BANK OF ENGLAND

KING'S College LONDON

Centre for Central Banking Studies

See it all: “One minute madness”

Rules & setting:

- Slides are on auto-forward.
- Each presenter as exactly 60 seconds to pitch her poster. Not a second longer!
- Posters are ordered and numbered as in the programme, such that the audience can easily refer to each in the open session & drinks reception directly afterwards.
- Enjoy the show!

0

Ready, steady, go …

Machine Learning, the Treasury Yield Curve and Recession Forecasting

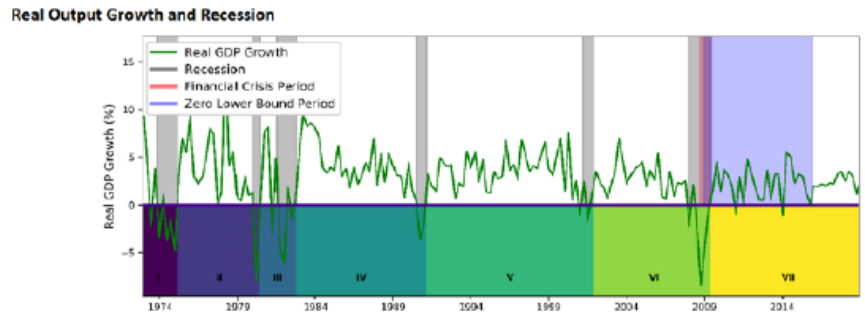
Michael Puglia and Adam Tucker, Federal Reserve Board of Governors

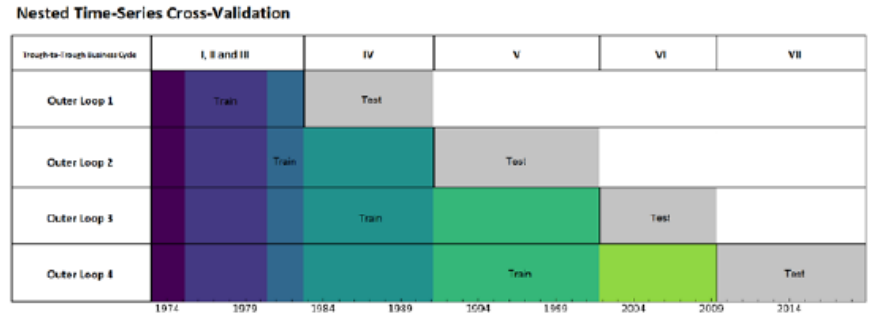
In a Nutshell

Applying ML to small macro/financial panels is difficult!

1. We propose a novel CV strategy to contend with the difficulties.
2. Our results suggest you might just use a probit regression next time you need to estimate recession probabilities from term spreads.
3. But then again, how seriously should we take any claim about ML classifier performance without a framework for statistical inference?
4. We use omnibus Cochrane's Q and post-hoc McNemar tests for model and algorithm selection.

1. Nested Time-Series CV



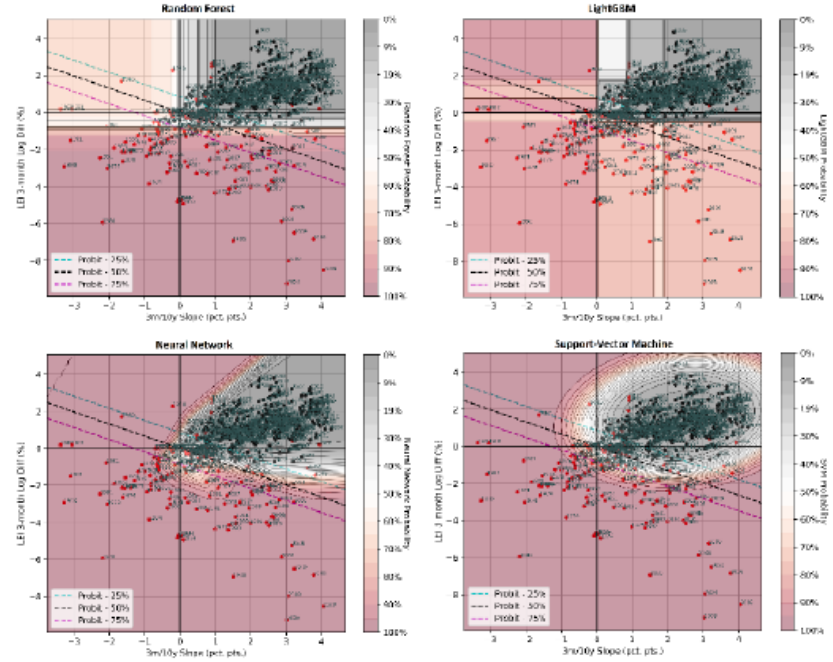


2. CV Results

Nested Time-Series CV
Rank
1. Probit
2. Neural Network
3. LightGBM
4. XGBoost
4. Random Forest
6. SVM

k-Folds CV
Rank
1. LightGBM
2. Random Forest
3. XGBoost
4. SVM
5. Neural Network
6. Probit

3. How Reliable Are Our Claims?



4. Let's Test That…

McNemar (1947) Within-Subjects Test for Classifiers

	c_2 correct	c_2 incorrect
c_1 correct	A	B
c_1 incorrect	C	D

$$\chi^2 = \frac{(|B - C| - 1)^2}{B + C}$$

Anomaly Detection in RTGS Systems

Performance Comparisons Between Shallow and Deep Neural Networks

L. Arciero, G. Bruno, S. Marchetti & J. Marcucci (Bank of Italy)

Aim

Monitor, measure and prevent materialization of systemic risk in the Italian financial system

Framework and Data

TARGET2-BdI: Payments exchanged within the Italian compo
TARGET2 (RTGS owned by Eurosystem)

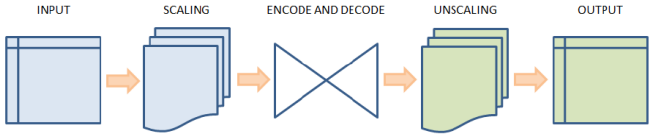
Focus on **top 20 banks** (each data point ⇒ 20×20 matrix)

583 working days @ 15-min. freq.: Jan. 2017-Apr. 2019

Methods

Deep Autoencoder ⇒

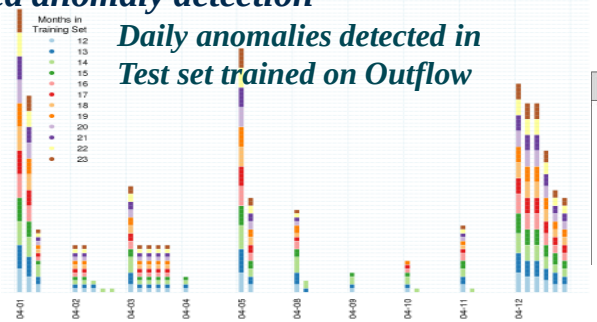
- ***knowledge representation*** and ***unsupervised anomaly detection***



Applications & Results

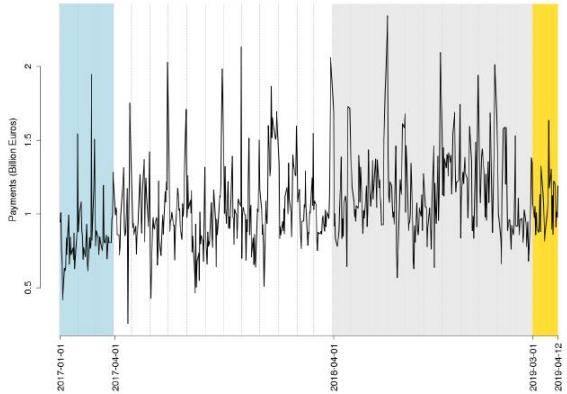
Real world case study ⇒

Simulated scenarios

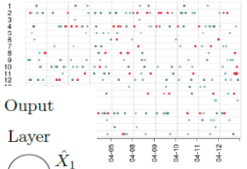


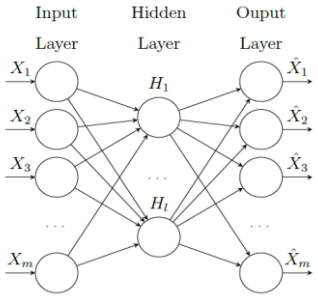
Daily anomalies detected in Test set trained on Outflow

Validation, Expanding Training & Test sets

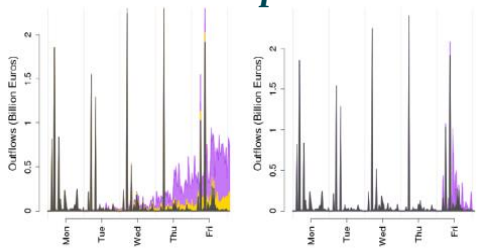


Contributions to RE

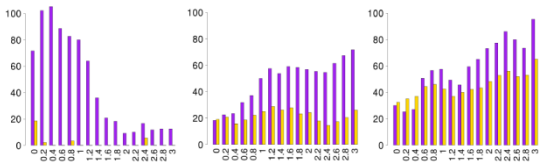




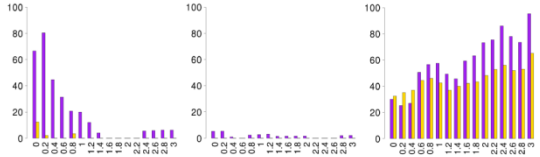
Gradual & abrupt anomalies



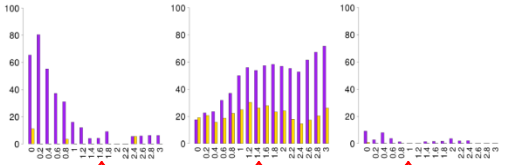
Additional anomalies: System



Additional anomalies: Outflow



Additional anomalies: Inflow



System ***Outflow*** ***Inflow***

Classification performance

	Extreme anomalies (%)	α	Sys	Out	In
Precision	50%	2	99.6%	94.5%	96.5%
		3	99.8%	95.6%	97.4%
	60%	2	99.7%	96.1%	97.7%
		3	99.9%	96.9%	98.3%
	40%	2	99.3%	92.4%	95.0%
		3	99.8%	[illegible]	[illegible]

$$RE(\mathbf{x}^t) \geq MRE^{\text{val}} + \alpha\sigma_{RE^{\text{val}}}$$

and $\alpha > 0$.

Bank of England Conference on «*Modelling with Big Data and Machine Learning*»
London, Nov. 4-5, 2019 – The usual disclaimers apply

Measuring and Understanding Patterns of Consumer Credit Stress
Gias Uddin and Brian Peterson, Bank of Canada

- **Data.** TransUnion consumer credit payment data for 25-29 million monthly active Canadian consumer loan records. Our analysis consists of all quarters between 2014 and 2018.
- **Motivation.** We find that at any given time 2-4.5% of the consumers show stress in at least one of their loans across the Canadian provinces. A loan is under stress, if its payment is behind at least 60 days. A consumer is under stress, if he shows stress in at least one of his loans at a time.
- **Detect Frequent Credit Stress Patterns.** We apply unsupervised machine learning technique, Frequent Itemset Mining to find frequent combinations of loans on which consumers show stress.
  1. The most frequent credit stress pattern is the combination of credit card and Line of credit, i.e., most number of stressed consumers show stress in both of these two loans together or one after another.
  2. We find that the probability of a stressed consumer to fall into mortgage stress increases, if the consumer shows stress in other loans in his loan profile.
- **Forecast Consumer Mortgage Default in Next Two Quarters.** We propose a model to forecast the mortgage default of stressed consumers by analyzing their loan profiles. A consumer is under mortgage default if he is behind mortgage payment for at least 90 days.
  1. The inclusion of credit stress information of other loans offers improvement in forecasting performance over a model that only uses consumer mortgage information.
  2. The best predictive variables are stress indicators based on consumer mortgage loans, followed by consumer credit risk rating. Among the other loans in consumer loan portfolios, stress information in credit cards, Line of Credits and HELOCs offer additional predictive power.
  3. The best model to predict potential defaults is Random Forest, followed by Lasso and Logit.

A Microstructure Study of Circuit Breakers in the Chinese Stock Markets

Steve Wang, Kuan Xu, and Hao Zhang

Abstract

Based on rare policy changes in the Chinese stock market in January 2016, we study the impacts of market-wide circuit breakers on market microstructure. To test if market-wide circuit breakers have the "cooling effect" and the magnet effect, we use high frequency transactions and limit order book data and Lasso IV models for endogenous market microstructure variables and exogenous policy variables based on a novel identification strategy. We find that market-wide circuit breakers have no "cooling effect" in decelerating falling prices (or returns) or reducing market volatility and order imbalance. Their presence does not affect bid-ask spreads but does reduces the large-, mid-, and small-sized trades in volume and trades. We also find that market-wide circuit breakers indeed induce significant magnet effects on stock returns, order imbalance, quote imbalance, and trades of various sizes.

Predicting bank distress in the UK with machine learning

Joel Suss & Henry Treitel

Motivation:

- Build early warning system for bank distress
- Two important criteria for regulators:
  1. Performance
  2. Transparency

Set-up:

- Input data: Regulatory returns; macro variables
- Outcome: subjective supervisory assessments of bank risk
- Compare 6 techniques (4 ML, 2 classical)
- Rigorous, double-block randomisation CV

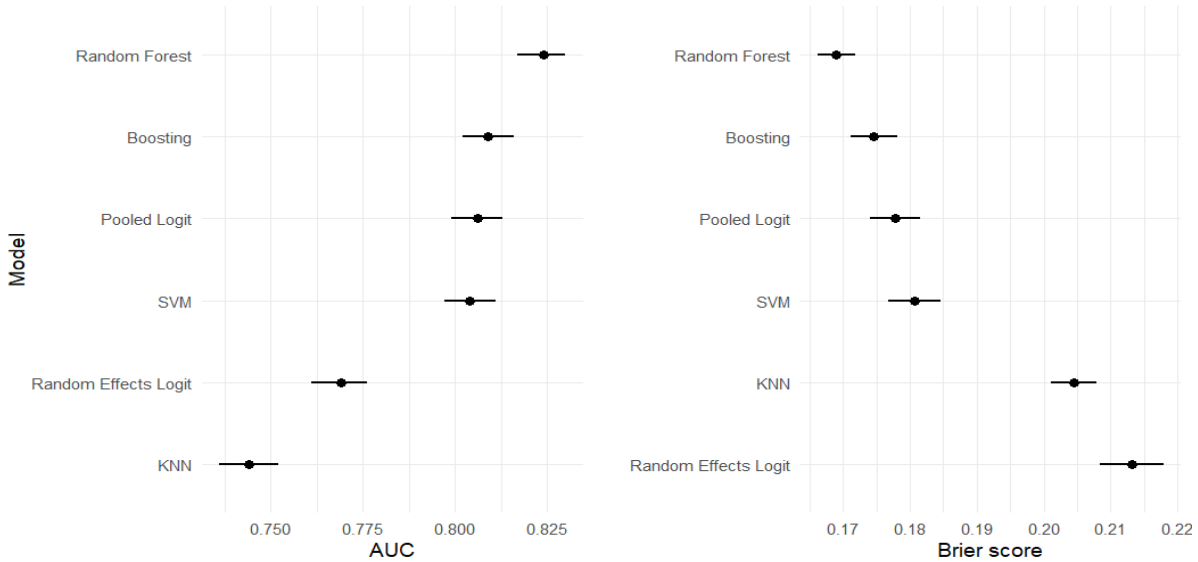
Performance:

- Random forest (RF) best standalone performer
  - AUC; Brier; false negative error rate
- Stack ensemble combining all models outperforms RF

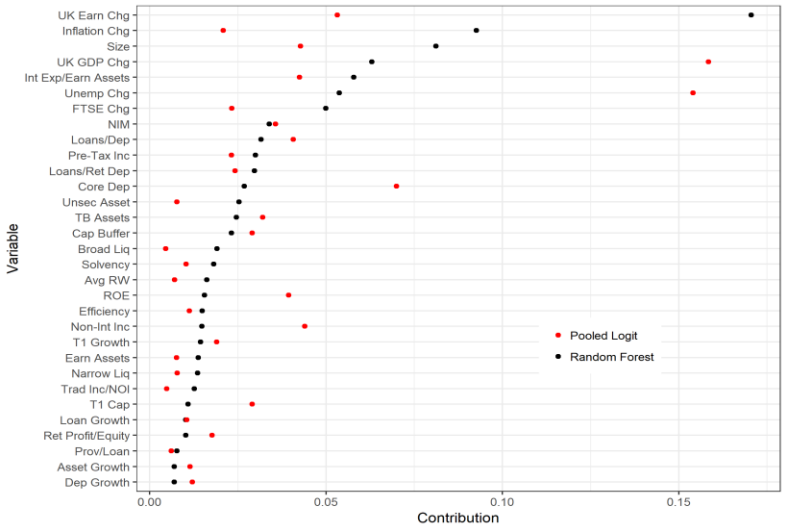
Transparency:

- Drivers of bank distress:
  - Shapley values (aggregated per predictor)
  - Shapley regression (statistical inference)
  - H-statistic (interaction strength)

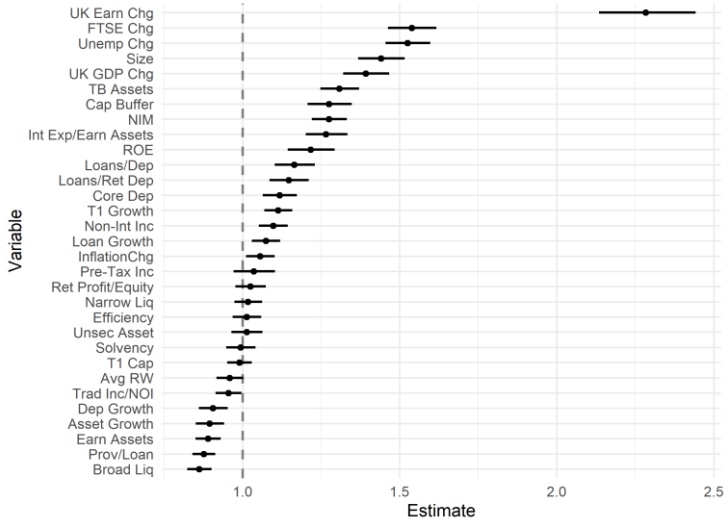
AUC and Brier score performance



Mean absolute Shapley values



Shapley regression coefficients



Risk Scorecards with Machine Learning

Fabrizio Russo[1], Torgunn Ringsjø[1], David Smith[1], James Woodcock[1], Thomas Pile[1], Lachezara Koteva[1]

[1]4most Europe Ltd.

fabrizio.russo@4-most.co.uk

4-most.co.uk

Abstract:

- The aim of the project was to develop a model that provides a better risk ranking than the scorecard for new business banking applications: improved discrimination would enable increased automation and accept rates as well as reduced bad rates.
- The experiment was designed including the following steps:
  - Horse Race (Algorithm Comparison)
  - Sample Design & Feature Engineering
  - Model Based Inference (based on Known-Good-Bad (KGB) and Accept-Reject (AR) models)
  - Through-the-Door (TTD) Model Build
  - Transparency through application level assessment
- Results: **8% increase in AUC using transparent XGBoost** - improved discrimination with a potential **increase in auto-decisions ($+33\%$)** and **accept rates ($+12\%$)** as well as **decrease in bad rates (-30%)**
- It was demonstrated that an ML model does not have to be a black box. Both variable and observation level analysis made the score **transparency comparable to a traditional scorecard**
  - XGB explainer was used to break down contributions of features to individual applications ensuring transparency and auditability
  - Stakeholders' review of sample applications was instrumental to model sign-off, informed the cut-off strategy and the underwriter guidelines
  - The implemented model was calibrated to align with the business' risk appetite
  - The business adopted the model's recommendations into their strategy with the confidence that they would be able to explain the decisions to both regulators and customers in line with GDPR
- The implementation of this ML Risk Score represents a **step towards acceptance of advanced analytics in a heavily regulated environment**

Boosting Crises over a Century - A Multi-country Perspective

Luca Alfieri (luca.alfieri@ut.ee)

- Models:
  - Logistic regression
  - Stepwise logistic regression (best performer)
  - Adaboost
- Macrohistory Database of Jordà-Schularick-Taylor from 1870 to 2017
- Contributions:
  - Data-rich environment and long time series
  - Machine learning method to detect crises in different countries

7

Financial Condition Index In an Incomplete Data Environment

Miguel C. Herculano

Abstract

This paper proposes a mixed-frequency factor-augmented vector autoregressive model with time-varying coefficients and stochastic volatility to construct a financial conditions index (FCI). This framework is extended to allow for different unbalanced panel techniques based on probabilistic principal components. In an incomplete data setting, with up to 62% of data missing, the approach yields a less noisy FCI that tracks the movement of the underlying financial variables more accurately and delivers better macroeconomic forecasts, on average.

Outline of the poster:

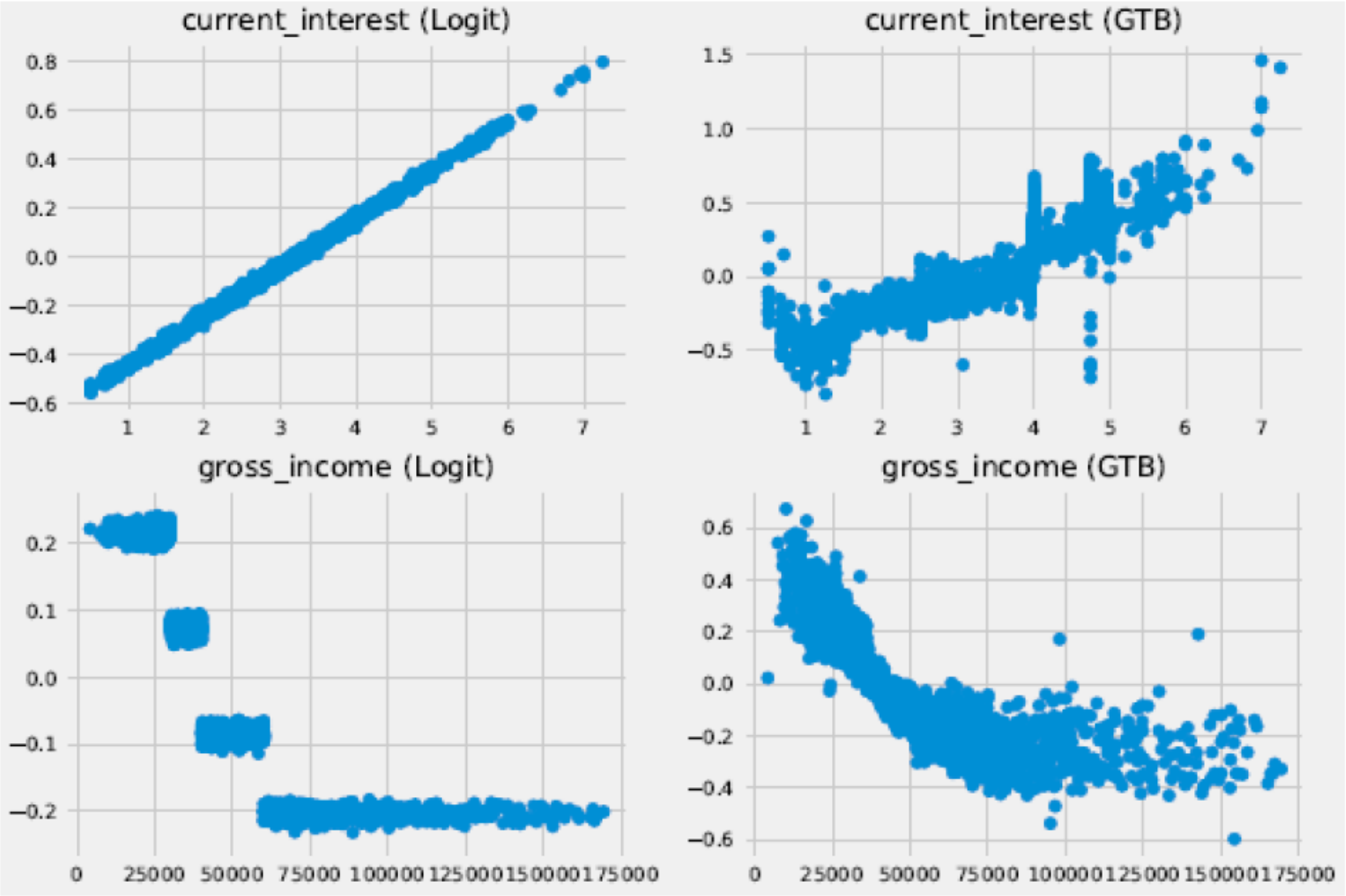
- What are Financial Condition Indexes ?
- Two main motivations to monitor financial conditions
- Common Critiques of Financial Condition Indexes
- A Mixed-Frequency Financial Conditions Index with Structural Instabilities
- Principal Component Analysis in the presence of missing values
- Results

Machine learning explainability in finance

- an application to mortgage risk analysis -

- Our approach brings out the drivers for individual predictions
- Helps make general statements of importance of variables
- The ‘workings’ of the simple model are easier to detect
- It is harder to extrapolate how the complex ML model would work under novel conditions

Simple, logit model | **More complex ML model**



Carsten Jung (Bank of England)*, Philippe Bracke (FCA), Anupam Datta (CMU), Shayak Sen (CMU)

Unsupervised anomaly detection in financial market infrastructures using neural networks: An application to the Canadian ACSS

Leonard Sabetti (1) and Ronald Heijmans (2,3)

(1) Bank of Canada, (2) De Nederlandsche Bank, (3) Payments Canada

Objectives

Can we analyze payments data by applying machine learning technics?

- Can a (deep) auto-encoder (AE) detect anomalous payment flows in a retail system?
- Can an AE detect a bank run automatically from the data?

Introduction

Figure 1:Daily net obligations.

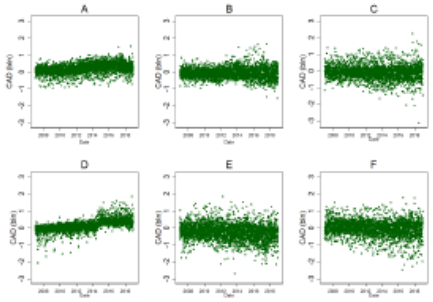
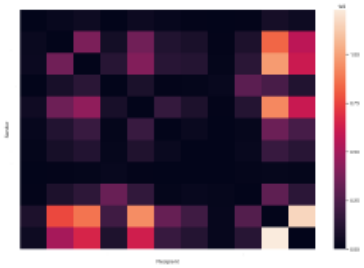
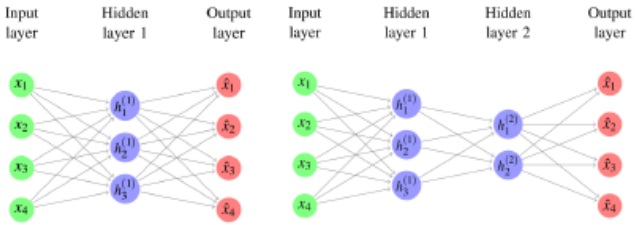


Figure 2:Heatmap bilateral flows.



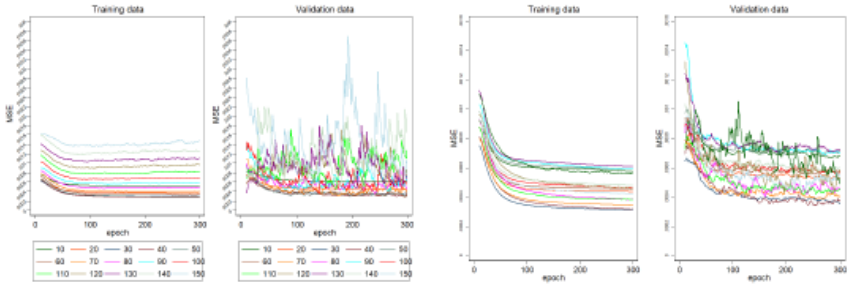
Method

Figure 3:AE with 1 (left) and 2 (right) hidden layers.



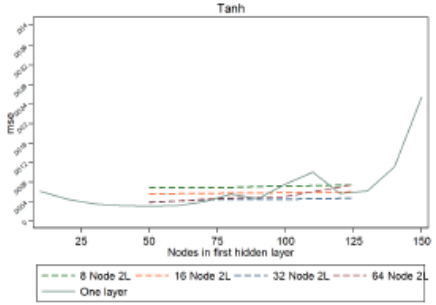
Optimizing number of neurons 1 HL (left) and 2 HL (right).

Figure 4:Daily net obligations.



Results

Figure 5:Model selection (tanh): 1 vs 2 HL.



Contact Information

- Email: Lsabetti@bank-banque-canada.ca, ronald.heijmans@dnb.nl

Econometrics at Scale: Spark Up Big Data in Economics

Benjamin Bluhm (benjaminbluhm@gmail.com) / Jannic Alexander Cutura (cutura@finance.uni-frankfurt.de)

- Ever increasing amount of data
- Ability to handle large datasets is crucial
- Little guidance for economists on how to handle big data
- Contribution:
  - **Practical guide** on how to use big data for economic research
  - **Accessible overview of concepts** and examples in **"econspeak"**
  - **R/Python source code** to reproduce examples

BANK OF ENGLAND

KING'S College LONDON

Thanks to all presenters.

Let's get out and chat!