NOWCASTING PRIVATE CONSUMPTION: TRADITIONAL INDICATORS, UNCERTAINTY MEASURES, AND THE ROLE OF INTERNET SEARCH QUERY DATA

María Gil, Javier J. Pérez, and Alberto Urtasun
Banco de España

A. Jesus Sánchez
Univ. Complutense

61st ISI World Statistics Congress
IFC sponsored session: The use of leading indicators to anticipate economic activity
Marrakech, 18 July 2017

Associate Directorate General Economics and Research
The views expressed in this presentation are those of the authors and should not be attributed to the Banco de España or the Eurosystem.
Outline

1. Motivation
2. Literature review
3. The data
4. Modeling approach
5. The empirical exercise
6. Selection of results and conclusions
1. Motivation

- Private consumption accounts for a significant share of GDP fluctuations, given its weight (some 60%) and the intensity of its variations.
- The accurate assessment of the current state of private consumption is thus key for real-time policy analysis (in particular the detection of turning points).
- The delays in the availability of statistical information pertaining to consumption in QNA and indicators conditions this assessment.
1. Motivation: traditional vs. “new” indicators

- There is more data around than what is typically exploited in many short term forecasting models

**Target variable:**
QNA private consumption

```
"Hard" indicators

- Available at the monthly frequency … but released with significant lags

"Soft" indicators

- More timely
- Typically more correlated with hard indicators than with QNA variables
- Contemporaneous, leading or lagging behavior?

Other (less standard) sources?

- Credit cards’ spending / withdrawals
- Indicators of “uncertainty”
- Google searches
- ...```
1. Motivation

- There is more data around than what is typically exploited in many short term forecasting models

**AIM OF THE PAPER:**

Explore the relative merits of...

... hard vs. soft

... traditional vs. “new” indicators

to “nowcast” Spanish real private consumption
Outline

1. Motivation
2. Literature review
3. The data
4. Modeling approach
5. The empirical exercise
6. Selection of results and conclusions
2. Literature review

- Traditionally, the literature on “nowcasting” has been quite focused on GDP.

- Some exceptions: papers in which GDP is modelled together with its demand and/or supply components:
  
  - For example, Burriel and García-Belmonte (2013) for the euro area or Arencibia, Gómez-Loscos, de Luis and Pérez-Quirós (2017) for the case of Spain.
  
  - The latter paper includes a block for QNA private consumption in which a number of hard/soft traditional indicators and credit cards’ transactions are included.

- More recently, the literature has started to explore “new” sources:
  
  - Not so much for the case of private consumption.
2. Literature review

- More recently, the literature has started to explore “new” sources

  ✔ GOOGLE SEARCHES

  Camacho and Pacce (2016): application to tourism – including these indicators improves forecasting performance over models in which they are omitted

  Artola, Pinto and de Pedraza-García (2015): application to tourism / useful

  Bortoli and Combes (2015): general economic activity / limited usefulness

  Vosen and Schmidt (2011, 2012): private consumption/ useful compared to soft

  ✔ ATM/POS DATA

  ✔ UNCERTAINTY MEASURES
2. Literature review

More recently, the literature has started to explore “new” sources

- **GOOGLE SEARCHES**

- **ATM/POS DATA**
  
  Duarte, Rodrigues and Rua (2016): private consumption
  
  Galbraith and Tkaz (2007): economic activity, consumption (non-durable)
  
  Gill, Perera and Sunner (2012): private spending

- **UNCERTAINTY MEASURES**
2. Literature review

- More recently, the literature has started to explore “new” sources

  ✓ GOOGLE SEARCHES

  ✓ ATM/POS DATA

  ✓ UNCERTAINTY MEASURES

  Conceptually, they should help

  Gil, Pérez, Urtasun (2017): indicators of uncertainty (in particular, “financial”) influence consumption and investment, in the case of Spain

Outline

1. Motivation
2. Literature review
3. The data
4. Modeling approach
5. The empirical exercise
6. Selection of results and conclusions
3. The data

- Quarterly private consumption,
  monthly indicators
  *(lower frequencies not exploited yet)*

- Quarterly private consumption
- Social sec. registrations
  - Retail trade index
  - Services activity index
  - PMI – services
  - Consumers confidence index
- Credit cards
  - Cash
  - Cash and cash equivalent
- Disagreement consumption
  - Disagreement about inflation
  - Uncertainty on unemployment
  - Economic Policy Uncertainty
    - IBEX-35 volatility index
- Google Trends
- “Traditional” indicators
  - ATM/POS
- Uncertainty
- Google search queries
3. The data: ATM/POS

- Widespread use of electronic payment systems
- Timeliness [daily/weekly frequency, in theory]
- Credit cards: payments by means of credit/debit cards in points of sales (in real euros) [seasonally adjusted and deflated by national CPI]
- Other options: Cash withdrawals, Cash and equivalents
3. The data: Google Trends

- Some evidence of usefulness in the literature (not robust) – risk of data mining
- Households use internet to buy goods and services: info easily available.
- “Google Trends” provides an index of the relative volume of search queries conducted through Google (daily/weekly)
  - It provides aggregated indexes of search queries which are classified into categories and sub-categories using an automated classification engine
  - We select consumption-relevant categories (~60) that match the product categories of personal consumption expenditures of the BEA's national income and product accounts
    - **Example:**
      | Classification by national product and income | Google categories |
      |------------------------------------------------|-------------------|
      | Durable goods                                   |                   |
      | Motor vehicles and parts                        |                   |
      | Furnishing and durable household equipment      |                   |
      | Recreational goods and vehicles                |                   |
      | Other durable goods                             |                   |
      | Automotive, auto financing, automotive parts, auto insurance, Seat, Mercedes Benz, Mercedes offer, second hand car, car, to buy a car |
      | Electrical appliance, home insurance, home remodel, home furnishing, interior decoration, interior design |
      | Online movie, to buy a movie, watch online movie, video games |
      | Telecommunications, router wifi, mobile phone, electronic book, novel |
  - Data available since January 2004, not seasonally adjusted [→ TRAMO-SEATS]
  - Distinguish durable/ non durable/ services
  - “Aggregation”: (i) Principal Components Analysis (literature); (ii) NA weights
3. The data: Uncertainty

- **Economic Policy Uncertainty Index (EPU)** (Baker, Bloom, Davis, 2016): it measures the frequency of news related to economic policy uncertainty in two of the most popular Spanish newspapers.

- **European Commission Business and Consumer Surveys**: “unemployment expectations for the next 12 months”, indicator computed as

\[ \sqrt{\text{Frac}_t^+ + \text{Frac}_t^- - (\text{Frac}_t^+ - \text{Frac}_t^-)^2} \]

where \( \text{Frac}^+/^- \) is the weighted fraction of consumers in the cross section with increase/decrease responses at time \( t \).

- **Indicators of disagreement about consumption and inflation forecasts**, calculated using the information provided by the FUNCAS panel of forecasters. At each point in time, this measure is computed as the standard deviation of such cross-section of forecasters

\[ \frac{1}{n} \sum_{i=1}^{n} (\hat{c}_i - \hat{c}_A)^2 \]
3. The data: Uncertainty

- Private consumption (NA)
- Disagreement about consumption forecasts (level, rhs)
- Uncertainty about unemployment expectations (level, rhs)
- Economic policy uncertainty index (EPU) (level, rhs)
- IBEX-35 volatility (level, rhs)
3. The data

- Preliminary exploration: regressions of $\Delta \log C$ on its lag and different indicators

<table>
<thead>
<tr>
<th>p-values</th>
<th>Dependent variable: consumption q-o-q growth rate</th>
<th>p-values</th>
<th>Dependent variable: consumption q-o-q growth rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eq1</td>
<td>0.00</td>
<td>Eq1</td>
<td>0.51</td>
</tr>
<tr>
<td>Eq2</td>
<td>0.00</td>
<td>Eq2</td>
<td>0.50</td>
</tr>
<tr>
<td>Eq3</td>
<td>0.01</td>
<td>Eq3</td>
<td>0.47</td>
</tr>
<tr>
<td>Eq4</td>
<td>0.00</td>
<td>Eq4</td>
<td>0.52</td>
</tr>
<tr>
<td>Eq5</td>
<td>0.01</td>
<td>Eq5</td>
<td>0.64</td>
</tr>
<tr>
<td>Eq6</td>
<td>0.00</td>
<td>Eq6</td>
<td>0.38</td>
</tr>
</tbody>
</table>

| Lag      | 0.00                                          | Google Trends (NA weights) | 0.02                                           |
| Google Trends (NA weights) | 0.00 | Google Trends (NA weights) | 0.75 |
| Google Trends Durables (NA weights) | 0.00 | Google Trends Durables (NA weights) | 0.37 |
| Google Trends 1st principal | 0.00 | Google Trends 1st principal | 0.83 |
| Credit cards | 0.00 | Credit cards | 0.00 |
| EPU      | 0.74                                          | EPU      | 0.04                                          |
| Social security registrations | 0.45 | Social security registrations | 0.00 |
| $R^2$    | 0.45                                          | $R^2$    | 0.71                                          |
|          | 0.50                                          |          | 0.70                                          |
|          | 0.56                                          |          | 0.70                                          |
|          | 0.54                                          |          | 0.70                                          |
|          | 0.69                                          |          | 0.76                                          |
|          | 0.42                                          |          | 0.72                                          |
Outline

1. Motivation
2. Literature review
3. The data
4. Modeling approach
5. The empirical exercise
6. Selection of results and conclusions
4. Modelling approach

- Different sampling frequency: monthly (indicators), quarterly (consumption)
- Publication delays cause missing values for some of the variables at the end of the sample ("ragged-end" problem)
4. Modelling approach

- Mixed-frequencies models, in the vein of Harvey and Chun (2000)
  - Multivariate setup: Unobserved Components Model
  - *Flexibility: aggregation and modelling using the State Space representation*
  - *Models in levels: no need to worry about ex ante stationarity or cointegration*
  - Seemingly Unrelated Structural Time Series Models (SUTSE)
  - Different sampling intervals: cumulator variable
  - Optimal interpolation using the Kalman Filter and the Smoothing Algorithm
  - Robust estimation
    - *Use of last estimation as initial condition up to the moment when there is no estimation improvement*
    - *Estimate using random points around the solution to ensure that we end up in the same solution*
4. Modelling approach

The basic model is of the Unobserved Component Model class known as the Basic Structural Model (Harvey 1989), that decomposes a set of time series in unobserved though meaningful components from an economic point of view (mainly trend, seasonal and irregular). The model is multivariate, and may be written as

\[
\begin{bmatrix}
    z_t \\
    u_t
\end{bmatrix} = T_t + S_t + \epsilon_t.
\]

State-Space representation

Trend: SRW (no seasonal)

\[
x_t = \Phi x_{t-1} + \epsilon w_t
\]

\[
\begin{bmatrix}
    z_t \\
    u_t
\end{bmatrix} = \begin{bmatrix}
    H \\
    H^u
\end{bmatrix} x_t + \begin{bmatrix}
    \epsilon_t \\
    \nu_t
\end{bmatrix}
\]

where

\[
w_t \sim N(0, \Sigma_{w_t}), \quad \epsilon_t \sim N(0, \Sigma_{\epsilon}),
\]

\[
\nu_t \sim N(0, \Sigma_{\nu_t}).
\]

Aggregation

Cumulator variable

\[
C_t = \begin{cases}
0, & t = \text{every January (monthly data)}/ \\
1, & \text{first quarter (quarterly data)}/ \\
\text{otherwise}.
\end{cases}
\]

Thus, the model turns out to be:

\[
\begin{bmatrix}
    z_t \\
    x_t
\end{bmatrix} = \begin{bmatrix} C_t \otimes I & H \Phi \\
    0 & \Phi
\end{bmatrix} \begin{bmatrix}
    z_{t-1} \\
    x_{t-1}
\end{bmatrix} + \begin{bmatrix}
    1 & H \\
    0 & E
\end{bmatrix} \begin{bmatrix}
    \epsilon_t \\
    \nu_t
\end{bmatrix}
\]

\[
\begin{bmatrix}
    z_t \\
    u_t
\end{bmatrix} = \begin{bmatrix}
    I & 0 \\
    0 & H^u
\end{bmatrix} \begin{bmatrix}
    z_t \\
    x_t
\end{bmatrix} + \begin{bmatrix}
    0 \\
    I
\end{bmatrix} \nu_t.
\]
Outline

1. Motivation
2. Literature review
3. The data
4. Modeling approach
5. The empirical exercise
6. Selection of results and conclusions
5. The empirical exercise

- Real-time database
- 3 different forecast origins (information sets) within each quarter: m1, m2, m3
- Out-of-sample evaluation over 2008Q1-2016Q4
## 4. Empirical exercise: database (m1)

<table>
<thead>
<tr>
<th>INFORMATION AVAILABLE AT NOWCASTING TIME (m1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous quarter</td>
</tr>
<tr>
<td>1st month</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Private consumption (QNA)</td>
</tr>
<tr>
<td>Social security registrations</td>
</tr>
<tr>
<td>Retail trade index</td>
</tr>
<tr>
<td>Services activity index</td>
</tr>
<tr>
<td>PMI. Services</td>
</tr>
<tr>
<td>Credit cards</td>
</tr>
<tr>
<td>Cash</td>
</tr>
<tr>
<td>Cash and cash equivalents</td>
</tr>
<tr>
<td>Disagreement about private consumption</td>
</tr>
<tr>
<td>Disagreement about inflation</td>
</tr>
<tr>
<td>Uncertainty about unemployment expectations</td>
</tr>
<tr>
<td>Economic policy uncertainty index (EPU)</td>
</tr>
<tr>
<td>Consumers confidence index</td>
</tr>
<tr>
<td>IBEX-35 volatility</td>
</tr>
<tr>
<td>Google Trends (1st principal component)</td>
</tr>
</tbody>
</table>

Previous quarter: 1st, 2nd, 3rd month
Current quarter: 1st, 2nd, 3rd month

- **Private consumption (QNA)**
- **Social security registrations**
- **Retail trade index**
- **Services activity index**
- **PMI. Services**
- **Credit cards**
- **Cash**
- **Cash and cash equivalents**
- **Disagreement about private consumption**
- **Disagreement about inflation**
- **Uncertainty about unemployment expectations**
- **Economic policy uncertainty index (EPU)**
- **Consumers confidence index**
- **IBEX-35 volatility**
- **Google Trends (1st principal component)**
4. Empirical exercise: database (m2)

<table>
<thead>
<tr>
<th>INFORMATION AVAILABLE AT NOWCASTING TIME (m2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous quarter</td>
</tr>
<tr>
<td>1st month</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Private consumption (QNA)</td>
</tr>
<tr>
<td>Social security registrations</td>
</tr>
<tr>
<td>Retail trade index</td>
</tr>
<tr>
<td>Services activity index</td>
</tr>
<tr>
<td>PMI. Services</td>
</tr>
<tr>
<td>Credit cards</td>
</tr>
<tr>
<td>Cash</td>
</tr>
<tr>
<td>Cash and cash equivalents</td>
</tr>
<tr>
<td>Disagreement about private consumption</td>
</tr>
<tr>
<td>Disagreement about inflation</td>
</tr>
<tr>
<td>Uncertainty about unemployment expectations</td>
</tr>
<tr>
<td>Economic policy uncertainty index (EPU)</td>
</tr>
<tr>
<td>Consumers confidence index</td>
</tr>
<tr>
<td>IBEX-35 volatility</td>
</tr>
<tr>
<td>Google Trends (1st principal component)</td>
</tr>
</tbody>
</table>
### 4. Empirical exercise: database (m3)

#### INFORMATION AVAILABLE AT NOWCASTING TIME (m3)

<table>
<thead>
<tr>
<th></th>
<th>Previous quarter</th>
<th>Current quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st month</td>
<td>2nd month</td>
</tr>
<tr>
<td><strong>Private consumption (QNA)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Social security registrations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Retail trade index</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Services activity index</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PMI. Services</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Credit cards</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cash</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cash and cash equivalents</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disagreement about private consumption</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disagreement about inflation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Uncertainty about unemployment expectations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Economic policy uncertainty index (EPU)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Consumers confidence index</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IBEX-35 volatility</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Google Trends (1st principal component)</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Outline

1. Motivation
2. Literature review
3. The data
4. Modeling approach
5. The empirical exercise
6. Selection of results and conclusions
6. Selection of results and conclusions

I. Selection within the groups of “new indicators”

Credit cards:

Credit Card (number) & Hard & Soft
Hard
Credit Card (value)
Credit Card (number) & Hard & Soft
Hard
Credit Card (number)
6. Selection of results and conclusions

I. Selection within the groups of “new indicators”

Uncertainty:

RMSE (relative to a RW)
6. Selection of results and conclusions

I. Selection within the groups of “new indicators”

Google Trends:

![Graph showing RMSE (relative to a RW) for m1, m2, and m3 with different combinations of Google (total), Google (components), Hard, and Hard&Soft categories.](image)
II. “Hard” versus “Soft”

Trade off between timeliness and accuracy

*Hard: released with lag*

*Soft: more timely; typically more correlated with “Hard” than with QNA var.*

![Graph showing RMSE (relative to a RW)]
III. Do “new” indicators improve the forecast performance of models that also include the “traditional” ones?

The curse of dimensionality
III. Do “new” indicators improve the forecast performance of models that also include the “traditional” ones?

Changes in the direction of growth

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>83.3%</td>
<td>77.8%</td>
<td>80.6%</td>
<td>66.7%</td>
<td>83.3%</td>
<td>83.3%</td>
<td>77.8%</td>
<td>77.8%</td>
<td>75.0%</td>
<td>72.2%</td>
<td>69.4%</td>
<td>72.2%</td>
</tr>
<tr>
<td>m2</td>
<td>86.1%</td>
<td>69.4%</td>
<td>75.0%</td>
<td>83.3%</td>
<td>77.8%</td>
<td>80.6%</td>
<td>75.0%</td>
<td>77.8%</td>
<td>80.6%</td>
<td>75.0%</td>
<td>83.3%</td>
<td>77.8%</td>
</tr>
<tr>
<td>m3</td>
<td>75.0%</td>
<td>69.4%</td>
<td>72.2%</td>
<td>69.4%</td>
<td>75.0%</td>
<td>80.6%</td>
<td>72.2%</td>
<td>75.0%</td>
<td>86.1%</td>
<td>75.0%</td>
<td>83.3%</td>
<td>86.1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Credit Cards (value)</th>
<th>Disagreement</th>
<th>Volatility</th>
<th>EPU</th>
<th>Google (total)</th>
<th>Google (components)</th>
<th>Hard &amp; Google (total)</th>
<th>Hard &amp; Google (components)</th>
<th>Hard &amp; Soft &amp; Google (components)</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>75.0%</td>
<td>63.9%</td>
<td>69.4%</td>
<td>61.1%</td>
<td>61.1%</td>
<td>63.9%</td>
<td>63.9%</td>
<td>80.6%</td>
<td>77.8%</td>
</tr>
<tr>
<td>m2</td>
<td>77.8%</td>
<td>72.2%</td>
<td>77.8%</td>
<td>75.0%</td>
<td>69.4%</td>
<td>77.8%</td>
<td>72.2%</td>
<td>77.8%</td>
<td>77.8%</td>
</tr>
<tr>
<td>m3</td>
<td>72.2%</td>
<td>75.0%</td>
<td>77.8%</td>
<td>75.0%</td>
<td>69.4%</td>
<td>77.8%</td>
<td>72.2%</td>
<td>72.2%</td>
<td>72.2%</td>
</tr>
</tbody>
</table>
6. Selection of results and conclusions

- **Summing up:**
  
  ✓ Most of the models outperformed the random walk
  
  ✓ “Hard” versus “Soft”: *There are some gains when combining.*
  
  ✓ “New” indicators:
    - *Credit cards*
    - *Uncertainty*
    - *Google Trends*
THANKS FOR YOUR ATTENTION
3. The data: additional slides
7. Further work

✓ Event analysis: other uses of “Google Trends”
Event studies: some illustrations

- Forecasters (and policy-makers) are occasionally confronted with unusual events (policy measures) that have the potential of affecting economic activity.

- The usual monthly-based indicators are in this case of little use, as the data are available only with a substantial lag.

- To assess the impact of these events on the economy, forecasters need to monitor indicators that are released on a timely basis: new high-frequency data (daily) sources that could provide accurate and timely information.

- Here we show that weekly google search data can be used to analyze the impact on consumer expenditure of different events (terrorist attacks, bankruptcies, political speeches etc.).

- We expect that the adjustment in the consumption of durables may be sharper after a VAT rate increase (2012).
Event studies: some illustrations

- Before the increase in VAT rates there was an increase in the consumption of durable goods (rational response to the expected increase in prices).
- This effect was transitory and was followed by a decline in consumption.
- In the case of the retail trade index you could see these effects with a lag of two months.
Event studies: some illustrations

- Although in this paper we are considering mainly monthly data, higher frequency data have also valuable information that can be used in the assessment of the response of consumption to a shock.