

Assessing the use of Google Trends to predict credit developments*

E. Burdeau, E. Kintzler
Banque de France

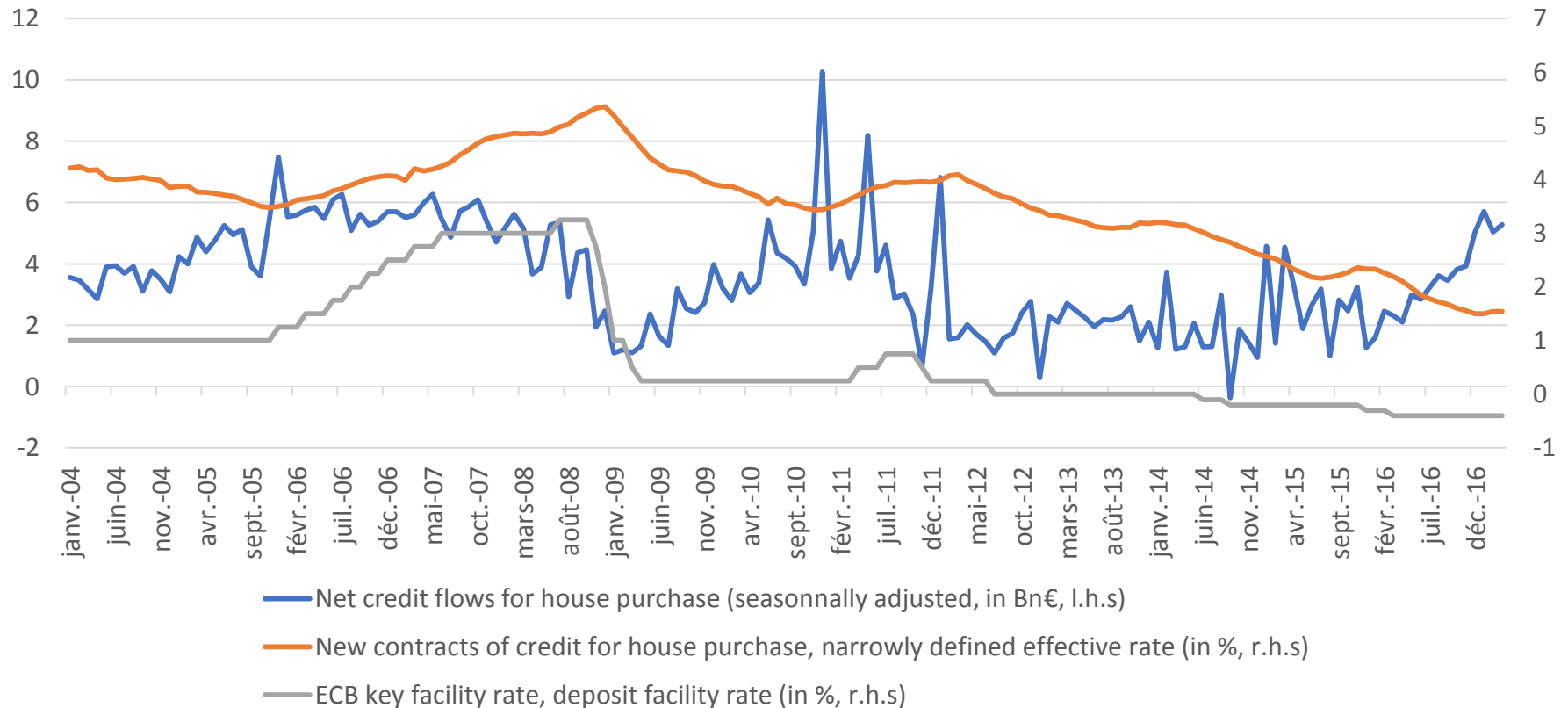
edwige.burdeau@banque-france.fr



*This article reflects the opinions of the authors and do not necessarily express the views of the Banque de France.

Motivation

Net credit flows for house purchase, narrowly defined effective rate of new contracts of credit for house purchase and ECB deposit facility rate



Sources: ECB, Banque de France



Google Trends, a widely used data source

Contrary to official statistics, Google Trends are made available **in near real time**.

A large part of the recent literature on forecasting has been devoted to the usefulness of Google Trends **to catch the demand side of the economy**:

- We identified **more than 50 articles** published since 2009 (without being exhaustive!)

From different fields, from finance to macroeconomics, especially on the housing sector:

- Nowcasting and forecasting of home sales, house prices, mortgage delinquency, building permits...

Generally relying on **“simple” models** such as autoregressive models augmented with Google Trends indicators, with some exceptions.

Our approach

- **Forecasts from 0 to 3 months in advance** of housing indicators: credit flows for house purchase ;
- With different explanatory indicators from Google Trends, we tried:
 - to define a **generic approach with Google Correlate** to identify appropriate indicators,
 - to identify **“robust” indicators**, by using principal component analysis.
- Relying on different **variable selection models, from the recent “machine learning” literature**, both linear and non-linear models.



Google Correlate – Automated query selection

In two ways:

By uploading external time series on credit for house purchase,

- Google Correlate returns the most highly correlated Google Trends and,
- **We selected manually 12 indicators, containing names of French credit institutions.**

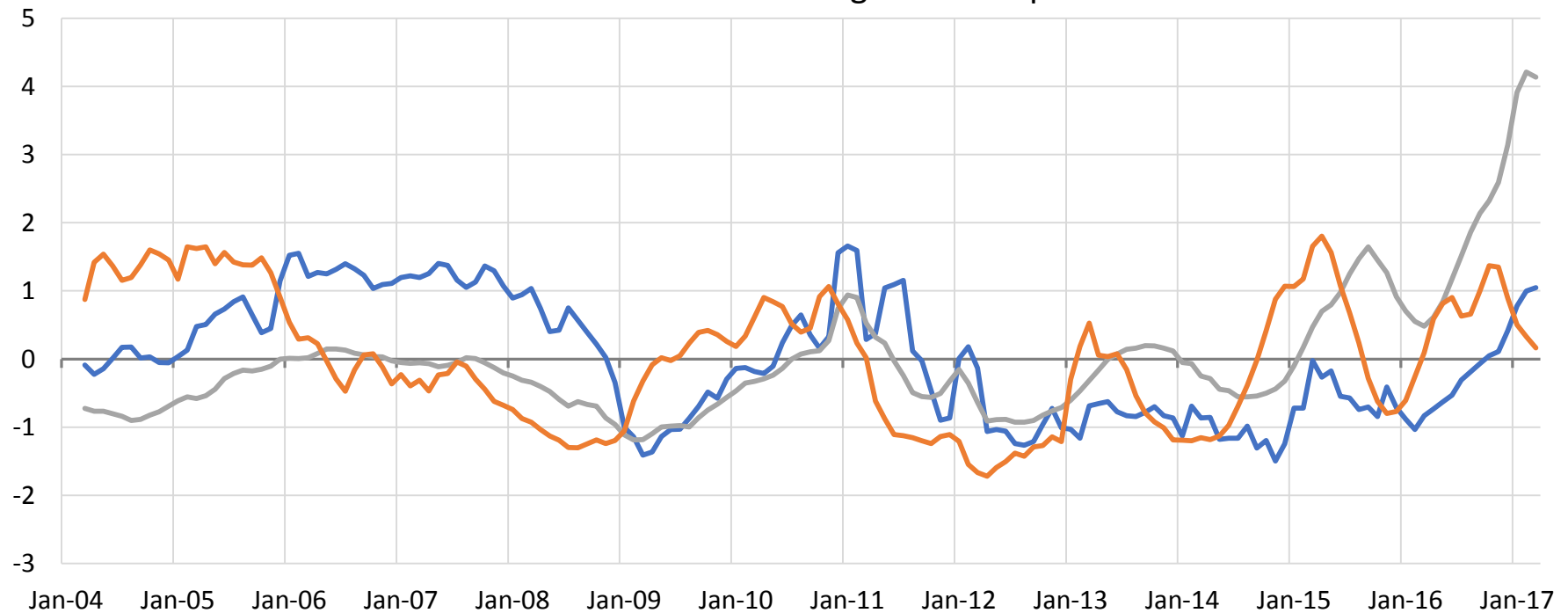
By giving a prespecified individual Google query (“credit for house purchase”),

- Google Correlate returns the 100 queries the most highly correlated,
- **We retained the first 5 principal components of this set of indicators.**

Set of potential explanatory variables: 17 variables from Google Trends + the 3 lagged terms of the endogenous variable.

Google Trends carry leading information on credit flows

Evolutions of credit flows for house purchase and the second principal component of the 100 correlates related to lending for house purchase



- Net credit flows for house purchase (seasonnally adjusted, standardized)
- New contracts of lending for house purchase (seasonnally adjusted, standardized)
- Second principal component of the 100 correlates of the term "lending for house purchase" (seasonnally adjusted, standardized)



A set of machine learning models have been tested

Different variable selection models to curb the noise of irrelevant preselected indicators:

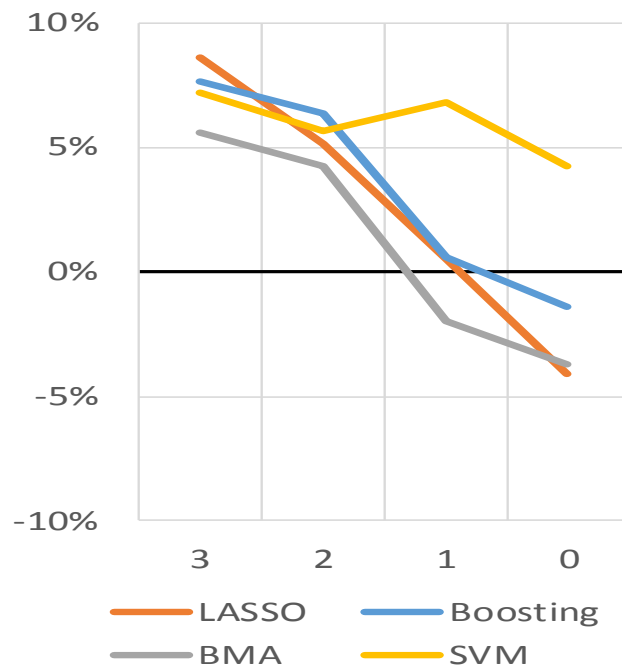
- **Penalized linear models (LASSO and Elastic net):** parameters of irrelevant variables are shrunk towards zero ;
- **L₂ Boosting approach:** an iterative approach in which only relevant variables are selected ;
- **Bayesian model averaging:** posterior probabilities of models containing irrelevant variables are low ;

A first experiment of a non-linear approach, called **support vector machine**, relying on a kernel approach (without variable selection).

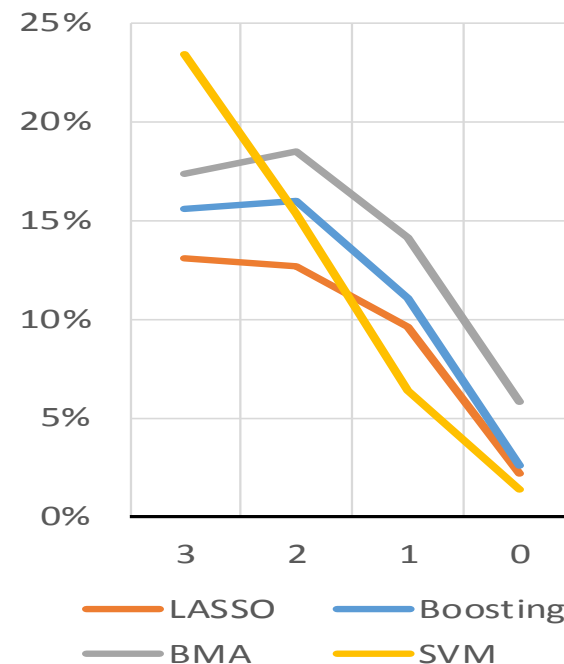
Google Trends indicators help forecasting credit flows several months in advance

Gain in RMSE compared to the AR model, in function of the number of months in advance forecasts are estimated

Forecasts of net credit flows for house purchase

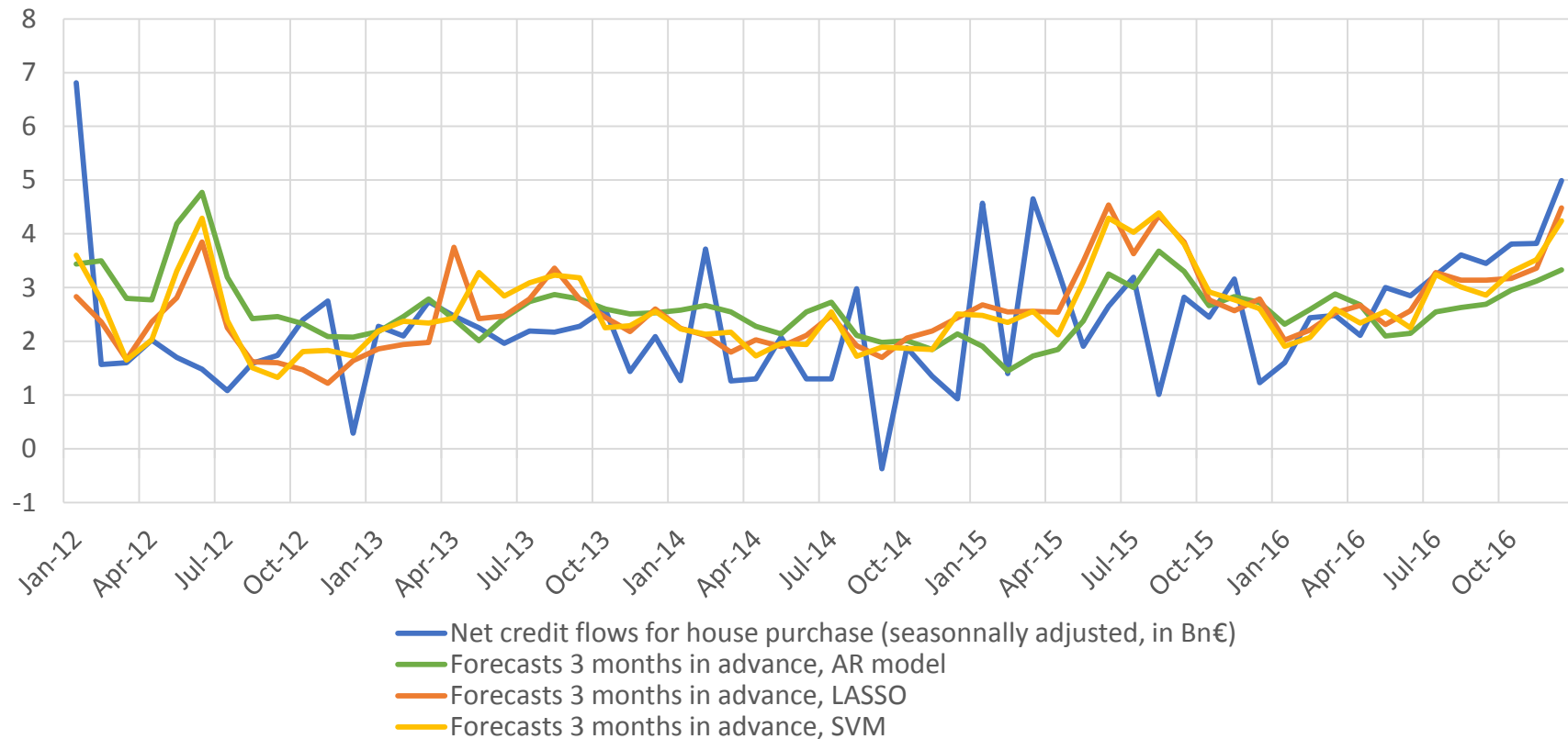


Forecasts of new contracts for house purchase



Google Trends indicators help forecasting credit flows several months in advance

Comparisons between net credit flows for house purchase and out-of-sample forecasts 3-months in advance for different type of models





Way forward

- Using both Google Trends and Google Correlate can significantly improve medium-term forecasts.
- We identified a good leading indicator of credit flows for house purchase :
 - the second principal component built with the 100 correlates of the term “lending for house purchase”.
- But, a frequency approach could confirm that Google Trends indicators are more appropriate to identify medium-term cycles.
- First experiment of a nonlinear approach is conclusive,
 - Other models, such as neural networks, should be tested.