Using the payment system data to forecast the Italian GDP

by Valentina Aprigliano, Guerino Ardizzi and Libero Monteforte
Temi di discussione

(Working papers)

Using the payment system data to forecast the Italian GDP

by Valentina Aprigliano, Guerino Ardizzi and Libero Monteforte

Number 1098 - February 2017
The purpose of the Temi di discussione series is to promote the circulation of working papers prepared within the Bank of Italy or presented in Bank seminars by outside economists with the aim of stimulating comments and suggestions.

The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.


Editorial Assistants: Roberto Marano, Nicoletta Olivanti.

ISSN 1594-7939 (print)
ISSN 2281-3950 (online)

Printed by the Printing and Publishing Division of the Bank of Italy
USING THE PAYMENT SYSTEM DATA TO FORECAST THE ITALIAN GDP

by Valentina Aprigliano*, Guerino Ardizzi** and Libero Monteforte*

Abstract

Payment systems track economic transactions and therefore could be considered important indicators of economic activity. This paper describes the available monthly data on the retail settlement system for Italy and selects some of them for short-term forecasting. Using a mixed frequency factor model to predict Italian GDP, we find that payment system flows stand out when compared to other standard business cycle indicators.

JEL Classification: C53, E17, E27, E32, E37, E42.
Keywords: short term forecasting, LASSO, mixed frequency models, Kalman smoothing, payment systems, TARGET2.

Contents

1. Introduction ................................................................. 5
2. Overview of the payment system for the Italian economy ................................................ 7
3. Data ................................................................................................. 10
   3.1 Payments data and macroeconomic aggregates .......................................................... 10
   3.2 Selection of the targeted predictors ............................................................................ 11
4. The model ................................................................................................. 15
5. Forecasting application ...................................................................................... 17
6. Conclusions ................................................................................................. 23
References .............................................................................................................. 24

* Bank of Italy, Directorate General for Economics, Statistics and Research.
** Bank of Italy, Directorate General for Markets and Payment Systems.
1 Introduction*

Following the global recession the interest for new macroeconomic forecasting tools, especially those based on monetary and financial information, has been increasing. On the back of the developments of computational tools for storing and elaborating large-scale datasets, the analysts are focusing on the search for new timely and reliable information in order to improve the forecasting ability in real time.

Data on payment instruments (cheques, credit transfers, direct debits, payment cards) could represent a unique source of information for short-term forecasting of the economic activity, as they trace economic transactions. This link was already clear at the beginning of the last century, when Irving Fisher described the seminal equations of the quantitative theory of money writing: “such elementary equations mean that the money paid in any transaction is the equivalent of the goods bought at the price of sale” (Fisher, 1912). Moreover the importance of payments, jointly with banking and asset markets, to understand the functioning of a monetary economics has recently received a new attention following the body of research called New Monetarist Economics (see Williamson and Wright 2010 and Schneider and Piazzesi 2015). Nevertheless, the use of payment data for forecasting purposes has been exploited only recently. Studies for Canada (Galbraith and Tkacz, 2007), (Galbraith and Tkacz, 2009), Portugal (Esteves, 2009) and Denmark (Carlsgaard and Storgaard, 2010) find that payment transactions can help with nowcasting and forecasting private consumption (and thereby also GDP) on a short-term basis. It is worth noting that this literature concentrates on payment cards (i.e. ATM/POS transactions as in Esteves 2009 and Carlsgaard and Storgaard 2010) or on cheques (Galbraith and Tkacz 2007), thus offering only a partial view of the economic activity: payments with cards can be related to private consumption but they represent a small share of the value of payments, especially in Italy.

To the best of our knowledge this paper is the first attempt to assess the ability of retail payment data to forecast the short-term development of the Italian economic activity. Differently from the previous literature for other countries, we use a comprehensive set of payment instruments, such as credit transfers, cheques, direct debits and debit cards. This extension allows us to conduct a more robust forecasting exercise compared to similar studies referred to other countries, which mainly consider a subset of payment instruments (debit cards or cheques). Such data are recorded electronically through clearing and settlement circuits managed by the Bank of Italy (for retail transactions, BI-COMP; for wholesale payments including customer transactions, BI-REL up to May 2008 and TARGET2-Bank of Italy subsequently) and they are not revised because they are recorded without errors by construction. Another appreciable feature of these data is their timeliness. Indeed, the payment data are available on a daily basis with a short delay.

We show that there is a close correlation between the retail payment series (PS, onward) and the main macroeconomic aggregates, such as GDP and its main components. In the empirical

---

*This paper represents the authors personal opinions and does not reflect the view of the Bank of Italy. Preliminary draft, please do not quote or circulate.

We are grateful to Luca Arciero, Alessandro Barbarino, Claudio Impenna, Massimiliano Marcellino, Antonio Rua, Roberta Zizza, Francesco Zollino, for useful comments. We also thank the participants to the seminar at the Banca d’Italia and the partecipants to the CIRET Conference, Copenhagen 2016, Iwce Conference, Rome 2016, RCEA Time Series Econometrics Workshop, Rimini 2016.

1Our data are extracted by the payment system infrastructures, while in some structural analysis data are collected by means of surveys or diaries (see Bagnall et al. 2016).
application, we develop a mixed frequency dynamic factor model to predict the quarter-on-quarter growth of the GDP by using a large-scale monthly dataset, which includes standard business cycle indicators; we perform out of sample forecasting simulations, including and excluding the PS. The contribution of the PS to the forecasts turns out to be appreciable and promising.

The rest of the paper is organized as follows. Section 2 gives an overview of the payment systems in Italy. Section 3 introduces the data, providing some descriptive evidence on the relationship between PS and the main macroeconomic aggregates (3.1) and showing that the Least Absolute Shrinkage Selector Operator (LASSO, onward) picks the PS amongst the first 13 predictors (out of the 50 indicators we compiled) to be included in the model (3.2). In Section 4 we describe the mixed frequency dynamic factor model and in section 5 we look at the forecasting contribution of the PS. Section 6 draws the conclusions.
2 Overview of the payment system for the Italian economy

A payment system is the set of instruments, rules, procedures and technologies used to settle money transfers among economic agents.

We can distinguish between wholesale payments and retail payments. The former typically involve the banking system handling large-value payments (interbank transactions), more connected with financial markets flows and refinancing operations with national central banks; the latter refer to transactions within the circuit of individuals and firms and mostly related to the economic activity (speech by Padoa Schioppa, Seoul May 2004).²

Before the launch of the euro in 1999, the payment system in Europe was highly fragmented. The Monetary Union posed the problem of harmonizing the infrastructure to transfer money among economic operators to foster the financial and commercial integration. The broad-based reorganization of the payment system has been mostly effective for wholesale payments, operated by two area-wide systems: TARGET2 (Trans-European Automated Real-Time Gross Settlement Express Transfer System; T2 onward), provided by the Eurosystem, and EURO1, privately owned. The Bank of Italy, along with Deutsche Bundesbank and Banque de France, contributed to develop T2, which settles the most part of the wholesale transfers, on a gross-basis in real-time.

As for the retail payments,³ the system is not fully integrated yet. However, since 2014 the Single Euro Payments Area (SEPA, onward) has strongly fostered the standardization and interoperability among different national clearing and settlement retail systems. The Bank of Italy manages the BI-Comp clearing system, which works in compliance with the rules of SEPA. BI-Comp clears the domestic payments on a multilateral net basis. These payments can be settled both in BI-Comp and in TARGET2. In fact, due to urgency and security reasons, banks may prefer to settle customers’ payments in TARGET2. This retail branch of TARGET2 is named TARGET2-retail from here onward (T2-retail for the sake of brevity).

The retail non-cash payments settled through BI-Comp and T2-retail add up to 5 trillion € on a yearly basis (about 60% of the total value of retail payments in Italy; 80% if we consider only electronic payments, excluding postal pre-printed processed and other paper-based credit transfers), about three times higher than the nominal value of the GDP.⁴ Although in Italy cash payment is still the most frequently used instrument of retail payments (about 80% of the number of retail payments), new information and communication technologies fostered non-cash payments, mainly those processed through electronic devices, allowing more flexibility and customization. If

³The retail payments system generally uses the Clearing and Settlement mechanism (CSM), in which one or more operators perform clearing (i.e. transmission, matching, confirmation of payments and calculation of a settlement position) and settlement (completion of the payment).
⁴Unlike cash payments, which are immediate vis-a-vis transfers of value between the payer and the payee through banknotes and coins, non-cash payments are exchanges of funds through accounts. Therefore, the relationship between the payer and the payee is mediated by authorized institutions (such as banks, postal offices), which actually process the payment before settling the transaction. It follows a crucial distinction depending on the party submitting the payment order: the credit-based instruments (i.e. credit transfers, card payments), submitted by the payer, and the debit-based instruments (i.e. direct debits, cheques), submitted by the payee. Credit transfers are the most common credit-based instrument, while debit-based instruments include direct-debits, card payments and cheques. For a detailed description of payment instruments see http://www.bancaditalia.it/compiti/sispaga-mercati/strumenti-pagamento/index.html and Payments, securities and derivatives, and the role of the Eurosystem, by ECB.
we consider the value of the transactions, on annual average, the share of the cash payments shrinks to about 45% of the consumer-to-business transactions (for instance, point of sales purchases) and to less than 10% of all transactions, including business-to-business payments. Figure 1 shows the monthly gross flow of retail non-cash payments settled through BI-Comp and T2-retail in Italy and it also depicts the upward trend in the use of electronic payment instruments (credit transfer, direct debit, debit cards) against the downtrend of the cheques.

The strong decrease observed in 2014 for credit transfers and for direct debit payments stems from the changes in the customer payment landscape following the migration to SEPA. Some participants reconsidered the routing policies for their customer payments and they ultimately chose in favour of SEPA-compliant automated clearing houses other than BI-Comp and T2. However, T2-retail has been less affected, because it meets some specific customers’s demands concerning urgency and assurance of the payments. The peak reached by direct debits in 2003 is due to interbank procedural changes and routing policies after the launch of the new domestic real-time

---

5These data refer to 2014.
gross settlement system (the so called “New Birel”). As regards cheques payments, the strong increase in 2009 is due to changes to the interbank procedures and to truncation limits for low-value cheques. More generally, transactions by cheques have been declining over the years because of the increasing use of electronic payments and the adoption of anti-money laundering restrictions, which have strongly limited their negotiability between private operators. As regards transactions by payment cards, the upward jump on July 2001 was caused by the inclusion of point of sale payments (POS); the peak in 2015 is due to more recent regulatory changes in the card market and to the inclusion of postal debit card and automated teller machine (ATM) transactions in the BI-Comp system.

Figure 2 depicts the share of transactions enabled by different payment instruments on the total gross flows settled in BI-Comp. Credit transfers represent the largest share (almost 55%); also direct debits have a sizeable weight (29%) since they provide a very handy solution to manage recurrent payments (e.g. utility bills, mortgage payments). Payment cards are far less used, accounting for only 3% of the total non-cash payments; the agents use payment cards for low-value transactions (50-100 € for POS; 100 - 200 € for ATM).

Figure 2: Total payments settled in BI-COMP

However, a large share of credit card payments are recorded as direct debits, since credit card statements are often charged on the payer’s bank current account. Italy is lagging behind for non-cash payment instruments compared to the other euro area major countries. Although retail cash transactions are not recorded in the clearing and settlement systems, ATM operations settled through the interbank procedures in BI-COMP may be a useful proxy of the demand for cash because there is a strong correlation between the amount of cash acquired at the ATM and the cash used at the point of sale.

---

6For business purposes credit transfers are the most popular and suitable payment instrument indeed (they account for 80% of the total value of the business-to-business transactions; direct debits add up to 10% as well as cheques).
3 Data

3.1 Payments data and macroeconomic aggregates

In this section we glance at some descriptive statistics which prove the tight relationship between the main macroeconomic aggregates (such as GDP, value added in service sector, private consumption and gross fixed investments) and the payment flows. In our analysis, we focus on retail payment instruments since they are used to settle commercial transactions as consumer-to-business and business-to-business payments. We find a strong empirical evidence underpinning the role of payment time series in tracking the economic activity.

We rely on the systems BI-Comp and T2-retail to collect timely and high-frequency data. Data are collected by the central bank for the individual transactions of the payment systems. These systems close each day, therefore the observation error is basically nil. We do not use the individual payments (that could be eventually relevant for a big data analysis) but the monthly time series stored in the Banca d’Italia’s database, publicly available.\(^7\) The monthly observations are simply the sum of the nominal value of the individual payments recorded in the month.

Some preliminary adjustments of the time series are needed. Firstly, we take the year-on-year (y-o-y, onward) growth rate to remove the seasonal pattern characterizing the payment flows\(^8\) as well as the macroeconomic variables and the high-frequency volatility. The breaks observed in Figure 1 are smoothed out when we consider the aggregated flow of payments. Finally, we divide the nominal values by the gdp national account deflator in order to obtain a measure of the volume of transactions.

The annual growth of GDP and of its main components comove closely with the annual growth of payment flows settled through BI-Comp and T2-retail (see Figure 3). The big drop at the end of the sample proves that BI-Comp was more affected than T2-retail by the launch of SEPA as of October 2014.\(^9\) Some interesting results come out of the correlation matrix (see Table 1).\(^10\) The correlation between the PS and the target variables is valuable and similar to the one shown by other activity indicators such as industrial production and business confidence, usually adopted in short-term forecasting. Results in Table 1 are consistent with the picture of the payment instruments composition in Italy. Credit transfers and direct debits are the most widespread payment tools; they are also more correlated with target variables than payment cards and cheques.

---

\(^7\)Data are issued on Supplements to the Statistical Bulletin for Payment System, \url{https://www.bancaditalia.it/pubblicazioni/sistema-pagamenti/index.html}.

\(^8\)Money transfers typically peak during Christmas and summer holidays.

\(^9\)The launch of SEPA has entailed the switching of all domestic credit transfers and direct debits from BI-COMP into STEP2, which is the pan-European infrastructure managed by a private body (SIA-Società Interbancaria per l’Automazione). Information on payment flows in STEP2 are not publicly available.

\(^10\)We consider the contemporaneous correlation.
3.2 Selection of the targeted predictors

In this section, we use the Least Absolute Shrinkage and Selection Operator (LASSO) to make a first screening of the variables depending on their ability to anticipate the quarterly GDP growth (for a review of Lasso estimator see Tibshirani 1996 and Hastie et al. 2009; Bai and Ng 2008 and Bulligan et al. 2012 for the selection of targeted predictors to make forecasts).

We compiled $N = 50$ variables providing a quite complete picture of the economic activity. This large sample (N-sample, onward) included indicators of the industrial and service activity (industrial production, electricity consumption, freight truck, business confidence), of the households’ consumption (retail sales of goods and services, consumers’ confidence), financial indexes, credit flows to firms other than time series from the payment systems T2-retail and BI-Comp.

The monthly indicators are transformed in quarterly variables and then LASSO regression is
Table 1: Correlation between payment flows and macroeconomic indicators; y-o-y percentage changes

<table>
<thead>
<tr>
<th>Payment flows (^a)</th>
<th>GDP</th>
<th>Private Consumption</th>
<th>Gross Fixed Investment</th>
<th>Value added service sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheques</td>
<td>-44.4</td>
<td>-2.1</td>
<td>-37.2</td>
<td>-35.6</td>
</tr>
<tr>
<td>Cards(^b)</td>
<td>1.4</td>
<td>-12.3</td>
<td>16.4</td>
<td>4.0</td>
</tr>
<tr>
<td>Credit Transfer</td>
<td>59.1</td>
<td>61</td>
<td>65.3</td>
<td>59.5</td>
</tr>
<tr>
<td>Direct Debts</td>
<td>52.1</td>
<td>39.4</td>
<td>48.7</td>
<td>54.4</td>
</tr>
<tr>
<td>BI-Comp</td>
<td>71.8</td>
<td>70.5</td>
<td>74.3</td>
<td>68.4</td>
</tr>
<tr>
<td>T2-retail</td>
<td>79.4</td>
<td>67.7</td>
<td>64</td>
<td>74.3</td>
</tr>
<tr>
<td>BI-Comp + T2-retail</td>
<td>90.7</td>
<td>82.5</td>
<td>82.3</td>
<td>85.9</td>
</tr>
</tbody>
</table>

Other indicators

<table>
<thead>
<tr>
<th></th>
<th>GDP</th>
<th>Private Consumption</th>
<th>Gross Fixed Investment</th>
<th>Value added service sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>M2</td>
<td>-38.1</td>
<td>-11.5</td>
<td>-20.4</td>
<td>-45.6</td>
</tr>
<tr>
<td>Industrial production</td>
<td>92.8</td>
<td>68.7</td>
<td>74.3</td>
<td>79.8</td>
</tr>
<tr>
<td>Business Confidence(^c)</td>
<td>68.5</td>
<td>74.9</td>
<td>67.3</td>
<td>64.3</td>
</tr>
<tr>
<td>Consumer Confidence(^d)</td>
<td>9.9</td>
<td>40.0</td>
<td>14.3</td>
<td>19.7</td>
</tr>
</tbody>
</table>

\(^a\) Correlations are computed on the sample 2000.Q1 - 2012.Q4 in order to exclude the break caused by the new standard SEPA.

\(^b\) Debit card.

\(^c\) Economic Sentiment Indicator (ESI) provided by Istat.

\(^d\) Consumer confidence survey by Istat.

performed. LASSO selects the regressors by solving the minimization problem

$$
\min \left( \sum_{t=1}^{T} (y_t - \sum_{j=1}^{N} \beta_j X_{jt})^2 \right) \quad s.t. \sum_{j=1}^{N} |\beta_j| \leq t
$$

where \( t \) is the tuning parameter.

LASSO selects \( n_L \) targeted variables, among all \( N \) collected:

$$
\hat{L}_{n_L} = \{ j \in \{1, 2, \ldots, N\} : |\hat{\beta}_{Lj}| > 0 \}
$$

including T2-retail along with other business cycle indicators, such as industrial production, electricity consumption, freight truck, retail sales and many business and consumers confidence indicators (mainly referred to the evolution of the economic situation, saving opportunities and expected new orders).

The final model \( \hat{I}_n \supset \hat{L}_{n_L} \) includes some additional indicators earlier discarded by Lasso, which contribute to improve forecast accuracy.\(^{11}\) They account for concerns of households about labor market, firms’ assessment of liquidity constraints and of the current level of production and orders, in particular for the sectors producing intermediate and investment goods.

We implement a large-scale dynamic factor model on \( \hat{I}_n \), which is more suitable when the cross-section size of the dataset, \( n_t \), is valuable in respect of the number of observations, \( T \). Working with large datasets proves convenient to detect whether PS contribute to improve the accuracy

\(^{11}\) The additional indicators were selected by minimizing the RMSFE in an out-of-sample forecasting exercise.
in the estimation of the economic growth. Their role in forecasting stands out when compared to many other indicators, often used to track the short-term dynamics of the economy.

Table (2) shows a detailed description of the N-sample (first column), as well as of the source and treatment of the data (second and third column, respectively). The fourth and the last columns indicate the variables \( \hat{L}_n \) picked by Lasso and finally included in our information set, \( \hat{I}_n \).
<table>
<thead>
<tr>
<th>Indicators</th>
<th>Source</th>
<th>Treatment</th>
<th>Added to Lasso selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total electricity consumption</td>
<td>Italian Electrical Net Society</td>
<td>(1-L)logSA</td>
<td>yes</td>
</tr>
<tr>
<td>Industrial production</td>
<td>Istat</td>
<td>(1-L)log</td>
<td>yes</td>
</tr>
<tr>
<td>Business Climate</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>CCS – Future General Economic sit.</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>CCS – Future Personal Economic sit.</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>CCS – Unemployment exp.</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>CCS - Saving opportunities, next 12 mth</td>
<td>Istat</td>
<td>SA</td>
<td>yes</td>
</tr>
<tr>
<td>CCS - Households balance sheet</td>
<td>Istat</td>
<td>(1-L)</td>
<td>yes</td>
</tr>
<tr>
<td>CCS - Current saving opportunities</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Current level of orders (intermediate goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Current level of production (int. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Exp. level of production (int. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>yes</td>
</tr>
<tr>
<td>BCS - Exp. level of orders (int. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>yes</td>
</tr>
<tr>
<td>BCS - Future general economic sit. (int. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Exp. level of liquidity (int. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Current level of liquidity (int. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Current level of orders (investment goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Current level of production (inv. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Current level of liquidity (inv. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Exp. level of orders (inv. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Exp. level of production (inv. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>yes</td>
</tr>
<tr>
<td>BCS - Future General Economic sit. (inv. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Current level of orders (consumer goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Current level of production (cons. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Exp. level of orders (cons. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>yes</td>
</tr>
<tr>
<td>BCS - Exp. level of production (cons. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Future General Economic sit. (cons. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>yes</td>
</tr>
<tr>
<td>BCS - Exp. level of liquidity (cons. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BCS - Current level of liquidity (cons. goods)</td>
<td>Istat</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>Current accounts deposits (stock)</td>
<td>Bank of Italy</td>
<td>(1-L)logSA</td>
<td>no</td>
</tr>
<tr>
<td>Credit flows to firms(^b)</td>
<td>Bank of Italy</td>
<td>(1-L)logSA</td>
<td>no</td>
</tr>
<tr>
<td>HICP</td>
<td>Istat</td>
<td>(1-L)logSA</td>
<td>yes</td>
</tr>
<tr>
<td>FTSE Italy (Banks)</td>
<td>Datastream</td>
<td>(1-L)log</td>
<td>no</td>
</tr>
<tr>
<td>FTSE Italy (Insurance)</td>
<td>Datastream</td>
<td>(1-L)log</td>
<td>no</td>
</tr>
<tr>
<td>FTSE Italy (Transport)</td>
<td>Datastream</td>
<td>(1-L)log</td>
<td>no</td>
</tr>
<tr>
<td>PMI Services - Business activity</td>
<td>Markit</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>PMI Services - New business</td>
<td>Markit</td>
<td>SA</td>
<td>yes</td>
</tr>
<tr>
<td>PMI Manufacturing - Output</td>
<td>Markit</td>
<td>SA</td>
<td>yes</td>
</tr>
<tr>
<td>PMI Manufacturing - New orders</td>
<td>Markit</td>
<td>SA</td>
<td>yes</td>
</tr>
<tr>
<td>PMI Manufacturing - Employment</td>
<td>Markit</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>PMI Manufacturing - New export orders</td>
<td>Markit</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>Freight truck</td>
<td>ASPI(^c)</td>
<td>(1-L)logSA</td>
<td>no</td>
</tr>
<tr>
<td>Retail trade - goods</td>
<td>Concommercio(^d)</td>
<td>SA</td>
<td>yes</td>
</tr>
<tr>
<td>Retail trade - services</td>
<td>Concommercio</td>
<td>SA</td>
<td>no</td>
</tr>
<tr>
<td>BI-COMP</td>
<td>Bank of Italy</td>
<td>(1-L)logSA</td>
<td>no</td>
</tr>
<tr>
<td>TARGET RETAIL</td>
<td>Bank of Italy</td>
<td>(1-L)logSA</td>
<td>yes</td>
</tr>
<tr>
<td>Payment system-total(^e)</td>
<td>Bank of Italy</td>
<td>(1-L)logSA</td>
<td>no</td>
</tr>
</tbody>
</table>

\(^a\) CCS: consumer survey; BCS: business survey.
\(^b\) Up to 1 mln €; 1-5 years maturity.
\(^c\) Italian concessionaire for the construction and management of toll highways.
\(^d\) Italian General Confederation of Enterprises, Professions and Self-Employment.
\(^e\) The sum of BI-COMP and T2-retail.
4 The model

Let $X_t$ be the $n$-vector of observable monthly variables earlier selected. They are assumed to be driven by $q$ common factors, $f_t = [f_{1t}, \ldots, f_{qt}]'$, and by an idiosyncratic component, $\xi_t$. The $n \times q$ matrices $\Lambda_i$, for $i = 1 \ldots s$, are the common-factor loadings. The dynamic factor model

$$X_t = \Lambda_s(L)f_t + \xi_t$$

$$= \Lambda_0 f_t + \Lambda_1 f_{t-1} + \Lambda_2 f_{t-2} + \cdots + \Lambda_s f_{t-s} + \xi_t$$

(3)

can be mapped to the static model

$$X_t = DF_t + \xi_t$$

(4)

where $F_t = [f_t', f_{t-1}', \ldots, f_{t-s}']'$ has dimension $r = q(s+1)$ and can be represented as a VAR(1) process\(^{12}\)

$$F_t = \mu + \Psi_1 F_{t-1} + \cdots + \Psi_r F_{t-r} + u_t$$

(5)

As stated in Bai and Ng (2007), $F_t$ is driven by $q < r$ common shocks if $u_t = R \epsilon_t$, where $\epsilon_t$ is a $q$-vector of mutually orthogonal shocks and the $q \times r$-matrix $R$ has rank $q$, then $E(u_t'u_t) = \Sigma_u = R \Sigma_R R$ has reduced rank $q$. We estimate $R$ as in Marcellino and Schumacher (2010). Given the OLS estimate $\hat{\Sigma}_u$ and its eigenvalue decomposition $\text{MPM}'$, let $M_s$ be the $r \times q$-matrix of the first $q$ eigenvectors and $P_s$ the diagonal matrix of the corresponding eigenvalues, then $\hat{R} = M_s P_s^{-1/2}$. We follow Bai and Ng (2002) and Bai and Ng (2007) to set the number of static ($r = 4$) and of dynamic factors ($q = 1$) and then we estimate the factors’ space $\mathcal{G}(F_t) = \text{span}(F_{1t}, \ldots, F_{rt})$ by principal components extracted from the balanced monthly dataset.

The quarterly growth rate of the real GDP, $y_{tq}$ with $t_q$ labeled by a multiple of the last month of each quarter (i.e. $t_q = 3, 6, \ldots, 3[T/3]$), is projected by an unrestricted MIDAS model on the monthly information (see Ghysels et al. 2004 for an extensive treatment of MIDAS model; we also refer to Foroni et al. 2015 for the unrestricted MIDAS model). For each quarter $t_q$ we have $m = 3$ values of the monthly regressors, therefore we apply the $L^{1/m}$ operator to obtain regressors lagged by $j$ months with respect to the quarter. Put formally:

$$y_{tq} = c + \beta_0 F_{tq} + \beta_1 F_{tq-1/m} + \cdots + \beta_p F_{tq-p/m} + \epsilon_{tq}$$

$$= c + \beta(L^{1/m})F_{tq} + \epsilon_{tq}$$

(6)

where the loadings $\beta_j$, for $j = 1, \ldots, p$ have dimension $1 \times r$. Equations from (4) to (6) are cast in a state-space form. The state equation is

$$X_t = \begin{bmatrix} D & 0_{n \times r+1} \end{bmatrix} \begin{bmatrix} F_t^s \end{bmatrix} + \xi_t$$

(7)

where

$$F_t^s = \begin{cases} [F_t', F_{t-1}', \ldots, F_{t-l+1}']' & \text{if } l \geq p \\ [F_t', F_{t-1}', \ldots, F_{t-p+1}']' & \text{otherwise} \end{cases}$$

(8)

\(^{12}\)According to Akaike information criteria, we set the order $l$ of the VAR equal to 4.
and \( \xi_t \sim N(0_{n \times 1}, \Sigma_\xi) \). To include the quarterly GDP growth into the state-space framework we constructed a monthly series \( y_t \) where \( y_t \equiv y_{t_q} \) when \( t \equiv t_q \) and missing otherwise.

If \( l \geq p \) the transition equation is

\[
\begin{bmatrix}
I_{rl} & 0_{rl \times 1} \\
-\beta_0 & 0_{1 \times r(t-1)} & 1
\end{bmatrix}
\begin{bmatrix}
F_t^* \\
y_t
\end{bmatrix}
= 
\begin{bmatrix}
[\Psi_i] & 0_{rl \times 1} \\
0_{1 \times r(t-p)+1}
\end{bmatrix}
\begin{bmatrix}
F_{t-1}^* \\
y_{t-1}
\end{bmatrix}
+ 
\begin{bmatrix}
\mu \\
0_{r(l-1) \times 1}
\end{bmatrix}
+ 
\begin{bmatrix}
u_t \\
\epsilon_t
\end{bmatrix}
\]  

(9)

where

\[
[\Psi_i] = \begin{bmatrix}
\Psi_1 & \Psi_2 & \cdots & \Psi_l \\
I_r & 0_{r \times r(t-1)} \\
\vdots \\
0_{r \times r(t-2)} & I_r & 0_{r \times r}
\end{bmatrix}
\]

while if \( l < p \)

\[
\begin{bmatrix}
I_{rp} & 0_{rp \times 1} \\
\beta_0 & 0_{1 \times r(p-1)} & 1
\end{bmatrix}
\begin{bmatrix}
F_t^* \\
y_t
\end{bmatrix}
= 
\begin{bmatrix}
[\Psi_i] & 0_{rl \times (p-l)+1} \\
0_{r(p-l) \times r(p-l) + 1}
\end{bmatrix}
\begin{bmatrix}
F_{t-1}^* \\
y_{t-1}
\end{bmatrix}
+ 
\begin{bmatrix}
\mu \\
0_{rp \times 1}
\end{bmatrix}
+ 
\begin{bmatrix}
u_t \\
\epsilon_t
\end{bmatrix}
\]  

(10)

We implement the Kalman recursions (Kim and Nelson, 1999) to extract the smoothed state variable \( \mathbf{FF}_t = [F_t^* \ y_t]' \). Let \( T_business \) be the vintage of the monthly series and \( h \) the forecasting horizon in terms of quarters, then the smoothed state variable is the estimate at time \( t \), for \( t = 1, 2, \ldots, T_business + 3h \), based on information up to \( T = T_business + 3h + 1 \), i.e. \( FF_{0:T} \). This is very appreciable because we use the latest information to make inference on the dynamics of the state variable. In fact, when \( T = T_business + 1, \ldots, T_business + 3h \) all observations are missing and they are given no weight by the Kalman filter, which basically forecasts the factors as elaborated in Giannone et al. (2008). In particular, we are interested in the forecast of the GDP at time \( T_{q_v} + h \), where \( T_{q_v} \) is the quarter corresponding to the month \( T_business \):

\[
y_{T_{q_v} + h} = c + \beta (L^{1/m}) \mathbf{FF}_{q_v + h | T} + \epsilon_{T_{q_v} + h}
\]  

(11)

One of the most common remarks to factor models concerns the possibility to disentangle the contribution of the observable variables to the forecasts. The variance of \( y_{t_q} \) is explained by the common factors, as shown in (6), which are not definitely given an economic meaning. As showed in Koopman and Harvey (2003), the output of the Kalman recursions can be used to size the weights attached to the observable variables in \( \mathbf{X}_j \) when forecasting \( y_{t_q} \), for \( j = 1, \ldots, T_business + 3h + 1 \). The smoothed state vector can be expressed as the weighted sum:

\[
\mathbf{FF}_{0:T} = \sum_{j=1}^{T} w_j(\mathbf{FF}_{t | T}) \mathbf{X}_j
\]  

(12)

where the weights \( w_j(\cdot) \) are function of the state vector. Each month \( t \), the weights are computed by the backward recursions (for \( j = t-1, t-2, \ldots, 1 \)) and the forward recursions (for \( j = t, t+1, \ldots, T \)) introduced in Koopman and Harvey (2003).\textsuperscript{13}

\textsuperscript{13}For a detailed description of the algorithms refer to Koopman and Harvey (2003), p. 1322.
5 Forecasting application

The latest available vintage for the empirical application ranges from January 2000 to November 2015. Because of the asynchronous release of the indicators, the dataset has a typical \textit{ragged-edge} structure and it is balanced on August 2015.

LASSO selected T2-retail among the first 13 most informative variables to fit the quarterly GDP growth, over the 50 indicators early collected. This result is a first evidence of the importance of the PS series to track the economic activity.

In Table 3 we assess how the forecasting performance of our model changes depending on whether we include or exclude the PS from the set of regressors. More precisely, we compare the benchmark model ($\hat{I}$), including all the variables listed in the fourth column of Table 2, with the model excluding the PS ($\hat{I}_{-PS}$) and with the model replacing PS with the first variable discarded by LASSO ($\hat{I}_{-PS,CLIMATE}$), i.e. Consumer survey – future economic climate (CLIMATE). We simulated the out-of-sample forecasts and we calculated the root mean squared forecast errors (RMSFE, onward) on two samples: the first one from 2008.Q2 until 2015.Q2; the second one from the bout of the sovereign-debt crisis in 2011.Q3 to 2015.Q2. The simulation was run in \textit{pseudo} real-time, since we use the latest available vintage of data and we cut it period by period being careful to replicate the missing values’ pattern at the end of the sample. The size of the estimation sample is fixed and it includes 78 months. The forecasting horizon is $h = -1, 0, 1$, depending on which month of the quarter is taken as the \textit{pseudo} vintage. For instance, in the first month of the quarter $t_q$, we backcast ($h = -1$) GDP of the previous quarter not yet released by the national statistical office (Istat),\footnote{The preliminary estimate of the Italian GDP is available 45 days later than the reference quarter.} other than nowcasting ($h = 0$) and forecasting 1 quarter ahead ($h = 1$).

<table>
<thead>
<tr>
<th>Sample</th>
<th>RMSFE</th>
<th>Backcast</th>
<th>Nowcast</th>
<th>Forecast 1-step</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample [2008.Q2 - 2015.Q2]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{I}$</td>
<td>0.910</td>
<td>0.985</td>
<td>0.991</td>
<td></td>
</tr>
<tr>
<td>$\hat{I}_{-PS}$</td>
<td>0.911</td>
<td>1.028</td>
<td>1.038</td>
<td></td>
</tr>
<tr>
<td>$\hat{I}_{-PS,CLIMATE}$</td>
<td>0.923</td>
<td>1.062</td>
<td>1.053</td>
<td></td>
</tr>
<tr>
<td><strong>Sample [2011.Q3 - 2015.Q2]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{I}$</td>
<td>0.517</td>
<td>0.529</td>
<td>0.544</td>
<td></td>
</tr>
<tr>
<td>$\hat{I}_{-PS}$</td>
<td>0.528</td>
<td>0.543</td>
<td>0.572</td>
<td></td>
</tr>
<tr>
<td>$\hat{I}_{-PS,CLIMATE}$</td>
<td>0.646</td>
<td>0.635</td>
<td>0.584</td>
<td></td>
</tr>
</tbody>
</table>

The PS improve the forecasting accuracy broadly. If we throw out only T2-retail the RMSFE rises throughout the forecasting horizons in both samples and in both competing models, $\hat{I}_{-PS}$ and $\hat{I}_{-PS,CLIMATE}$. This is a first rough proof of the relevance of the PS to make forecast. This result is even more remarkable if we consider that within a factor-model framework the marginal
contribution of the single indicator to the covariance of the common components typically fades as the cross-section dimension of the dataset becomes sizeable.

In order to disentangle the predictive contribution of the single indicator, we estimate the weights in (12), which are implicitly assigned to the observations when estimating the smoothed vector \( \mathbf{FF}_{t|T} \). Figures from 4 to 6 show the weights attached to T2-retail (red bars) and to some other indicators (blue bars), which track very well the short-term evolution of the economic activity (as electricity consumption, industrial production index, freight truck, manufacturing PMI - output component, expectations about liquidity of the manufacturing firms producing intermediate goods). Let us assume that the current month is October 2015. Therefore, we need to anticipate the GDP growth on 2015.Q3 (backcast; Figure 4), 2015.Q4 (nowcast; Figure 5) and 2016.Q1 (one-step-ahead forecast; Figure 6).

Let us consider the case of nowcasting (2015.Q4). In general, we observe that past observations contribute less to the forecasts, as expected; since November 2015 onward the weights are exactly zero because no observation is available yet. The weights of the PS are comparable to those of electricity consumption and liquidity expectations throughout the forecasting horizons considered. PS have sizeable weights compared to those of the manufacturing PMI (output component). This means that past information on PS contributes more than that on the PMI to anticipate the GDP growth. When forecasting 1-step ahead, the latest available data on PS weight as much as those on the freight traffic and they contribute to make up for the missing information on industrial production at the end of the sample.
Figure 4: Weights for backcasting (PS red bar; other variable blue bar)
Figure 5: Weights for nowcasting (PS red bar; other variable blue bar)
Figure 6: Weights for 1-step ahead forecasting (PS red bar; other variable blue bar)
Figure 7: Weights for 2-steps ahead forecasting (PS red bar; other variable blue bar)
6 Conclusions

Our findings show that payment data help with tracking the economic activity in Italy. We look at different aggregates of payment system flows in Italy, jointly with other indicators usually adopted in GDP forecasting, and we see that they maintain some additional information content. We start from a large database of short term monthly indicators and Lasso selects a series of retail payment, jointly with other standard business cycle indicators such as industrial production and business surveys. Moreover, an out of sample forecasting application using a mixed frequencies MIDAS model shows that the model including retail payments flows outperforms the one based on standard short-term indicators only, in terms of forecasting accuracy. In order to disentangle the contribution of the observable variables to the forecasts of GDP, we estimate the weights proposed in Koopman and Harvey (2003) and we find that the weights attached to PS are comparable to those of some of the most important short-term indicators generally used to track the economic activity.

Our analysis is based on pseudo real-time data, but given that payment system data are available with a short delay and are not affected by revisions this result looks relevant also for real time forecasting.

The payment data indicator used in this model is the aggregate of a huge number of transactions. The forecasting ability of this indicator paves the way for a future work exploring the big data base of the individual transactions.
References


RECENTLY PUBLISHED “TEMI” (*)

N. 1073 – *Search costs and the severity of adverse selection*, by Francesco Palazzo (July 2016).


N. 1075 – *Quantifying the productivity effects of global sourcing*, by Sara Formai and Filippo Vergara Caffarelli (July 2016).


N. 1079 – *Parents, schools and human capital differences across countries*, by Marta De Philippis and Federico Rossi (September 2016).

N. 1080 – *Self-fulfilling deflations*, by Roberto Piazza, (September 2016).

N. 1081 – *Dealing with student heterogeneity: curriculum implementation strategies and student achievement*, by Rosario Maria Ballatore and Paolo Sestito, (September 2016).


N. 1083 – *BTP futures and cash relationships: a high frequency data analysis*, by Onofrio Panzarino, Francesco Potente and Alfonso Puorro, (September 2016).

N. 1084 – *Women at work: the impact of welfare and fiscal policies in a dynamic labor supply model*, by Maria Rosaria Marino, Marzia Romanelli and Martino Tasso, (September 2016).

N. 1085 – *Foreign ownership and performance: evidence from a panel of Italian firms*, by Chiara Bentivogli and Litterio Mirenda (October 2016).

N. 1086 – *Should I stay or should I go? Firms’ mobility across banks in the aftermath of financial turmoil*, by Davide Arnaudo, Giacinto Micucci, Massimiliano Rigon and Paola Rossi (October 2016).

N. 1087 – *Housing and credit markets in Italy in times of crisis*, by Michele Loberto and Francesco Zollino (October 2016).

N. 1088 – *Search peer monitoring via loss mutualization*, by Francesco Palazzo (October 2016).

N. 1089 – *Non-standard monetary policy, asset prices and macroprudential policy in a monetary union*, by Lorenzo Burlon, Andrea Gerali, Alessandro Notarpietro and Massimiliano Pisani (October 2016).

N. 1090 – *Does credit scoring improve the selection of borrowers and credit quality?*, by Giorgio Albareto, Roberto Felici and Enrico Sette (October 2016).

N. 1091 – *Asymmetric information and the securitization of SME loans*, by Ugo Albertazzi, Margherita Bottero, Leonardo Gambacorta and Steven Ongena (December 2016).

N. 1092 – *Copula-based random effects models for clustered data*, by Santiago Pereda Fernández (December 2016).

N. 1093 – *Structural transformation and allocation efficiency in China and India*, by Enrica Di Stefano and Daniela Marconi (December 2016).

N. 1094 – *The bank lending channel of conventional and unconventional monetary policy*, by Ugo Albertazzi, Andrea Nobili and Federico M. Signoretti (December 2016).

N. 1095 – *Household debt and income inequality: evidence from Italian survey data*, by David Loschiavo (December 2016).

(*) Requests for copies should be sent to:


FANTINO D., A. MORI and D. SCALISE, *Collaboration between firms and universities in Italy: the role of a firm’s proximity to top-rated departments*, Rivista Italiana degli economisti, v. 1, 2, pp. 219-251, TD No. 884 (October 2012).


BARONE G. and S. MOCETTI, Inequality and trust: new evidence from panel data, Economic Inquiry, v. 54, pp. 794-809, TD No. 973 (October 2014).


FORTHCOMING


DE BLASIO G. and S. POY, The impact of local minimum wages on employment: evidence from Italy in the 1950s, Regional Science and Urban Economics, TD No. 953 (March 2014).

FEDERICO S. and E. TOSTI, Exporters and importers of services: firm-level evidence on Italy, The World Economy, TD No. 877 (September 2012).


RIGGI M., Capital destruction, jobless recoveries, and the discipline device role of unemployment, Macroeconomic Dynamics, TD No. 871 July 2012).
