Data of Sectoral Financial Flows as a High-Frequency Indicator of Economic Activity¹

Natalia Turdyeva, *Bank of Russia*

turdyevana@mail.cbr.ru

Anna Tsvetkova, *Bank of Russia*

tsvetkovaan@mail.cbr.ru

Levon Movsesyan, *Bank of Russia*

movsesyanls@mail.cbr.ru

Alexey Porshakov, *Bank of Russia*

porshakovas@cbr.ru

Dmitriy Chernyadyev, *Bank of Russia*

chernyadevdn@mail.cbr.ru

In times of crisis, events are moving fast and standard macroeconomic statistics published with a lag cannot quite keep pace with the changing situation. During such periods, there is an increasing need to use high-frequency indicators that allow virtually real-time monitoring of economic activity. In many countries, this is achieved by using financial transaction data. In this paper, we present a methodology for the current analysis of sectoral financial flows in the Russian economy based on data from the Bank of Russia payment system. We use the information on the dynamics of average daily payments for each class of OKVED 2 (the Russian National Classifier of Economic Activities) to develop high-frequency indicators of economic activity, which have been published on the Bank of Russia website since April 2020. We also tentatively discuss the potential of financial transaction data in terms of improving the tools for short-term forecasting of business activity dynamics and solutions to other research problems.

Keywords: *Bank of Russia payment system, high-frequency data, sectoral financial flows, nowcasting, coronacrisis*

JEL Codes: C80, C81, E42

Citation: Turdyeva, N., Tsvetkova, A., Movsesyan, L., Porshakov, A. and Chernyadyev, D. (2021). Data of Sectoral Financial Flows as a High-Frequency Indicator of Economic Activity. *Russian Journal of Money and Finance*, 80(2), pp. 28–49.

doi: 10.31477/rjmf.202102.28

¹ The authors are grateful to Konstantin Styryn and Ramis Khabibullin for their help and comments.

1. Introduction: COVID-19 shock and the need for high-frequency data for assessing current economic activity

In many ways, the economic shock caused by the pandemic is unique. This uniqueness was displayed not only in the scale of the economic downturn but also in the speed at which the crisis spread and in its very heterogeneous impact on various sectors of the economy. The negative consequences of the crisis are certainly not limited to the primary effects experienced by those sectors of the economy directly affected by the restrictive measures imposed in 2020 (primarily the service sector). Through intersectoral links, they spread to other sectors, and these secondary effects are significant in scale and quite long-lasting (see, in particular, Ponomarenko et al., 2020). Thus, in the context of the coronacrisis, even when the most acute phase of the pandemic has passed and the epidemiological situation is gradually returning to normal, the issues related to the long-term implications of the crisis and the peculiarities of the structural adjustment of the economy represent an important area for economic analysis.

The last 15 years have been marked by a substantial increase in the capabilities and extent of using the real-time assessment (nowcasting) of economic indicators. Nowcasting becomes especially important during economic crises. The global nature of the coronacrisis and the unprecedented high uncertainty associated with it amid extremely rapid changes in the situation when countries introduced anti-epidemic restrictions have required governments and central banks around the world to adopt new analytical approaches in their decision-making on anti-crisis measures to support the economy which are based as much as possible on live data streams. As a result, in addition to traditional macroeconomic indicators, which are not always able to keep up with rapid changes in the economic situation, the world began to use high-frequency data.

Financial transaction data, which reflects both consumer activity and financial flows between firms, has gained widespread use in various countries. In Russia, one such tool for rapid assessment of economic activity is the analysis of data from the Bank of Russia payment system (BoR PS), which enables almost real-time monitoring of the situation in economic sectors on a daily basis.

It should be noted that this analysis can be used to address several applied problems. First, the use of sector-specific data on financial flows helps to identify turning points in business activity dynamics as quickly as possible and to monitor the situation at the level of individual sectors and their combined groups. Second, as the impact of the pandemic gradually weakens and the economy starts out on the path of recovery, data on sectoral payments, along with other current indicators, is potentially useful in terms of analysing structural changes in the economy and identifying the sectors that will later become the drivers of recovery growth. This is extremely important for forecasting and developing various measures of macroeconomic policy.

Our paper aims to achieve two key objectives. The **first objective** is to look through the prism of existing literature in order to provide a detailed overview of the main methodological issues related to assessing the dynamics of business activity using high-frequency data and, in particular, data from payment systems. The **second objective** is to use the example of the Bank of Russia to clearly demonstrate the opportunities for using daily data from the payment system both for current monitoring of business activity in the acute phase of the crisis and for analysing the path of subsequent economic recovery and the structural shifts associated with such recovery.

In this paper, we present a methodology for building a high-frequency indicator we have developed for the average daily incoming payment in each class of OKVED 2 (the Russian National Classifier of Economic Activities). When analysing the financial flows, we aggregate the volume of incoming payments to the sectoral level (OKVED 2 classes) and carry out preliminary processing of data, including seasonal adjustment of time series based on the model developed by the Bank of Russia.

The resulting sectoral series of data on financial transactions allow us to assess the dynamics of payments in the economy as a whole and in its individual sectors. In 2020, we used the deviation of the average daily incoming payment from its 'normal' level, by which we mean the average daily incoming payment in the period from 20 January to 13 March 2020, as the key weekly indicator of economic activity based on transaction data.

Along with the indicators for each individual sector, we aggregate the results in order to generate composite indicators for the entire economy. To do this, we weight the resulting deviations from the 'normal' level using the weights corresponding to the shares of sectors in the structure of gross value added (GVA). We also develop several composite indicators by adjusting the weights of the sectors based on their categories of use. They reflect the dynamics of economic activity in sectors focused on export, consumer, intermediate, investment, and government demand.

The highly granular sectoral monitoring of financial payments allows us to assess the impact of the oil sector on the overall dynamics of economic activity. In particular, financial transaction data enable us to identify positive trends in most sectors during the summer of 2020, when the overall recovery trend was restrained by cuts in oil production.

We use BoR PS data to identify the sectors that fared the best during the pandemic, as well as those whose revenues sharply declined. The resulting indicators are consistent with the dynamics of economic activity in these sectors according both to Rosstat estimates and estimates from other alternative sources.

Since April 2020, the Bank of Russia has been publishing indicators based on financial flow data of economic activity by sectors, groups of sectors and the economy as a whole on a weekly basis. Financial transaction data provide a large amount of information. They can be used to address various research and applied problems, which helps improve short-term forecasting.

The remaining part of the paper is structured as follows. In Section 2, we describe the global experience of using high-frequency data to assess economic activity. In Section 3, we present a description of BoR PS data and the methodology for its preprocessing. In Section 4, we demonstrate the dynamics of composite indicators based on payment data for the economy as a whole and for groups of sectors. In Section 5, we provide examples of our analysis for individual sectors. We compare them both with traditional statistical indicators and alternative indicators of economic activity in these sectors. Section 6 discusses the potential uses of financial payment data in economic analysis and research.

2. Global experience: Real-time monitoring of economic activity using payment system data

Researchers and analysts began using payment system data several years ago, long before the recession associated with the COVID-19 pandemic.

Galbraith and Tkacz (2015) describe the nowcasting of Canada's GDP based on electronic payment data. The paper uses aggregated information on transactions that pass through the payment system, including the amounts of transactions with debit and credit cards, as well as those using bank cheques. This data is collected electronically and, thus, is convenient, as it can be accessed immediately and has virtually no errors. The authors conclude that the three reviewed time series are informative for monitoring economic activity, especially during recessions (the data clearly shows the recession of 2008–2009). They claim that, among the reviewed payment data, debit card transactions give the highest accuracy of forecasts.

Duarte et al. (2016) focus on forecasting and nowcasting households' consumption in Portugal by using high-frequency data on ATM withdrawals and point-of-sale payments collected by the Bank of Portugal. With regard to households' consumption, the reviewed data excludes durable goods, such as vehicles, since this type of expense is commonly paid using bank transfers or loans.

In addition to data from ATMs and cash terminals, several other typical predictors, such as monthly retail turnover and indexes of consumer confidence, are considered in forecasting quarterly households' consumption. These indicators are among the most popular and make it possible to evaluate the benefits of using the data from ATMs and points of sale for the established objectives. The authors found that the use of high-frequency data improves the quality of forecasting and nowcasting and the data from ATMs and cash terminals ensures the greatest gain in accuracy.

Aprigliano et al. (2017) build forecasts and nowcast of economic activity in Italy using data from payment systems (retail payment data). The authors examine the entire range of payment instruments, such as credit transfers, bank cheques, direct debiting, and debit cards. Information about them is recorded electronically through clearing and settlement schemes managed by the Bank of Italy (BI-COMP and T2-retail) and is characterised by efficiency, high frequency, and the absence of errors

in data collection. The aggregated monthly time series used for the analysis are stored in the Bank of Italy's publicly available database. The authors note that the use of data on the entire range of payment instruments increases the reliability of forecasting compared to similar studies that consider only a subset of such instruments.

The researchers conclude that changes in the dynamics of GDP and its main components are closely related to changes in payment flows through payment systems. The correlation between dynamics of GDP and changes in payment flows is similar to the correlation between the main macroeconomic aggregates and such indicators as the industrial production index and the business confidence index, which are commonly used in short-term forecasting of GDP and its components.

The paper describes a mixed-frequency dynamic factor model for predicting quarterly GDP growth using a wide set of monthly data, which includes standard indicators of the economic cycle, such as electricity consumption, industrial production index, consumer and business surveys data, current account deposits, loans to firms, consumer price index, stock indexes, Purchasing Managers' Index (PMI), indicators of freight transportation and trade, and some others. When the results of out-of-sample forecasting using a model that includes payment data are compared with the results of similar forecasting without such data, it reveals that the contribution of payment data to the accuracy of forecasts is substantial.

The coronacrisis led to a large number of new studies on monitoring and forecasting economic activity, including by using financial transaction data.

Carvalho et al. (2020) used the data from credit and debit card transactions and the point-of-sale terminals of Spain's second-largest bank to nowcast aggregate consumption, which is the main component of GDP. In addition, the data on card spending is compared with the results of the official household budget survey conducted by the national statistical office. Finally, the study analyses the mobility of the population based on information about the transport expenses of bank card holders.

The authors of the study argue that credit and debit card expenses are good indicators of consumption both in general and within certain categories of consumption. The cross-country analysis, as well as the analysis at the level of Spanish provinces, suggest a significant and sharp decline in spending following the introduction of coronavirus lockdown measures. The data also indicate a subsequent V-shaped recovery in spending.

A comparison of consumer spending based on bank card data and a survey of household budgets demonstrates that these transactions are highly informative. When official household budget survey data is not available, information on credit and debit card spending can replace consumer surveys.

The authors used the high geographic granularity of the data set to find out that the residents of the wealthiest neighbourhoods in Madrid (based on their postal codes) experienced the greatest decline in spending. During the lockdown, consumption was reallocated from luxury goods to basic necessities and, while the

restrictions were in force, a representative consumer had a consumer basket more like the basket of a relatively poor consumer before the lockdown.

The authors consider data from the Google Mobility Report for Spain (where changes in mobility are tracked with user logins to Google services) as a traditional indicator of population mobility. They found that changes in the alternative indicator, transport expenses based on bank card data, very accurately track changes in population mobility. The disaggregation of transport expenses by Madrid postal code suggests that, during mobility restrictions, cardholders in low-income areas travel more during the working week than those in higher-income areas, but weekend travel expenses are more similar among different income groups. This is consistent with the fact that low-income households have to travel to work during the lockdown, while higher-income households have more opportunities for self-isolation, as they belong to professions more likely to work from home.

Barlas et al. (2020) analyse the feasibility of using data on bank transactions between firms in Turkey in short-term forecasting. The authors test whether the inclusion of the aggregated indexes of consumption and investment based on bank transaction data in the dynamic factor model improves the accuracy of nowcasting of GDP and investments in Turkey.

The nowcasting model of GDP dynamics includes variables related to ‘hard’ data (quantitative data collected by statistical agencies), such as the GDP itself (quarterly), industrial production (monthly), non-metallic mineral production (monthly), electricity production (daily), car sales (monthly), employment growth (monthly), unemployment growth (monthly), car exports (monthly) and car imports (monthly). The model also includes ‘soft’ data (surveys of business executives, households or experts), such as the Manufacturing PMI (monthly) and real sector confidence index (monthly). Finally, the alternative model includes the daily consumption and investment indexes proposed by the authors of the study.

The investment nowcasting model includes ‘hard’ data, such as gross fixed capital formation (quarterly), capital goods production (monthly), non-metallic mineral production (monthly), import volume of capital goods (monthly) and sales of commercial vehicles (monthly). The model also includes ‘soft’ data, such as the real sector confidence index (monthly) and financial variables, including the growth of corporate loans in real terms (monthly). Finally, an investment index based on financial flow data was added to the alternative model.

The results of the analysis indicate increased accuracy in nowcasting investments and GDP (in terms of reducing the out-of-sample root-mean-square error) when using investment index based on financial flow data. For GDP, the increase in the accuracy of nowcasting is smaller, since investments are only one component of GDP, and a smaller one than consumption.

The authors also used the example of two recent crises in Turkey (the financial crisis in August–September 2018 and COVID-19 crisis in March 2020) to analyse the extent to which the investment (and consumption) index based on big data

helped the benchmark model (not involving the use of big data) to predict dramatic changes in economic activity. The researchers conclude that the investment index based on big data plays an important role in forecasting the effects of the crisis in the summer of 2018. The predictive role of the investment index during the COVID-19 crisis was insignificant, but the model that includes the consumption index has its benefits, as it predicted a downturn in the economy almost a month earlier (at the end of March) than the benchmark model.

In addition, when considering the contribution of various factors to nowcasting with the same time interval for all variables (big data on investments is available only since 2013), it is clear that information on investments has very high relevance. Only industrial production, for which the statistics are published with a significant lag, has a greater contribution.

The authors demonstrated that the contribution of investment and consumption data based on financial flow data is particularly important at the beginning of the quarter, when the availability of information on real variables is low. Their study indicates the increasing contribution of indexes based on financial flow data to GDP nowcasting. Finally, the authors argue that the relevance of big data is particularly high during crises, which makes it an important tool for predicting downturns in the economy, enabling a prompt response to changing conditions.

The authors specifically note that investment data provides additional benefits besides providing real-time information about the investment cycle. Since information is recorded at a very granular level, this enables analysis of investment data by sector, which, in turn, makes it possible to track the response to the pandemic shock. Investments in sectors of the Turkish economy that are highly integrated into global value chains (such as the textile and automotive industries) experienced a temporary shock, as they were affected by the closure of foreign markets. Tourism and the entertainment industry have been hit particularly hard by the mobility restrictions.

Another important benefit of big data is the geographical details of information. This makes it possible to track investment inflows both by sector and by region and to study the geographical distribution of investments and the response of different regions within the country to the investment shock. For example, in Turkey, major cities, such as Istanbul or Ankara, have not been affected as badly as other regions. The strongest decline in investment was observed in the western regions of the country, where the manufacturing industry and export-oriented production facilities are located.

Chetty et al. (2020) create a database that can be used to quickly estimate consumer spending, business income, employment levels, and other key indicators that help to assess the state of the US economy. The researchers use high-frequency, granular and anonymised data collected by private companies to monitor and analyse the effects of the COVID-19 pandemic and the subsequent economic recovery virtually in real time.

Figure 1. Investment dynamics and the index built with financial transaction data, %, YoY, Turkey

Source: Barlas et al., 2020

Information on consumer spending is provided by credit and debit card data collected by the private company to support various financial services, such as the loyalty programmes of banks. The raw data is provided by county, zip code, income group, sector, and date. To account for cash transactions, the authors use transaction data from the company that enables people to earn rewards by uploading photos of their receipts to a mobile app (only those receipts that specify a cash payment are selected). Small business revenue and transaction data is provided by the firm that aggregates data on credit card spending to provide analytics to its customers. The advantage of this data is that it uses the location of the companies rather than the place of the cardholder's residence.

The statistics collected by the authors shows that high income individuals sharply reduced their spending in mid-March 2020, especially in regions with high rates of COVID-19 and in sectors where physical interaction is required. This led to a significant decline in the income of small businesses located in affluent areas and increased unemployment, especially among low-paid workers. It is also noted that the effect of government measures was insignificant. Additional payments to low-income households significantly stimulated consumer spending. However, only a small fraction of this spending went to the businesses most affected by the coronavirus shock and, as a result, these measures failed to make a positive impact on employment.

In the future, the created database can be used for the current analysis of regional heterogeneity, since different areas and sectors of the economy are often exposed to different shocks, and they are affected by different measures of local economic policy. The analysis of such heterogeneity can help to quickly diagnose key factors that lead to economic downturns.

3. The Bank of Russia's experience: BoR PS data

3.1. Data description

In Russia, the analysis of financial flow data has become a tool for the rapid assessment of economic activity. The source of such data is the BoR PS. The source data is presented with a daily frequency and is broken down by region and by sector in accordance with the all-Russian classifier of types of economic activity (86 sectors assigned with two-digit OKVED 2 class codes). Historical data is available from 1 January 2016.

Financial flow data has a number of features that differentiate it from indicators of real economic activity. First, the payment volumes include financial transactions that are not related to the sale of goods or payment for services. Examples of such transactions include mergers and acquisitions, securities transactions, rouble transfers during currency conversion, etc. It should be clarified that we do not monitor transactions where the Taxpayer Identification Numbers (known in Russia as INNs) of the payer and recipient are the same. In other words, we do not consider transactions within a company or between the units and branches of the same organisation. Second, the volumes of financial payments are nominal indicators that reflect changes not only in the physical volumes of production, the turnover of goods, or the volume of provided services, but also the dynamics of their prices.

3.2. Data preprocessing: Seasonal adjustment and winsorisation

Before using the data for analysis, we pre-process it. As a rule, financial payments are not processed in BoR PS on weekends and public holidays. However, due to technical reasons, certain payments are registered on such days. The volume of such transactions is very small compared to the flows on working days and, at the first stage of data preprocessing, such observations were excluded from the analysis. The second and third stages of data preprocessing include seasonal adjustment and winsorisation.

At the second stage of preprocessing, we aggregate data by region and obtain time series corresponding to 86 sectors. To these series we apply the procedure for seasonal adjustment, which is just as necessary for daily data as it is for traditional monthly macroeconomic statistics.

The methods widely used for seasonal adjustments can be applied to data of monthly or even quarterly frequency (Lee, 2018). In recent years, packages for seasonal adjustment of daily data have become common (Taylor and Letham, 2018; Ollech, 2018). The Prophet package (Taylor and Letham, 2018) addresses the tasks that are the closest to the goals of the Bank of Russia. However, given the characteristics of BoR PS data, such as the values missed on weekends and holidays, the use of the Prophet package creates additional operational complexity

in building and maintaining the model due to the effects of calendarisation and the variability in the choice of a regressor for missed values. The Bank of Russia developed a procedure for seasonal adjustment of daily financial data that accounts for their specifics (Seleznev et al., 2020). This methodology of adjustment makes it possible to feed BoR PS data into the model input without additional processing.

The procedure (Seleznev et al., 2020) assumes that the time series includes a set of multiplicative components, such as a non-stationary time trend tr_t , a stationary component e_t and several seasonal components. In turn, the set of seasonal components includes an intra-year component (s_t^y), an intra-month component (s_t^m), and an intra-week component (s_t^w).

The intra-week component s_t^w is modeled as a set of dummy variables. The intra-month s_t^m and intra-year s_t^y components are modeled using a set of trigonometric functions that approximate periodic functions. This set was selected to be large enough in order to prevent seasonal variations from persisting in the adjusted series. Bayesian regression (Sparse Bayesian Learning, SBL) with automatic selection of hyperparameters (Tipping, 2001) was used to prevent overfitting. A variational Bayes algorithm with mean-field approximation (Khabinullin and Seleznev, 2020) was used to calculate the maximum likelihood. This makes it possible to easily modify the SBL model for other tasks.

When there are no signs of overfitting or underfitting, the developed SBL model accounts for the characteristics of financial data and demonstrates a similar or better quality of forecasts compared to the already known models used for seasonal adjustment of daily data (see Seleznev et al., 2020, Table 1).

There is a number of limitations in the basic procedure of seasonal adjustment that must be taken into account when interpreting the results. First, the basic procedure cannot consider the evolution of the seasonal component since, for the model, only repeated patterns can be viewed as seasonality.

Second, floating seasonality may also be taken into account incorrectly. An example of floating seasonality can be the peaks of payments that fall on a certain day of the month. However, if that day falls on a weekend, the payments are made on the previous or subsequent working days and, in rare cases, they are postponed to the next month. In some situations, the postponement of payments to another day can be interpreted as a significant decrease or increase in the volume of payments in the adjusted series. However, in fact, these are only seasonal fluctuations.

Third, as with other seasonal adjustment models, there is a tail-wagging effect. At the last points of the time series, the fluctuations in the volume of payments can be interpreted as an increase or decrease that is not related to seasonality. Although it is likely that the observed fluctuations reflect changes in seasonal patterns and, accordingly, should be excluded from the adjusted series.

In the next, third stage of data preprocessing, we apply winsorisation in each of the 86 sectors. The purpose of this procedure is to exclude the impact of individual very large payments. They are often purely financial in nature, such as transfers in

mergers and acquisitions, currency conversions, and so on, and do not reflect current economic activity. For this purpose, in each sector, we defined a minimum level of 0.5%² of the largest total daily payments for the period from 1 January 2016 to 31 December 2020. On days when the volume of payments exceeds this level, the initial values are replaced with threshold values. After preprocessing data for each of the 86 sectors, we get seasonally adjusted and winsorised series of financial flows.

3.3. Assessing the dynamics of economic activity

The analysis of data on incoming financial flows that has been preprocessed as described above makes it possible to assess the dynamics of economic activity in each sector corresponding to a two-digit OKVED 2 code (86 sectors).

We calculate the average daily incoming payment for each industry on a weekly basis. This allows us to maintain comparability between weeks. When using the total weekly volumes of financial flows, we always observed a decline in those weeks where the number of working days is less than five. Also, for each sector, we calculated the average daily incoming payment for the period from 20 January to 13 March 2020. We took this value as the 'normal' level. This choice was caused by the need to carry out operational analysis of the state of mutual settlements at the sectoral level and promptly make results public during the first wave of the novel coronavirus infection. The 'normal' period covers those weeks when the effect of the first days of the year has already subsided, while the implications of the pandemic have not yet been felt.

As the COVID-19 shock faded and we moved away in time from the 'normal' period of 2020, it became clear that the chosen comparison metric was losing its relevance. As a result, in 2021, we decided to abandon comparing the volume of incoming payments with the 'normal' period and began to use quarterly dynamics. We continue to improve our analysis approaches, in particular by taking into account the change in the trend after the COVID-19 shock. However, in this paper, we describe the practice of calculating the indicators in 2020.

As the key sectoral indicator for monitoring the sectoral financial flows in 2020, we use the deviation of the calculated average daily volume of incoming payments from the 'normal' level in the given sector. The analysis of data on outgoing payments is for reference only.

With available estimates of deviations in the volume of incoming revenues from the 'normal' level for each sector, we also develop aggregated indicators of economic activity. The key composite indicator is the one 'for the economy as a whole'. In addition, to assess the intensity of incoming financial flows, excluding the impact of the most volatile components, we also use such indicators as 'for the economy as a whole excluding the extraction of crude oil and gas and the production of petroleum products' and 'for the economy as a whole excluding the extraction and production of petroleum products and the activities of state authorities'.

² The cut-off level in the winsorisation procedure is the same for all sectors; it was determined using the expert method based on the distribution of outliers across all classes of OKVED 2.

The aggregated indicators are built as a weighted average deviation from the ‘normal’ level of all 86 sectors (OKVED 2 classes):

$$Income_flow_index_{total} = \sum_{i=1}^{86} w_i \times income_flow_index_i,$$

where $Income_flow_index_{total}$ is the composite indicator of deviation in the volume of incoming payments from the ‘normal’ level for the economy as a whole, and $income_flow_index_i$ is the deviation from the ‘normal’ level in sector i . The 86 sectors include OKVED 2 classes from 1 to 97. w_i are the weights of the sectors, which we calculate by using the GVA structure for 2018 as the basis.

In some cases, the source data on GVA structure published by Rosstat combines several classes of OKVED 2 into one subgroup (86 sectors are combined into 61 subgroups).³ Therefore, the total weight of a specific OKVED 2 class is derived by multiplying the weight of that class in the subgroup by the weight of the subgroup in the total GVA:

$$w_i = W_l \times w_i^{within\ group},$$

where W_l is the weight of a subgroup of sectors l which includes the sector i in the GVA structure in 2018, and $w_i^{within\ group}$ is the weight of a specific OKVED 2 class in the group derived from other data sources. To calculate the weights $w_i^{within\ group}$ in the extraction sectors (5–9)⁴ and manufacturing industries (10–12 and 13–15), we use the added value structure for industrial production in 2018. In the case of financial services (64–66), we use data on the GVA for the three types of services (financial intermediation services, insurance, and non-governmental pension services other than compulsory social insurance services, and auxiliary services in the area of financial intermediation) from the symmetric input-output table for 2016:

$$w_i^{within\ group} = \frac{VA_i}{\sum_{i=1}^3 VA_i},$$

where VA_i is the value added for the type of service i . Similarly, we calculate intra-group weights $w_i^{within\ group}$ in other sectors using data on the revenue structure for 2018 (from information on financial results of organisations for 2018 by type of economic activity).

The weights for the composite indicators ‘for the economy as a whole excluding the extraction of crude oil and gas and the production of petroleum products’ and ‘for the economy as a whole excluding the extraction and production of petroleum products and the activities of state authorities’ were obtained by reweighting the already calculated weights up to 100% after excluding the corresponding industries.

We also grouped the dynamics of financial payments by sector in accordance with the categories of usage. As a result, we have five indicators for sectors oriented on external demand, final consumption, intermediate consumption, investment

³ See <https://rosstat.gov.ru/storage/mediabank/tDfpHWR9/str2.xls>

⁴ Hereinafter, the OKVED 2 code of a sector is indicated in parentheses after the name of the sector.

demand, and government demand. The weights of sectors are based on their shares in GVA and data on the structure of usage for goods and services in basic prices obtained from the system of input-output tables for 2017. The source data in the table of goods and services usage is presented by subgroups corresponding to the subgroups of sectors in the GVA structure:

$$W_l^j = \frac{W_l \times k_l^j}{\sum_{l=1}^{61} W_l \times k_l^j},$$

where W_l^j is the weight of the l -th subgroup of sectors in j -th use category; W_l is the weight of a subgroup of sectors in the GVA structure; and k_l^j is the share of the j -th category of usage in the total usage of products of the subgroup of sectors l in 2017.⁵ After calculating the weights for the subgroups of sectors, we drill down the data to the individual sectors (OKVED 2 classes) in the same way as we calculated the weights for the aggregate indicator of the economy as a whole.

It is worth noting that many industries are included in the calculation of multiple aggregated indicators (rather than a single one) by usage category, since their products are used in various areas, including exports, processing, or consumption by end users.

Thus, by using financial flow data, we can estimate the dynamics of payment volumes in individual sectors, as well as in the economy as a whole, based on the structure of GVA.

4. Composite indicators

The analysis of sectoral financial flows clearly shows the scale of the coronacrisis, which peaked in early April 2020. Figure 2 illustrates the dynamics of the indicator describing the deviation in the volume of incoming financial flows from the 'normal' level for the economy as a whole (total for GVA), as well as the dynamics of a modified indicator that excludes the most volatile components (i.e. excluding crude oil and gas extraction, petroleum products production, and public administration).

In the summer, the gap of incoming payments from its 'normal' level began to close but, in the fall, this process slowed down somewhat. By the end of the year, the volume of incoming payments had almost reached the 'normal' level. However, it should be borne in mind that this is a nominal indicator and, in normal times, a positive trend over the course of the year is a normal phenomenon. In other words, the return of incoming payments to near pre-pandemic levels does not yet indicate a full recovery of economic activity.

⁵ The five use categories (external demand, final consumption, intermediate consumption, investment demand, and government demand) correspond to the columns 70, 63, 62, 69, and 64 + 65 in the table of goods and services usage, respectively. The total use of goods and services corresponds to column 72, which is the sum of intermediate and final use.

Figure 2. The dynamics of deviation of incoming payments (weighted by the share of the sectors in GVA) from the 'normal' level, %

Source: Bank of Russia, *Monitoring of Sectoral Financial Flows*⁶

Figure 3. Weekly deviations of incoming payments from their 'normal' level for aggregated groups of sectors, %

Source: Bank of Russia, authors' calculations

For Russia as a whole, the relatively rapid growth of incoming financial flows after the peak of the crisis was extremely heterogeneous. The analysis of financial payments makes it possible to assess the sectoral restructuring in 2020 that was triggered by the coronacrisis. Figure 3 illustrates the differences in the dynamics of composite indicators by category of usage based on data about the structure of value added and the structure of goods and services usage. At the peak of the crisis, the largest drop among the aggregated groups of sectors was observed in export-oriented industries, while the consumer goods sectors suffered the least. The composite

⁶ See <https://www.cbr.ru/analytics/finflows/>

indicator for the groups of the consumer goods sector shows that this sector became the main driver of growth in the period after the most acute phase of the crisis.

5. Sectoral dynamics

We can use financial flow data to assess the impact of the crisis on individual sectors of the Russian economy with even greater granularity. We compare the dynamics of financial flows by sector and nominal gross value added (GVA). To do this, we calculate the quarterly growth rate of the average daily incoming payment year-on-year for each sector (OKVED 2 class). Then we calculated the weighted average growth rate for the entire sector:

$$g_t^J = \sum_{i \in J} w^i \times \left(\frac{fin_inflows_t^i / wd_t}{fin_inflows_{t-4}^i / wd_{t-4}} - 1 \right),$$

where g_t^J is the year-on-year growth rate in the sector J in the quarter t ; $fin_inflows_t^i$ is the volume of incoming financial flows in the sector (OKVED 2 class) i related to the sector J in the quarter t ; wd_t is the number of working days in a quarter t ; w^i is the weight of the sector (OKVED 2 class) in the sector J . The weights w^i were calculated using the data on GVA structure, as well as the weights calculated for the composite indicator of the economy as a whole (see Section 3.3).

Our analysis reveals a fairly close relationship between the indicators of nominal GVA growth rates and growth rates of incoming payments by sector (in 2020Q2, the pair correlation was 56%) (see Figure 4). This suggests that the data on these sectoral financial flows can be used as the leading indicators of business activity for certain types of activity.

Figure 4. Increase in nominal GVA and incoming payments by sector, YoY, %

Source: Rosstat, Bank of Russia, authors' calculations

5.1. Dynamics of financial revenues in shock-resistant sectors

Amid the slump in economic activity across most sectors, some areas, on the contrary, experienced an increase of financial flows. Shock-resistant sectors include the production of textiles (13) and the production of medicines (21). As an indicator to describe the dynamics of these sectors, we are using the volume of financial flows, as well as Rosstat indicators representing the basic indexes. Therefore, we also use the data on the volume of incoming payments to build the basic indexes. To do this, we calculate the average daily volume of payments for each month based on seasonally unadjusted data. Then seasonality is adjusted by the standard TRAMO/SEATS method (which is possible because we are moving to the monthly frequency of data). The time series of financial flows in each sector are reduced to their basic form, where the average seasonally adjusted payment in 2019 is taken as 100%.

In the production of textiles (13) and in the pharmaceutical industry (21), an increase in incoming payments preceded the growth of the industrial production index, which we compare with the volume of incoming payments (Figure 5). It is worth noting that the dynamics of incoming financial flows in the pharmaceutical industry (21) indicated that the positive trend let up in September–October. This was due to difficulties in implementing a system for monitoring the circulation of medicines.⁷ The production statistics for these months confirmed the temporary drop. However, the last months of the year saw a resumption of growth in both financial flows and output.

5.2. Dynamics of financial revenues in the sectors with the largest decrease in incoming payments

Two groups can be identified among the sectors that were hit the hardest by the coronacrisis. The first group includes export-oriented sectors. After the most acute phase of the crisis, the volume of revenues in the export-oriented group remained at a lower level virtually until the end of 2020.

The main contribution was made by the extraction of crude oil and gas (6) following the decline of world prices and subsequent production cuts under the OPEC+ agreement. Since the extraction of crude oil and gas (6) contributes significantly to GVA for Russia as a whole (its share is 10.6%), the sector had a substantial impact on the composite indicator of economic activity which we publish in Monitoring of Sectoral Financial Flows (deviation of the volume of financial flows from its 'normal' level) (Figure 6). The financial payment data made it possible to quickly identify the positive trends that began during the summer in other sectors and separate them from the negative ones observed in the extraction of crude oil and gas (6).

⁷ See <https://www.kommersant.ru/doc/4575783>

The second group of sectors most severely affected by the coronacrisis includes services closely associated with the presence of a large number of people. This includes travel agencies (79) and public catering (56). For this sectors, we also built basic indexes describing the volume of incoming payments and compared them with the Rosstat indicators, the scale of the fall and subsequent recovery in the dynamics of financial flows and Rosstat indicators are similar.

Figure 5. Dynamics of the industrial production index and average incoming payments, 100 = average in 2019, seasonally adjusted

Production of textiles (13)

Production of medicines and materials used for medical purposes (21)

Sources: Rosstat, Bank of Russia, authors' calculations

In the tourism sector, the decline of financial flows in May adjusted for seasonality reached 91% compared to February. According to Rosstat, the decrease in the volume of paid services in this sector was 98% when adjusted for seasonality. The sector then experienced a recovery. However, by December, the volume of incoming revenues reached only 50% compared to February when adjusted for

seasonality, and the volume of paid services was at 60% of its level in February when adjusted for seasonality (Figure 7).

In 2020, the slump in the business activity in public catering (56) was also significant, although smaller in scale than in the tourism sector. In April, financial flows decreased by 38% compared to February when adjusted for seasonality and the turnover of the public catering sector fell by 44% when adjusted for seasonality. After the recovery in the summer, the volume of revenues in the sector reported in the fall was relatively stable (91% of the February level in December, adjusted for seasonality), but the turnover of the public catering sector began to fall again in the last months of the year (77% of the February level in December, adjusted for seasonality) (Figure 7).

Figure 6. Contributions of deviations from the 'normal' level to the dynamics of the consolidated indicator weighted with the shares of GVA, %

Sources: Rosstat, Bank of Russia, authors' calculations

In addition to traditional Rosstat indicators, other alternative indicators also can give an idea of economic activity in certain areas. One such indicator is SberIndex,⁸ which reflects the annual growth rate in the spending volume of consumers who pay for their purchases using Sberbank financial products (Figure 8).

Another similar indicator of economic activity is the Tinkoff CoronaIndex.⁹ It reflects the annual growth rate in the turnover of businesses that use the services of Tinkoff Bank.

All three indicators – the volume of incoming payments, consumer spending and business turnover – in the area of tourism (79) in April–May lagged behind

⁸ See <https://www.sberindex.ru/en>

⁹ See <https://index.tinkoff.ru/> [In Russian].

the previous year's values by about 90%. SberIndex and Tinkoff CoronaIndex give a more optimistic assessment of the recovery in the sector since July than the BoR PS data. In late August and early December, there was a sharp increase in consumer spending. However, in the last weeks of the year, their volume fell back to the level of about 50% of the previous year's figures. For July–December 2020, the estimated decrease in incoming payments was more stable than the dynamics of consumer spending and stayed in the range of 50–60% year-on-year.

Figure 7. Dynamics of Rosstat indicators and average incoming payments, 100 = average in 2019, seasonally adjusted

Activities of travel agencies and other organisations providing services in tourism (79)

Activities of public catering (56)

Source: Rosstat, Bank of Russia, authors' calculations

The volume of consumer spending in cafes and restaurants during the most severe restrictions (April–May) decreased more significantly (65–77%) than business turnover (38–62%) and the volume of incoming payments (31–49%) in the provision of food and beverage services (56). The summer months have seen a

narrower gap from last year's levels. According to Tinkoff CoronaIndex, in August, the turnover in restaurants even exceeded the level of the previous year. However, by the end of the year, the dynamics of both consumer spending and incoming payments in this area deteriorated again.

Figure 8. Growth rate of incoming payments according to BoR PS data and alternative indicators of business activity in tourism (79), %, YoY

Source: Bank of Russia, SberIndex, Tinkoff CoronaIndex

Figure 9. The growth rate of incoming payments according to BoR PS data and other alternative indicators of business activity in public catering (56), %, YoY

Source: Bank of Russia, SberIndex, Tinkoff CoronaIndex

The dynamics of financial flows estimated in accordance with BoR PS data is consistent both with Rosstat indicators on the volume services provided by travel agencies and the turnover of the public catering sector and with alternative indicators built using the payment data from individual commercial banks. All of them reflect a strong decline, recovery and the subsequent fading of recovery growth towards the end of the year.

6. Further research areas

BoR PS data is highly valuable for economic analysis. On the one hand, there is a relationship between the consolidated high-frequency indicators of financial flows and economic activity in general, which makes it possible to assess its fluctuations virtually in real time. On the other hand, BoR PS data provides a detailed view of the sectoral dynamics and structural shifts in the economy. Therefore, BoR PS data has the potential to be used in forecasting economic activity.

BoR PS data can also be used in a wide range of research and applied tasks. In particular, such data, along with the indication of the payment purpose, can serve as a source for building tables similar to the input-output tables that are used in a wide range of economic models but are published by Rosstat with a several-year lag. BoR PS data can also be used to assess interregional relations and the level of economic activity in Russian regions.

Payment data can serve as a source for building indicators that represent an alternative to traditional statistical indicators. For example, payments of personal income taxes and contributions to the Social Insurance Fund can serve as an indicator of the labour market, including at the regional level. Such an indicator can take into account not only changes in the number of employed persons but also a decline in the number of hours worked or a decrease in employee remuneration. Data on payments aimed at acquiring shares of capital in other companies can become an indirect indicator of investment activity.

Financial transaction data can provide more accurate information about the date when a business ended its economic activity than data from the Single State Register of Legal Entities. This will make it possible to assess the life cycle of businesses and events and factors that determine their exit from the market much more accurately.

The results of processing BoR PS data in the current version are just the start of a journey in using high-frequency data for economic activity analysis. In terms of practical policy conclusions, it appears that the monitoring of payments has a quite extensive scope of application, as it will help not only in the current analysis of business activity and the balance of short- and medium-term risks for price dynamics, but also in long-term studies of the real and financial sectors. Therefore, the findings presented in this paper should become an important impetus for the research community to continue actively exploring the possibilities of using financial payment data.

References

- Aprigliano, V., Ardizzi, G. and Monteforte, L.** (2017). *Using the Payment System Data to Forecast the Italian GDP*. Banca d'Italia Temi di Discussione (Bank of Italy Working Papers), N 1098.
- Barlas, A. B., Guler, S. M., Ortiz, A. and Rodrigo, T.** (2020). *Investment in Real Time and High Definition: A Big Data Approach*. BBVA Research Working Papers, N 13.
- Carvalho, V., Garcia, J. R., Hansen, S., Ortiz, A., Rodrigo, T. and More, J. V. R.** (2020). *Tracking the COVID-19 Crisis with High-Resolution Transaction Data*. Cambridge Working Papers in Economics, N 2030.
- Chetty, R., Friedman, J. N., Hendren, N., Stepner, M. and The Opportunity Insights Team** (2020). *The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data*. NBER Working Paper, N 27431. <https://doi.org/10.3386/w27431>
- Duarte, C., Rodrigues, P. M. M. and Rua, A.** (2017). A Mixed Frequency Approach to the Forecasting of Private Consumption with ATM/POS Data. *International Journal of Forecasting*, 33(1), pp. 61–75. <https://doi.org/10.1016/j.ijforecast.2016.08.003>
- Galbraith, J. W. and Tkacz, G.** (2015). *Nowcasting GDP with Electronic Payments Data*. ECB Statistics Paper Series, N 10.
- Khabibullin, R. and Seleznev, S.** (2020). *Stochastic Gradient Variational Bayes and Normalizing Flows for Estimating Macroeconomic Models*. Bank of Russia Working Paper Series, N 61.
- Lee, K.** (2018). Seasonal Adjustment. In: *Quarterly National Accounts Manual (2017 Edition)*. Washington, DC: International Monetary Fund, pp. 127–164.
- Ollech, D.** (2018). Seasonal Adjustment of Daily Time Series. Deutsche Bundesbank Discussion Paper, N 41.
- Ponomarenko, A., Popova, S., Sinyakov, A., Turdyeva, N. and Chernyadyev, D.** (2020). Assessing the Consequences of the Pandemic for the Russian Economy Through an Input-Output Model. *Russian Journal of Money and Finance*, 79(4), pp. 3–17. <https://doi.org/10.31477/rjmf.202004.03>
- Seleznev, S., Turdyeva, N., Khabibullin, R. and Tsvetkova, A.** (2020). *Seasonal Adjustment of the Bank of Russia Payment System Financial Flows Data*. Bank of Russia Working Paper Series, N 65.
- Taylor, S. J. and Letham, B.** (2018). Forecasting at Scale. *The American Statistician*, 72(1), pp. 37–45. <https://doi.org/10.1080/00031305.2017.1380080>
- Tipping, M. E.** (2001). Sparse Bayesian Learning and the Relevance Vector Machine. *Journal of Machine Learning Research*, 1, pp. 211–244.