# Discussion: Advances in Machine Learning for Economic Forecasting

David Kohns & Arnab Bhattacharjee

Heriot-Watt University

November 05, 2019

# An Overview

- The general theme of the presentation was to show advances in existing ML methods to time series-contexts
- The two main problems tackled in these articles are (i) time variation in parameters and (ii) time dependence
- Without (ii): there is a single (or multiple) structural break(s)
- Without (i): existing tools suffice
- Contribution of the first part: how to deal with time variation in the informational content of unstructured data sources
- Contribution second part: extend non-parametric ML methods to time varying parameters through kernel smoothing

# Discussion: Assumption of Sparsity

- My research focus at Heriot-Watt lies in model selection consistency under uncertainty over the degree of sparsity in the data
- In a time series context the notion of sparsity becomes even more subtle as it may change over time: some (groups) of variables become more or less important
- Main comment: Understanding which variables shift in their relevance for the model can give new insights into business cycle modelling and forecasting, especially using new Big Data sources
- Main question: What do these methods presented today assume about the underlying degree of sparsity over time? Can these methods be used to yield interpretations as to when certain variables drive forecasts and why?

# Further Comments & Questions

- Discussion of the 1st part:
  - Justification for ordinary factor structure: do we assume that the informational content in the articles are dense? Have these results been compared to sparse PCA?
  - How computationally complex is the extraction of sentiments and factor estimation?
  - Have there been comparisons to web-scrapped data from social media and web-searches?
- Discussion of the 2nd part:
  - What is the interpretation behind the relative performance of the SVRs compared to NN and TV-SVRs and TV-NN?
  - Is there a method which the practitioner can use to decide which estimator to use?
  - Is there justification for smooth transitions in the data? Have the results been compared to the same ML methods based on discontinuous structural breaks?
  - What architecture was used for the NNs? What is the computational complexity of these methods?