Nowcasting economic activity with electronic payments data: a predictive modeling approach

16th Payment and Settlement System Simulator Seminar

Bank of Finland

Helsinki, August 30-31, 2018

Carlos León

Financial Infrastructure Oversight Dept.

Banco de la República (Colombia)

cleonrin@banrep.gov.co

Tilburg University

CentER, Tilburg University

Disclaimer

This presentation is based on :

- León, C. & Ortega, F. (2018). Nowcasting economic activity with electronic payments data: A predictive modeling approach, *Borradores de Economía, No. 1037, Banco de la República.*

  http://www.banrep.gov.co/sites/default/files/publicaciones/archivos/borradores_de_economia_1037.pdf

As usual, the opinions and statements in these articles and this presentation are the sole responsibility of the author(s), and do not represent neither those of Banco de la República nor of its Board of Directors. Any remaining errors are our own.


iltapäivä | TIISTAI


iltapäivä | TIISTAI

But… how do they do it?

Well, it's sunny…

Today, I don't care!

Introduction

- The lack of timely information about the current state of the economy is a well-recognized problem among policy makers (Evans, 2005).
- *Nowcasting*, defined as *current-period estimates* (Galbraith & Tkacz, 2017) or *the prediction of the present, the very near future and the very recent past* (Banbura et al., 2017), has become a standard activity for central banks (see Tiffin, 2016, Hinds et al., 2017).
- Economic activity is one among many lagged key macro variables.
- Quarterly GDP figures are released with an approximate … week lag:*

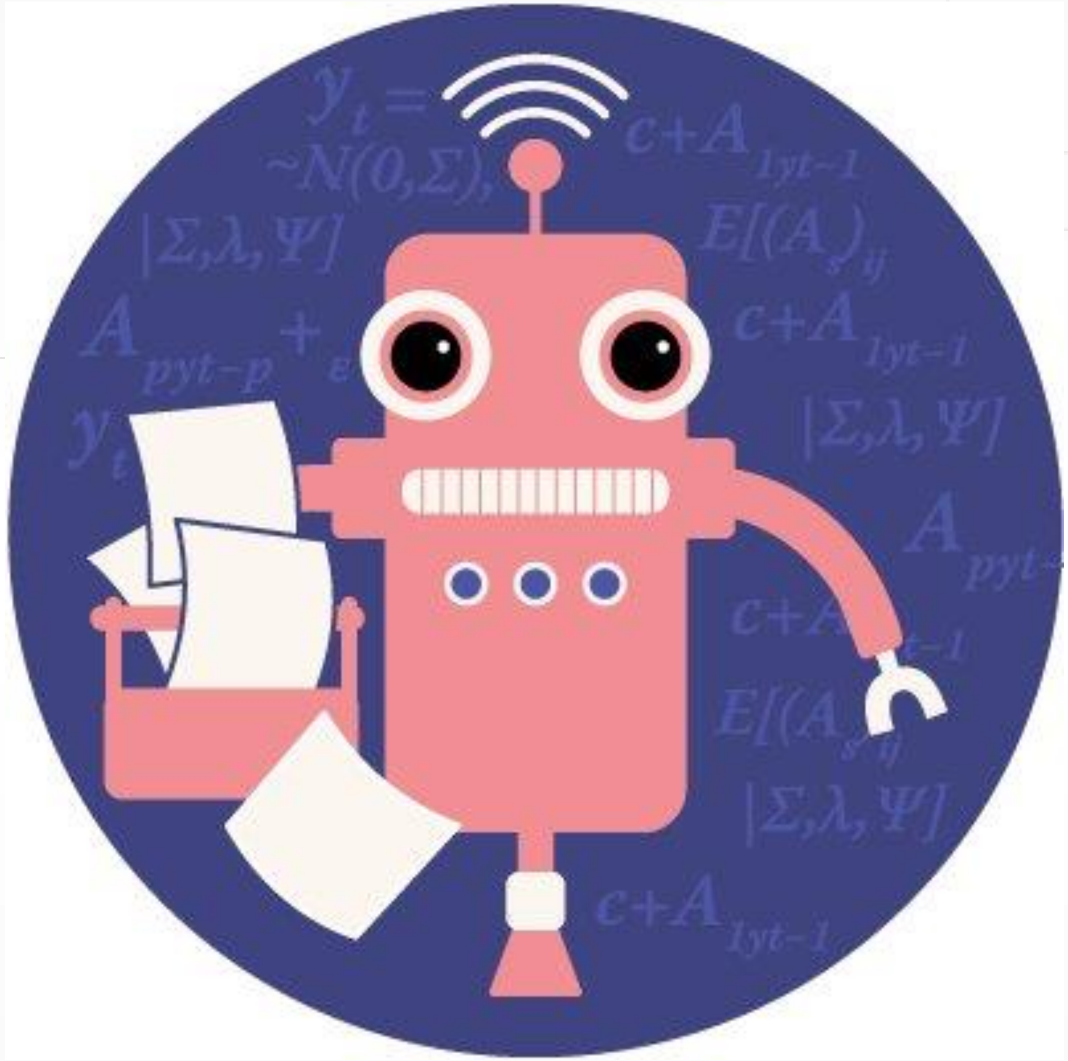
(*) see Banbura et al., 2013, Tiffin, 2016, Bragoli, 2017, Galbraith & Tkacz, 2017, DANE, 2017.

Introduction

- Our aim is to answer a single question: **is it possible to nowcast changes in the leading short-term activity indicator (ISE) with a dataset of electronic payment instruments*?**
- As nowcasting is better understood as a prediction task, we undertake a ***predictive modeling* approach** (see Shmueli, 2010, Varian, 2014, Mullainathan & Spiess, 2017); **neither a explanatory model nor a policy model are intended.**
- An accurate current-period estimate of ISE could provide the central bank, financial market participants, and other economic agents (e.g. the government, real sector) with **better tools for decision-making**.
- Closest related work by Galbraith and Tkacz (2017), who claim to be the first to make economic activity nowcasting based on payments data.

(*) **A payment instrument enables the holder or user to transfer funds, such as cash, cheques, debit and credit cards, and electronic transfers.** Cheques are to be considered electronic payments because they are cleared electronically. We focus on electronic transfers and cheques only.

@SPNowcast



y t =
~N(0,Σ),
c+A 1yt-1
|Σ,λ,Ψ]
E[(A s) ij
A pyt-p + e
c+A 1yt-1
|Σ,λ,Ψ]
y t
A pyt-
E[(A s) ij
|Σ,λ,Ψ]
c+A 1yt-1

Typical hat instead of the antenna
(no plans to tweet –for now)

Emeralds jewelry

Coffee beans ready to export

Contents

Data

- Nowcasting: exploiting the information that is available early and at higher frequencies (i.e. the input) to make a current-period or early estimate of a lagged or low-frequency variable (i.e. the output).
- In our case, two main challenges:
  - Finding a suitable set of **high-frequency indicators (i.e. inputs)** from which reliable signals of **economic activity (i.e. outputs)** are to be extracted.
  - Selecting an appropriate **prediction method** for extracting reliable signals of economic activity that serve as fair current-period or early estimates of the short-term economic activity indicator.

Data - Output

- Most literature sets GDP as output variable.
- We choose to nowcast ISE (*Indicador de Seguimiento a la Economía*)
  - A monthly indicator of economic activity that combines information about the production of goods and services pertaining to the most important economic activities in Colombia*.
  - Estimated and released by the Colombian bureau of statistics (DANE) with an approximate 2-month lag.
- As the set of electronic payments data is available from 2001, working on GDP quarterly figures would restrict our sample to about 48 observations, whereas working on ISE provides 192 observations –*still a non-large data set*.
- As electronic payments data is non-seasonally adjusted, we use the non-seasonally adjusted ISE.

(*) It corresponds to 111 indicators that track the dynamics of nine economic activities that compose the GDP, and they are weighted according to their contribution to the value added to economic activity. See DANE (2016).

DATA - OUTPUT

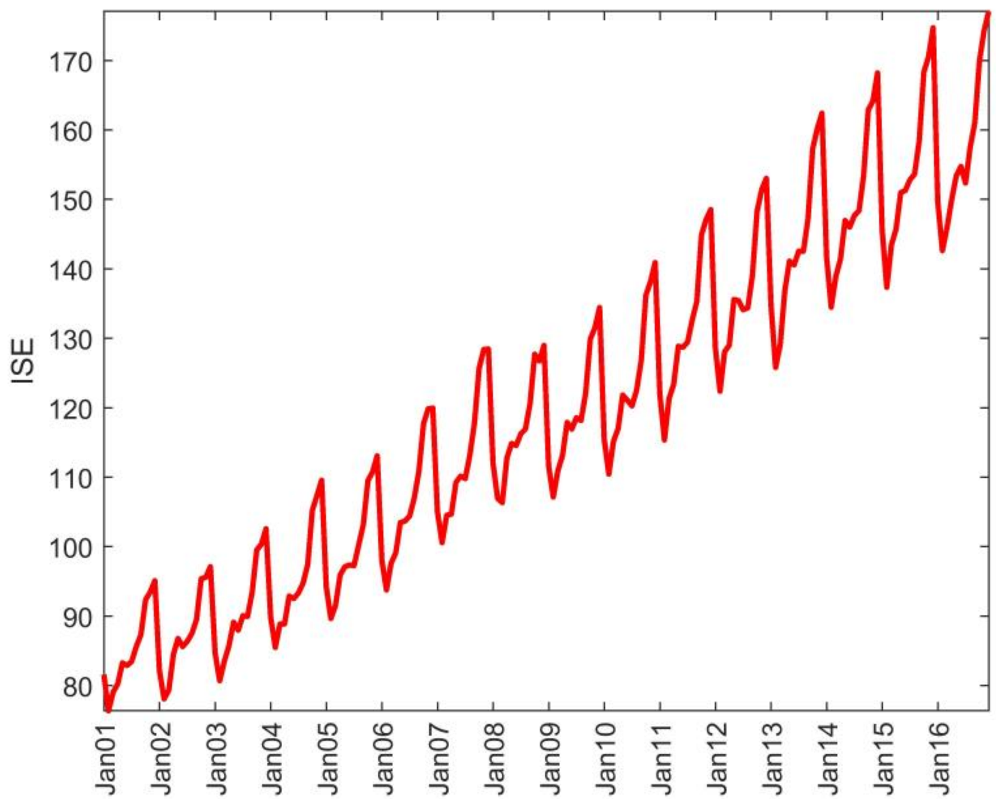


Figure 1. ISE. Non-seasonally adjusted, at constant prices, revised index. Source: DANE.

Contents

Data - Inputs

- Most economic activity nowcasting literature relies on macroeconomic data, financial variables, surveys, or a mixture of them (see Evans, 2005, Bell et al., 2014, Bec & Mogliani, 2015, Bragoli, 2017), but with a non-negligible lag (days, weeks, months).
- Economic activity from online data has surfaced. McLaren (2011) and Choi and Varian (2012) illustrate how Google query indexes may be used for short-term economic activity prediction.
- We choose to use electronic payments data.
- First implemented by Galbraith & Tkacz (IJF, 2017) for the Canadian case:
  - *The payments system variables capture a broad range of spending activity and are available on a very timely basis, making them suitable current indicators.*
  - Electronic payments are *available quickly* […] and *virtually free of sampling error.*
  - They use the value and number of operations from debit card transactions (i.e. point-of-sale payments) and small cheques (i.e. under $50,000).

Data - Inputs

- Our choice of payment instruments (value and number of operations):
  - Electronic transfers by individuals, firms, and the central government, processed and cleared by both ACHs:
    - ACH Colombia: owned and managed by banks, to process small repetitious payments (e.g. **payrolls, supplies**, mortgages, insurance, utility bills).
    - ACH Cenit: owned and managed by the central bank, to process central government's payments (e.g. **payrolls, transfers to municipalities**, public social security).
  - Cheques processed and cleared in central bank's local clearing house (Cedec).
- Unlike Galbraith and Tkacz (2017)…
  - We do not consider card transactions (2- or 3-month lagged; non-large contribution to total value of electronic payments, below 5%).
  - We do not limit cheques to being "small".
- Akin to Galbraith and Tkacz (2017), cash, and payments cleared and settled in the books of banks are unavailable –*and we know they are sizeable*.

Data - Inputs

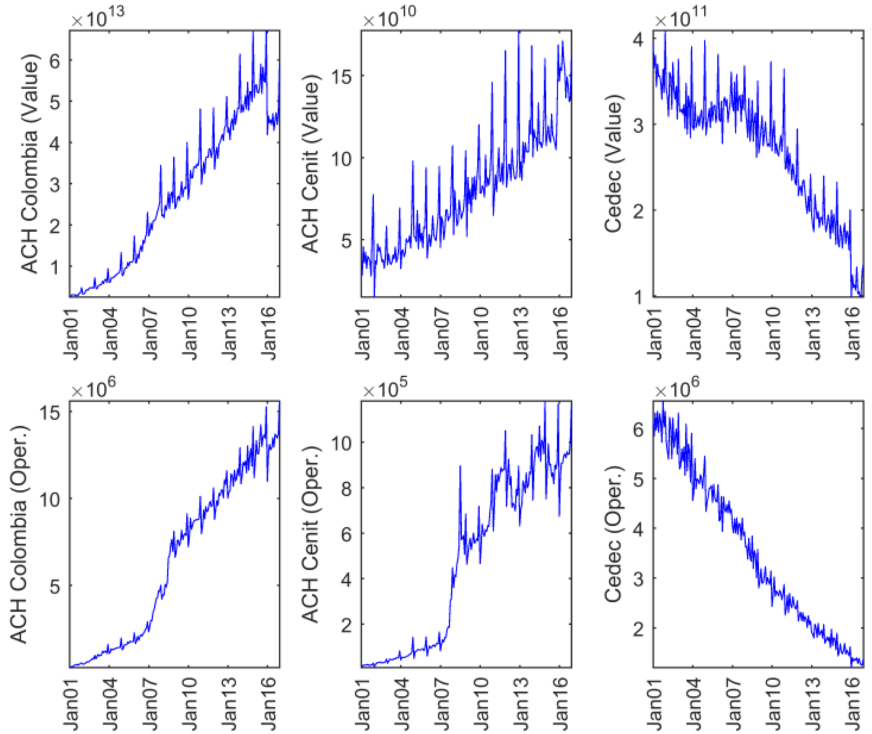


Figure 2. Electronic payment instruments data. Non-seasonally adjusted. First row presents the value (in constant Colombian pesos, deflated by the consumer price index) and the second the number of operations. Source: ACH Colombia and Banco de la República.

Contents

Prediction method

- Forecasting and nowcasting are, by definition, **prediction problems**.
- *Machine learning* methods (i.e. *predictive modeling*) are concerned with developing high-performance systems that can provide useful **out-of-sample predictions** (Varian, 2014).
- Accordingly, machine learning approaches have been used to forecasting and nowcasting economic activity:
  - ANNs (Tkacz and Hu (1999) for Canada, Aiken (2000) for U.S., Zukime & Junoh (2004) for Malaysia).
  - Elastic net regression (Tiffin (2016) for Lebanon, Hinds et al. (2017) for UK).
  - Random forests (Tiffin (2016) for Lebanon).

Prediction method

- Our choice, Artificial Neural Networks (ANN). Why?
  - ANNs can closely approximate any function, thus they are **able to deal with non-linear relationships** between factors in the data (see Bishop, 1995, Han & Kamber, 2006, Demyanyk & Hasan, 2009, Sarlin, 2014, Hagan et al., 2014). *This is the most plausible reason for enhanced GDP forecasting according to Tkacz and Hu (1999) and Zukime and Junoh (2004)*.
  - ANNs make **no assumptions about the statistical properties of the data** (see Zhang et al., 1999, McNelis, 2005, Demyanyk & Hasan, 2009, Sarlin, 2014).
  - ANNs have proven to be **very effective in time series prediction problems**, even better than standard econometrics (see Kohzadi et al., 1995, Zhang et al., 1999, Misas et al., 2003, McNelis, 2005, Jalil & Misas, 2006, Han & Kamber, 2006, Chaudhuri & Ghosh, 2016, Di Piazza et al., 2016).

Prediction method

- *Black box criticism*: ANNs' results are **opaque and lack interpretability** (see Han & Kamber, 2006, Witten et al., 2011, Chakraborty & Joseph, 2017).
- *Black box criticism* comes from a desire to tie down empirical estimation with an underlying economic theory (McNelis, 2005).
- We do not care about the *black box criticism* because **we have no underlying economic theory to test**.
- We do not aim at producing *good estimates of parameters* that underlie the relationships between inputs and outputs: **we aim at *good output estimation***.

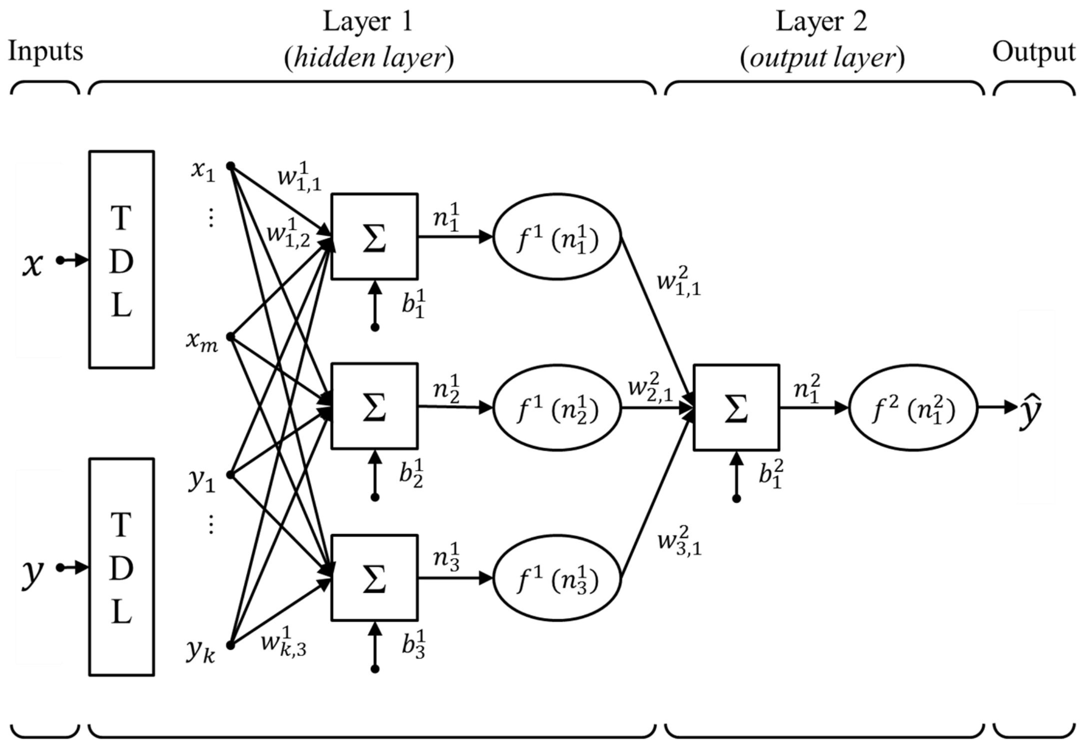
NARX ANN

- Based on Lin et al. (1996), let Ω represent a non-linear function, and k and m the output and inputs lag order, respectively, a Non-linear AutoRegressive eXogenous (NARX):

$$y_t = \Omega[\underbrace{x_t, x_{t-1}, \dots, x_{t-m}}_{\textit{Exogenous}}, \underbrace{y_{t-1}, \dots, y_{t-k}}_{\textit{Autoregressive}}]$$

- The output variable (y_t) is modeled as a non-linear function Ω of k lagged y_t variables and m lagged exogenous variables (x_t).
- For the NARX model, Lin et al. (1996) suggest that the non-linear function Ω may be approximated by an ANN structure, i.e. a NARX-ANN model.
- The NARX ANN model can efficiently be used for modeling non-stationary and non-linear dynamic systems (see Mahmoud et al., 2013, Chaudhuri & Ghosh, 2016, Di Piazza et al., 2016).

NARX ANN



$$\hat{y} = f^2(\mathrm{w}^2\mathrm{f}^1(\mathrm{w}^1(x,y) + \mathrm{b}^1) + \mathrm{b}^2)$$

Figure 3. Sample NARX ANN with two layers. The first (hidden) layer consists of two *tapped delayed lines* (TDL) that contain the vectors of lagged inputs; three bias scalar terms (b_1^1, b_2^1, b_3^1) with a constant input of 1; all w^1 weights that connect each input with the three neurons (i.e. the sum operators and their corresponding activation functions, f^1). The second layer consists of a single neuron with an operator that sums the product of the weights ($w_{1,1}^2, w_{2,1}^2, w_{3,1}^2$) and the result from the neurons in the first layer ($f^1(n_1^1), f^1(n_2^1), f^1(n_3^1)$); one bias scalar term, b_1^2, with a constant input of 1; and the second activation function ($f^2(n_1^2)$), which yields the output. Based on Hagan et al. (2014) and Di Piazza et al. (2016).

A *COMMITTEE* OF ANN

- Our choice of NARX ANN model:
  - **Two layers**, as the one exhibited (*the standard*).
  - Outputs & inputs as log-returns. We **nowcast changes in economic activity**.
  - **Lag of output variable (y_t) is 12 months** (i.e. seasonality), starting in the second lag.
  - **Exogenous inputs (x_t) are not lagged** (i.e. analogous accuracy, slower process).
  - Random allocation of train (70%), validation (15%) and test datasets (15%); i.e. performance is robust to changes –block allocation, interleaved allocation.
  - Four scenarios of number of neurons (10, 20, 30, 40), for ***empirical tuning*** (see Mullainathan & Spiess, 2017).
  - **1,000 training processes** (1,000 ANNs)…
    - A single training may not produce optimal performance (i.e. local optima).
    - This is akin to a ***bootstrap method*** –convenient as dataset is not large.
    - Enable to obtain a **density forecast** (Chakraborty & Joseph, 2017).
    - Enable to use a ***committee* method** (see Bishop, 1995, Mitchell, 1997, Hastie et al., 2013, Hagan et al., 2014); the **joint output will usually achieve higher performance** than any single network used in isolation.

Contents

Main results and validations

- Two standard measures of out-of-sample (i.e. on the test set) performance:
  - Root mean square error ($RMSE$) between the observed and the predicted logarithmic returns of ISE.
  - Correlation (r) between the observed and the predicted logarithmic returns of ISE.
- Estimated on two different stages:
  - ~~Average $RMSE$ and average r, on 1,000 predictions, denoted as $\overline{RMSE}$ and $\bar{r}$.~~
  - **$RMSE$ and correlation on the average prediction (i.e. the committee prediction), denoted as $\overline{\overline{RMSE}}$ and $\bar{\bar{r}}$.**
- A naïve benchmark (UK's quarterly GDP nowcast, 1992-2017):

[…] being able to predict GDP growth with an RMSE of 0.3% is relatively impressive.

Hinds et al. (2017)

UK's average GDP absolute change: 0.67%

With data from UK's Office for National Statistics

Main results and validations

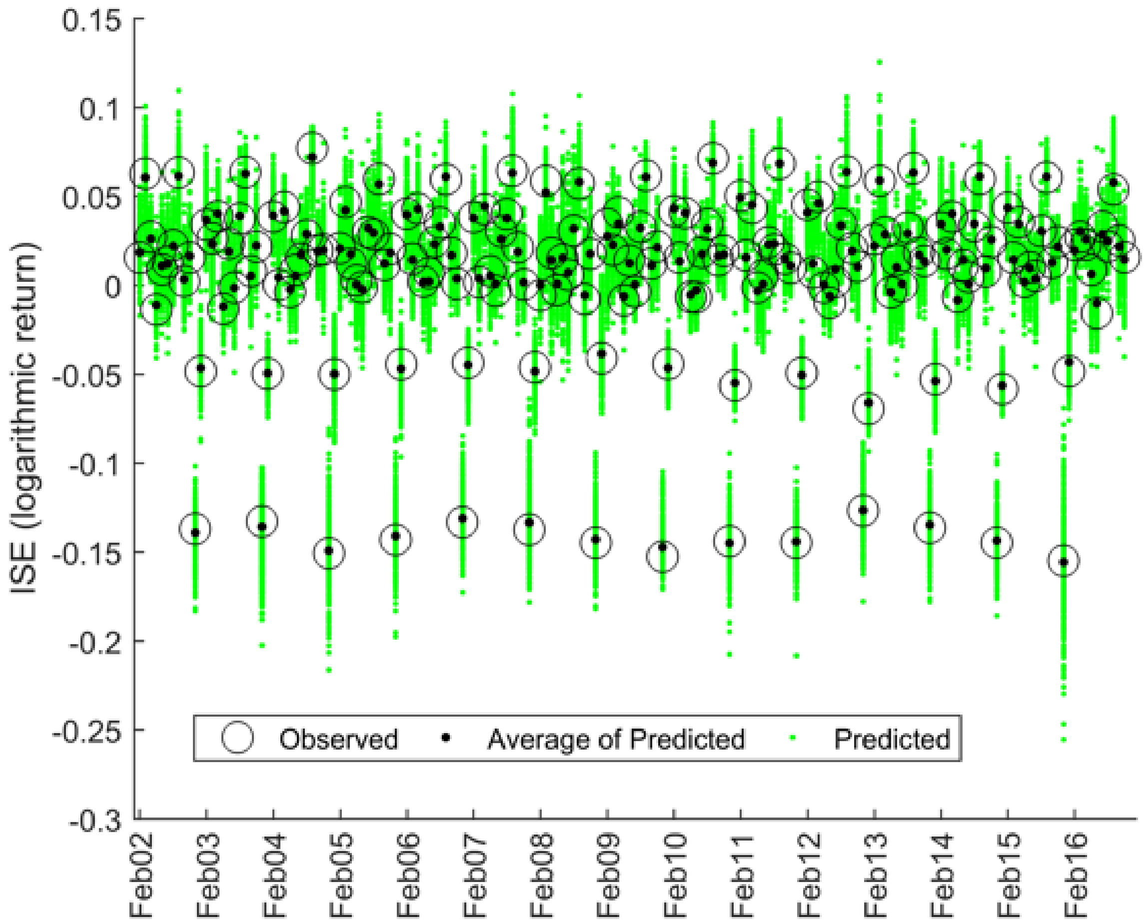


Figure 4. Observed (empty circles), predicted (green dots), and average of predicted (black dots) ISE logarithmic returns. Predicted ISE series correspond to 1,000 independent training processes of the scenario corresponding to 30 neurons. Source: DANE and authors' calculations.

Main results and validations

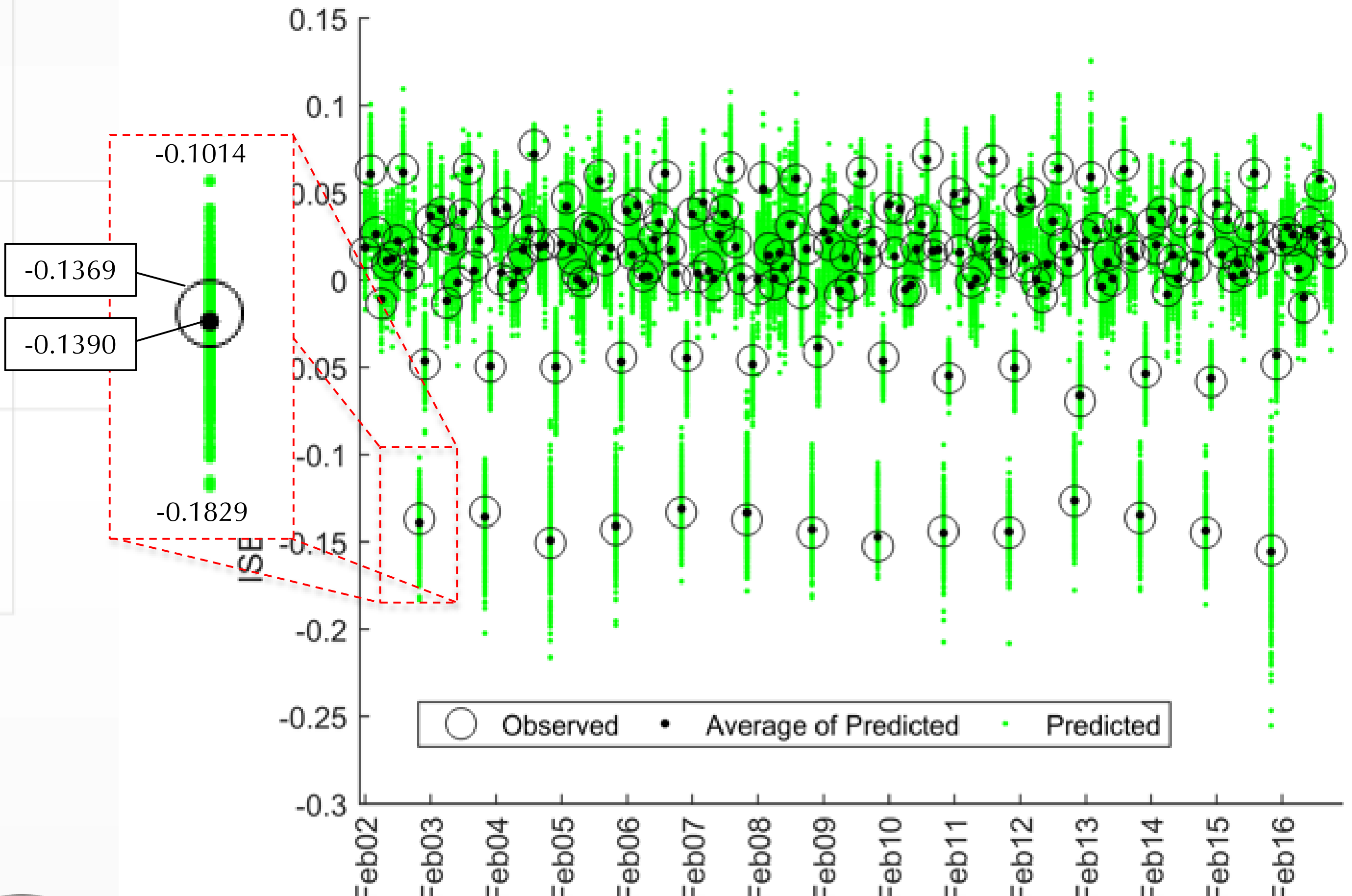


Figure 4. Observed (empty circles), predicted (green dots), and average of predicted (black dots) ISE logarithmic returns. Predicted ISE series correspond to 1,000 independent training processes of the scenario corresponding to 30 neurons. Source: DANE and authors' calculations.

Main results and validations

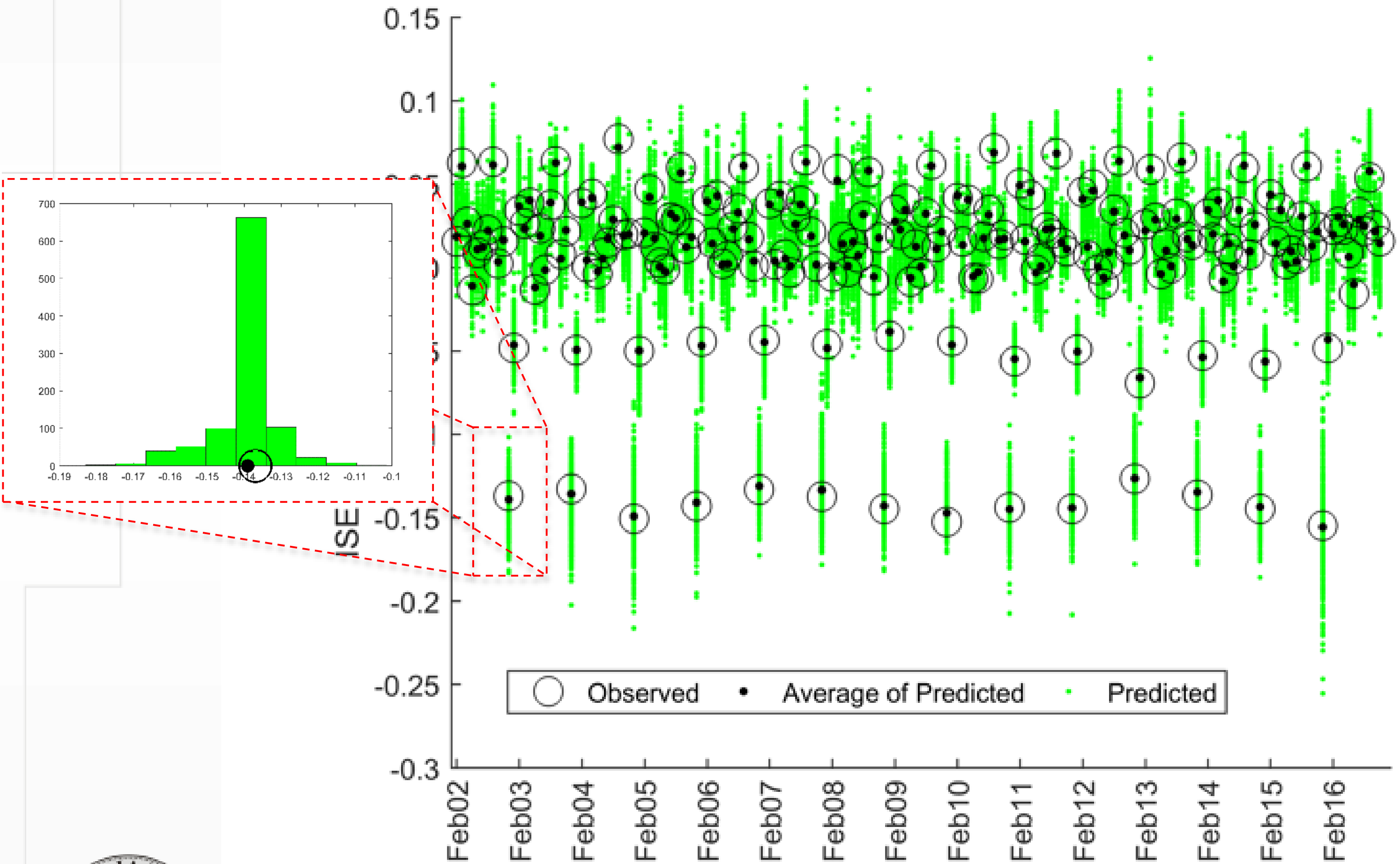


Figure 4. Observed (empty circles), predicted (green dots), and average of predicted (black dots) ISE logarithmic returns. Predicted ISE series correspond to 1,000 independent training processes of the scenario corresponding to 30 neurons. Source: DANE and authors' calculations.

Main results and validations

	Number of neurons			
	10	20	**30**	40
$\overline{\overline{RMSE}}$	.0042	.0029	**.0026**	.0026
$\bar{\bar{r}}$	.9971	.9984	**.9987**	.9987

Table 3. Committee of networks' predictive (out-of-sample) performance for selected scenarios of number of neurons in the hidden layer. $\overline{\overline{RMSE}}$ corresponds to the $RMSE$ calculated with the average predicted log-return on the 1,000 independent training processes of each scenario. $\bar{\bar{r}}$ corresponds to the correlation coefficient between the observed and the average predicted log-return on the 1,000 independent training processes of each scenario. The highest performance scenario is in bold. Source: authors' calculations.

On the average prediction for each observation (i.e. *the committee*)

- A benchmark from UK:
  - RMSE / Avg.Abs.Change = .0030 / .0067 = **.4476**
- In our case (30-neuron):
  - Average MoM ISE abs. change (01-2016) = .0361
  - $\overline{\overline{RMSE}}$ = .0026
  - $\overline{\overline{RMSE}}$/ Avg.Abs.Change = .0026 / .0361 = **.0720**

Main results and validations

- Validation:
  - Does electronic payments data contribute to reducing economic activity nowcast error?
  - That is, does the NARX-ANN improve the nowcasting performance of a NAR-ANN benchmark model?

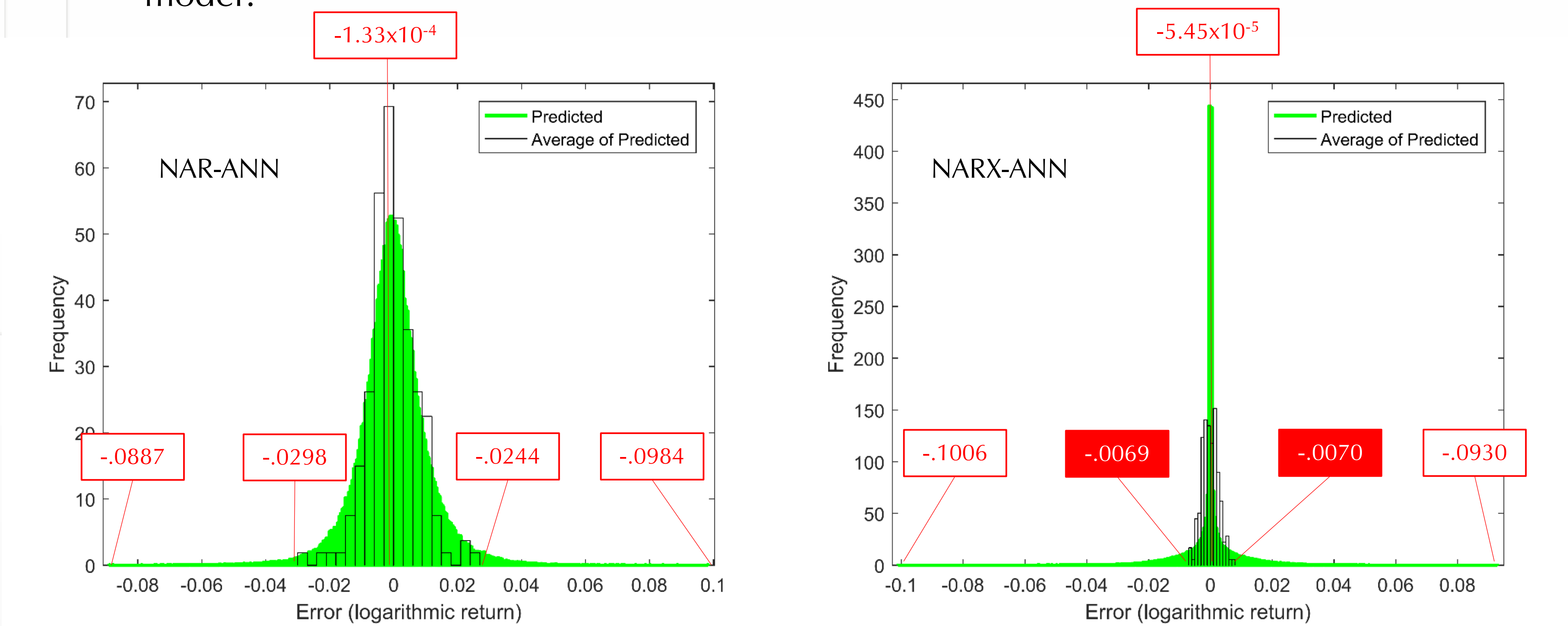
	NAR [a]	NARX [b]	Change (%) [c]
$\overline{\overline{RMSE}}$	.0054	.0026	-51.85

Table 5. Comparison of predictive out-of-sample performance for the best performing NAR-ANN and NARX-ANN models. $\overline{RMSE}$ corresponds to the average $RMSE$ estimated on 1,000 independent training processes for each scenario. $\overline{\overline{RMSE}}$ corresponds to the $RMSE$ calculated with the average predicted log-return on the 1,000 independent training processes of each scenario. [a] The best performing NAR-ANN by $\overline{RMSE}$ and $\overline{\overline{RMSE}}$ are the 10-neuron and 40-neuron models, respectively (see Table 8, in Appendix). [b] The best performing NARX-ANN is the 30-neuron model (see tables 2 and 3). Source: authors' calculations.

- The best NARX-ANN outperforms the best NAR-ANN
- Moreover, the worst NARX-ANN outperforms the best NAR-ANN

Main results and validations

- Validation:
  - Does electronic payments data contribute to reducing economic activity nowcast error?
  - That is, does the NARX-ANN improve the nowcasting performance of a NAR-ANN model?



Contents

Final remarks

- Concurrent with the work of Galbraith and Tkacz (2017), we rely on electronic payment instruments data as high-frequency indicators from which reliable signals of economic activity are to be extracted to nowcast ISE.
- Results suggest that electronic payment instruments data and a NARX ANN enable us to nowcast economic activity with fair accuracy.
- The NARX-ANN model outperforms a benchmark non-linear autoregressive artificial neural network model (NAR-ANN); electronic payments data does contribute to reduce nowcast error.
- So, results suggest that **it is possible to nowcast changes in the ISE with a set of electronic payment instruments data.**

Final remarks

- Extensions:
  - From nowcasting to forecasting.
  - High-frequency (e.g. daily, weekly) nowcast of economic activity (as suggested in Evans, 2005).
  - Using nowcasts to feed other models.
  - Including other variables compatible with nowcast objective (e.g. interest rates, stock indexes).
  - Debit and credit card data (i.e. the dynamics of payment habits and technology).


TIEMPO