Macroeconomic Predictions using Payments Data and Machine Learning*

James Chapman and Ajit Desai

November 5, 2021 (paper-link)

Advanced Analytics: new methods and applications for macroeconomic policy (virtual conference)

* The opinions here are of the authors and do not necessarily reflect the ones of the Bank of Canada

Objective

Demonstrate the usefulness of payments data and machine learning (ML):

- Use payments data from Canada's retail and large value payments systems
- Use ML models: elastic net, neural network, random forest, and gradient boosting
- Estimate current period (**nowcast**) GDP, retail, and wholesale trade sales

Address the associated challenges: interpretability and overfitting

- Shapley value-based approach to interpret ML model predictions
- Improved cross-validation strategy to alleviate the overfitting

Motivation

Macroeconomic Nowcasting:

- **Delay**: official estimates are released with a substantial lag
- **Uncertainty**: undergo multiple revisions sometime after years
- **Crisis**: nonlinear impacts and unconventional policies

Payments Data & Machine Learning:

- **Timely & Precise**: available immediately, no measurement or sampling error
- **High-frequency & Broad**: daily aggregates, 15+ years, 20+ streams
- **Handle Big Data**: non-traditional, high-frequency, wide and large
- **Nonlinearity**: flexible in capturing nonlinear relationships

Results Preview

Payments system data and ML models can lower nowcast errors significantly[1]:

- 35 to 40% reduction in RMSE for predicting GDP, retail and wholesale sales over a benchmark[2] and 15 to 25% reduction over payments data with factor model
- Out-of-sample model performance is relatively higher during the COVID-19 crisis period than the pre-COVID "normal" period
- Model interpretation reveals that, a few payments streams are important over entire nowcasting periods and their importance increases during crisis periods
- Proposed cross-validation strategy help to reduce nowcasting RMSEs (6-12%)

[1] Gradient boosting model performed consistently better than other models
[2] Our benchmark is mixture of a few lagged and timely indicators in a linear model

Literature

Payments data for macroeconomic prediction:

- Galbraith & Tkacz (2018): Nowcasting with payments system data
- Aprigliano et al. (2019): Payment system data to forecast the economic activity
- Chapman and Desai (2020): Nowcasting with retail payments data during crisis

Machine learning for macroeconomic prediction:

- Richardson et al. (2020): Nowcasting GDP using machine learning
- Maehashi and Shintani (2020): GDP prediction using factor models and ML
- Coulombe et al. (2020): How is ML useful for macroeconomic forecasting?

Machine learning interpretability and overfitting:

- Lundberg et al. (2017): SHAP-unified approach to interpret ML model predictions
- Buckmann et al. (2021): ML interpretability tool for economic forecasting
- Bergmeir and Benitez (2012): On the use of CV for time series predictions

Outline

1. Data
2. Methodology
3. Interpretability
4. Overfitting
5. Results

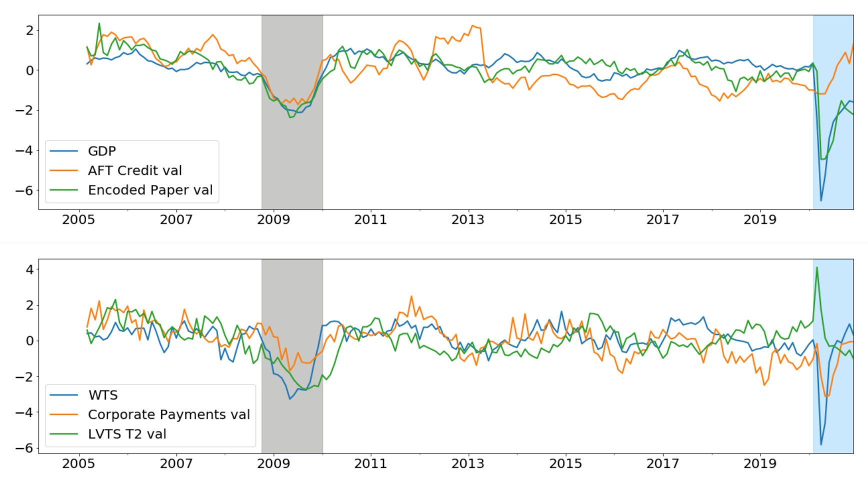
Data

Canadian ACSS and LVTS Data

Stream	Short Description
AFT Credit	Direct Deposit: payroll, account transfers, social security
AFT Debit	Pre-authorized debit (PAD): bills, mortgages, utility
Encoded Paper	Paper bills: cheques, bank drafts, paper PAD, etc.
Shared ABM	Debit card payments to withdraw cash at ABM
POS Payments	Point of sale (POS) payments using debit card
Corporate Payments	Exchange of Corporate-to-Corporate and bill payments
Allstream	It is the sum of all payments streams settled in the ACSS
LVTS-T1	Time critical payments and payments to Bank of Canada
LVTS-T2	Security settlement, foreign exchange and other obligations

Automated clearing settlement system (ACSS) and the large-value transfer system (LVTS)
First six streams are representative of twenty payments instruments processed separately in ACSS

Payments Data for Prediction



Standardization year-over-year growth comparisons of monthly targets and payments streams

Methodology

Models

Dynamic Factor Model (DFM): Captures dynamics of large set of predictors into small number of latent factors

$$X_t = \Lambda f_t + \epsilon_t,$$
$$f_t = A_1 f_{t-1} + \cdots + A_p f_{t-p} + u_t$$

Gradient Boosting Regression (GBR): Sequence of small trees are built on a repeatedly modified training dataset

$$\hat{\mathrm{y}}_i = \sum_{m=1}^{M} h_m(\mathrm{x_i})$$

Elastic Net, Support Vector Machines, Neural Network, and Random Forest

Payments Data & ML Models for Nowcasting

Opportunities:

- **Payments Data**: timely, precise, high-frequency and broad
- **ML models**: handle big data and nonlinearity; focus on prediction accuracy

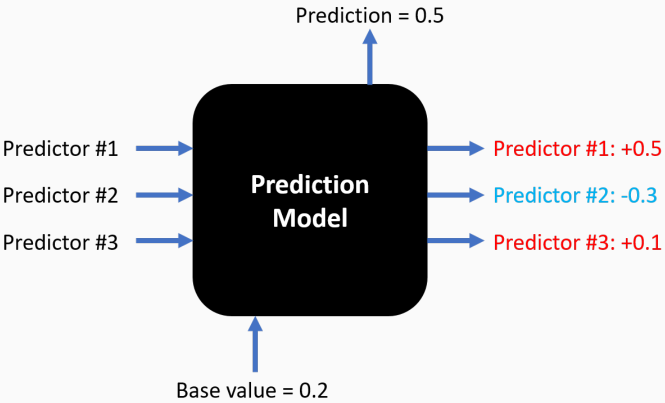
Challenges:

- **Missing information**: Not all payment schemes captured (credit card, on-us)
- **Many changes** in the streams: policy changes or technological advancements
- **Strong** seasonality, colinearity and non-stationary
- **Interpretability**: black-box nature, no causal relationships
- **overfitting**: high error-susceptibility, model selection

Interpretability

Shapley Values: SHAP[3]

Example: Consider nowcasting is a "game" then the Shapley values can be used to fairly distribute the *payout* (= the prediction) among the *players* (= the predictors)



[3] Lundberg et al. (2017). SHAP: A unified approach to interpreting model predictions.

SHAP: Advantages & Disadvantages

Advantages:

- Theoretical foundation
- Model independent
- Local and global interpretation

Disadvantages:

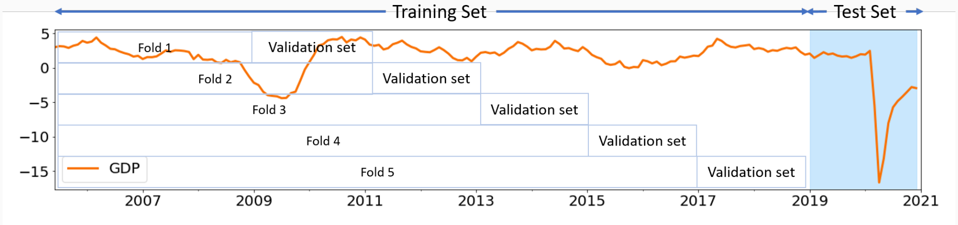
- Computationally expensive with increasing number of predictors
- Parametric models suffer from collinearity in the predictors
- Sensitive and prone to adversarial attacks (misleading interpretations)[4]

[4] Alvarez-Melis and Jaakkola (2018): On the robustness of ML interpretability methods for prediction problems
Slack, Dylan, et al. (2019): Fooling LIME and SHAP: adversarial attacks on post hoc explanation methods

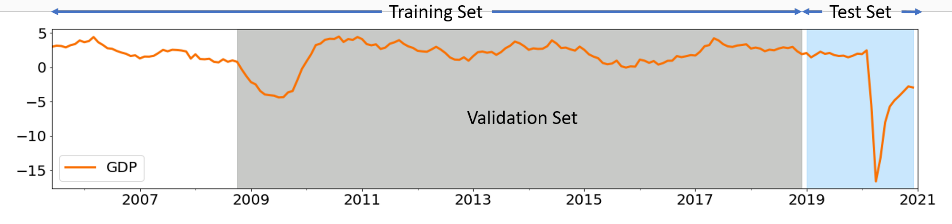
Overfitting

K-fold Cross-validation: Traditional vs randomized expanding window

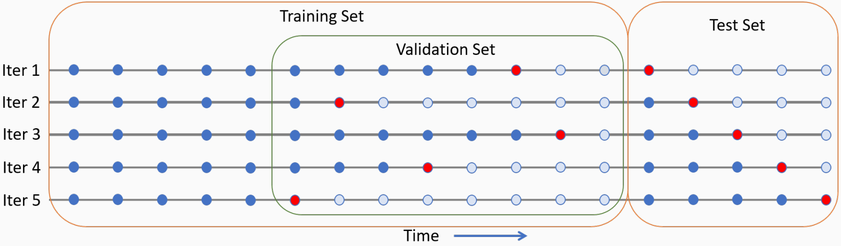
Standard approach for time-series:



Proposed approach for macroeconomic time series:



Randomized Expanding Window



Advantages & Disadvantages:

- Distribution of each validation set is similar to the test set
- Help avoid breaking the order of data (autocorrelation)
- Could sample many validation sets (no constrains)
- Some observations may get selected more than once, and some may never get selected in the validation set (theoretical and empirical analysis needed)

Results

Nowcasting Models: Case specifications

Prediction horizons $(t, t+1, t+2)$ are based on payments data availability t

E.g.: To predict May's GDP growth rates on June 1^{st}, i.e., at $t+1$, we use **payments data** for May (at t), and other latest available macro indicators:

- Base case (benchmark): OLS (5 predictor)[5]

$$\widehat{GDP}_{t+1} = \mathcal{F}(GDP_{t-2},\ CPI_{t-1},\ UNE_{t-1},\ CFSI_{t},\ CBCC_{t})$$

- Main case (of interest): DFM, ENT, RFR, GBR, ANN[6] (**23** predictors)

$$\widehat{GDP}_{t+1} = \mathcal{F}(GDP_{t-2},\ CPI_{t-1},\ UNE_{t-1},\ CFSI_{t},\ CBCC_{t},\ Payments_{t}).$$

[5] CPI: Consumer Price Index, UNE: Unemployment, CFSI: Canadian Financial Stress Indicator, CBCC: Consumer Board's Confidence Index

[6] OLS: Ordinary Least Squares, DFM: Dynamic Factor, ENT: Elastic Net, RFR: Random Forest, GBR: Gradient Boosting, ANN: Neural Network

Nowcasting Models: Results

RMSE on out-of-sample testing period[a] at $t + 1$ prediction horizon:

Target	Benchmark[b]	Main-DFM[c]	Main-ML[d]	% RMSE Reduction[e]
GDP	3.97	2.98	2.43	39*
RTS	8.47	6.36	5.44	36*
WTS	7.17	6.18	4.28	41*

[a] Training: Mar 2005 to Dec 2018 and testing: Jan 2019 to Dec 2020

[b] Benchmark: OLS using first available lagged target and other base case variables

[c] Main-DFM: Payments data along with the benchmark variables in the DFM model

[d] Main-ML: Payments data along with the benchmark variables in the ML model (only the best among ENT, RFR, GBR, ANN is showed)

[e] % Reduction in RMSE using ML model with payments data over the benchmark model

* Denote statistical significance at the 10% over benchmark

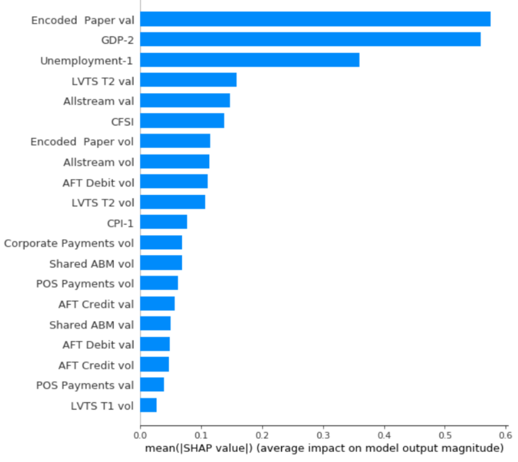
Model Interpretation and Payments Data Contribution

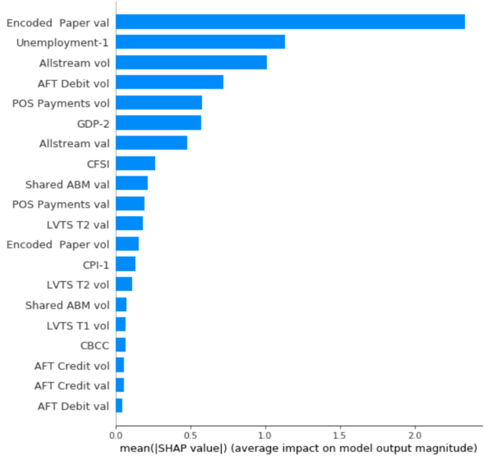
ML Models: Local interpretation

Force plots: provide insights into marginal contributions for each month's predictions

ML Models: Global interpretation

Feature importance plots: payments data importance increase during crisis periods

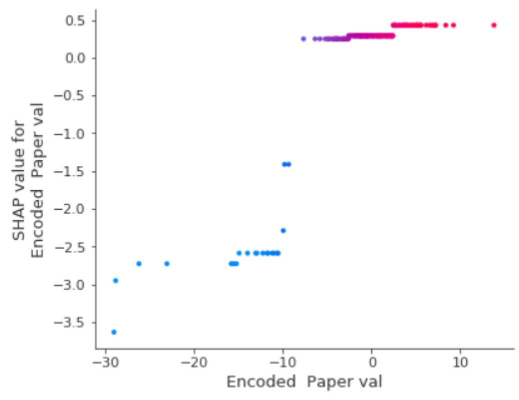


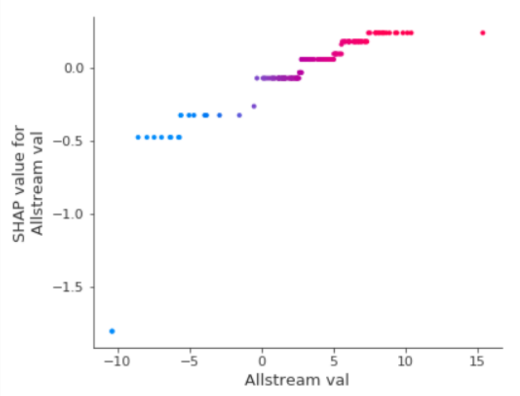


Left: full sample and Right: Covid-19 period (Mar to Dec 20)

ML Models: Dependence plots

Contribution of some of the payments streams is asymmetrical and nonlinear





Conclusions

This paper substantiates the use of payments data and ML models for macroeconomic prediction and provides a set of tools to overcome associated challenges:

- Payments data provide economic information in real-time and help reduce dependence on lagged variables (during both normal times and crisis periods)
- Machine learning provide set of econometric tools to effectively process various payments streams and capture sudden and large effects of the economic crisis
- Shapley value-based SHAP approach is useful to get insights into the ML model predictions (local and global interpretations)
- Proposed cross-validation technique can help reduce overfitting and improve prediction accuracy in macroeconomic prediction models

Thank you!