

Use of Google Trends Data in Banque de France Monthly Retail Trade Survey*

François Robin

Key Question

Google Trends data provide the number of searches performed by internet users on the Google search engine for a given theme. They are free and available in quasi-real time, which are properties of *a priori* interest. However, do these data provide additional information on online retail sales data for the previous month in comparison with survey data and the expected change based on the historical time series?

Methodology

The methodology in construction of Google Trends indices is unclear. This means that outliers must be corrected automatically. Furthermore, the huge range of possible variables favours variable selection, using the adaptive lasso process. Lastly, using three sources of information enables comparison of the combination of single models, *through* Bayesian aggregation, to the global model.

Main Results

- Three single models enable us to forecast online retail sales over a one-month horizon: the SARIMA and two adaptive lasso models, using Google Trends data and quantitative survey trend indices from the monthly retail trends survey. The combination of these three models and the global model (adaptive lasso applied to all variables) uses all information.
- For all six estimations (total, shoes, consumer electronics, clothing, furniture and household appliances), the forecasts for the five models are close. The contribution of exogenous data is clearest for the total.
- The Google Trends model is parsimonious and stable in terms of selection and coefficients. However, output from models using Google Trends are sensitive to the sampling method used by Google in their construction.
- The model combination performs better overall than the global model.

Model performance in estimation of total online retail

Total	Google Trends	Retail	SARIMA	Global model	Model combination
Root Mean Square Forward Error	4.8	5.2	5.0	5.5	4.8

Note: A lower mean quadratic error indicates better performance.
Sources: Google Trends, FEVAD, Banque de France DGS SEEC.

Message

Free and available in quasi-real time, Google Trends data offer valuable properties in forecasting online retail sales data at one-month horizons. However, their use requires prior analysis of suitable search themes, automated outlier correction and multiple searches for identical terms to ensure the reliability of results. In particular, due to the changeability of online retail and changing internet user patterns, Google Trends data are useful but not conclusive in short-term online retail forecasting.