

Using Payment System Data to Forecast Economic Activity*

Valentina Aprigliano, Guerino Ardizzi, and Libero Monteforte
Banca d'Italia

Payment systems trace economic transactions; they could therefore be considered important indicators of economic activity. This paper describes the monthly data available on Italy's retail settlement system and selects some of them for nowcasting and short-term forecasting. Using a mixed-frequency factor model based on a large-scale data set to predict Italian GDP and its main components, the contribution of payment system flows to improving forecasting accuracy is found to be non-negligible. Moreover, the timeliness of the data improves nowcasting accuracy throughout the quarter.

JEL Codes: C53, E17, E27, E32, E37, E42.

1. Introduction

Ever since the global recession, interest in 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 data sets, analysts are focusing on the pursuit of new, timely, and reliable information in order to improve the forecasting ability in real time.

Data on payment instruments (checks, credit transfers, direct debits, payment cards) could represent a unique source of information for short-term forecasting of 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

*This paper represents the authors' personal opinions and does not reflect the views of the Bank of Italy. We are grateful to Luca Arciero, Alessandro Barbarino, Claudio Impenna, Massimiliano Marcellino, Antonio Rua, Roberta Zizza, Francesco Zollino, and two anonymous referees for their useful comments. We also thank the participants who attended the CIRET Conference, Copenhagen 2016, RCEA Workshop, Rimini 2016, IAAE Conference, Sapporo 2017.

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, together with banking and asset markets, for understanding how monetary economics works has recently received renewed attention from the growing body of research called New Monetarist Economics (see Williamson and Wright 2010 and Schneider and Piazzesi 2015). Nevertheless, the use of payment data for macroeconomic forecasting has only recently been exploited.

Studies for Canada (Galbraith and Tkacz 2007, 2009, 2018), Portugal (Esteves 2009; Duarte, Rodrigues, and Rua 2017), Denmark (Carlsen and Storgaard 2010), and the United States (Barnett et al. 2016) find that payment transactions can help with nowcasting and with forecasting GDP and private consumption in the short term.

It is worth noting that this strand of the literature concentrates on payment cards (as in Esteves 2009; Carlsen and Storgaard 2010; Duarte, Rodrigues, and Rua 2017; Galbraith and Tkacz 2018; and Barnett et al. 2016, who also consider checks) and mainly targets GDP and consumption, thus offering only a partial view of the whole system of payments and of economic activity.

To the best of our knowledge, this paper is the first attempt to assess the ability of retail payment data to make accurate short-term forecasts, focusing on Italy and drawing on a broad set of aggregated payments. We target both GDP and its main domestic components (households’ consumption, HC, gross fixed investments, GFI, and value added in the service sector, VAS) and we use a comprehensive set of payment instruments, including credit transfers, checks, direct debits, and debit cards. This approach allows us to conduct a more robust empirical application than similar studies for other countries, which mostly consider a subset of payment instruments and macroeconomic aggregates. Our 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.¹ Another noteworthy

¹Our data are extracted from the payment system infrastructures, while in some structural analysis data are collected by means of surveys or diaries (see Bagnall et al. 2016).

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 retail payment series (hereinafter PS) and the main macroeconomic aggregates. In the empirical application, we select the indicators by means of LASSO (least absolute shrinkage and selection operator) and then we set up a mixed-frequency dynamic factor model to predict the quarter-on-quarter growth of GDP and its main components by using a large-scale monthly data set, which includes standard business cycle indicators other than payment data; we perform out-of-sample forecasting simulations, including and excluding PS. The contribution of PS turns out to be appreciable and promising throughout the forecasting horizon considered (from one quarter backwards up to two quarters ahead). Moreover, the timeliness of the payment data lets us refine the nowcasting throughout the quarter in real time.

The rest of the paper is organized as follows. Section 2 gives an overview of the payment system in Italy. Section 3 introduces the data, providing some descriptive evidence on the relationship between PS and the main macroeconomic aggregates. Section 4 deals with the empirical application: subsection 4.1 shows that LASSO picks PS amongst the first fourteen predictors to be included in the model (out of the fifty indicators compiled earlier); subsection 4.2 describes the forecasting exercise and the results. Section 5 concludes.

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), usually 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 closely related to economic activity (Padoa-Schioppa 2004).

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

As for retail payments,² the system is not yet fully integrated. However, since 2014 the Single Euro Payments Area (hereinafter SEPA) has strongly promoted the standardization and interoperability of different national clearing and settlement retail systems. The Bank of Italy manages the BI-Comp clearing system, which works in accordance with the rules of SEPA. BI-Comp clears domestic payments on a multilateral net basis. These payments can be settled both in BI-Comp and in T2. In fact, due to urgency and for security reasons, banks may prefer to settle customers' payments in T2. This retail branch of T2 is named T2-retail from here on in.

Retail non-cash payments settled through BI-Comp and T2-retail add up to €5 trillion on a yearly basis (about 60 percent of the total value of retail payments in Italy—80 percent if we only consider electronic payments, excluding postal pre-printed processed and other paper-based credit transfers), about three times higher than the nominal value of GDP.³ Although in Italy cash payments

²The 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).

³Unlike cash payments, which amount to immediate transfers of value between the payer and the payee through bank notes and coins, non-cash payments are exchanges of funds through accounts. It follows that the relationship between the payer and the payee is mediated by authorized institutions (such as banks and postal offices), which actually process the payment before settling the transaction. This brings us to a crucial distinction depending on the party submitting the payment order: that between *credit-based* instruments (i.e., credit transfers, card payments), submitted by the payer, and *debit-based* instruments (i.e., direct

remain the most frequently used instrument for retail payments (about 80 percent of the *number* of retail payments), new information and communication technologies have encouraged non-cash payments, mainly those processed through electronic devices, allowing greater flexibility and customization. If we consider the *value* of the transactions, on average per year,⁴ the share of cash payments shrinks to about 45 percent of consumer-to-business transactions (for instance, point-of-sales, or POS, purchases) and to less than 10 percent of all transactions, including business-to-business payments. It is worth noting that the payment data recorded in BI-Comp include ATM cash withdrawals, which may represent a good proxy of the transactions paid with cash (see Schmiedel, Kostova, and Ruttenberg 2012). Data on cash withdrawals also allow us to take into account the long-run trends on consumer payment habits (cash versus non cash) within our forecasting approach. For the Italian economy, the percentage of cash withdrawals on the total value of payments is fairly stable (below 5 percent) in the time span considered in the empirical application. In the same period, the figure is also broadly stable for the euro area. Figure 1 shows the monthly gross flow of retail payments settled through BI-Comp and T2-retail in Italy.

The sharp decrease observed in 2014 for BI-Comp stems from changes in the customer payment landscape following the migration to SEPA.⁵ Some participants reconsidered the routing policies for their customer payments, and they ultimately opted for SEPA-compliant automated clearinghouses other than BI-Comp and T2. However, T2-retail has been less affected, because it meets some specific customers' demands concerning urgency and assurance of payments.

debits, checks), submitted by the payee. Credit transfers are the most common credit-based instrument, while debit-based instruments include direct debits, card payments, and checks. For a detailed description of payment instruments, see <http://www.bancaditalia.it/compiti/sispaga-mercati/strumenti-pagamento/index.html?com.dotmarketing.htmlpage.language=1> and *The Payment System: Payments, Securities and Derivatives, and the Role of the Eurosystem*, by the European Central Bank.

⁴These data refer to 2014.

⁵SEPA represents a harmonization of procedures and platforms in the euro zone for processing retail credit transfers and direct debits.

Figure 1. Total Payments Settled in BI-Comp and TARGET2-Retail (values)

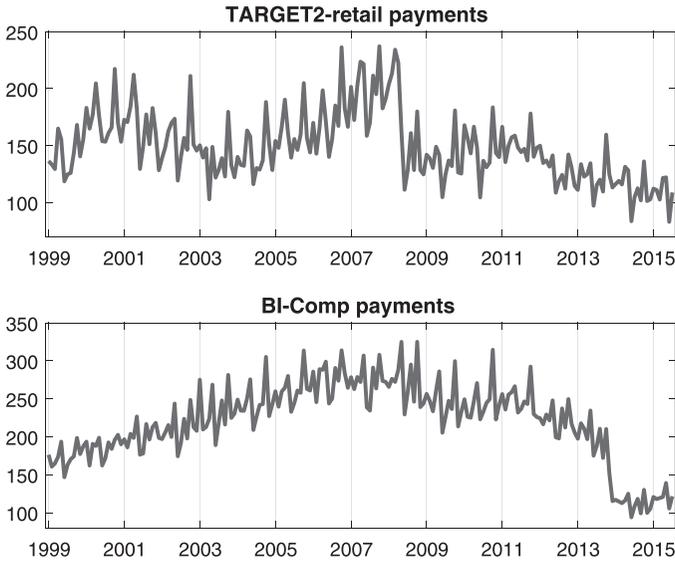
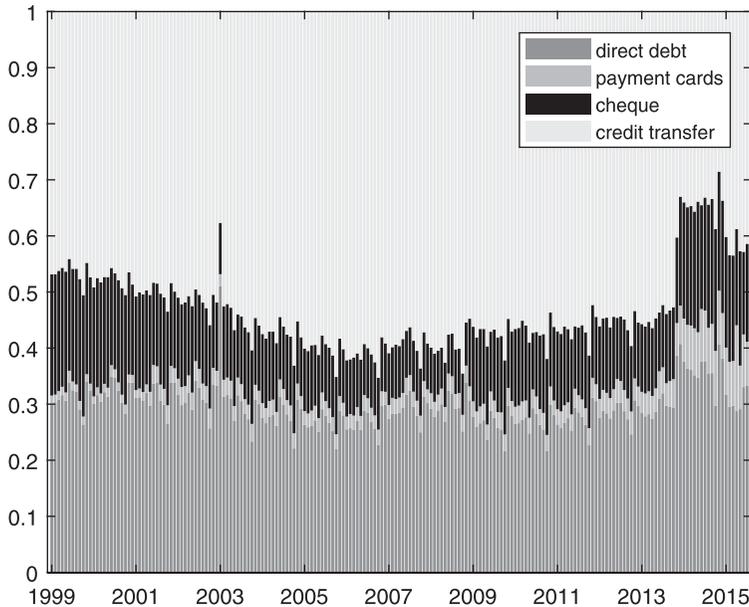


Figure 2 depicts the share of transactions enabled by different payment instruments over the total gross flows settled in BI-Comp. Credit transfers represent the largest share (almost 55 percent);⁶ that of direct debits is also sizable (29 percent) since they provide a very neat solution to managing recurrent payments (e.g., utility bills, mortgage payments). Payment cards are far less used in Italy, accounting for only 3 percent of the total non-cash payments; cards are used for low-value transactions (€50–€100 for POS; €100–€200 for ATM).

However, a large share of credit card payments are recorded as direct debits, since credit card statements are often charged to a payer's current account.

⁶For business purposes credit transfers are the most popular and suitable payment instrument, accounting for 80 percent of the total value of business-to-business transactions; direct debits and checks account for 10 percent, respectively.

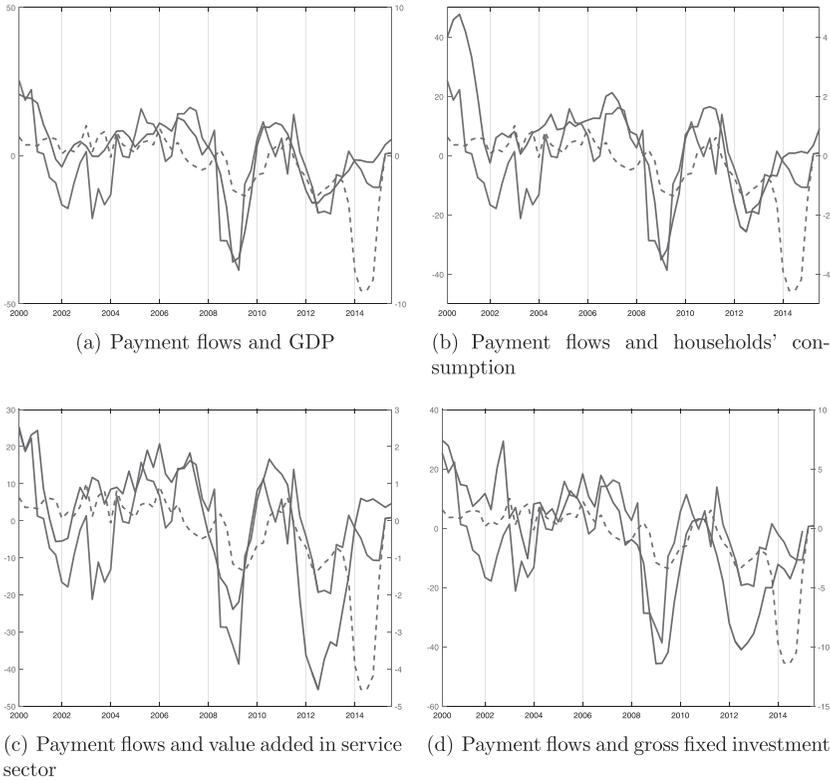
Figure 2. Total Payments Settled in BI-Comp

3. Payments Data and Macroeconomic Aggregates

In this section we look at some descriptive statistics which prove the close relationship between the main macroeconomic aggregates (such as GDP, value added in the 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 clear empirical evidence underpinning the role of payment time series in tracking economic activity.

We rely on the BI-Comp and T2-retail systems, described in section 2, 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 every day, therefore the observation error is basically nil. We do not use individual payments (which eventually could be relevant for a big data analysis) but instead use the aggregated monthly time series stored in Bank of

Figure 3. BI-Comp (dashed line) and T2-Retail (solid gray line; left-side axis) Flows Compared with Macroeconomic Aggregates (solid black line; right-side axis): Year-on-Year Percentage Changes



Italy's database, which is publicly available.⁷ The monthly observations are simply the sum of the nominal value of all the individual payments recorded in the month.

The annual growth of GDP and of its main components co-moves closely with the annual growth of payment flows settled through BI-Comp and T2-retail (see figure 3).⁸ The big drop at the end of the

⁷Data may be found at <https://www.bancaditalia.it/statistiche/tematiche/statistiche-sistema-pagamenti/index.html>.

⁸BI-Comp and T2-retail nominal flows are divided by the GDP national account deflator in order to obtain a measure of the volume of transactions. In

Table 1. Correlation between Payment Flows and Macroeconomic Indicators (year-on-year percentage changes)

Payment Flows ^a	GDP	Private Consumption	Gross Fixed Investment	Value Added Service Sector
BI-Comp	71.8	70.5	74.3	68.4
T2-Retail	79.4	67.7	64.0	74.3
BI-Comp + T2-Retail	90.7	82.5	82.3	85.9
<i>Other Indicators</i>				
M2	-38.1	-11.5	-20.4	-45.6
Industrial Production	92.8	68.7	74.3	79.8
Business Confidence ^b	68.5	74.9	67.3	64.3
Consumer Confidence ^c	9.9	40.0	14.3	19.7
^a Contemporaneous unconditional correlations are computed on the sample 2000:Q1–2012:Q4 in order to exclude the break caused by the new standard SEPA. ^b Economic Sentiment Indicator (ESI) provided by Istat. ^c Consumer confidence survey by Istat.				

sample proves that BI-Comp was more affected than T2-retail by the launch of SEPA in October 2014.⁹ Some interesting results also emerge from the correlation matrix (see table 1).¹⁰ The correlation between PS and the target variables is valuable and similar to the one shown by other indicators such as industrial production and business confidence, usually adopted in short-term forecasting. The

the empirical application (section 4), we consider nominal value, but we include the harmonized consumer price index as a variable to control for the price effect instead of deflating the payment series.

⁹The launch of SEPA has entailed the switching of some domestic credit transfers and direct debits from BI-Comp into STEP2, which is the pan-European infrastructure managed by a private body (Interbank Society for Automation). Data on payment flows in STEP2 are not publicly available. Nonetheless, removing the outlier corresponding to the launch of SEPA turns out to be suited to preserving the predictive ability of BI-Comp.

¹⁰We consider the contemporaneous correlation.

results in table 1 are consistent with the picture of the payment instruments composition in Italy. Credit transfers and direct debits are the payment tools most relevant in terms of values; they are also more closely correlated with the target variables than payment cards and checks.

4. Empirical Application

4.1 *Selection of the Targeted Predictors*

We compiled $N = 50$ variables, which provide a fairly complete picture of economic activity, including indicators of industrial and service activity (industrial production, electricity consumption, freight truck, business confidence), households' consumption (retail sales of goods and services, consumers' confidence), financial indexes, and credit flows to firms other than time series from the payment systems T2-retail and BI-Comp. All the variables are seasonally adjusted by monthly seasonal dummies; as for PS, we also remove some outliers pinned down as the peaks and troughs larger than 1.5 times the standard deviation of the time series. These outliers are replaced by the mean of the neighboring observations. The payment series refer to the nominal value of the transactions.¹¹ We also use nominal estimates of GDP and of its main components. The construction of the appropriate deflator for both PS and the macroeconomic aggregates would require a specific analysis, therefore we decided to include the price index¹² in the final model in order to control for the price effect.

Factor models provide a parsimonious way of handling such a large-scale data set. Nevertheless, some recent studies have demonstrated that using a large number of variables to estimate the common factors may worsen the forecasting performance (see Boivin and Ng 2006). Therefore, we use LASSO (for a review of the LASSO estimator, see Tibshirani 1996 and Hastie, Tibshirani, and Friedman 2009) to screen the variables depending on their ability to anticipate our targets, in the same spirit as Bai and Ng (2008), who use both

¹¹The number of transactions is also available, but in this paper we focus on the values.

¹²We use the harmonized index of consumer prices for Italy.

hard and soft thresholding methods to select the so-called targeted predictors. Moreover, the selection made by LASSO provides initial evidence of whether PS can be considered good predictors of the short-term evolution of economic activity, compared with other workhorse indicators for conjunctural analysis.

The monthly indicators are transformed into quarterly variables and then the LASSO regression is performed. LASSO selects the regressors by solving the minimization problem

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

where τ is the tuning parameter. LASSO selects n_L targeted variables, among all N collected:

$$\hat{L}_{n_L} = \{j \in \{1, 2, \dots, N\} : |\hat{\beta}_{Lj}| > 0\}. \quad (2)$$

We revise the LASSO selection by introducing and discarding a few variables as documented below.¹³ With respect to the GDP model, we get rid of the credit flows and PS-total (the sum of BI-Comp and T2-retail); we deem the latter redundant once BI-Comp and T2-retail are included. As for households' consumption model, LASSO would pick BI-Comp and PS-total, but we find it reasonable to replace the latter (which carries redundant information if BI-Comp is already included) with T2-retail; we discard consumers' survey on "future personal economic situations" and two series from the business surveys ("expected level of orders" and "expected level of liquidity" in the consumer goods sector) which track industrial activity more effectively. In the model for gross fixed investments, the "expected level of production" in the intermediate goods sector is less volatile than the "expected level of orders." We include BI-Comp and T2-retail otherwise discarded by LASSO in both the model for investments and for value added in the service sector. In the latter we also include the price index, excluded by LASSO. Table 2 shows the variables finally included in our information set, \hat{I}_n .

¹³As a robustness check, we estimated three alternative models. More precisely, in the first model LASSO picks the targeted predictors and PS are in the pool of candidates; in the second model PS are not among the candidates; and the third model expands the second model with PS.

Table 2. The Final Model (\hat{I}_n)

Indicators	GDP	HC	GIF	VAS
Total Electricity Consumption	No	Yes	Yes	Yes
Industrial Production	Yes	Yes	Yes	Yes
Business Climate	No	No	Yes	Yes
CCS—Future General Economic Situations	No	No	Yes	Yes
CCS—Future Personal Economic Sit.	No	No	Yes	Yes
CCS—Unemployment Exp.	No	No	Yes	Yes
CCS—Saving Opportunities, Next Twelve Months	Yes	Yes	Yes	Yes
CCS—Households Balance Sheet	Yes	Yes	Yes	Yes
CCS—Current Saving Opportunities	No	No	Yes	No
BCS—Current Level of Orders (Intermediate Goods)	No	No	Yes	No
BCS—Current Level of Production (Int. Goods)	No	No	Yes	No
BCS—Expected Level of Production (Int. Goods)	No	No	Yes	Yes
BCS—Exp. Level of Orders (Int. Goods)	No	No	No	No
BCS—Future General Economic Sit. (Int. Goods)	No	No	No	No
BCS—Exp. Level of Liquidity (Int. Goods)	No	No	No	No
BCS—Current Level of Liquidity (Int. Goods)	No	No	No	No
BCS—Current Level of Orders (Investment Goods)	No	No	No	No
BCS—Current Level of Production (Inv. Goods)	No	No	No	No
BCS—Current Level of Liquidity (Inv. Goods)	No	No	No	No
BCS—Exp. Level or Orders (Inv. Goods)	No	No	No	No
BCS—Exp. Level of Production (Inv. Goods)	No	No	No	No
BCS—Future General Economic Sit. (Inv. Goods)	No	No	No	No
BCS—Exp. Level of Liquidity (Inv. Goods)	No	No	No	No
BCS—Current Level of Orders (Consumer Goods)	No	No	No	No
BCS—Current Level of Production (Cons. Goods)	No	No	No	No
BCS—Exp. Level of Orders (Cons. Goods)	Yes	No	Yes	No
BCS—Exp. Level of Production (Cons. Goods)	No	No	No	No
BCS—Future General Economic Sit. (Cons. Goods)	Yes	No	No	No
BCS—Exp. Level of Liquidity (Cons. Goods)	No	No	No	No
BCS—Current Level of Liquidity (Cons. Goods)	No	Yes	No	No
Current Accounts Deposits (Stock)	No	No	No	No
Credit Flows to Firms	No	No	No	No
HICP	Yes	Yes	Yes	Yes
FTSE Italy (Banks)	No	No	No	No
FTSE Italy (Insurance)	No	No	No	No
FTSE Italy (Transport)	No	No	No	No
PMI Services—Business Activity	No	No	No	No
PMI Services—New Business	Yes	No	No	No
PMI Manufacturing	No	No	No	No
PMI Manufacturing—Output	No	No	No	No
PMI Manufacturing—New Orders	Yes	Yes	Yes	Yes
PMI Manufacturing—Employment	No	No	No	No
PMI Manufacturing—New Export Orders	No	Yes	No	No
Freight Truck	Yes	No	No	Yes
Retail Trade—Goods	Yes	Yes	No	Yes
Retail Trade—Services	No	Yes	No	No
BI-COMP	Yes	Yes	Yes	Yes
TARGET RETAIL	Yes	Yes	Yes	Yes
Payments System—Total	No	No	No	No

We implement a large-scale dynamic factor model on \hat{I}_n , which is more suitable when the cross-section size of the data set, n , is valuable with respect to the number of observations, T . Working with large data sets proves a robust strategy for detecting whether PS contribute to improving accuracy in the estimation of economic growth, since they are compared to many other indicators, often used to track the short-term dynamics of the economy.

4.2 Forecasting Exercise

4.2.1 The Model

Let \mathbf{X}_t be the n -vector of observable monthly variables selected earlier and \mathbf{X} the $T \times n$ matrix. They are driven by q common factors, $\mathbf{f}_t = [f_{1t}, \dots, f_{qt}]'$, and by n idiosyncratic components, $\xi_t = [\xi_{1t}, \dots, \xi_{nt}]'$, which are assumed to be uncorrelated. The $n \times q$ matrices Λ_i , for $i = 1 \dots s$ lags, are the common-factor loadings. The dynamic factor model

$$\begin{aligned} \mathbf{X}_t &= \Lambda_s(L)\mathbf{f}_t + \xi_t \\ &= \Lambda_0\mathbf{f}_t + \Lambda_1\mathbf{f}_{t-1} + \Lambda_2\mathbf{f}_{t-2} + \dots + \Lambda_s\mathbf{f}_{t-s} + \xi_t \end{aligned} \quad (3)$$

can be mapped to the static model

$$\mathbf{X}_t = \mathbf{D}\mathbf{F}_t + \xi_t, \quad (4)$$

where $\mathbf{F}_t = [\mathbf{f}'_t, \mathbf{f}'_{t-1}, \dots, \mathbf{f}'_{t-s}]'$ has the dimension $r = q(s+1)$ and can be represented as a VAR(1) process,¹⁴

$$\mathbf{F}_t = \mu + \Psi_1\mathbf{F}_{t-1} + \dots + \Psi_l\mathbf{F}_{t-l} + \mathbf{u}_t. \quad (5)$$

As stated in Bai and Ng (2007), \mathbf{F}_t is driven by $q < r$ common shocks if $\mathbf{u}_t = \mathbf{A}\epsilon_t$, where ϵ_t is a q -vector of mutually orthogonal shocks with variance equal to one and the $r \times q$ -matrix \mathbf{A} has rank q ; then $E(\mathbf{u}_t\mathbf{u}'_t) = \Sigma_u = \mathbf{A}\Sigma_\epsilon\mathbf{A}'$ has reduced rank q . We estimate \mathbf{A} as in

¹⁴We set the order l of the VAR equal to 4.

Marcellino and Schumacher (2010): given the OLS estimate $\hat{\Sigma}_u$ and its eigenvalue decomposition \mathbf{MPM}' , let \mathbf{M}_* be the $r \times q$ -matrix of the first q eigenvectors and \mathbf{P}_* the diagonal matrix of the corresponding eigenvalues; then $\hat{\mathbf{A}} = \mathbf{M}_* \mathbf{P}_*^{-1/2}$ and the reduced-rank estimate of Σ_u is equal to $\Sigma_u^r = \hat{\mathbf{A}} \hat{\mathbf{A}}'$. We follow Bai and Ng (2002, 2007) to set the number of static ($r = 4$) and dynamic ($q = 1$) factors and then we estimate factors' space $\mathcal{G}(F_t) = \overline{\text{span}}(F_{1t}, \dots, F_{rt})$ by principal components extracted from the balanced monthly data set. More specifically, we extract from the covariance matrix of \mathbf{X} the first r eigenvalues and the corresponding $n \times 1$ eigenvectors, \mathbf{V}_i for $i = 1, \dots, r$, and then we compute the $T \times r$ matrix of the common static factors $\mathbf{F} = \mathbf{XV}$. By construction, the covariance matrix of the common static factors is the diagonal matrix of the first r eigenvalues, while $\hat{\Sigma}_\xi$ is a diagonal matrix whose entries are extracted by the diagonal of the covariance matrix of $\mathbf{X} - \mathbf{F}\hat{\mathbf{D}}'$, where $\hat{\mathbf{D}}$ is the OLS estimate of the $n \times r$ matrix of static-factors loadings.

The quarterly growth rate of the target variable, y_{t_q} with t_q labeled by a multiple of the last month of each quarter (i.e., $t_q = 3, 6, \dots, 3\lceil T/3 \rceil$), is projected by an unrestricted mixed-data sampling (MIDAS) model on the monthly information $\mathcal{G}(F)$ (see Ghysels, Santa-Clara, and Valkanov 2004 for an extensive treatment of the MIDAS model and refer to Foroni, Marcellino, and Schumacher 2015 for the unrestricted MIDAS model). For each quarter t_q we have $m = 3$ values of the monthly regressors; therefore, we apply the $L^{j/m}$ operator to obtain regressors lagged by j months with respect to the quarter. Put formally,

$$\begin{aligned} y_{t_q} &= c + \beta_0 \mathbf{F}_{t_q} + \beta_1 \mathbf{F}_{t_q-1/m} + \dots + \beta_p \mathbf{F}_{t_q-p/m} + \epsilon_{t_q} \\ &= c + \beta(L^{1/m}) \mathbf{F}_{t_q} + \epsilon_{t_q}, \end{aligned} \quad (6)$$

where the loadings β_j for $j = 1, \dots, p$ have dimension $1 \times r$ and are estimated by a simple OLS. We choose $p = 3$. By way of example, let us suppose Q1 is the current quarter and $t_q = 3\lceil t/3 \rceil$. Therefore, the estimate of y_{t_q} will rely on information on December, January, February, and March. Equations (4)–(6) are cast in a state-space form. To include the quarterly growth rate of the target variable into the state-space framework, we construct a monthly series y_t ,

where $y_t \equiv y_{t_q}$ when $t \equiv t_q$, and is missing otherwise. If $l \geq p$, the state equation is

$$\mathbf{X}_t = \underbrace{[\mathbf{D} \quad 0_{n \times r(l-1)+1}]}_{\mathbf{DD}} \cdot \begin{bmatrix} \mathbf{F}_t^* \\ y_t \end{bmatrix} + \mathbf{e}_t, \tag{7}$$

where

$$\mathbf{F}_t^* = \begin{cases} [\mathbf{F}'_t, \mathbf{F}'_{t-1}, \dots, \mathbf{F}'_{t-l+1}]' & \text{if } l \geq p \\ [\mathbf{F}'_t, \mathbf{F}'_{t-1}, \dots, \mathbf{F}'_{t-p+1}]' & \text{otherwise} \end{cases} \tag{8}$$

and $\mathbf{e}_t = [\xi_t, 0]' \sim N(0_{(n+1) \times 1}, \mathbf{R})$ with $\mathbf{R} = \begin{bmatrix} \Sigma_\xi & 0 \\ 0 & 0 \end{bmatrix}$; the transition equation is

$$\underbrace{\begin{bmatrix} \mathbf{I}_{rl} & 0_{rl \times 1} \\ -\beta_0 & 0_{1 \times r(l-1)} \quad 1 \end{bmatrix}}_H \cdot \begin{bmatrix} \mathbf{F}_t^* \\ y_t \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathbf{Psi} \\ \mathbf{I}^* \end{bmatrix} & 0_{rl \times 1} \\ \underbrace{[\beta_1 \dots \beta_p]}_G & 0_{1 \times r(l-p)+1} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{F}_{t-1}^* \\ y_{t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} \mu \\ 0_{r(l-1) \times 1} \\ c \end{bmatrix}}_d + \begin{bmatrix} \mathbf{u}_t \\ \epsilon_t \end{bmatrix}, \tag{9}$$

where

$$\begin{bmatrix} \mathbf{Psi} \\ \mathbf{I}^* \end{bmatrix} = \begin{bmatrix} \Psi_1 & \Psi_2 & \dots & \Psi_l \\ \mathbf{I}_r & & 0_{r \times r(l-1)} & \\ \vdots & & & \\ 0_{r \times r(l-2)} & \mathbf{I}_r & & 0_{r \times r} \end{bmatrix},$$

while if $l < p$, the state equation is

$$\mathbf{X}_t = [\mathbf{D} \quad 0_{n \times r(p-1)+1}] \cdot \begin{bmatrix} \mathbf{F}_t^* \\ y_t \end{bmatrix} + \mathbf{e}_t \tag{10}$$

and the transition equation is

$$\begin{aligned}
 \begin{bmatrix} \mathbf{I}_{rp} & 0_{rp \times 1} \\ -\beta_0 & 0_{1 \times r(p-1)} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{F}_t^* \\ y_t \end{bmatrix} &= \begin{bmatrix} \begin{bmatrix} \mathbf{Psi} \\ \mathbf{I}^* \end{bmatrix} & 0_{rl \times r(p-l)+1} \\ & 0_{r(p-l) \times rp+1} \\ [\beta_1 \cdots \beta_p] & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{F}_{t-1}^* \\ y_{t-1} \end{bmatrix} \\
 &+ \begin{bmatrix} \mu \\ 0_{r(p-1) \times 1} \\ c \end{bmatrix} + \begin{bmatrix} \mathbf{u}_t \\ 0_{r(p-1) \times 1} \\ \epsilon_t \end{bmatrix} \tag{11}
 \end{aligned}$$

with $\begin{bmatrix} \mathbf{u}_t \\ 0 \\ \epsilon_t \end{bmatrix} \sim N(\underline{0}, \mathbf{Q})$ and $\mathbf{Q} = \begin{bmatrix} \Sigma_u^r & 0 \\ 0 & \sigma_\epsilon^2 \end{bmatrix}$. We implement the

Kalman recursions (Kim and Nelson 1999) to extract the smoothed state variable $\mathbf{F}\mathbf{F}_t = [\mathbf{F}_t^{*'} \ y_t]'$. Let T_ν 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 $t = 1, 2, \dots, T_\nu + 3h$, based on information up to $T = T_\nu + 3h + 1$, i.e., $\mathbf{F}\mathbf{F}_{t|T}$. This is very appreciable because we use the latest information to infer the dynamics of the state variable. In fact, when $t = T_\nu + 1, \dots, T_\nu + 3h$, all observations are missing and they are given no weight by the Kalman filter, which basically forecasts the factors as elaborated in Giannone, Reichlin, and Small (2008). In practice, we impose infinite variance to the i -th idiosyncratic component if the i -th observable variable is missing. In so doing, the weight of the missing observation will fade when computing the Kalman gain, which projects the \mathbf{X} s onto the space spanned by the idiosyncratic factors. In particular, we are interested in forecasting the target variable at time $T_{q_\nu} + h$, where T_{q_ν} is the quarter corresponding to the month T_ν :

$$y_{T_{q_\nu}+h} = c + \beta(L^{1/m})\mathbf{F}_{T_{q_\nu}+h|T} + \epsilon_{T_{q_\nu}+h}. \tag{12}$$

4.2.2 Out-of-Sample Simulation

We conduct an out-of-sample forecasting exercise at different horizons ($h = -1, 0, 1, 2$)¹⁵ to assess how the forecasting performance of our models changes depending on whether we include or exclude PS from the set of regressors.

We run a pseudo real-time simulation; therefore, 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.¹⁶ This data set ranges from January 2000 to November 2015, and it is balanced on August 2015.

The exercise is carried out on two different time intervals. The first estimation sample goes from January 2000 to April 2008 and expands until the last balanced date (i.e., August 2015). The current period is assumed to be two months after the balanced date, therefore the first *pseudo* vintage is June 2008 and we nowcast 2008:Q2 and make two-steps-ahead forecasts for 2008:Q3 and 2008:Q4. We end up with the first sample of forecasting errors, which is 2008:Q2–2015:Q2. The second estimation sample goes from January 2000 to July 2011 and the sample of the forecasting errors is 2011:Q3–2015:Q2.

The benchmark model includes all the variables listed in the column “Final Model” of table 2 and is compared with the model replacing PS with the first variable discarded by LASSO for each target \hat{I}_{PS} .¹⁷ In fact, we consider two benchmark models: the first includes only T2-retail such as payment series (\hat{I}_{T2}) while the second

¹⁵The forecasting horizon depends on which month of the quarter is taken as the *pseudo* vintage. For instance, in the first month of quarter t_q , we backcast ($h = -1$) the target variable of the previous quarter not yet released by the national statistical office (Istat); we also nowcast ($h = 0$) and forecast one quarter and two quarters ahead ($h = 1, 2$).

¹⁶It is reasonable that the final estimate of GDP is related more closely to the payment series as well as to the final revised versions of other macroeconomic variables. If the latter were in real time, then the payment series may have been given an unfair advantage. Therefore, by making everything observed as their final revised values, we prevent this risk.

¹⁷Electricity consumption for GDP; business climate for HC; orders' expectations—intermediate goods for GFI; households' current saving convenience for VAS. As for GFI and VAS, PS are initially discarded by LASSO and we reintroduce them in place of orders' expectations—intermediate goods and households' current saving convenience, respectively.

Table 3. Relative RMSFE of the Model including only T2-Retail

		Backcast	Nowcast	Forecast One Step	Forecast Two Steps
HC	2008:Q2–2015:Q2	1.10	1.14	1.15	1.14
	2011:Q3–2015:Q2	0.97	0.98	1.04	1.06
GFI	2008:Q2–2015:Q2	1.00	0.93	1.06	1.07
	2011:Q3–2015:Q2	0.80	0.82	1.17	1.12
VAS	2008:Q2–2015:Q2	1.00	1.08	1.11	1.12
	2011:Q3–2015:Q2	1.01	1.03	0.99	0.98
GDP	2008:Q2–2015:Q2	0.74	0.86	1.04	1.11
	2011:Q3–2015:Q2	1.28	1.10	0.91*	0.94

Notes: This table shows the relative RMSFE of the model including only T2-retail (\hat{I}_{T2}) vis-à-vis the model replacing payment series with the first variable discarded by LASSO for each target (\hat{I}_{PS}). For GDP, the Diebold-Mariano test (with HAC variance estimators) is computed; the null hypothesis is that the out-of-sample errors of the competing models are equal. * shows p-values < 0.05, ** p-values < 0.01, and *** p-values < 0.001.

includes both T2-retail and BI-Comp ($\hat{I}_{T2,BC}$). Tables 3 and 4 show the root mean square forecasting error (RMSFE) of the competing models \hat{I}_{PS} relative to the RMSFE of the two benchmark models. A figure greater than one means that PS improve the forecasting performance. PS improve the forecasting accuracy broadly.¹⁸ In general, information on retail payments plays a role when the targets are projected one quarter and two quarters ahead, irrespective of the time interval, and it tracks well the dynamics of households' consumption.¹⁹ In particular, we record a forecasting gain throughout the horizons in $\hat{I}_{T2,BC}$.

¹⁸We compared the forecast accuracy of non-nested models (when we exclude the payment variable we include the first excluded indicator in LASSO); therefore, we used the Diebold-Mariano test (with heteroskedasticity and autocorrelation consistent, HAC, variance estimators) of forecast accuracy. We computed this test for GDP forecasts, but we did not find a clear pattern of statistical significance. We interpret this finding as a result of the small out-of-sample window in our forecasting application.

¹⁹Regarding the three alternative specifications described in footnote 13, the results on the role of payment series in forecasting are mixed. However, these models generally worsen the forecasting performance compared with the benchmark $\hat{I}_{T2,BC}$ (the results are available on request).

Table 4. Relative RMSFE of the Model including both T2-Retail and BI-Comp

		Backcast	Nowcast	Forecast One Step	Forecast Two Steps
HC	2008:Q2–2015:Q2	1.23	1.21	1.13	1.10
	2011:Q3–2015:Q2	1.09	1.07	1.09	1.06
GFI	2008:Q2–2015:Q2	1.10	1.07	1.24	1.22
	2011:Q3–2015:Q2	0.80	0.82	1.17	1.12
VAS	2008:Q2–2015:Q2	1.08	1.13	1.09	1.09
	2011:Q3–2015:Q2	1.10	1.08	1.01	1.01
GDP	2008:Q2–2015:Q2	0.93	0.98	1.07	1.09
	2011:Q3–2015:Q2	1.79	1.58**	1.05	1.01

Notes: This table shows the relative RMSFE of the model including both T2-retail and BI-Comp ($\hat{I}_{T2,BC}$) vis-à-vis the model replacing payment series with the first variable discarded by LASSO for each target (\hat{I}_{PS}). For GDP, the Diebold-Mariano test (with HAC variance estimators) is computed: the null hypothesis is that the out-of-sample errors of the competing models are equal. * shows p-values < 0.05, ** p-values < 0.01, and *** p-values < 0.001.

Model $\hat{I}_{T2,BC}$ performs better than \hat{I}_{T2} ; however, the latter also yields some valuable results. T2 shrinks the RMSFE for backcasting and nowcasting GDP by 28 percent and 10 percent, respectively, during the turmoil sparked by the sovereign debt crisis; in the same period, we gain 17 percent of predictive accuracy when forecasting investments one step ahead.

Some remarkable results come out of $\hat{I}_{T2,BC}$. The forecasting performance improves for HC and GFI throughout the horizons when we consider the longest sample. As for activity in the service sector, in the same period PS lowers the RMSFE for nowcasting by 13 percent. The most notable outcome regards the backcast and the nowcast of GDP during the last four years. In this case, using PS lowers the RMSFE by 79 percent and 58 percent, respectively.

These results bear out the suitability of PS to making predictions of real activity. The model we used, which belongs to the class of models with large data sets, strengthens our claim further. Indeed, 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 data set becomes sizable.

Table 5. Relative RMSFE for the Second and the Third Month of the Quarter vis-à-vis the First Month (percentage values)

	2005:Q2–2015:Q2	
	$\hat{I}_{T2,BC}$	\hat{I}_{-PS}
First Month		
Second Month	–5.9	–2.9
Third Month	–21.0	–18.0

Note: The figure is the ratio between the RMSFE in both the second and the third month of the quarter and in the first month.

One of the most appealing characteristics of the payment system data is their timeliness, since they are available in the reference month. Therefore, we can assess how much we benefit from the monthly real-time information on PS to improve the nowcast during the quarter. This experiment is conducted for GDP only. PS are assumed to be the most timely regressors of the GDP model, while the other explicative variables are two months late with respect to the reference period. Using the same *pseudo* real-time simulation exercise, we update the nowcast each month of the quarter and we observe how the RMSFE changes. We compare the nowcast made in the first month of the quarter with the nowcast made in the second and in the third month by $\hat{I}_{T2,BC}$ and \hat{I}_{-PS} , where the latter replaces PS with the total electricity consumption, which is as timely as PS. In so doing, we lose two-thirds of the observations, so we use a longer sample, from 2005:Q2 to 2015:Q2. In table 5 we observe a monotonic improvement of the forecasting accuracy from the first to the last month of the quarter for both models, as expected. However, $\hat{I}_{T2,BC}$ performs somewhat better than \hat{I}_{-PS} . When we are in the second month of the quarter, the nowcast by $\hat{I}_{T2,BC}$ is 5.9 percent more accurate than the nowcast made in the first month (2.9 percent by \hat{I}_{-PS}). The accuracy improves further in the third month of the quarter (the nowcast is 21 percent more precise than in the first month—18 percent by \hat{I}_{-PS}).

4.2.3 Observation Weights

One of the most common remarks made about factor models concerns the possibility of disentangling 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 given any definite economic meaning. As showed in Koopman and Harvey (2003), the output of the Kalman recursions can be used to measure the weights attached to the observable variables in \mathbf{X}_j when forecasting y_{t_q} , for $j = 1, \dots, T_\nu + 3h + 1$. The smoothed state vector can be expressed as the weighted sum:

$$\mathbf{FF}_{t|T} = \sum_{j=1}^T w_j(\mathbf{FF}_{t|T})\mathbf{X}_j, \tag{13}$$

where the weights $w_j(\cdot)$ are a function of the state vector. Each month t , the weights are computed by the backward recursions (for $j = t - 1, t - 2, \dots, 1$) and the forward recursions (for $j = t, t + 1, \dots, T$) introduced in Koopman and Harvey (2003).²⁰ Let us define $\mathbf{L}_t = \mathbf{Z} - \mathbf{K}_t\mathbf{D}\mathbf{D}$, $\mathbf{N}_t = \mathbf{D}\mathbf{D}'\mathbf{S}_t^{-1}\mathbf{D}\mathbf{D} + \mathbf{L}_t'\mathbf{N}_t\mathbf{L}_t$ for $t = T, \dots, 1$ and $\mathbf{N}_T = 0$; $\mathbf{C}_t = \mathbf{D}\mathbf{D}'\mathbf{J}_t - \mathbf{Z}'\mathbf{N}_t\mathbf{K}_t$ where $\mathbf{J}_t = \mathbf{S}_t^{-1} + \mathbf{K}_t'\mathbf{N}_t\mathbf{K}_t$ for $t = T, \dots, 1$. The weights for filtering are given by the following recursions:

$$w_j(\mathbf{FF}_{t|t-1}) = B_{t,j}\mathbf{K}_j, \quad B_{t,j-1} = B_{t,j}\mathbf{Z} - w_j(\mathbf{FF}_{t|t-1})\mathbf{D}\mathbf{D} \tag{14}$$

for $j = t - 1, \dots, 1$ with $B_{t,t-1} = \mathbf{I} - \mathbf{P}_{t|t-1}\mathbf{N}_{t-1}$, while the weights for smoothing are given by

$$w_j(\mathbf{FF}_{t|T}) = (\mathbf{I} - \mathbf{P}_{t|t-1}\mathbf{N}_{t-1})w_j(\mathbf{FF}_{t|t-1}), \quad j < t \tag{15}$$

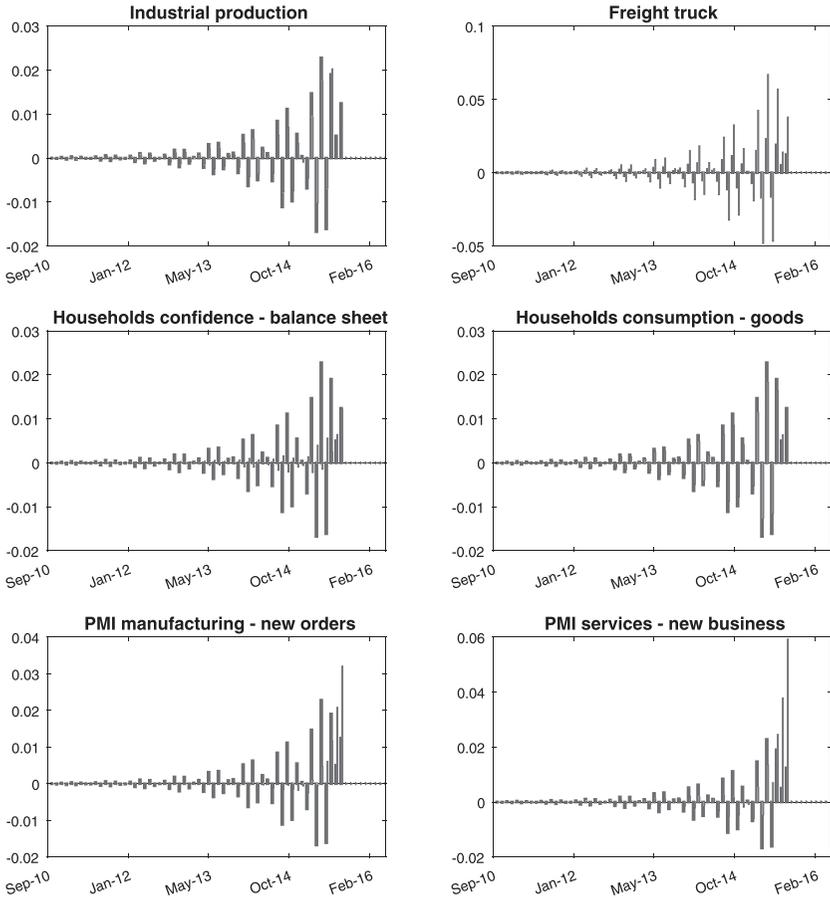
$$w_j(\mathbf{FF}_{t|T}) = B_{t,j}^*\mathbf{C}_j \quad B_{t,j+1}^* = B_{t,j}^*\mathbf{L}_j', \quad j = t, \dots, T \tag{16}$$

with $B_{t|t}^* = P_{t|t-1}$.

Figure 4 shows the weights attached to T2-retail (black bars) and to some other indicators (gray bars), which track the short-term evolution of economic activity very well (such as the

²⁰A detailed description of the algorithms may be found in Koopman and Harvey (2003, p. 1322).

Figure 4. Weights for Nowcasting (PS: black bar; other variable: gray bar)



industrial production index; freight truck; households' confidence about their own balance sheet; households' consumption of goods; manufacturing PMI—new orders; and service PMI—new business). Let us assume that the current month is October 2015. Therefore, we need to anticipate GDP growth 2015:Q4 (nowcasting).

The weights of PS are comparable to those of industrial production and of households' consumption of goods; the latest available data on PS help make up for the missing information on industrial

production and households' consumption at the end of the sample. PS have sizable weights compared with those of the qualitative surveys (PMIs and households' confidence about the balance sheet); this means that past information on PS is more effective than that on very timely and cyclical indicators in anticipating GDP growth.

5. Conclusions

Our findings show that payment data track economic activity. We look at different aggregates of payment system flows in Italy, together with other indicators usually adopted in macroeconomic 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 the payments jointly with other standard business cycle indicators (e.g., industrial production and business surveys), for both GDP and households' consumption. Moreover, an out-of-sample forecasting application using a mixed-frequency factor model shows that the model including retail payment flows generally outperforms the one based on standard short-term indicators only, in terms of forecasting accuracy. The results are shown not only for GDP but also for consumption, investments, and value added in the service sector. 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 economic activity. Using the mixed-frequency feature of our model, we show that the timeliness of the payment data improves the forecast accuracy during the quarter, more than other comparable short-term indicators.

The way we pay is changing following the digitalization of retail payments, and this process fosters the production of big data, which the analysts can rely on. Payment data indicators used in this model are the aggregate of a huge number of transactions. However, the forecasting ability of these indicators paves the way for future research exploring the big data structure of the individual transactions.

References

- Bagnall, J., D. Bounie, K. P. Huynh, A. Kosse, T. Schmidt, S. Schuh, and H. Stix. 2016. "Consumer Cash Usage: A Cross-Country Comparison with Payment Diary Survey Data." *International Journal of Central Banking* 12 (4): 1–61.
- Bai, J., and S. Ng. 2002. "Determining the Number of Factors in Approximate Factor Models." *Econometrica* 70 (1): 191–221.
- . 2007. "Determining the Number of Primitive Shocks in Factor Models." *Journal of Business and Economic Statistics* 25 (1): 52–60.
- . 2008. "Forecasting Economic Time Series Using Targeted Predictors." *Journal of Econometrics* 146 (2): 304–17.
- Barnett, W., M. Chauvet, D. Leiva-Leon, and L. Su. 2016. "Nowcasting Nominal GDP with the Credit-Card Augmented Divisia Monetary Aggregates." Working Papers Series in Theoretical and Applied Economics No. 201605, University of Kansas, Department of Economics.
- Boivin, J., and S. Ng. 2006. "Are More Data Always Better for Factor Analysis?" *Journal of Econometrics* 132 (1): 169–94.
- Carlsen, M., and P. E. Storgaard. 2010. "Dankort Payments as a Timely Indicator of Retail Sales in Denmark." Technical Report.
- Duarte, C., P. M. Rodrigues, and A. Rua. 2017. "A Mixed Frequency Approach to the Forecasting of Private Consumption with ATM/POS Data." *International Journal of Forecasting* 33 (1): 61–75.
- Esteves, P. S. 2009. "Are ATM/POS Data Relevant When Nowcasting Private Consumption?" Technical Report.
- Fisher, I. 1912. *The Purchasing Power of Money*. New York: Macmillan Company.
- Forni, C., M. Marcellino, and C. Schumacher. 2015. "Unrestricted Mixed Data Sampling (MIDAS): MIDAS Regressions with Unrestricted Lag Polynomials." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 178 (1): 57–82.

- Galbraith, J. W., and G. Tkacz. 2007. "Electronic Transactions as High-Frequency Indicators of Economic Activity." Technical Report.
- . 2009. "A Note on Monitoring Daily Economic Activity via Electronic Transaction Data." CIRANO Working Paper No. 2009s-23.
- . 2018. "Nowcasting with Payments System Data." *International Journal of Forecasting* 34 (2): 366–76.
- Ghysels, E., P. Santa-Clara, and R. Valkanov. 2004. "The MIDAS Touch: Mixed Data Sampling Regression Models." CIRANO Working Paper No. 2004s-20.
- Giannone, D., L. Reichlin, and D. Small. 2008. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Journal of Monetary Economics* 55 (4): 665–76.
- Hastie, T. J., R. J. Tibshirani, and J. H. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics. New York: Springer.
- Kim, C.-J., and C. R. Nelson. 1999. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, Vol. 1. The MIT Press.
- Koopman, S. J., and A. Harvey. 2003. "Computing Observation Weights for Signal Extraction and Filtering." *Journal of Economic Dynamics and Control* 27 (7): 1317–33.
- Marcellino, M., and C. Schumacher. 2010. "Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP." *Oxford Bulletin of Economics and Statistics* 72 (4): 518–50.
- Padoa-Schioppa, T. 2004. "Shaping the Payment System: A Central Bank's Role." Speech delivered at the Bank of Korea's Conference on Payment Systems, Seoul, May 13. Available at https://www.ecb.europa.eu/press/key/date/2004/html/sp040513_1.en.html.
- Schmiedel, H., G. Kostova, and W. Ruttenberg. 2012. "The Social and Private Costs of Retail Payment Instruments: A European Perspective." ECB Occasional Paper No. 137.
- Schneider, M., and M. Piazzesi. 2015. "Payments, Credit and Asset Prices." 2015 Meeting Paper No. 133, Society for Economic Dynamics.

- Tibshirani, R. 1996. "Regression Shrinkage and Selection Via the Lasso." *Journal of the Royal Statistical Society: Series B (Methodological)* 58 (1): 267–88.
- Williamson, S., and R. Wright. 2010. "New Monetarist Economics: Models." In *Handbook of Monetary Economics*, Vol. 3, ed. B. M. Friedman and M. Woodford, 25–96 (chapter 2). Elsevier.