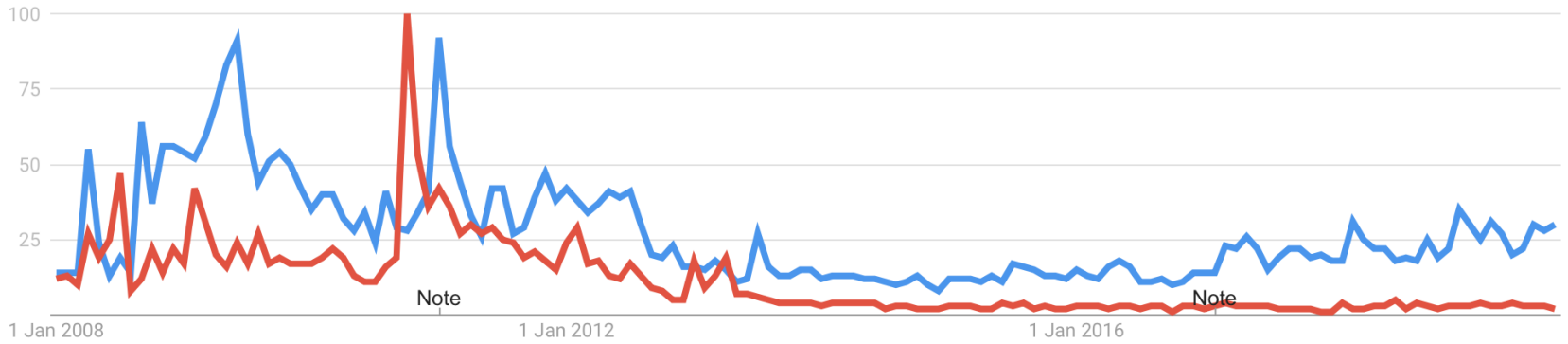
- This could be useful when searches are directly related to the variables of interest, for e.g.,
  - "file for unemployment" correlates with unemployment rate.
  - "Sales/shopping" correlates with monthly retail sales.

Comments:

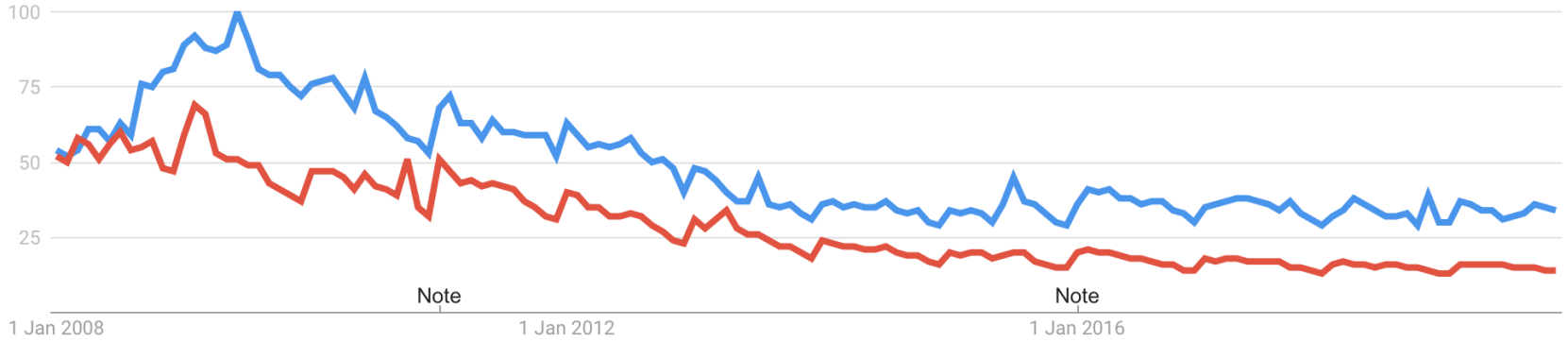
- What is the causal nexus between searching for "real GDP growth" and the actual growth rate?
- Where is the true source of information?
- Shall we look at "web search" or "news search"?

Comment 1: Google Search Volume Index

Figure: Foreclosure and Bankruptcy search



(a) Web Search



(b) News Search

Comment 2: Spike-and-Slab prior

Hierarchical mixture of normals
(George and McCulloch 1993), i.e.,

$$\beta_i|\gamma_i \sim (1-\gamma_i)\, N\left(0, \tau^2\right) + \gamma_i N\left(0, c^2\tau^2\right)$$

Comment 2: Spike-and-Slab prior

Hierarchical mixture of normals
(George and McCulloch 1993), i.e.,

$$\beta_i|\gamma_i \sim (1-\gamma_i)\, N\left(0, \tau^2\right) + \gamma_i N\left(0, c^2\tau^2\right)$$
$$\gamma_i|\pi \sim \mathrm{Bernoulli}\left(\pi\right), \qquad i = 1, \ldots, p$$

Comment 2: Spike-and-Slab prior

Hierarchical mixture of normals
(George and McCulloch 1993), i.e.,

$$\begin{aligned}
\beta_i|\gamma_i &\sim (1-\gamma_i)\, N\left(0, \tau^2\right) + \gamma_i N\left(0, c^2\tau^2\right) \\
\gamma_i|\pi &\sim \mathrm{Bernoulli}\left(\pi\right), \qquad i = 1, \ldots, p \\
\pi &\sim p\left(d\pi\right)
\end{aligned}$$

Comment 2: Spike-and-Slab prior

Hierarchical mixture of normals
(George and McCulloch 1993), i.e.,

$$\beta_i|\gamma_i \sim (1-\gamma_i)\, N\left(0, \tau^2\right) + \gamma_i N\left(0, c^2\tau^2\right)$$
$$\gamma_i|\pi \sim \mathrm{Bernoulli}\left(\pi\right), \qquad i = 1, \ldots, p$$
$$\pi \sim p\left(d\pi\right)$$

Pros:

- No assumption of sparsity;
  the parameter π is
  estimated from the data
- Easy to interpret
- Embeds model and
  parameter uncertainty.

Comment 2: Spike-and-Slab prior

Hierarchical mixture of normals
(George and McCulloch 1993), i.e.,

$$\beta_i|\gamma_i \sim (1-\gamma_i)\, N\left(0,\tau^2\right) + \gamma_i N\left(0, c^2\tau^2\right)$$
$$\gamma_i|\pi \sim \text{Bernoulli}\left(\pi\right), \qquad i = 1, \ldots, p$$
$$\pi \sim p\left(d\pi\right)$$

Pros:

- No assumption of sparsity; the parameter π is estimated from the data
- Easy to interpret
- Embeds model and parameter uncertainty.

Cons:

- crucial the choice of $N\left(0,\tau^2\right)$ vs t-student, inverse-Gamma, Laplace, etc.
- crucial the choice of $p\left(d\pi\right)$; uniform, Beta, etc.
- it is static.

Comment 2: Spike-and-Slab prior

In the paper:

1. π is fixed
2. Slab distribution $N\left(0, c^2\tau^2\right)$
3. Parameters are not recursively estimated.

Comment 2: Spike-and-Slab prior

That is:

1. Sparsity is assumed a priori (!?)
2. Slab distribution $N\left(0, c^2\tau^2\right)$
3. Parameters are not recursively estimated.

Comment 2: Spike-and-Slab prior

That is:

1. Sparsity is assumed a priori (!?)
2. Normal prior is not robust to large and noisy signals. (see Carvalho et al. 2010, Biometrika)
3. Parameters are not recursively estimated.

Comment 2: Spike-and-Slab prior

That is:

1. Sparsity is assumed a priori (!?)
2. The prior is not robust to large and noisy signals. (see Carvalho et al. 2010, Biometrika)
3. Sparsity is assumed constant OOS.

Comment 2: Spike-and-Slab prior

That is:

1. Sparsity is assumed a priori (!?)
2. The prior is not robust to large and noisy signals.
   (see Carvalho et al. 2010, Biometrika)
3. Sparsity is assumed constant OOS.

Comments:

– Why not fully exploit the hierarchical structure?
  e.g., $\pi \sim \mathrm{Beta}\,(a_0, b_0)$.

Comment 2: Spike-and-Slab prior

That is:

1. Sparsity is assumed a priori (!?)
2. The prior is not robust to large and noisy signals.
   (see Carvalho et al. 2010, Biometrika)
3. Sparsity is assumed constant OOS.

Comments:

- Why not fully exploit the hierarchical structure?
  e.g., $\pi \sim \mathrm{Beta}\,(a_0, b_0)$.
- Alternative mixtures could deliver different results, e.g., Frühwirth-Schnatter and Wagner (2010, JoE).

Comment 2: Spike-and-Slab prior

That is:

1. Sparsity is assumed a priori (!?)
2. The prior is not robust to large and noisy signals.
   (see Carvalho et al. 2010, Biometrika)
3. Sparsity is assumed constant OOS.

Comments:

- Why not fully exploit the hierarchical structure?
  e.g., $\pi \sim \mathrm{Beta}\,(a_0, b_0)$.
- Alternative mixtures could deliver different results, e.g., Frühwirth-Schnatter and Wagner (2010, JoE).
- Dynamic spike-and-slab
  (e.g., Korobilis and Koop 2019, Bianchi et al. 2019).

Conclusion

Very interesting paper, recommended reading!!